environmental microbiology

Environmental Microbiology (2010) **12**(5), 1293–1303 doi:10.1111/j.1462-2920.2010.02173.x

Relationship between cystic fibrosis respiratory tract bacterial communities and age, genotype, antibiotics and *Pseudomonas aeruginosa*

Vanja Klepac-Ceraj,[1]* Katherine P. Lemon,[2] Thomas R. Martin,[3] Martin Allgaier,[4] Steven W. Kembel,[5] Alixandra A. Knapp,[3] Stephen Lory,[1] Eoin L. Brodie,[6] Susan V. Lynch,[4] Brendan J. M. Bohannan,[5] Jessica L. Green,[5] Brian A. Maurer[7] and Roberto Kolter[1]

[1]*Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02115, USA. Divisions of* [2]*Infectious Diseases and* [3]*Respiratory Diseases, Children's Hospital, Harvard Medical School, Boston, MA 02115, USA.*

[4]*Department of Anesthesia and Perioperative Care, University of California San Francisco, San Francisco, CA 94143, USA.*

[5]*Center for Ecology & Evolutionary Biology, University of Oregon, Eugene, OR 97403, USA.*

[6]*Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.*

[7]*Michigan State University, Department of Fisheries and Wildlife, East Lansing, MI 48824, USA.*

Summary

Polymicrobial bronchopulmonary infections in cystic fibrosis (CF) cause progressive lung damage and death. Although the arrival of *Pseudomonas aeruginosa* often heralds a more rapid rate of pulmonary decline, there is significant inter-individual variation in the rate of decline, the causes of which remain poorly understood. By coupling culture-independent methods with ecological analyses, we discovered correlations between bacterial community profiles and clinical disease markers in respiratory tracts of 45 children with CF. Bacterial community complexity was inversely correlated with patient age, presence of *P. aeruginosa* and antibiotic exposure, and was related to CF genotype. Strikingly, bacterial communities lacking *P. aeruginosa* were much more similar to each other than were those containing *P. aeruginosa*, regardless of antibiotic exposure. This suggests that community composition might be a better predictor of disease progression than the presence of *P. aeruginosa* alone and deserves further study.

Received 22 October, 2009; accepted 19 December, 2009. *For correspondence. E-mail vanjakle@gmail.com; Tel. (+1) 617 892 8592; Fax (+1) 617 262 4021.

© 2010 Society for Applied Microbiology and Blackwell Publishing Ltd

Introduction

Cystic fibrosis (CF) is the most common inherited chronic disease in Caucasians (Gibson *et al*., 2003). In this autosomal recessive disorder an abnormal epithelial cell chloride channel (the cystic fibrosis transmembrane conductance regulator or CFTR) results in dehydrated airway surface liquid (Rubin, 2009). This impairs normal airway clearance mechanisms resulting in chronic bronchopulmonary infections (Ratjen and Doring, 2003; Ratjen, 2006). The chronic infections incite a persistent neutrophilic inflammatory response, ultimately destroying lung tissue (Chmiel and Davis, 2003). Bacterial cultivation has shown that multiple species, including anaerobes, usually coexist in the respiratory tract of CF patients (Wahab *et al*., 2004; Harrison, 2007; Tunney *et al*., 2008). Moreover, recent culture-independent studies have detected many bacterial species not previously noticed by cultivation (Rogers *et al*., 2003; 2004; Harris *et al*., 2007), and suggested that some species such as *Prevotella* spp. and *Streptococcus milleri* might contribute to worsening lung function (Harris *et al*., 2007; Sibley *et al*., 2008). Based on cultivation data, there is a general pattern of pathogen succession starting with *Haemophilus influenzae* and *Staphylococcus aureus*, both common constituents of healthy upper respiratory tract microbiota, and progressing to *Pseudomonas aeruginosa*, *Stenotrophomonas maltophilia* and *Burkholderia* spp., which are rarely present in the microbiota of healthy children (Wilson, 2005; Goss and Burns, 2007). While the arrival of *P. aeruginosa* is a harbinger of more rapid disease progression in most individuals (Henry *et al*., 1992; Jacques *et al*., 1998), there remains substantial variation in the subsequent rate of pulmonary decline, with some patients dying in adolescence and others surviving into their forties and beyond.

While it is clear that several cultivable bacterial species, e.g. *S. aureus* and *P. aeruginosa*, are important CF pathogens (Henry *et al*., 1992), the relationship between the total bacterial community in this polymicrobial infection

and CF lung disease remains unexplored. General ecological theory predicts that the probability of acquiring an invasive species in a complex community, for example acquiring a potential pathogen in the polymicrobial community of the CF respiratory tract, will be influenced by current community composition and stability (Elton, 1958; Tilman, 1982; Case, 1990). It is very likely that the clinical outcome of chronic CF lung infections is greatly influenced by the many interactions among the microorganisms present and between these microorganisms and the host. Indeed, it has been demonstrated that the virulence of *P. aeruginosa* is enhanced by some avirulent constituents of healthy oropharyngeal microbiota in a rat lung infection model (Duan *et al.*, 2003). Also, antibiotics that decrease or eliminate some community members might facilitate pathogen colonization, similar to the way in which disturbances can impact other ecological communities (Brown and Gurevitch, 2004; Sekirov *et al.*, 2008; Evanno *et al.*, 2009). In addition, multiple compounding factors, including host genetics, environment and treatment variations, also appear to contribute to the rate of decline in CF (Ratjen and Doring, 2003; Mckone *et al.*, 2006; Accurso and Sontag, 2008).

Despite mounting evidence of links between the diversity of human-associated microbiota and human health (Dethlefsen *et al.*, 2007), the influence of total bacterial community composition, including uncultivable organisms, on CF lung disease has not been previously investigated. Here, we assess the influence of various factors on respiratory tract bacterial community composition in children with CF across the time frame of disease progression. These include both factors exogenous to the bacterial community, such as host's age, CFTR genotype and antibiotic therapy, and factors endogenous to the community, such as colonization by *P. aeruginosa.* Further, we explore potential interactions among these different parameters on community composition and CF clinical disease markers.

Results and discussion

To explore the relationship between respiratory tract bacterial community composition and CF lung disease in children we analysed oropharyngeal swab samples from a cohort of 45 patients ranging in age from 2 to 16 years (mean age 8.6 years). Oropharyngeal swabs are standard practice in outpatient clinics to detect CF lung pathogens by cultivation in younger children who cannot reliably provide sputum. This single sampling technique could be applied across the age range of our participants. While bronchoalveolar lavage (BAL) is a direct way to sample bacteria present in the CF lung, this invasive procedure poses the inherent risk of adverse health events and is therefore only used when clinically necessary. To ascertain how well the bacterial community isolated from an oropharyngeal swab corresponded to that from a BAL sample, we compared paired samples taken from each of three children with CF who underwent a clinically indicated BAL independent of our study. (These three children were not participants in this study and separate IRB approval was obtained for collection of these samples.) Culture-independent analysis was accomplished using a high-density 16S rRNA gene microarray, PhyloChip (Brodie *et al.*, 2007; Desantis *et al.*, 2007). We observed a concordance between the overall bacterial communities detected using the two sampling methods (Fig. S1). Thus, oropharyngeal swabs are a reasonable proxy for lung pathogens. They are a safe, non-invasive and readily available method to explore the ecology of the CF respiratory tract microbiota in children.

