Comparison of the airway microbiota in children with chronic suppurative lung disease

Ahmed, Bushra; Cox, Michael J.; Cuthbertson, Leah; James, Phillip; Gardner, Laura; Cookson, William; Davies, Jane; Moffatt, Miriam; Bush, Andrew

DOI:
10.1136/bmjresp-2021-001106

BMJ Open Respiratory Research

Comparison of the airway microbiota in children with chronic suppurative lung disease

Bushra Ahmed,[1,2] Michael J Cox,[1] Leah Cuthbertson,[3] Phillip James,[1] Laura Gardner,[1,2] William Cookson,[3] Jane Davies,[2,4] Miriam Moffatt,[3] Andrew Bush[2]

To cite: Ahmed B, Cox MJ, Cuthbertson L, *et al*. Comparison of the airway microbiota in children with chronic suppurative lung disease. *BMJ Open Resp Res* 2021;**8**:e001106. doi:10.1136/bmjresp-2021-001106

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjresp-2021-001106).

Received 13 September 2021
Accepted 19 November 2021

[1]National Heart and Lung Institute, Imperial College London, London, UK
[2]Department of Respiratory Paediatrics, Royal Brompton Hospital, London, UK
[3]Genomic Medicine, Imperial College London, London, UK
[4]Gene Therapy, Imperial College London, London, UK

Correspondence to
Dr Bushra Ahmed;
Bushra.Ahmed1@nhs.net

ABSTRACT

Rationale The airway microbiota is important in chronic suppurative lung diseases, such as primary ciliary dyskinesia (PCD) and cystic fibrosis (CF). This comparison has not previously been described but is important because difference between the two diseases may relate to the differing prognoses and lead to pathological insights and potentially, new treatments.

Objectives To compare the longitudinal development of the airway microbiota in children with PCD to that of CF and relate this to age and clinical status.

Methods Sixty-two age-matched children (age range 0.5–17 years) with PCD or CF (n=31 in each group) were recruited prospectively and followed for 1.1 years. Throat swabs or sputum as well as clinical information were collected at routine clinical appointments. 16S rRNA gene sequencing was performed.

Measurements and main results The microbiota was highly individual and more diverse in PCD and differed in community composition when compared with CF. While *Streptococcus* was the most abundant genus in both conditions, *Pseudomonas* was more abundant in CF with *Haemophilus* more abundant in PCD (P_{adj}=0.0005). In PCD only, an inverse relationship was seen in the relative abundance of *Streptococcus* and *Haemophilus* with age.

Conclusions Bacterial community composition differs between children with PCD and those with CF. *Pseudomonas* is more prevalent in CF and *Haemophilus* in PCD, at least until infection with *Pseudomonas* supervenes. Interactions between organisms, particularly members of *Haemophilus*, *Streptococcus* and *Pseudomonas* genera appear important. Study of the interactions between these organisms may lead to new therapies or risk stratification.

Key messages

- The airway microbiota is highly individual and differs between PCD and CF. Study of the different interactions between components of the microbiota, especially *Haemophilus*, *Streptococcus* and *Pseudomonas*, may potentially lead to new therapies.
- This is the first study to describe the longitudinal development of the airway microbiota in children with PCD, and how it differs from CF to help to further our understanding of microbiological processes when bacterial elimination is compromised.

INTRODUCTION

Airway infections in chronic suppurative lung diseases (CSLD), such as primary ciliary dyskinesia (PCD) and cystic fibrosis (CF), are important. Previously, based on bacterial cultures, it was suggested that airway infections in CSLD are caused by only a handful of pathogens. Bacterial sequencing, however, has identified rich bacterial communities ('microbiota') in the airways of patients with CSLD. This includes 'traditional' pathogens, such as *Pseudomonas aeruginosa*, as well as many other micro-organisms particularly anaerobes. The role of these polymicrobial communities in disease progression is at present unclear.[1–3]

The relationship between airway bacterial diversity and disease severity has received much attention. CF has been relatively well studied, with several longitudinal studies in adults and large, cross-sectional studies across all ages. These have described an inverse relationship between diversity and spirometry (forced expired volume in 1 s (FEV_1)) with an increase in diversity seen at least until school age[4] and a decrease in adulthood.[5–11] Antibiotic use was found to be a primary driver of reduced diversity in adults,[8] suggesting perhaps antibiotics should be used more judiciously. Several longitudinal studies of the airway microbiota in infants with CF have however demonstrated either a decrease[12 13] or no change in diversity with age.[14–16] In infants and preschool aged children, diversity was not associated with FEV_1 at 6 years[13] or antibiotic usage.[12] Whether reduced diversity is a marker of end-stage disease, an effect of the frequent courses of antibiotics that many patients with CF receive, or an innocent bystander, has yet to be determined.

Longitudinal studies in infants with CF have demonstrated changes in community composition with age.[14 17] Previously, we

described a significant inverse relationship between *Streptococcus* and *Haemophilus* in oropharyngeal samples until 2 years of age in infants with CF diagnosed on newborn screening (NBS).[16] Similarly, a comparison of bronchoalveolar lavage (BALF) between clinically stable children with CF and controls aged between one and 6 years of age demonstrated a trend towards an increase in the abundance of Proteobacteria (which includes *Pseudomonas, Neisseria* and *Haemophilus*) and a significant decrease in *Streptococcus* and other anaerobes with age in CF.[12] The clinical significance of this relationship is unknown but there is evidence of a significant reduction in airway inflammation (BALF total cell and neutrophil counts) with an increased relative abundance of anaerobes *Streptococcus, Veillonella* and *Prevotella*.[18] Due to high interpatient variability,[6 18] longitudinal studies coupling FEV_1 with diversity of the airway microbiota in children have the potential to clarify this relationship.

The airway microbiota is thought to be governed by three factors: bacterial immigration (eg, through microaspiration from the oropharynx and seeding from the upper airways), elimination (eg, through mucociliary clearance) and regional growth conditions.[19] PCD is caused by genetically inherited dysfunction of motile cilia. It shares certain pathological features with CF including impaired mucociliary clearance, chronic airway infections and neutrophilic airway inflammation.[20] Important differences, however, exist between these two conditions, including differences in sputum composition, which may influence regional growth conditions and subsequent airway microbiology. *Haemophilus influenzae* is frequently cultured in CF and PCD, but *Streptococcus pneumoniae* is more prevalent in PCD while *P. aeruginosa* is more prevalent in CF.[21] PCD however, for unknown reasons, rarely leads to as steep a decline in spirometry as CF (FEV_1 annual decline 0.8% in PCD and 3.6% in CF[21]). There is much less information about the longitudinal development of the microbiota in PCD. To date there is only a single cross-sectional published study of the microbiota in 24 patients with PCD across a wide age range (between 4 and 73 years of age).[22] Understanding the variation in the microbiota between these two forms of CSLD is important to further our understanding of microbiological processes when bacterial elimination is compromised and may lead to development of novel therapies.

