



Article

Bile Acid Signal Molecules Associate Temporally with Respiratory Inflammation and Microbiome Signatures in Clinically Stable Cystic Fibrosis Patients

Stephanie Flynn ^{1,†}, F. Jerry Reen ^{1,†,‡}, Jose A. Caparrós-Martín ^{2,3,†} , David F. Woods ¹ , Jörg Peplies ⁴, Sarath C. Ranganathan ^{5,6,7}, Stephen M. Stick ^{2,8,9}, Fergal O’Gara ^{1,2,3,*} and On behalf of AREST CF [§]

¹ BIOMERIT Research Centre, School of Microbiology, University College Cork, T12 K8AF Cork, Ireland; stephflynn282@gmail.com (S.F.); j.reen@ucc.ie (F.J.R.); david.woods@ucc.ie (D.F.W.)

² Wal-yan Respiratory Research Centre. Telethon Kids Institute, 6009 Perth, Western Australia, Australia; jose.caparros-martin@curtin.edu.au (J.A.C.-M.); Stephen.Stick@telethonkids.org.au (S.M.S.)

³ Curtin Health Innovation Research Institute (CHIRI), Curtin University, 6845 Perth, Western Australia, Australia

⁴ Ribocon GmbH, Fahrenheitstraße. 1, 28359 Bremen, Germany; jpeplies@ribocon.com

⁵ Department of Respiratory Medicine, The Royal Children’s Hospital, 3052 Melbourne, Australia; sarath.ranganathan@rch.org.au

⁶ Infection and Immunity, Murdoch Children’s Research Institute, 3052 Melbourne, Australia

⁷ Department of Paediatrics, University of Melbourne, 3010 Melbourne, Australia

⁸ Telethon Kids Institute, The University of Western Australia, 6009 Perth, Western Australia, Australia

⁹ Department of Respiratory Medicine and Sleep Medicine, Perth Children’s Hospital, 6009 Perth, Western Australia, Australia

* Correspondence: f.ogara@ucc.ie

† These authors contributed equally to the manuscript.

‡ Present address: School of Microbiology, University College Cork, T12 K8AF Cork, Ireland.

§ Membership of AREST CF is provided in the Acknowledgments.

Received: 2 October 2020; Accepted: 1 November 2020; Published: 6 November 2020



Abstract: Cystic fibrosis (CF) is a congenital disorder resulting in a multisystemic impairment in ion homeostasis. The subsequent alteration of electrochemical gradients severely compromises the function of the airway epithelia. These functional changes are accompanied by recurrent cycles of inflammation–infection that progressively lead to pulmonary insufficiency. Recent developments have pointed to the existence of a gut–lung axis connection, which may modulate the progression of lung disease. Molecular signals governing the interplay between these two organs are therefore candidate molecules requiring further clinical evaluation as potential biomarkers. We demonstrate a temporal association between bile acid (BA) metabolites and inflammatory markers in bronchoalveolar lavage fluid (BALF) from clinically stable children with CF. By modelling the BALF-associated microbial communities, we demonstrate that profiles enriched in operational taxonomic units assigned to supraglottic taxa and opportunistic pathogens are closely associated with inflammatory biomarkers. Applying regression analyses, we also confirmed a linear link between BA concentration and pathogen abundance in BALF. Analysis of the time series data suggests that the continuous detection of BAs in BALF is linked to differential ecological succession trajectories of the lung microbiota. Our data provide further evidence supporting a role for BAs in the early pathogenesis and progression of CF lung disease.

Keywords: cystic fibrosis; bile acids; lung microbiota; inflammation; gut–lung axis

1. Introduction

Cystic fibrosis (CF, OMIM 219700) is a monogenetic autosomal recessive condition caused by biallelic pathogenic variants in the *Cystic Fibrosis Conductance Transmembrane Regulator* (CFTR) gene [1]. The CF disease phenotype arises from alterations in ion secretion pathways as a result of impaired chloride homeostasis [2]. These phenotypic manifestations involve many organ systems and progressive lung damage is the main driver of mortality and morbidity [2,3].

Cystic fibrosis lung disease starts early in life with an evident deterioration of the lung parenchyma even before the first respiratory manifestations are noticeable [4]. The most recent evidence suggests that this initial damage of the lungs is driven by neutrophilic inflammation [5]. Secondary to this exacerbated inflammatory response, colonisation of the lower airways by opportunistic pathogens initiates the self-reinforcing cycles of inflammation–infection that mediate the progressive functional deterioration of the CF lung epithelia [6,7]. Based on these evidence sets, early childhood interventions should therefore be the focus of attention. Key processes governing the deregulated inflammatory responses, and the microbial succession events favouring the establishment of opportunistic bacteria need to be better understood. However, the molecular mechanisms involved in the initiation of these pathological processes, as well as their implications in disease progression remain largely unknown.

Given the well-known immunomodulatory role played by the gut microbiota, a concept has been proposed suggesting that the microbial communities from the gastrointestinal tract could influence the progression of chronic pulmonary disorders [8,9]. Thus, the activity of this gut–lung communication axis may help to explain some of the clinical heterogeneity observed amongst CF patients. These observations also raise the possibility of complementary therapies targeting the gut to regulate the inflammatory component characteristic of chronic pulmonary conditions [9], i.e., high-fibre meals or probiotic supplements to modulate the gut microbiota [10].

The microorganisms that appear to be associated with an inflammatory response early in life are typically described as “oral flora” [6], raising the prospect of seeding the lower airway by microaspiration [11]. This mechanism may help to deliver gut-generated effector molecules and microorganisms into the lungs. Reflux is common in infants, therefore microaspiration of gastric refluxate is likely to occur. Bile acids (BAs) have emerged as potential modulators of the gut–lung axis, capable of inducing both an immune response in the host cells and a behavioural response in pathogens [12,13]. We previously reported that the presence of BAs in bronchoalveolar lavage fluid (BALF) associated with inflammatory markers, and we provided evidence linking these metabolites with early disease progression trajectories [14].

In this work we carried out a longitudinal analysis to explore whether the detection of BAs was temporally linked to an increase in the production of inflammatory markers, and to alterations in the compositional profiles of the lung microbiota. Using BALF as surrogate of the lower airway environment, we demonstrated a temporal association between BAs and inflammatory markers. Applying regression models, we also showed that (i) BA concentration is linked to the abundance of CF respiratory pathogens, and that, (ii) continuous BA positivity could alter the ecological succession trajectories of the CF lung microbiota.

2. Materials and Methods

2.1. BALF Cohort

Samples in this study were collected from infants and pre-school aged CF patients over the course of seven years (age range from 1–7 years). Collection of the bronchoalveolar lavage fluid (BALF) specimens, cytologic profile, cytokine quantification and clinical microbiology, were performed accordingly to the AREST-CF standard operating procedures [14].

2.2. Bile Acid Profiling and Bacterial DNA Extraction

A BALF supernatant aliquot from the two middle washes of the right middle lobe was centrifuged at 13,000 rpm for 10 min. Supernatants were profiled for 12 different bile acids as we previously described [14]. Microbial DNA was then extracted from the bacterial pellets using the Genra PureGene DNA Extraction kit (QIAGEN) as per the manufacturer's recommendations. To have an unbiased representation of the microbial DNA signatures in BALF, a minor modification was included, consisting of an overnight digestion with proteinase K.

2.3. Profiling and Analysis of the BALF-Associated Microbial Communities

Microbial profiling using 16S-based amplicon sequencing was performed at LGC (Germany). A pre-amplification step was performed for 20 cycles using the primers GM3 and 1061R. One μL of this PCR was used as template in a second PCR reaction targeting the V3-V4 region of the 16S rRNA gene using the forward primer 341F 5'-NNNNNNNNNTCCTACGGGNGGCWGCAG and reverse primer 785R 5'-NNNNNNNNNTGACTACHVGGGTATCTAAKCC as previously described [15]. The experimental conditions for this second PCR, library preparation and sequencing steps using Illumina V3 Chemistry are extensively explained in our previous publication [15].

Marker-gene sequencing data were processed using the SILVAngs analytical pipeline [16,17] as we previously described [15]. Operational taxonomic units (OTUs) representing low-count taxa (abundance lower than 0.01% across all samples) were removed as previously described [18]. Taxonomic profiles were then pre-filtered to remove technical-associated contaminant counts obtained after sequencing negative controls using the functions of the R package decontam [19]. To account for differences in library size, the pre-filtered taxonomy table was then re-scaled using the cumulative sum scaling normalisation method implemented in the R package metagenomeSeq [20].

