

## Comparison of the upper and lower airway microbiota in children with chronic lung diseases

Ahmed, Bushra; Cox, Michael J.; Cuthbertson, Leah; James, Phillip L.; Cookson, William O. C.; Davies, Jane C.; Moffatt, Miriam F.; Bush, Andrew

DOI:

[10.1371/journal.pone.0201156](https://doi.org/10.1371/journal.pone.0201156)

License:

Creative Commons: Attribution (CC BY)

### Document Version

Publisher's PDF, also known as Version of record

### Citation for published version (Harvard):

Ahmed, B, Cox, MJ, Cuthbertson, L, James, PL, Cookson, WOC, Davies, JC, Moffatt, MF & Bush, A 2018, 'Comparison of the upper and lower airway microbiota in children with chronic lung diseases', *PLoS ONE*, vol. 13, no. 8, e0201156, pp. 1-15. <https://doi.org/10.1371/journal.pone.0201156>

[Link to publication on Research at Birmingham portal](#)

### General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

### Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

RESEARCH ARTICLE

# Comparison of the upper and lower airway microbiota in children with chronic lung diseases

Bushra Ahmed<sup>1,2\*</sup>, Michael J. Cox<sup>1</sup>, Leah Cuthbertson<sup>1</sup>, Phillip L. James<sup>1</sup>, William O. C. Cookson<sup>1</sup>, Jane C. Davies<sup>1,2</sup>, Miriam F. Moffatt<sup>1‡</sup>, Andrew Bush<sup>1,2‡</sup>

**1** National Heart and Lung Institute, Imperial College London, London, United Kingdom, **2** Department of Respiratory Paediatrics, Royal Brompton Hospital, London, United Kingdom

‡ These authors are joint senior authors.

\* [b.ahmed12@imperial.ac.uk](mailto:b.ahmed12@imperial.ac.uk)



## Abstract

### Rationale

The lower airway microbiota is important in normal immunological development and chronic lung diseases (CLDs). Young children cannot expectorate and because of the uncertainty whether upper airway samples reflect the lower airway microbiota, there have been few longitudinal paediatric studies to date.

### Objectives

To assess whether throat swabs (TS) and cough swabs (CS) are representative of the lower airway microbiota.

### Methods

TS, CS, bronchoalveolar lavage and bronchial brushings were prospectively collected from 49 children undergoing fiberoptic bronchoscopy for CLDs. Bacterial DNA was extracted and the 16S rRNA gene V4 region sequenced using the Illumina MiSeq.

### Results

5.97 million high quality reads were obtained from 168 samples (47 TS, 37 CS, 42 BALF and 42 bronchial brushings). CS sequenced poorly. At a community level, no difference in alpha diversity (richness, evenness or Shannon Diversity Index) was seen between lower airway samples and TS ( $P > 0.05$ ). Less than 6.31% of beta diversity variation related to sampling method for TS ( $P = 0.001$ ). Variation between pathologies and individual patients was greater (20%, 54% respectively  $P \leq 0.001$ ) than between TS and lower airway samples. There was strong correlation in the relative abundance of genera between samples ( $r = 0.78$ ,  $P < 0.001$ ). Similarity between upper and lower airway samples was observed to be less for individuals where one sample type was dominated by a single organism.

## OPEN ACCESS

**Citation:** Ahmed B, Cox MJ, Cuthbertson L, James PL, Cookson WOC, Davies JC, et al. (2018) Comparison of the upper and lower airway microbiota in children with chronic lung diseases. PLoS ONE 13(8): e0201156. <https://doi.org/10.1371/journal.pone.0201156>

**Editor:** Alexander Larcombe, Telethon Institute for Child Health Research, AUSTRALIA

**Received:** February 13, 2018

**Accepted:** July 10, 2018

**Published:** August 2, 2018

**Copyright:** © 2018 Ahmed et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Sequence data is available at the European Nucleotide Archive (Accession number: PRJEB14074).

**Funding:** The project was supported by the NIHR Respiratory Disease Biomedical Research Unit at the Royal Brompton and Harefield NHS Foundation Trust, Imperial College London, the Asmarley Trust and the Wellcome Trust. AB and WOC are NIHR Senior Investigators whilst MFM and WOC are Joint WT Senior Investigators. BA was supported by an NIHR PhD studentship. AB was supported by

the NIHR Respiratory Disease Biomedical Research Unit at the Royal Brompton and Harefield NHS Foundation Trust and Imperial College London. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Conclusions

At the community structure level, TS correlate with lower airway samples and distinguish between different CLDs. TS may be a useful sample for the study of the differences in longitudinal changes in the respiratory microbiota between different CLDs. Differences are too great however for TS to be used for clinical decision making.

## Introduction

Complex microbial communities inhabit the healthy lower airways, that were once thought to be sterile, and are important in chronic lung diseases such as cystic fibrosis (CF)[1]. Using molecular techniques, chronic suppurative airway diseases are now believed to be more polymicrobial than previously appreciated[2–4] with increasing interest in the role the microbiota has to play in shaping normal and pathological airway immune responses.

Whilst studies in adults have highlighted the importance of the microbiota in disease progression in CF, comparatively less is known about the early development of the airway microbiota in children. This is in part because longitudinal study of the airway microbiota in children is challenging due to the difficulties in obtaining repeated lower airway samples. Most children cannot spontaneously expectorate and obtaining bronchoalveolar lavage fluid (BALF) frequently is neither ethical nor feasible. In the only two longitudinal studies of the infant lower airway microbiota reported to date, BALF was collected infrequently (at most every 6 months)[5,6] thus limiting the utility of BALF to track concurrent changes in the microbiota with symptoms. Nasopharyngeal sampling has also been used to track monthly changes in the airway microbiota in infants[7–9], but this has been shown to be a poor surrogate for the lower airway microbiota[10]. Thus, there is a need to identify a reliable surrogate for lower airway sampling which can be collected frequently and is well tolerated by children.

In clinical practice, cultures of either throat swabs (TS) or cough swabs (CS) are regularly used as surrogates for lower airway samples. Culture dependent results of TS and CS are however unreliable due to variability in sensitivity when compared with lower airway results: e.g. conventional culture sensitivity for *Pseudomonas aeruginosa* ranges from 35.7–71%[11,12]. Whether the poor sensitivity of upper airway samples is due to the limitations of microbial cultures or the sampling method however remains unclear.

Similarities in the microbiota between throat swabs and lower airway samples have been reported in older expectorating children with CF (TS compared with sputum)[10] as well as in adults (N = 6 adults, TS compared with BALF)[13]. In younger children and infants with chronic lung diseases, the use of upper airway samples, particularly CS, has not been validated and was the objective of the present study. We hypothesised that upper airway samples (CS and TS) would be a reliable surrogate for the lower airway microbiota. We aimed to compare TS and CS with lower airway samples obtained at bronchoscopy in children with CLDs undergoing a clinically indicated procedure.

## Methods

For additional details please see [S1 Appendix](#).

## Subjects and sampling

Children undergoing a clinical bronchoscopy at the Royal Brompton Hospital (RBH) between December 2012 and May 2013 were recruited. Ethical approval was granted by the RBH NIHR

Biomedical Research Unit Advanced Lung Disease Biobank (NRES reference 10/H0504/9). Parental written consent and age-appropriate assent from the child was obtained.

At least one upper airway (TS or CS) and a paired lower airway sample (BALF or bronchial brushing) were collected from each child (see [S1 Appendix](#)). Bacterial culture of BALF was performed as per standard clinical practice.

