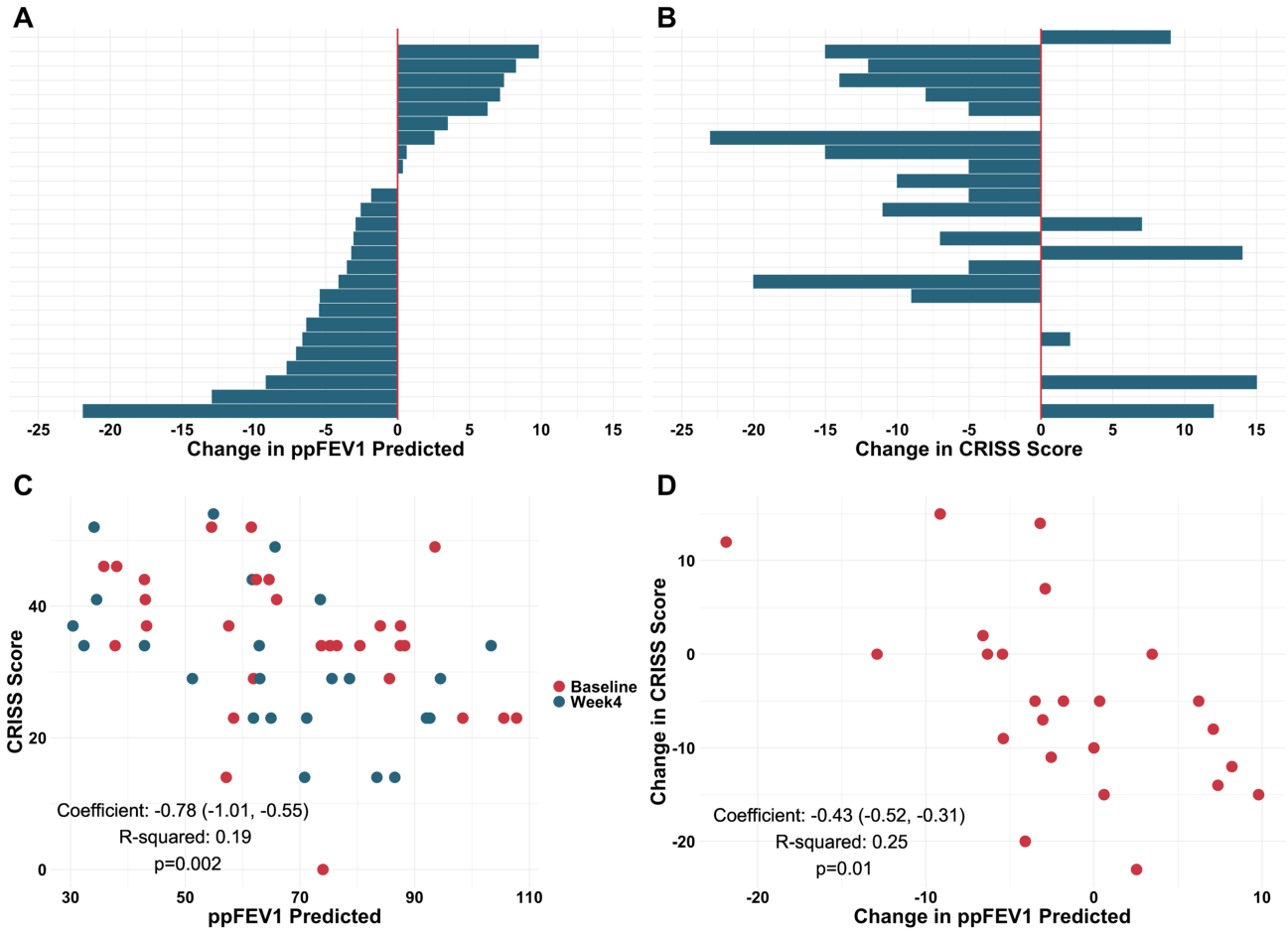

Figure S1. Therapeutic response to one month of maintenance tobramycin therapy. $ppFEV_1$ and symptom scores (CRISS) were collected from all participants at baseline and after one month of inhaled tobramycin. Change in $ppFEV_1$ (A) and CRISS score (B) after one month of therapy is presented for each subject. Each horizontally aligned bar represents the same individual in both A and B. Positive numbers in (A) indicate improvements in lung function, while negative numbers in (B) represent improvements in symptoms. (C) Scatter plot comparing $ppFEV_1$ and CRISS score for each individual at baseline and at Week 4. (D) Scatter plot of changes in CRISS score and $ppFEV_1$. R^2 and p values result from linear regression analysis, controlling for week on therapy in 1C. As a comparison, a common estimate of repeatability of standard spirometry is ~6 $ppFEV_1$ [1], and pulmonary exacerbations frequently involve a reduction of ≥10 $ppFEV_1$ [2], while a change in CRISS score of 16 has been found to indicate a CF respiratory exacerbation.

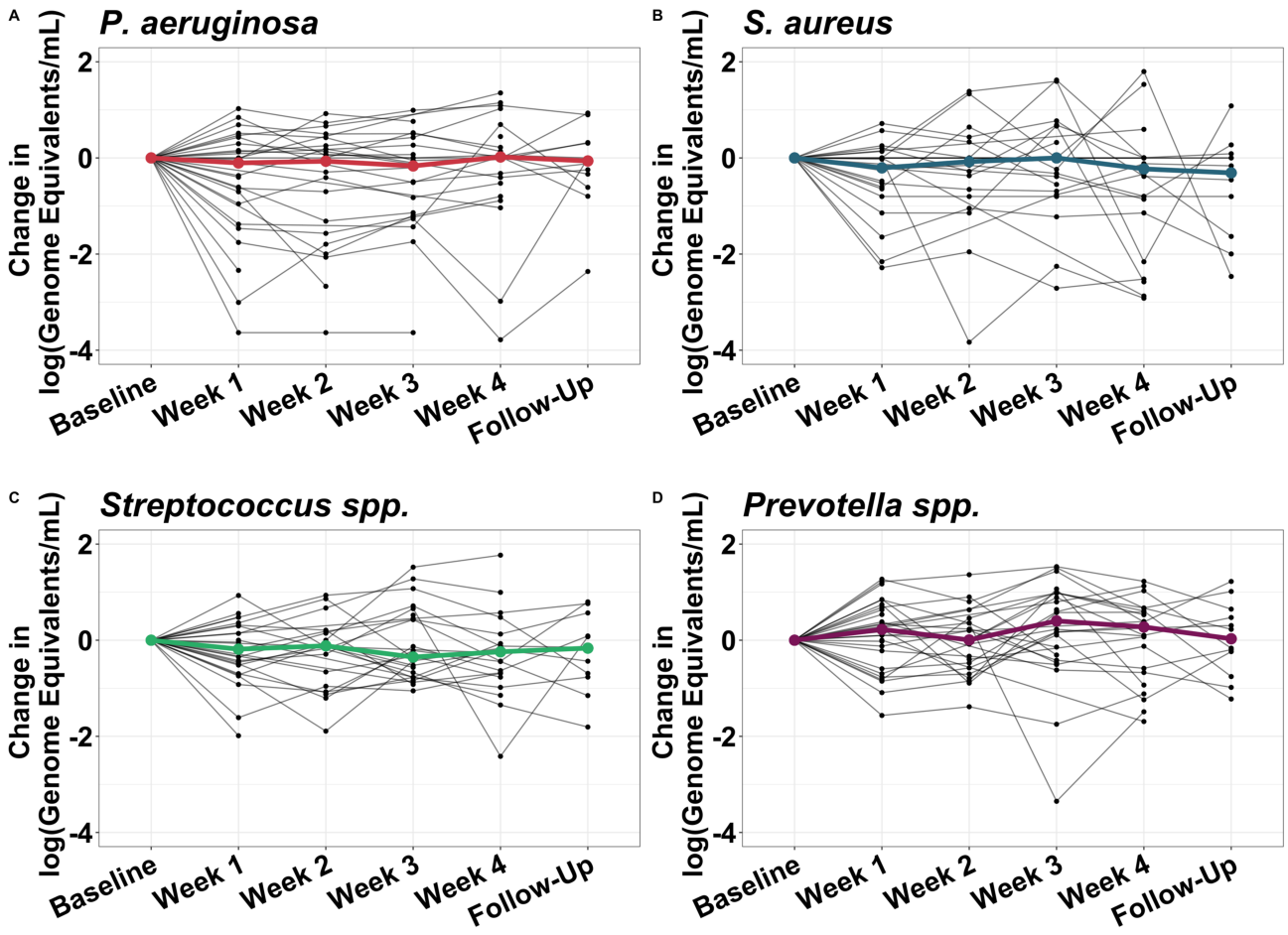

Figure S2. Change in absolute viable load of select taxa by quantitative PCR. Change in absolute load of P. aeruginosa (A), S. aureus (B), Streptococcus spp. (C) and Prevotella spp. from baseline by week on therapy. Each black line represents a single individual and the colored lines represent the median at each time point. Loads were calculated using a quadruplex quantitative PCR assay and were performed after removal of extracellular DNA from dead cells.