2.4. Methodological Strategy to Minimize the Effect of Environmental Contaminants in the BALF-Associated Microbial Profiles

Sequence-based studies using low-biomass samples such as those involved in this work are extremely sensitive to potential bias introduced by laboratory contaminants [21]. In practice, this issue has sparked controversy with regards to whether the detection of microbial communities in specified body niches can be robustly interpreted biologically in the absence of rigorous technical and analytical controls [22,23]. To evaluate the possible impact of potential artificial communities in our study, BALF specimens were randomly assigned to 5 different processing batches. A DNA extraction blank control was processed alongside the samples within each batch. Libraries for these batch extraction controls were generated in parallel with the patients' samples and sequenced. This approach allowed us to identify and deconfound the environmental noise present within each sample batch (Tables S1 and S2). Following recommended publication standards, taxa detected in negative controls are provided in Table S1 [24]. To filter out microbial communities heavily influenced by the presence of putative contaminants within each batch, the following sample exclusion criteria was used. This strategy was in agreement with internationally accepted standards to mitigate the effect of contaminants in low microbial biomass studies [21,24]. Firstly, OTUs tables were pre-filtered to remove low-count taxa representing potential PCR artefacts as previously described [18]. Pre-filtered OTU count tables were then converted to proportion tables, and correlated to the composition of the corresponding batch-associated extraction blank control (Table S3). Samples with a correlation coefficient equal to or higher than 0.7, which suggests a strong influence of environmental noise on sample microbial profiles, were excluded from further analysis. This was done independently of sequencing depth, as it is not a surrogate for total bacterial biomass. According to this criterion, we filtered out 14 samples, which were strongly related to the corresponding extraction control. Secondly, we performed a batch-wise contaminant OTU removal from the biological samples. Restrictive approaches recommend eliminating all the OTUs observed in negative controls from the actual samples. We considered that the application of this constrained filtering step would result in the removal of important biological

information, since OTUs typically observed in negative controls could be truly representative of BALF microbial communities [24]. We therefore applied the prevalence method implemented in the R package *decontam* to detect, and subtract, putative contaminant taxa present in the negative controls (Table S4) [19]. Accordingly with this approach, OTUs represented by the *Capnocytophaga*, *Ralstonia*, and *Saccharibacteria* taxa in batch 1; *Stenotrophomonas*, *Methylobacterium* and *Bacillus* taxa in batch 2; *Bacteroides*, *Inquilinus*, *Bacillus*, *Lactobacillus* and *Acinetobacter* taxa in batch 3; and *Saccharibacteria* taxon in batch 4, were fully removed from the associated biological specimens. Finally, we subtracted the counts per OTU in the batch negative extraction control from each of the biological specimens in the same batch. Samples resulting with fewer than 1000 reads were also removed from the final dataset. The existence of a good correlation between the reported clinical microbiology and the 16S-based microbial findings (Pearson's r 0.68, t -test 0.00012), provide support on the reliability of the reported BALF-associated microbial communities and the adequacy of the filtering steps described (Table S5). After performing these filtering steps, we normalized the OTU tables of the 59 remaining BALF specimens, applying the cumulative sum scaling (CSS) approach as implemented in the R package *metagenomeSeq* [20].

2.5. Statistical Analysis

Statistical analysis was conducted in R (version 4.0.2) [25]. Median-like quantile normalisation was performed with the functions contained in the R package *metagenomeSeq* as described above [20]. Model-based clustering was carried out with the *Mclust* function of the R package *mclust* [26]. The number of mixing components and the model covariance structure were selected using penalized-likelihood criteria [26]. For differential abundance testing we used the function *fitFeatureModel* as implemented in the R package *metagenomeSeq* [20]. Linear models were fitted with the built-in function *lm*. We used regression diagnostic plots to evaluate whether the assumptions of the linear regression were met, and to confirm the absence of high leverage points in the dataset. Spearman correlations and hierarchical clustering were calculated with the built-in functions *cor* and *hclust*, respectively. Functional capabilities were predicted using Tax4Fun [27]. KEGG Enrichment analysis was calculated using a global test algorithm as implemented in *MicrobiomeAnalyst* [28,29].

Time series data were modelled using linear mixed models with cubic spline regression following the *timeOmics* framework [30]. For this analysis, the pre-filtered OTU abundance table was transformed to centred log ratios following the package vignette. A double strategy was applied to evaluate the quality of clustering achieved using *timeOmics*; (i) to select the number of components and the optimum number of clusters within each principal component we used the average silhouette coefficient and, (ii) to evaluate the degree of association between features assigned to the same cluster, we calculated the proportionality distance [30].

The packages *ggplot2*, *factoextra*, *superheat*, *mixOmics* and *sjPlot* were used for data visualisation [31–35]. The probability value threshold for statistical significance was set at 0.05.

2.6. Ethics, Consent and Permissions

The ethical aspects of the AREST-CF program were reviewed and approved by the Princess Margaret Hospital for Children, Perth ethics committee (10 December 2009) (Ethical approval Ref. 1762/EPP). Written informed consent for publication was obtained from the parents/guardians. All methods performed in this study were carried out in accordance with the relevant guidelines and regulations.

2.7. Data Availability

The marker-gene sequencing data are accessible through the Sequencing Read Archive (BioProject PRJNA662556).

3. Results

3.1. Study Cohort

We evaluated 77 longitudinal BALF samples (average samples per patient, 3.67) from 21 CF patients (14 females, 7 males. Age range 1–7 years; median 3 years, interquartile range (IQR) 2–5 years) enrolled in the AREST CF Early Surveillance Program (Perth and Melbourne, Australia). Diagnosis was performed through new-born genetic screening in most of the cases (17 subjects). Six patients were asymptomatic at diagnosis, and in two cases information on symptoms was not recorded. Clinical presentation in the rest of the subjects was variable, but included typical CF-associated gastrointestinal manifestations and clinical features of failure to thrive. All the study subjects carried the common c.1521_1523delCTT (p.Phe508del) genetic variant in the homozygous (8 patients) or heterozygous (13 patients) state. For 10 of the heterozygous carriers a second disease-causing allele was identified. Sweat chloride testing confirmed CFTR dysfunction in the remaining three cases. Most of the BALF samples were collected annually during asymptomatic periods, in which computed tomography scans of the chest were also performed. Sample acquisition in four specimens was conducted after *Pseudomonas aeruginosa* eradication therapy, and one sample during an exacerbation episode. The later biological specimen represents the unique sample of our cohort collected at the age of seven. For 50 out of the 77 BALF samples, ongoing antibiotic regimen was reported at the time of the collection of the bronchial wash (of which seven specimens were associated with two or more type of antibiotics). Amoxicillin/clavulanate was the most prescribed therapeutic (38 samples), followed by azithromycin (9 samples), cephalexin (6 samples), tobramycin (3 samples), co-trimoxazole (2 samples) and ciprofloxacin (1 sample). For 10 samples, information on antibiotic therapy was not available. Additional medical procedures during the follow-up visit included assessment of inflammation and infection in the lungs. Detailed protocols for the collection of BALF, measurements of inflammatory markers, clinical microbiology, evaluation of respiratory function, and quantification of structural lung disease were as previously described [14,36–39]. Clinical metadata for this patient cohort are provided as Supplementary Materials.

3.2. Temporal Associations between Bile Acids and Inflammatory Markers

We and others have reported that BAs both stimulate the release of pro-inflammatory cytokines *in vitro*, and may induce airway inflammation and modulate disease progression trajectories in CF [13,14,40–43]. Confirming our previous observations [14], total BA concentration was significantly associated with several inflammatory markers in the cohort BALF samples: interleukin (IL)-1 β (Spearman's ρ 0.6, Spearman's test 0.001, false discovery rate correction (FDR) 0.01); IL-6 (Spearman's ρ 0.52, Spearman's test 0.005, FDR 0.01); IL-8 (Spearman's ρ 0.32, Spearman's test 0.003, FDR 0.01); and percentage of neutrophils (Spearman's ρ 0.26, Spearman's test 0.02, FDR 0.048) (Table S6). After controlling the false discovery rate, we did not observe significant associations between the indicated markers in BALF (including total BA concentration) and any antibiotic regimen reported at the time of collection of the bronchial washing (Table S7).

Interleukin 8 is a powerful neutrophil-attracting signal molecule that recruits neutrophils at foci of inflammation and infection [44]. Given the relevant role played by IL-8 in the migration and activation of neutrophils in CF, we evaluated the co-occurrence of neutrophils and IL-8 levels with BA concentration across the longitudinal BALF samples in our patient cohort (Figure 1). We observed a consistent strong significant relationship between neutrophil and BA load in the BALF specimens obtained from four-year-old (Spearman's ρ 0.77, Spearman's test <0.001, FDR 0.0008) and five-year-old patients (Spearman's ρ 0.71, Spearman's test 0.009, FDR 0.009). Co-occurrence between IL-8 and BA burden was found to be significant only in BALF samples obtained from the five-year-old CF children (Spearman's ρ 0.718, Spearman's test 0.005, FDR 0.009) (Figure 1). It is worth noting that, although the small sample size of this study did not allow us to reject the null hypothesis of no relationship, the effect size of the bivariate relationship between BA concentration and IL-8 levels in BALF at year 1 was also

notable (Spearman's ρ 0.499, Spearman's test 0.08, FDR 0.16) (Figure 1). Interestingly, structural lung disease as determined by quantifying tomographic images using the PRAGMA scoring method [36], was similar over the analysis period in our patient cohort (Figure 1).