## 16S rRNA gene library preparation and sequencing

A maximum of 2 x 2ml aliquots of BALF (median 3.2ml, range 0.5–4ml) were centrifuged at 21,000g for 30 minutes and the cell pellet retained for DNA extraction. Frozen swab and brushing heads were transferred directly into a Lysing Matrix E tube containing sodium phosphate buffer. DNA was extracted using the MP Bio FastDNA Spin Kit for Soil (<http://www.mpbio.com>) according to the manufacturer's instructions.

Quadruplicate PCRs of the 16S rRNA gene V4 region were performed using a custom indexed forward primer S-D-Bact-0564-a-S-15 (5' AYT GGG YDT AAA GNG 3'), reverse primer S-D-Bact-0785-b-A-18 (5' TAC NVG GGT ATC TAA TCC 3')[14] and a high fidelity *Taq* polymerase master mix (Q5, New England Biolabs). A mock community was included to assess sequencing quality. PCR cycling conditions were: annealing at 95°C for 2 minutes followed by 35 cycles at 95°C for 20 seconds, 50°C for 20 seconds and 72°C for 5 minutes. Amplicons were purified, quantified and equi-molar pooled to form a DNA library for paired-end sequencing using the Illumina MiSeq V2 reagent kit [15] ([S1 Fig](#) for laboratory workflow). With the exception of 10 samples (details given in [S1 Appendix](#)), all samples from an individual patient were run on the same plate to limit the impact of any batch effect.

## Data analysis

Sample size was opportunistic in the absence of data to inform a power calculation at the inception of the study. Upstream processing was performed using Quantitative Insights into Microbial Ecology (QIIME) Version 1.9.0 (see [S1 Appendix](#)). Downstream analyses, to assess community level differences in diversity and Operational Taxonomic Unit (OTU) level differences, were performed using Phyloseq in R version 3.2.0 ([S2 Fig](#)). Within samples, alpha-diversity differences in richness (the number of different species), evenness (the spread of species) and the Shannon Diversity Index were calculated using paired t-tests and Wilcoxon sign-ranked tests for parametric and non-parametric data respectively. The Shannon Diversity Index is a composite measure of richness and evenness which quantifies the uncertainty in predicting species identity when randomly sampling from a community. Bland-Altman plots were constructed to analyse the agreement in alpha diversity between samples.

Between sample, beta-diversity differences were tested using the Bray Curtis dissimilarity, the unweighted UniFrac and weighted UniFrac scores using a permutational multivariate ANOVA (PERMANOVA)[16], in which the  $r^2$  value represents the degree of variance in community composition explained by the variables tested in the model. Block designs were used for paired comparisons between samples from the same patient. The Bray-Curtis index measures dissimilarity between two samples by estimating the number of shared organisms from the total number of organisms present in both samples; the higher the Bray Curtis dissimilarity score, the greater the difference between two samples. The UniFrac score is a qualitative measure of the phylogenetic distance between two communities, with the weighted UniFrac score additionally accounting for the relative abundance of organisms within these samples. Quantitative measures of beta diversity, such as Bray Curtis dissimilarity and weighted UniFrac scores, provide information on community differences due to the relative abundance of

species. The UniFrac score provides information on community differences driven by selective pressures on the community[17].

OTU level differences were assessed using Spearman’s rank correlation. The Multtest package in R was used to perform multiple t-tests with Benjamini-Hochberg correction to compare differences in the relative abundance of genera between upper and lower airway samples. A *P* value of less than 0.05 was considered to be statistically significant. Sequence data is available at the European Nucleotide Archive (Accession number: PRJEB14074).

## Results

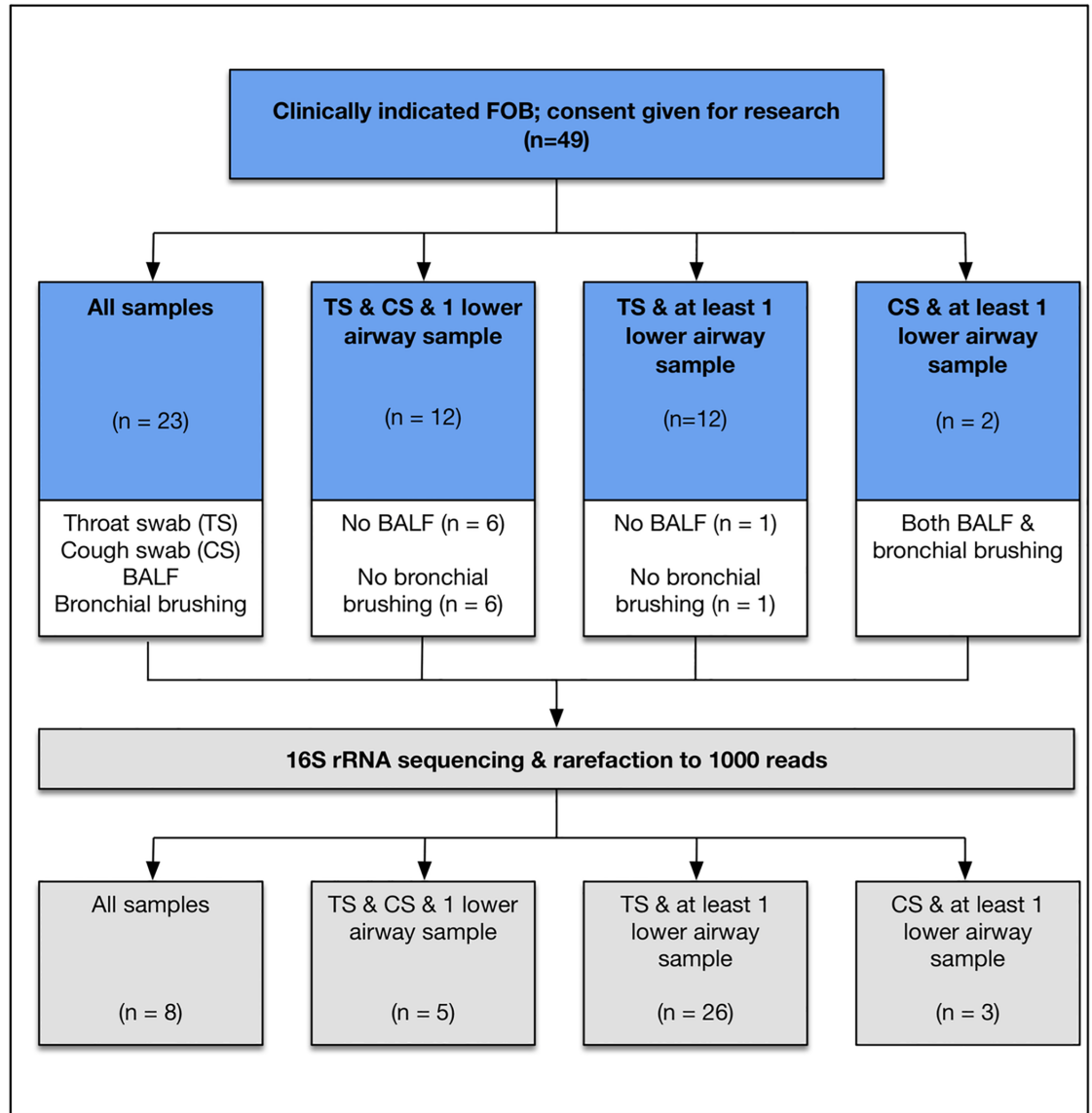
### Patient demographics & sampling

Patient characteristics are detailed in Table 1. In total 47 TS, 37 CS, 42 BALF samples and 42 bronchial brushings were collected, with 23 of the 49 patients providing all sample types and a median of 3 samples collected per patient. Reasons for missed samples included: infants who were too young to perform a CS (*n* = 12); children who were unable to tolerate TS sampling due to gagging and vomiting (*n* = 2); lack of parental consent for bronchial brushings (*n* = 7), and insufficient BALF remaining for research after aliquoting for clinical purposes (*n* = 7). For the remaining 26 patients, at least one pair of upper and lower airway samples was collected

**Table 1. Summary of patient characteristics (n = 49) in the comparison of the upper and lower airway microbiota.** CSLD—chronic suppurative lung disease; CF—cystic fibrosis; NBS—newborn screened; PCD—Primary ciliary dyskinesia.