Cumulatively from all 45 oropharyngeal samples we identified 2051 bacterial taxa belonging to 43 phyla. As is the case for microbiota from other human body sites (Ley *et al.*, 2006; Dethlefsen *et al.*, 2007), we observed a high degree of interpersonal variation at the taxon and subfamily level. Similar to preliminary data from healthy adult throats, *Firmicutes*, *Proteobacteria* and *Bacteroidetes* were the dominant three phyla (K.P. Lemon, V. Klepac-Ceraj, H.K. Schiffer, E.L. Brodie, S.V. Lynch and R. Kolter, unpubl. results). The local mucosal environment in the CF respiratory tract differs from that of healthy individuals, in that the transepithelial potential difference due to sodium uptake is two to three times as great in patients with CF as in persons without the disease (Knowles *et al.*, 1981). Our goal was to determine if within the specific environment of the CF respiratory tract there were changes in community composition that correlated with clinical markers of disease. The bacterial diversity ranged from 206 to 1329 species per person, with the most dominant organism never accounting for more than 12% of the total population. These complex microbial communities contained a large number of bacteria commonly found in adjacent human locations, such as gut or mouth (Aas *et al.*, 2005; Eckburg *et al.*, 2005), and included many anaerobes, as well as a number of known human pathogens not previously associated with CF.

The CF oropharyngeal bacterial communities clustered into groups that correlated with the mean number of taxa per community

To identify potential correlates of community composition, we performed hierarchical clustering of the 45 pediatric CF respiratory tract bacterial communities (Fig. 1A). Three distinct groups were readily apparent in this analysis (labelled A, B and C). Communities in group A contained a significantly greater number of taxa per community, lacked *P. aeruginosa* by cultivation (Fig. 1A),

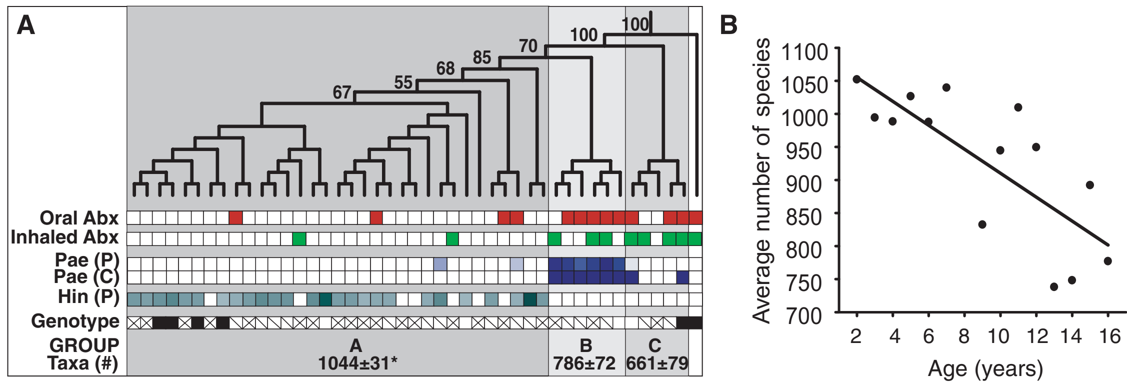

Fig. 1. A. Hierarchical clustering of the CF oropharyngeal bacterial communities was performed in MeV (Saeed *et al.*, 2006) using centered Pearson correlation with average linkage. *The mean number of taxa in group A is significantly different from groups B and C (one-way ANOVA, d.f. = 2, $F = 13.8171$, $P \leq$ 1E-4). Abbreviations are as follows: Abx is antibiotics, Pae is *P. aeruginosa*, Hin is *H. influenzae*, P is PhyloChip and C is culture. Differential shading of boxes represents relative abundance based on hybridization intensity data with darker boxes indicating greater abundance. Genotypes are denoted as follows: box with X is Δ508/Δ508, box with \ is Δ508/OC, white box is OC/OC and black box is genotype unknown. Numbers above branches are booststrap values.
B. Taxon/species richness averaged across patients of the same age plotted against age ($r^2 = 0.61$, d.f. = 12, P = 9.7E-4).

and tended to come from younger individuals (data not shown). Samples that tested positive for *P. aeruginosa* by cultivation, and were exposed to antibiotics, tended to cluster together (Fig. 1A, group B). Similarly, microbial communities from participants who were taking antibiotics, but did not have *P. aeruginosa* tended to cluster together (Fig. 1A, group C). There was very strong concordance between detection of *P. aeruginosa* by culture [Fig. 1A, Pae (C)] and by microarray [Fig. 1A, Pae (P)]. Detection of *P. aeruginosa* by cultivation is the standard in clinical practice; therefore, for the purposes of all our analyses, we considered communities to be positive for *P. aeruginosa* only if it was detected by cultivation. Based on this clustering, we hypothesized that exogenous factors such as age, CFTR genotype and antibiotic therapy might affect community composition to varying degrees. Moreover, the presence of key CF pathogen *P. aeruginosa* appeared to be a key determinant of community structure. We next sought to test the importance of these parameters individually.

An inverse correlation between age and CF oropharyngeal community richness

We divided the exogenous factors into those intrinisic to the host, age and CFTR genotype, and that extrinsic to the host, antibiotic exposure. On average, the cohort in this study still had lung function comparable to healthy individuals based on their forced expiratory volume in one second (FEV1) values (mean 97.6%; Table S1). There was a trend, which did not reach statistical significance, towards a correlation between bacterial community richness and FEV1 (data not shown). Strikingly, however, we observed an inverse correlation between age and bacterial community richness (number of taxa per community) (Fig. 1B). Not only did communities from younger children have higher absolute richness, but they also had greater phylogenetic diversity and evenness (Fig. S2). In ecology, the time hypothesis posits that multi-species communities tend to diversify over time (Ricklefs and Miller, 1999). When the opposite is observed, such as in this case, one probable explanation is that the community has experienced disturbances. For example, the frequent disturbance of logging on forests in Madagascar has resulted in decreased diversity (Brown and Gurevitch, 2004). Likewise, the decline in CF respiratory tract bacterial community complexity with age might be explained by frequent disturbances via antibiotics and/or by *P. aeruginosa* invasion.

