Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/4/9/eaau1908/DC1

Supplementary Materials for

Niche partitioning of a pathogenic microbiome driven by chemical gradients

Robert A. Quinn*, William Comstock, Tianyu Zhang, James T. Morton, Ricardo da Silva, Alda Tran, Alexander Aksenov, Louis-Felix Nothias, Daniel Wangpraseurt, Alexey V. Melnik, Gail Ackermann, Douglas Conrad, Isaac Klapper, Rob Knight, Pieter C. Dorrestein*

*Corresponding author. Email: pdorrestein@ucsd.edu (P.C.D.); quinnr1234@gmail.com (R.A.Q.)

Published 26 September 2018, *Sci. Adv.* **4**, eaau1908 (2018)
DOI: 10.1126/sciadv.aau1908

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

(available at advances.sciencemag.org/cgi/content/full/4/9/eaau1908/DC1)

Table S1 (Microsoft Excel format). Patient samples and information collected in this study.

Table S2 (Microsoft Excel format). ANOVA of qualitative and quantitative variables measured during the WinCF pH gradient experiments.
Table S3 (Microsoft Excel format). Metabolites that most changed with the WinCF gas production gradient according to an RF variable importance plot from the untreated samples.
Table S4 (Microsoft Excel format). Mean abundance through the depth gradient (1 to 10 mm) of *P. aeruginosa* virulence factor metabolites detected in the WinCF depth experiments and the corresponding Pearson's correlation (*r*).
Table S5 (Microsoft Excel format). Deblurred OTUs and their sequences that most changed with the WinCF depth gradient according to an RF variable importance plot from the untreated samples.
Table S6 (Microsoft Excel format). Results of the assessment of bias in WinCF system.
Table S7 (Microsoft Excel format). Confusion matrix and out-of-bag error from an RF classification of the pH experiment metabolomics data based on patient source.

Supplementary Data

Experimental Design.

