White Paper

Dynamics in the airway microbiome in children with Cystic fibrosis as revealed by 16S rRNA gene sequencing

Winnie Ridderberg, Ph.D.

QIAGEN Bioinformatics, Microbial Genomics, Aarhus Denmark.

Introduction

Cystic fibrosis (CF) is an autosomal recessive disorder caused by dysfunction of the cystic fibrosis transmembrane conductance regulator (CFTR), an ion channel that functions to transport chloride and thiocyanate ions across epithelial cell membranes. Accordingly, CF is a multi-organ disease. However, the majority of the associated morbidity and mortality is attributed to pulmonary symptoms. In the airways, loss of CFTR function leads to thickened mucus and impaired ciliary clearance creating an environment favouring bacterial growth. CF is thus characterised by recurrent pulmonary infections and chronic inflammation leading to progressive lung destruction and, in many cases, ultimately respiratory failure (Elborn J S, 2016).

Culture-independent approaches based on next generation sequencing methodologies have revealed that the airway microbiome is complex and constantly changing. Microbial diversity has been shown to be increased in young patients, and often progressively decreases with age (de Koff et al. 2016). In young patients with CF, obtaining samples representative of the lower airways is challenging as they often do not expectorate sputum. Consequently, the lower airway microbiome of young patients with CF is insufficiently characterised (Hoppe and Zemanick 2017). The purpose of this study was to investigate the dynamics of the lower airway microbiota during a period of six months in children with CF, using samples obtained by laryngeal suction. The study was done in collaboration with the Department of Clinical Microbiology and CF Centre at Aarhus University Hospital, Denmark.

In this white paper, we demonstrate the analysis of 16S rRNA amplicon sequencing data from complex microbial communities using tools from the CLC Microbial Genomics Module.

Materials and Methods

The original study included serial airway samples from 62 children under 11 years of age. Samples were collected at monthly routine visits to the CF clinic during a period of six months. For the purposes of the analysis methodologies presented in this white paper, we have included only a small subset of 22 samples from four children with CF at ages 0, 5, 6 and 10 years. Patient characteristics are listed in Table 1.

Table 1. Patients and sample meta-data.

Patient No.	Age	Sample No.	Sampling date	Exacerbation at sampling time	Antimicrobial treatment at sampling time
CF21	5	21-1	18.09.15	-	-
		21-2	17.10.15	-	-
		21-3	30.10.15	-	-
		21-4	15.12.15	-	-
		21-5	29.01.16	-	-
		21-6	26.02.16	-	-
CF28	6	28-1	16.09.15	-	+
		28-2	21.10.15	-	+
		28-3	16.11.15	-	-
		28-4	16.12.15	-	-
		28-5	27.01.16	-	-
CF39	0	39-1	16.09.15	-	-
		39-2	25.09.15	+	-
		39-3	28.10.15	-	-
		39-4	01.12.15	+	-
		39-5	02.02.16	+	-
CF58	10	58-1	04.09.15	+	-
		58-2	07.10.15	-	-
		58-3	13.11.15	-	-
		58-4	15.12.15	-	+
		58-5	05.01.16	-	-
		58-6	10.02.16	-	-

Bacterial V3-V4 16S rRNA gene sequences were amplified following the Illumina 16S Metagenomic Sequencing Library Preparation protocol (Illumina). Amplicons were sequenced on the Illumina MiSeq using 2x300 bp paired-end chemistry.

Bioinformatics analysis was conducted using the CLC Microbial Genomics Module. Figure 1 illustrates the tools and workflow used for data analysis. Bacterial reads were clustered against the SILVA reference database (v128) at 97% identity threshold (Quast et al. 2013).

Results

Patient specific microbiotas

Clustering reads at phylum level revealed large fluctuations in the microbiome composition within each patient and distinct differences between individual patients. The most represented phyla were Proteobacteria and Firmicutes (Figure 2A). Across all 22 samples, a total of 118 genera were detected, but ten most abundant genera accounted for ~90% of the total relative abundance; *Neisseria* (20%), *Streptococcus* (18%), *Bordetella* (13%), *Haemophilus* (10%), *Veillonella* (9%), *Moraxella* (7%), *Gemella* (4%), *Rothia* (4%), *Prevotella* (4%), *Fusobacterium* (2%) (Figure 2B).