CF oropharyngeal community composition was related to CFTR genotype

The composition of a multi-species community can also be influenced by the immediate physical environment (Ricklefs and Miller, 1999). Unlike age, CFTR genotype does not vary over time. A number of mutations have been identified in the CFTR gene encoding the chloride channel that is altered in individuals with CF. Certain CFTR genotypes are associated with an increased probability of morbidity and mortality at a younger age (McKone *et al.*, 2006), yet the effect of CFTR genotype on bacterial community composition has not been previously examined. The most common CFTR genotype of CF patients is homozygous ΔF508/ΔF508, and ΔF508 is categorized clinically as one of the 'severe' mutations (McKone *et al.*, 2006). To compare community composition based on genotype, we classified all participants into three catego-

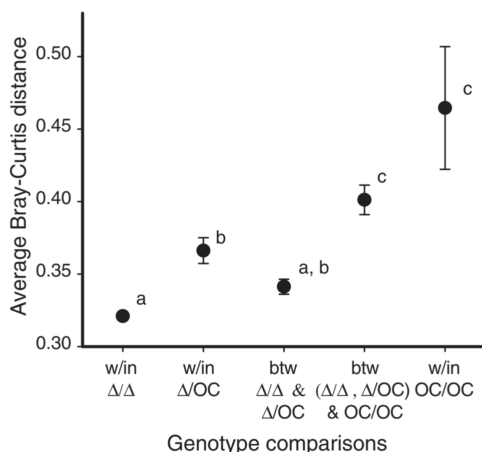

Fig. 2. Comparison of Bray-Curtis distances of bacterial communities based on CFTR genotype. A higher Bray-Curtis number indicates that communities are more distant/diverse from each other: (w/in Δ/Δ) within ΔF508/ΔF508; (w/in Δ/OC) within ΔF508/other CFTR mutation (OC); (btw Δ/Δ & Δ/OC) between ΔF508/ΔF508 and ΔF508/OC; (btw (Δ/Δ, Δ/OC) & OC/OC) between the presence of at least a single Δ508 allele (Δ508/Δ508 and Δ508/OC) and OC/OC; (w/in OC/OC) within OC/OC group (mean ± s.e.m.). The letters in the figures indicate significant differences among means, as determined by randomization tests.

ries: ΔF508/ΔF508, ΔF508/other CFTR mutation (OC) and OC/OC (Fig. 1A and Table S1). These three categories were relatively evenly distributed across the groupings derived from cluster analysis of bacterial community composition shown in Fig. 1A and there was no statistical difference in age, antibiotic exposure or presence of *P. aeruginosa* between these three CFTR genotype groups (data not shown). We then measured the variation in taxa diversity between samples (beta diversity). Pair-wise comparison of bacterial community composition within and between these three genotypic categories revealed that communities from participants who carry at least one ΔF508 allele (Δ/Δ and Δ/OC in Fig. 2) are more similar to each other than to those from participants without this allele (OC/OC in Fig. 2). This suggests the surprising possibility that the ΔF508 allele might impact the CF respiratory tract environment in a manner distinct from other CFTR mutant alleles and deserves further study within a much larger cohort.

Bacterial communities exposed to both inhaled and oral antibiotics had significantly fewer taxa than communities receiving either one of the two or no antibiotics

We next explored the influence of antibiotics on bacterial community composition. Six of the 45 participants were receiving long-term antibiotics with a combination of oral agent(s) and an inhaled antibiotic, in this case tobramycin (TOBI; Fig. 3A). In addition, four participants were receiving only a long-term inhaled antibiotic (TOBI or amikacin) and seven only long-term oral antibiotics. While there were several oral agents used, for the purposes of this ecological analysis, we considered any oral antibiotic as an equivalent disturbance. The long-term oral antibiotics included three times per week azithromycin as both an antibacterial and anti-inflammatory agent (Saiman *et al.*, 2003; Florescu *et al.*, 2009). Although inhaled aminoglycoside antibiotics are optimized for high dose drug delivery to the lower airways, our data reveal a clear impact on oropharyngeal bacteria with the loss of taxa likely to be susceptible (light green highlighting in Fig. S3). Only communities exposed to both long-term oral and inhaled antibiotics had a significant decrease in taxonomic richness (alpha diversity; Fig. 3A). In contrast, bacterial communi-

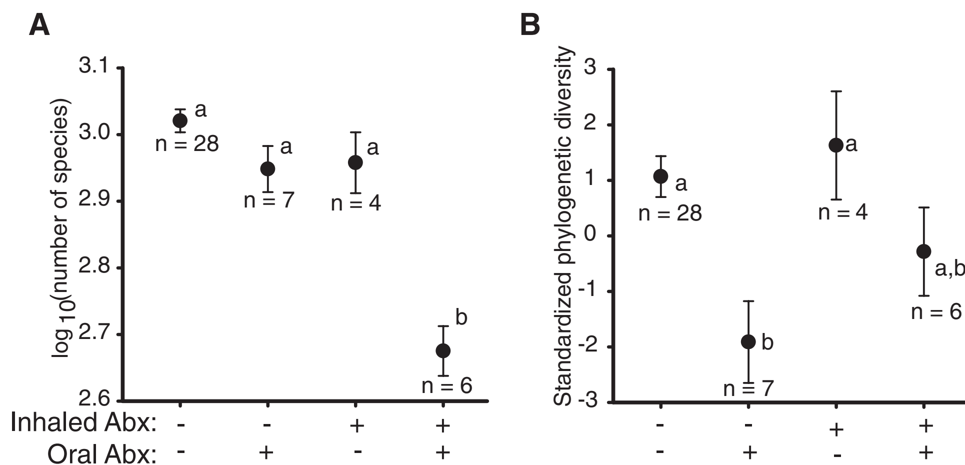

Fig. 3. A. Comparison of the richness of bacterial communities exposed to both inhaled and oral antibiotics (one-way ANOVA; d.f.1 = 3, d.f.2 = 41, $F = 23.52$, $P \leq 0.0001$, mean ± s.e.m.).
B. Comparison of standardized phylogenetic diversity of communities exposed to oral antibiotics and inhaled antibiotics.
The letters in the figures indicate significant differences among means, determined by the least square mean test with Bonferroni adjustment for multiple comparisons.

