



ELSEVIER

Contents lists available at ScienceDirect

Journal of Cystic Fibrosis

journal homepage: www.elsevier.com/locate/jcf

Original Article

The lung microbiota in children with cystic fibrosis captured by induced sputum sampling

Rebecca Weiser^{a,*}, Juliette Oakley^{b,c}, Katherine Ronchetti^{b,d}, Jo-Dee Tame^{b,d}, Sven Hoehn^e, Tomasz P. Jurkowski^e, Eshwar Mahenthiralingam^a, Julian T. Forton^{b,c}^a Microbiomes Microbes and Informatics Group, Organisms and Environment Division, School of Biosciences, Cardiff University, Sir Martin Evans Building, Park Place, Cardiff, UK^b Department of Paediatric Respiratory Medicine, Noah's Ark Children's Hospital for Wales, Cardiff, UK^c School of Medicine, Cardiff University, UK^d Department of Paediatric Physiotherapy, Noah's Ark Children's Hospital for Wales, Cardiff, UK^e Molecular Biosciences Division, School of Biosciences, Cardiff University, Sir Martin Evans Building, Park Place, Cardiff, UK

ARTICLE INFO

Article history:

Received 24 October 2021

Revised 15 December 2021

Accepted 9 January 2022

Available online xxx

Keywords:

Microbiota

Microbiome

Cystic fibrosis

Induced sputum

Sputum induction

ABSTRACT

Background: Spatial topography of the cystic fibrosis (CF) lung microbiota is poorly understood in childhood. How best to sample the respiratory tract in children for microbiota analysis, and the utility of microbiota profiling in clinical management of early infection remains unclear. By comparison with bronchoalveolar lavage (BAL), we assessed the ability of induced sputum (IS) sampling to characterise the lower airway microbiota.

Methods: Sample sets from IS and two or three matched BAL compartments were obtained for microbiota analysis as part of the CF-Sputum Induction Trial (UKCRN_14615, ISRCTNR_12473810). Microbiota profiles and pathogen detection were compared between matched samples.

Results: Twenty-eight patients, aged 1.1–17.7 years, provided 30 sample sets. Within-patient BAL comparisons revealed spatial heterogeneity in 8/30 (27%) sample sets indicating that the lower airway microbiota from BAL is frequently compartmentalised in children with CF. IS samples closely resembled one or more matched BAL compartments in 15/30 (50%) sets, and were related in composition in a further 9/30 (30%). IS detected 86.2% of the Top 5 genera found across matched BAL samples. The sensitivity of IS to detect specific CF-pathogens identified in matched BAL samples at relative abundance $\geq 5\%$ varied between 43 and 100%, with negative predictive values between 73 and 100%.

Conclusions: Spatial heterogeneity of the lower airway microbiota was observed in BAL samples and presents difficulties for consistent lung sampling. IS captured a microbiota signature representative of the lower airway in 80% of cases, and is a straightforward, non-invasive intervention that can be performed frequently to aid pathogen diagnosis and understand microbiota evolution in children with CF.

© 2022 The Authors. Published by Elsevier B.V. on behalf of European Cystic Fibrosis Society. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

Lung infection and progressive lung disease is the major cause of morbidity and mortality in cystic fibrosis (CF) [1]. Identifying and treating lower airway infection is therefore critical. This is comparatively straightforward in adults and older children, who have a high bacterial load, frequently have pathogen-dominated disease and can spontaneously expectorate. Younger children with

CF have milder disease, less pathogen-dominance, lower bacterial load and are often incapable of spontaneously expectorating sputum even if actively coughing during an exacerbation.

Oropharyngeal cough swabs or throat swabs are a well-tolerated approach to bacterial surveillance but are poorly representative of the lower airway [2]. Bronchoalveolar lavage (BAL) is considered to be the gold standard [3] but is an invasive procedure and the potential detrimental effects of repeated general anaesthetic on child development suggests these procedures should be minimised [4]. Sputum induction is a simple, well tolerated and frequently repeatable approach to sampling the lower airway in children who are unable to expectorate spontaneously [5–9].

* Corresponding author at: Microbiomes Microbes and Informatics Group, Organisms and Environment Division, School of Biosciences, Cardiff University, Sir Martin Evans Building, Park Place, Cardiff, CF10 3AX, UK.

E-mail address: WeiserR@Cardiff.ac.uk (R. Weiser).

<https://doi.org/10.1016/j.jcf.2022.01.006>

1569-1993/© 2022 The Authors. Published by Elsevier B.V. on behalf of European Cystic Fibrosis Society. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

In the CF Sputum Induction Trial (CF-SpIT) [6] where induced sputum samples (IS) were compared to matched six-lobe BAL, IS performed as well as the gold standard two-lobe BAL in a sensitivity analysis to detect all pathogens identified on all samples using culture microbiology. These observations suggested that IS is capable of sampling widely from the bronchoalveolar compartment and is a composite reflection of multiple lower airway niches. To extend these investigations, we applied culture-independent microbiota analysis to comprehensively evaluate the bacterial diversity in samples from the CF-SpIT trial. This unique collection of within-patient matched samples originating from four lower airway niches (IS and two or three BAL samples) enabled us to interrogate the spatial topography of the lower airway microbiota in detail, quantify the ability of IS to correctly describe it and identify broad age-related microbiota trends.

