OPEN ACCESS

ORIGINAL ARTICLE

Longitudinal profiling of the lung microbiome in the AERIS study demonstrates repeatability of bacterial and eosinophilic COPD exacerbations

David Mayhew,[1] Nathalie Devos,[2] Christophe Lambert,[2] James R Brown,[1] Stuart C Clarke,[3,4] Viktoriya L Kim,[4] Michal Magid-Slav,[1] Bruce E Miller,[5] Kristoffer K Ostridge,[4] Ruchi Patel,[6] Ganesh Sathe,[6] Daniel F Simola,[1] Karl J Staples,[3,4,7] Ruby Sung,[5] Ruth Tal-Singer,[5] Andrew C Tuck,[3] Stephanie Van Horn,[6] Vincent Weynants,[2] Nicholas P Williams,[4] Jeanne-Marie Devaster,[2] Tom M A Wilkinson,[3,4,7] on behalf of the AERIS Study Group

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/thoraxjnl-2017-210408).

For numbered affiliations see end of article.

Correspondence to
Professor Tom M A Wilkinson, Clinical and Experimental Sciences, University of Southampton Faculty of Medicine, Southampton General Hospital, Southampton SO16 6YD, UK;
t.wilkinson@soton.ac.uk

DM and ND contributed equally.

Received 16 April 2017
Revised 29 November 2017
Accepted 5 December 2017
Published Online First
31 January 2018

ABSTRACT

Background Alterations in the composition of the lung microbiome associated with adverse clinical outcomes, known as dysbiosis, have been implicated with disease severity and exacerbations in COPD.

Objective To characterise longitudinal changes in the lung microbiome in the AERIS study (Acute Exacerbation and Respiratory InfectionS in COPD) and their relationship with associated COPD outcomes.

Methods We surveyed 584 sputum samples from 101 patients with COPD to analyse the lung microbiome at both stable and exacerbation time points over 1 year using high-throughput sequencing of the 16S ribosomal RNA gene. We incorporated additional lung microbiology, blood markers and in-depth clinical assessments to classify COPD phenotypes.

Results The stability of the lung microbiome over time was more likely to be decreased in exacerbations and within individuals with higher exacerbation frequencies. Analysis of exacerbation phenotypes using a Markov chain model revealed that bacterial and eosinophilic exacerbations were more likely to be repeated in subsequent exacerbations within a subject, whereas viral exacerbations were not more likely to be repeated. We also confirmed the association of bacterial genera, including *Haemophilus* and *Moraxella*, with disease severity, exacerbation events and bronchiectasis.

Conclusions Subtypes of COPD have distinct bacterial compositions and stabilities over time. Some exacerbation subtypes have non-random probabilities of repeating those subtypes in the future. This study provides insights pertaining to the identification of bacterial targets in the lung and biomarkers to classify COPD subtypes and to determine appropriate treatments for the patient.

Trial registration number Results, NCT01360398.

Key messages

What is the key question?

► What is the composition and stability of the lung microbiome in patients with COPD when longitudinally sampled at stable and exacerbation events?

What is the bottom line?

► The composition of the lung microbiome shows unique profiles within subtypes of COPD exacerbations, and the exacerbations experienced by an individual over time are non-random.

Why read on?

► These results describe a unique examination of the stability of the lung microbiome in COPD and predictability of future exacerbations available within a large cohort containing longitudinal sampling and clinical measurements over 1 year.

INTRODUCTION

COPD is a chronic inflammatory disorder resulting in irreversible decline in lung function as a consequence of inhalation of tobacco smoke or other irritants.[1] One of the difficulties in treating and managing COPD is the heterogeneity of this complex disease in terms of severity, progression, exercise tolerance and nature of symptoms.[2 3] This complexity is also evident in acute exacerbations of COPD (AECOPD), which are transient and apparently stochastic periods of increased COPD symptoms requiring additional medical treatment and, often, hospitalisation.[4] Known subtypes of exacerbations are classified by the nature of key triggers including bacterial or viral infections, and/or high eosinophil levels, and these events are typically treated with a combination of antibiotics and steroids in a non-specific manner.[5]

The lung microbiome represents an emerging opportunity to understand COPD heterogeneity and exacerbations. The healthy human lung contains a variety of commensal microbiota throughout the respiratory tract, and these bacteria can show substantial heterogeneity between individuals, across regions within the lung and over time within an individual.[6–8] Alterations in the taxonomic composition of the lung microbiome, known

To cite: Mayhew D, Devos N, Lambert C, *et al*. *Thorax* 2018;**73**:422–430.

as dysbiosis, have been associated with multiple lung diseases, and in particular may play a functional role in disease severity and exacerbations in COPD.[6 9]

Multiple studies have reported differences in the microbiome between healthy and disease states, differences correlated with both COPD severity and exacerbation states within an individual, and interactions between the microbiome and host immune response.[7 9–12] Previous studies have observed changes in the lung microbiome between AECOPD events relative to stable events,[9 13–15] but the stability of the lung microbiome longitudinally sampled over longer time frames and multiple exacerbations remains poorly understood.

The Acute Exacerbation and Respiratory InfectionS in COPD (AERIS) observational cohort study allows for a unique examination of the lung microbiome with a rich set of microbiology and clinical measurements longitudinally observed in stable time points and exacerbation events in 104 patients with COPD (a subset of the full cohort of 127 patients).[16 17] Results have already been described for subjects followed over 1 year for the primary study objective (estimation of the incidences of all-cause AECOPD and AECOPD with sputum containing bacterial pathogens detected by culture) and for secondary objective results on the incidences of bacterial and viral pathogens detected in AECOPD by PCR and in stable-state COPD by culture (bacteria only) and PCR.[17] Here we present 1-year data from the 16S rRNA gene sequencing analysis of the AERIS patient cohort. Using this deeply phenotyped cohort, we tested whether subtypes of COPD and AECOPD have different lung microbiome profiles and whether these compositions are stable over multiple clinical visits, particularly at exacerbations.

METHODS

Study design

The AERIS study (ClinicalTrials.gov: NCT01360398) was a prospective, observational cohort study based at University Hospital Southampton (UHS). The study assessed the contribution of changes in the COPD airway microbiome to the incidence of AECOPD in patients aged 40–85 years with a confirmed diagnosis of COPD, categorised as moderate, severe or very severe,[1 16] recruited from UHS and referring practices from June 2011 to June 2012. The study protocol has been described in detail.[16] AERIS was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice. All participants provided written informed consent. The protocol summary is available at www.gsk-clinicalstudyregister.com (study identifier, 114378).