In this study, we used sputum or throat swabs (TS), the latter which we have previously confirmed are a surrogate for the lower airway,[23] to describe the airway microbiota longitudinally in children with CSLD. We hypothesised that the microbiota would differ between PCD and CF and would show differences between the two disease with age and clinical status.

METHODS

Subjects and sampling

Children were recruited and followed up opportunistically during routine clinical appointments (on average every 6 months for PCD and 3 months for CF) at the Royal Brompton Hospital (RBH) between December 2012 and March 2014. The inclusion criteria were regular attendance at RBH for clinical appointments and diagnoses of PCD and CF made in accordance with standard guidelines[24–26] (see online supplemental file 1). The exclusion criteria were patients who were receiving CFTR modulating medications (of which only ivacaftor was available at the time of the study).

Sputum samples, or TS in non-expectorating patients, were collected alongside detailed clinical information and handled as previously described (see online supplemental file 1 and online supplemental table S1).[23] Where possible, paired sputum samples and TS were collected to compare the microbiota between these sample types. Technical control samples using blank swabs were collected to test for contamination at the point of sampling. Simultaneous microbial culture was performed as per standard clinical practice according to the CF Trust Guidelines[27] in the Clinical Microbiology Department at RBH.

Pulmonary exacerbations were defined as a change in clinical features necessitating treatment with intravenous antibiotics as determined by an independent clinician. *P. aeruginosa* chronic status was defined according to the Leeds criteria.[28]

Patients or the public were not involved in the design, conduct, reporting or dissemination of our research.

16S rRNA gene library preparation and sequencing

For DNA extraction, whole sputum samples were defrosted and 300 µl added directly to a Lysing Matrix E (LME) tube. For TS, frozen swab heads were transferred directly to an LME tube containing sodium phosphate buffer. Further DNA extraction steps, 16S rRNA gene V4 region library preparation and sequencing (using the Illumina MiSeq V2 reagent kit) were performed as described previously.[23]

Data analysis

Sequence processing was performed in Quantitative Insights into Microbial Ecology (QIIME) V.1.9.0.[23] After combining forward and reverse reads, de-multiplexing, quality filtering and removal of PhiX, 35 923 028 reads were obtained with a median of 18 324 reads per patient sample (range 82–348 657 reads). Operational taxonomic unit (OTU) picking was performed in UCLUST (V.1.2.22q) with a threshold of 97% sequence similarity.[29] *Undibacterium* spp, *Comamonadaceae* and an individual *Chlamydophila* OTU (ID 4987) were found to be highly abundant in control samples and were removed as contaminants[30] (online supplemental figure S1). Samples were rarefied to 1000 reads to standardise sequencing depth resulting in 76 (18.6%) of the 409 samples being removed prior to further data analysis (online supplemental figure S2). The composition of mock communities was compared with ensure sequencing consistency

between sequencing runs (online supplemental figure S3).

Downstream analyses were performed using Phyloseq (V.1.20.0)[31] in R (V.3.4.0).[16] Barplots were constructed using ggplot2 (V.2.2.1)[32] and gridExtra (V.2.2.1).[33] Non-linear mixed effects modelling using a negative binomial distribution was performed using glmmADMB (V.0.8.3.3)[34] controlling for patient to assess the relationship between age range (1–2 years, 2–3 years, etc) and (a) alpha (within subject) diversity, measured by richness, evenness, Shannon and Inverse Simpson's diversity indices (see online supplemental file 1) and (b) the relative abundance of the most common genera.

Beta (between subject) diversity differences (Bray Curtis dissimilarity) were tested using a permutational multivariate analysis of variance (PERMANOVA)[35] blocked by participant study number, using the adonis function in vegan (V.2.4.3).[35] OTU level changes were assessed using multiple correlation testing using Spearman's rank with a false discovery rate correction (Benjamini-Hochberg) and adjusted p values reported.

To assess the relationship between clinical variables (online supplemental table S1) and alpha diversity, t-tests and Wilcoxon signed-rank tests were used for parametric and non-parametric binary variables and ANOVAs and Kruskal-Wallis test similarly for variables with multiple values. Pearson correlation was used to test changes in continuous variables and alpha diversity. A PERMANOVA was used to test the relationship between beta diversity and clinical variables within each patient group. All variables were initially tested independently to determine which were significant. Age was analysed as a categorical variable (as age ranges, see above).

Significant variables were then tested in a combined model. For CF, the variables included in the final model were: *P. aeruginosa* chronic infection status at baseline, growth of *P. aeruginosa* during the study, azithromycin prophylaxis and treatment courses (oral co-trimoxazole, nebulised colomycin and nebulised tobramycin). For PCD, the variables included were gender and *H. influenzae* growths. A p value <0.05 was considered statistically significant. Sequence data is available at the ENA Accession number: PRJEB26618.

Agreement between microbial cultures and 16S rRNA gene sequencing was assessed by comparing the frequency with which organisms grown on culture were also the most abundant organism identified by molecular techniques and calculating Cohen's kappa statistic.

RESULTS

Patient demographics

Each disease group contained 31 children. Baseline demographics are given in table 1. Follow-up was for a mean of 1.1 years for both groups (SD 0.48 and 0.46 years, PCD and CF, respectively). Medians of five (PCD, range 2–15) and seven (CF, range 2–23) samples per subject were collected (online supplemental figure S2 and S4).

Clinical microbiological cultures during the study period differed between PCD and CF (table 2). *S. pneumoniae* (p=0.01) and *H. influenzae* (p<0.0001) were cultured more frequently in PCD, with *P. aeruginosa* (p=0.03) and *Aspergillus fumigatus* (p=0.001) being more frequent in CF.

Comparing TS and sputum microbiota

Sixteen patients had paired sputum and TS samples (online supplemental figures S5 and S6). There was no significant difference in alpha or beta diversity or genera. A subanalysis of longitudinal data was performed using sputum samples only. Although there was a trend towards greater relative abundance of *Pseudomonas* in sputum (mean 4.0% (SD 13.5%)) in sputum c.f. 0.5% (SD 0.8%) in TS, P_{adj}>0.05), overall the trends in the most common genera, alpha and beta diversity remained unchanged (online supplemental file 1). Both TS and sputum were therefore used for longitudinal comparisons of PCD with CF. TS was analysed, where available, and sputum if no TS sample had been collected at that timepoint.