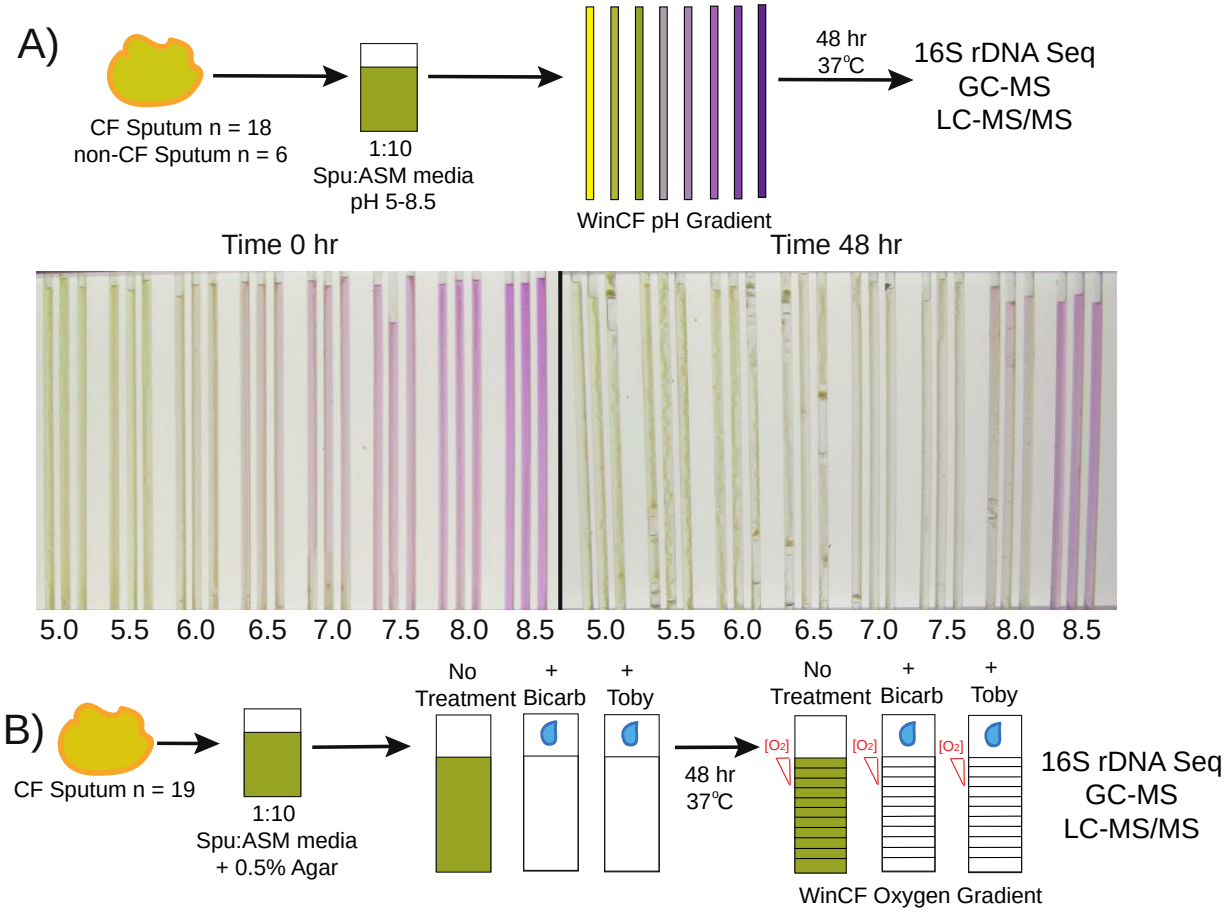


Fig. S1. Schematic of the experimental design with the WinCF system modified for A) the pH experiments and B) the oxygen experiments.

Random Forests Regression and Classification Analysis. A random forests analysis was run on both the metabolomic and microbiome data for the pH, gas production and depth linear variables to determine how strongly these variables impacted the omics data. All three variables had strong impacts on the data structure as seen by the close relationship between the actual measured variable and the predicted variable from the random forests linear model (fig. S2). The percent variance explained indicates how much of the variability seen in the data is explained by the linear variable.