Using the Bray-Curtis dissimilarity metric we found that samples cluster by patient (Figure 3A) and that lung microbiota profiles were significantly different between patients (PERMANOVA p=0.0001). Samples from individual patients resemble each other more than samples from other patients, indicating that the airway microbiota is personal to each CF patient. The difference between each clustered group of patient samples was reduced, however, when beta-diversity was estimated using the Weighted UniFrac metric, which also accounts for the phylogenetic relationship of taxa (Figure 3B). Here, the microbiota of patient CF39 was not statistically significantly different from

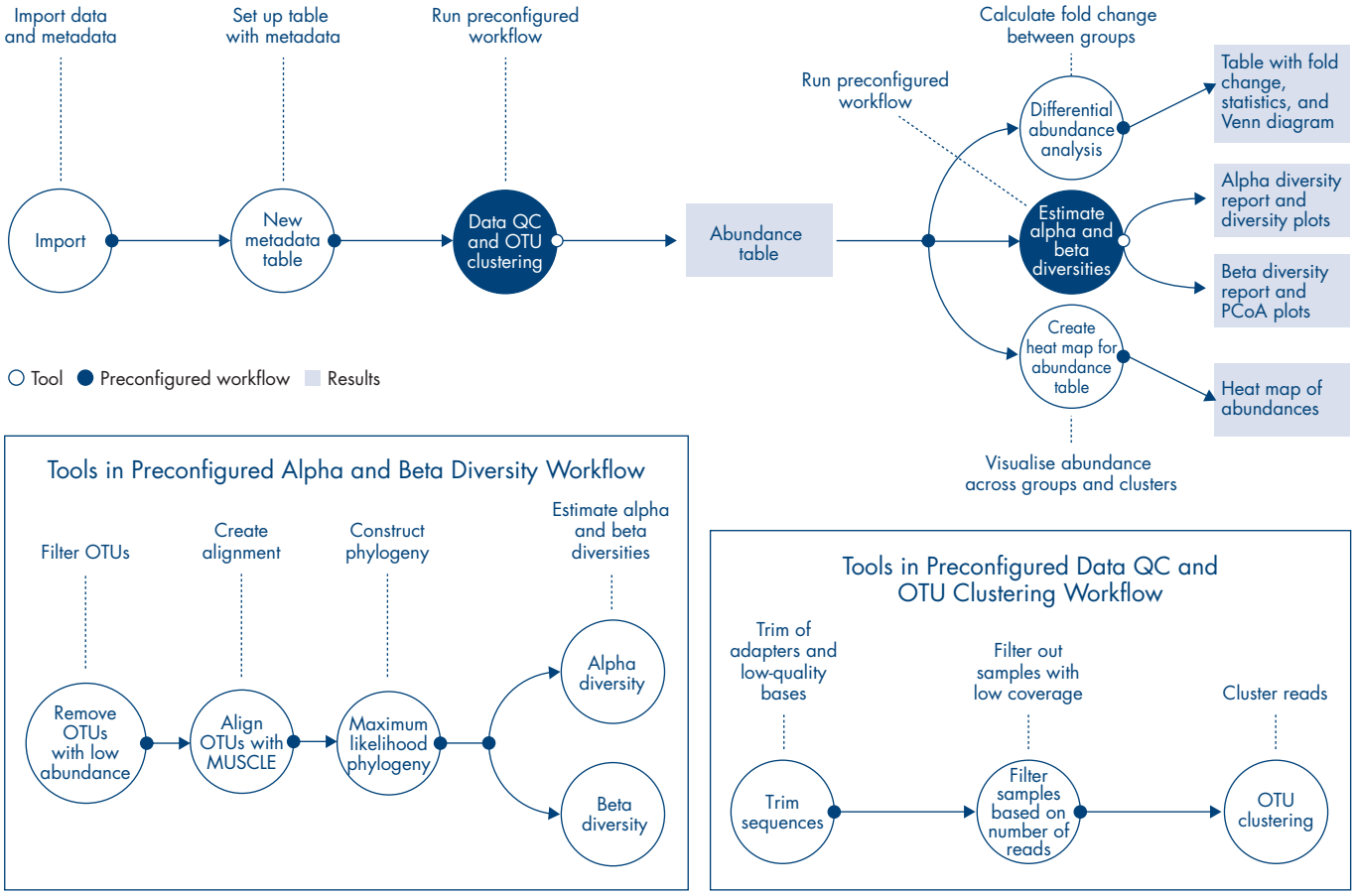

Figure 1. Data analysis workflow using tools from CLC Microbial Genomics Module

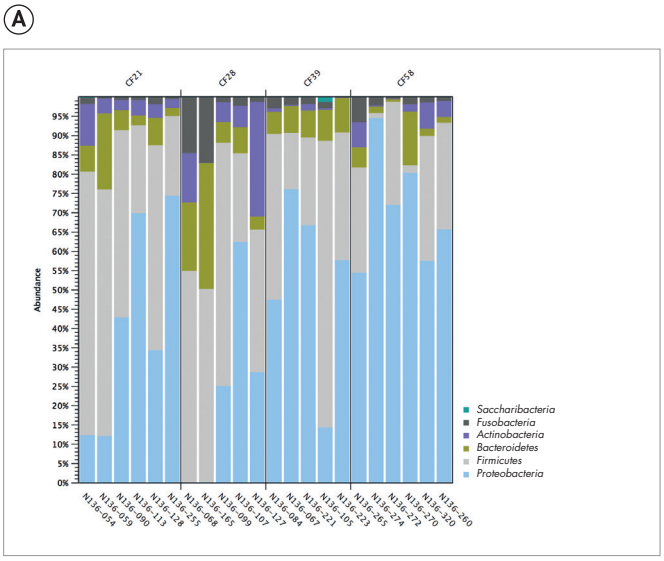

Figure 2. Taxonomic profile of samples. Taxa were clustered at (A) phylum level and (B) genus level. igure legend modified for clarity.

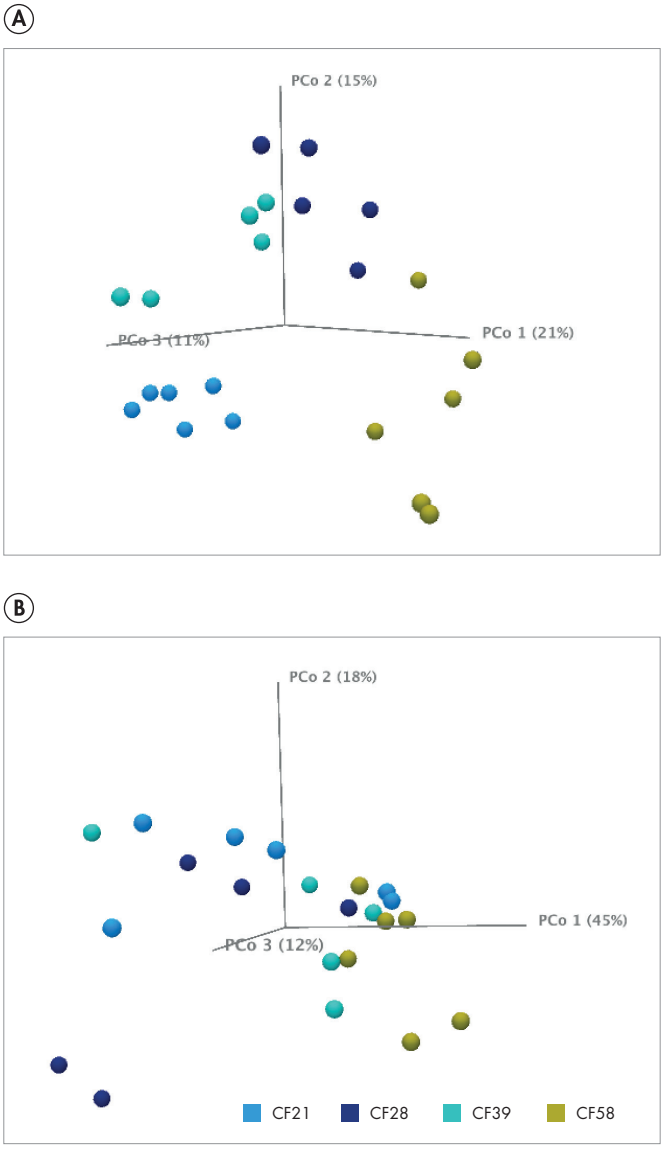

Figure 3. Principal Component Analysis. Beta diversity estimated by (A) Bray-Curtis dissimilarity and (B) Weighted UniFrac.

that of patients CF21 (p=0.25758) and CF28 (p=0.13492) (PERMANOVA). The microbiotas of the remaining patients were significantly different from each other.