Demographic	Median (range) or n (%)
Age (years)	5.4 (0.1–16.2)
Gender (female)	29 (59)
<b>Underlying pathology</b>	
• CF (Newborn screened [NBS] infants)	• 12 (25)
• CF (not NBS)	• 5 (10)
• PCD	• 5 (10)
• Recurrent LRTI	• 24 (49)
• Upper airway pathology	• 2 (4)
• Haemoptysis	• 1 (2)
<b>Current symptoms</b>	
• Lower respiratory	• 10 (20)
• Upper respiratory	• 21 (42)
<b>Antibiotic usage</b>	
Current treatment course	14 (29)
• Intravenous (IV)	• 3 (6)
• Oral	• 8 (16)
• Nebulised	• 6 (12)
Oral prophylaxis	20 (41)
• Augmentin	• 2 (4)
• Azithromycin	• 9 (18)
• Flucloxacillin	• 9 (18)
<b>Inhaled corticosteroids</b>	12 (25)
<b>Bronchoscopy route</b>	
• Oral (Via laryngeal mask airway)	• 11 (22)
• Nasal	• 38 (78)

<https://doi.org/10.1371/journal.pone.0201156.t001>



**Fig 1. Illustration of the combinations of upper and lower airway samples taken.** Forty-nine patients were recruited and in forty-seven patients, a throat swab and at least 1 lower airway sample was obtained for comparison. Samples for forty-two patients remained after rarefying to 1000 reads: 8 patients remained in whom all 4 samples were available; 39 had TS and at least a paired lower airway sample (either BALF, bronchial brushing or both), and 17 had CS and a paired lower airway sample. n—number of subjects. After rarefaction, some subjects had fewer samples for comparison and therefore moved into a different category.

<https://doi.org/10.1371/journal.pone.0201156.g001>

thereby allowing comparisons to be made (Fig 1). Twenty-six of the 42 BALFs obtained were culture positive.

Indications for bronchoscopy included: chronic suppurative lung disease (CSLD—namely CF and Primary Ciliary Dyskinesia [PCD]) (45%) diagnosed on standard criteria [18], [19], and non-CSLD controls (55%). The majority of the controls (24/27, 89%) were being investigated for recurrent lower respiratory tract infections (LRTI). For those with CSLD, twenty (25%) were infants with CF diagnosed on newborn screening (NBS) undergoing a routine bronchoscopy at 3–5 months of age [20]. The age distribution of patients was skewed with a median of 5.4 years (0.1–16.2 years). Thirty-nine patients (80%) were clinically stable at the



time of bronchoscopy. Eighteen patients (37%) had received a treatment course of antibiotics in the previous 30 days, with 29% receiving antibiotics at the time of bronchoscopy (either intravenous [IV], oral, nebulised or a combination).

## Sequencing

16S rRNA gene sequencing was performed on 168 extracted DNA samples. After initial processing and quality control, a total of 5.97 million reads were obtained with an average of 31,117 reads per sample (range 118–192,812 reads). Technical and PCR negative controls were examined to identify potential contaminant OTUs. Three individual OTUs were found to be highly abundant in technical and PCR control samples and were removed: *Burkholderia* (OTU ID 1606), *Undibacterium* (OTU ID 1727) and *Ralstonia* (OTU ID 1703). In addition, the genera *Bradyrhizobium* spp., *Sediminibacterium* spp. and *Methylobacterium* spp. were removed as these have previously been identified as common reagent contaminants[21]. A small significant batch effect was found between the two sequencing runs using Bray Curtis dissimilarity ( $r^2 = 0.03$ ,  $P = 0.001$ ) and unweighted UniFrac scores ( $r^2 = 0.01$ ,  $P = 0.03$ ), but not the weighted UniFrac score ( $r^2 = 0.01$ ,  $P = 0.69$ ).

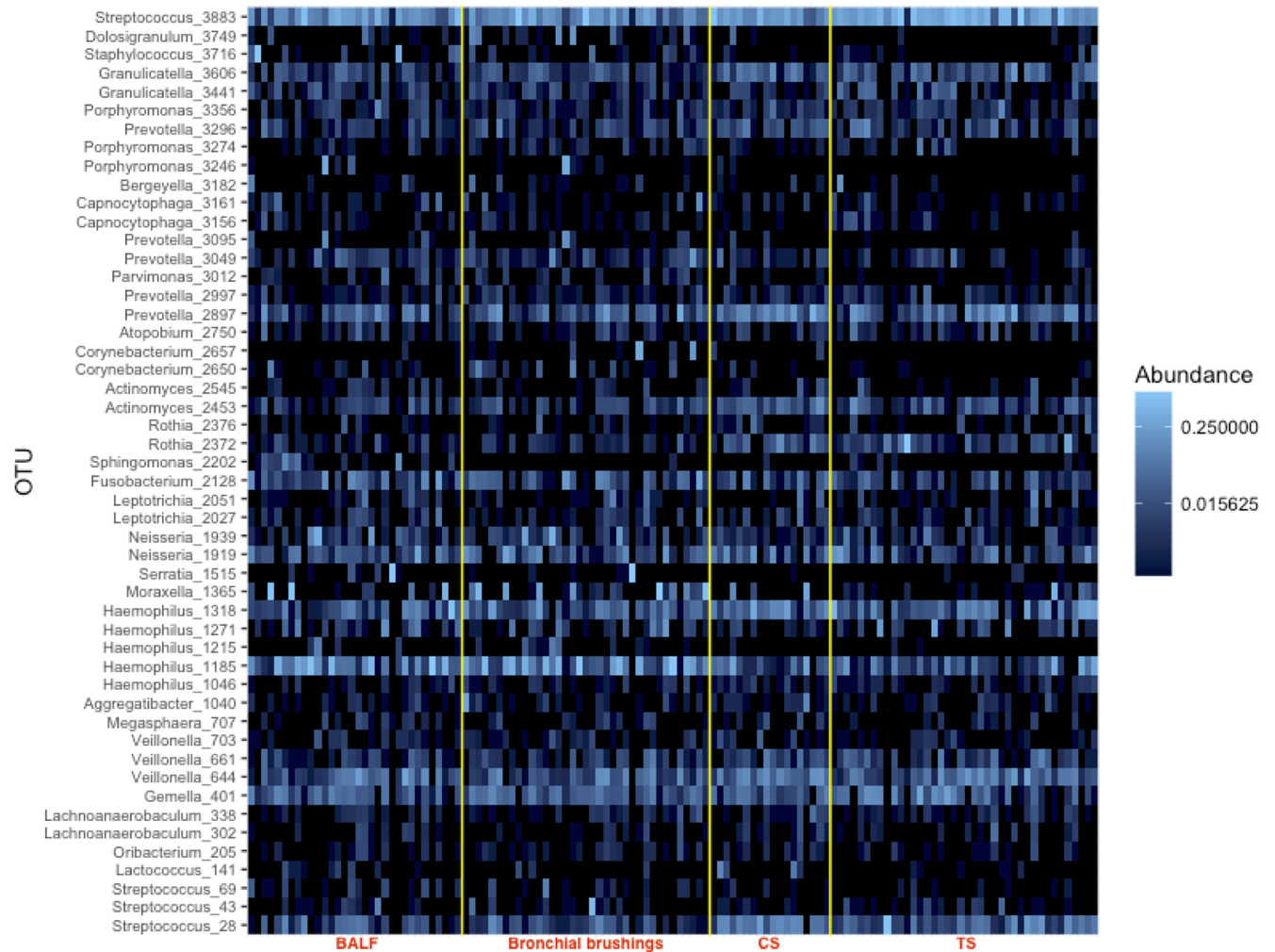
Prior to downstream processing, data was filtered to remove OTUs with less than 20 total reads. Samples were rarefied to 1,000 reads and any sample with fewer than 1,000 reads removed (rationale for rarefaction level chosen detailed in the [S1 Appendix](#) and [S3 Fig](#)). After rarefaction 34/42 (81%) of the BALFs, 36/42 (86%) bronchial brushings, 44/47 (94%) TS and 17/37 (46%) CS remained. CS sequenced poorly, mostly likely due to having a very low biomass, and consequently post rarefaction only 8 patients remained in whom all 4 samples (TS, CS, BALF and bronchial brushing) were available for comparison. In contrast, post rarefaction 39 patients had a TS and at least a paired lower airway sample remaining with 17 having a CS and paired lower airway sample ([Fig 1](#)).