Processing of sputum samples

All study procedures for sputum sampling, the detection of exacerbations and pathogen detection were described previously.[16 17] Briefly, patients were followed monthly in the stable state and reviewed within 72 hours of onset of AECOPD symptoms. Sputum samples were obtained by spontaneous expectoration or induced and were processed according to standard methods (and were collected prior to additional antibiotic treatment for the exacerbation). Bacterial respiratory pathogens were identified using techniques described in the online supplementary methods.

COPD exacerbation subtypes were classified using previously defined criteria.[5] The online supplementary methods provide more details on bacterial, viral and eosinophilic subtype classification and DNA extraction.

16S rRNA gene amplification and sequencing

The V4 hypervariable region of the 16S rRNA gene was amplified with specific primers (515F/806R), including Illumina sequencing adapters and sample-specific barcodes, and sequenced on an Illumina MiSeq sequencer (see online supplementary methods). Sequence data are deposited in the National Center for Biotechnology Information's Sequence Read Archive (SRA) (BioProject PRJNA377739, SRA accession SRP102629).

16S rRNA gene sequence analysis

Paired-end sequence reads were filtered for quality, assembled using PEAR (paired-end read merger)[18] and then processed using the QIIME pipeline[19] (see online supplementary methods). The flow chart for the analysis of FASTQ sequence files is shown in online supplementary figure S1. Because these samples were generated from low biomass material, we performed experimental and bioinformatics controls and concluded that the probability of contaminating DNA amplicons was low (online supplementary figure S2).

Samples were rarified to a fixed depth (30,419 reads) corresponding to the fewest number of reads in a sample which was sufficiently amplified and sequenced. Shannon diversity (evenness of bacteria within a sample), relative abundances and UniFrac distances (fraction of the genetic distance that is unique for one of the compared samples) were calculated using QIIME. Comparisons of Shannon diversity and relative abundance between non-longitudinal groups were performed with Mann-Whitney or Kruskal-Wallis tests (after averaging repeated measures within a subject at that condition to a single value), and comparisons of UniFrac distances within and between groups were performed with a one-way analysis of variance after randomly dividing individuals into equally sized subsets for each comparison group and only using one distance measurement from each individual to ensure independence of the test. Longitudinal comparisons of relative abundances between stable and exacerbation events were made using a linear mixed-effects model and treating the subject as a random effect. We assumed that multiple exacerbations within an individual are independent. Additional information on the statistical analyses is described in the online supplementary methods.

Markov chain modelling

In order to describe the probability of each exacerbation repeating the same phenotype in its subsequent exacerbation, longitudinal exacerbations within individuals were modelled as a Markov chain. Only subjects with multiple exacerbations were included and each exacerbation type was modelled as a state classified using the same criteria[5] or modifications as listed. Transition probabilities were calculated by counting the relative frequency of observed transitions between temporally adjacent exacerbations within an individual. Differences in Markov chain transition frequencies between observed frequencies and expected independent frequencies from incidence of each phenotype were tested with a χ^2 test, and comparisons of frequencies between nodes were tested with a Fisher's exact test.

RESULTS

Population and sampling

Samples for 16S rRNA gene sequencing were collected from 104 subjects with available sputum samples in the first year of the study (figure 1). Characteristics of the cohort used for microbiome analysis were similar to those of the full cohort[17] (table 1 for summary) (online Supplementary file 3 for full information).

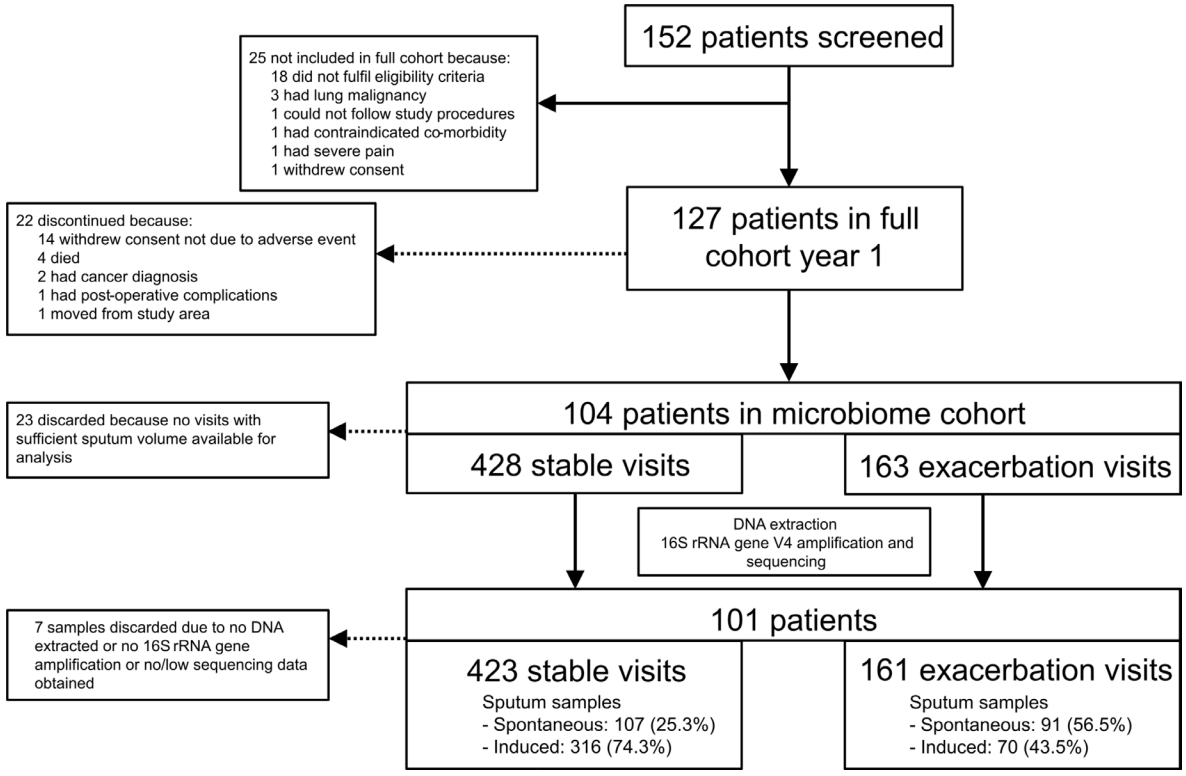

Figure 1 Flow chart of subject enrolment, sputum sampling and selection samples for microbiome analysis for AERIS (Acute Exacerbation and Respiratory InfectionS in COPD).