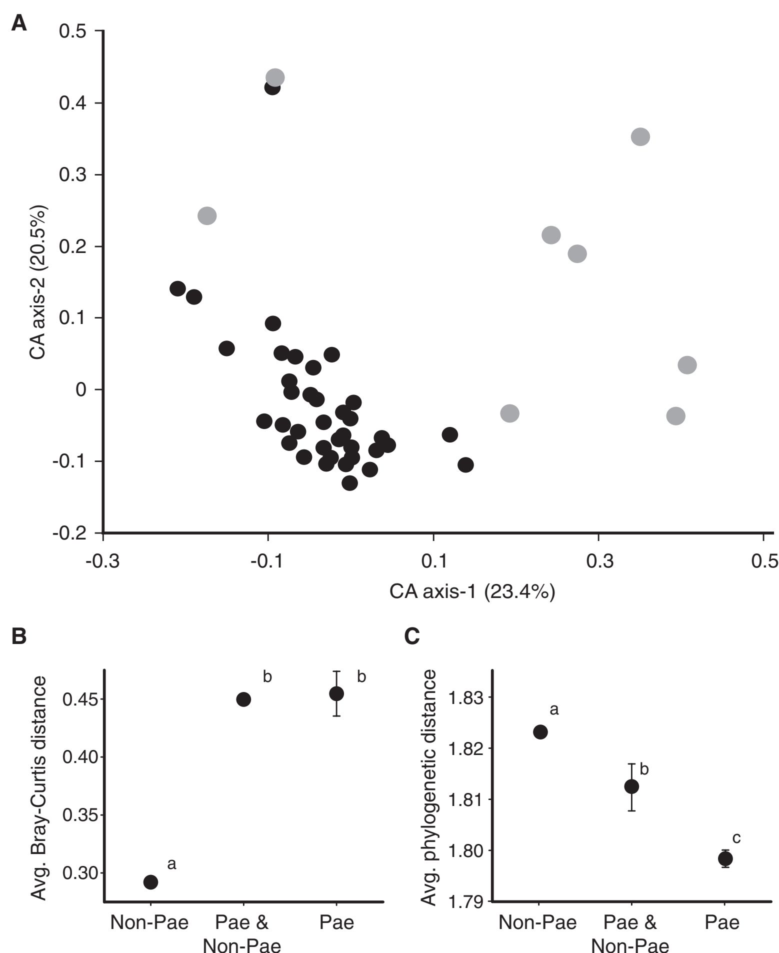

Fig. 4. A. Plot of correspondence analysis (CA) of the oropharyngeal bacterial community (PhyloChip data) relatedness from 45 children. Each circle represents one community from one child. Communities without *P. aeruginosa* are labelled with black circles and communities with *P. aeruginosa* are labelled with grey circles. The percentages of variation described by the correspondence analysis coordinates are shown in parentheses.
B and C. (B) Comparison of average Bray-Curtis distances and (C) of standardized phylogenetic distances among bacterial communities without *P. aeruginosa* (non-Pae), among communities with and without *P. aeruginosa* (Pae & Non-Pae) and among communities with *P. aeruginosa* (Pae). Higher values indicate more communities that were more diverse from each other (mean ± s.e.m.). The letters in the figures indicate significant differences among means, as determined by randomization tests (i.e. a is not different from a, but is from b).

ties from patients exposed to long-term oral antibiotics, without an inhaled antibiotic, had significantly reduced phylogenetic diversity, whereas those exposed to either both or inhaled antibiotic alone did not (Fig. 3B). Bacterial taxa that were significantly reduced (*t*-test with correction for multiple comparisons, *P*- and *q*-value < 0.05, > 10-fold change) in communities exposed to an inhaled aminoglycoside antibiotic (TOBI or amikacin) were primarily from the phylum *Proteobacteria* (Fig. S3). In contrast, communities from patients treated with oral antibiotics had significantly reduced number of taxa (*t*-test, *P*- and *q*-value < 0.05, > 10-fold change) from a broader spectrum of the bacterial phyla (Fig. S3).

CF oropharyngeal communities containing P. aeruginosa *did not group with other communities nor with each other*

In addition to exogenous factors, in this case host age, CFTR genotype and antibiotic exposure, endogenous factors, such as the presence of a specific bacterium, or group of bacteria, might impact CF respiratory tract bacterial community composition. This is true in other multi-species communities where the presence or arrival of an individual species can strongly affect community structure, which in turn may impact the ecological processes and the entire ecosystem (Sanders *et al.*, 2003). To identify major sources of variation in the dataset, we used correspondence analysis, a technique for identifying potential relations between variables when there are no *a priori* expectations as to the nature of those relations. This revealed that most bacterial communities grouped together (black dots in Fig. 4A), except those from individuals with *P. aeruginosa* (grey dots in Fig. 4A). Perhaps more importantly, the *P. aeruginosa*-containing communities did not cluster with each other, implying that the shift in community composition with *P. aeruginosa* might be either idiosyncratic or due to selection for a community metagenome that can be achieved by different sets of taxa. This is supported by our observation that the variation in taxa diversity between samples (beta diversity) was smallest when comparing communities without *P. aeruginosa*, and largest when comparing communities with *P. aeruginosa* (Fig. 4B). Additionally, the breadth of phyloge-

netic diversity was largest within communities lacking *P. aeruginosa* and smallest within communities containing *P. aeruginosa* (Fig. 4C). The apparent contrast between the observed taxonomic divergence (Fig. 2B) and phylogenetic convergence (Fig. 4C) in communities with *P. aeruginosa* likely reflects losses at deep phylogenetic branches in communities with, as compared with without, *P. aeruginosa* (Fig. S3). Various perturbations might account for such an observation, including differences in antibiotic exposure and in *P. aeruginosa*-produced small molecule exposure. Thus, the presence of *P. aeruginosa* was associated with significant differences in overall bacterial community composition, suggesting that communities diverge from each other upon the arrival of *P. aeruginosa*. This might be related to a number of factors that can vary between communities, including preceding community composition, bacterial interference, *P. aeruginosa* strain variation, variations in host environment and/or differences in antibiotic therapy.

CF oropharyngeal communities exposed to antibiotics and containing P. aeruginosa *exhibited decreased richness and phylogenetic diversity*

All eight participants with *P. aeruginosa* were also receiving some form of long-term antibiotic therapy. This is likely to be a common occurrence as the presence of *P. aeruginosa* is one clinical indication for initiating long-term antibiotics. These two factors may interact with each other in influencing community composition and it is thus difficult to disentangle the effect of this key CF pathogen from that of antibiotics. The combination of the presence of *P. aeruginosa* and long-term antibiotic treatment significantly reduced taxonomic richness within samples (alpha diversity; Fig. 5A). This pattern existed at the taxa/species, family, order and class taxonomic levels ($P < 0.01$), as determined by Unifrac analysis (Lozupone *et al.*, 2006). A similar decrease in species richness has been observed in intubated patients with ventilator-associated pneumonia (VAP) when both *P. aeruginosa* and antibiotics were present, and it was postulated that an antibiotic-induced decline in diversity inadvertently permits *P. aeruginosa* to flourish (Flanagan *et al.*, 2007). In addition to decreased richness, we observed that the presence of both *P. aeruginosa* and long-term antibiotic treatment also significantly reduced phylogenetic diversity within each community (Fig. 5B). Furthermore, pair-wise community comparison revealed that bacterial communities with *P. aeruginosa* were phylogenetically more distant from each other than were those lacking this key CF pathogen (beta diversity; Fig. 4B).