## Comparison of the upper and lower airway microbiota

For BALF and bronchial brushings, the three most common genera were identical: *Haemophilus* spp. (23.6% and 24.6% of total reads respectively), *Streptococcus* spp. (20.3% and 20.5%) and *Prevotella* spp. (6.5% and 7.9%). Similar genera were seen in TS but with *Streptococcus* spp. the most common (39.5%) followed by *Haemophilus* spp. (15.4%) and *Prevotella* spp. (8.7%). At an OTU level, there was no difference observed between upper and lower airway samples ([Fig 2](#)).

Comparing alpha diversity between paired BALF and bronchial brushings ( $N = 26$ ), no significant differences were found in richness (median for BALF = 57.5 (range 12–87), for bronchial brushings = 50, (range 14–101),  $P = 0.578$ ), evenness (median for BALF = 0.60 [range 0.16–0.75], for bronchial brushings = 0.55 [range 0.13–0.70]  $P = 0.084$ ), or by the Shannon Diversity Index (median for BALF 2.46 [range 0.39–3.22], for bronchial brushings 2.18 [range 0.37–3.19],  $P = 0.173$ ). Similarly no differences were observed for beta diversity (Bray Curtis dissimilarity index [ $P = 0.650$ ,  $r^2 = 0.005$ ], unweighted UniFrac score [ $P = 0.114$ ,  $r^2 = 0.02$ ], weighted UniFrac score [ $P = 0.255$ ,  $r^2 = 0.008$ ]). Consequently where both BALF and bronchial brushings were available for comparison with TS, bronchial brushings were used.

Looking at individual patient barplots the airway microbiota was revealed to be highly individual ([S4 Fig](#)), with some similarities observed between TS and lower airway samples ([Fig 3](#)). *Streptococcus* spp., however, appears to be more abundant in TS compared to lower airway samples. Communities with low evenness were in most cases dominated by a single organism. There was a trend for the dominant organism in the lower airway sample to be the same bacterium isolated by routine clinical culture of BALF ([S1 Appendix](#) and [S4 Fig](#)).



**Fig 2. Heatmap showing similarities in the relative abundances of the top 50 OTUs present between upper (throat swabs [TS] and cough swabs [CS]) and lower airway samples (bronchoalveolar lavage fluid [BALF] and bronchial brushings).**

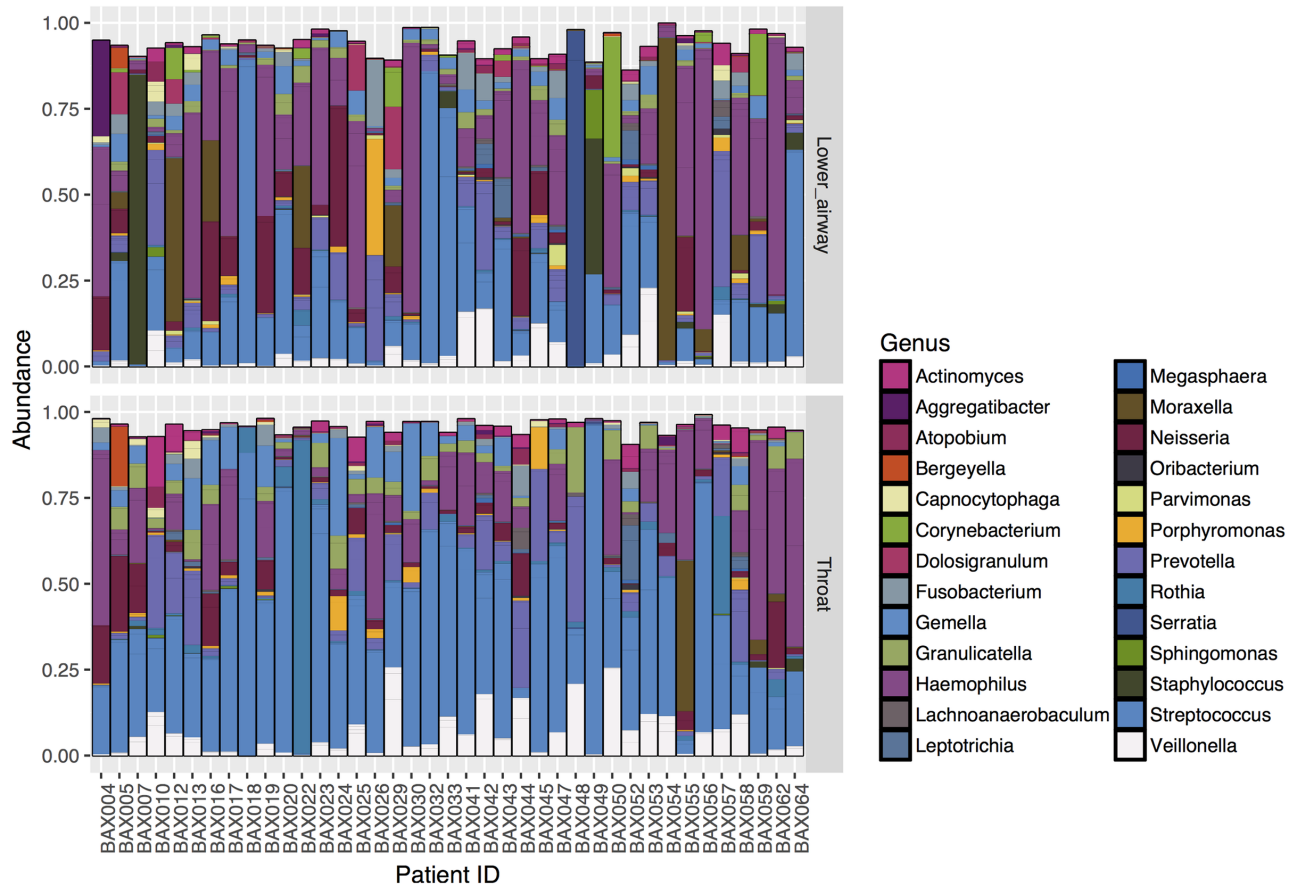
<https://doi.org/10.1371/journal.pone.0201156.g002>

Comparing TS and either BALF or bronchial brushings (N = 39), no significant difference was found in richness (median for lower airway samples = 51 [range 14–101], for TS = 47 [range 25–70],  $P = 0.120$ ), evenness (median for lower airway samples = 0.55 [range 0.13–0.76], for TS = 0.56 [range 0.09–0.75],  $P = 0.809$ ) or by the Shannon Diversity Index (median for lower airway samples = 2.18 [range 0.37–3.29], for TS = 2.13 [range 0.31–3.17],  $P = 0.777$ , Fig 4).