To understand which bacterial genera differentiated the microbiota of individual patients, we investigated the relative abundance of the 15 most abundant genera across all patients. As seen from the Venn diagrams of the differential abundance (Figure 4), very few taxa set the microbiota of the individual patients apart. Only the relative abundance of *Veillonella* was differentially represented between CF21 and the remaining patients. Similarly, the only taxa significantly differentially represented between CF58 and all the other patients was *Bordetella*. Patient CF28 differentiated from CF21 and CF39 by the relative abundance of *Moraxella*, and from CF39 and CF58 by the relative abundance of *Alloprevotella*. CF39 differentiated from CF58 by taxa *Gemella*, *Moraxella*, and *Fusobacteria*, and from CF28 by taxa *Alloprevotella*, *Moraxella*, and *Staphylococcus* (Figure 5).

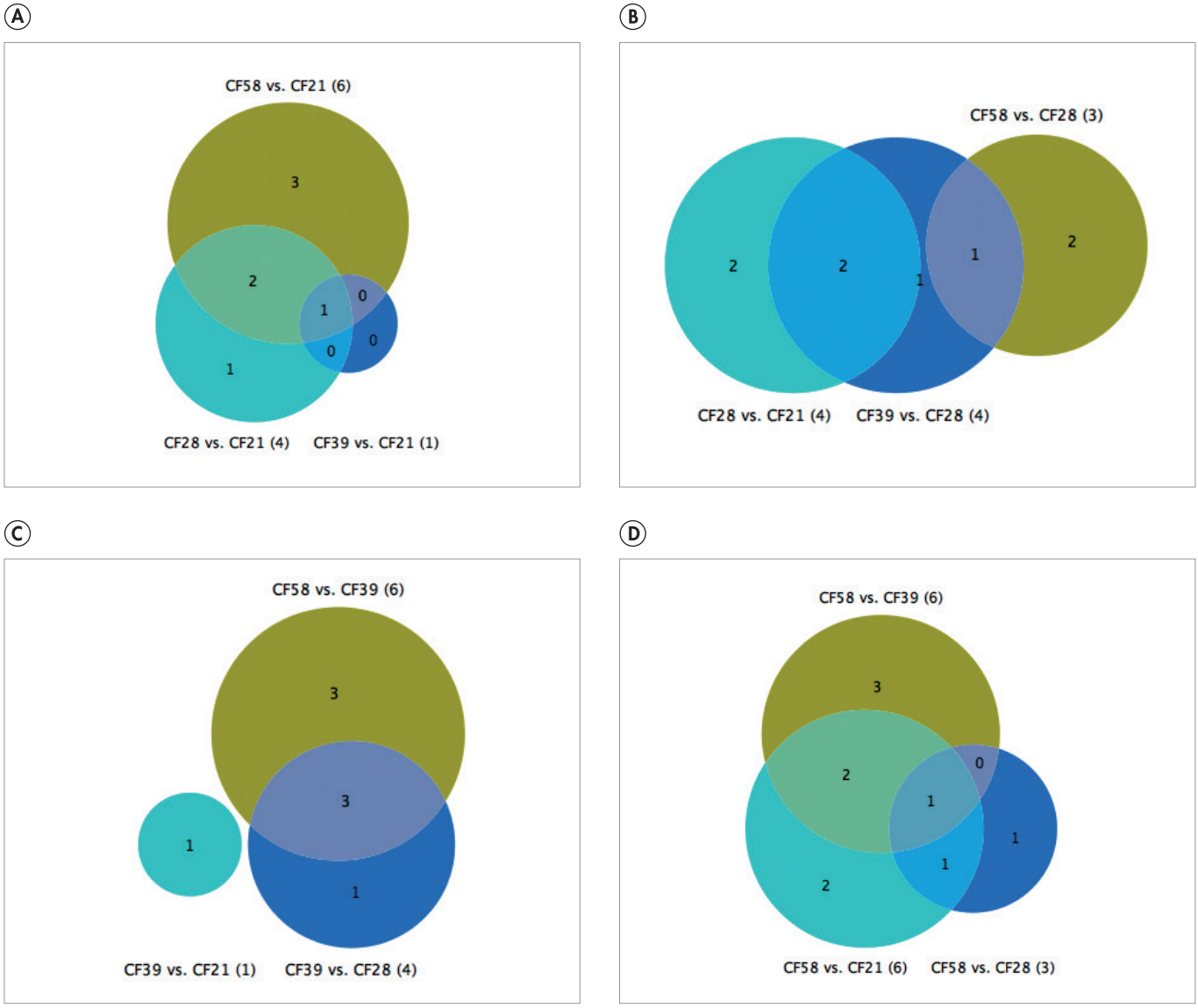

Figure 4. Venn diagrams. Comparison of top 15 most abundant genera across all patients – (A) CF21 vs all, (B) CF28 vs all, (C) CF39 vs all, (D) CF58 vs all.

Microbial community dynamics

As the microbiota was shown to fluctuate substantially within patients during the six-month study period (Figure 2), we sought to investigate the underlying causes of such perturbations in the microbial community. For all patients, antimicrobial treatment regimens administered and exacerbations of disease occurring during sampling were registered. Figure 5 shows the alpha diversity, estimated by phylogenetic diversity, of individual samples for each of the four patients. The alpha diversity of the airway microbiota was found to be lower if patients were receiving antimicrobial treatment. Clinical exacerbations, did not affect alpha diversity of microbial communities.

We calculated the beta diversity (Weighted UniFrac) of longitudinal samples to evaluate the microbiota changes over time and in response to antimicrobial treatment and exacerbation of disease. Patient CF21 was clinically stabile during the study period with no exacerbations and no antimicrobial treatments. However, we see that the microbiota changes considerably during the six months (Figure 6A). The differential abundance analysis (Figure 7A) showed that the lung microbiota was