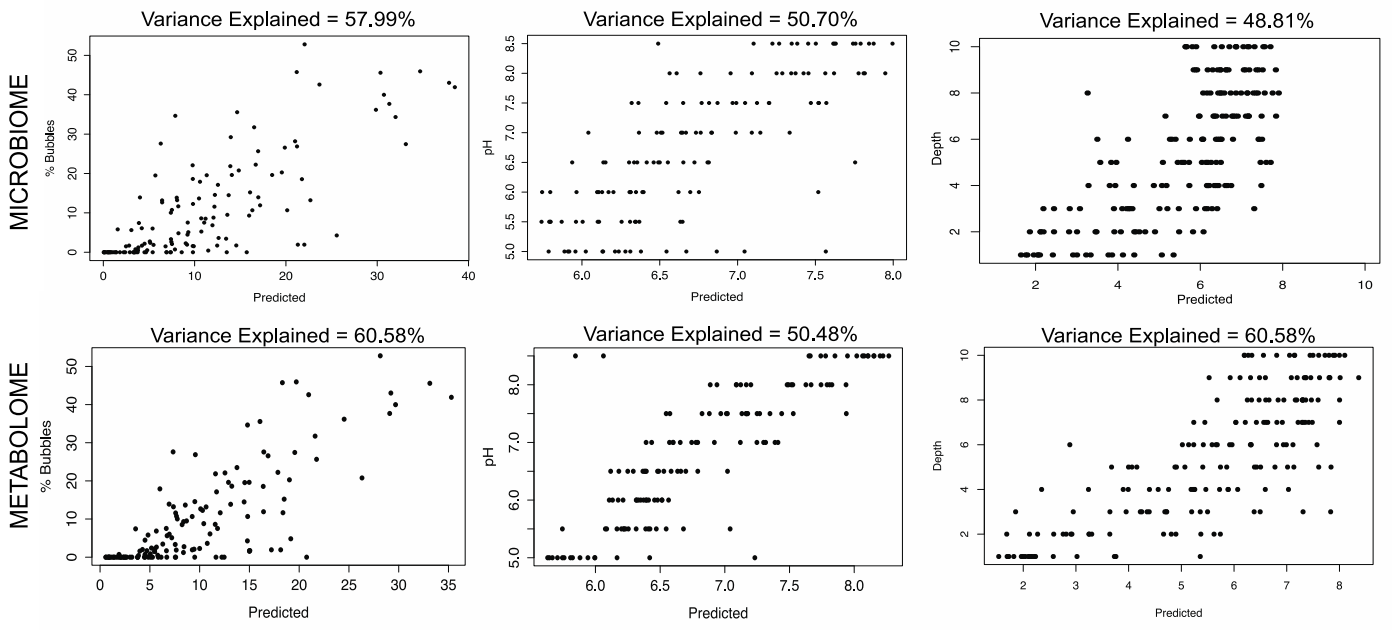


Fig. S2. Actual measurement versus predicted value from RF machine learning algorithm on the microbiome and metabolome data through pH, gas production, and depth variables. The actual value was imputed into the algorithm for each sample and the output was the predicted value based on the entire data set. The RF contained 5000 trees and the percent variance of each regression model is shown.

Balance Trees and Niche Partitioning. Due to challenges with microbiome data compositionality (*46*), balance trees were used to provide an overall assessment of niche partitioning through the chemical gradients and support the random forests analysis above. Balance trees showed that pH had a highly significant effect on the community structure (fig. S3a,d). There was a partitioning of the community into those of extreme high pH, high pH and low pH (fig. S3b). A histogram of the mean niche distribution of the most diverse community was present around 6.5, whereas the pH extremes, particularly higher pH, were occupied by fewer taxa (fig. S3c).

Balance trees revealed that the oxygen gradient was also strongly affected the microbiome data partitioning the community into those with a niche for the upper layers of the mucus (1-4 mm), the middle layers (4-7 mm) and the deepest layers (7-10 mm, fig. S3d,e). The genera were normally distributed around a depth of 4.5 mm, with few bacteria exclusively preferring the aerobic layers (1-2 mm, fig. S3f).