For those individuals showing a difference in alpha diversity between upper and lower airway samples, Bland-Altman plots of alpha diversity measurements revealed somewhat less agreement between TS and lower airway samples with low evenness (S5 Fig). This again suggests that there is less concurrence between TS and lower airway samples in the presence of a dominant organism. The mean Shannon Diversity Index was 2.09 (standard deviation [SD] 0.64) for TS and 2.12 [SD 0.78] for lower airway samples). A small significant difference ( $P = 0.001$ ) was found in beta-diversity between TS and lower airway samples using the Bray Curtis, unweighted and weighted UniFrac scores ( $r^2 = 0.06$ , 0.03 and 0.06 respectively [Fig 5]).

Considering genera, 80% (72 out of 90) were in common between lower airway samples and TS, whilst 15.6% (14 out of 90) were unique to BALF and 4.4% (4 out of 90) were unique





**Fig 3. Individual patient barplots (N = 39) comparing paired lower airway samples and TS for the 50 most common OTUs.** The relative abundance of OTUs for each sample is shown with TS below and their corresponding lower airway sample above. Bars are of uneven heights due to the presence of low abundance “other” OTUs which have not been included in the plot. This illustrates that the airway microbiota is highly individual with variable degrees of similarity between TS and lower airway samples.

<https://doi.org/10.1371/journal.pone.0201156.g003>

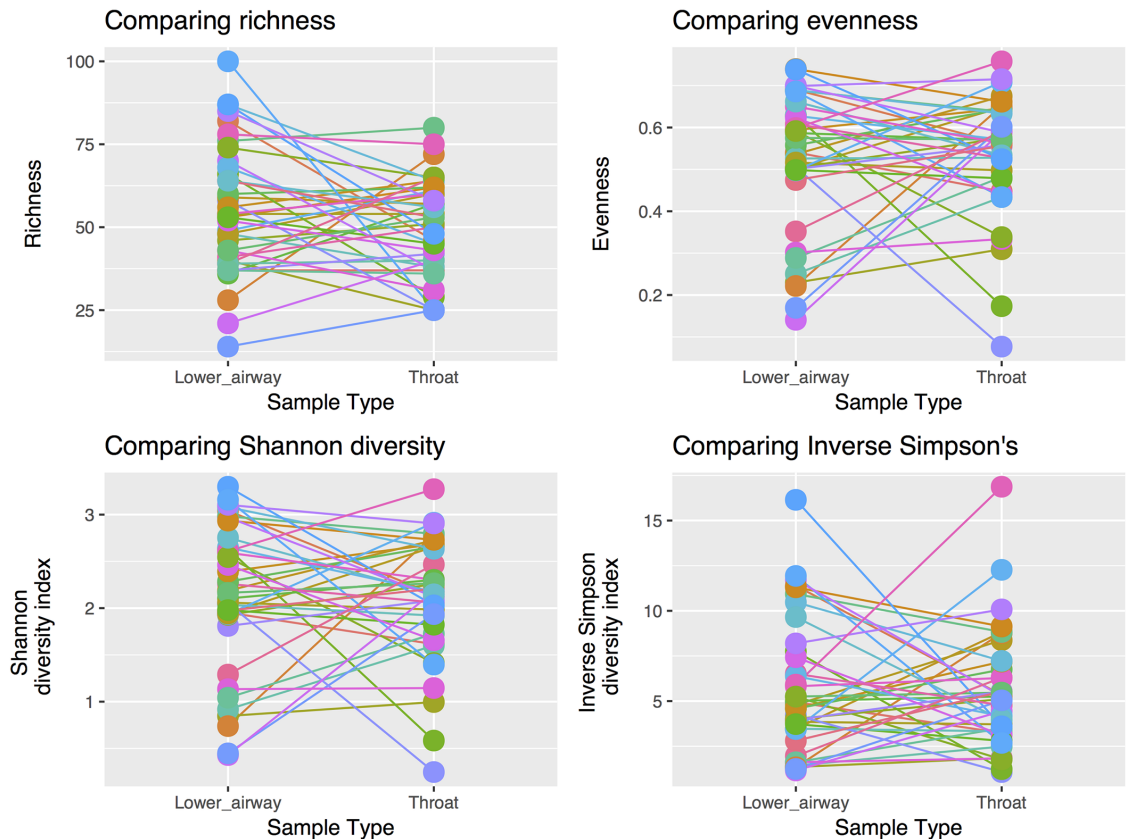
to TS. None of the genera unique to BALF were grown on bacterial culture of the same sample. The few genera different between upper and lower airway samples were present in very low relative abundance (< 0.1%).

Spearman’s rank correlation testing showed good correlation between the relative abundances of genera and OTUs present when comparing lower airway samples with TS (genera level:  $r = 0.776$ ,  $P < 0.001$ ; OTU level:  $r = 0.557$ ,  $P < 0.001$ ). Using multiple paired t-tests with Benjamini-Hochberg correction to compare the relative abundance of genera, a significantly higher relative abundance of *Streptococcus* spp. was seen ( $t = -182$ ,  $P < 0.0001$ ,  $P_{adj} = 0.0009$ ) for TS. No other genus was significantly different between TS and lower airway samples ( $P_{adj} > 0.05$ ) (Table 2).

Although CS sequenced poorly, when successfully sequenced they showed similarities in diversity with lower airway samples (see S1 Appendix).

### Clinical variables influencing correlation between the upper and lower airway microbiota

To determine whether specific clinical features underpinned the differences in similarity between upper and lower airway samples (Fig 3), beta-diversity testing was performed using a



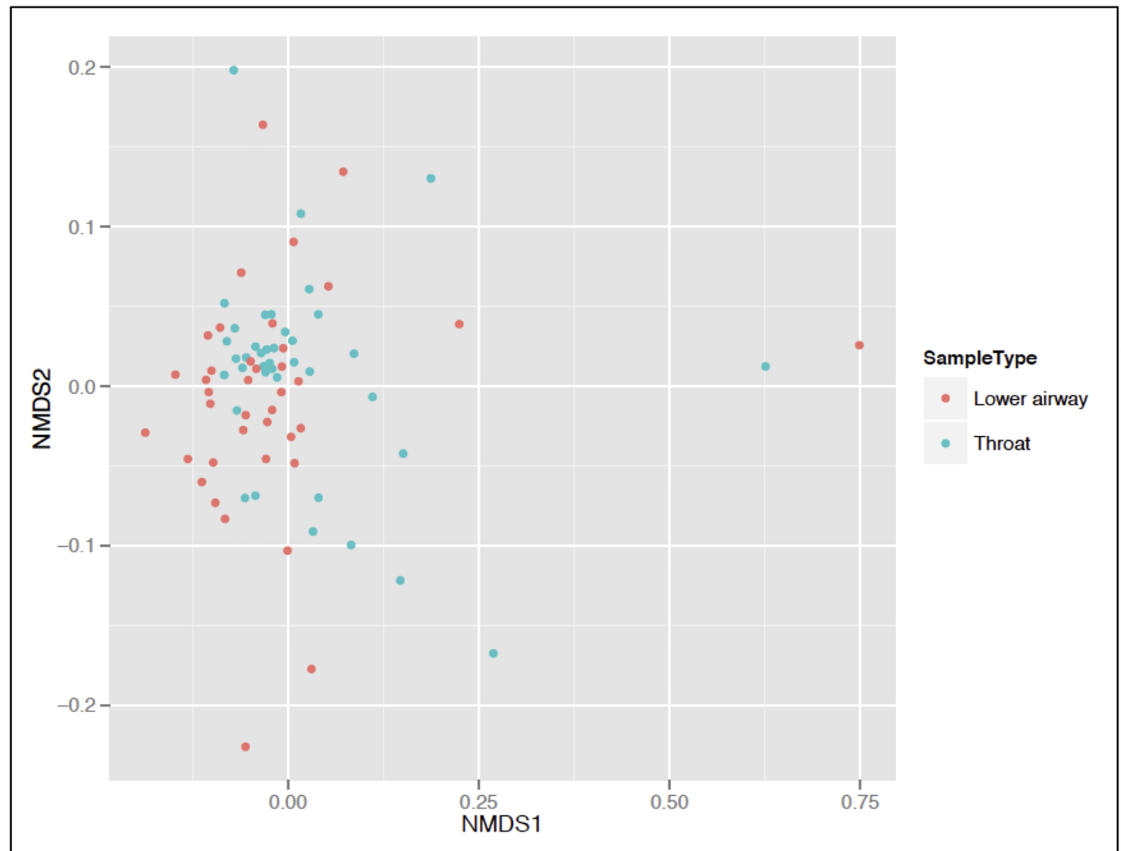
**Fig 4. Alpha-diversity comparisons between lower airway samples and throat swabs.** No significant difference was seen in: (a) richness ( $t_{(38)} = 1.6523, P = 0.107$ ); (b) evenness ( $W_{(38)} = 367, P = 0.756$ ), and (c) Shannon diversity index ( $W_{(36)} = 384, P = 0.940$ ).