2. Methods

2.1. Study design and participants

CF-SpIT is a prospective internally-controlled interventional trial performed at the Children's Hospital for Wales, Cardiff, UK, in children with CF aged between 6 months and 18 years. Samples from sputum induction, and single-lobe, two-lobe, and six-lobe BAL were matched for within-patient comparison, with the aim of testing IS as an infection diagnostic. Results from CF-SpIT using conventional microbiology are published elsewhere [6]. This study is subject to Institutional Review by the Cardiff and Vale Research Review Service (CaRRS; Project-ID-11-RPM-5216) and approved by the South Wales Research Ethics Committee (11/WA/0334). This study is registered with the UK Clinical Research Network (14,615) and with the International Standard Randomised Controlled Trial Network Registry (12,473,810).

2.2. Respiratory sample collection and DNA extraction

Sputum induction and BAL procedures were performed as previously described [6]. BAL fluid from the right middle lobe was collected and labelled as BAL1, from the left lingula as BAL2, and from the combination of right upper lobe, right lower lobe, left upper lobe and left lower lobe as BAL3. All samples were divided immediately and one aliquot frozen at -80°C within 30 min of collection. A DNA extraction protocol adapted from a previous study [6] was applied to respiratory samples as described in Supplementary Materials.

2.3. Quantitative PCR (qPCR)

Quantification of total bacterial load was performed using a SYBR Green qPCR assay and primers targeting the 16S rRNA gene [10] to determine total bacterial load for removal of contaminant sequence reads (see Supplementary Materials).

2.4. 16S rRNA gene sequencing and bacterial diversity analysis

Sample library preparation and 16S rRNA gene sequencing was performed at the Cardiff University Genomics Research Hub. The V4-region of the 16S rRNA gene was amplified and sequenced as previously described [11], generating 250 bp paired-end reads on the Illumina MiSeq platform. Bioinformatic analysis was performed using Mothur [12] v1.39.5 using the MiSeq SOP pipeline. After removal of contaminant sequence reads, downstream statistical analyses were performed using R statistical software [13], Microsoft Excel and SPSS. OTUs were consolidated into groups at the lowest level of taxonomic hierarchy assigned by Mothur (genus or family)

as the 16S rRNA hypervariable regions lack resolution for species-level identification. *Haemophilus*, *Pseudomonas*, *Staphylococcus*, *Enterobacteriaceae*, *Stenotrophomonas* and *Burkholderia* were considered relevant CF lung pathogens. We acknowledge that this approach could lead to a mixture of different species (including 'non-typical' pathogens) being consolidated into these pathogen groups. See Supplementary Materials for full details. Bioinformatics scripts and R code for all analyses is available at <https://github.com/Beky-Weiser/CFSpIT-16S-Microbiota-Analysis>.

2.5. Data availability

Raw sequence data have been submitted to the European Nucleotide Archive under project number PRJEB34389.

3. Results

3.1. Patient demographics and sample sets

A total of 120 respiratory samples in 30 matched BAL1-BAL2-BAL3-IS sets were obtained from 28 paediatric patients aged between 1.1 and 17.7 years of age (Supplementary Table S1). Microbiota profiles were obtained for 116 respiratory samples which resulted in 26 matched sets of BAL1-BAL2-BAL3-IS samples and 4 matched sets with 2 BAL samples and 1 IS sample. Each individual possessed a unique microbiota profile (see Supplementary Figure S1).

3.2. Bacterial communities in IS samples were higher diversity than in BAL but overlapped in composition

The bacterial diversity in BAL samples was highly variable. This broad distribution of diversity was similar for niches BAL1, BAL2 and BAL3 (alpha diversity as determined by the Shannon index; Fig. 1a). The bacterial diversity of IS samples was less variable and significantly higher than BAL samples (Fig. 1a; $P < 0.001$). IS samples shared a baseline of 'non-typical' pathogens such as *Prevotella*, *Veillonella* and *Neisseria*, but also overlapped in composition with BAL1, BAL2 and BAL3 samples (beta diversity as determined by Bray-Curtis dissimilarity; Fig. 1b), indicating that other bacterial genera, including 'typical' pathogens such as *Pseudomonas*, *Haemophilus* and *Staphylococcus*, were shared between the different sample types.

3.3. Analysis of BAL samples revealed microbiota compartmentalisation in the lower airways

In order to further investigate the variation in bacterial diversity seen in BAL samples (Fig. 1), we initially limited the analysis to the 78 samples in the 26 matched BAL1-BAL2-BAL3 sets. Pathogen detection was evaluated using different relative abundance thresholds for any one BAL sample (Supplementary Figure S2a) or for all 3 BAL samples in a given sample set (Supplementary Figure S2b). Known pathogens (*Haemophilus*, *Pseudomonas*, *Staphylococcus*, *Enterobacteriaceae*, *Stenotrophomonas* and *Burkholderia*) were detectable at high prevalence with one or more pathogen detectable in 100% of BAL sample sets at the presence/absence level, and *Pseudomonas* was specifically detectable in 23/26 (88%) sample sets (Supplementary Figure S2a). Pathogen detection at increasing relative abundance thresholds remained high with 25/26 (96%) sample sets positive for at least one pathogen at relative abundance $\geq 5\%$ (Supplementary Figure S2a). In total, 40 pathogens were identified at $\geq 5\%$ relative abundance, although pathogen identification in matched BAL samples often differed. In 12/40 (30%) cases, the pathogen was present in one of three matched BAL niches, in 9/40