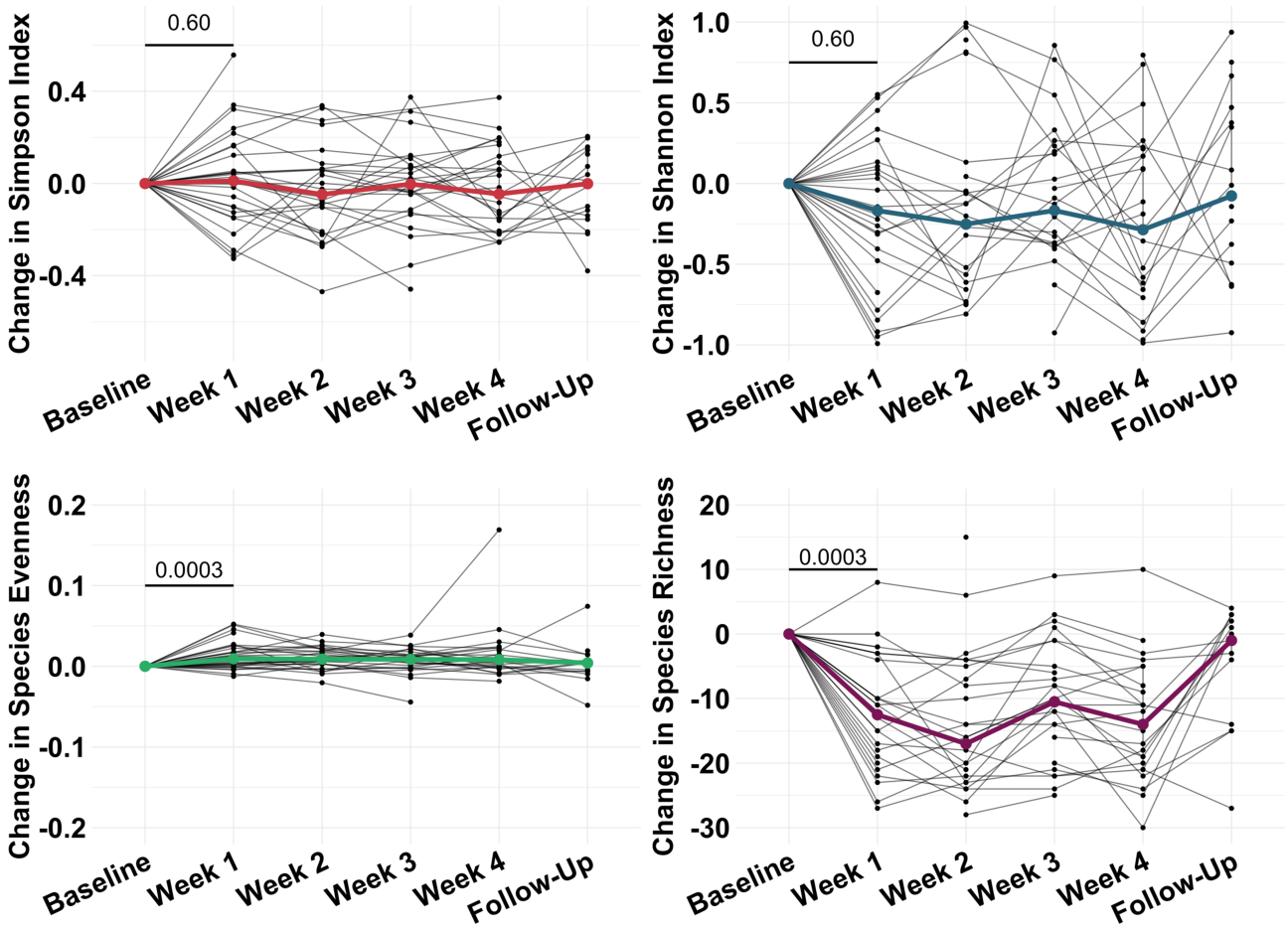

Figure S3. Diversity metrics of all samples by week on therapy. Change in Simpson index (A), Shannon index (B), evenness (C) and Richness (D) by week on therapy. Each black line represents a single individual and colored lines represents the median value. All values were calculated at the species level. Wilcoxon signed rank test was used to assess difference between groups.

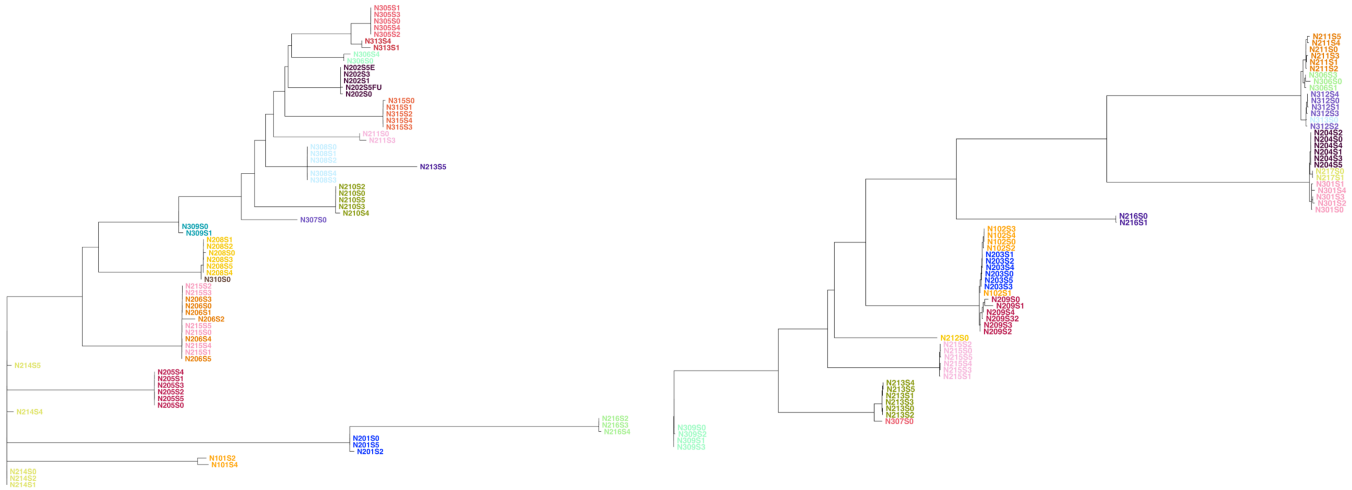

Figure S4. Dominant strains vary between individuals but not with antibiotic therapy. Phylogenetic trees of the dominant strain of P. aeruginosa (left) and S. aureus (right) in each individual using StrainPhlAn2. Samples collected from the same individual are similarly colored.

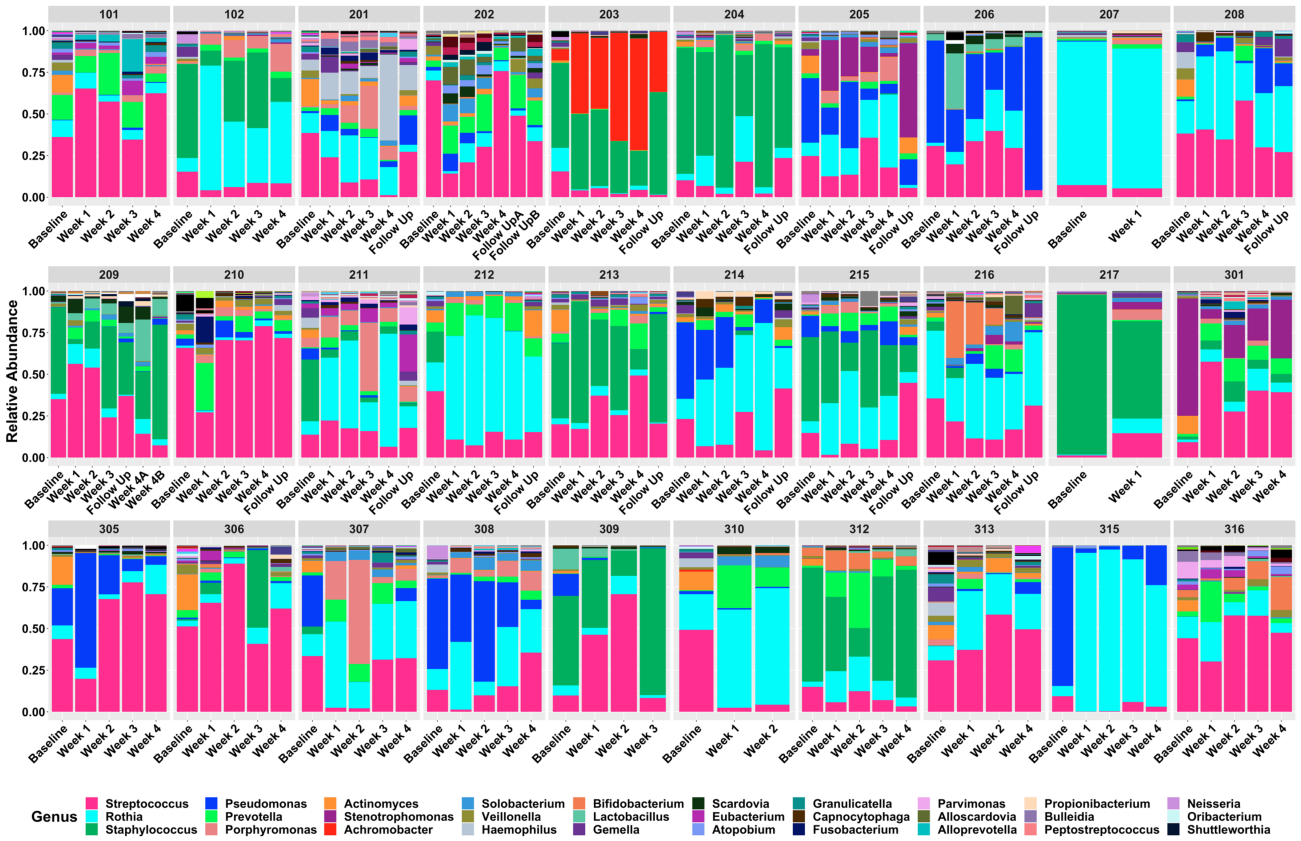

Figure S5. Taxonomic profile of all samples at the genus level. All samples were subjected to metagenomic shotgun sequencing followed and taxonomy was determined using MetaPhlAn2. Reads from the same genus were summed before plotting profiles for all samples, separated by individual. Each color represents a different species. Two subjects (202 and 209) provided an extra sample at specific weeks, which are designated “A” and “B” here.