<https://doi.org/10.1371/journal.pone.0201156.g004>

PERMANOVA including a total of 11 variables—those listed in Table 1 as well as “Patient ID” (i.e. which patient was sampled) and “Sample type”. The influence of FOB route and underlying pathology were tested in separate PERMANOVAs using FOB samples or TS only respectively. Two variables were found to exert a large influence on community structure. “Patient ID” had the largest influence accounting for 47.0–53.8% of the variance (for the unweighted UniFrac and Bray Curtis dissimilarity scores respectively,  $P = 0.001$ ), whilst the underlying pathology accounted for 18.7%–22.5% (unweighted UniFrac and Bray Curtis dissimilarity scores respectively,  $P \leq 0.006$ ) (Table 3). FOB route accounted for 5.43%–8.16% of the variance (unweighted and weighted UniFrac respectively,  $P < 0.05$ ), with fewer *Corynebacterium* spp. (OTU 2657), *Moraxella* spp. (OTU 1365) and *Dolosigranulum* spp. (OTU 349) present in FOB performed nasally as opposed to FOB through the oral route. Other significant variables, but similarly of more minor influence, included the use of prophylactic antibiotics and nebulized antibiotics. Patient age was not a significant influence on community structure.

## Discussion

In summary, the present study has revealed that TS sequence well (94% of samples sequenced) in contrast to CS that sequenced poorly (44%). We attribute this to the limited biomass and recoverable microbial DNA for CS and the use of CS as a sample for sequencing is therefore not recommended.



**Fig 5. Non-metric multidimensional scaling (NMDS) plot comparing the UniFrac score between patients.** This shows similarity in the clustering pattern between lower airway samples and throat swabs.

<https://doi.org/10.1371/journal.pone.0201156.g005>

When comparing TS with lower airway samples (either BALF or bronchial brushings), a strong correlation was seen in the relative abundance of genera and there was no significant difference in within sample alpha diversity. Whilst a significant difference in beta diversity was seen between sample types, the degree of variation due to sample site and other clinical variables, such as antibiotic usage, was small. This was consistent when testing whether similar organisms were present between samples (Bray Curtis dissimilarity) and when testing the relative abundance and phylogenetic relationships between samples (UniFrac and weighted UniFrac scores).

There was greater variability between disease states (up to 22.5%) than between upper and lower airway samples. However, the number of children with discordant lower and upper airway samples precludes the use of TS for clinical decision making in individuals. Nevertheless TS contain valuable research information that could be used to explore the development of the lower airway microbiota in groups of children with different diseases. For example, TS could be used to determine whether the more benign clinical course of PCD compared to that of CF relates to different temporal evolution of the lower airway microbiota.

Discordant dominance of an OTU in either the upper or lower airway sample led to greater dissimilarity between TS and lower airway samples. For those samples showing the greatest dissimilarity, there was a trend for the dominant organism in BALF to be the same organism grown on bacterial cultures of the same fluid. This could suggest overgrowth of an organism in an individual sample, or low biomass in that sample allowing artificial dominance of an

**Table 2. Comparing the mean relative abundance (percent) of the most common or clinically important genera between TS and paired lower airway samples (BALF or bronchial brushing).** Using multiple paired t-tests with Benjamini-Hochberg correction, only *Streptococcus* spp. was significantly different ( $P_{(adj)} = 0.0009$ ). TS—throat swab; NS—non-significant ( $P_{(adj)} > 0.05$ ). SD—standard deviation.

Genus	Lower airway mean (SD)	TS mean (SD)	$P_{(adj)}$ -value
<i>Haemophilus</i> spp.	25.2 (24.4)	15.6 (16.2)	0.35
<i>Streptococcus</i> spp.	21.5 (22.2)	39.7 (22.8)	<b>0.0009</b>
<i>Prevotella</i> spp.	7.7 (9.9)	8.3 (10.0)	0.89
<i>Neisseria</i> spp.	6.9 (10.1)	4.6 (6.3)	0.47
<i>Moraxella</i> spp.	6.2 (17.4)	1.4 (7.0)	0.47
<i>Veillonella</i> spp.	4.2 (5.6)	6.8 (7.1)	0.35
<i>Staphylococcus</i> spp.	4.1 (16.3)	0.2 (0.7)	0.47
<i>Gemella</i> spp.	3.0 (2.6)	5.3 (7.3)	0.42
<i>Fusobacterium</i> spp.	2.6 (3.8)	1.3 (2.0)	0.42
<i>Serratia</i> spp.	2.6 (15.9)	0.05 (0.1)	0.55
<i>Corynebacterium</i> spp.	2.2 (6.6)	0.03 (0.1)	0.35
<i>Granulicatella</i> spp.	1.7 (1.7)	4.2 (3.9)	0.05
<i>Dolosigranulum</i> spp.	1.6 (4.2)	0.03 (0.1)	0.35
<i>Porphyromonas</i> spp.	1.6 (5.4)	1.2 (2.5)	0.79
<i>Leptotrichia</i> spp.	1.1 (2.7)	1.0 (2.6)	0.83
<i>Rothia</i> spp.	0.5 (1.0)	4.5 (15.6)	0.42
<i>Burkholderia</i> spp.	0.1 (0.1)	0.00 (0.02)	0.35
<i>Pseudomonas</i> spp.	0.04 (0.1)	0.04 (0.1)	0.90
<i>Ralstonia</i> spp.	0.04 (0.1)	0.01 (0.03)	0.42

<https://doi.org/10.1371/journal.pone.0201156.t002>

organism. Quantitative PCR was not performed as part of this study and may have helped determine whether low biomass contributed to this observation. Therefore, where samples are dominated by an individual genus, the results should be treated with caution and repeat upper airway sampling paired with culture dependent microbiology is wise. Upper airway samples may therefore fail to detect dominant pathogens in the lower airways and vice versa, limiting clinical use in an individual child but also a consideration for the design of research studies.

Two main factors influenced the degree of variation in the airway microbiota. The patient sampled had the greatest influence (53.8% of variation, Bray Curtis dissimilarity score). This confirms that the microbiota is individual to a patient[22–24]. The underlying pathology accounted for up to 22.5% of variation in TS, suggesting TS detect disease differences

**Table 3. Beta diversity summaries of significant clinical variables influencing community structure.** Adonis (PERMANOVA) results shown are for those variables which were statistically significant ( $P < 0.05$ ) using the Bray Curtis dissimilarity score. IV—intravenous. FOB—fiberoptic bronchoscopy.