Differences in the airway microbiota between PCD and CF

Streptococcus was the most common genus in both PCD and CF (31.3% and 55.0% of total reads, respectively, p<0.0001, P_{adj}<0.0001) (figure 1, online supplemental figure S7 and table 3). *Haemophilus* was notably more common in PCD (20.5%) than CF (3.2%, p<0.0001, P_{adj}<0.0001). *Pseudomonas* was the second most common genus in CF (8.1%) but was less prevalent in PCD (1.2%, p=0.0001, P_{adj}=0.01). Other common genera were similar between PCD and CF except *Staphylococcus* which was uncommon (<0.1%) in PCD (p=0.015, P_{adj}>0.05).

Concordance between bacterial cultures and molecular genus identification was better with PCD than with CF (tables 2 and 3). *Haemophilus* was the most commonly cultured organism in PCD and showed moderate agreement with molecular identification (Cohen's kappa 0.43). *Pseudomonas* was the most commonly cultured organism in CF and showed fair agreement with molecular identification (Cohen's kappa 0.38). Alpha diversity was higher in PCD than CF (PCD, Shannon diversity median 2.12 (range 0.190–3.137) and 1.94 (CF, range 0.167–3.158), p=0.04) (figure 2). Combining all samples (all patients and timepoints), diagnosis accounted for 5.5% of variance in community structure (p=0.001, online supplemental figure S8).

Changes in the airway microbiota with age

In PCD, in the majority of patients, *Streptococcus* had a high relative abundance in early childhood in contrast to relative abundance of *Haemophilus* that showed a low relative abundance (66.7% vs 1.3% at 0.5 years, online supplemental figure S9). With age, there was an inverse relationship between the relative abundances of *Streptococcus* and *Haemophilus: Streptococcus* decreased until 5 years of age

Table 1 Baseline patient demographics comparing CF and PCD (n=31 in each group)

Demographic	CF	PCD	P value
Age (years)	9.3 (4.8)	9.9 (4.6)	NS
Gender (female)	18 (58%)	20 (65%)	NS
BMI (z-score)	−0.222 (0.969)	−0.149 (1.29)	NS
Median FEV_1 (%)	80.0 (40–119)	79.5 (39–123)	NS
Median FVC (%)	87.0 (40–132)	87.0 (43–111)	NS
CFTR genotype p.Phe508del homozygous p.Phe508del heterozygous* Other	17 (55%) 8 (26%) 6 (19%)	N/A	N/A
Ciliary structure ODA and IDA defect ODA defect IDA and microtubular defect Partial IDA defect Normal electron microscopy	N/A	9 (29%) 8 (26%) 6 (19%) 1 (3%) 7 (23%)	N/A
GORD	13 (42%)	1 (3%)	**0.0003**
P. aeruginosa infection status Chronic Intermittent Free Never	6 (19%) 6 (19%) 13 (42%) 6 (19%)	0 (0%) 2 (6%) 4 (13%) 25 (81%)	**<0.0001**
Exacerbations	11 (35%)	14 (45%)	**NS**
Inhaled steroids	11 (35%)	8 (26%)	**NS**
Antibiotic prophylaxis ▶ Azithromycin ▶ Co-amoxiclav ▶ Flucloxacillin	25 (81%) 14 (45%) 3 (10%) 8 (26%)	18 (58%) 17 (55%) 1 (3%) 0 (0%)	**NS** **NS** **NS** **0.01**

Values are given as mean (SD) or number (percentage, %). Significant differences between groups (p<0.05) are shown in bold.
Prophylactic antibiotic doses used: Azithromycin (given once daily, three times a week):if patient weight <15 kg, 10 mg/kg; if 15–40 kg, 250 mg;if >40 kg, 500 mg. Co-amoxiclav 400/57 (given two times a day): if patient aged 2 months–2 years, 0.15 mL/kg; if 2–6 years, 2.5 mL;if 7–12 years, 5 mL. Flucloxacillin (given two times a day): if patient aged <3years, 125 mg; if >3 years, 25 mg/kg.
*Where known, all non-p.Phe508Del genes were class I–III (minimal function mutations). Full list given in online supplemental file 1.
BMI, body mass index; FEV_1, forced expired volume in 1 s; FVC, forced vital capacity; GORD, gastro-oesophageal reflux disease; IDA, inner dynein arm; N/A, not applicable; NS, non-significant; ODA, outer dynein arm.

(17.2%, p=0.003) while *Haemophilus* increased (55.8%, p=0.005). The reverse was seen between 6 and 9 years of age. *Streptococcus* then decreased in relative abundance between the ages of 10 and 17 years (10.2%, p=0.03 at 17 years) while *Haemophilus* increased (79.3%, p=0.003 at 17 years). The relative abundance of *Neisseria* fluctuated peaking at 12 years of age (relative abundance 0.1% and 17% at 0.5 and 12 years, respectively, p<0.0001). No significant changes were seen in the relative abundances of *Prevotella* or *Veillonella*. There were no significant changes in the relative abundance of the five most common genera with age in CF (p>0.05) (figure 3).

In CF, at OTU level two OTUs were found to have a small (r≤0.3) but significant (P_{adj}<0.05) change in relative abundance with age: *Haemophilus* (OTU ID 7005) decreased with age (r=−0.254, P_{adj}=0.011) while *Actinomyces* (OTU ID 2413) increased (r=0.237, P_{adj}=0.046). In PCD, two OTUs changed significantly with age: *Streptococcus* (OTU ID 10797), decreasing with age (r=−0.323, P_{adj}=0.002) and *Haemophilus* (OTU ID 5932), increasing (r=0.244, P_{adj}=0.011).

Only alpha diversity changed significantly with age (both diseases). Richness increased in early childhood, peaking at age 7 years for PCD (mean richness 33.5 at 0.5 years, 67.5 at 7 years, p=0.006) and age 4 years for CF (mean richness 29 at 0.5 years, 68 at 4 years, p=0.034, figure 4). There was no significant difference in beta diversity with age (p>0.05) for either group.

From examining individual patient barplots (online supplemental figure S9), while the airway microbiota was highly individual, *Streptococcus* were abundant throughout for both CF and PCD. *Pseudomonas* was abundant in several children with CF aged above 8 years. The changes in relative abundance of *Haemophilus* in PCD at an individual level were consistent with the age-related changes described overall for PCD above. There were no consistent changes in genera seen with changes in symptoms, spirometry, growth on bacterial cultures or antibiotic use.