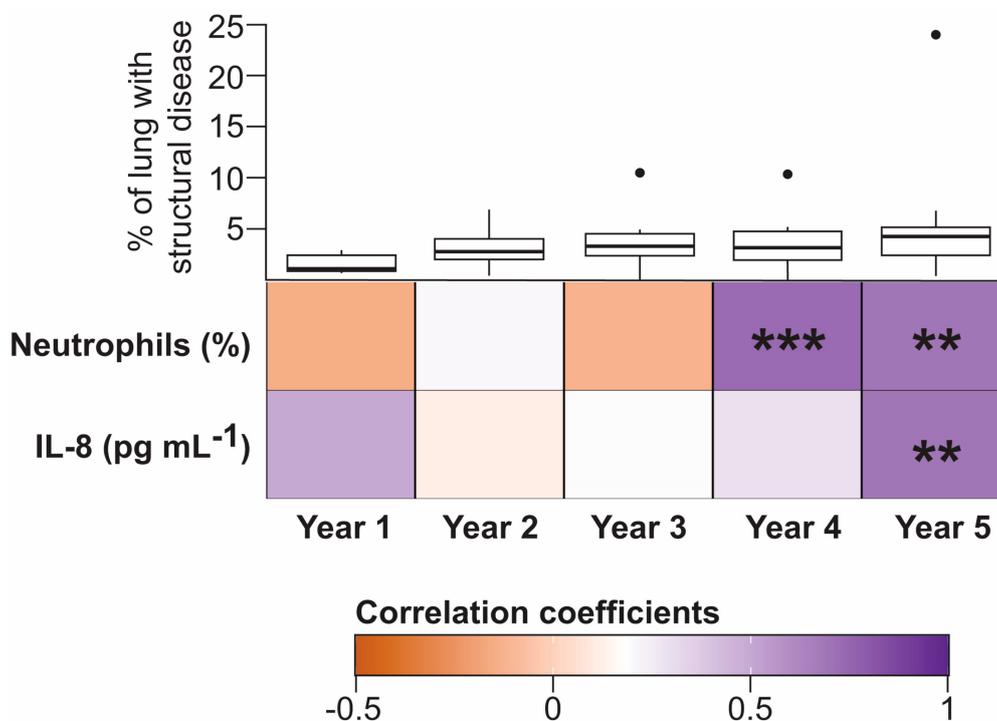


Figure 1. Temporal associations between bile acids (BAs) and inflammatory markers in bronchoalveolar lavage fluid (BALF). The heatmap represents the Spearman rank-order correlation coefficients between total BA concentration (log₁₀ transformed) and the indicated inflammatory biomarkers in BALF. Temporal co-occurrence is represented in samples obtained from one- to five-year-old cystic fibrosis (CF) subjects (Number of independent samples: year 1, $n = 13$; year 2, $n = 14$; year 3, $n = 13$; year 4, $n = 16$; year 5, $n = 13$). Boxplots in the top row represent the percentage of structural lung disease over the five-year period quantified using the PRAGMA protocol [36]. No significant differences in airway remodelling were observed in the context of Dunnett's test between year 1 and each of the following four years. Asterisks indicate statistical significance for the indicated correlation per year after false discovery rate (FDR) correction. **, $p < 0.01$; ***, $p < 0.001$.

Previous research has demonstrated that the detection of inflammatory markers is robustly associated with the gradual acquisition of pathogen-dominated microbial communities in CF [6,45,46]. This observation suggests that the progressive establishment of bacterial pathogens in the lung microbiota is a major trigger of chronic inflammation in CF. In order to address whether the presence of BAs was also associated with alterations in the BALF-associated microbiota, we carried out a 16S amplicon sequencing approach from the bacterial pellets obtained after high speed centrifugation of the same BALF cohort aliquots used for profiling BAs.

3.3. Characterisation of the 16S-Based BALF Microbial Structures and Their Association with Disease Outcomes

In agreement with previous publications, sequencing read data from BALF pellets were assigned to taxonomic identities representing anaerobes typically observed in the oropharyngeal cavity, and classical CF pathogens (Figure S1) [6,46–48]. Visual inspection of the BALF-associated microbial profiles showed that qualitatively, most of the patients demonstrated a relatively stable microbial profile over the time. Interestingly, for some patients an age-associated crash in the structure of the community was noticeable. The drastic remodelling of the microbial profile in these patients was not reversed during

the progression of the study, and it was consistently linked to the acquisition of dominant pathogens. This was, for example, the case in four of the CF children where the BALF-associated microbial assemblages became dominated by OTUs represented by *Pseudomonas*, *Stenotrophomonas*, *Haemophilus* or *Achromobacter* taxa (Figure S1).

Recent evidence suggests that translocation of oral microorganisms into the lower airways through microaspiration triggers pulmonary inflammation [49,50]. These oral-enriched microbial assemblages may also favour pathogen colonisation in CF through a cross-feeding interaction associated with the anaerobic fermentation of mucins [6]. We therefore evaluated whether (i) the markers of disease progression established in our cohort were linked to the presence of BALF-associated microbial ecotypes, and (ii) whether detection of BAs could be associated with the existence of specific 16S-based bacterial clusters. For this purpose, we performed a model-based clustering approach by fitting a Gaussian Mixture Model (GMM) to the CSS-transformed taxonomic profiles. This approach allowed us to classify the BALF-associated communities on the basis of their bacterial profiles in an unbiased manner. A diagonal covariance structure with equally shaped and varying volume ellipses aligned to the axes (VEI model) [26], associated with three components of mixtures, was identified as the most parsimonious model based on the Bayesian Information Criterion (Figure S2). Three microbial ecotypes with different functional capabilities were therefore inferred from the dataset representing the study subjects (Figure 2a, Figure S3 and Table S8). Examination of the associated clinical metadata confirmed that these microbial-based clusters have medical relevance (Figure 2b–d). Thus, microbial clusters 2 and 3 were characterized by a progressive increase in neutrophil counts and neutrophil elastase (NE) levels in BALF (Figure 2b,c). Interestingly, both microbial ecotypes were associated with a noticeable increase in lung disease, and cluster 3 showed a positive trend towards higher levels of BAs in BALF (Figure 2d,e). Notably, these phenotypic differences between clusters were not associated with age, cell burden or the presence of IL-8 levels in BALF (Figure S4), and the odds of being on antibiotics were not significantly different across the different microbial clusters (Figure S4E). Given that clustering of the BALF-associated microbiota was reflecting contrasting clinical outcomes, we proceeded to characterize the differences in composition between microbial clusters. Differential abundance testing was performed using a zero-inflated log-normal mixture model to identify OTUs that differed between the three components of mixtures [20]. This analysis showed that, compared to cluster 1, clusters 2 and 3 demonstrated a progressive increase in the abundance of OTUs assigned to oral taxa and known CF pathogens (Figure 2f,g). This change in community composition was accompanied by a gradual increase in inflammatory markers (Figure 2b,c). Interestingly, compared to cluster 2, cluster 3 was characterized by a higher abundance of pathogens, which may help to explain the high NE levels observed in this microbiota-based group (Figure 2c and Figure S5).

The results derived from the model-based analyses may suggest that BA detection in BALF could be associated with specific bacterial groups rather than with single microbial ecotypes. In order to address this possibility, we collapsed the OTU profiles into three categories: pathogens (containing conventional and emergent CF pathogens) [51–53], oral (representing oral-associated taxa) [54–56] and others (the rest of the members of each community) (Figure S6). In our cohort, regression analyses demonstrated a positive association between pathogens and oral taxa, with NE levels in BALF (Figure S6A,D). The data show that both microbial groups had a moderate-to-large effect on this inflammatory marker (Pathogens: Spearman's ρ 0.5, Spearman's test <0.0001 , FDR 0.00019. Oral: Spearman's ρ 0.35, Spearman's test 0.005, FDR 0.028) (Table 1 and Table S9). Similarly, but not surprisingly, the percentage of structural lung disease was a significant predictor of pathogen abundance in BALF (Figure S6C). Interestingly, pathogen count was associated with the concentration of BAs. Accordingly to the regression coefficients, concentrations of BAs associate with pathogen abundance by increasing their population by 8.8 (standard error SE 3.64, adjusted R^2 0.07, F -statistic (1,57) = 5.848, p -value 0.018) (Figure S6D). No interaction effect of age on any of the predictors was detected (Figure S6F–I).

Altogether, these associations suggest a potential role for (i) oral taxa in the development of early airway inflammation in CF and, (ii) BAs in the selection of opportunistic pathogens in CF lungs.