Variable & diversity measurement	$r^2$	P-value
Sample Type	0.06	0.001
Patient ID	0.54	0.001
Underlying pathology	0.20	0.001
Lower respiratory tract symptoms	0.02	0.003
Nebulised antibiotics	0.02	0.003
Current IV antibiotics	0.01	0.012
FOB route	0.08	0.001

<https://doi.org/10.1371/journal.pone.0201156.t003>

which may be useful in longitudinal studies of the airway microbiota when comparing different patient groups.

Our findings in 49 children are similar to those of Charlson *et al.*[13], who demonstrated similarity of the microbiota along the respiratory tree in 6 adults (including smokers), and Boutin *et al.*[10] who demonstrated similarities between TS and sputum samples in 20 adults and children with CF with a mean age of 16.1 years. Similarly, Marsh *et al.*[25] compared upper airway samples (nasopharyngeal and oropharyngeal swabs) with BALF sampled from a single lobe and found upper airway samples were a reliable surrogate in 69% of children with either idiopathic bronchiectasis, protracted bacterial bronchitis or healthy airways. Our current study we believe is the first to compare the microbiota using both TS and CS and lower airways samples in young children with CSLD (CF or PCD) and non-CSLD controls (mainly recurrent LRTI), which is representative of the pathologies frequently encountered in Paediatric Respiratory Medicine. This is important since children are less able to spontaneously expectorate and in whom finding a reliable surrogate for lower airway sampling which can be obtained frequently is particularly pertinent.

Our study may be underpowered as we were unable to perform an *a priori* power calculation (S1 Appendix). A significant difference in beta-diversity was however observed meaning differences between groups of patient samples could be detected. Nonetheless caution should be taken in relation to the interpretation of the CS data given the low sample size ( $N = 17$ ) and their very low biomass, rendering them at increased risk of contamination by spurious OTUs. Low biomass also will have contributed to the variation in sequencing depth between samples and the need to rarefy to a level balancing capturing OTUs and retaining sufficient samples for paired comparisons. However, higher sequencing depths do not lead to improved identification of ecological patterns[26].

Eighteen patients (37%) had received a course of antibiotics in the previous 30 days and 41% were prescribed prophylactic antibiotics. Whilst this could potentially introduce a bias in our results by influencing both the upper and lower airway microbiota so that they show greater similarity, it would not have been ethically permissible to stop a clinically indicated treatment for the purposes of this study, and indeed many of these patients are prescribed antibiotics in routine clinical practice. Nonetheless, we accept that extrapolation of our results to antibiotic-naïve children should be cautious.

Ideally, all bronchoscopies would have been performed either via an LMA or endotracheal tube in order to minimize risk of contamination of the bronchoscope. The route was determined by clinical considerations. Bronchoscopy route accounted for up to 8.16% variance in community structure, much less than the variance due to the individual patient (53.8%) or the underlying disease (22.5%). Greater abundance of organisms associated with the nasal passages such as *Corynebacterium* was seen in samples where bronchoscopy was performed transnasally. Similarly, contamination of upper airway samples with the oral microbiota cannot be excluded, although care was taken to limit for this by avoiding contact with the oral cavity during sampling and not using suction until the bronchoscope was below the vocal cords. Nasopharyngeal samples were not collected and have previously been found to be highly diverse in young children[27].

BALF was only collected and pooled from two lung lobes in this study. As lobar differences exist in bacterial distribution[28], ideally all six lobes would have been sampled to determine geographical consistency. More work is needed on intra-lobar differences and the relationship with upper airway cultures.

A small significant batch effect was seen between sample plates using the Bray Curtis and weighted UniFrac scores ( $r^2 = 3.32\%$  and  $3.82\%$  respectively). Except for 10 samples, all

samples from an individual patient were however, run on the same plate thereby limiting the impact of any batch effect.

In summary, CS sequenced poorly and their use cannot be recommended for non-culture based microbiota studies. Considering TS as a surrogate for lower airway samples, although representative at a community level and at this level can demonstrate disease differences, throat swabs can show substantial differences at the individual patient level. Notably larger differences are observed when samples are dominated by an individual organism, the latter being identified by both molecular and culture-dependent techniques. A combination of these two techniques may be an important consideration and advantageous in order to obtain a comprehensive assessment of a patient's airway bacterial community. Consequently TS do not have utility for individual clinical decision making but they provide an opportunity as a research tool for tracking longitudinal changes in groups of patients with, for example, CF and PCD.

## Supporting information

**S1 Appendix. Supplementary methods and results for the comparison of the upper and lower airway microbiota in children with chronic lung diseases.**

(DOCX)

**S1 Fig. Illustration of the methodological steps in sample processing from DNA extraction to 16S rRNA gene sequencing using the Illumina MiSeq.**

(TIF)

**S2 Fig. Diagram illustrating the analysis pipeline for sequences obtained from the Illumina MiSeq.** Upstream analyses were performed in QIIME and downstream analyses were performed in Phyloseq in R.

(TIF)

**S3 Fig. Rarefaction curves with yellow lines denoting the number of OTUs sampled at 1,000 reads and 3,000 reads.** This illustrates that an asymptote is reached by 1,000 reads. At this threshold, the majority of OTUs have been sampled and little additional information is obtained at higher rarefaction levels. Consequently a rarefaction level of 1,000 reads was chosen.

(TIF)

**S4 Fig. Illustrating individual patient barplots (N = 40) organised from those showing the greatest similarity between upper and lower airway samples to those showing the least similarity (determined by Bray Curtis dissimilarity).** Only BALF samples were sent for bacterial culture as part of routine clinical care. The results of BALF culture and disease group are also detailed.

(PDF)

**S5 Fig. Bland Altman plots showing agreement between TS and lower airway samples in alpha diversity measurements illustrated by (a) richness, (b) evenness and (c) Shannon Diversity Index.** Overall agreement is seen between samples, apart from at low levels of evenness and Shannon Diversity.

(TIF)

**S1 Table. Summary of sequencing quality statistics.**

(DOCX)



## Author Contributions

**Conceptualization:** Bushra Ahmed, Michael J. Cox, William O. C. Cookson, Jane C. Davies, Miriam F. Moffatt, Andrew Bush.

**Data curation:** Bushra Ahmed, Michael J. Cox, Leah Cuthbertson, Phillip L. James.

**Formal analysis:** Bushra Ahmed, Michael J. Cox, Leah Cuthbertson, Phillip L. James.

**Funding acquisition:** William O. C. Cookson, Jane C. Davies, Miriam F. Moffatt, Andrew Bush.

**Investigation:** Bushra Ahmed.

**Methodology:** Bushra Ahmed, Michael J. Cox, Leah Cuthbertson, Phillip L. James, Jane C. Davies, Andrew Bush.

**Project administration:** Bushra Ahmed, Michael J. Cox, William O. C. Cookson, Jane C. Davies, Miriam F. Moffatt, Andrew Bush.

**Supervision:** Michael J. Cox, William O. C. Cookson, Jane C. Davies, Miriam F. Moffatt, Andrew Bush.

**Writing – original draft:** Bushra Ahmed.

**Writing – review & editing:** Bushra Ahmed, Michael J. Cox, Leah Cuthbertson, Phillip L. James, William O. C. Cookson, Jane C. Davies, Miriam F. Moffatt, Andrew Bush.