Table 2 Organisms grown on clinical microbiology during study period for patients with cystic fibrosis (CF) and primary ciliary dyskinesia (PCD)

Cultured organism	CF	PCD	P value
P. aeruginosa	10 (32%)	3 (10%)	**0.03**
Mucoid *P. aeruginosa*	3 (10%)	0 (0%)	NS
MSSA	4 (13%)	4 (13%)	NS
MRSA	1 (3%)	0 (0%)	NS
Streptococcus pneumoniae	0 (0%)	6 (19%)	**0.01**
H. influenzae	0 (0%)	14 (45%)	**<0.0001**
Stenotrophomonas maltophilia	1 (3%)	1 (3%)	NS
Serratia marcescens	2 (6%)	0 (0%)	NS
NTM	2 (6%)	0 (0%)	NS
'URT flora'	1 (3%)	23 (74%)	**<0.0001**
Moraxella catarrhalis	0 (0%)	1 (3%)	NS
Coliforms	1 (3%)	1 (3%)	NS
Aspergillus fumigatus	9 (29%)	0 (0%)	**0.001**

Frequency of organisms grown in each group given as number of patients (percentage frequency, %). Clinical microbiology was performed on either oropharyngeal swabs or spontaneously expectorated sputum samples, collected as part of routine clinical care. Fisher's exact test used to compare proportions of positive cultures between CF and PCD for each organism. Significant differences (p<0.05) are shown in bold. NS – non-significant (p>0.05). *H. influenzae*, *Haemophilus influenzae*; MRSA, methicillin resistant *S. aureus*; MSSA, methicillin sensitive *S. aureus*; NTM, non-tuberculous mycobacterium; *P. aeruginosa*, *Pseudomonas aeruginosa*; URT, upper respiratory tract.

There was no apparent skewing by, for example, an individual who contributed many samples to the study.

Influence of clinical variables, pulmonary exacerbations and antibiotics

The airway microbiota was highly individual, with the patient sampled accounting for 37% (PCD) and 39% (CF) of variance in community structure (p≤0.01).

For PCD, growths of *H. influenzae* accounted for 5.7% of variance in community structure (p=0.001). Alpha diversity was reduced in *H. influenzae* culture positive samples compared with those that were culture negative (median Inverse Simpson's diversity index 3.23 (range 1.10–15.9) and 5.16 (range 1.11–19.04), respectively, p=0.009). Gender had a weak relationship with community structure (r^2=0.046, p=0.001) but none with alpha diversity (p>0.05). There was no significant relationship between alpha and beta diversity and lung function (FEV_1 or forced vital capacity % predicted) for either disease.

For CF, all alpha diversity indices were higher in patients who were intermittently infected with *P. aeruginosa* infection compared with those who had never grown or had chronic *P. aeruginosa* at baseline (n=6 in each group, p=0.02, online supplemental table S3). Growth of *P. aeruginosa* during the study associated with reduced alpha diversity (p=0.01). Alpha diversity was higher in heterozygous p.Phe508del patients with CF than those who were homozygous or had other *CFTR* mutations (p=0.009). In the combined PERMANOVA, only *P. aeruginosa* chronic infection status had a small but significant relationship with community structure (r^2=0.017, p=0.014).

Fourteen patients with PCD and 11 patients with CF experienced a pulmonary exacerbation (necessitating intravenous antibiotics) of whom 12 PCD and 7 patients with CF had sequential samples collected at baseline (B), exacerbation (E), treatment (T) and recovery (R) as previously described.[16] Significant differences were seen only in PCD (see online supplemental file 1 and online supplemental figures S10 and S11).

There were no significant differences with oral or intravenous treatment antibiotic courses, prophylactic or nebulised antibiotics for either condition.

DISCUSSION

Here, we report for the first time the development of the airway microbiota in children with PCD, and compare it to that in a similar but more severe disease, CF.

Our study shows that the airway microbiota is highly individual but differs between groups of children with PCD and CF, a novel finding. Several previous studies in CF have also demonstrated high interpatient variability in the airway microbiota,[6 18] strengthening arguments for personalised medicine. There was greater diversity in PCD than CF with a significant difference in both alpha and beta diversity. Since most patients with PCD studied had genotypes associated with a milder phenotype, this could suggest increased diversity of the microbiota is associated with milder disease. The significance of this is unclear, however, given that the median FEV_1% and hence one marker of disease severity was similar between both groups.

Interpretation of diversity with disease severity is complex. Changes in diversity reflect changes in the relative abundance of organisms. Samples with an even spread of different organisms have greater diversity whereas those dominated by an individual organism have lower diversity. Therefore, changes in diversity may reflect presence or absence of a dominant pathogen, which drives disease severity.

There was no significant relationship between diversity (alpha or beta) and spirometry for either disease. One interpretation is that diversity of the microbiota does not influence disease severity. Disease severity in both groups was however relatively mild, reducing the ability to detect such a relationship. Furthermore, there was no significant difference in FEV_1 between PCD and CF with the median FEV_1 80% predicted in both groups. This is slightly lower than the median for our centre (RBH) from the UK CF Registry data for that year (FEV_1 86.5% predicted in 2013) and a selection bias cannot be excluded. Thus, whether differences in the microbiota between PCD and CF relate to differences in disease prognosis could not be fully determined. FEV_1, however, is not a sensitive marker for tracking particularly early lung disease in children[36] and may be difficult under 5 years of age. In future studies, relating more sensitive tests of distal airway disease such as lung clearance index, forced