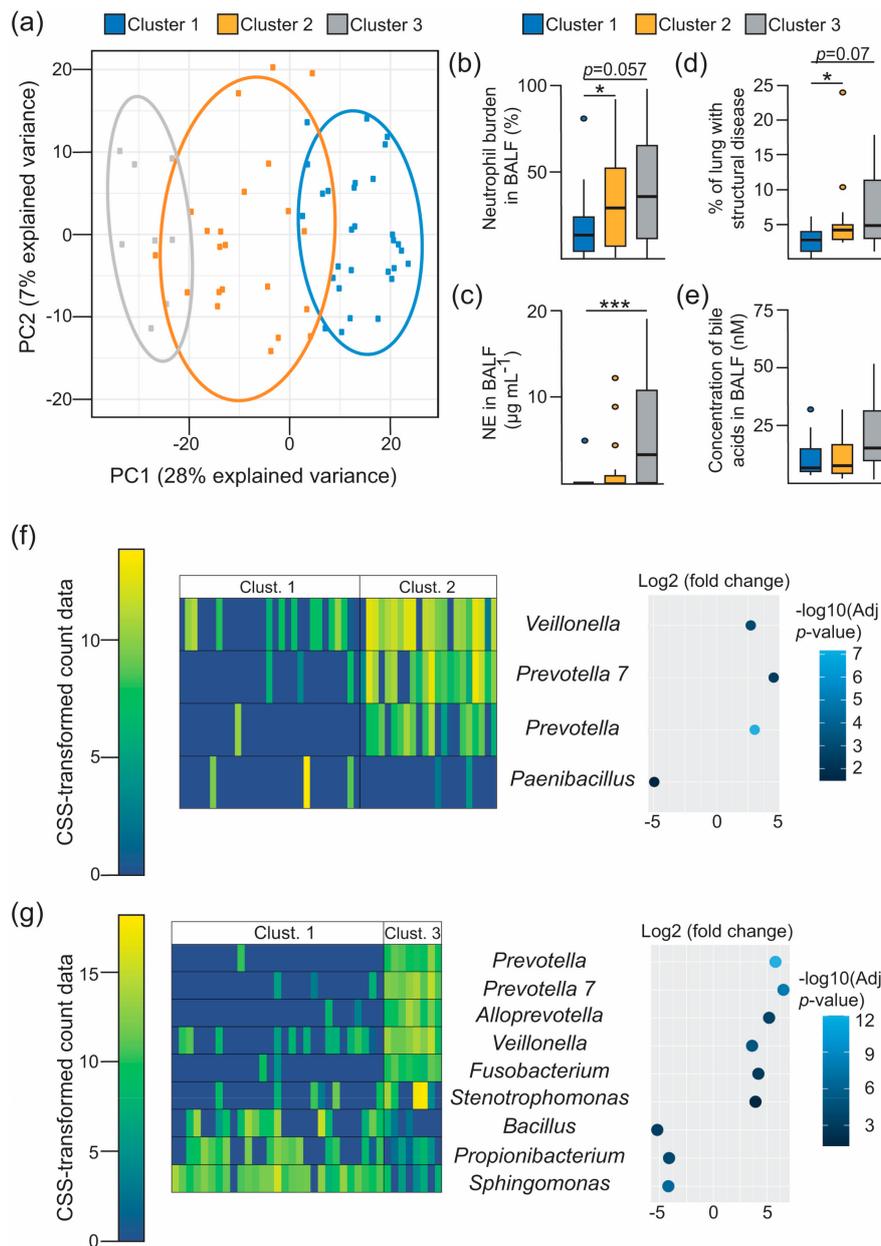


Figure 2. Inflammatory surrogates in BALF are linked to specific microbial assemblages. (a) Principal component analysis decomposition of the OTU counts transformed using the cumulative sum scaling (CSS) normalization method. Mean-centering was performed to better represent the direction of variability across group centroids. BALF-associated taxonomic profiles are projected onto the first two components of the model. Dots represent each BALF specimens, which are coloured accordingly with the Gaussian Mixture Model-based cluster membership (cluster 1, $n = 29$ BALF specimens; cluster 2, $n = 22$; cluster 3, $n = 8$). Ellipses represent the 86% confidence region. (b–e) Boxplots showing between-cluster comparisons of the indicated markers of disease progression (b–d) and bile acid concentration (e) in BALF. Multiple comparisons were assessed for significance in the context of Dunnett’s test. *, $p < 0.05$; ***, $p < 0.001$. (f,g) Heatmaps (left) and bivariate dot plot (right) representing the CSS-normalised count data and the fold change (\log_2 scale) for the selected OTUs between the indicated clusters (Clust.), respectively. Differentially abundant features between clusters were estimated from a zero-inflated log-normal model as implemented in the *fitFeatureModel* method [20]. Features with a \log_2 (fold change) $> |2|$ and an FDR-corrected (adjusted p -value) p -value < 0.05 are depicted. Colour legend for the heatmap (left), and the dot plot (right), represent the CSS-transformed count data and the $-\log_{10}$ (adjusted p -value) respectively. Each column in the heatmaps represents the profiles of individual BALF samples, and rows represent the indicated OTUs.

Table 1. The table represents the correlation between CSS-transformed pathogen counts, and the detection of the indicated cytology and inflammatory markers and bile acid levels in BALF. We used Spearman's ρ to test for associations between variables. Uncertainty level is reported as p -values calculated using Spearman's test. To control the false discovery rate (FDR), p -values were adjusted by applying Benjamini & Hochberg correction method (adjusted p -value). Significant associations are highlighted in bold. Dis (%) represents the percentage of structural lung disease quantified using the PRAGMA scoring method [36].

Variable	Spearman's ρ	p -value	Adjusted p -Value
Log10(Bile acid concentration) (μM)	0.339976	0.008425	0.02074
Interleukin 8 (pg mL^{-1})	0.216353	0.099796	0.12474
Neutrophil Elastase (ng mL^{-1})	0.508734	3.886×10^{-5}	0.00019
Neutrophils burden (%)	0.170807	0.199858	0.19985
Dis (%)	0.38695	0.012444	0.02074

3.4. Temporal Dynamics of the Lung Microbiota Associated with the Detection of Bile Acids in BALF

Next, we determined whether BAs were linked to ecological succession in microbial trajectories in the CF lung. To address this question, we firstly separated the patients into groups based on the total concentration of BAs in BALF. For classification, we applied a hierarchical clustering approach to the first three available BALF specimens for each child (Figure S7). This strategy allowed us to group the patients into two categories: (1) subjects with a continuous detection of high concentrations of BAs over a three-year period ($n = 6$, hereafter high BA (HBA) group), and (2) individuals with no or intermittent detection of high levels of BAs across the longitudinal samples ($n = 10$, hereafter low BA (LBA) group) (Figure S7A). No relation between bile acid-based clustering and antibiotic intake was observed (Chi-square test, χ^2 (degrees of freedom 1, sample size 41) = 0.49, p -value 0.48).

To evaluate whether the temporal evolution of the BALF-associated microbiota was affected by the presence of BAs, we made use of the recently published timeOmics framework [30]. This method implements modelling of temporal OTUs profiles using linear mixed models with spline regression, in combination with multivariate analysis of clustered profiles [30]. Modelling of the temporal trajectories of the 98 OTUs across the time in each of the HBA and LBA patient subsets was performed using linear mixed regression with cubic splines (Figure S8A,B). To avoid the inclusion of linear models representing noise, the quality of the profiles was subsequently assessed using two methods: a Breusch–Pagan test to check that the residuals were equally distributed; and a threshold based on the predictive power of the linear models in comparison to the more complex fitted models of the resulting profiles [30]. A total of 90 and 89 profiles were retained from the HBA and the LBA patient subsets, respectively. To select the more informative profiles from each group of patients, we analysed the predicted OTU trajectories using a sparse principal component analysis model (sPCA). The average silhouette coefficient was used to select the optimal number of components in the sPCA model and the number of features to keep for each component (Figure S8C,D). Based on these criteria, two clusters of profiles were selected from the HBA patient subset. The first cluster (sPCA-1 negative) showed a decrease in the abundance of two OTUs assigned to two pathobiont taxa (*Rothia* and *Streptococcus*) (Figure 3a and Figure S8E). Conversely, the second cluster (sPCA-1 positive) demonstrated a linear increase in the relative abundance of both pathobionts and opportunistic pathogens (*Stenotrophomonas*, *Ralstonia* and *Moraxella*) (Figure 3a and Figure S8E). In the case of the LBA subset, only one cluster of features with similar expression profiles was selected (Figure 3c and Figure S8F). This cluster showed a non-linear increase in oral pathobionts (*Tannerella* and *Prevotella 2*) over time in the LBA subset. The proportionality distance within all of these clusters was low, suggesting a strong association between the OTU profiles belonging to the same cluster (Figure 3b,d). Based on these results it is tempting to speculate that a continuous presence of BAs in the CF airways could influence lung colonisation patterns, by accelerating the acquisition of opportunistic pathogens.