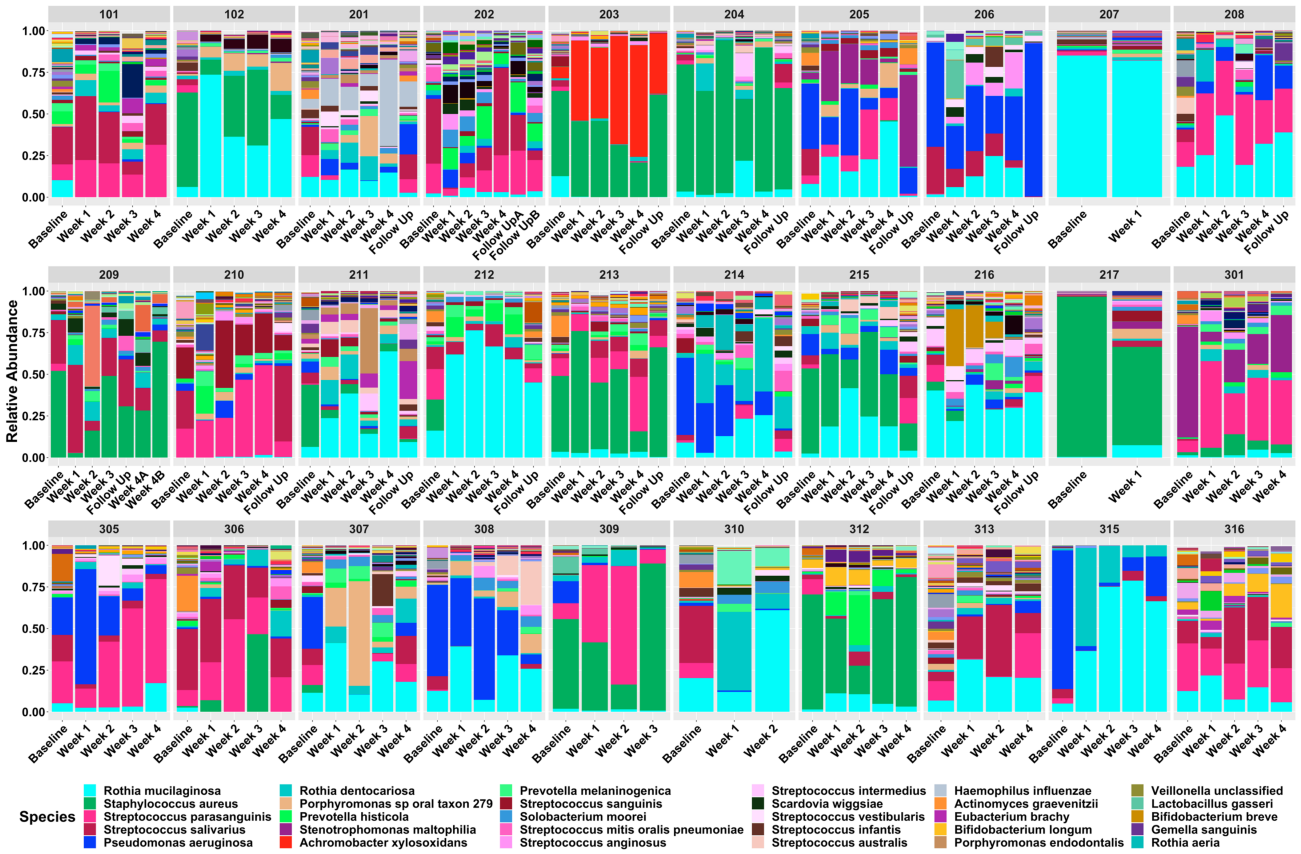

Figure S6. Taxonomic profile of all samples at the species level. All samples were subjected to metagenomic shotgun sequencing followed and taxonomy was determined using MetaPhlAn2. Reads from the same species were summed before plotting profiles for all samples, separated by individual. Each color represents a different species. Two subjects (202 and 209) provided an extra sample at specific weeks, which are designated “A” and “B” here. Species names for most abundant 15 taxa are presented for ease of viewing.

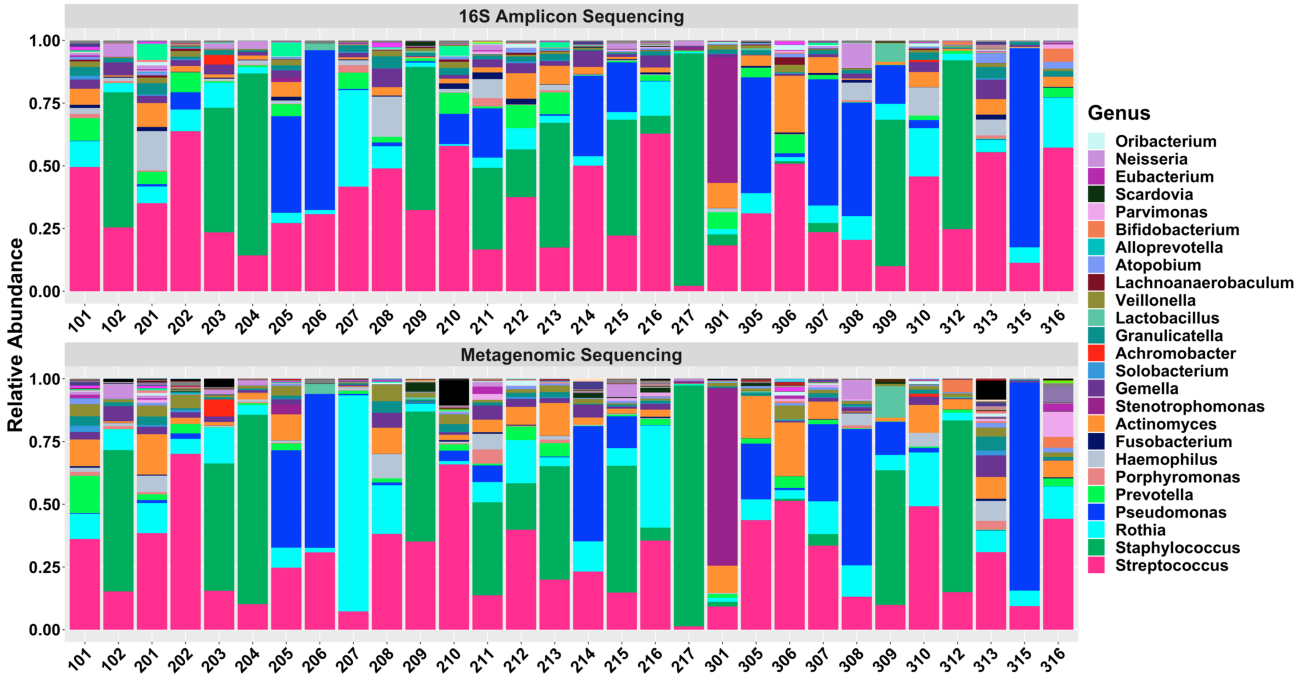

Figure S7. Comparison of taxonomic profiles analyzed via metagenomic shotgun sequencing vs. 16S amplicon sequencing. Taxonomic profiles of all 30 baseline samples at the genus level, determined using metagenomic shotgun sequencing (top) or 16S amplicon sequencing (bottom). Vertically aligned bars represent the same sample. Note that 16S amplicon sequencing does not reliably identify taxa at the species level, so only genera are displayed.