CF oropharyngeal communities exposed to antibiotics, but not containing P. aeruginosa*, exhibited decreased richness*

Nine of the 45 participants were receiving antibiotics but did not have *P. aeruginosa*, and their communities had significantly fewer taxa than those from participants who were not receiving antibiotics (Fig. 5A; one-way ANOVA, d.f.1 = 2, d.f.2 = 42, $F = 10.92$, $P = 0.0002$). Antibiotic treatment of communities lacking *P. aeruginosa* seems to have had uniform effects throughout the phylogenetic tree, as the observed decline in taxon number per com-

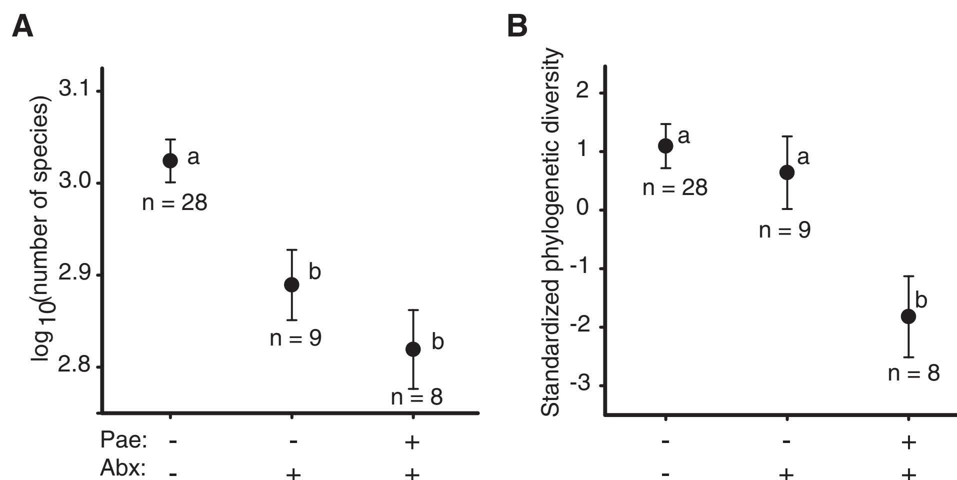

Fig. 5. A. Comparison of the richness of bacterial communities exposed to antibiotics and/or containing *P. aeruginosa* (one-way ANOVA; d.f.1 = 2, d.f.2 = 42, $F = 10.92$, $P = 0.0002$, mean ± s.e.m.).
B. Comparison of standardized phylogenetic diversity of communities exposed to antibiotics and/or containing *P. aeruginosa* (one-way ANOVA; d.f.1 = 2, d.f.2 = 42, $F = 6.87$, $P = 0.0026$, mean ± s.e.m.).
The letters in the figures indicate significant differences among means, as determined by randomization tests (i.e. a is not different from a, but is from b).

munity was relatively evenly distributed across the phylogenetic tree (Fig. S3, Fig. 5B).

An inverse relationship between the presence of P. aeruginosa *and members of the family* Pasteurellaceae

To further explore how the CF respiratory tract bacterial communities were altered in the presence of *P. aeruginosa*, we examined the co-occurrence of particular taxa with *P. aeruginosa* (Fig. S4 and Table S2). Communities with *P. aeruginosa* lacked the family *Pasteurellaceae* (Figs S2 and S4), which includes the early CF pathogen *H. influenzae.* In these communities, the abundance of several taxa from the family *Pasteurellaceae* was reduced over 500-fold based on estimates of 16S rRNA gene copy number derived from hybridization intensity data (*t*-test, *t*-score = −7.083, *P*-value < 1.80E-05, *q*-value < 2.61E-07). The inverse relationship between *P. aeruginosa* and *H. influenzae* was also observed in the culture data from each participant (Fig. 1A). Of note, *H. influenzae* was absent in the communities of all participants receiving an inhaled aminoglycoside independent of *P. aeruginosa* presence or absence (Fig. 1A). Since all of the individuals with *P. aeruginosa* were also receiving antibiotics, it is unclear whether the disappearance of *H. influenzae* is caused by repeated antibiotic exposure, by the arrival of *P. aeruginosa* or a combination of both (Tables S2–S4). Distinguishing between these possibilities will require a long-term longitudinal study following a cohort of children starting at a young age, prior to both antibiotic exposure and the arrival of *P. aeruginosa*.

Concluding remarks

Using a sensitive culture-independent method that permits relative quantification, we have demonstrated that the CF respiratory tract bacterial community composition is correlated with several clinical disease markers, including the presence of the principal CF pulmonary pathogen *P. aeruginosa*, increased antibiotic exposure and CFTR genotype. We also observed an inverse correlation between age and the following community characteristics: bacterial taxon richness, evenness and phylogenetic diversity within communities. This relationship might reflect cumulative antibiotic exposure, which is likely to increase with age, and the effects of successive CF pulmonary pathogens over time. The correlation between community assembly and CFTR genotype was independent of age, antibiotics and *P. aeruginosa*. This observation suggests the possiblity that the presence of the ΔF508 allele may uniquely alter the CF respiratory tract environment. We believe this possiblity deserves further exploration within a larger cohort.

Multiple factors influence the progression of CF lung disease and many of these are related to the key CF pathogen *P. aeruginosa.* Culture-dependent studies have established a role for pathogen succession with the acquisition of *P. aeruginosa* being a sentinel event (Wahab *et al*., 2004), for variations in *P. aeruginosa* strains (Head and Yu, 2004) and, in an animal model of infection, for increased virulence of *P. aeruginosa* in the presence of some members of the oropharyngeal microbiota (Duan *et al*., 2003). The results presented here are the first demonstration that in CF patients the respiratory tract bacterial community composition shifts in the presence of *P. aeruginosa.* Importantly, this divergence of *P. aeruginosa*-containing communities is not solely due to variability in number of taxa, but is also due to differences in phylogenetic composition. We hypothesize that these differences in bacterial community composition account for some of the variation in the rate of clinical deterioration among children with CF and *P. aeruginosa*. To establish potential cause and effect with regard to community composition will require a longitudinal study of CF patients that are enrolled prior to their acquiring *P. aeruginosa*. Understanding how microbial communities change over time in CF patients will provide new tools for monitoring and assessing the underlying causes of lung function decline. Current clinical practices for treating CF patient lung infections have resulted in significant increases in life expectancy (FitzSimmons, 1993). They also create microbial communities that are continually responding to massive disturbances from airway clearance treatments and antibiotic exposure. In this cross-sectional analysis, it was not possible to fully disentangle the effect of antibiotics and *P. aeruginosa*. In other ecosystems, such as forests experiencing insect outbreaks, management to maintain favourable community structure in the face of disturbances often requires an adaptive approach, where ecosystem monitoring and flexible management plans allow conservation professionals to respond to changing conditions (Walters, 1986; Oglethorpe, 2002). Likewise, accurate and comprehensive monitoring of microbial diversity in CF patients will allow clinical professionals to design flexible treatment protocols that can be adapted to the needs of individual patients over the course of their lives.