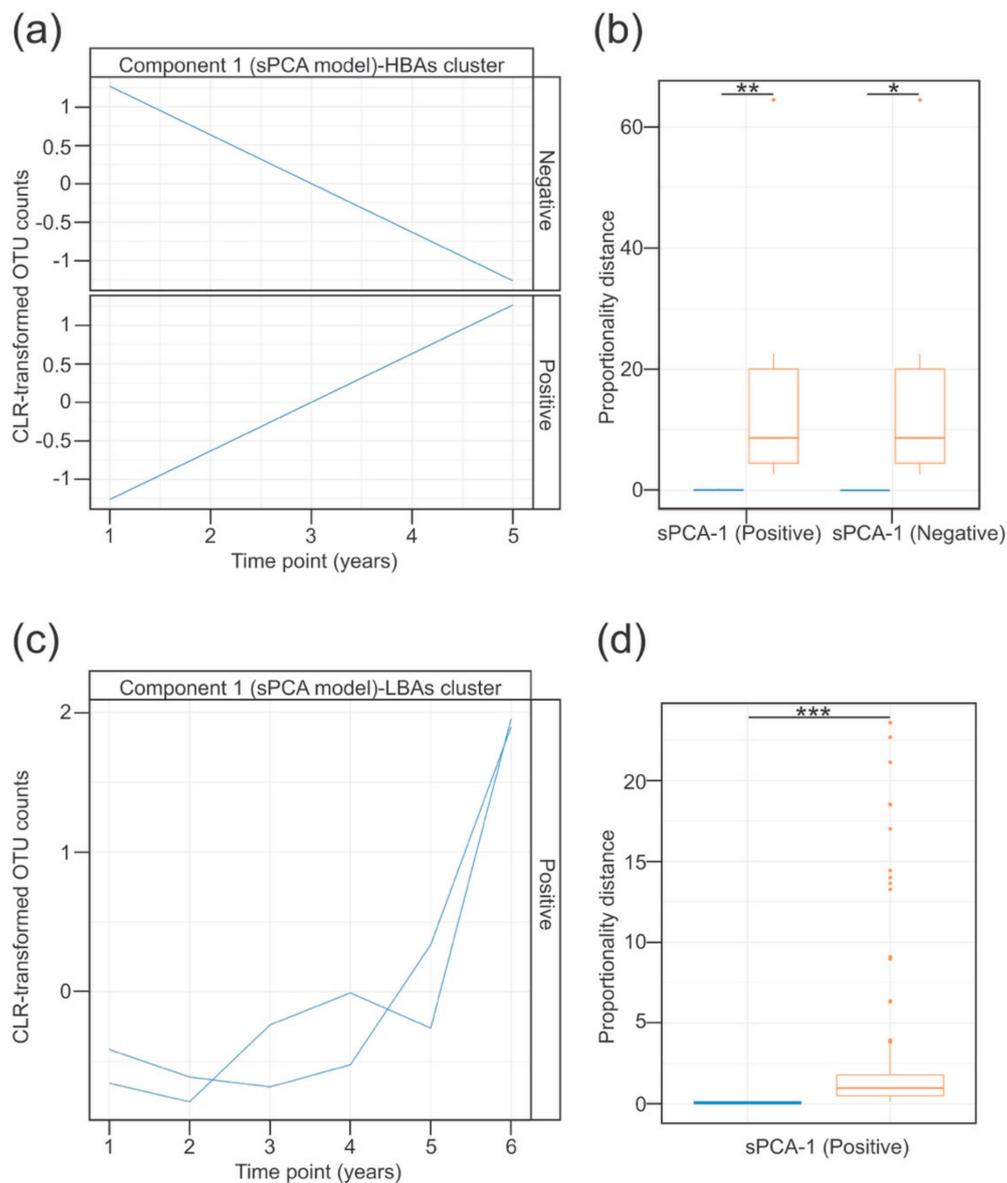


Figure 3. Longitudinal modelling of the BALF-associated microbial profiles in patients clustered based on BA detection patterns. (a–d) Temporal profiles of the OTUs grouped using sPCA in the HBA (A) or LBA (C) patient clusters. Blue traces represent the modelled abundance (Centred Log Ratio (CLR)-transformed) of the OTUs selected in the sPCA model across time. Within each component, the selected OTU profiles were subclustered into positive or negative according to the direction of the correlation between each OTU profile, and the indicated latent variable of the sPCA model. (b,d) Association between the features assigned to the same cluster was evaluated using proportionality distances. Bar plot in (b) (for the HBA subset) shows the proportionality distance of features from the same cluster (blue), and the distance of OTUs inside that cluster with every feature in the other cluster (orange). Given that only one cluster was selected in the LBA group, the bar plot in (d) represents the proportionality distance between pairs of features from the cluster (blue), and between each feature in that cluster and the rest of features in the entire dataset (orange). In both cases, the intra-cluster distance was significantly lower than the distance outside the cluster, suggesting a strong association between the selected OTU profiles clustered together. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$ in the context of the Wilcoxon test. Type I error rate was controlled with the false discovery rate method.

4. Discussion

Evidence suggests that both inflammation and infection can occur in CF patients even in the absence of obvious symptomatology [5,37]. The airway microenvironment is therefore likely to play a significant role in permitting pathogens to colonise the lower airways, while at the same time facilitating immune cell recruitment and inflammation. The gut–lung axis is an emerging concept that is capturing broad interest in respiratory disease because of its potential involvement in the progression of chronic lung disease [8,9]. This communication axis has been highlighted based on the role played by the gut microbiota and its associated metabolites in the maintenance of homeostasis in the host immune system [9]. In this regard, the exacerbated and uncontrolled immune responses typically observed in patients with chronic lung disorders could be mediated in part by the disruption of the host-microbiota crosstalk in the gut. Given their nutritional requirements, CF patients are recommended a high-calorie, high-fat diet. This dietary intervention affects the production of immune-regulatory molecules by the gut microbiota, such as short chain fatty acids and secondary BAs [57]. Studies using animal models have shown that the gut microbiota modulates BA synthesis [58,59]. This is also the case with lipid-enriched diets, which change the composition and function of the intestinal flora. This can happen directly, and additionally by dysregulating BA metabolism [60]. We and others have demonstrated the presence of BAs in the lower airways of CF patients [14,41,43,61,62]. Specifically, we have demonstrated that detection of BAs is associated with inflammatory markers and less desirable disease trajectories [14]. Current evidence suggests that BAs are likely to reach the lower airways through duodenogastroesophageal reflux and microaspiration [14,43,61]. The outcome of this action could constitute a communication channel between the gut and the lungs. In this work, we report a temporal association of BAs with inflammatory markers. These associations link the presence of BAs in BALF to the expression of IL-8 and the recruitment of neutrophils into the airways of clinically stable children with CF. Several lines of evidence as outlined below suggest that these associations are specific, rather than secondary to the aspiration of gastric content. Firstly, independent laboratories have demonstrated BA-mediated expression of inflammatory markers in vitro in respiratory epithelial cells [13,40]. In our laboratory, we have also shown that this response requires the BA receptor FXR [13]. Secondly, our longitudinal profiles revealed a serial correlation between BAs and the neutrophil load in BALF. Finally, monitoring relationships between airway inflammation and markers of gastric aspiration such as pepsin in CF children has yielded contrasting results even within the same research programme [63,64]. At this stage we are not focusing on the underlying molecular mechanism(s) through which BAs stimulate inflammation in CF lungs. One possibility is via epithelial signalling involving a self-amplifying loop between a combination of neutrophils, dendritic cells, or other lymphoid lineage cells. Another interesting possibility could involve an imbalanced T_H17/T_{reg} cell differentiation process from naive $CD4^+$ T cells in the presence of specific BA signatures in the lungs [65–67]. This BA-mediated process could provide a hypothetical mechanism for the proposed role of T_H17 cells in initiating the inflammatory response in the early stages of CF [68]. Importantly, the mere temporary presence of immune regulatory signals such as BAs in the airways, could be sufficient *per se* to interfere with defence responses and subsequently promote inflammation. Our future research will address these possible mechanisms.

Model-based clustering allowed us to identify structurally similar microbial groups. These microbial ecotypes showed contrasting expression patterns of disease progression markers and were not associated with antibiotic regimen. In general, we found that the more pro-inflammatory microbial communities were composed of taxa from both the oral cavity and opportunistic pathogens. This observation is in line with previous reports in healthy people, linking the presence of oral flora in the airways, with an inflammatory phenotype [49,50]. Despite the data presented in the former reports, controversy exists as to whether bacteria associated with the supraglottic region play any role in early CF disease [6,69]. The recent work by Jorth and colleagues argues against this possibility [69]. A rigorously well-controlled marker-gene sequencing experiment supported their conclusions. Nevertheless, the outcomes drawn from the study could also be confounded by the

age-range of the patient cohort under investigation (mean, 13.1 years old; age-range, 6–21 years) [69], as the airways of older CF children/teenagers are usually dominated by pathogens [70]. Moreover, the patients enrolled in that specific study underwent bronchoscopy for clinical indications, which may also have involved underlying infectious processes [69]. In a separate investigation, Muhlebach and colleagues reported that, in a clinically stable cohort of CF children, the lungs are initially colonised by microorganisms associated with supraglottal niches. These microbes are successively replaced over time by pathogen-enriched microbial assemblages [6]. The authors also proposed that this shift in community composition could be driven by a cross-feeding interaction. This would involve oral flora releasing nutrients from mucins in the mucous layer that could eventually facilitate pathogen colonisation [71]. Our results fit with this hypothesis and point to a pro-inflammatory role for the oral bacteria. Furthermore, we suggest that nutritional interaction(s) between microbes may dictate colonisation patterns in CF lungs [71]. If this concept is well founded, it would significantly question whether “mixed oral flora” communities are in fact harmless. Rather, this oral flora could actually be responsible for the “sterile inflammation” observed in the absence of pathogen detection in CF infants [5,37].

The observed linear association between BA load and pathogen abundance is intriguing. Due to their antimicrobial character, BAs can shape microbial communities [72]. Thus, environmental selection would allow the stable settlement of bile-tolerant microorganisms. Interestingly, we have reported that several classical CF pathogens tolerate BAs, adopting chronic lifestyles in the presence of these molecules [12,13,73]. This adaptive process could lead to an increase in their ecological resilience behaviour, and ultimately facilitate the establishment of persistent chronic infections in the CF lungs. Due to their clinical relevance, our findings justify additional targeted investigation in further studies.