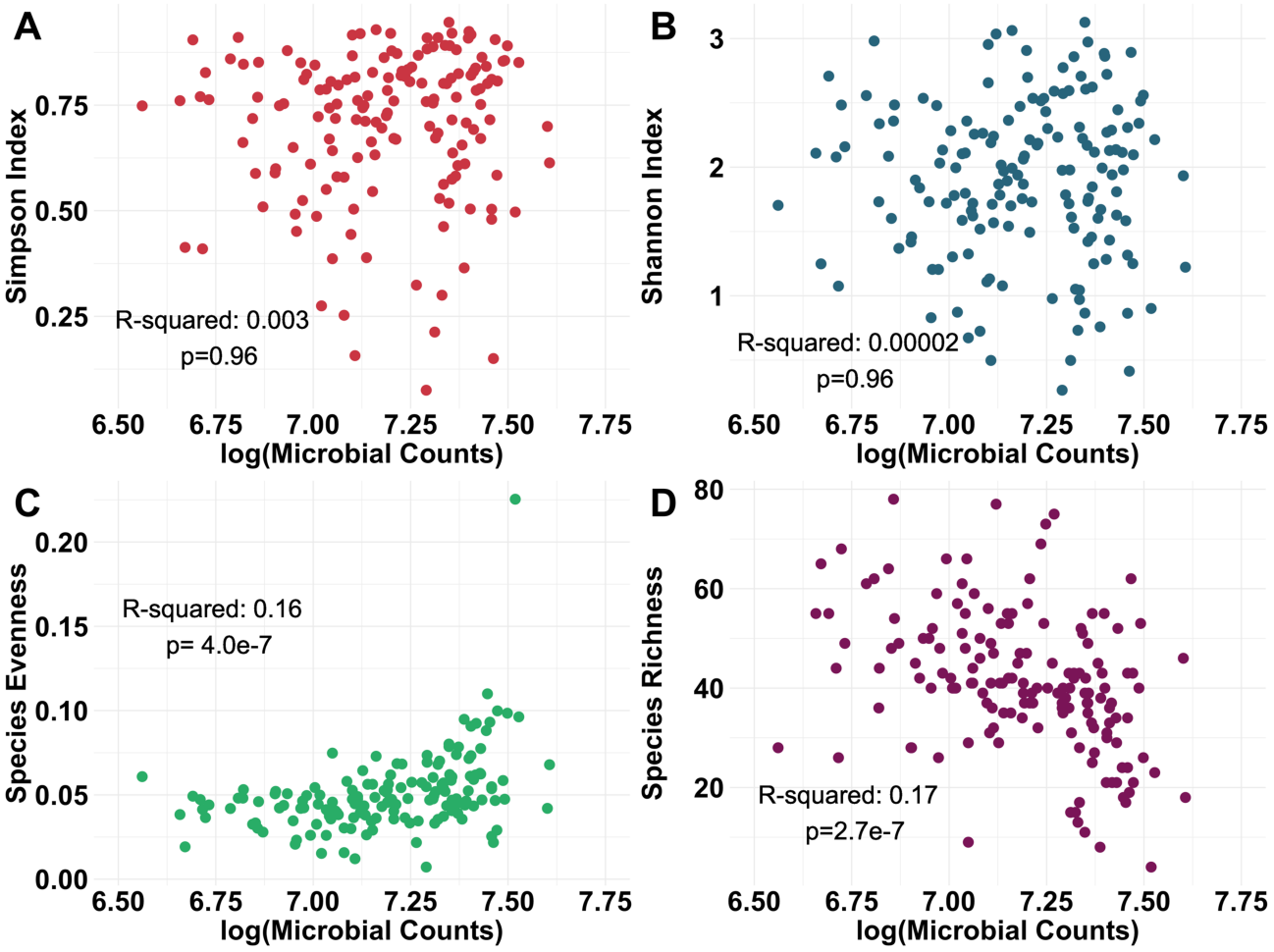

Figure S8. The effect of microbial sequencing depth on community diversity. Dot plots comparing log microbial reads and Simpson index (A), Shannon index (B), evenness (C) and richness (D). Note that microbial reads, rather than total reads, are presented as percent human DNA differed between samples. R^2 and p values result from linear regression analysis. A Benjamini-Hochberg correction performed to account for multiple comparisons.

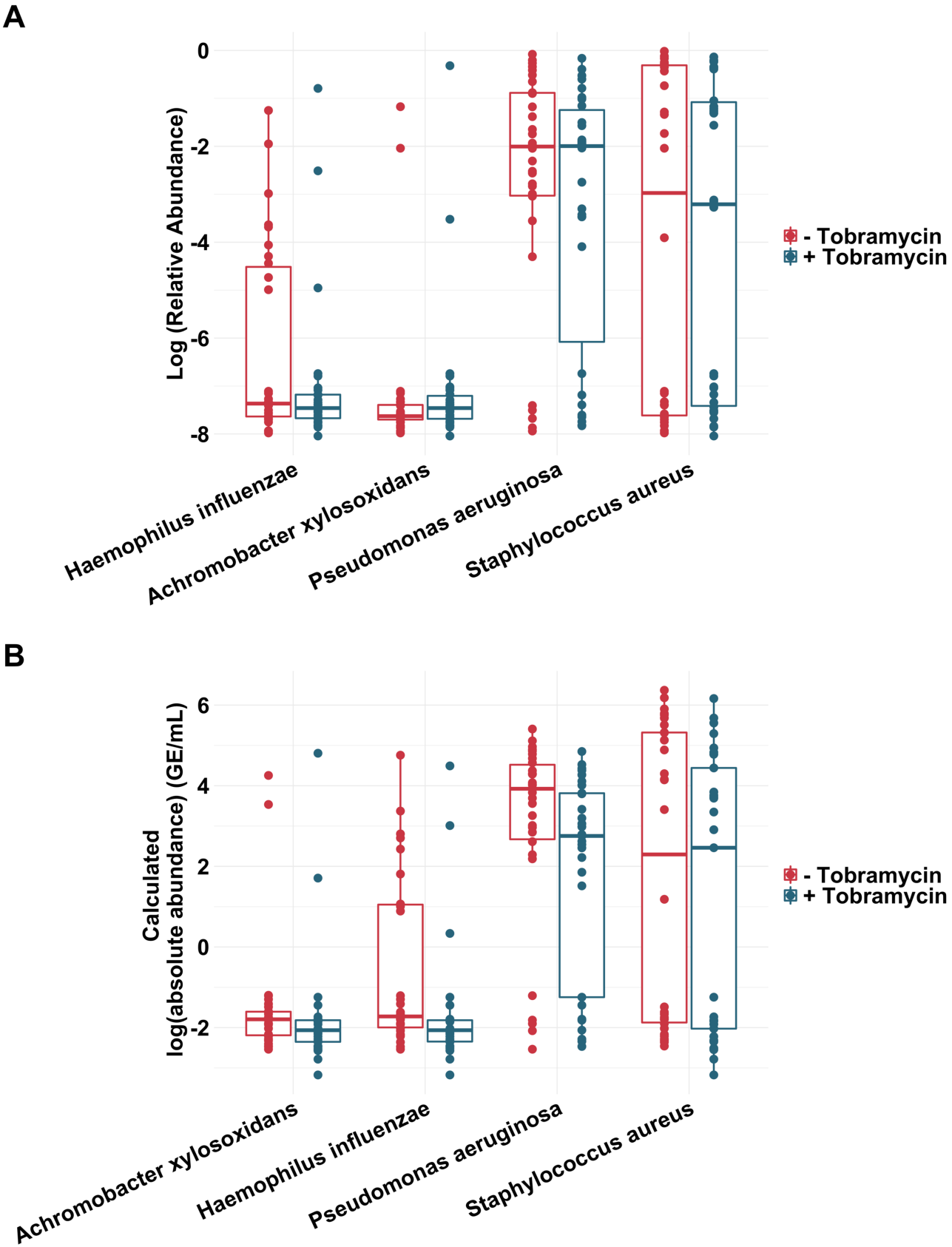
A
Log (Relative Abundance)
0
-2
-4
-6
-8
- Tobramycin
+ Tobramycin
Haemophilus influenzae
Achromobacter xylosoxidans
Pseudomonas aeruginosa
Staphylococcus aureus
B
Calculated log(absolute abundance) (GE/mL)
6
4
2
0
-2
- Tobramycin
+ Tobramycin
Achromobacter xylosoxidans
Haemophilus influenzae
Pseudomonas aeruginosa
Staphylococcus aureus

Figure S9. Change in relative abundance of classic CF pathogens after one week of tobramycin therapy. A) Log-relative abundances and B) calculated absolute abundances of classic CF pathogens after one week of tobramycin therapy. Absolute abundances were calculated by multiplying relative abundances via MetaPhlAn2 by total bacterial loads as determined via universal 16S qPCR. Compare to Fig. 3c-d. Boxes represent interquartile region and middle represents the median.

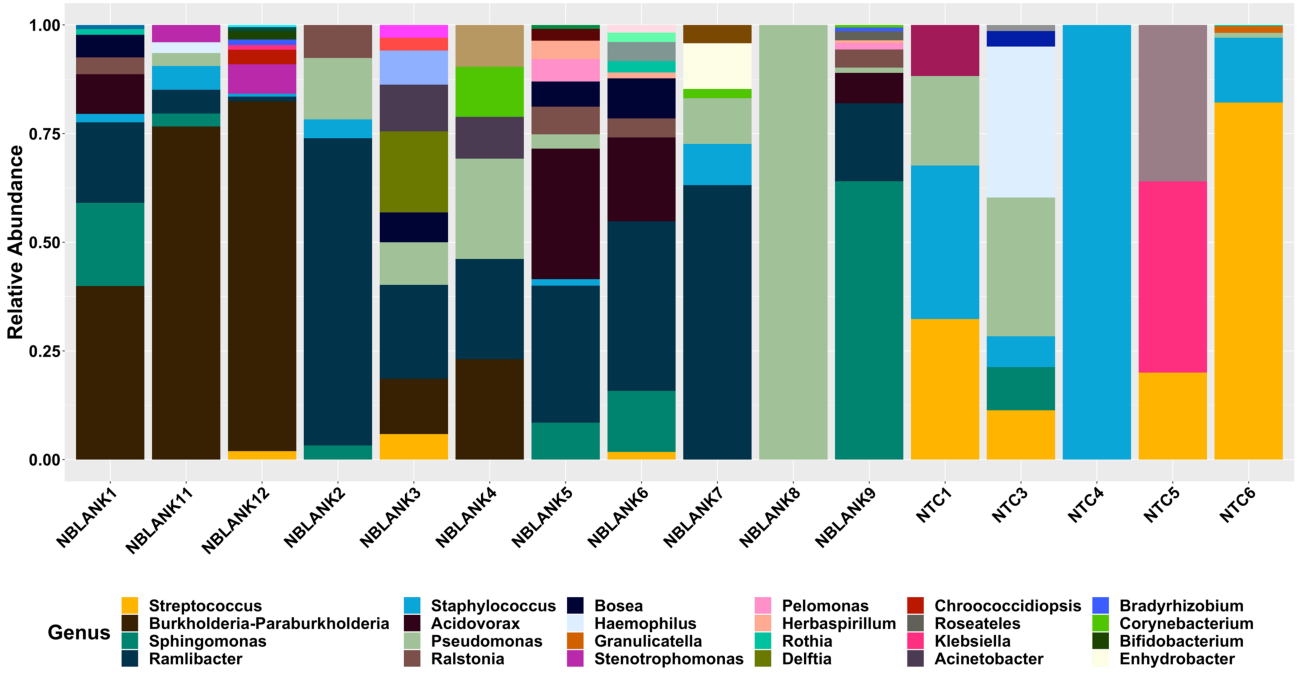

Figure S10. Phylogenetic profile of extraction blanks and 16S amplicon PCR no-template-controls. Extraction blanks (NBLANK) and no template controls (NTC) produced an average of 526 reads and 126 reads on average. NTC2 did not produce any reads and thus is not displayed.