## References

1. Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, et al. (2010) Disordered Microbial Communities in Asthmatic Airways. *PLoS ONE* 5: e8578. <https://doi.org/10.1371/journal.pone.0008578> PMID: 20052417
2. van der Gast CJ, Walker AW, Stressmann FA, Rogers GB, Scott P, Daniels TW, et al. (2011) Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *ISME J* 5: 780–791. <https://doi.org/10.1038/ismej.2010.175> PMID: 21151003
3. Tunney MM, Klem ER, Fodor AA, Gilpin DF, Moriarty TF, McGrath SJ, et al. (2011) Use of culture and molecular analysis to determine the effect of antibiotic treatment on microbial community diversity and abundance during exacerbation in patients with cystic fibrosis. *Thorax* 66: 579–584. <https://doi.org/10.1136/thx.2010.137281> PMID: 21270069
4. Rogers GB, Carroll MP, Serisier DJ, Hockey PM, Kehagia V, Jones GR, et al. (2005) Bacterial activity in cystic fibrosis lung infections. *Respir Res* 6: 49. <https://doi.org/10.1186/1465-9921-6-49> PMID: 15929792
5. Laguna TA, Wagner BD, Williams CB, Stevens MJ, Robertson CE, Welchlin CW, et al. (2016) Airway Microbiota in Bronchoalveolar Lavage Fluid from Clinically Well Infants with Cystic Fibrosis. *PLoS One* 11: e0167649. <https://doi.org/10.1371/journal.pone.0167649> PMID: 27930727
6. Frayman KB, Armstrong DS, Carzino R, Ferkol TW, Grimwood K, Storch GA, et al. (2017) The lower airway microbiota in early cystic fibrosis lung disease: a longitudinal analysis. *Thorax* 72: 1104–1112. <https://doi.org/10.1136/thoraxjnl-2016-209279> PMID: 28280235
7. Mika M, Korten I, Qi W, Regamey N, Frey U, Casaulta C, et al. (2016) The nasal microbiota in infants with cystic fibrosis in the first year of life: a prospective cohort study. *Lancet Respir Med* 4: 627–635. [https://doi.org/10.1016/S2213-2600\(16\)30081-9](https://doi.org/10.1016/S2213-2600(16)30081-9) PMID: 27180018
8. Prevaes SM, de Winter-de Groot KM, Janssens HM, de Steenhuijsen Piters WA, Tramper-Stranders GA, Wyllie AL, et al. (2016) Development of the Nasopharyngeal Microbiota in Infants with Cystic Fibrosis. *Am J Respir Crit Care Med* 193: 504–515. <https://doi.org/10.1164/rccm.201509-1759OC> PMID: 26492486
9. Biesbroek G, Tsivtsivadze E, Sanders EA, Montijn R, Veenhoven RH, Keijsers BJ, et al. (2014) Early respiratory microbiota composition determines bacterial succession patterns and respiratory health in children. *Am J Respir Crit Care Med* 190: 1283–1292. <https://doi.org/10.1164/rccm.201407-1240OC> PMID: 25329446

10. Boutin S, Graeber SY, Weitnauer M, Panitz J, Stahl M, Clausnitzer D, et al. (2015) Comparison of Microbiomes from Different Niches of Upper and Lower Airways in Children and Adolescents with Cystic Fibrosis. *PLoS ONE* 10: e0116029. <https://doi.org/10.1371/journal.pone.0116029> PMID: 25629612
11. Armstrong DS, Grimwood K, Carlin JB, Carzino R, Olinsky A, Phelan PD (1996) Bronchoalveolar lavage or oropharyngeal cultures to identify lower respiratory pathogens in infants with cystic fibrosis. *Pediatr Pulmonol* 21: 267–275. [https://doi.org/10.1002/\(SICI\)1099-0496\(199605\)21:5<267::AID-PPUL1>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1099-0496(199605)21:5<267::AID-PPUL1>3.0.CO;2-K) PMID: 8726151
12. Jung A, Kleinau I, Schonian G, Bauernfeind A, Chen C, Griese M, et al. (2002) Sequential genotyping of *Pseudomonas aeruginosa* from upper and lower airways of cystic fibrosis patients. *European Respiratory Journal* 20: 1457–1463. PMID: 12503704
13. Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, et al. (2011) Topographical continuity of bacterial populations in the healthy human respiratory tract. *Am J Respir Crit Care Med* 184: 957–963. <https://doi.org/10.1164/rccm.201104-0655OC> PMID: 21680950
14. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41: e1. <https://doi.org/10.1093/nar/gks808> PMID: 22933715
15. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79: 5112–5120. <https://doi.org/10.1128/AEM.01043-13> PMID: 23793624
16. Dixon P (2003) VEGAN, a package for R functions for community ecology. *Journal of Vegetation Science* 14: 927–930.
17. Lozupone CA, Hamady M, Kelley ST, Knight R (2007) Quantitative and Qualitative  $\beta$  Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. *Applied and Environmental Microbiology* 73: 1576–1585. <https://doi.org/10.1128/AEM.01996-06> PMID: 17220268
18. Farrell PM, Rosenstein BJ, White TB, Accurso FJ, Castellani C, Cutting GR, et al. (2008) Guidelines for diagnosis of cystic fibrosis in newborns through older adults: Cystic Fibrosis Foundation consensus report. *J Pediatr* 153: S4–S14. <https://doi.org/10.1016/j.jpeds.2008.05.005> PMID: 18639722
19. Lucas JS, Barbato A, Collins SA, Goutaki M, Behan L, Caudri D, et al. (2017) European Respiratory Society guidelines for the diagnosis of primary ciliary dyskinesia. *Eur Respir J* 49.
20. Staffer P, Davies JC, Balfour-Lynn IM, Rosenthal M, Bush A (2011) Bronchoscopy in cystic fibrosis infants diagnosed by newborn screening. *Pediatr Pulmonol* 46: 696–700. <https://doi.org/10.1002/ppul.21434> PMID: 21365781
21. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. (2014) Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 12: 87. <https://doi.org/10.1186/s12915-014-0087-z> PMID: 25387460
22. Oh J, Byrd AL, Deming C, Conlan S, Program NCS, Kong HH, et al. (2014) Biogeography and individuality shape function in the human skin metagenome. *Nature* 514: 59–64. <https://doi.org/10.1038/nature13786> PMID: 25279917
23. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, et al. (2009) A core gut microbiome in obese and lean twins. *Nature* 457: 480–484. <https://doi.org/10.1038/nature07540> PMID: 19043404
24. Utter DR, Mark Welch JL, Borisy GG (2016) Individuality, Stability, and Variability of the Plaque Microbiome. *Frontiers in Microbiology* 7: 564. <https://doi.org/10.3389/fmicb.2016.00564> PMID: 27148241
25. Marsh RL, Kaestli M, Chang AB, Binks MJ, Pope CE, Hoffman LR, et al. (2016) The microbiota in bronchoalveolar lavage from young children with chronic lung disease includes taxa present in both the oropharynx and nasopharynx. *Microbiome* 4: 37. <https://doi.org/10.1186/s40168-016-0182-1> PMID: 27388563
26. Kuczynski J, Liu Z, Lozupone C, McDonald D, Fierer N, Knight R (2010) Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. *Nat Methods* 7: 813–819. <https://doi.org/10.1038/nmeth.1499> PMID: 20818378
27. Bogaert D, Keijsers B, Huse S, Rossen J, Veenhoven R, van Gils E, et al. (2011) Variability and Diversity of Nasopharyngeal Microbiota in Children: A Metagenomic Analysis. *PLOS ONE* 6: e17035. <https://doi.org/10.1371/journal.pone.0017035> PMID: 21386965
28. Gutierrez JP, Grimwood K, Armstrong DS, Carlin JB, Carzino R, Olinsky A, et al. (2001) Interlobar differences in bronchoalveolar lavage fluid from children with cystic fibrosis. *European Respiratory Journal* 17: 281–286. PMID: 11334132