Experimental procedures

Enrolment of subjects and sample collection

Both the Committee on Clinical Investigation at Children's Hospital Boston and Harvard Medical School's Institutional Review Board approved the protocol for this study. Informed consent was obtained from all participants. Participants were recruited during previously scheduled outpatient visits to the CF clinic at Children's Hospital Boston. Inclusion criteria were age from 2 through 16 and a diagnosis of CF. The posterior

wall of each participant's oropharynx was swabbed with a sterile double-tipped swab (BBL CultureSwab, Becton, Dickinson and Company, MD) during a routine previously scheduled clinic visit. One tip was sent to a clinical microbiology laboratory for culture analysis and the other was used for molecular analysis. Oropharyngeal swabs are standard practice in outpatient clinics to detect CF lung pathogens in younger children who cannot reliably provide sputum (Rosenfeld *et al.*, 1999). This single technique could be applied across the age range of our participants, 2–16 years. While BAL represents a direct way to sample bacteria present in the CF lung, this invasive procedure poses an inherent risk of adverse health events and is therefore only used when clinically necessary. Swabs are unlikely to be a perfect proxy for lung pathogens, but they are a safe, non-invasive, and readily available method to explore the ecology of the CF respiratory tract microbiota in children. Oropharyngeal swabs were stored at −80°C within minutes of collection. All samples were coded to protect patient identity and clinical data were obtained from medical record databases. Clinical data included information on each participant's age, lung function, CFTR genotype, pulmonary exacerbation, administered antibiotics, and colonization by microorganisms identified by standard clinical laboratory cultivation techniques.

DNA extraction, amplification and PhyloChip hybridization

Total genomic DNA was extracted from swabs using a bead-beating protocol (Lysis Matrix B, MP Bio) combined with the AllPrep Qiagen RNA/DNA Isolation Kit (Qiagen). 16S rRNA genes were PCR amplified using universal bacterial primers 27F and 1492R (Lane, 1991). Eight replicate reactions across a temperature gradient (48–58°C) were performed for each sample to minimize potential PCR amplification bias (Polz and Cavanaugh, 1998). The pooled amplicon of each sample (250 ng) was spiked with internal standards to permit normalization of inter-array variation (Brodie *et al.*, 2006). This mix was fragmented, biotin labelled and hybridized to the PhyloChip (Affymetrix, CA) as previously described (Brodie *et al.*, 2006). The PhyloChip can identify up to 8434 distinct bacterial taxa/species from a sample without the need for cultivation (Brodie *et al.*, 2007) and has been shown to identify microbial taxa that constitute as little as 0.01% of a polymicrobial population (Desantis *et al.*, 2007). Using this method, we sampled bacterial communities collected by swabbing the oropharynx of 45 children with CF. Data analysis was performed using custom software to scale each array such that the average intensity of the probe sets responding to the internal controls equalled 2500 fluorescent intensity units. The positive fraction (PosFrac) was calculated for each probe set as the number of positive probe pairs divided by the total number of probe pairs in a probe set. A taxon was considered 'present' when its PosFrac was 0.9 or greater. On the PhyloChip, a taxon is defined as a cluster of 16S rRNA gene sequences with $\leq$ 3% divergence (Desantis *et al.*, 2007).

Analyses of microbial community profiles

To determine bacterial community structural relatedness across patient samples, hierarchical clustering using centred Pearson correlation with average linkage and bootstrap analysis were performed in MeV_4.02 on the $\log_2$-transformed hybridization data for each sample with a filter cut-off of 0.02% (Saeed *et al.*, 2006). (Similar results were obtained with a variety of other clustering methods.) Correspondence analysis of sample hybridization profiles was performed in MeV. All taxa detected by the PhyloChip cumulatively from all of the samples were used to construct a phylogenetic tree based on the ARB parsimony tree delivered with the Greengenes ARB database (October 2006 release) (Desantis *et al.*, 2006). Unweighted and weighted UniFrac analyses were performed on the resulting tree (Lozupone *et al.*, 2006). To determine if the UniFrac distances were on average significantly different for groups (*P. aeruginosa* culture positive or negative, oral antibiotics positive or negative, inhaled antibiotic positive or negative), we performed a *t*-test on the UniFrac distance matrix. The dendogram in Fig. 1 and the circular tree in Fig. S1 were drawn using the online software package iTOL (Letunic and Bork, 2007).

Phylogenetic diversity measurements

Phylogenetic diversity within individual patients was measured as the mean pair-wise phylogenetic distance (MPD) separating randomly drawn pairs of microbial taxa occurring in that patient (Webb *et al.*, 2008). Because MPD is correlated with taxa richness, we standardized observed MPD values for each patient by comparing them with the MPD expected for randomly assembled communities of the same number of taxa drawn from the pool of all microbial taxa observed across all samples. Standardized phylogenetic diversity was calculated as observed MPD minus the mean MPD across 999 random communities, divided by the standard deviation of the MPD across those random communities. Positive values of standardized phylogenetic diversity thus indicate that the microbial community contained taxa that were more distantly related than expected by chance given the taxa richness of that community, while negative values indicate the opposite. Phylogenetic diversity among samples, or phylobetadiversity, was measured in analogous fashion, as the mean pair-wise phylogenetic distance separating a randomly drawn taxon from each of the two samples. Phylogenetic diversity was calculated using the Picante package (Kembel *et al.*, 2009) for the R statistical programming environment (R Development Core Team, 2008).

Statistical analyses

Two-tailed *t*-tests were used to identify statistically significant differences between two groups (*P. aeruginosa* present/absent, oral antibiotics or TOBI present/absent) of significant taxa. These were performed on log-transformed fluorescence values using MATLAB 7.1. To account for the multiple *t*-tests performed, we calculated a measure of statistical significance called *q*-value, which measures significance in terms of the false discovery rate (Storey and Tibshirani, 2003). ANOVA analysis and the Bonferroni correction were used to correct for multiple hypothesis and were performed in SAS.

Acknowledgements

We thank the children, and their families, who participated in this study. We thank the Children's Hospital Boston Division of Respiratory Diseases research coordinators for patient enrolment and data collection. We thank I. Ceraj and M.L. Coleman for help with data analysis, comments and discussion. We thank M.F. Polz, J.R. Thompson, D.K. Newman, H.C. Vlamakis, R.J. Case, C.N. Peterson, P. Klepac and E.A. Gontang for critical reading of the manuscript. We thank members of the Kolter lab for thoughtful discussions and Todd DeSantis, Yvette Piceno and Gary Andersen for their assistance with the PhyloChip. The Cystic Fibrosis Foundation funded V.K.-C.

References

Aas, J.A., Paster, B.J., Stokes, L.N., Olsen, I., and Dewhirst, F.E. (2005) Defining the normal bacterial flora of the oral cavity. *J Clin Microbiol* **43:** 5721–5732.