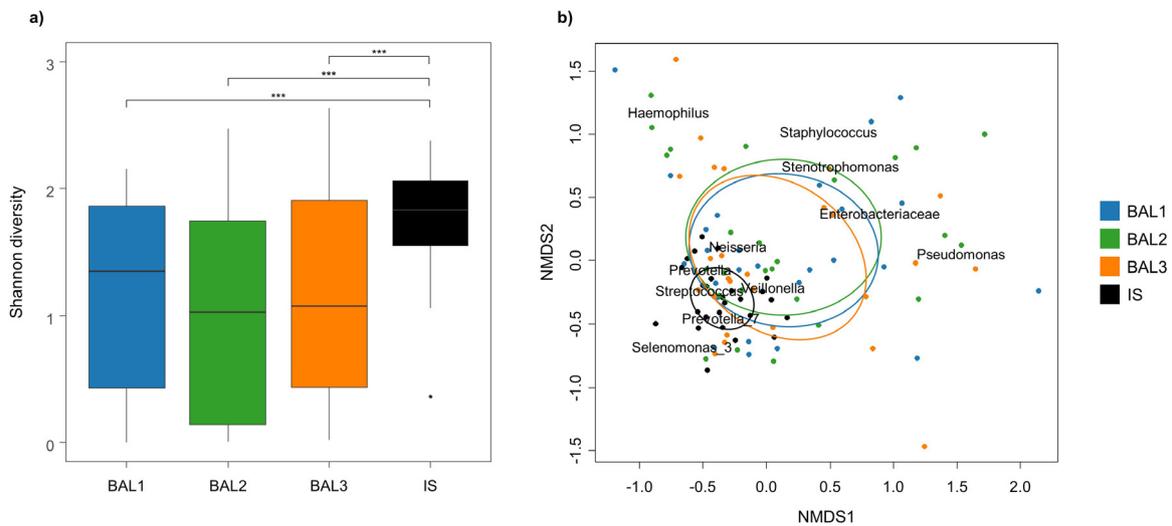


Fig. 1. Bacterial diversity and community composition of BAL and IS sample types. (a) Comparison of the Shannon diversity index (alpha diversity) between BAL and IS sample types. BAL samples exhibit a wider variation in diversity than IS. Diversity was significantly higher for IS compared to BAL samples ($***p < 0.001$). (b) NMDS ordination plot of Bray-Curtis dissimilarity distances (beta diversity) for BAL and IS sample types. BAL samples exhibit similar overall community compositions, but samples across the group of individuals were widely distributed. IS samples clustered closely together and overlapped with all 3 BAL sample types. Points represent individual samples, ellipses are standard deviations of points scores for each sampling type and the top 7.5% genera/families based on total abundance across the dataset are shown.

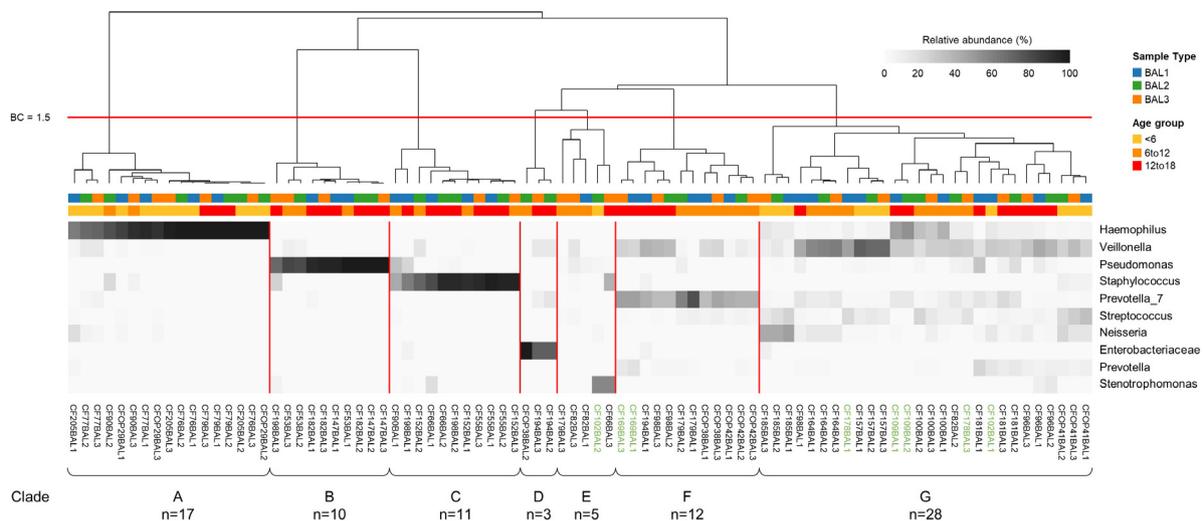


Fig. 2. Hierarchical cluster analysis of the microbiota in BAL samples and its association with age group. Hierarchical cluster analysis of Bray-Curtis dissimilarity distances was performed using Ward's method. Based on a Bray-Curtis dissimilarity distance cut-off of 1.5, the 86 BAL samples segregated into 7 clades (Clades A-G). Clades A-D are pathogen-dominant, characterised by known CF pathogens. Clade E contains a mix of sample profile types, and clades F and G are diverse. The heatmap illustrates the relative abundance of the top 10 most abundant genera/families identified across the dataset for each BAL sample, with the gradient colour scale shown at the top right and the number of samples belonging to each clade at the bottom of the heatmap. Heatmap column annotations indicate the sample type and patient age and a key is given at the top right of the figure. Sample codes shown in green indicate sample sets with only 2 BAL samples.