Taking advantage of the longitudinal nature of our study, we also performed a temporal analysis on the BALF-associated microbial profiles in patients with contrasting patterns of BA detection. This analysis allowed us to capture strongly associated taxa correlating in time within the different BA-based groups. Thus, patients with a continuous record of BA detection exhibited contrasting co-occurrent microbial profiles compared to patients not sharing this pattern. These differences were specifically related to the reduction in oral bacteria (*Rothia*, *Streptococcus*), and the acquisition of emergent opportunistic pathogens (*Stenotrophomonas*, *Ralstonia*, *Moraxella*), which were only observed in the HBA group. This analysis suggests that the continuous presence of BAs in the airways could play a relevant role in the microbial colonisation process of the CF lungs, by mediating changes in the pattern of ecological interaction. This could again be related to the recognised antimicrobial character of BA molecules [72]. The long time interval between the collection of the samples (one year), the limited time points included in the analysis (three per patient), and the reduced number of patients evaluated (13 patients), does not justify drawing any inferences from these results to a broader population. However, our data further highlight the potential role of BAs in CF pathogenesis, suggesting that these molecules could mediate the rearrangement of the taxon associations in the CF lung microbiota.

Our study provides observational data related to the effect of airway exposure to BAs early in life in CF patients. As such, this research cannot exclude the possibility of confounding variables affecting the outcome of our analysis. For example, the use of proton pump inhibitors would prevent the microaspiration of acidic gastric contents, which can independently induce damage in the airways. Furthermore, the absence of non-CF individuals does not allow us to evaluate the translatability of our results to a broader population. Therefore, we cannot address whether the presence of BAs in the airways of non-CF subjects would also associate with inflammatory responses, as in the case of microaspiration of oral flora [49,50]. These limitations do not permit us to draw definitive conclusions. Nevertheless, they do not invalidate the hypothesis emerging from our study, which suggests an important role for BAs in early CF lung disease. Due to the clinical benefits of early childhood interventions in CF, the data presented certainly provide an important stimulus to further explore this hypothesis with future intervention studies.

Supplementary Materials: Supplementary materials can be found at <http://www.mdpi.com/2076-2607/8/11/1741/s1>. Supplementary Files: Clinical metadata. Figure S1: Relative composition of the BALF-associated microbial communities collapsed at genus level. Figure S2: Bayesian Information Criterion for the model-based clustering. Figure S3: Uncertainty plot showing the probability of an observation being from a specific cluster. Figure S4: Boxplot representing comparisons between the GMM-based clusters and clinical descriptors in BALF. Figure S5: Differentially regulated features between cluster 2 and cluster 3. Figure S6: Associations between the BALF-associated microbiota and clinical outcomes. Figure S7: Classification of BALF samples based on the longitudinal bile acid profiles. Figure S8: Time modelling of the OTU profiles in the HBAs and LBAs patient clusters. Table S1: Raw counts of the taxa detected in the negative extraction controls. Table S2: Frequency in biological specimens of taxa detected in negative controls. Table S3: Pearson correlation coefficients between the microbial profiles of the BALF specimens with their corresponding batch negative extraction control. Table S4: Putative contaminant taxa associated with each extraction batch. Table S5: Correspondence between the clinical microbiology results and the 16S amplicon sequencing profiles. Table S6: Correlation between the total concentrations of bile acids and the detection of cytology and inflammatory markers in BALF. Table S7: Associations between markers in BALF and antibiotic intake at the time of the collection of the BALF specimen. Table S8: KEGG Enrichment analysis. Table S9: Correlation between CSS-transformed oral counts, and the detection of cytology and inflammatory markers and bile acid levels in BALF.

Author Contributions: Conceptualisation, F.J.R. and F.O.; Methodology, F.J.R., S.F. and D.F.W.; Formal Analysis, J.A.C.-M., J.P., F.J.R., S.F. and D.F.W.; Investigation, F.J.R., S.F. and D.F.W.; Resources, S.C.R. and S.M.S.; Writing—Original Draft Preparation, J.A.C.-M. and F.O.; Writing—Review and Editing, J.A.C.-M., F.J.R., S.F., D.F.W., J.P., S.C.R., S.M.S., and F.O.; Visualisation, J.A.C.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by grants awarded by the European Commission (EU-634486-INMARE), The Irish Research Council for Science, Engineering and Technology (GOIPG/2014/647–IRCSET), Science Foundation Ireland (SSPC3-Pharm5: & 15/TIDA/2977), Health Research Board/Irish Thoracic Society (MRCG-2014-6 and MRCG-2018-16) and HRB-ILP-POR-2019-004, The Marine Institute (Beaufort award C2CRA 2007/082), The Glenn Brown Memorial Grant 2017 (The Institute for Respiratory Health, Perth, Australia), U.S. CF Foundation (Grant CFF 1710) and Australian NHMRC 2020-(Synergy APP1183640). S.M.S. is a National Health and Medical Research Council Practitioner Fellow.

Acknowledgments: We acknowledge the patients and their families for their participation in this study. The full list of the AREST members is available at <https://arestcf.telethonkids.org.au/our-people/>.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Riordan, J.R.; Rommens, J.M.; Kerem, B.; Alon, N.; Rozmahel, R.; Grzelczak, Z.; Zielenski, J.; Lok, S.; Plavsic, N.; Chou, J.L.; et al. Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science* **1989**, *245*, 1066–1073. [[CrossRef](#)] [[PubMed](#)]
2. Elborn, J.S. Cystic fibrosis. *Lancet* **2016**, *388*, 2519–2531. [[CrossRef](#)]
3. Kreda, S.M.; Davis, C.W.; Rose, M.C. CFTR, mucins, and mucus obstruction in cystic fibrosis. *Cold Spring Harb. Perspect. Med.* **2012**, *2*, a009589. [[CrossRef](#)] [[PubMed](#)]
4. Ranganathan, S.C.; Hall, G.L.; Sly, P.D.; Stick, S.M.; Douglas, T.A.; Australian Respiratory Early Surveillance Team for Cystic, F. Early Lung Disease in Infants and Preschool Children with Cystic Fibrosis. What Have We Learned and What Should We Do about It? *Am. J. Respir. Crit. Care Med.* **2017**, *195*, 1567–1575. [[CrossRef](#)] [[PubMed](#)]
5. Sly, P.D.; Gangell, C.L.; Chen, L.; Ware, R.S.; Ranganathan, S.; Mott, L.S.; Murray, C.P.; Stick, S.M.; Investigators, A.C. Risk factors for bronchiectasis in children with cystic fibrosis. *N. Engl. J. Med.* **2013**, *368*, 1963–1970. [[CrossRef](#)] [[PubMed](#)]
6. Muhlebach, M.S.; Zorn, B.T.; Esther, C.R.; Hatch, J.E.; Murray, C.P.; Turkovic, L.; Ranganathan, S.C.; Boucher, R.C.; Stick, S.M.; Wolfgang, M.C. 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. [[CrossRef](#)]
7. Cuthbertson, L.; Walker, A.W.; Oliver, A.E.; Rogers, G.B.; Rivett, D.W.; Hampton, T.H.; Ashare, A.; Elborn, J.S.; De Soyza, A.; Carroll, M.P.; et al. Lung function and microbiota diversity in cystic fibrosis. *Microbiome* **2020**, *8*, 45. [[CrossRef](#)] [[PubMed](#)]