Accurso, F.J., and Sontag, M.K. (2008) Gene modifiers in cystic fibrosis. *J Clin Invest* **118:** 839–841.

Brodie, E.L., Desantis, T.Z., Joyner, D.C., Baek, S.M., Larsen, J.T., Andersen, G.L., *et al.* (2006) Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. *Appl Environ Microbiol* **72:** 6288–6298.

Brodie, E.L., DeSantis, T.Z., Parker, J.P., Zubietta, I.X., Piceno, Y.M., and Andersen, G.L. (2007) Urban aerosols harbor diverse and dynamic bacterial populations. *Proc Natl Acad Sci USA* **104:** 299–304.

Brown, K.A., and Gurevitch, J. (2004) Long-term impacts of logging on forest diversity in Madagascar. *Proc Natl Acad Sci USA* **101:** 6045–6049.

Case, T.J. (1990) Invasion resistance arises in strongly interacting species-rich model competition communities. *Proc Natl Acad Sci USA* **87:** 9610–9614.

Chmiel, J.F., and Davis, P.B. (2003) State of the art: why do the lungs of patients with cystic fibrosis become infected and why can't they clear the infection? *Respir Res* **4:** 8.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., *et al.* (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72:** 5069–5072.

DeSantis, T.Z., Brodie, E.L., Moberg, J.P., Zubieta, I.X., Piceno, Y.M., and Andersen, G.L. (2007) High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. *Microb Ecol* **53:** 371–383.

Dethlefsen, L., McFall-Ngai, M., and Relman, D.A. (2007) An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* **449:** 811–818.

Duan, K.M., Dammel, C., Stein, J., Rabin, H., and Surette, M.G. (2003) Modulation of *Pseudomonas aeruginosa* gene expression by host microflora through interspecies communication. *Mol Microbiol* **50:** 1477–1491.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., *et al.* (2005) Diversity of the human intestinal microbial flora. *Science* **308:** 1635–1638.

Elton, C.S. (1958) *The Ecology of Invasions by Animals and Plants*. London, UK: Methuen.

Evanno, G., Castella, E., Antoine, C., Paillat, G., and Goudet, J. (2009) Parallel changes in genetic diversity and species diversity following a natural disturbance. *Mol Ecol* **18:** 1137–1144.

FitzSimmons, S.C. (1993) The changing epidemiology of cystic fibrosis. *J Pediatr* **122:** 1–9.

Flanagan, J.L., Brodie, E.L., Weng, L., Lynch, S.V., Garcia, O., Brown, R., *et al.* (2007) Loss of bacterial diversity during antibiotic treatment of intubated patients colonized with *Pseudomonas aeruginosa*. *J Clin Microbiol* **45:** 1954–1962.

Florescu, D.F., Murphy, P.J., and Kalil, A.C. (2009) Effects of prolonged use of azithromycin in patients with cystic fibrosis: a meta-analysis. *Pulm Pharmacol Ther* **22:** 467–472.

Gibson, R.L., Burns, J.L., and Ramsey, B.W. (2003) Pathophysiology and management of pulmonary infections in cystic fibrosis. *Am J Respir Crit Care Med* **168:** 918–951.

Goss, C.H., and Burns, J.L. (2007) Exacerbations in cystic fibrosis • 1: epidemiology and pathogenesis. *Thorax* **62:** 360–367.

Harris, J.K., De Groote, M.A., Sagel, S.D., Zemanick, E.T., Kapsner, R., Penvari, C., *et al.* (2007) Molecular identification of bacteria in bronchoalveolar lavage fluid from children with cystic fibrosis. *Proc Natl Acad Sci USA* **104:** 20529–20533.

Harrison, F. (2007) Microbial ecology of the cystic fibrosis lung. *Microbiology* **153:** 917–923.

Head, N.E., and Yu, H.W. (2004) Cross-sectional analysis of clinical and environmental isolates of *Pseudomonas aeruginosa*: biofilm formation, virulence, and genome diversity. *Infect Immun* **72:** 133–144.

Henry, R.L., Mellis, C.M., and Petrovic, L. (1992) Mucoid *Pseudomonas aeruginosa* is a marker of poor survival in cystic fibrosis. *Pediatr Pulmonol* **12:** 158–161.

Jacques, I., Derelle, J., Weber, M., and Vidailhet, M. (1998) Pulmonary evolution of cystic fibrosis patients colonized by *Pseudomonas aeruginosa* and/or *Burkholderia cepacia*. *Eur J Pediatr* **157:** 427–431.

Kembel, S.W., Ackerly, D.D., Blomberg, S., Cowan, P.D., Helmus, M., Morlon, H., and Webb, C.O. (2009) Picante: R tools for integrating phylogenies and ecology. *R package version 0.6-0*.

Knowles, M., Gatzy, J., and Boucher, R. (1981) Increased bioelectric potential difference across respiratory epithelia in cystic fibrosis. *N Engl J Med* **305:** 1489–1495.

Lane, D.J. (1991) 16S/23S rRNA sequencing. In *Nucleic Acid Techniques in Bacterial Systematics*. Stackebrandt, E., and Goodfellow, M. (eds). New York, USA: John Wiley and Sons, pp. 115–148.

Letunic, I., and Bork, P. (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23:** 127–128.

Ley, R.E., Peterson, D.A., and Gordon, J.I. (2006) Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124:** 837–848.

Lozupone, C., Hamady, M., and Knight, R. (2006) UniFrac – an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* **7:** 371.

McKone, E.F., Goss, C.H., and Aitken, M.L. (2006) CFTR genotype as a predictor of prognosis in cystic fibrosis. *Chest* **130:** 1441–1447.

Oglethorpe, J.A.E. (ed.) (2002) *Adaptive Management: From Theory to Practice*. Gland, Switzerland and Cambridge, UK: IUCN, The World Conservation Union.

Polz, M.F., and Cavanaugh, C.M. (1998) Bias in template-to-product ratios in multitemplate PCR. *Appl Environ Microbiol* **64:** 3724–3730.

R Development Core Team (2008) R: a language and environment for statistical computing. Vienna, Austria.

Ratjen, F. (2006) Restoring airway surface liquid in cystic fibrosis. *N Engl J Med* **354:** 291–293.

Ratjen, F., and Doring, G. (2003) Cystic fibrosis. *Lancet* **361:** 681–689.

Ricklefs, R.E., and Miller, G.L. (1999) *Ecology*. New York, NY, USA: W.H. Freeman.

Rogers, G.B., Hart, C.A., Mason, J.R., Hughes, M., Walshaw, M.J., and Bruce, K.D. (2003) Bacterial diversity in cases of lung infection in cystic fibrosis patients: 16S ribosomal DNA (rDNA) length heterogeneity PCR and 16S rDNA terminal restriction fragment length polymorphism profiling. *J Clin Microbiol* **41:** 3548–3558.