Of the 104 subjects included, 101 had at least one sputum sample, which was successfully amplified and sequenced for inclusion in the microbiome analysis. Comparison of the nature of sputum sampling on microbial profiles revealed no significant differences between spontaneous and induced samples (online supplementary figure S3).

Table 1 Characteristics of the cohort for microbiome analysis

Characteristics	N=101
Age (years) at enrolment, mean±SD	67.1±8.4
Female sex, n (%)	42 (41.6)
Body mass index at enrolment, mean±SD	27.6±5.4
Current smokers, n (%)	40 (39.6)
Medication for COPD, n (%)	101 (100)
Inhaled corticosteroids, n (%)	94 (93.1)
COPD status, GOLD stage, n (%)	
Mild	0 (0)
Moderate	45 (44.6)
Severe	40 (39.6)
Very severe	16 (15.8)
Bronchiectasis status, n (%)	10 (9.9)
Number of exacerbations experienced by subject in 12 months, n (%)	
One exacerbation	31 (22.0)
Two exacerbations	23 (29.1)
Three or more exacerbations	47 (19.7)
FEV_1 after bronchodilator use (% predicted), mean±SD	47.1±12.8

GOLD, Global Initiative for Chronic Obstructive Lung Disease; N, number of subjects in the microbiome cohort; n, number of subjects corresponding to characteristics.

Lung microbiome composition shows differences with COPD severity

An analysis of the relative abundances of bacterial taxa identified in the set of 584 microbiome samples passing quality control revealed bacteria commonly observed in the lung microbiome, with Firmicutes, Proteobacteria and Bacteroidetes representing the three most abundant phyla, and *Veillonella*, *Haemophilus*, *Streptococcus*, *Prevotella* and *Moraxella* representing the five most abundant genera (online supplementary figure S4 and Supplementary file 4). The number of successfully sequenced microbiome sputum samples averaged 5.7 per subject, with 2.1 collected during an exacerbation.

We first compared the composition and diversity in observed bacteria (measured by the Shannon Diversity Index) of the lung microbiome with trends observed in previous studies. As described in other studies describing the lung microbiome in COPD, we observed a shift towards increasing Proteobacteria with increasing disease severity.[20 21] More specifically this shift included a significant increase in *Haemophilus* (Proteobacteria) and decreases in *Prevotella* (Bacteroidetes) and *Veillonella* (Firmicutes), as well as decreased Shannon diversity (P_{adj} <0.05 for each, Mann-Whitney) with increasing disease severity (figure 2A). Current smokers did not show significant differences in diversity or composition (online supplementary figure S5).

Changes in the lung microbiome in exacerbation states

Testing for differences between stable and exacerbation states allows for inclusion of longitudinal data from patients with samples at both visit types with a linear mixed-effects model.