8. Marsland, B.J.; Trompette, A.; Gollwitzer, E.S. The Gut-Lung Axis in Respiratory Disease. *Ann. Am. Thorac. Soc.* **2015**, *12* (Suppl. 2), S150–S156. [[CrossRef](#)]
9. Budden, K.F.; Gellatly, S.L.; Wood, D.L.; Cooper, M.A.; Morrison, M.; Hugenholtz, P.; Hansbro, P.M. Emerging pathogenic links between microbiota and the gut-lung axis. *Nat. Rev. Microbiol.* **2017**, *15*, 55–63. [[CrossRef](#)]
10. Anderson, J.L.; Miles, C.; Tierney, A.C. Effect of probiotics on respiratory, gastrointestinal and nutritional outcomes in patients with cystic fibrosis: A systematic review. *J. Cyst. Fibros. Off. J. Eur. Cyst. Fibros. Soc.* **2017**, *16*, 186–197. [[CrossRef](#)]
11. Dickson, R.P.; Erb-Downward, J.R.; Freeman, C.M.; McCloskey, L.; Falkowski, N.R.; Huffnagle, G.B.; Curtis, J.L. Bacterial Topography of the Healthy Human Lower Respiratory Tract. *mBio* **2017**, *8*. [[CrossRef](#)]
12. Reen, F.J.; Woods, D.F.; Mooij, M.J.; Adams, C.; O’Gara, F. Respiratory pathogens adopt a chronic lifestyle in response to bile. *PLoS ONE* **2012**, *7*, e45978. [[CrossRef](#)]
13. Reen, F.J.; Flynn, S.; Woods, D.F.; Dunphy, N.; Chroinin, M.N.; Mullane, D.; Stick, S.; Adams, C.; O’Gara, F. Bile signalling promotes chronic respiratory infections and antibiotic tolerance. *Sci. Rep.* **2016**, *6*, 29768. [[CrossRef](#)]
14. Caparros-Martin, J.A.; Flynn, S.; Reen, F.J.; Woods, D.F.; Agudelo-Romero, P.; Ranganathan, S.C.; Stick, S.M.; O’Gara, F. The Detection of Bile Acids in the Lungs of Paediatric Cystic Fibrosis Patients Is Associated with Altered Inflammatory Patterns. *Diagnostics* **2020**, *10*, 282. [[CrossRef](#)]
15. Caparros-Martin, J.A.; Lareu, R.R.; Ramsay, J.P.; Peplies, J.; Reen, F.J.; Headlam, H.A.; Ward, N.C.; Croft, K.D.; Newsholme, P.; Hughes, J.D.; et al. Statin therapy causes gut dysbiosis in mice through a PXR-dependent mechanism. *Microbiome* **2017**, *5*, 95. [[CrossRef](#)]
16. Klindworth, A.; Pruesse, E.; Schweer, T.; Peplies, J.; Quast, C.; Horn, M.; Glockner, F.O. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **2013**, *41*, e1. [[CrossRef](#)]
17. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glockner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, D590–D596. [[CrossRef](#)]
18. Le Cao, K.A.; Costello, M.E.; Lakis, V.A.; Bartolo, F.; Chua, X.Y.; Brazeilles, R.; Rondeau, P. MixMC: A Multivariate Statistical Framework to Gain Insight into Microbial Communities. *PLoS ONE* **2016**, *11*, e0160169. [[CrossRef](#)]
19. Davis, N.M.; Proctor, D.M.; Holmes, S.P.; Relman, D.A.; Callahan, B.J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **2018**, *6*, 226. [[CrossRef](#)]
20. Paulson, J.N.; Stine, O.C.; Bravo, H.C.; Pop, M. Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods* **2013**, *10*, 1200–1202. [[CrossRef](#)]
21. Salter, S.J.; Cox, M.J.; Turek, E.M.; Calus, S.T.; Cookson, W.O.; Moffatt, M.F.; Turner, P.; Parkhill, J.; Loman, N.J.; Walker, A.W. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **2014**, *12*, 87. [[CrossRef](#)]
22. De Goffau, M.C.; Lager, S.; Sovio, U.; Gaccioli, F.; Cook, E.; Peacock, S.J.; Parkhill, J.; Charnock-Jones, D.S.; Smith, G.C.S. Human placenta has no microbiome but can contain potential pathogens. *Nature* **2019**, *572*, 329–334. [[CrossRef](#)]
23. Aagaard, K.; Ma, J.; Antony, K.M.; Ganu, R.; Petrosino, J.; Versalovic, J. The placenta harbors a unique microbiome. *Sci. Transl. Med.* **2014**, *6*, 237ra265. [[CrossRef](#)]
24. Eisenhofer, R.; Minich, J.J.; Marotz, C.; Cooper, A.; Knight, R.; Weyrich, L.S. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends Microbiol.* **2019**, *27*, 105–117. [[CrossRef](#)] [[PubMed](#)]
25. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017.
26. Scrucca, L.; Fop, M.; Murphy, T.B.; Raftery, A.E. mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. *R J.* **2016**, *8*, 289–317. [[CrossRef](#)] [[PubMed](#)]
27. Asshauer, K.P.; Wemheuer, B.; Daniel, R.; Meinicke, P. Tax4Fun: Predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* **2015**, *31*, 2882–2884. [[CrossRef](#)]
28. Goeman, J.J.; van de Geer, S.A.; de Kort, F.; van Houwelingen, H.C. A global test for groups of genes: Testing association with a clinical outcome. *Bioinformatics* **2004**, *20*, 93–99. [[CrossRef](#)] [[PubMed](#)]

29. Chong, J.; Liu, P.; Zhou, G.; Xia, J. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. *Nat. Protoc.* **2020**, *15*, 799–821. [[CrossRef](#)]
30. Bodein, A.; Chapleur, O.; Droit, A.; Le Cao, K.A. A Generic Multivariate Framework for the Integration of Microbiome Longitudinal Studies with Other Data Types. *Front. Genet.* **2019**, *10*, 963. [[CrossRef](#)]
31. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016.
32. Barter, R.; Yu, B. Superheat: A Graphical Tool for Exploring Complex Datasets Using Heatmaps. R Package Version 0.1.0. 2017. Available online: <https://CRAN.R-project.org/package=superheat> (accessed on 1 June 2020).
33. Kassambara, A.; Mundt, F. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. 2020. Available online: <https://CRAN.R-project.org/package=factoextra> (accessed on 1 June 2020).
34. Lüdtke, D. sjPlot: Data Visualization for Statistics in Social Science. 2020. Available online: <https://CRAN.R-project.org/package=sjPlot> (accessed on 1 September 2020).
35. Rohart, F.; Gautier, B.; Singh, A.; Le Cao, K.A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* **2017**, *13*, e1005752. [[CrossRef](#)]
36. Rosenow, T.; Oudraad, M.C.; Murray, C.P.; Turkovic, L.; Kuo, W.; de Bruijne, M.; Ranganathan, S.C.; Tiddens, H.A.; Stick, S.M.; Australian Respiratory Early Surveillance Team for Cystic, F. PRAGMA-CF. A Quantitative Structural Lung Disease Computed Tomography Outcome in Young Children with Cystic Fibrosis. *Am. J. Respir. Crit. Care Med.* **2015**, *191*, 1158–1165. [[CrossRef](#)]
37. Sly, P.D.; Brennan, S.; Gangell, C.; de Klerk, N.; Murray, C.; Mott, L.; Stick, S.M.; Robinson, P.J.; Robertson, C.F.; Ranganathan, S.C.; et al. Lung disease at diagnosis in infants with cystic fibrosis detected by newborn screening. *Am. J. Respir. Crit. Care Med.* **2009**, *180*, 146–152. [[CrossRef](#)]
38. Mott, L.S.; Park, J.; Murray, C.P.; Gangell, C.L.; de Klerk, N.H.; Robinson, P.J.; Robertson, C.F.; Ranganathan, S.C.; Sly, P.D.; Stick, S.M.; et al. Progression of early structural lung disease in young children with cystic fibrosis assessed using CT. *Thorax* **2012**, *67*, 509–516. [[CrossRef](#)] [[PubMed](#)]
39. Pillarisetti, N.; Williamson, E.; Linnane, B.; Skoric, B.; Robertson, C.F.; Robinson, P.; Massie, J.; Hall, G.L.; Sly, P.; Stick, S.; et al. Infection, inflammation, and lung function decline in infants with cystic fibrosis. *Am. J. Respir. Crit. Care Med.* **2011**, *184*, 75–81. [[CrossRef](#)]
40. Aldhahrani, A.; Verdon, B.; Ward, C.; Pearson, J. Effects of bile acids on human airway epithelial cells: Implications for aerodigestive diseases. *ERJ Open Res.* **2017**, *3*. [[CrossRef](#)]
41. Blondeau, K.; Dupont, L.J.; Mertens, V.; Verleden, G.; Malfroot, A.; Vandenzplas, Y.; Hauser, B.; Sifrim, D. Gastro-oesophageal reflux and aspiration of gastric contents in adult patients with cystic fibrosis. *Gut* **2008**, *57*, 1049–1055. [[CrossRef](#)]
42. Legendre, C.; Reen, F.J.; Woods, D.F.; Mooij, M.J.; Adams, C.; O’Gara, F. Bile acids repress hypoxia-inducible factor 1 signaling and modulate the airway immune response. *Infect. Immun.* **2014**, *82*, 3531–3541. [[CrossRef](#)]
43. Pauwels, A.; Decraene, A.; Blondeau, K.; Mertens, V.; Farre, R.; Proesmans, M.; Van Bleyenbergh, P.; Sifrim, D.; Dupont, L.J. Bile acids in sputum and increased airway inflammation in patients with cystic fibrosis. *Chest* **2012**, *141*, 1568–1574. [[CrossRef](#)] [[PubMed](#)]
44. Downey, D.G.; Bell, S.C.; Elborn, J.S. Neutrophils in cystic fibrosis. *Thorax* **2009**, *64*, 81–88. [[CrossRef](#)]
45. Dakin, C.J.; Numa, A.H.; Wang, H.; Morton, J.R.; Vertzyas, C.C.; Henry, R.L. Inflammation, infection, and pulmonary function in infants and young children with cystic fibrosis. *Am. J. Respir. Crit. Care Med.* **2002**, *165*, 904–910. [[CrossRef](#)]
46. Frayman, K.B.; Armstrong, D.S.; Carzino, R.; Ferkol, T.W.; Grimwood, K.; Storch, G.A.; Teo, S.M.; Wylie, K.M.; Ranganathan, S.C. The lower airway microbiota in early cystic fibrosis lung disease: A longitudinal analysis. *Thorax* **2017**, *72*, 1104–1112. [[CrossRef](#)]
47. Prevaes, S.M.; de Steenhuijsen Pipers, W.A.; de Winter-de Groot, K.M.; Janssens, H.M.; Tramper-Stranders, G.A.; Chu, M.L.; Tiddens, H.A.; van Westreenen, M.; van der Ent, C.K.; Sanders, E.A.; et al. Concordance between upper and lower airway microbiota in infants with cystic fibrosis. *Eur. Respir. J.* **2017**, *49*. [[CrossRef](#)]
48. Marsh, R.L.; Kaestli, M.; Chang, A.B.; Binks, M.J.; Pope, C.E.; Hoffman, L.R.; Smith-Vaughan, H.C. The microbiota in bronchoalveolar lavage from young children with chronic lung disease includes taxa present in both the oropharynx and nasopharynx. *Microbiome* **2016**, *4*, 37. [[CrossRef](#)]