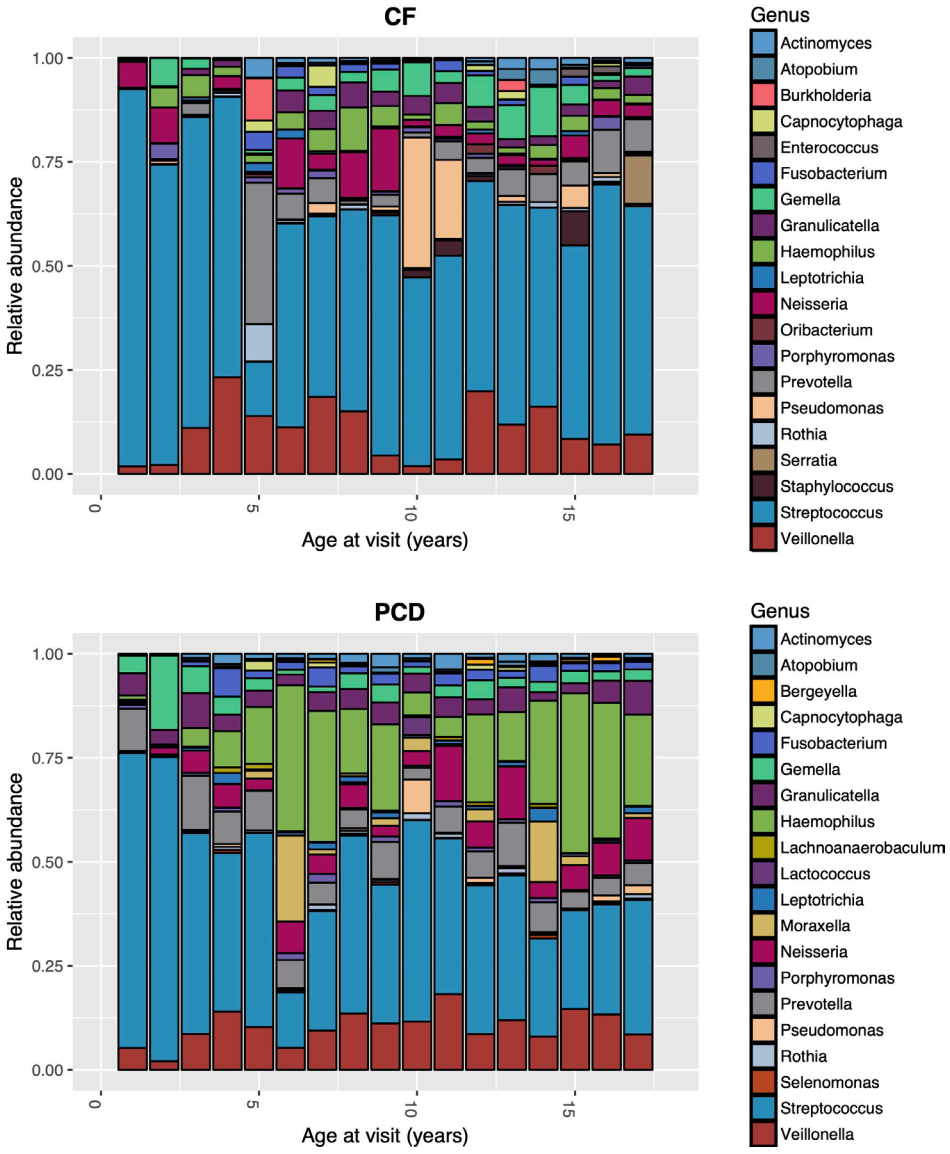

Figure 1 Stacked barplot illustrating changes in the relative abundance of the 20 most common genera with age in (A) CF and (B) PCD. Each bar shows the mean relative abundance of genera identified for all patients within each age range by patient group. This figure shows that *Streptococcus* was the most abundant genus in both CF and PCD throughout childhood. *Pseudomonas* was more abundant in CF (non-significant, p>0.05) while *Haemophilus* was more abundant in PCD (P_{adj}<0.0001). CF, cystic fibrosis; PCD, primary ciliary dyskinesia.

oscillation or imaging to airway microbiota diversity could potentially be more revealing.

An inverse relationship was seen between *Streptococcus* and *Haemophilus* with age only in PCD. This trend parallels that seen in CF NBS infants in the first 2 years of life.[16] Notably the microbiota in PCD and CF NBS appears to have greater similarities than with older children with CF. We speculate the microbiota in PCD is like that of the early microbiota in CF. Whether this is of direct pathological significance, or a marker of another underlying disease process is unclear.

Previous cross-sectional studies in CF identified an increase in diversity during childhood, at least until school age,[4] and a decrease in adulthood.[6 7] In our study, with the exception of species richness (the number of different organisms), there was no significant difference with age in either alpha or beta diversity in either disease. Many of the children in this study were on prophylactic antibiotics (81% in CF and 58% in PCD). While a treatment effect of prophylactic antibiotics was not seen in either group, given the high proportion of patients receiving prophylaxis, particularly in CF, it is possible that this suppressed changes in the microbiota over time. There were significant differences, however, at genera level: *Streptococcus* was the most common genus in both groups (31% PCD and 55% CF); *Haemophilus* was sixfold more abundant in PCD than CF (21% vs 3.2%), whereas *Pseudomonas* was the second most common genus (8.1%) in CF but not PCD. Several anaerobes were prevalent in PCD as also reported for CF here and elsewhere.

Table 3 Genera comparison (mean % relative abundance) between cystic fibrosis (CF) and primary ciliary dyskinesia (PCD) for all samples at all timepoints

Genus	CF (%)	PCD (%)	P value	Adjusted p value
Streptococcus	55	31.3	<0.0001	<0.0001
Haemophilus	3.2	20.5	<0.0001	<0.0001
Pseudomonas	8.1	1.2	0.0001	0.012
Veillonella	8.1	10.4	0.018	NS
Granulicatella	4.2	4	0.771	NS
Gemella	4.2	2.8	0.039	NS
Neisseria	3.8	6.3	0.004	NS
Prevotella	3.3	5.2	0.016	NS
Rothia	1.5	0.6	0.128	NS
Staphylococcus	1.1	<0.1	0.015	NS
Fusobacterium	1.1	2.3	0.003	NS
Moraxella	0.5	3.3	0.044	NS

Multiple t-tests with a Bonferroni correction were used to test differences in relative abundance of genera between CF and PCD. NS—non-significant (p>0.05).

There was reduced diversity with growth of *H. influenzae* in PCD and *P. aeruginosa* in CF, suggesting these organisms have important influences on community structure. Previous studies in PCD have demonstrated that while *H. influenzae* is the most common organism cultured in children, *P. aeruginosa* becomes the dominant pathogen with age.[37] This increase in *P. aeruginosa* occurs later in PCD than in CF, predominantly after 30 years of age, when transition to a mucoid phenotype also occurs.[38] A cross-sectional study of non-CF bronchiectasis found a competitive effect between *Haemophilus* and *Pseudomonas;* when *Haemophilus* were dominant, *Pseudomonas* were either absent or present in very low

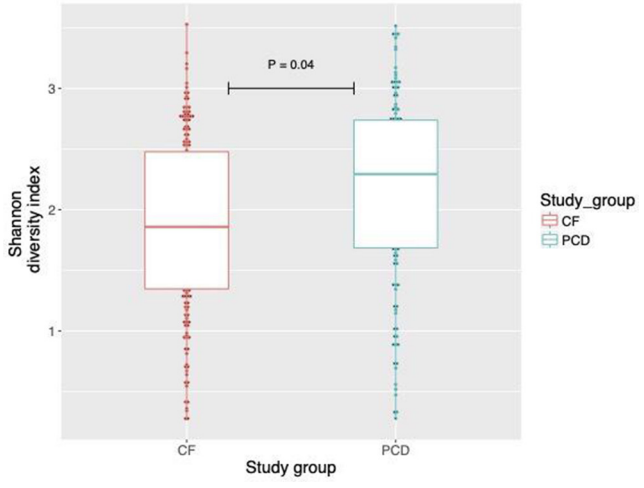

Figure 2 Boxplot comparing alpha (Shannon) diversity between CF (N=171) and PCD (N=129). Each boxplot shows the median and IQR for the Shannon diversity index for each group. The boxplot shows that Shannon diversity was higher in PCD than CF (p=0.04). CF, cystic fibrosis; PCD, primary ciliary dyskinesia.