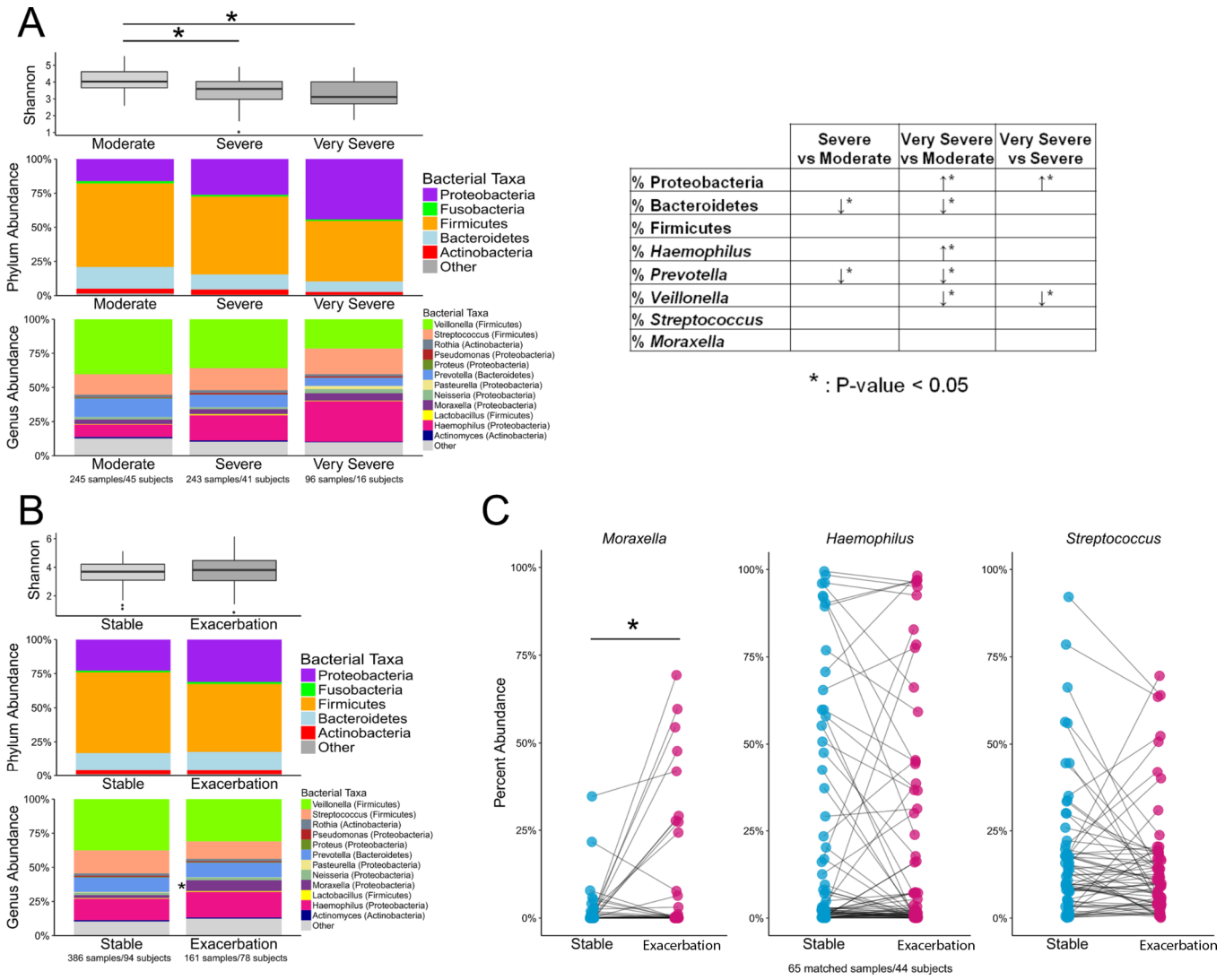

Figure 2 Microbiome differences in disease severity and stable or exacerbation visits. (A) The Shannon Diversity Index and relative abundances of bacteria labelled at the phylum and genus level of samples grouped by COPD disease severity. The bar graphs show the mean relative abundance at the subject level after averaging for multiple measures for that subject. Significant differences in relative abundances between groups are labelled with arrows indicating the relative change in abundance; *P<0.05 (Mann-Whitney). (B) The same alpha diversity and relative abundances grouped by stable or exacerbation status showed fewer differences overall except for *Moraxella*; *P<0.05 (linear mixed-effects model). (C) Paired analysis of changes in relative abundances of key genera between matched stable and subsequent exacerbation events; *P<0.05 (paired Student's t-test).

The differences between these states were less pronounced than differences between disease severities. No significant changes in Shannon diversity or core taxa relative abundances were observed, with the exception of the genus *Moraxella* (Proteobacteria), which showed a significant increase in relative abundance in exacerbations (P=0.0012) (figure 2B). To confirm these results we also used a paired t-test to compare matched stable and exacerbation events within an individual. Again, *Moraxella* showed a significant increase in exacerbation (P=0.0227) (figure 2C). The probability of a sample being observed in an exacerbation event compared with stable is 2.6-fold higher if the relative abundance of *Moraxella* is greater than 10% (95% CI 2.1 to 38.1).

Clinical and microbiology data have been used as biomarkers to stratify subtypes of COPD and AECOPD,[5] and some of these have revealed distinct lung microbiome profiles.[9] We compared the composition of previously defined exacerbation subtypes characterised by sputum potentially pathogenic bacterial culture, viral PCR or eosinophil percentage. The composition of bacterial exacerbations was dissimilar from those of viral and eosinophilic, with a higher proportion of Proteobacteria in bacterial exacerbations (online supplementary figure S6). Another classification of COPD with a unique microbiome profile is that of bronchiectasis, where we observed a substantial increase in *Haemophilus* (P=1.2E-5), which was evident in both stable and exacerbation events (online supplementary figure S7).

Longitudinal stability of the lung microbiome

Longitudinal sampling within the AERIS study allowed us to assess the relative stability of the lung microbiome within an individual. To analyse temporal microbiome stability, we computed UniFrac distances (weighted and unweighted) between all pairs of microbiome profiles within a subject, and stratified results based on comparisons between stable–stable, stable–exacerbation and exacerbation–exacerbation comparisons. This metric measures the similarity in bacterial composition between samples, and higher distances indicate more dissimilar communities. In all groups, we found UniFrac distance to be significantly lower