Category	Characteristic	Number of subjects (%)*
Study site	Seattle Children's Hospital	2 (7)
	University of Washington	17 (56)
	University of Michigan	11 (37)
Mean age in years (range)		35 (9-75)
Gender		
	Female	16 (53)
	Male	14 (47)
Genotype		
	Homozygous F508del	12 (40)
	Heterozygous F508del	16 (53)
	Other	2 (7)
Clinical characteristics		
	Mean Baseline ppFEV$_1$ (range)	69.1 (35.8 – 107.8)

	Mean Baseline CRISS Score (range)	34 (0 – 52) 90
Medical History		
	Impaired glucose tolerance[#]	9 (30)
	Liver dysfunction[$]	2 (7)
	Allergic bronchopulmonary aspergillosis	1 (3)
	Gastroesophageal reflux	3 (10)
	Asthma	2 (7)
Culture History (one year prior to baseline)		
	Pseudomonas aeruginosa	30 (100)
	Staphylococcus aureus, MSSA	15 (50)
	Staphylococcus aureus, MRSA	2 (7)
	Burkholderia cepacia complex	1 (3)
	Haemophilus influenzae	1 (3)
	Stenotrophomonas maltophila	5 (17)
	Achromobacter xylosoxidans	1 (3)
	Nontuberculous Mycobacterium	2 (7)
	Candida spp.	4 (13)
	Aspergillus spp.	10 (33)
Other medications		
	Azithromycin	19 (64)
	Inhaled DNAse	23 (77)
	Hypertonic saline	13 (43)
	Bronchodilators	28 (93)
	CFTR modulators	8 (27)

1 Includes history of abnormal liver function tests
2 * Unless otherwise noted in the “Characteristic” column
3 # Includes any history of impaired glucose tolerance, including CF-related diabetes.
4
5 **Table S1. Subject demographic table.** Aggregate data for select clinical characteristics for all study
6 participants.
7
8 **File S1. Low abundance taxa.** Those taxa less than 1% relative abundance in all samples, pooled into the
9 "other" category for both metagenomic sequencing (left column) and 16S amplicon sequencing (right column)
0

MetaPhlAn2	**16S Amplicon Sequencing**
Achromobacter unclassified	*Acidovorax*
Acinetobacter pittii calcoaceticus nosocomialis	*Acinetobacter*
Actinobacillus unclassified	*Actinobaculum*
Actinobaculum unclassified	*Actinomycetospora*
Actinomyces georgiae	*Aerococcus*
Actinomyces oris	*Afipia*
Actinomyces sp oral taxon 171	*Aggregatibacter*
Actinomyces sp oral taxon 448	*Alcanivorax*

Actinomyces sp ph3
Actinomyces timonensis
Actinomyces viscosus
Afipia unclassified
Aggregatibacter aphrophilus
Aggregatibacter unclassified
Alloprevotella rava
Alloprevotella unclassified
Anaerococcus lactolyticus
Anaerococcus obesiensis
Anaerococcus sp PH9
Anaeroglobus geminatus
Aspergillus terreus
Bacteroides fragilis
Bilophila unclassified
Bilophila wadsworthia
Brevibacterium unclassified
Burkholderia cenocepacia
Burkholderia lata
Burkholderia multivorans
Burkholderia unclassified
C2likevirus unclassified
Campylobacter concisus
Campylobacter gracilis
Campylobacter rectus
Campylobacter showae
Candida unclassified
candidate division TM7 single cell isolate TM7b
candidate division TM7 single cell isolate TM7c
Candidatus Prevotella conceptionensis
Capnocytophaga granulosa
Capnocytophaga ochracea
Capnocytophaga sp oral taxon 338
Cardiobacterium hominis
Cardiobacterium valvarum
Catonella morbi
Clavispora lusitaniae
Collinsella unclassified
Comamonas unclassified
Corynebacterium accolens
Corynebacterium matruchotii
Cryptobacterium curtum
Deinococcus unclassified
Anaeroglobus
Anoxybacillus
Arsenicicoccus
Aspergillus
Bergeyella
Bilophila
Blautia
Bordetella
Bosea
Bradyrhizobium
Brevundimonas
Brochothrix
Burkholderia-Paraburkholderia
Butyrivibrio
Caenimonas
Campylobacter
Candidatus_Saccharimonas
Cardiobacterium
Catonella
Centipeda
Chroococcidiopsis
Chryseobacterium
Cloacibacterium
Comamonas
Conchiformibius
Cryptobacterium
Defluviitaleaceae_UCG-011
Delftia
Dietzia
Distigma
Dolosigranulum
Eikenella
Enhydrobacter
Enterococcus
Erysipelotrichaceae_UCG-006
Escherichia/Shigella
Ezakiella
Fastidiosipila
Filifactor
Flavobacterium
Fretibacterium
Gardnerella
Halomonas

Delftia unclassified
Dialister micraerophilus
Dietzia unclassified
Dolosigranulum pigrum
Eggerthia catenaformis
Eikenella corrodens
Enterococcus faecalis
Eremothecium unclassified
Escherichia coli
Escherichia unclassified
Eubacteriaceae bacterium CM5
Eubacteriaceae noname unclassified
Eubacterium infirmum
Eubacterium saphenum
Eubacterium yurii
Facklamia unclassified
Finegoldia magna
Fretibacterium fastidiosum
Fusobacterium periodonticum
Gardnerella vaginalis
Gemella unclassified
Haemophilus aegyptius
Haemophilus haemolyticus
Haemophilus parahaemolyticus
Haemophilus paraphrohaemolyticus
Haemophilus sputorum
Herbaspirillum unclassified
Human herpesvirus 4
Human herpesvirus 7
Johnsonella ignava
Jonquetella anthropi
Jonquetella unclassified
KI polyomavirus
Kingella denitrificans
Kingella oralis
Kingella unclassified
Lachnospiraceae oral taxon 107
Lactobacillus crispatus
Lactobacillus phage J 1
Lactobacillus phage Lc Nu
Lactobacillus vaginalis
Lautropia mirabilis
Leptotrichia buccalis
Helcococcus
Herbaspirillum
hgcI_clade
Howardella
Hydrogenophaga
Hymenobacter
Janthinobacterium
Johnsonella
Jonquetella
Kingella
Klebsiella
Lactococcus
Lautropia
Lawsonella
Leuconostoc
Mannheimia
Megasphaera
Mesorhizobium
Methylobacillus
Methylobacterium
Microvirga
Mitochondrion
Mobiluncus
Mogibacterium
Moraxella
Mycobacterium
Olsenella
Ornithobacterium
Paracoccus
Parascardovia
Pelomonas
Peptococcus
Photobacterium
Phyllobacterium
Prevotellaceae_YAB2003_group
Pseudopropionibacterium
Pseudoramibacter
Psychrobacter
Ralstonia
Ramlibacter
Rheinheimera
Rhizobium
Rhodococcus

Leptotrichia goodfellowii
Leptotrichia hofstadii
Leptotrichia shahii
Leptotrichia unclassified
Leptotrichia wadei
Lymphocryptovirus unclassified
Meyerozyma guilliermondii
Mitsuokella unclassified
Mycobacterium avium
Naumovozyma unclassified
Neisseria elongata
Neisseria macacae
Neisseria meningitidis
Neisseria subflava
Olsenella unclassified
Oribacterium sp ACB1
Oribacterium sp oral taxon 078
Parvimonas sp oral taxon 110
Peptoniphilus harei
Peptoniphilus sp oral taxon 836
Porcine type C oncovirus
Porphyromonas asaccharolytica
Porphyromonas catoniae
Porphyromonas uenonis
Prevotella baroniae
Prevotella bivia
Prevotella buccae
Prevotella dentalis
Prevotella intermedia
Prevotella loescheii
Prevotella marshii
Prevotella micans
Prevotella saccharolytica
Prevotella sp oral taxon 473
Propionibacterium acnes
Propionibacterium phage P100D
Propionibacterium phage P101A
Propionibacterium phage PAD20
Pseudomonas geniculata
Pseudomonas phage F116
Pseudomonas unclassified
Pseudoramibacter alactolyticus
Pusillimonas unclassified
Rikenellaceae_RC9_gut_group
Roseateles
Ruminococcaceae_UCG-014
Selenomonas
Senegalimassilia
Shinella
Slackia
Sneathia
Sphingomonas
Sphingopyxis
Streptobacillus
Suttonella
Tannerella
Thermus
Thioalkalispira
Ureaplasma
Weissella
Xanthomonas
Yersinia