(23%) it was present in two of the three niches and in 19/40 (47%) it was present in all three niches.

Hierarchical cluster analysis of all BAL microbiota profiles ($n = 86$) using Bray-Curtis dissimilarity distances revealed clades of related microbiota profiles (Fig. 2). The microbiota segregated into 7 broad clades, 5 of which contained samples that were characterised by the presence of known CF lung pathogens (Fig. 2; Clades A-E), and two of which exhibited high microbiota diversity and were largely pathogen free (Clades F and G). The pathogen positive clades were characterised by high relative abundance of *Haemophilus* (Clade A), *Pseudomonas* (Clade B), *Staphylococcus* (Clade C), *Enterobacteriaceae* (Clade D) and *Stenotrophomonas* (Clade E). The pathogen-negative diverse groupings were associated with the presence of *Prevotella*, *Veillonella*, *Streptococcus* and *Neisseria*, and segregated specifically into clades characterised by

high relative abundance of *Prevotella* (Clade F) and *Veillonella* (Clade G), respectively.

Pathogen-dominated clades (Fig. 2; Clades A-E) were significantly associated with patient age ($p = 0.001$), with *Haemophilus* predominantly present in the younger age group (less than 6 years), and *Pseudomonas*, *Staphylococcus* and *Enterobacteriaceae* in the intermediate and older age groups (6–12 and 12–18 years). These associations were also apparent when NMDS was used to compare community composition between individuals in different age groups; microbiota profiles in younger children were often characterised by *Haemophilus* together with non-traditional CF genera *Neisseria* and *Veillonella* (Supplementary Figure S3). The microbiota in older children was characterised by *Pseudomonas*, *Staphylococcus*, *Enterobacteriaceae* and *Stenotrophomonas* (Supplementary Figure S3).