Rogers, G.B., Carroll, M.P., Serisier, D.J., Hockey, P.M., Jones, G., and Bruce, K.D. (2004) Characterization of bacterial community diversity in cystic fibrosis lung infections by use of 16S ribosomal DNA terminal restriction fragment length polymorphism profiling. *J Clin Microbiol* **42:** 5176–5183.

Rosenfeld, M., Emerson, J., Accurso, F., Armstrong, D., Castile, R., Grimwood, K., *et al.* (1999) Diagnostic accuracy of oropharyngeal cultures in infants and young children with cystic fibrosis. *Pediatr Pulmonol* **28:** 321–328.

Rubin, B.K. (2009) Mucus, phlegm, and sputum in cystic fibrosis. *Respir Care* **54:** 726–732. discussion 732.

Saeed, A.I., Bhagabati, N.K., Braisted, J.C., Liang, W., Sharov, V., Howe, E.A., *et al.* (2006) TM4 microarray software suite. *Methods Enzymol* **411:** 134–193.

Saiman, L., Marshall, B.C., Mayer-Hamblett, N., Burns, J.L., Quittner, A.L., Cibene, D.A., *et al.* (2003) Azithromycin in patients with cystic fibrosis chronically infected with *Pseudomonas aeruginosa*: a randomized controlled trial. *JAMA* **290:** 1749–1756.

Sanders, N.J., Gotelli, N.J., Heller, N.E., and Gordon, D.M. (2003) Community disassembly by an invasive species. *Proc Natl Acad Sci USA* **100:** 2474–2477.

Sekirov, I., Tam, N.M., Jogova, M., Robertson, M.L., Li, Y., Lupp, C., and Finlay, B.B. (2008) Antibiotic-induced perturbations of the intestinal microbiota alter host susceptibility to enteric infection. *Infect Immun* **76:** 4726–4736.

Sibley, C.D., Parkins, M.D., Rabin, H.R., Duan, K., Norgaard, J.C., and Surette, M.G. (2008) A polymicrobial perspective of pulmonary infections exposes an enigmatic pathogen in cystic fibrosis patients. *Proc Natl Acad Sci USA* **105:** 15070–15075.

Storey, J.D., and Tibshirani, R. (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* **100:** 9440–9445.

Tilman, D. (1982) Resource competition and community structure. *Monogr Popul Biol* **17:** 1–296.

Tunney, M.M., Field, T.R., Moriarty, T.F., Patrick, S., Doering, G., Muhlebach, M.S., *et al.* (2008) Detection of anaerobic bacteria in high numbers in sputum from patients with cystic fibrosis. *Am J Respir Crit Care Med* **177:** 995–1001.

Wahab, A.A., Janahi, I.A., Marafia, M.M., and El-Shafie, S. (2004) Microbiological identification in cystic fibrosis patients with CFTR I1234V mutation. *J Trop Pediatr* **50:** 229–233.

Walters, C.J. (1986) *Adaptive Management of Renewable Resources*. New York, USA: MacMillan.

Webb, C.O., Ackerly, D.D., and Kembel, S.W. (2008) Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* **24:** 2098–2100.

Wilson, M. (2005) *Microbial Inhabitants of Humans: Their Ecology and Role in Health and Disease*. Cambridge, UK: Cambridge University Press.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Correspondence analysis (CA) plot: relatedness of BAL and oropharyngeal bacterial communities identified by the PhyloChip (pairs indicated by the same colour; square BAL and circle oropharyngeal sample). Black circles represent all oropharyngeal bacterial community from 45 children with CF. This analysis reveals that lower CF respiratory tract (sampled using BAL) and upper respiratory tract bacterial communities (sampled using oropharyngeal swab) from the same patient, sampled at the same time, generally group together. Each circle represents one community from one child. The percentages of variation described by the correspondence analysis coordinates are shown in parentheses.

Fig. S2. A. Diversity as measured by Shannon index (*H*) averaged across participants of the same age plotted against age.

B. Evenness [Shannon index (*H*)/natural log(species richness)] averaged across participants of the same age plotted against age.

C. Phylogenetic diversity within each community (alpha diversity) averaged across participants of the same age plotted against age. Here phylogenetic diversity within individual communities was measured as the mean pair-wise phylogenetic distance (MPD) separating randomly drawn pairs of microbial taxa occurring in that community (*8*).

Fig. S3. A phylogenetic tree of cumulative bacterial taxa from all samples whose abundance changed significantly based on *P. aeruginosa* or antibiotic presence/absence was constructed in ARB using parsimony. Coloured boxes represent the following, with darker hue indicating an increase and lighter a decrease: blue for *P. aeruginosa*, red for oral antibiotics and green for inhaled antibiotics. Only taxa/species that changed by > 10-fold (*t*-test with *q*-value < 0.01) between the groups were plotted.

Fig. S4. Average frequency of bacterial taxa/species in communities that were negative (top) and positive for *P. aeruginosa* by culture (bottom). Peaks with average frequency greater than 0.05% were: (1) *Synergistes* unknown (AF229792.1), (2) *Fusobacterium* unknown (AY274839.1), (3) unknown Chloroflexi bacterium (AJ306793), (4) *Cyanobacterium sp.* (X82156.1), (5) uncultured *Chromatiales*

(AJ518784.1), (6) *Pasteurellaceae* group (AB088998.1), (7) *Helicobacter* group, (8) *Micrococcae* group, (9), (10) and (11) *Clostridiales*, (12) *Streptococcus* group (AF009482.1), (13) *Clostridium sp.* unclassified, (14) *Rihanellaceae*, (15) *Prevotellaceae* group (AF371893.1), (16) *Xanthomonas* (AB110496), (17) *Pseudomonadeaceae* (AE004501.1), (18) *Pseudoalteromonas ruthenica* (AF316891.1), (19) *Staphylococcae* group, (20) *Streptococcus constellatus* (AF104676.1) and *Streptococcus gordonii* (AF003931.1). An average frequency of 0 indicates there was no change in taxon presence with and without *P. aeruginosa*.

Table S1. Overview of enrolled study subjects by age, CF genotype and FEV1.

Table S2. *T*-test for bacterial taxa/species that varied significantly (q-value < 0.01 and > 10-fold change) between communities that had culture-proven *P. aeruginosa* and those that were culture-negative for *P. aeruginosa*. Taxa ordered by fold change starting with greatest.

Table S3. *T*-test for bacterial taxa/species that varied significantly (q-value < 0.01 and > 10-fold change) between communities currently exposed to oral antibiotics and those that were not. Taxa ordered by fold change starting with greatest.

Table S4. *T*-test for bacterial taxa/species that varied significantly (q-value < 0.01 and > 10-fold change) between communities currently exposed to inhaled tobramycin (TOBI) and those that were not. Taxa ordered by fold change starting with greatest.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.