Ralstonia pickettii
Riemerella unclassified
Roseolovirus unclassified
Ruminococcus torques
Saccharomyces cerevisiae
Scardovia unclassified
Selenomonas artemidis
Selenomonas flueggei
Selenomonas noxia
Selenomonas sputigena
Serratia marcescens
Slackia exigua
Slackia unclassified
Staphylococcus hominis
Staphylococcus phage PVL
Staphylococcus warneri
Streptococcus agalactiae
Streptococcus cristatus
Streptococcus massiliensis
Streptococcus oligofermentans
Streptococcus phage Cp 1
Streptococcus pseudopneumoniae
Streptococcus sp F0442
Subdoligranulum unclassified
Tannerella forsythia
Thiomonas unclassified
Torque teno virus
Treponema denticola
Treponema maltophilum
Treponema medium
Treponema socranskii
Treponema vincentii
Ureaplasma unclassified
Ureaplasma urealyticum
Veillonella dispar
Veillonella sp 3 1 44
Veillonella sp 6 1 27
Veillonella sp oral taxon 780

File S2. Code used to determine differentially abundant taxa and functional categories

```
#load required packages
```

```
library(stringr)
library(ALDEx2)
library(ggpubr)
library(vegan)
library(compositions)
library(randomForest)
library(zCompositions)
library(rfUtilities)
library(MASS)
library(caret)
set.seed(100)

#############################################################################
#Aldex2 Calculations-ID's taxa/KOs significantly different between groups#
#Here I'm using the comparison of samples before/after TIP as an example##
#############################################################################
#Make a vector noting which samples were collected on vs. off antibiotic
#"Meta_ag_count_round" is a table with samples as columns and taxa as rows
Week <-str_extract(colnames(Meta_ag_count_round[,1:157]),"S\\d{1,2}")
Abx <- as.factor(ifelse(Week %in% c("S0","S5"),"A","B"))
#make Species designations rownames (makes output of aldex easier to interpret)
rownames(Meta_ag_count_round) <- Meta_ag_count_round$Species
#perform Aldex2 calculation to ID significantly different taxa
x.all_abx <- aldex(
  Meta_ag_count_round[,which(grepl("S0",colnames(Meta_ag_count_round))
|grepl("S1",colnames(Meta_ag_count_round)))],#only use Baseline and Week 1 samples

as.character(Abx[which(grepl("S0",colnames(Meta_ag_count_round))|grepl("S1",colnames(Meta_ag_count_r
ound)))]), #compare samples off vs. on Abx (A vs. B), using only baseline and week 1 samples
  mc.samples=128, #number Monte Carlo samples to use to estimate distribution, this number is
recommended by aldex2 doc.
  test="t", #Run Welch's t and Wilcoxin test
  effect=TRUE, #print out effect sizes
    include.sample.summary=FALSE) #don't include mean clr values for each sample

#order Species by effect size and extract those taxa with p value less than 0.05
x.all_abx <- x.all_abx[order(abs(x.all_abx$effect),decreasing = T),]
ald_abx <-rownames(x.all_abx)[which(x.all_abx$wi.eBH<0.05)] #wi.eBH is the wilcoxin rank sum test with a
benjamini hochbert correction for multiple comparisons

#############################################################################
####################PCoA using Aitchison Distance#######################
#######Here using the PCoA from Fig. 3a as an example#################
#############################################################################
#first need to replace zeros using Bayesian-multiplicative replacement, bc Aitchison involves logs
```

```
tip_prop_zr <-
t(cmultRepl(as.data.frame(t(Meta_ag_count_round[,which(grepl("S0",colnames(Meta_ag_count_round))
|grepl("S1",colnames(Meta_ag_count_round)))])),method="CZM"))
#center log transform data and transpose
tip_prop_clr <- as.data.frame(clr(tip_prop_zr))
clr_t <- t(tip_prop_clr)

#create distance matrix. Aitchison distance is euclidean distance on centered log-transformed data
mod_dist <- vegdist(as.matrix(t(tip_prop_clr[,1:60])),method="euclidean")
#Make vector for samples collected on and off antibiotic
Week <-str_extract(colnames(tip_prop_clr[,1:60]),"S\\d{1,2}")
Abx <- as.factor(ifelse(Week %in% c("S0","S5"),"A","B"))
#Perform PERMANOVA using the Abx vector as a grouping variable
mod_perm <- adonis(clr_t~Abx, permutations=1000, method="euclidean")
mod_perm
#Homogeneity of Variance using Abx as grouping variable -> this also calculates coordinates for biplot
x<-betadisper(mod_dist, group=factor(Abx))
permutest(x)
```

SUPPLEMENTAL METHODS

Study design and Description of Subjects

The TIP study population comprised 30 people with CF who were spontaneously expectorating at three clinical sites, the University of Washington Medical Center, Seattle Children's Hospital and University of Michigan Health System. Spontaneously expectorated sputum was collected prospectively from these subjects at commencement of TIP ("baseline") and weekly during the month-long treatment period, which occurred after at least 28 days without exposure to any antibiotics other than maintenance azithromycin. A subset of participants also provided a follow-up sample one month after completion of the 28-day-long TIP treatment cycle ("follow-up"). Study inclusion criteria were as follows:

- Documented diagnosis of CF, based on standard criteria established by the Cystic Fibrosis Foundation
- Age > 6 years
- Able to reproducibly perform spirometry for clinical care
- Able to expectorate at least daily
- History of recurrent P. aeruginosa positive respiratory cultures, defined as 2 or more occurrences prior to enrollment
- Respiratory culture positive for P. aeruginosa at least once within the 12 months prior to enrollment
- Abstinence or documented use of contraception by females of child-bearing potential at baseline
- Signed informed consent, and assent, as applicable
- FEV_1 greater than 25% predicted at baseline

Study exclusion criteria were as follows:

- History of intolerance of inhaled tobramycin
- In the opinion of CF care provider, judged not to be able to tolerate 28 days without antibiotics other than chronic azithromycin
- Treatment with any antibiotics (other than chronic azithromycin) within the prior 28 days to first sample collection
- Initiation of chronic azithromycin within the prior 28 days of first sample collection
- Requiring hospitalization or treatment with antibiotics in addition to inhaled tobramycin during the study period
- Positive urine or serum pregnancy test at baseline
- Participating in an investigational drug trial that is determined by the investigators to conflict with this study's goals

Of the 30 participants recruited to the study, 4 did not complete the study: one was lost to follow up, one decided to leave the study early for unknown reasons, one did not tolerate tobramycin and was switched to alternative inhaled and oral antibiotics, one was started on intravenous antibiotics due to increased symptoms. We excluded all samples collected during periods of clinical instability requiring new antibiotics other than tobramycin. We chose these exclusion criteria to ensure all participants were clinically stable at the time of sample collection (limiting potential microbiota changes not attributable to antibiotic therapy). Sputum samples not collected during stability were excluded from the study.

Sputum specimen processing

All samples were shipped on ice within 48 hours of collection and sputum was homogenized with dithiothreitol to liquefy the specimen as described below. After culture, the remainder of the specimen was frozen for DNA analysis. Storage at 4C has been shown to preserve CF sputum microbiota profiles compared with -80C [3]. The 48-hour window was chosen to allow participants to send in samples from home (maximizing recruitment in the study) as well as to allow samples to be available by mail from three different clinical sites in two cities, (University of Washington Medical Center, Seattle Children's Hospital and University of Michigan Health System), broadening generalizability. Sputum was also frozen in glycerol to allow for future culturing of non-canonical organisms where volume allowed. Subject details are present in Table S1.

Sputum culture for standard pathogens

An aliquot of each sample was cultured on differential media according to standard practice [4] to calculate total culturable loads of gram negative organisms (predominantly P. aeruginosa) and S. aureus. Viable loads of each taxon between baseline and Week 1 were compared using a paired Wilcoxon signed-rank sum test.

Initial sample processing and bacterial culture

All samples were either collected at a study visit followed by transfer to the lab or mailed directly the lab by the study participant on ice. An aliquot of each sample was cultured on differential media according to standard practice [4] to calculate total culturable loads of gram negative organisms and of S. aureus as follows. Samples were diluted 1:1 in 0.1% Sputolysin (DTT; Millipore Sigma), serially diluted in PBS and plated on blood agar (Fisher), MacConkey agar (BD) and mannitol salt agar (BD) according to standard practice [4]. Plates were incubated at 35°C for 24-48 hrs prior to quantifying the abundances of both S. aureus and select gram negative organisms (MacConkey agar, predominantly P. aeruginosa but counts also including Achromobacter xylosoxidans, Stenotrophomonas maltophila, Serratia marcescens and an unidentified yeast taxon determined by microbiota analysis). Colonies on mannitol salt agar were confirmed as S. aureus either by mannitol fermentation or nucA PCR [5]. All colonies on MacConkey agar were counted as gram-negative organisms. Representative isolates were collected and frozen for future analysis.

Spirometry and symptom scores

Standard spirometric measures of lung function, including forced expiratory volume in 1 second (FEV_1) and forced vital capacity (FVC), were recorded before and after therapy according to standard methods [6]. Participants' subjective responses were collected before, during and one month after treatment, using the standardized Chronic Respiratory Infection Symptom Score (CRISS) [7], which quantifies symptoms such as cough and fatigue.