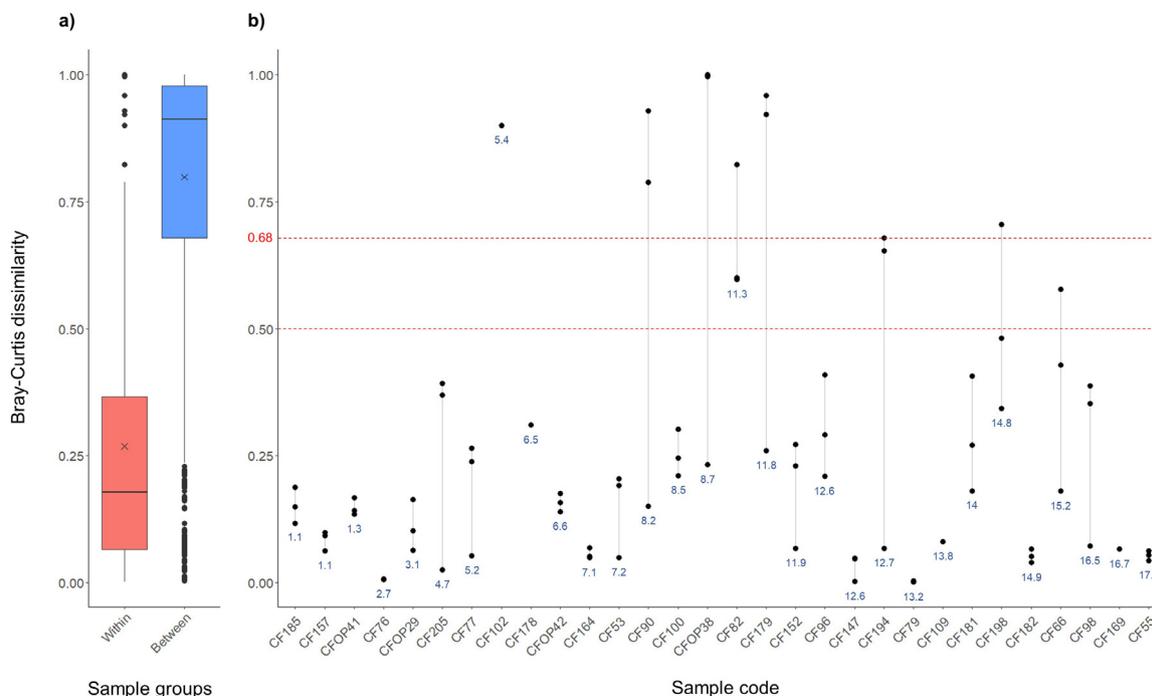


Fig. 3. Comparison of within-patient and between-patient microbiota diversity of BAL samples. Bray-Curtis dissimilarity distances for within-patient and between-patient comparisons were made and shown as follows. (a) A summary of BAL within-patient comparisons (median = 0.18 (IQR 0.06–0.37) and between-patient comparisons (median = 0.91 (IQR 0.68 – 0.98). The mean values are shown as a cross inside the boxplot (within = 0.27; between = 0.80). (b) Bray-Curtis dissimilarity distances between BAL samples taken from individual patients are visualised using strip charts; three values are plotted per patient (BAL1-BAL2, BAL1-BAL3 and BAL2-BAL3 distances), points are joined to highlight the spread. Patient age increases from left to right and is displayed in years in blue underneath the Bray-Curtis dissimilarity scores. When all three values are ≤ 0.5 (highlighted by the red dotted line), BAL1, BAL2 and BAL3 microbiota profiles are considered to be concordant. The first quartile of the between-patient data (0.68) has also been highlighted by a red dotted line to illustrate that 7 within-patient comparisons fall on or above this line and are highly discordant.

Matched BAL samples from the same individual only segregated into the same clade in hierarchical cluster analysis in 21/30 (70%) cases, suggesting heterogeneity in the lower airway microbiota in these children (Fig. 2). To comprehensively investigate the relatedness of microbiota profiles between samples taken from different niches in the same patient, Bray-Curtis dissimilarity distances between BAL samples within- and between-patients were compared (Fig. 3). The degree of similarity observed between matched BAL samples taken from the same patient was generally high (i.e. low Bray-Curtis dissimilarity values; Fig. 3). Overall, the within-patient Bray-Curtis scores (median 0.18; IQR 0.06–0.37) were lower than the between-patient Bray-Curtis scores (median 0.91; IQR 0.68–0.98). However, in 7/30 (23%) cases, within-patient Bray-Curtis scores were highly dissimilar, and in the range seen for between-patient sample comparisons (>0.68), highlighting a high level of disparity and suggesting compartmentalisation of the lower airway (Fig. 3; CF102, CF90, CFOP38, CF82, CF179, CF194 and CF198). The microbiota from different lobes was generally concordant in children of less than 5 years old, but appeared more compartmentalised in older children between the ages of 5.4 and 14.8 (Fig. 3).

3.4. IS captured a meaningful representation of the lower airway microbiota in 80% of cases

Having established that the lower airway microbiota from BAL sampling may be non-uniform across multiple compartments, we next compared microbiota from all matched IS and BAL samples ($n = 120$). The Bray-Curtis dissimilarity distance between BAL and IS pairs from each patient was used to define concordant (Bray-Curtis dissimilarity ≤ 0.5) and discordant (Bray-Curtis dissimilarity > 0.5) microbiota profiles (Supplementary Table S2). Based on this analysis, 8 microbiota profile types were identified (Table 1). Ex-

amples of each microbiota profile type are given in Fig. 4 and the full set is displayed in Supplementary Figure S1.

In 12/30 (40%) matched sample sets where BAL profiles were concordant, IS captured the pattern of bacterial diversity seen in BAL with examples of both microbiota-diverse and pathogen-dominated states (Table 1, profile types 1 and 3). In a further 3/30 (10%) cases, BAL diversity profiles were discordant with one another, but IS microbiota profiles matched one or two of the BAL compartments (Table 1, profile 6). In 9/30 (30%) matched sample sets, IS was discordant with BAL by Bray-Curtis dissimilarity measures but contained similar taxa to BAL at different relative abundances, and identified the dominant taxon (Table 1, profile type 4) or the most prevalent taxa (Table 1, profile types 2 and 7). In 6/30 (20%) matched samples, IS diversity profiles were dissimilar to all matched BAL samples (Table 1, profile types 5 and 8).

To analyse this further, we assessed the ability of IS to capture the Top 5 and Top 10 genera present across matched 6-lobe BAL samples (BAL1+BAL2+BAL3). The top 5 and top 10 most prevalent genera respectively accounted for an average of 90.0% (standard deviation 9.1) and 97.0% (standard deviation 3.4) of the total bacterial abundance in any BAL sample. Using presence or absence as detection criteria, IS captured 86.2% (standard deviation 19.4) of the Top 5 and 78.8% (standard deviation 18.2) of the Top 10 genera found across the 26 sets of matched 6-lobe BAL samples.

3.5. Identification of specific lower airway pathogens on IS was variable

We next assessed the ability of IS to detect specific lower airway pathogens identified on 6-lobe BAL (BAL1+BAL2+BAL3) for the 26 complete matched sets. For this analysis, we arbitrarily defined a threshold of detection for any pathogen at $\geq 5\%$ relative abundance on BAL to be clinically relevant. At this threshold, 25/26

Table 1

IS and BAL profile definitions based on comparison of Bray-Curtis dissimilarity distances and shared taxa between sample types.