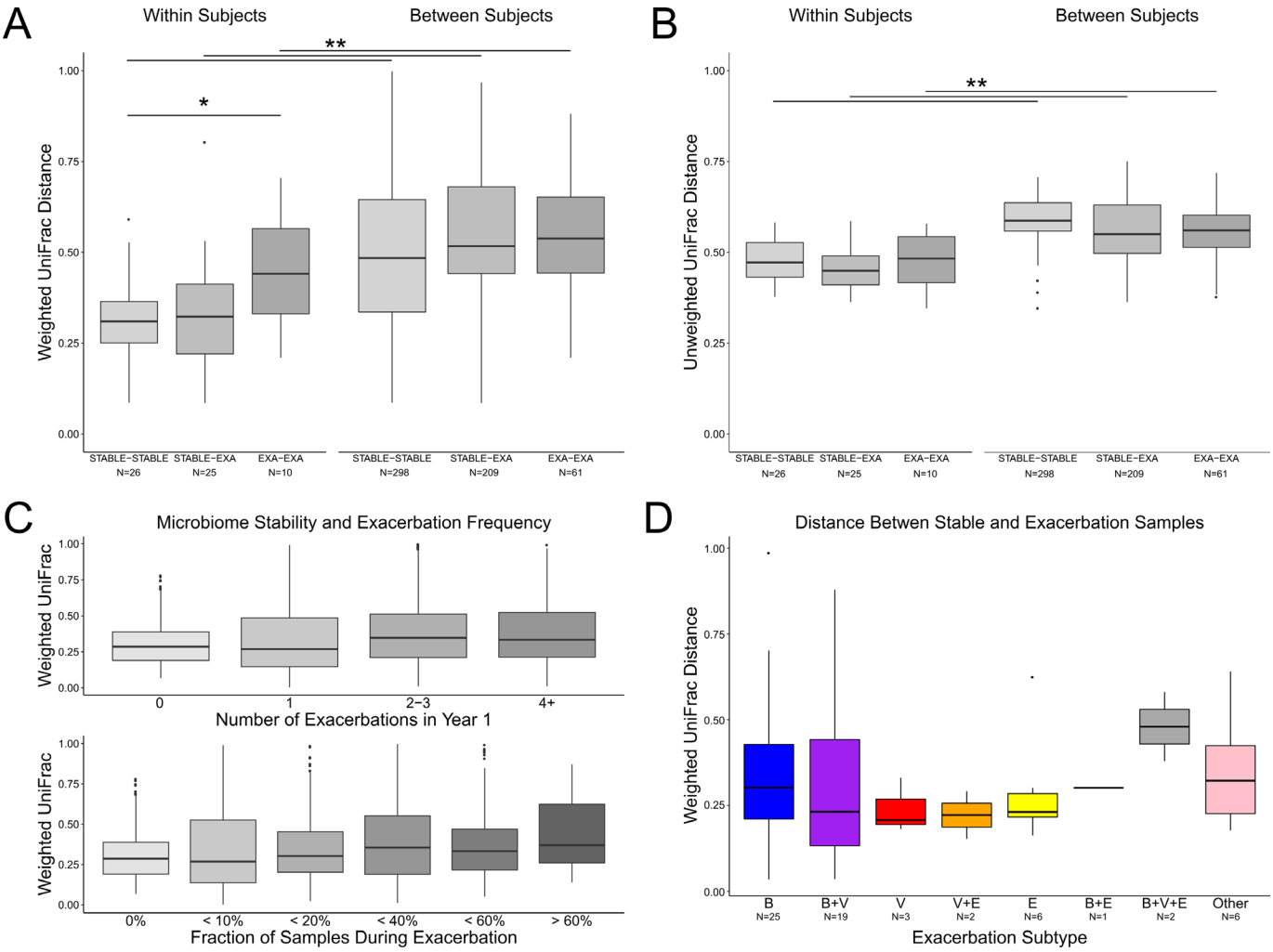

Figure 3 Lung microbiome stability. (A) Weighted UniFrac distances measured within and between subjects and comparing stable and exacerbation events after randomly dividing individuals into equal-sized subsets to ensure independence; *P<0.05, **P<0.01 (one-way analysis of variance (ANOVA)). (B) Unweighted UniFrac distances measured within and between subjects and comparing stable and exacerbation events on the same subsets; **P<0.01 (one-way ANOVA). (C) Weighted UniFrac distances for all within-subject samples as a function of exacerbation frequency defined by the number of exacerbation events and the fraction of samples within an individual taken during an exacerbation. (D) Paired weighted UniFrac distances between an exacerbation sample and its previous stable sample from that subject. Exacerbation subtypes labelled as B (bacterial), V (viral), E (eosinophilic), other or mixed. There was not a significant difference in UniFrac distances between these groupings of stable-to-exacerbation transitions (P=0.38, one-way ANOVA). EXA, exacerbation.

within an individual compared with distances between individuals (figure 3A,B). This result suggests that individuals have somewhat distinct lung microbiomes from each other. Moreover, weighted UniFrac distances were significantly higher for exacerbation–exacerbation comparisons relative to stable–stable comparisons (P<0.05) (comparisons using unweighted UniFrac distance not significant) (figure 3B). This measure-specific result suggests that dysbiosis events in the lung may typically result from changes in the relative abundance of pre-existing bacteria (detected by weighted UniFrac) rather than complete removal or appearance of novel species (detected by unweighted UniFrac).

Moreover, given that weighted UniFrac distances involving exacerbations were higher than stable–stable, it appears exacerbation events are most likely to be associated with dysbiosis within an individual. While the lung microbiome may have a degree of within-subject stability, there remains a large degree of variation between longitudinal samples, especially when comparing an individual's exacerbation events (see online supplementary figure S4B). One possible explanation is that the frequency of exacerbation events experienced by an individual may contribute to destabilisation of the lung microbiome, such that frequent exacerbators may be associated with greater dysbiosis than infrequent exacerbators. To evaluate this hypothesis, we analysed an individual's UniFrac distance as a function of exacerbation frequency. Because of incomplete sampling of all exacerbation events, we conservatively estimated exacerbation frequency in two ways, by counting total reported exacerbation events and by the proportion of microbiome samples obtained from an exacerbation event relative to the total number of microbiome samples obtained for that individual. We found that the lung microbiome became more distinct with greater exacerbation frequency using either definition, affecting bacterial abundance in both stable and exacerbation states (r=0.14, P<0.001, Pearson) (figure 3C) (online Supplementary file 5). To identify specific taxa associated with exacerbation frequency, we computed the correlation between each taxon's average abundance with exacerbation frequency across subjects. The genus with the highest positive correlation was *Moraxella* (r=0.23,

P=0.016, Pearson) (online supplementary figure S8), consistent with our observation of its increased abundance in exacerbations relative to stable states. In contrast, the genus *Lactobacillus* showed the most negative correlation with exacerbation frequency (r=−0.37, P=0.02, Pearson).

To test whether any of the AECOPD phenotypes are more likely to experience dysbiosis, or large changes between stable and exacerbation events, we performed a paired analysis comparing the weighted UniFrac distance between each exacerbation and its previous stable event in that subject (figure 3D). While 18% of all matched stable–exacerbation transitions have a UniFrac distance greater than 0.5, and some of the largest occurred in bacteria-driven exacerbations, none of the exacerbation classifications had significantly higher changes in UniFrac distance compared with exacerbations lacking that phenotype. With 46% of transitions having a UniFrac distance less than 0.25, many exacerbations show little evidence of dysbiosis.

Stochastic modelling of COPD exacerbation phenotypes

Having observed that some COPD exacerbation phenotypes had different lung microbiome profiles, we used the longitudinal attributes of the AERIS study to model the exacerbations experienced by an individual over time as a stochastic process. We employed a Markov chain analysis and defined each exacerbation event as a discrete state of being positive or negative for bacterial, viral or eosinophilic status independently for each type (inclusive of mixed-type exacerbations). To estimate the state transition probabilities for each Markov model, we counted the number of exacerbations of a given phenotype that were chronologically followed by another exacerbation with the same phenotype (figure 4A). We found significantly non-random transition probabilities for the bacterial and eosinophilic Markov models with P values of 9.25E-11 and 1.42E-3 (χ^2 test, df=3), respectively. In contrast, the viral Markov model was not significant (P=0.141). These results indicate that for bacterial and eosinophilic exacerbations, the phenotype of the next exacerbation experienced by an individual may be more likely to repeat the prior exacerbation phenotype than expected by chance. Interestingly, we did not detect a significant difference in the times between exacerbations for any of the Markov transitions (online Supplementary file 2).

Next, a Markov model of bacterial exacerbation phenotype was built to examine the potential role of *Haemophilus influenzae* (HI). Most of the *Haemophilus* observed from sputum is likely non-typeable *Haemophilus influenza*e (NTHI) a non-capsulated bacterium that commonly infects the airways and whose carriage is commonly associated with COPD and inflammation.[12] In our analysis of microbiome profiles in different COPD phenotypes, *Haemophilus* was the dominant genus observed in patients with bronchiectasis, both in stable and exacerbation events. We hypothesised that patients with a positive HI culture in a positive bacterial exacerbation may be even more likely to repeat an HI-positive culture in their next exacerbation. The Markov chain of bacterial exacerbations was modified by dividing the bacterial-positive exacerbation state in two separate states of positive and negative HI. After calculating the new transition probabilities, this Markov chain was non-random (P=1.42E-12, χ^2 test, df=5), and the HI-positive bacterial exacerbations were most likely to repeat a subsequent HI-positive exacerbation and less likely to transition non-bacterial exacerbation compared with HI-negative bacterial exacerbations (P=2.62E-4, Fisher's exact test) (figure 4B). The repetition of HI-positive exacerbations suggests a persistence of *Haemophilus* in certain subtypes