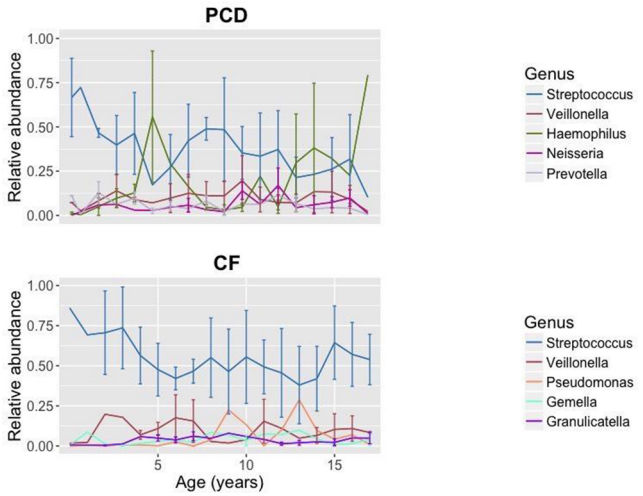

Figure 3 Changes in the relative abundance of the five most common genera with age throughout childhood in CF and PCD. The five most common genera were identified as those with the highest relative abundance across all samples and all time points. (A) Trends in the mean relative abundance with SE bars of *Streptococcus*, *Haemophilus*, *Veillonella*, *Prevotella* and *Neisseria* with increasing age until 17 years of age in PCD. An inverse relationship was seen with age between *Streptococcus* and *Haemophilus* (p≤0.005). (B) Trends in the mean relative abundance with SE bars of *Streptococcus*, *Pseudomonas*, *Veillonella*, *Granulicatella* and *Gemella* with increasing age until 17 years of age in CF. All changes were non-significant (p>0.05). CF, cystic fibrosis; PCD, primary ciliary dyskinesia.

abundance and vice versa.[39] Possibly the high abundance of *Haemophilus* in the early CF microbiota and PCD leads to a milder phenotype, with replacement of *Haemophilus* by *Pseudomonas* in CF resulting in a more severe phenotype. Interventions aimed at preventing this switch, focussing on the mechanisms whereby *Haemophilus* is associated with improved prognosis in CF, are needed to further explore this potential relationship and hopefully delineate new therapeutic options.

This study has several strengths. It is the first study of the airway microbiota in children with PCD followed up for over a year. The children with PCD and CF in this study were well matched both by age and disease severity and were sampled frequently (median interval 61 and 42 days in PCD and CF, respectively). This has allowed meaningful comparisons and a detailed investigation of the microbiota to be made. The longitudinal design is important to account for the high interpatient variability in the microbiota seen both in this study and previously.

One limitation is that different sampling methods (TS and sputum) were used at different timepoints. Ideally the same sample type would have been collected at all times, but this proved unfeasible. Comparison of paired TS and sputum (n=16) did not demonstrate a difference in the microbiota and repeating the longitudinal analyses with sputum samples only did not change any conclusions. The relative abundance of *Pseudomonas* was lower in TS than in sputum samples as in previous