Profile type	BAL profile description	IS profile description	N°sets	Figure S1 codes
1	BAL1, BAL2 and BAL3 concordant (Bray-Curtis dissimilarity distance ≤ 0.5)	Microbiota diverse	10	A1-A10
2	BAL1, BAL2 and BAL3 discordant (Bray-Curtis dissimilarity distance > 0.5 for ≥ 1 BAL pairwise comparisons)	Microbiota diverse	1	B1
3	BAL1, BAL2 and BAL3 discordant (Bray-Curtis dissimilarity distance > 0.5 for ≥ 1 BAL pairwise comparisons)	Pathogen dominated	2	B2-B3
4	BAL1, BAL2 and BAL3 discordant (Bray-Curtis dissimilarity distance > 0.5 for ≥ 1 BAL pairwise comparisons)	Pathogen dominated	6	B4-B9
5	BAL1, BAL2 and BAL3 discordant (Bray-Curtis dissimilarity distance > 0.5 for ≥ 1 BAL pairwise comparisons)	Concordant to ≥ 1 BAL (Bray-Curtis dissimilarity distance ≤ 0.5)	3	B10-C2
6	BAL1, BAL2 and BAL3 discordant (Bray-Curtis dissimilarity distance > 0.5 for ≥ 1 BAL pairwise comparisons)	Discordant to all BALs (Bray-Curtis dissimilarity distance > 0.5)	3	B10-C2
7	BAL1, BAL2 and BAL3 discordant (Bray-Curtis dissimilarity distance > 0.5 for ≥ 1 BAL pairwise comparisons)	Concordant to ≥ 1 BAL (Bray-Curtis dissimilarity distance ≤ 0.5)	3	C3-C5
8	BAL1, BAL2 and BAL3 discordant (Bray-Curtis dissimilarity distance > 0.5 for ≥ 1 BAL pairwise comparisons)	Discordant to all BALs (Bray-Curtis dissimilarity distance > 0.5)	2	C6-C7
		Discordant to all BALs (Bray-Curtis dissimilarity distance > 0.5)	3	C8-C10

The 'Figure S1 codes' column relates to the positions of the profiles belonging to each profile type in Supplementary Figure S1 which shows all 30 matched BAL-IS sets.

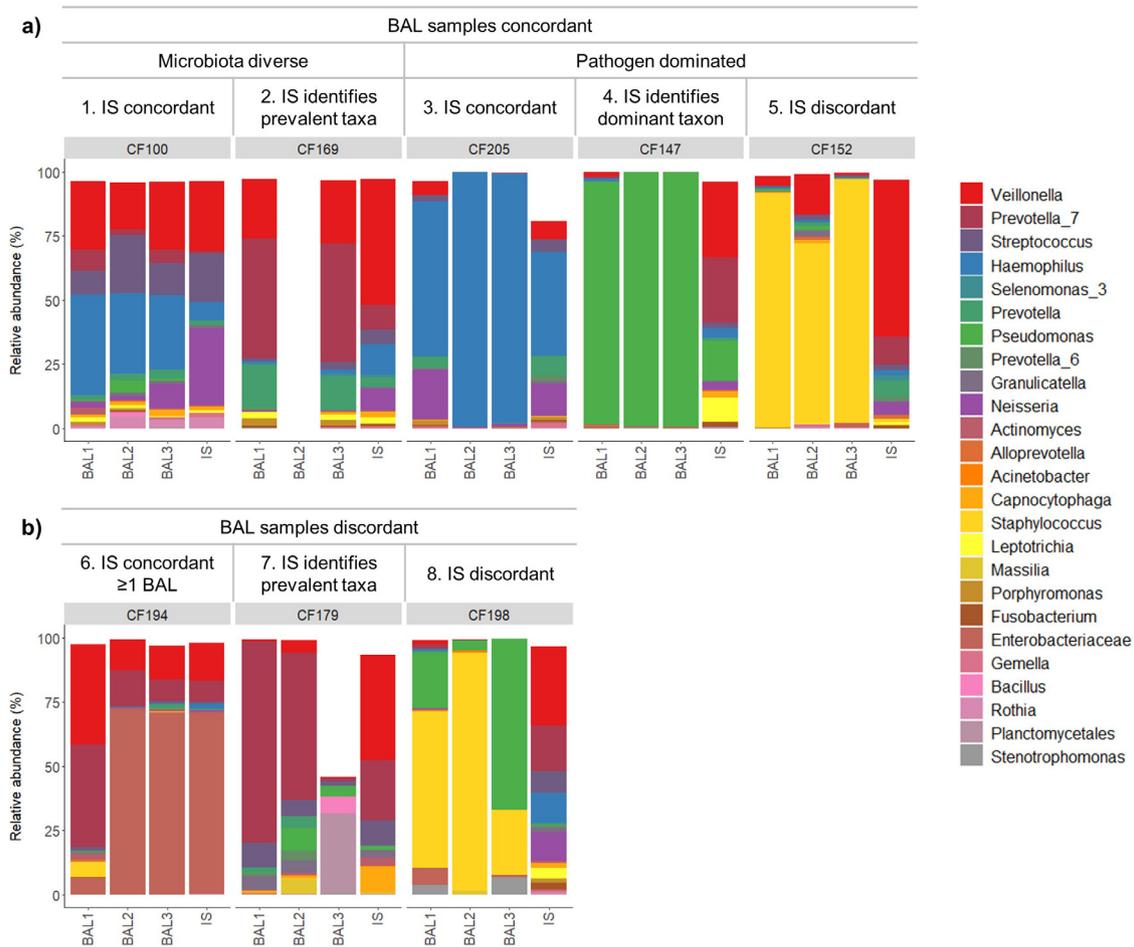


Fig. 4. Microbiota profiles of BAL1, BAL2, BAL3 and IS sample sets illustrating the 8 microbiota profile type definitions. Microbiota profiles are given for each matched sample set using relative abundance of the top 25 genera/families identified across the dataset. Samples are grouped by sample set number, and whether BAL samples are (a) concordant (≤ 0.5 Bray-Curtis dissimilarity), or (b) discordant (> 0.5 Bray-Curtis dissimilarity). Microbiota profile definition numbers are shown above each stacked bar chart and relate to Table 1. The taxa colour key is displayed at the right of the figure.