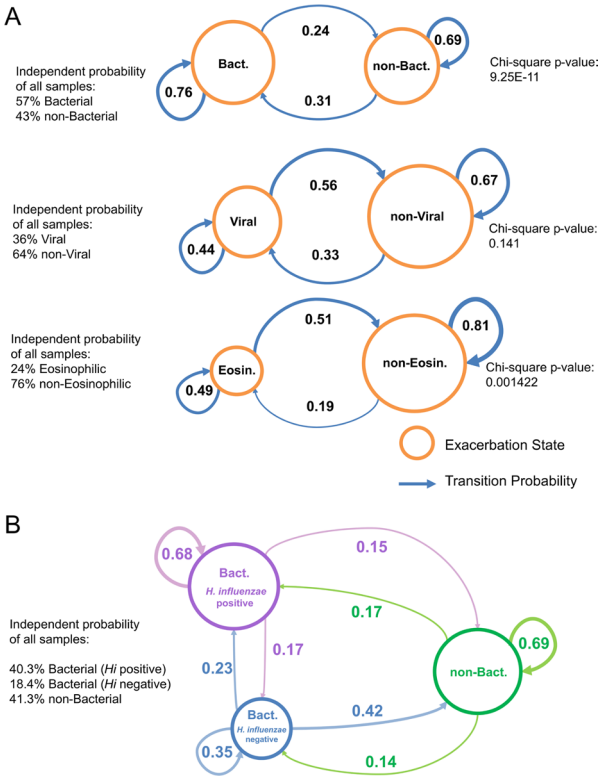

Figure 4 Markov chain analysis of transitions between exacerbation states. (A) Markov chain analysis from longitudinal exacerbation sampling within individuals identifies non-random transition probabilities for bacterial and eosinophilic exacerbations, but not viral. The size of each node is proportional to abundance of that exacerbation type, and the width of the edges is proportional to the transition probabilities. (B) Markov chain analysis of the bacterial exacerbation identifies significantly different transition probabilities for bacterial exacerbations that were positive or negative for the presence of *Haemophilus influenzae* (HI).

of COPD over time and observed through multiple exacerbations. We also expanded the Markov chain analysis of eosinophilic exacerbations by dividing the eosinophilic-positive state into high eosinophilic (>6% in sputum) and moderate eosinophilic (>3% and <6%) positive exacerbations.[22] The revised model showed that the higher the eosinophil levels in the exacerbation the more likely it was to repeat the same high eosinophil phenotype (P=0.02, Fisher's exact test) (online supplementary figure S9).

DISCUSSION

This study has confirmed previous findings of lung microbiome heterogeneity with distinct patterns of bacterial abundance in COPD subtypes,[9 20 21] and, for the first time, described the stability of the lung microbiome in COPD and the non-random nature of exacerbations experienced by an individual over time by modelling exacerbations as a Markov chain. Our analysis shows that grouping samples by specific phenotypes could yield distinct microbiome populations or probabilities of repeating that type of exacerbation. These subtypes underscore the importance of sample size and stratification in generating reproducible results in studying the microbiome of a heterogeneous disease.

Using the unique repeated longitudinal sampling of the AERIS study design, we found that the lung microbiome shows significantly less variation within an individual than between

individuals, indicating some degree of temporal stability of an individual's lung microbiome. Nevertheless, we also observed notable dysbiosis events within individuals. Exacerbations within individuals showed higher microbiome variability than stable time points, and frequent exacerbators are more likely to experience significant changes in lung microbiome patterns. These findings will require further examination to determine the causes and consequences of lung dysbiosis, as likely variables include bacterial composition, host immune system and the use of antibiotics. One appealing model is that lung microbiome composition can functionally drive host inflammatory signals via bacterial proteins or metabolites; specific examples have already been identified, and models to investigate their mechanism are being created and tested.[9 11 23 24]

An appreciation of an individual's lung microbiome may influence the future clinician's choice of appropriate therapy, especially for exacerbations that are commonly treated with antibiotics. The increasing prevalence of recurrent *Clostridium difficile* infections after antibiotic use has highlighted the risks of disrupting the healthy microbiome when treating primary infections.[25] Therefore a move towards selective-spectrum antimicrobials, which may be less likely to disturb commensal species and hence minimise treatment-related risk of future infection or exacerbations, needs to be explored. The relationship of NTHI with bronchiectasis and its association with repeating bacterial exacerbations may represent a unique treatment challenge. *Haemophilus* is known to produce biofilms[26] protecting it from the immune system and antibiotics, which may explain its persistence in these COPD subtypes and may offer another avenue for a therapeutic target. Biofilm formation is a component of antibiotic resistance in *Moraxella catarrhalis* and *Pseudomonas aeruginosa* as well,[27] indicating potentially common difficulties in eliminating pathogenic bacteria in the lung. The recurring motifs of key pathogenic bacteria such as *Moraxella* and NTHI identified in this and other COPD lung microbiome studies support the potential of a vaccine or targeted antibacterial drug against these pathogens in order to minimise a wider disruption of the lung microbiome.

The ability to model exacerbation phenotypes as stochastic processes has important implications for diagnosis and treatment of AECOPD if the phenotypes of future clinical events can be accurately predicted. Exacerbation events are typically diagnosed and treated as independent phenomena, as and when they are experienced by a patient with COPD.[28] If clinical data from the previous exacerbation can inform the likely phenotype of the next event, it can enable a more rapid administration of the appropriate therapy.[29] While bacterial and eosinophilic exacerbations are most likely to repeat the same phenotype in our Markov chain model, there is also evidence from other studies that viral infections may predispose the respiratory tract to subsequent secondary bacterial infections,[30] indicating additional longitudinal relationships between the infections and colonisation of viruses and bacteria.

A major strength of this study is the longitudinal information within individuals over time at both stable and exacerbation time points, which has allowed for us to study the dynamics of the lung microbiome and its relationships with clinical phenotypes in ways that have to date not been possible. While we have microbiome data sampled over 1 year in the study, they do represent a short period of time in the typical timeline of COPD progression. An analysis of how the stability of exacerbation phenotype probabilities or how microbiome diversities change over longer periods of time will require longer term studies.