DNA extraction

All samples were extracted from frozen, DTT-treated sputum using a method designed to minimize human DNA and extracellular bacterial DNA as detailed previously [8]. 300-450mL of DTT-treated sputum was suspended in 7 mL dH_2O and incubated at room temperature for one hour with gentle agitation. 10x strength Benzonase buffer (200 mM Tris-HCl, 10 mM $MgCl_2$) to a final 1X and 250U Benzonase (Sigma E-1014) were added and the sample was incubated at 37°C for 2 h with gentle agitation. 1.5M NaCl (150 mM final) was added and the resulting solution was centrifuged at 8,000g for 10min and the pellet washed once with PBS, centrifuged at 13,000g for 4min, the pellet resuspended in 400µL TE, and 0.5M EDTA (5 mM final) was added to inactivate the endonuclease. A mixture of 1 mm and 0.1 mm silica:zirconia beads and a single tungsten-carbide bead was added to the TE solution, followed by bead-beating for one minute in a BioSpec MiniBeadBeater. The resulting solution was boiled for 5 min at 95°C. Lysozyme (Sigma L6876, 3 mg/mL final) and lysostaphin (Ambi LSPN, 0.14 mg/mL final) were added, and the sample was incubated for one hour at 37°C. Proteinase K (Invitrogen 25530049, 1.4 mg/mL final) and SDS (1.8% final) were added, and the resulting solution incubated at 56°C for 30 min before cooling to room temperature. The solution was removed to a separate tube and 5 M NaCl was added (2 M final) before adding phenol:chloroform:isoamylalcohol (25:24:1)

at a 1:1 volume. The solution was then incubated for 20 min at room temperature, centrifuged at 13,000 g for 20 min [17] and the top aqueous layer was collected. 0.133 volume equivalent of 7.5 M ammonium acetate was added to the aqueous layer, and the resulting solution diluted 1:1 with cold 100% ethanol to precipitate DNA. DNA product was cleaned with a spin column (BioBasic BS423, skipping the cell lysis steps and proceeding directly to DNA clean-up steps). Reagent blanks consisting of PBS alone were processed and sequenced via 16S amplicon sequencing to assess contamination.

Quantitative PCR

Total bacterial load was determined by universal 16S quantitative PCR with PowerUp SYBR Green Master mix (Applied Biosystems A25742) using previously published primers and reaction conditions [9]. Proportions of human and bacterial DNA in samples was determined by calculating the genome equivalents (GE) of each, using $5x10^6$ bp for the average microbial genome size. Based on the following commonly-used definition of limit of detection (LOD) for qPCR [10,11], "the lowest copy number associated with the serial dilution that gave a positive PCR response on 95% of occasions," the LOD of our universal 16S qPCR assay was approximately 8.24 GE per uL. In addition to all sputum samples, we calculated total bacterial loads of 12 extraction blanks. Average loads in extraction blanks were ~4 logs lower than bacterial loads in samples and the lowest sample bacterial load was >2 logs higher than the highest extraction blank bacterial load.

Taxon-specific qPCR was determined using previously published primers against the femA gene for S. aureus [12], the gyrB gene for P. aeruginosa [13] and genus-specific primers against the 16S genes for both Streptococcus [14] and Prevotella [15].

Data were analyzed with the Bio-Rad CFX Maestro 1.1 (Version 4.1.2434.0124), using software-defined Cq thresholds and total bacterial load between baseline and Week 1 was compared using a paired Wilcoxon signed-rank sum test.

Phylogenetic composition from metagenomic shotgun sequencing

Next-generation sequencing libraries were prepared for all samples using the Nextera DNA Sample Prep Kit (Illumina FC-121-1031) following manufacturer's instructions. All samples were sequenced on the Illumina HiSeq platform, producing an average of 3.3 $x10^7$ total reads per sample after quality filtering. Sequencing data from all samples were de-duplicated using SeqUniq (version 0.1) https://github.com/standage/sequniq) and quality filtered using KneadData (version 0.6.1) and Trimmomatic (version 0.33) [16], producing an average of 2.9 $x10^7$ reads per sample. Human reads, comprising 12%-95% of our samples based on the efficiency of the human DNA depletion step in our DNA extraction method, were identified and removed with BMTagger (version 3.101), resulting in $1.3x10^7$ microbial reads per sample on average. Due to the range of human DNA load across our samples, we were unable to use this diploid genome to normalize our microbial read abundances as has been proposed elsewhere [17]. Community phylogenetic composition was determined using MetaPhlAn2 [18,19] (version 2.2.0) to produce a MetaPhlAn2 taxa table. MetaPhlAn2 utilizes a small set

of taxon-specific bacterial marker genes for taxonomic assignment and excludes genes from accessory genomic elements or paralogs, which cannot be reliably attributed to a specific taxon; thus it prioritizes specificity over sensitivity in contrast to other bioinformatic pipelines which determine the taxonomic origin of all reads. We utilized such a marker-gene approach for taxonomic designation to reduce the spurious detection of low-abundance taxa. [20–22] All commands were executed with default settings, with the exception of KneadData, which was used with the "–run-bmtagger" flag. For all samples, the Achromobacter genus represents reads originally annotated either as Achromobacter or Bordetella, the latter of which proved to reflect misidentified Achromobacter reads: Using mock communities, we found MetaPhlAn2 consistently splits this taxon into Achromobacter and Bordetella, even when this latter taxon was not present, as confirmed by 16S amplicon sequencing of the same samples [8].

As the lowest sample bacterial load was >2 logs higher than the highest extraction blank bacterial load, we pooled all taxa below 1% in an “other” category. Those taxa pooled into the other category are listed in File S1.

Phylogenetic composition from 16S amplicon sequencing

As a control for our metagenomic sequencing, we performed 16S amplicon sequencing (the most common microbiota sequencing method) for 155 out of the 157 study sputum samples as well as 12 extraction blank controls, one for each day of DNA extractions. The V4 region of the 16S rRNA gene was amplified using primers from the Earth Microbiome Project [18] and barcodes adapted from Kozich et al [23]. (detailed at https://github.com/SchlossLab/MiSeq_WetLab_SOP/blob/master/MiSeq_WetLab_SOP_v4.md). 16S amplicons were made under the following conditions: 94°C for 3min, 30 cycles of the following sequence: [94°C for 45s, 50°C for 60s, 72°C for 90s], and then 72°C for 10min. Libraries were constructed by pooling equimolar amounts of each sample or of each blank at the volume of the least concentrated sample. Libraries were sequenced on the Illumina MiSeq platform producing paired 300 bp reads.

16S amplicon sequencing data were analyzed using the denoising program DADA2 [24] (version 1.6.0) and the complete code is listed in Supplementary File 1. Briefly, we computationally trimmed 10bp off the beginning of the both the forward and reverse reads and truncated the forward read to 200bp and the reverse read to 100bp. We used our entire data set to define an error rate at each base pair and then denoised all sequences. Forward and reverse reads were merged and any pair without perfect overlap was removed. Finally, chimeric sequences were removed. This program produces a list of “Amplicon Sequence Variants” (ASV) analogous to OTUs generated with a 97% clustering method. Each ASV was annotated with the RDP Bayesian classifier [24] against the SILVA database [25] to produce a 16S amplicon taxa table. ASVs identified as Pseudomonas, Staphylococcus and Achromobacter were analyzed with BLASTn to determine species identity.

After quality filtering, sputum samples produced an average of 83,332 reads per sample, extraction blanks produced an average of 575 reads and amplicon blanks produced an average of 126 reads. Most reads in the blanks were taxa in common among samples with similar sequencing barcodes. As the absolute abundance of

these reads was 2-3 logs lower in blanks than in neighboring samples during sequencing, we found it unlikely they contributed to the taxonomic profiles of our samples, but that these reads were more likely due to errors in barcode reading [26,27]. The remaining taxa were those noted by Salter et al. to be common reagent contaminants [28] and were found at <1% relative abundance in samples. As the number of reads from blanks were substantially outnumbered by those detected in sputum samples, we did not analyze the reads from these controls further.

As the lowest sample bacterial load was >2 logs higher than the highest extraction blank bacterial load, we pooled all taxa below 1% in an "other" category. Those taxa pooled into the other category and those taxa identified in extraction blanks are listed in File S1. The phylogenetic profiles of extraction blanks and amplicon PCR no template controls are shown in Fig. S10.

Microbiota composition analysis

Analysis of ecological sputum microbiota composition was performed in R [29] (version 3.4.2) and visualized using ggplot2 [30] (version 2.2.1). For sputum samples, all taxa with <1% relative abundance in all samples or >1% relative abundance solely in extraction blanks were pooled into the "Other" category, as the largest difference in bacterial loads between samples and extraction blanks was 2 logs.

As MetaPhlAn2 output data are expressed as proportions, we multiplied the proportion of each taxon in a sample by that sample's total number of microbial reads, and rounded to the nearest integer value, to obtain calculated absolute number of reads per sample contributed by each taxon. Pseudocounts were added to all reads using the cmultRepl function in the zCompositions [31] package to remove zero values, and a centered log transformation using the clr function was then performed in the compositions package [32]. PERMANOVA calculations and PCA plots using Euclidean distance of transformed data were performed using the adonis function in the vegan package [33]. Homogeneity of variance was calculated and PCA coordinates determined using the betadisper function in the vegan package. Differences in taxonomic abundances between sample groups were identified by paired Wilcoxon signed-rank test performed using the aldex function in the ALDEx2 package, a probabilistic model that implements statistical testing using a Dirichlet distribution designed for compositional data [34,35]. Code for these analyses is presented in File S2.