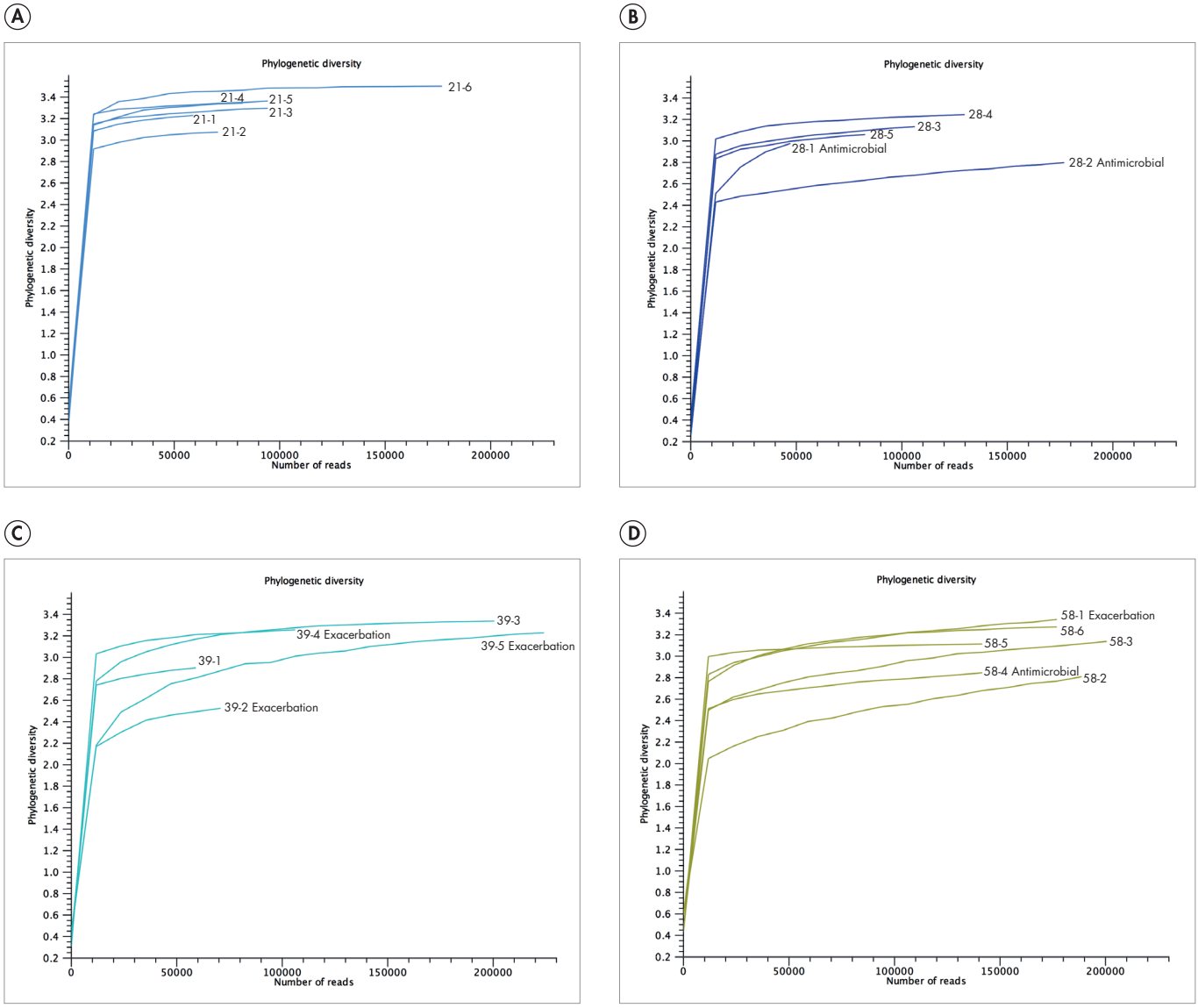

Figure 5. Alpha-diversity estimates (phylogenetic distance) for patient samples. (A) CF21, (B) CF28, (C) CF39, (D) CF58. Samples taken during antimicrobial treatment or disease exacerbation are marked.

initially dominated by normal flora or oral cavity related taxa (*Prevotella, Veillonella, Rothia, Granulicatella, Streptococcus*), but over time the relative abundance of these taxa was reduced and the relative abundance of opportunistic pathogens *Haemophilus, Moraxella*, and *Aggregatibacter* gradually increased.

The lung microbiota of patient CF28 changed dramatically from the first two samples, taken in September and October, to the remaining samples from November, December and January (Figure 6B). This significant alteration of the microbiota was most likely due to the antimicrobial treatment the patient was given during the first two sampling time points. During antimicrobial treatment, the microbial community was dominated by *Staphylococcus, Fusobacteria* and *Streptococcus*, and the normal flora was suppressed by the dominant pathogen (*Staphylococcus*) and treatment. After treatment commenced, the relative abundance of normal flora *Prevotella, Rohtia, Gemella, Neisseria*, and *Veillonella* increased in the range of 6603-fold to 424-fold (Figure 7B).