A limitation of the study was the exclusive use of sputum as source material for 16S rRNA profiling. The diversity of microbiome compositions at different locations within the respiratory pathway was not fully represented. Sufficient sputum for microbiome analysis was not always available at every exacerbation or stable visit, which blinded the analyses to these points in time. Samples in the study included both induced and spontaneous sputum, which may have confounding effects on the measurement of microbiome composition.[31] Using sputum samples (which may contain relatively low abundances of bacterial DNA) increased the risk of sample contamination and warrants including additional negative controls.[32] While massively parallel 16S rRNA gene sequencing is a powerful tool in surveying bacterial populations, using a single hypervariable region for PCR limited the resolution of our bacterial analysis to the genus level. This resolution prevented more subtle examination of meaningful differences in pathogenic versus non-pathogenic species or strains within the same genus, which can have very different effects on lung health and COPD (eg, HI compared with *Haemophilus haemolyticus*).[33] However, HI-specific quantitative PCR confirmed the high prevalence of this pathogen in the lung of AERIS subjects.[17]

A final caveat was that our study was conducted at a single site, which limited the geographical scope of the finding, but avoided the potential for additional confounding variables between different sites. This particular population limited our ability to explore the relations of the lung microbiome with different maintenance COPD medications as has been observed in other studies,[13 34] as the vast majority of sampled participants (94 out of 101) were receiving inhaled corticosteroid treatment at the time of the study. It is worth noting that sputum samples from exacerbation events were collected before antibiotic treatment, which likely simplifies another confounding treatment variable. Interestingly, despite being a single-site study, many of our microbiome observations mirror trends observed in other studies. The changes in *Moraxella* abundance in exacerbations and different compositions of exacerbations mirror findings by Wang and colleagues.[9] Of course, *Moraxella* dysbiosis can only explain part of the variance in lung microbiome abundance, as in both studies only a minority of exacerbation events show a substantial increase in *Moraxella*, indicating other factors are involved. *Haemophilus*, while generally more stable (especially in bronchiectasis), can change in relative abundance in a minority of subjects and may also explain some of the dysbiosis in bacteria-driven exacerbations. We and others have observed commensal lung bacteria that correlate with positive outcomes, including *Lactobacillus*. *Lactobacillus* is often overlooked in the lung microbiome given its modest abundance compared with the previously mentioned taxa, but it has been reported to have anti-inflammatory effects in COPD[35] and convey protection against viral infections.[36] This study therefore raises interesting questions around the role of manipulating the lung microbiome which go beyond the eradication of key pathogens to the broader consideration of correcting the dysbiosis associated with poor clinical outcomes.

Observational microbiome studies suffer from the weakness of identifying correlations between the microbiome and clinical features, and therefore carry with them the temptation to confuse correlation with causation. Changes in the microbiome are not necessarily just a function of disease progression, but also of environment, medication and lifestyle factors, which can confuse the analysis.[37] Management of COPD introduces

these same covariates, and importantly the recurrent use of antibiotics to treat exacerbations may have adverse consequences by driving loss of diversity, which may lead to greater risk of future exacerbations or indeed to disease progression itself. Additional functional studies of how the lung microbiome interacts with the host immune system and metabolic milieu of the lung will be necessary to translate models into robust biomarkers to better target treatments and to identify new therapeutic strategies aimed at the microbiome.

Author affiliations
[1]Computational Biology, Target Sciences, GSK R&D, King of Prussia, Pennsylvania, USA
[2]GSK Vaccines, Rixensart, Belgium
[3]Clinical and Experimental Sciences, University of Southampton Faculty of Medicine, Southampton General Hospital, Southampton, UK
[4]Southampton NIHR Respiratory Biomedical Research Unit, Southampton General Hospital, Southampton, UK
[5]Respiratory Therapy Area Unit, GSK R&D, King of Prussia, Pennsylvania, USA
[6]Target and Pathway Validation, Target Sciences, GSK R&D, King of Prussia, Pennsylvania, USA
[7]Wessex Investigational Sciences Hub, University of Southampton Faculty of Medicine, Southampton General Hospital, Southampton, UK

Acknowledgements We are grateful to Zhang Wang and Nicholas Locantore for their advice on the analysis of this study, to Judith Dupont and Sylvie Vanderschrick for their technical assistance, and to Yousef El Aalamat for his contribution to the statistical analysis of the study. The authors would like to also thank Geraldine Drevon and Regis Azizieh (XPE Pharma & Science, on behalf of GSK Vaccines) for coordination and editorial support.

Collaborators AERIS Study Group: J Alnajar, R Anderson, E Aris, WR Ballou, A Barton, S Bourne, M Caubet, SC Clarke, D Cleary, C Cohet, N Coombs, K Cox, J-M Devaster, V Devine, N Devos, E Dineen, T Elliott, R Gladstone, S Harden, J Jefferies, V Kim, S Mesia Vela, P Moris, K Ostridge, TG Pascal, M Peeters, S Schoonbroodt, KJ Staples, A Tuck, L Welch, V Weynants, TMA Wilkinson, AP Williams, N Williams, C Woelk, M Wojtas, S Wootton.

Contributors Conceived and designed the experiments: ND, JRB, GS, J-MD, TMW. Subject recruitment and sample processing: SC, VK, KO, KJS, ACT, NW. Performed the experiments: ND, RP, SVH. Data analysis: DM, ND, CL, MM-S, DFS, VW, J-MD. Data interpretation and writing: DM, ND, CL, JRB, MM-S, BEM, DFS, KJS, RT-S, VW, J-MD, TMW. The investigators obtained data and cared for the study participants. The authors had full access to all data in the study, contributed to the writing of the report and had final responsibility for the decision to submit for publication. All members of the AERIS Study Group were involved in the planning, conduct and/or reporting of the work described in the article.

Funding The study funder, GlaxoSmithKline Biologicals SA, designed the study in collaboration with the investigators, and coordinated collection, analysis and interpretation of data.

Competing interests TMW has received reimbursement for travel and meeting attendance from Boehringer Ingelheim and AstraZeneca, outside of the submitted work. SC received a grant from Pfizer, outside of the submitted work. KJS received grants from Asthma UK (08/026) and BMA HC Roscoe Award, outside of the submitted work, and he has a patent PCT/GB2010/050821 'Ex Vivo Modelling of Therapeutic Interventions' pending. BEM, CL, DFS, DM, GS, J-MD, JRB, ND, MM-S, RS, RT-S, SVH and VW are employees of the GSK group of companies. RP was an employee of the GSK group of companies at the time the study was conducted. BEM, JRB, J-MD, ND, RT-S and VW own shares/restricted shares in the GSK group of companies. KJS, VK, KO, ACT, SC and TMW received an institutional grant from the GSK group of companies to conduct this study.