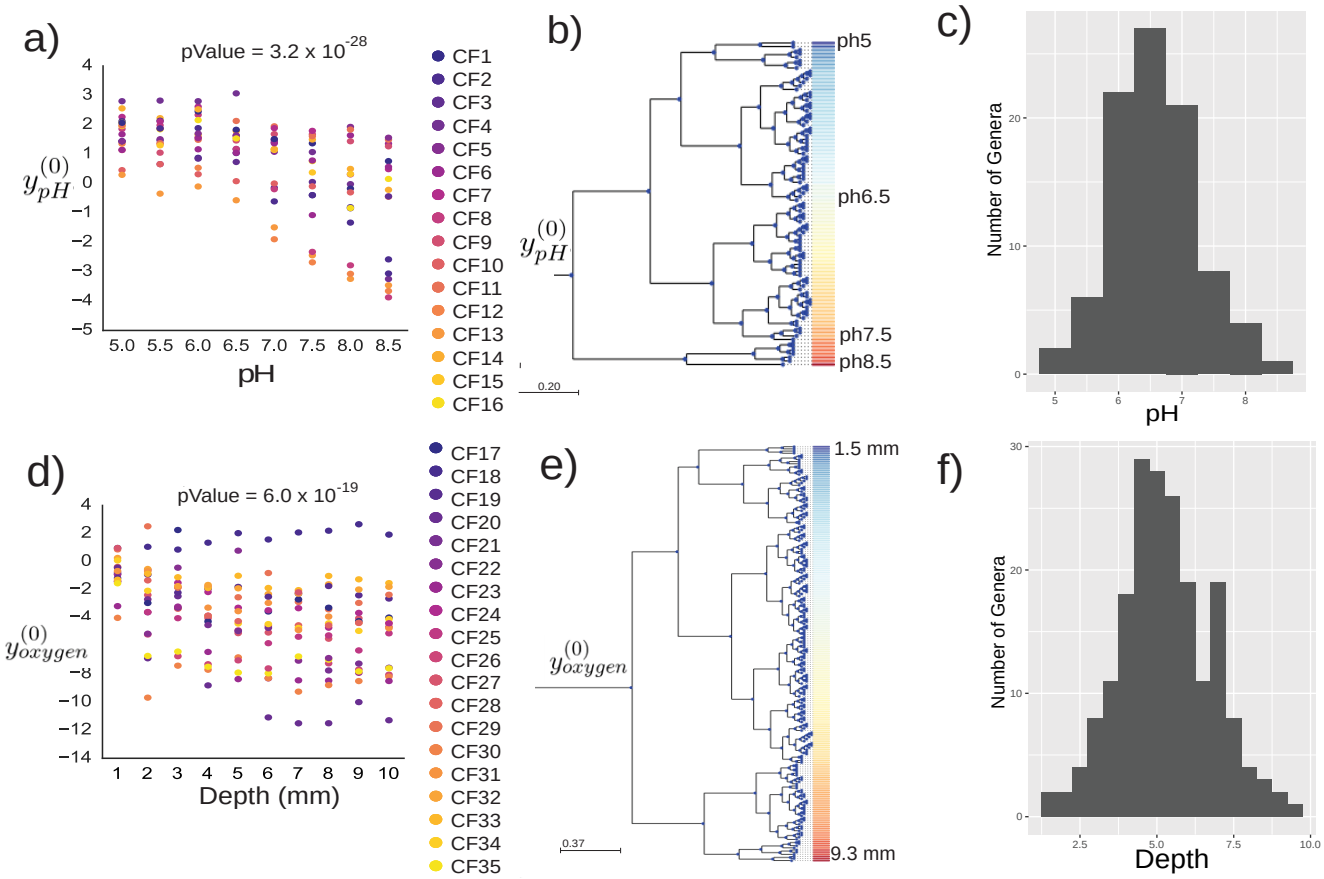


Fig. S3. Niche partitioning of CF lung microbiota in the pH and oxygen experiments. Balance trees (a-dc) and bacterial genera distributions (e and f) in the WinCF pH and oxygen experiments. Balances were calculated based on the mean a) pH and c) depth of each 16S rDNA microbiome profiles from the WinCF experiments. The tree represents a clustering of OTUs based on their mean pH or depth. Balances were constructed from this clustering using the isometric log-ration (ilr) transform. The trees on the right (b,d) show the largest division of microbes according to the pH and depth ranges that they occupy. The scatter plots on the left represent the log ratios of these partitions compared to pH (p-value = 3.2×10^{-28}) and depth (p-value=6×10^{-28}).

Taxa associated with Chemical Gradients. For the pH gradient, *Pseudomonas* (p=0.02), *Haemophilus* (p<0.0001), *Achromobacter* (p=0.0001) and the anaerobe *Fusobacterium* (p<0.0001) were positively correlated with elevated pH (Fig. 1b). The anaerobic genera *Veillonella* (p<0.0001), *Prevotella* (p<0.0001), *Actinomyces* (p<0.0001) and *Streptococcus* (p<0.0001) were negatively correlated with pH (Fig. 1b). *Fusobacterium*, also an obligate anaerobe, was found as one of the highest pH organisms. This result is supported in studies of oral microbial communities, which have found *Fusobacterium* to

exist in a high pH niche space (*47*). One patient had the opposite trend in the balance tree analysis with depth where *P. aeruginosa* increased (Fig. 1c, fig. S5). This was the only sample that did not contain anaerobes in the sequencing data used as the starter inoculant; instead, this patient's community was dominated by three different aerobic pathogens (fig. S5).