49. Segal, L.N.; Alekseyenko, A.V.; Clemente, J.C.; Kulkarni, R.; Wu, B.; Gao, Z.; Chen, H.; Berger, K.I.; Goldring, R.M.; Rom, W.N.; et al. Enrichment of lung microbiome with supraglottic taxa is associated with increased pulmonary inflammation. *Microbiome* **2013**, *1*, 19. [[CrossRef](#)]
50. Segal, L.N.; Clemente, J.C.; Tsay, J.C.; Koralov, S.B.; Keller, B.C.; Wu, B.G.; Li, Y.; Shen, N.; Ghedin, E.; Morris, A.; et al. Enrichment of the lung microbiome with oral taxa is associated with lung inflammation of a Th17 phenotype. *Nat. Microbiol.* **2016**, *1*, 16031. [[CrossRef](#)]
51. Parkins, M.D.; Floto, R.A. Emerging bacterial pathogens and changing concepts of bacterial pathogenesis in cystic fibrosis. *J. Cyst. Fibros. Off. J. Eur. Cyst. Fibros. Soc.* **2015**, *14*, 293–304. [[CrossRef](#)]
52. Lyczak, J.B.; Cannon, C.L.; Pier, G.B. Lung Infections Associated with Cystic Fibrosis. *Clin. Microbiol. Rev.* **2002**, *15*, 194–222. [[CrossRef](#)]
53. Lipuma, J.J. The changing microbial epidemiology in cystic fibrosis. *Clin. Microbiol. Rev.* **2010**, *23*, 299–323. [[CrossRef](#)]
54. Sutter, V.L. Anaerobes as normal flora. *Rev. Infect. Dis.* **1984**, *6*, S62–S66. [[CrossRef](#)]
55. Espinoza, J.; Harkins, D.; Torralba, M.; Gomez, A.; Highlander, S.; Jones, M.; Leong, P.; Saffery, R.; Bockmann, M.; Kuelbs, C.; et al. Supragingival Plaque Microbiome Ecology and Functional Potential in the Context of Health and Disease. *mBio* **2018**, *9*, e01631-18. [[CrossRef](#)]
56. Aas, J.A.; Paster, B.J.; Stokes, L.N.; Olsen, I.; Dewhirst, F.E. Defining the normal bacterial flora of the oral cavity. *J. Clin. Microbiol.* **2005**, *43*, 5721–5732. [[CrossRef](#)]
57. Makki, K.; Deehan, E.C.; Walter, J.; Backhed, F. The Impact of Dietary Fiber on Gut Microbiota in Host Health and Disease. *Cell Host Microbe* **2018**, *23*, 705–715. [[CrossRef](#)]
58. Sayin, S.I.; Wahlstrom, A.; Felin, J.; Jantti, S.; Marschall, H.U.; Bamberg, K.; Angelin, B.; Hyotylainen, T.; Oresic, M.; Backhed, F. Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist. *Cell Metab.* **2013**, *17*, 225–235. [[CrossRef](#)] [[PubMed](#)]
59. Li, F.; Jiang, C.; Krausz, K.W.; Li, Y.; Albert, I.; Hao, H.; Fabre, K.M.; Mitchell, J.B.; Patterson, A.D.; Gonzalez, F.J. Microbiome remodelling leads to inhibition of intestinal farnesoid X receptor signalling and decreased obesity. *Nat. Commun.* **2013**, *4*, 2384. [[CrossRef](#)]
60. Zheng, X.; Huang, F.; Zhao, A.; Lei, S.; Zhang, Y.; Xie, G.; Chen, T.; Qu, C.; Rajani, C.; Dong, B.; et al. Bile acid is a significant host factor shaping the gut microbiome of diet-induced obese mice. *BMC Biol.* **2017**, *15*, 120. [[CrossRef](#)]
61. Aseeri, A.; Brodli, M.; Lordan, J.; Corris, P.; Pearson, J.; Ward, C.; Manning, N. Bile acids are present in the lower airways of people with cystic fibrosis. *Am. J. Respir. Crit. Care Med.* **2012**, *185*, 463. [[CrossRef](#)]
62. Reen, F.J.; Woods, D.F.; Mooij, M.J.; Chroinin, M.N.; Mullane, D.; Zhou, L.; Quille, J.; Fitzpatrick, D.; Glennon, J.D.; McGlacken, G.P.; et al. Aspirated bile: A major host trigger modulating respiratory pathogen colonisation in cystic fibrosis patients. *Eur. J. Clin. Microbiol. Infect. Dis.* **2014**, *33*, 1763–1771. [[CrossRef](#)]
63. McNally, P.; Ervine, E.; Shields, M.D.; Dimitrov, B.D.; El Nazir, B.; Taggart, C.C.; Greene, C.M.; McElvaney, N.G.; Grealley, P. High concentrations of pepsin in bronchoalveolar lavage fluid from children with cystic fibrosis are associated with high interleukin-8 concentrations. *Thorax* **2011**, *66*, 140–143. [[CrossRef](#)]
64. Clarke, D.; Gorman, I.; Ringholz, F.; McDermott, M.; Cox, D.W.; Grealley, P.; Linnane, B.; Mc Nally, P. Pulmonary aspiration in preschool children with cystic fibrosis. *Respir. Res.* **2018**, *19*, 255. [[CrossRef](#)]
65. Hang, S.; Paik, D.; Yao, L.; Kim, E.; Trinath, J.; Lu, J.; Ha, S.; Nelson, B.N.; Kelly, S.P.; Wu, L.; et al. Bile acid metabolites control TH17 and Treg cell differentiation. *Nature* **2019**, *576*, 143–148. [[CrossRef](#)] [[PubMed](#)]
66. Campbell, C.; McKenney, P.T.; Konstantinovskiy, D.; Isaeva, O.I.; Schizas, M.; Verter, J.; Mai, C.; Jin, W.B.; Guo, C.J.; Violante, S.; et al. Bacterial metabolism of bile acids promotes generation of peripheral regulatory T cells. *Nature* **2020**, *581*, 475–479. [[CrossRef](#)] [[PubMed](#)]
67. Song, X.; Sun, X.; Oh, S.F.; Wu, M.; Zhang, Y.; Zheng, W.; Geva-Zatorsky, N.; Jupp, R.; Mathis, D.; Benoist, C.; et al. Microbial bile acid metabolites modulate gut RORgamma(+) regulatory T cell homeostasis. *Nature* **2020**, *577*, 410–415. [[CrossRef](#)]
68. Tan, H.L.; Regamey, N.; Brown, S.; Bush, A.; Lloyd, C.M.; Davies, J.C. The Th17 pathway in cystic fibrosis lung disease. *Am. J. Respir. Crit. Care Med.* **2011**, *184*, 252–258. [[CrossRef](#)]
69. Jorth, P.; Ehsan, Z.; Rezayat, A.; Caldwell, E.; Pope, C.; Brewington, J.J.; Goss, C.H.; Benschoter, D.; Clancy, J.P.; Singh, P.K. Direct Lung Sampling Indicates That Established Pathogens Dominate Early Infections in Children with Cystic Fibrosis. *Cell Rep.* **2019**, *27*, 1190–1204.e1193. [[CrossRef](#)]

70. Zemanick, E.T.; Wagner, B.D.; Robertson, C.E.; Ahrens, R.C.; Chmiel, J.F.; Clancy, J.P.; Gibson, R.L.; Harris, W.T.; Kurland, G.; Laguna, T.A.; et al. Airway microbiota across age and disease spectrum in cystic fibrosis. *Eur. Respir. J.* **2017**, *50*. [[CrossRef](#)]
71. Flynn, J.M.; Niccum, D.; Dunitz, J.M.; Hunter, R.C. Evidence and Role for Bacterial Mucin Degradation in Cystic Fibrosis Airway Disease. *PLoS Pathog.* **2016**, *12*, e1005846. [[CrossRef](#)]
72. Duboc, H.; Rajca, S.; Rainteau, D.; Benarous, D.; Maubert, M.A.; Quervain, E.; Thomas, G.; Barbu, V.; Humbert, L.; Despras, G.; et al. Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. *Gut* **2013**, *62*, 531–539. [[CrossRef](#)]
73. Ulluwishewa, D.; Wang, L.; Pereira, C.; Flynn, S.; Cain, E.; Stick, S.; Reen, F.J.; Ramsay, J.P.; O’Gara, F. Dissecting the regulation of bile-induced biofilm formation in *Staphylococcus aureus*. *Microbiology* **2016**, *162*, 1398–1406. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).