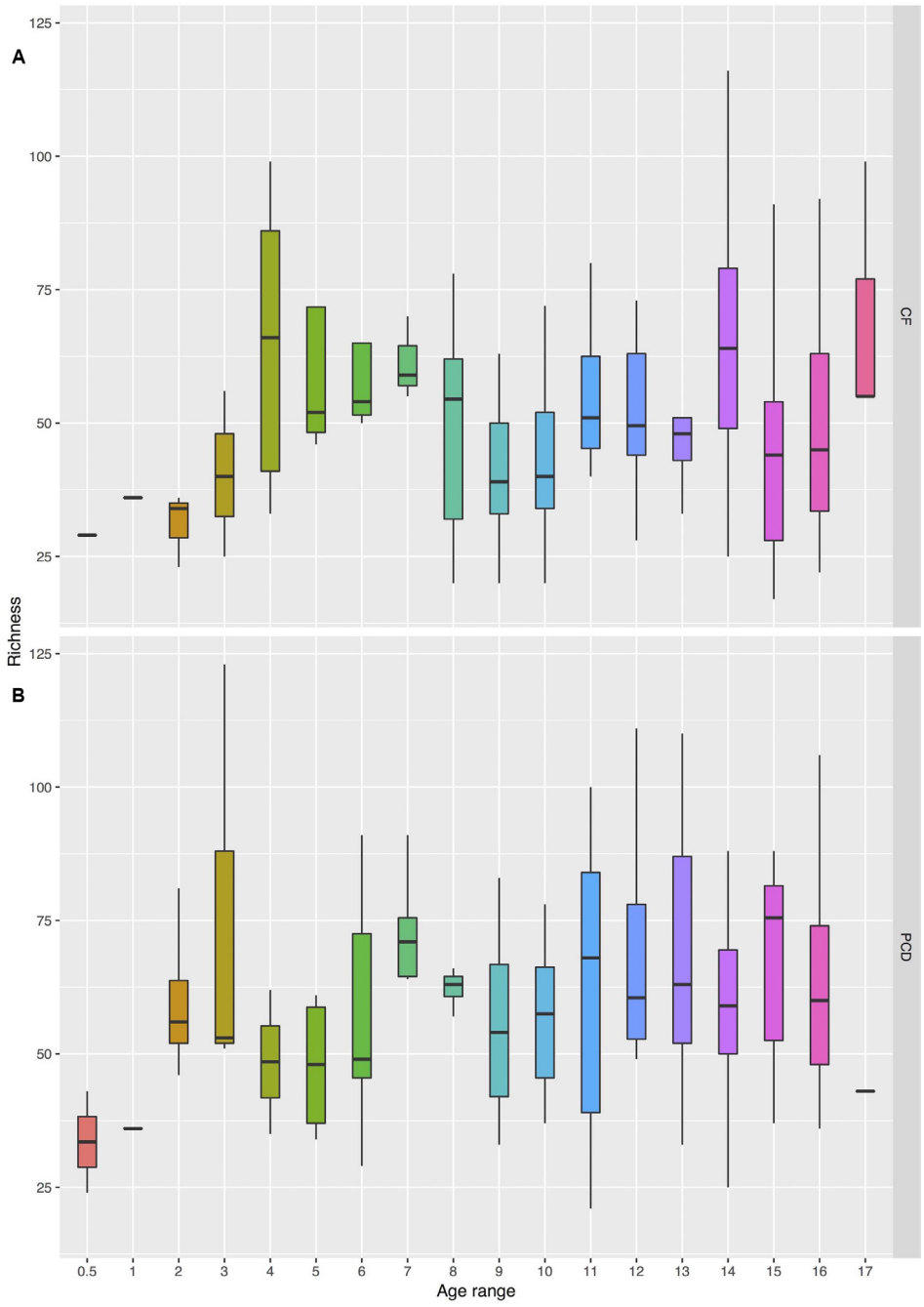

Figure 4 Boxplots comparing changes in richness with age between CF and PCD. (A) Changes in richness with age in CF. This shows an increase in richness until 4 years of age (p=0.034) with largely little change in richness thereafter. (B) Changes in richness with age in PCD. This shows an increase in richness until 7 years of age (p=0.006) followed by a decrease until 10 years of age before a gradual increase during adolescence (p=0.01). CF, cystic fibrosis; PCD, primary ciliary dyskinesia.

studies.[40 41] *Pseudomonas*, however, remained the second most abundant organism whether analysing sputum alone or both sputum and TS combined and the overall trends in community composition and diversity remained unchanged.

While TS are an attractive candidate as a surrogate for lower airway sampling, their accuracy in representing the lower airway microbiota remains debatable. A study of 16 children with CF demonstrated that for many participants, communities were similar between oropharyngeal samples and sputum. In those samples that diverged, there was a higher relative abundance of *Pseudomonas*, *Staphylococcus*, *Enterobacteriaceae* and *Haemophilus* in sputum samples.[41] Similarly, a study of 20 children with CF found high concordance between TS and sputum samples in patients who were *Pseudomonas* negative.[40] In the current study, 94% of patients with PCD and 61% of patients with CF were either 'free' or had never grown *Pseudomonas* at baseline. This may explain why similarity was seen when comparing paired sputum and TS from the children in this study. Previously we have shown

that TS correlate with the lower airway microbiota. Although differences exist precluding their use for individual decision making, TS can distinguish between disease groups and are useful for longitudinal study of the microbiota between groups of patients with CSLD.[23]

More children with CF had gastro-oesophageal reflux disease (GORD) requiring treatment with proton pump inhibitors (PPIs) than children with PCD. GORD and specifically treatment with PPIs may affect the lung microbiota.[42] Thus, it is possible that the differences between CF and PCD in this study could relate to GORD. That said we detected no effect of treatment with PPIs.

One limitation of 16S rRNA gene sequencing is that it is often unable to discern species level changes and cannot assess functional changes. It is plausible that while overall community composition may not have changed, functional changes occurred in gene expression, metabolite production or interactions between different species with disease status.[43] To this end, analysing longitudinal differences in CSLD with disease status using metagenomics or other functional approaches, as well as including digital droplet PCR to distinguish between absolute and relative abundance changes, would be an important area for future work.

In conclusion, we have demonstrated that the PCD microbiota differs from that of CF. The airway microbiota is highly individual in these diseases, further strengthening arguments for personalised approaches to patient management. Diversity was higher in PCD and community composition differed. There is a greater similarity between PCD and early CF diagnosed on NBS than with later established CF. *Pseudomonas* is more prevalent in CF contrasting with *Haemophilus* in PCD. An inverse relationship was seen between *Streptococcus* and *Haemophilus* with age in PCD but not CF. A similar relationship has been seen previously in NBS infants with CF.[16] This could suggest a switch from a less pathogenic to a more pathogenic community composition later in childhood in CF, potentially influenced by a competitive relationship between *Haemophilus* and *Pseudomonas*. Further study to better understand this relationship may lead to new therapeutic approaches.

Contributors BA planned the project, recruited participants, designed and performed experiments, analysed the data, wrote the manuscript and is the guarantor for the overall content; MJC planned the project, designed experiments, analysed the data and wrote the manuscript; LC and PJ designed experiments, analysed the data and wrote the manuscript; LG collected data and wrote the manuscript; WC, JD, MM and AB planned the project, designed experiments, analysed the data and wrote the manuscript. All authors reviewed, revised and approved the manuscript for submission.

Funding The authors thank Professor Claire Hogg for helpful discussions and her assistance in patient recruitment and the NIHR Biomedical Research Unit at Royal Brompton and Harefield NHS Foundation Trust. AB was supported by the NIHR Respiratory Disease Biomedical Research Unit at the Royal Brompton and Harefield NHS Foundation Trust and Imperial College London. AB is an emeritus NIHR Senior Investigator. BA was supported by NIHR. MM and WC were supported by the Asmarley Trust and Wellcome Trust. JD is an NIHR Senior Investigator.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the RBH Biomedical Research Unit Advanced Lung Disease Biobank (NRES reference 10/H0504/9). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Sequence data is available at the ENA Accession number: PRJEB26618.