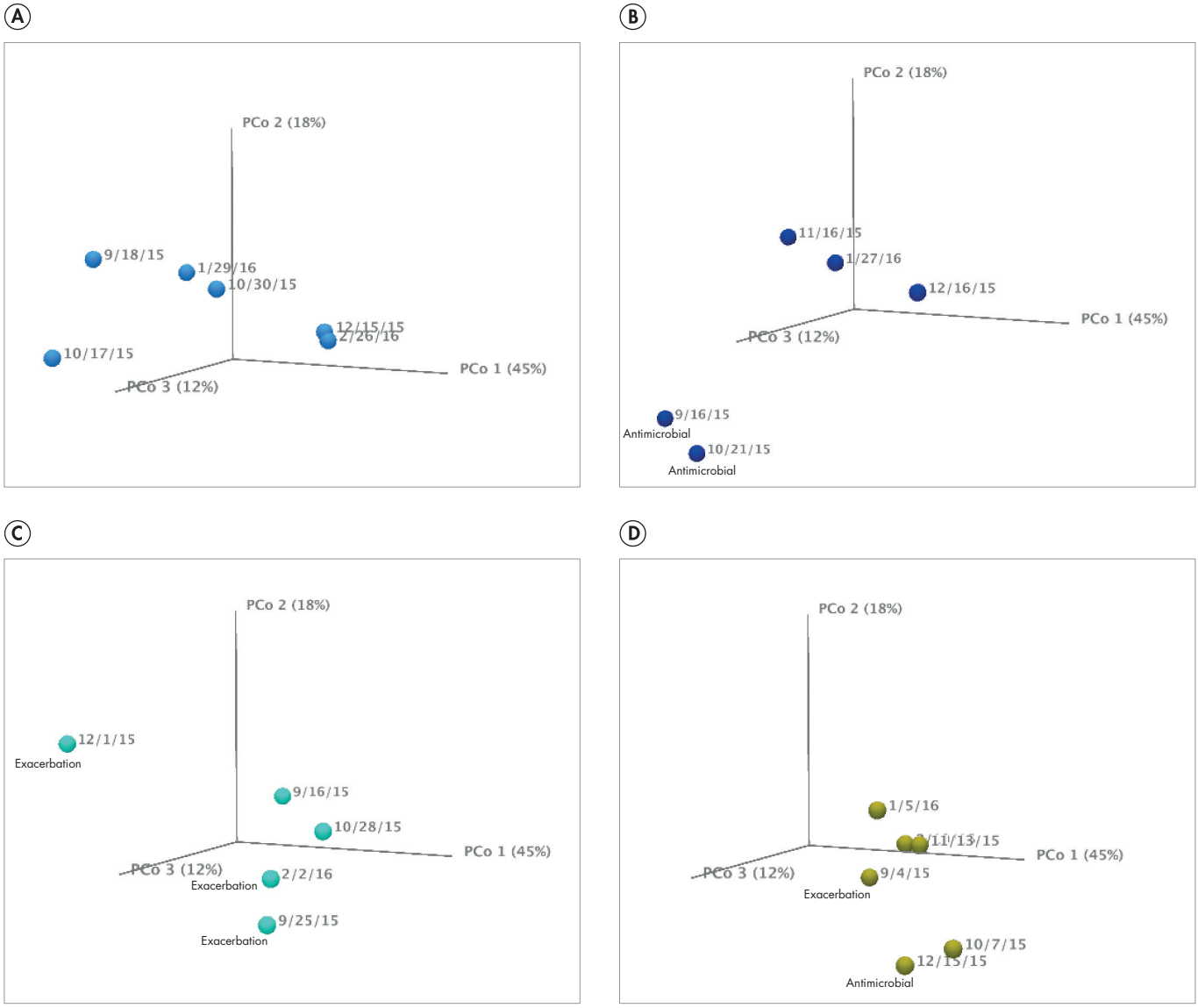

Figure 6. Principal component analysis of beta diversity estimations using Weighted UniFrac metric. (A) CF21, (B) CF28, (C) CF39, (D) CF58. Sampling dates are marked and samples taken during antimicrobial treatment and disease exacerbation are indicated.

Patient CF39 experienced three episodes of exacerbation during the study period and the PCoA plot reveal large fluctuations in the microbial community (Figure 6C). Considering the first sample from September (16-09-2015) the base line, then the second sample also from September (25-09-2015) represent a disturbance in the microbiota. In the third sample (28-10-2015) the microbiota return to baseline, and in the following two samples (12-01-2016 and 02-02-2016) the microbiota is disturbed again. Such perturbations were also revealed by the differential abundance analysis (Figure 7C), showing that during the first exacerbation the levels of Prevotella, Streptococcus, and Moraxella were increased, during the second exacerbation the relative abundance of Veillonella and Prevotella increased, and during the final exacerbation the relative levels of Moraxella and Alloprevotella were increased.

The airway microbiota of CF58 fluctuates substantially during the six-month study period. The patient experienced a single episode of disease exacerbation and had one antimicrobial treatment during sampling. As seen from the PCoA plot (Figure 6D) samples cluster in two groups – one cluster containing four samples all dominated by normal flora bacteria, the other cluster containing the sample taken during antimicrobial treatment and the sample from 07-10-2015 (first sample after exacerbation). The microbiota of these two samples were