Ethics approval Southampton and South West Hampshire Research Ethics Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Vestbo J, Hurd SS, Agustí AG, *et al*. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2013;187:347–65.
2 Roca J, Vargas C, Cano I, *et al*. Chronic obstructive pulmonary disease heterogeneity: challenges for health risk assessment, stratification and management. *J Transl Med* 2014;12 (Suppl 2):S3.
3 Agusti A, Calverley PM, Celli B, *et al*. Characterisation of COPD heterogeneity in the ECLIPSE cohort. *Respir Res* 2010;11:122.
4 Sethi S, Murphy TF. Infection in the pathogenesis and course of chronic obstructive pulmonary disease. *N Engl J Med Overseas Ed* 2008;359:2355–65.
5 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;184:662–71.
6 Dickson RP, Erb-Downward JR, Martinez FJ, *et al*. The microbiome and the respiratory Tract. *Annu Rev Physiol* 2016;78:481–504.
7 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:e16384.
8 Dickson RP, Erb-Downward JR, Huffnagle GB. The role of the bacterial microbiome in lung disease. *Expert Rev Respir Med* 2013;7:245–57.
9 Wang Z, Bafadhel M, Haldar K, *et al*. Lung microbiome dynamics in COPD exacerbations. *Eur Respir J* 2016;47:1082–92.
10 Sze MA, Dimitriu PA, Hayashi S, *et al*. The lung tissue microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012;185:1073–80.
11 Yadava K, Pattaroni C, Sichelstiel AK, *et al*. Microbiota promotes chronic pulmonary inflammation by enhancing IL-17A and autoantibodies. *Am J Respir Crit Care Med* 2016;193:975–87.
12 Staples KJ, Taylor S, Thomas S, *et al*. Relationships between mucosal antibodies, non-typeable *Haemophilus influenzae* (NTHi) infection and airway inflammation in COPD. *PLoS One* 2016;11:e0167250.
13 Huang YJ, Sethi S, Murphy T, *et al*. Airway microbiome dynamics in exacerbations of chronic obstructive pulmonary disease. *J Clin Microbiol* 2014;52:2813–23.
14 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;33:1101–11.
15 Molyneaux PL, Mallia P, Cox MJ, *et al*. Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2013;188:1224–31.
16 Bourne S, Cohet C, Kim V, *et al*. Acute exacerbation and respiratory infections in COPD (AERIS): protocol for a prospective, observational cohort study. *BMJ Open* 2014;4:e004546.
17 Wilkinson TMA, Aris E, Bourne S, *et al*. A prospective, observational cohort study of the seasonal dynamics of airway pathogens in the aetiology of exacerbations in COPD. *Thorax* 2017;72:919–27.
18 Zhang J, Kobert K, Flouri T, *et al*. PEAR: a fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics* 2014;30:614–20.
19 Caporaso JG, Kuczynski J, Stombaugh J, *et al*. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;7:335–6.
20 Galiana A, Aguirre E, Rodriguez JC, *et al*. Sputum microbiota in moderate versus severe patients with COPD. *Eur Respir J* 2014;43:1787–90.
21 Garcia-Nuñez M, Millares L, Pomares X, *et al*. Severity-related changes of bronchial microbiome in chronic obstructive pulmonary disease. *J Clin Microbiol* 2014;52:4217–23.
22 Pascoe S, Locantore N, Dransfield MT, *et al*. Blood eosinophil counts, exacerbations, and response to the addition of inhaled fluticasone furoate to vilanterol in patients with chronic obstructive pulmonary disease: a secondary analysis of data from two parallel randomised controlled trials. *Lancet Respir Med* 2015;3:435–42.
23 Sze MA, Dimitriu PA, Suzuki M, *et al*. Host response to the lung microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2015;192:438–45.
24 Avgousti DC, Herrmann C, Kulej K, *et al*. A core viral protein binds host nucleosomes to sequester immune danger signals. *Nature* 2016;535:173–7.
25 Buffie CG, Bucci V, Stein RR, *et al*. Precision microbiome reconstitution restores bile acid mediated resistance to clostridium difficile. *Nature* 2015;517:205–8.
26 Murphy TF, Kirkham C. Biofilm formation by nontypeable haemophilus influenzae: strain variability, outer membrane antigen expression and role of pili. *BMC Microbiol* 2002;2:7.
27 Kyd JM, McGrath J, Krishnamurthy A. Mechanisms of bacterial resistance to antibiotics in infections of COPD patients. *Curr Drug Targets* 2011;12:521–30.
28 Hillas G, Perlikos F, Tzanakis N. Acute exacerbation of COPD: is it the "stroke of the lungs"? *Int J Chron Obstruct Pulmon Dis* 2016;11:1579–86.
29 Woodruff PG, Agusti A, Roche N, *et al*. Current concepts in targeting chronic obstructive pulmonary disease pharmacotherapy: making progress towards personalised management. *Lancet* 2015;385:1789–98.
30 Bellinghausen C, Rohde GG, Savelkoul PH, *et al*. Viral-bacterial interactions in the respiratory tract. *J Gen Virol* 2016;97:3089–102.
31 Tangedal S, Aanerud M, Grønseth R, *et al*. Comparing microbiota profiles in induced and spontaneous sputum samples in COPD patients. *Respir Res* 2017;18:164.

32 Salter SJ, Cox MJ, Turek EM, *et al*. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 2014;12:87.
33 Murphy TF, Brauer AL, Sethi S, *et al*. Haemophilus haemolyticus: a human respiratory tract commensal to be distinguished from haemophilus influenzae. *J Infect Dis* 2007;195:81–9.
34 Pragman AA, Kim HB, Reilly CS, *et al*. The lung microbiome in moderate and severe chronic obstructive pulmonary disease. *PLoS One* 2012;7:e47305.
35 Mortaz E, Adcock IM, Ricciardolo FL, *et al*. Anti-inflammatory effects of *lactobacillus* rahmnosus and *bifidobacterium* breve on cigarette smoke activated human macrophages. *PLoS One* 2015;10:e0136455.
36 Gabryszewski SJ, Bachar O, Dyer KD, *et al*. *Lactobacillus*-mediated priming of the respiratory mucosa protects against lethal pneumovirus infection. *J Immunol* 2011;186:1151–61.
37 Forslund K, Hildebrand F, Nielsen T, *et al*. Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* 2015;528:262–6.