REFERENCES

1 Tunney MM, Field TR, Moriarty TF, *et al*. Detection of anaerobic bacteria in high numbers in sputum from patients with cystic fibrosis. *Am J Respir Crit Care Med* 2008;177:995–1001.
2 van der Gast CJ, Walker AW, Stressmann FA, *et al*. Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *Isme J* 2011;5:780–91.
3 Rogers GB, Carroll MP, Serisier DJ, *et al*. Bacterial activity in cystic fibrosis lung infections. *Respir Res* 2005;6:49.
4 Muhlebach MS, Zorn BT, Esther CR, *et al*. Initial acquisition and succession of the cystic fibrosis lung microbiome is associated with disease progression in infants and preschool children. *PLoS Pathog* 2018;14:e1006798.
5 Cuthbertson L, Walker AW, Oliver AE, *et al*. Lung function and microbiota diversity in cystic fibrosis. *Microbiome* 2020;8:45.
6 Coburn B, Wang PW, Diaz Caballero J, *et al*. Lung microbiota across age and disease stage in cystic fibrosis. *Sci Rep* 2015;5:10241.
7 Cox MJ, Allgaier M, Taylor B, *et al*. Airway microbiota and pathogen abundance in age-stratified cystic fibrosis patients. *PLoS One* 2010;5:e11044.
8 Zhao J, Schloss PD, Kalikin LM, *et al*. Decade-long bacterial community dynamics in cystic fibrosis airways. *Proc Natl Acad Sci U S A* 2012;109:5809–14.
9 Stressmann FA, Rogers GB, van der Gast CJ, *et al*. Long-term cultivation-independent microbial diversity analysis demonstrates that bacterial communities infecting the adult cystic fibrosis lung show stability and resilience. *Thorax* 2012;67:867–73.
10 Delhaes L, Monchy S, Fréalle E, *et al*. The airway microbiota in cystic fibrosis: a complex fungal and bacterial community--implications for therapeutic management. *PLoS One* 2012;7:e36313.
11 Hunter RC, Klepac-Ceraj V, Lorenzi MM, *et al*. Phenazine content in the cystic fibrosis respiratory tract negatively correlates with lung function and microbial complexity. *Am J Respir Cell Mol Biol* 2012;47:738–45.
12 Linnane B, Walsh AM, Walsh CJ, *et al*. The lung microbiome in young children with cystic fibrosis: a prospective cohort study. *Microorganisms* 2021;9. doi:10.3390/microorganisms9030492. [Epub ahead of print: 26 02 2021].
13 Frayman KB, Armstrong DS, Carzino R, *et al*. The lower airway microbiota in early cystic fibrosis lung disease: a longitudinal analysis. *Thorax* 2017;72:1104–12.
14 Prevaes SMPJ, de Winter-de Groot KM, Janssens HM, *et al*. Development of the nasopharyngeal microbiota in infants with cystic fibrosis. *Am J Respir Crit Care Med* 2016;193:504–15.
15 Mika M, Korten I, Qi W, *et al*. The nasal microbiota in infants with cystic fibrosis in the first year of life: a prospective cohort study. *Lancet Respir Med* 2016;4:627–35.
16 Ahmed B, Cox MJ, Cuthbertson L, *et al*. Longitudinal development of the airway microbiota in infants with cystic fibrosis. *Sci Rep* 2019;9:5143.

17 Mika M, Korten I, Qi W, *et al*. The nasal microbiota in infants with cystic fibrosis in the first year of life: a prospective cohort study. *Lancet Respir Med* 2016;4:627–35.
18 Zemanick ET, Wagner BD, Robertson CE, *et al*. Airway microbiota across age and disease spectrum in cystic fibrosis. *Eur Respir J* 2017;50:1700832.
19 Dickson RP, Martinez FJ, Huffnagle GB. The role of the microbiome in exacerbations of chronic lung diseases. *Lancet* 2014;384:691–702.
20 Bush A, Payne D, Pike S, *et al*. Mucus properties in children with primary ciliary dyskinesia: comparison with cystic fibrosis. *Chest* 2006;129:118–23.
21 Noone PG, Leigh MW, Sannuti A, *et al*. Primary ciliary dyskinesia: diagnostic and phenotypic features. *Am J Respir Crit Care Med* 2004;169:459–67.
22 Rogers GB, Carroll MP, Zain NMM, *et al*. Complexity, temporal stability, and clinical correlates of airway bacterial community composition in primary ciliary dyskinesia. *J Clin Microbiol* 2013;51:4029–35.
23 Ahmed B, Cox MJ, Cuthbertson L, *et al*. Comparison of the upper and lower airway microbiota in children with chronic lung diseases. *PLoS One* 2018;13:e0201156.
24 Farrell PM, White TB, Ren CL, *et al*. Diagnosis of cystic fibrosis: consensus guidelines from the cystic fibrosis Foundation. *J Pediatr* 2017;181S:S4–15.
25 Bush A, Hogg C. Primary ciliary dyskinesia: recent advances in epidemiology, diagnosis, management and relationship with the expanding spectrum of ciliopathy. *Expert Rev Respir Med* 2012;6:663–82.
26 Lucas JS, Barbato A, Collins SA, *et al*. European respiratory Society guidelines for the diagnosis of primary ciliary dyskinesia. *Eur Respir J* 2017;49:1601090.
27 Cystic Fibrosis Trust. *Standards of care: laboratory standards for processing microbiological samples from people with cystic fibrosis*. London, UK, 2010.
28 Lee TWR, Brownlee KG, Conway SP, *et al*. Evaluation of a new definition for chronic Pseudomonas aeruginosa infection in cystic fibrosis patients. *J Cyst Fibros* 2003;2:29–34.
29 Edgar RC. Search and clustering orders of magnitude faster than blast. *Bioinformatics* 2010;26:2460–1.
30 Salter SJ, Cox MJ, Turek EM, *et al*. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 2014;12:87.
31 McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;8:e61217.
32 Wickham H. *ggplot2: elegant graphics for data analysis*. New York: Springer, 2009.
33 Auguie B. gridExtra: Miscellaneous Functions for "Grid" Graphics, 2016. Available: https://CRAN.R-project.org/package=gridExtra
34 Fournier DA, Skaug HJ, Ancheta J, *et al*. Ad model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software* 2012;27:233–49.
35 VEGAN DP. A package for R functions for commnity ecology. *Journal of Vegetation Science* 2003;14:927–30.
36 Davies JC, Cunningham S, Alton EWFW, *et al*. Lung clearance index in CF: a sensitive marker of lung disease severity. *Thorax* 2008;63:96–7.
37 Alanin MC, Nielsen KG, von Buchwald C, *et al*. A longitudinal study of lung bacterial pathogens in patients with primary ciliary dyskinesia. *Clin Microbiol Infect* 2015;21:1093.e1–1093.e7.
38 Wijers CD, Chmiel JF, Gaston BM. Bacterial infections in patients with primary ciliary dyskinesia: comparison with cystic fibrosis. *Chron Respir Dis* 2017;14:392–406.
39 Rogers GB, van der Gast CJ, Serisier DJ. Predominant pathogen competition and core microbiota divergence in chronic airway infection. *Isme J* 2015;9:217–25.
40 Boutin S, Graeber SY, Weitnauer M, *et al*. Comparison of microbiomes from different niches of upper and lower airways in children and adolescents with cystic fibrosis. *PLoS One* 2015;10:e0116029.
41 Zemanick ET, Wagner BD, Robertson CE, *et al*. Assessment of airway microbiota and inflammation in cystic fibrosis using multiple sampling methods. *Ann Am Thorac Soc* 2015;12:221–9.
42 Rosen R, Hu L, Amirault J, *et al*. 16S community profiling identifies proton pump inhibitor related differences in gastric, lung, and oropharyngeal microflora. *J Pediatr* 2015;166:917–23.
43 Khanolkar RA, Clark ST, Wang PW, *et al*. Ecological succession of polymicrobial communities in the cystic fibrosis airways. *mSystems* 2020;5. doi:10.1128/mSystems.00809-20. [Epub ahead of print: 01 12 2020].