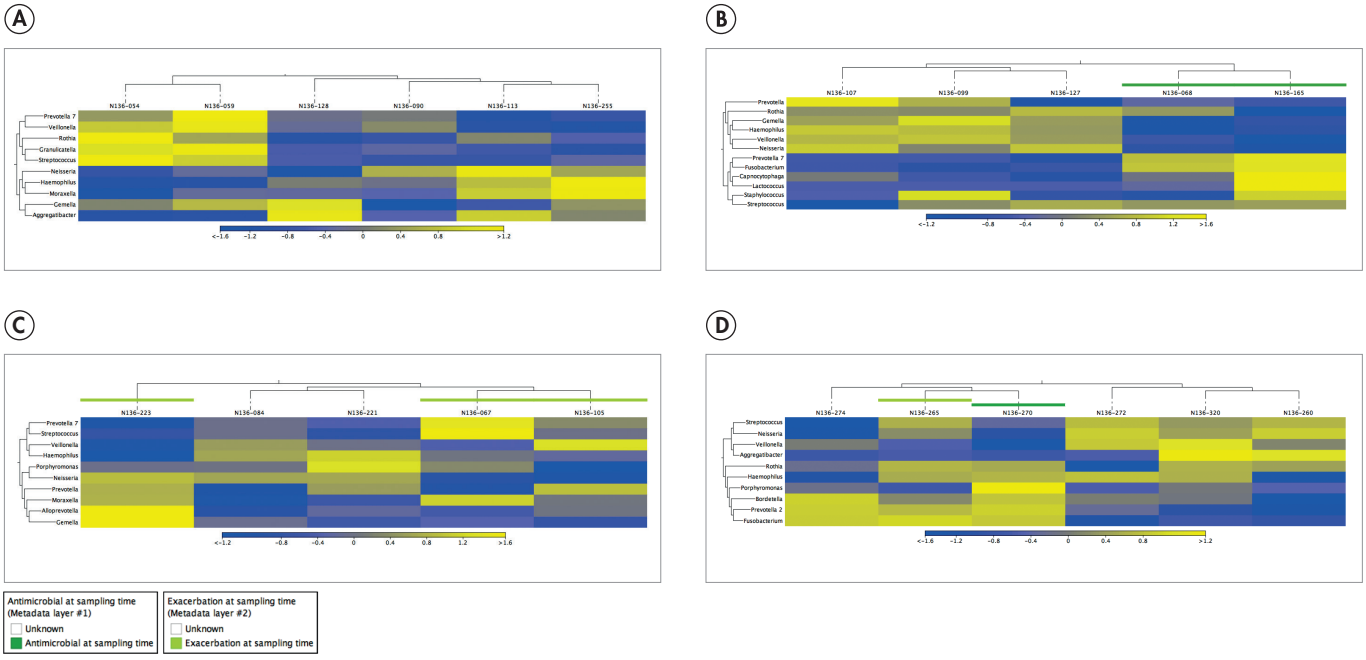

Figure 7. Heat map of differentially represented taxa. (A) CF21, (B) CF28, (C) CF39, (D) CF58.

characterised by increasing levels of *Bordetella, Prevotella* and *Fusobacteria*, and decreasing levels of *Streptococcus, Neisseria* and *Veillonella* (Figure 7D).

Conclusions

To gain better understanding of the progressive nature of CF it is important to investigate how the microbial community of the lung changes though time in relation to clinical state and antimicrobial treatment. In this white paper, we investigated the longitudinal dynamics of the CF lung microbiota from an early point in CF and found that the microbial communities were highly individual and specific to each patient. Large fluctuations in the microbiota composition were observed over time – even in the absence of disease exacerbation and influence from antimicrobial treatment. As expected, we found antimicrobial treatment administered during sampling to have a profound effect on the microbiota composition. However, the microbial communities were able to return to pre-treatment state. Additionally, we found pulmonary exacerbation to be associated with alterations in the microbiota composition.

Using the CF airway microbiome as a model, we have demonstrated the utility of CLC Microbial Genomics Module for studying complex microbial communities and identifying both common and distinct members of those communities across a range of microbiome samples. In addition, the underlying CLC Genomics Workbench software, on which the CLC Microbial Genomics Module extends, enabled end-to-end processing, analysis and insight into microbiome data collected for ongoing cystic fibrosis research.

To learn more – visit qiagenbioinformatics.com

Tutorials

OTU clustering step by step

http://resources.qiagenbioinformatics.com/tutorials/OTU_Clustering_Steps.pdf

OTU clustering using workflows

http://resources.qiagenbioinformatics.com/tutorials/OTU_Clustering_Workflows.pdf

References

Elborn J S. (2016). Cystic fibrosis. Lancet 388, 2519-31.

Hoppe J E and Zemanick E T (2017). Lessons learnt from the lower airway microbiome in early CF. Thorax, 72 (12), 1063-1064.

de Koff E M, de Winter-de Groot K M, and Bogaert D (2016). Development of the respiratory tract in cystic fibrosis. Curr Opin Pulm Med 22, 623-628.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucl. Acids Res. 41 (D1), D590-D596.

16S Metagenomic Sequencing Library Preparation, Illumina (www.illumina.com)