Estimated absolute abundances for all taxa were calculated by multiplying relative abundance data, after addition of pseudocounts as detailed above, by the total bacterial load determined by universal 16S qPCR.

Strain-level analysis of P. aeruginosa and S. aureus was performed using StrainPhlAn2 [36] applied to metagenomic data, using default settings with the exception of setting the "--bootstrap_raxml" flag to 999.

Functional Metagenomic Analysis

After quality filtering, all metagenomic sequencing reads were mapped to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [37] and normalized using MUSiCC [38], a computational program that uses single-copy genes to normalize functional abundances in each sample, as described previously [39]. Samples

were collapsed at the KEGG module level. All modules with a variance less than 0.2 and those present in less than 25% of samples were removed. All samples were processed in the same manner as taxonomic data as detailed above.

Statistical analysis: Comparison of changes in sputum microbiota and clinical outcome measures

For a priori comparison of all microbial community characteristics to therapeutic response, we calculated change in $ppFEV_1$ and CRISS symptom score from baseline to week 4 for all subjects who had week 4 visits (26 subject total). To maximize analytical power, we treated subjective and objective clinical responses as continuous, rather than dichotomous, variables. Log-transformed viable counts of gram-negative organisms and *S. aureus*, and total bacterial load via 16S qPCR, both at baseline and change in these features from baseline to week 4, were compared to clinical outcomes using linear regression. Centered log transformation of relative abundance of all taxa, as detailed above, at baseline and change in each taxon from baseline to week 4 were compared to clinical outcomes using linear regression. Shannon diversity at baseline and change in this metric from baseline to week 4 were compared to clinical outcomes using linear regression. Each KEGG module as baseline was compared to clinical outcomes using EdgeR [40]. A Benjamini-Hochberg correction for multiple comparisons was performed for all analyses described above, and p-values smaller than 0.05 were considered significant. Relative abundances of all taxa, normalized abundances of functional groups and diversity metrics were compared to change in cultured viable count of *P. aeruginosa* and *S. aureus* from baseline to week 1 in the same manner.

1 Enright PL, Beck KC, Sherrill DL. Repeatability of Spirometry in 18,000 Adult Patients. *American Journal of Respiratory and Critical Care Medicine* 2004;**169**:235–8. doi:10.1164/rccm.200204-347oc

2 Fuchs HJ, Borowitz DS, Christiansen DH, *et al.* Effect of aerosolized recombinant human DNase on exacerbations of respiratory symptoms and on pulmonary function in patients with cystic fibrosis. *New England Journal of Medicine* 1994;**331**:637–42.

3 Zhao J, Li J, Schloss PD, *et al.* Effect of sample storage conditions on culture-independent bacterial community measures in cystic fibrosis sputum specimens. *Journal of clinical microbiology* 2011;**49**:3717–8.

4 Saiman L, Siegel JD, LiPuma JJ, *et al.* Infection Prevention and Control Guideline for Cystic Fibrosis: 2013 Update. *Infection Control & Hospital Epidemiology* 2014;**35**:S1–67. doi:10.1086/676882

5 Brakstad OG, Aasbakk K, Maeland JA. Detection of Staphylococcus aureus by polymerase chain reaction amplification of the nuc gene. *Journal of clinical microbiology* 1992;**30**:1654–60.

6 Szczesniak R, Heltshe SL, Stanojevic S, *et al. Use of FEV1 in cystic fibrosis epidemiologic studies and clinical trials: A statistical perspective for the clinical researcher*. 2017. doi:10.1016/j.jcf.2017.01.002

7 Goss CH, Edwards TC, Ramsey BW, *et al.* Patient-reported respiratory symptoms in cystic fibrosis. *Journal of Cystic Fibrosis* 2009;**8**:245–52. doi:10.1016/j.jcf.2009.04.003

8 Nelson MT, Pope CE, Marsh RL, *et al.* Human and Extracellular DNA Depletion for Metagenomic Analysis of Complex Clinical Infection Samples Yields Optimized Viable Microbiome Profiles. *Cell Reports* 2019;**26**:2227-2240.e5.

9 Nadkarni MA, Martin FE, Jacques NA, *et al.* Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology* 2002;**148**:257–66.

10 Burns M, Valdivia H. Modelling the limit of detection in real-time quantitative PCR. *European Food Research and Technology* 2007;**226**:1513–24. doi:10.1007/s00217-007-0683-z

11 Bustin SA, Benes V, Garson JA, *et al.* The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry* 2009;**55**:611–22. doi:10.1373/clinchem.2008.112797

12 Johnson EJ, Zemanick ET, Accurso FJ, *et al.* Molecular Identification of Staphylococcus aureus in Airway Samples from Children with Cystic Fibrosis. *PLOS ONE* 2016;**11**:e0147643. doi:10.1371/journal.pone.0147643

13 Anuj SN, Whiley DM, Kidd TJ, *et al.* Identification of Pseudomonas aeruginosa by a duplex real-time polymerase chain reaction assay targeting the ecfX and the gyrB genes. *Diagnostic Microbiology and Infectious Disease* 2009;**63**:127–31. doi:10.1016/j.diagmicrobio.2008.09.018

14 Olson AB, Sibley CD, Schmidt L, *et al.* Development of Real-Time PCR Assays for Detection of the Streptococcus milleri Group from Cystic Fibrosis Clinical Specimens by Targeting the cpn60 and 16S rRNA Genes. *Journal of Clinical Microbiology* 2010;**48**:1150–60. doi:10.1128/jcm.02082-09

15 Martin FE, Nadkarni MA, Jacques NA, *et al.* Quantitative microbiological study of human carious dentine by culture and real-time PCR: association of anaerobes with histopathological changes in chronic pulpitis. *Journal of clinical microbiology* 2002;**40**:1698–704.

16 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**:2114–20. doi:10.1093/bioinformatics/btu170

17 Chouvarine P, Wiehlmann L, Losada PM, *et al.* Filtration and normalization of sequencing read data in whole-metagenome shotgun samples. *PloS one* 2016;**11**.

18 Thompson LR, Sanders JG, McDonald D, *et al.* A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 2017;**551**:457 EP-.

19 Truong DT, Franzosa EA, Tickle TL, *et al.* MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods* 2015;**12**:902–3. doi:10.1038/nmeth.3589

20 Simon HY, Siddle KJ, Park DJ, *et al.* Benchmarking metagenomics tools for taxonomic classification. *Cell* 2019;**178**:779–94.

21 Teo YV, Neretti N. A comparative study of metagenomics analysis pipelines at the species level. *bioRxiv* 2016;:081141.

22 Tamames J, Cobo-Simón M, Puente-Sánchez F. Assessing the performance of different approaches for functional and taxonomic annotation of metagenomes. *BMC genomics* 2019;**20**:1–16.

23 Kozich JJ, Westcott SL, Baxter NT, *et al.* Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Applied and Environmental Microbiology* 2013;**79**:5112–20. doi:10.1128/aem.01043-13

24 Wang Q, Garrity GM, Tiedje JM, *et al.* Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology* 2007;**73**:5261–7. doi:10.1128/aem.00062-07

25 Quast C, Pruesse E, Yilmaz P, *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 2013;**41**:D590–6. doi:10.1093/nar/gks1219

26 Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Research* 2012;**40**:e3–e3. doi:10.1093/nar/gkr771

27 Sinclair L, Osman OA, Bertilsson S, *et al.* Microbial Community Composition and Diversity via 16S rRNA Gene Amplicons: Evaluating the Illumina Platform. *PLoS ONE* 2015;**10**:e0116955. doi:10.1371/journal.pone.0116955

28 Salter SJ, Cox MJ, Turek EM, *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC biology* 2014;**12**:1.

29 R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: 2017. https://www.R-project.org/

30 Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York 2009. http://ggplot2.org

31 Palarea-Albaladejo J, Martín-Fernández JA. zCompositions—R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics and Intelligent Laboratory Systems* 2015;**143**:85–96.

32 van den Boogaart KG, Tolosana-Delgado R. "Compositions": a unified R package to analyze compositional data. *Computers & Geosciences* 2008;**34**:320–38.

33 Oksanen J, Blanchet FG, Friendly M, *et al. vegan: Community Ecology Package*. 2017. https://CRAN.R-project.org/package=vegan

34 Fernandes AD, Reid JN, Macklaim JM, *et al.* Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome* 2014;**2**:15.

35 Fernandes AD, Macklaim JM, Linn TG, *et al.* ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-Seq. *PLoS One* 2013;**8**:e67019.

36 Truong DT, Tett A, Pasolli E, *et al.* Microbial strain-level population structure and genetic diversity from metagenomes. *Genome research* 2017;**27**:626–38.

37 Kanehisa M, Goto S, Sato Y, *et al.* KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research* 2012;**40**:D109–14. doi:10.1093/nar/gkr988

38 Manor O, Borenstein E. MUSiCC: a marker genes based framework for metagenomic normalization and accurate profiling of gene abundances in the microbiome. *Genome Biology* 2015;**16**. doi:10.1186/s13059-015-0610-8

39 Manor O, Levy R, Pope CE, *et al.* Metagenomic evidence for taxonomic dysbiosis and functional imbalance in the gastrointestinal tracts of children with cystic fibrosis. *Scientific reports* 2016;**6**:22493. doi:10.1038/srep22493

40 Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* 2010;**26**:139–40. doi:10.1093/bioinformatics/btp616