(96%) BAL sample sets were positive for a pathogen and all three BAL samples were positive in 20/26 sets (77%) (Supplementary Figure S2). Given that IS samples were of greater diversity than BAL samples (Fig. 1a) and consequently generally reported genera in lower relative abundances, we considered the presence of a pathogen on IS at any relative abundance to be potentially clinically important. At this presence/absence level, IS identified one or more pathogens in 25/26 (96%) samples. Sensitivities of IS to detect *Haemophilus*, *Pseudomonas* and *Staphylococcus* on BAL samples at $\geq 5\%$ relative abundance were 100%, 64% and 43% respectively. Negative predictive values for *Haemophilus*, *Pseudomonas* and *Staphy-*

lococcus were 100%, 73% and 79% respectively (see Supplementary Table S3 for further details). Unique and overlapping pathogen identification for matched samples is shown in Supplementary Figure S4.

Discussion

In this study we used samples collected in the CF-SpIT trial [6] to interrogate the applicability of molecular detection techniques in assessing lower airway microbiology in children with CF. This unique collection of within-patient, time-matched samples

from four lower airway niches (BAL and IS) enabled us to study the spatial variation in lower airway microbiota in children with CF, and to investigate the role of microbiota analysis in defining meaningful lower airway infection.

Microbiota from matched BAL samples from the same individual generally showed similar bacterial diversity profiles. Certain individuals exhibited diverse microbiota and others showed established pathogen-dominated microbiota across all matched BAL samples. However, there were exceptions, where within the same individual, certain BAL compartments were diverse, whereas others showed pathogen-dominance. In pathogen-positive individuals, the pathogen was identified at $\geq 5\%$ relative abundance in all three BAL niches in only 47% of instances, suggesting compartmentalisation of the lower airway microbiota is common in children. In a study of young children between 1 and 12 months old, microbiota analysis identified that BAL samples from the same individual were largely concordant [14]. Our study suggests that the microbiota from different lobes is generally concordant in children of less than 5 years old, but appears more compartmentalised in older children between the ages of approximately 5 to 15 years (Fig. 3). The lower airway compartmentalisation observed in this study potentially represents a transition stage in the evolution of the lower airway microbiota from the uniform balanced and diverse community seen in young children, to the uniform pathogen-dominated community seen in adults [15], and is therefore a key period where intervention may have greatest benefit.

Whilst we observed a general trend towards pathogen-dominance with increasing age, we identified individuals with compartmentalisation of the lower airway in children as young as 5 years of age. These findings suggest that evolution of the microbiota to pathogen-dominance in children with CF is not necessarily associated with chronological age, but could be associated with factors that impose selective pressure on the microbiota, such as infection, inflammation or antibiotic therapy. This raises the possibility of using microbiota composition as a biomarker of disease progression, to identify at an early stage, those children who are relatively accelerated in their evolution towards pathogen-dominance. To further establish this concept would require longitudinal surveillance of the evolving microbiota through childhood with regular, accurate sampling.

The CF-SpIT study validated induced sputum as an accurate, repeatable and reliable approach to sample the lower airway using conventional microbiology [6]. In this follow-on study, we have extended these findings by using culture-independent microbiota analysis. IS samples were more diverse and collected contributions from a wider respiratory niche than BAL samples, but there was clear overlap in composition of the bacterial communities detected by the different sample types. In 50% of individuals, bacterial composition and diversity of IS closely resembled all or at least one matched BAL sample (Table 1). In a further 30%, IS contained similar taxa to BAL but differed in the relative abundance of those bacteria present (Table 1), and performed well in a paediatric population where matched lower airway BAL samples themselves exhibit heterogeneity. From this analysis it is clear that the lower airway is a complex and complicated environment to describe, and capturing the full lung microbiota will require multiple sampling techniques. Longitudinal studies will help understand how IS performs in relation to identifying the emergence of pathogen-dominant lung infections.

Remarkably, in our analyses, DNA from known CF pathogens was present in 100% BAL sample sets. Culture-independent technologies are capable of detecting bacterial DNA at extremely low concentrations, and DNA from known pathogens may be detected where culture is negative [16,17]. This may represent early infection or low-level bystander, and presents a new problem in how we should define a clinically important infection using such sen-

sitive technologies. We pragmatically defined a pathogen relative abundance threshold of $\geq 5\%$ on BAL as representing a clinically meaningful lower airway infection, but there is currently no guidance as to what threshold of detection constitutes infection, or how microbiota analysis can be applied to the clinical management of polymicrobial infections such as those observed in CF. The long-term implications of low level pathogen carriage may become clearer through longitudinal sampling and relation to clinical outcomes.

Respiratory samples from children with mild disease are low biomass, and close to the threshold of reliable molecular detection. Analysis is therefore vulnerable to significant distortion by low biomass contamination [18] and decontamination protocols will be necessary (see Supplementary Materials for details of our decontamination protocols). In this study, we compared multiple time-matched samples with fine granularity and reassuringly found patterns replicated in multiple samples from the same individual, suggesting a genuine biological signal.

We have shown that in children with CF, the lower airway microbiota can be highly compartmentalised. IS in this study captured an informative representation of the lower airway microbiota in 80% of cases thereby validating it as an effective approach to sampling the lower airway using molecular detection techniques. IS sampling can be used to obtain frequent serial samples and is a candidate approach for future large-scale microbiota studies in children with CF, and in the clinical application of microbiota profiling in children.

Funding

Health and Care Research Wales-Academic Health Science Collaboration, Wellcome Trust Institutional Strategic Support Fund, US Cystic Fibrosis Foundation Grant (MAHENT20G0).

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the manuscript.

Declaration of Competing Interest

The authors have no conflicts of interest to disclose.

CRediT authorship contribution statement

Rebecca Weiser: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **Juliette Oakley:** Validation, Formal analysis, Investigation. **Katherine Ronchetti:** Methodology, Investigation, Writing – review & editing, Project administration. **Jo-Dee Tame:** Methodology, Investigation, Writing – review & editing, Project administration. **Sven Hoehn:** Methodology, Software, Validation, Investigation, Data curation, Writing – review & editing, Project administration. **Tomasz P. Jurkowski:** Methodology, Software, Validation, Resources, Writing – review & editing, Supervision, Project administration. **Eshwar Mahenthiralingam:** Conceptualization, Methodology, Validation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Julian T. Forton:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources,

Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Acknowledgements

Cerith Jones is acknowledged for technical assistance with organising DNA samples for microbial diversity analysis. Lena Thia and Iolo Doull are acknowledged for their contributions to sample collection.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.jcf.2022.01.006](https://doi.org/10.1016/j.jcf.2022.01.006).

References

- [1] Ramsey KA, Ranganathan S, Park J, et al. Early respiratory infection is associated with reduced spirometry in children with cystic fibrosis. *Am J Respir Crit Care Med* 2014;190(10):1111–16.
- [2] Rosenfeld M, Emerson J, Accurso F, et al. Diagnostic accuracy of oropharyngeal cultures in infants and young children with cystic fibrosis. *Pediatr Pulmonol* 1999;28(5):321–8.
- [3] Brennan S, Gangell C, Wainwright C, et al. Disease surveillance using bronchoalveolar lavage. *Paediatr Respir Rev* 2008;9(3):151–9.
- [4] Schneuer FJ, Bentley JP, Davidson AJ, et al. The impact of general anesthesia on child development and school performance: a population-based study. *Paediatr Anaesth* 2018;28(6):528–36.
- [5] Forton J. Induced sputum in young healthy children with cystic fibrosis. *Paediatr Respir Rev* 2015;16(1):6–8 Suppl.
- [6] Ronchetti K, Tame JD, Paisey C, et al. The CF-Sputum Induction Trial (CF-SpIT) to assess lower airway bacterial sampling in young children with cystic fibrosis: a prospective internally controlled interventional trial. *Lancet Respir Med* 2018;6(6):461–71.
- [7] Al-Saleh S, Dell SD, Grasemann H, et al. Sputum induction in routine clinical care of children with cystic fibrosis. *J Pediatr* 2010;157(6):1006–11 e1.
- [8] Blau H, Linnane B, Carzino R, et al. Induced sputum compared to bronchoalveolar lavage in young, non-expectorating cystic fibrosis children. *J Cyst Fibros* 2014;13(1):106–10.
- [9] D'Sylva P, Caudri D, Shaw N, et al. Induced sputum to detect lung pathogens in young children with cystic fibrosis. *Pediatr Pulmonol* 2017;52(2):182–9.
- [10] Nadkarni MA, Martin FE, Jacques NA, et al. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology* 2002;148(1):257–66.
- [11] Kozich JJ, Westcott SL, Baxter NT, et al. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 2013;79(17):5112–20.
- [12] Schloss PD, Westcott SL, Ryabin T, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 2009;75(23):7537–41.
- [13] R-Core-Team R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
- [14] Prevaes SMPJ, de Steenhuijsen Pijters WAA, de Winter-de Groot Karin M, et al. Concordance between upper and lower airway microbiota in infants with cystic fibrosis. *European Respiratory Journal* 2017;49(3):1602235.
- [15] Boutin S, Dalpke AH. Acquisition and adaptation of the airway microbiota in the early life of cystic fibrosis patients. *Mol Cell Pediatr* 2017;4(1):1.
- [16] van Belkum A, Renders NH, Smith S, et al. Comparison of conventional and molecular methods for the detection of bacterial pathogens in sputum samples from cystic fibrosis patients. *FEMS Immunol Med Microbiol* 2000;27(1):51–7.
- [17] Flight WG, Smith A, Paisey C, et al. Rapid detection of emerging pathogens and loss of microbial diversity associated with severe lung disease in cystic fibrosis. *J Clin Microbiol* 2015;53(7):2022–9.
- [18] Eisenhofer R, Minich JJ, Marotz C, et al. Contamination in low microbial biomass microbiome studies: issues and recommendations. *Trends Microbiol* 2019;27(2):105–17.