O_2 Penetration Into WinCF Media. The concentration of O_2 was measured through the vertical depth of the WinCF columns using Clark-type O_2 microsensors (tip diameter approx. 50 µm, OX-50, Unisense A/S Aarhus, Denmark). The O_2 sensors had a response time of <5 s, a stirring sensitivity of <2% and a detection limit of 0.3 µm O_2. Details on the experimental set-up and measurement procedure can be found in (*48*). Briefly, O_2 microsensors were linearly calibrated in air saturated fresh water (temperature = 25ºC) and in anoxic water (flushed with N_2). O_2 sensors were connected to a pA-meter (Unisense A/S) that was connected to an A/D converter (Unisense A/S) and interfaced with a PC running dedicated data acquisition and positioning software (ProFix, Pyroscience, Gmbh). O_2 gradients through the WinCF columns were measured from the surface of the WinCF column (depth = 0mm) until a depth of about 5mm with a step size of 50 µm to 200 um using a motorized micromanipulator (MU-1, Pyroscience, GmbH). O_2 gradients were measured in triplicates for each of two separate patients and one control (no sputum added).
The media was oxic from 0 to 2000 µm (O_2 saturation 90-100%), transitioned to anoxia from 2000 µm to 2200 µm and was anoxic through the rest of the 10 mm column depth (fig. S4).

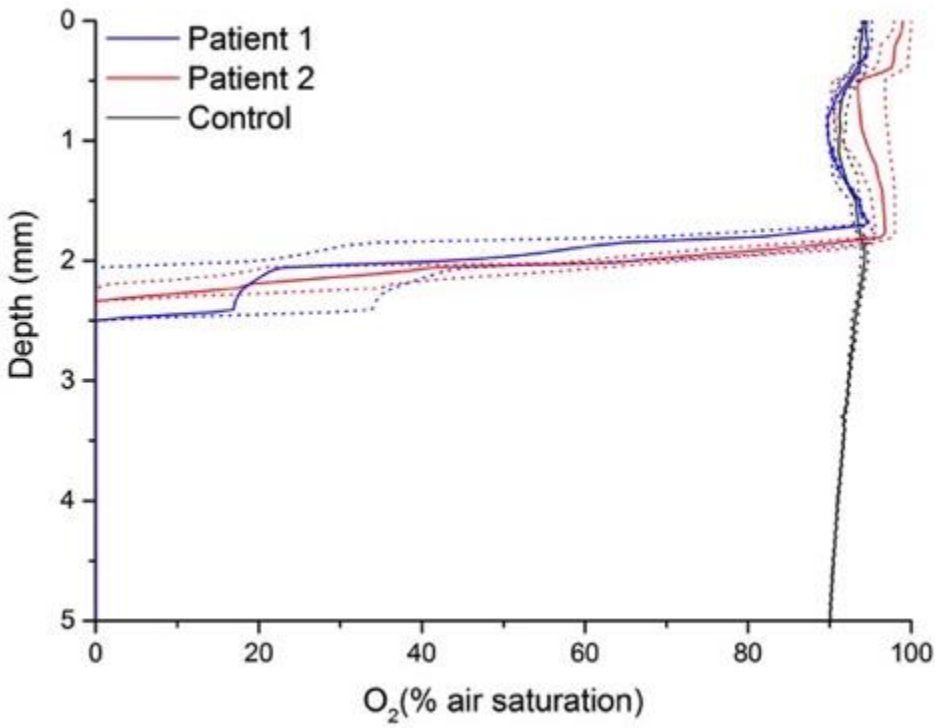


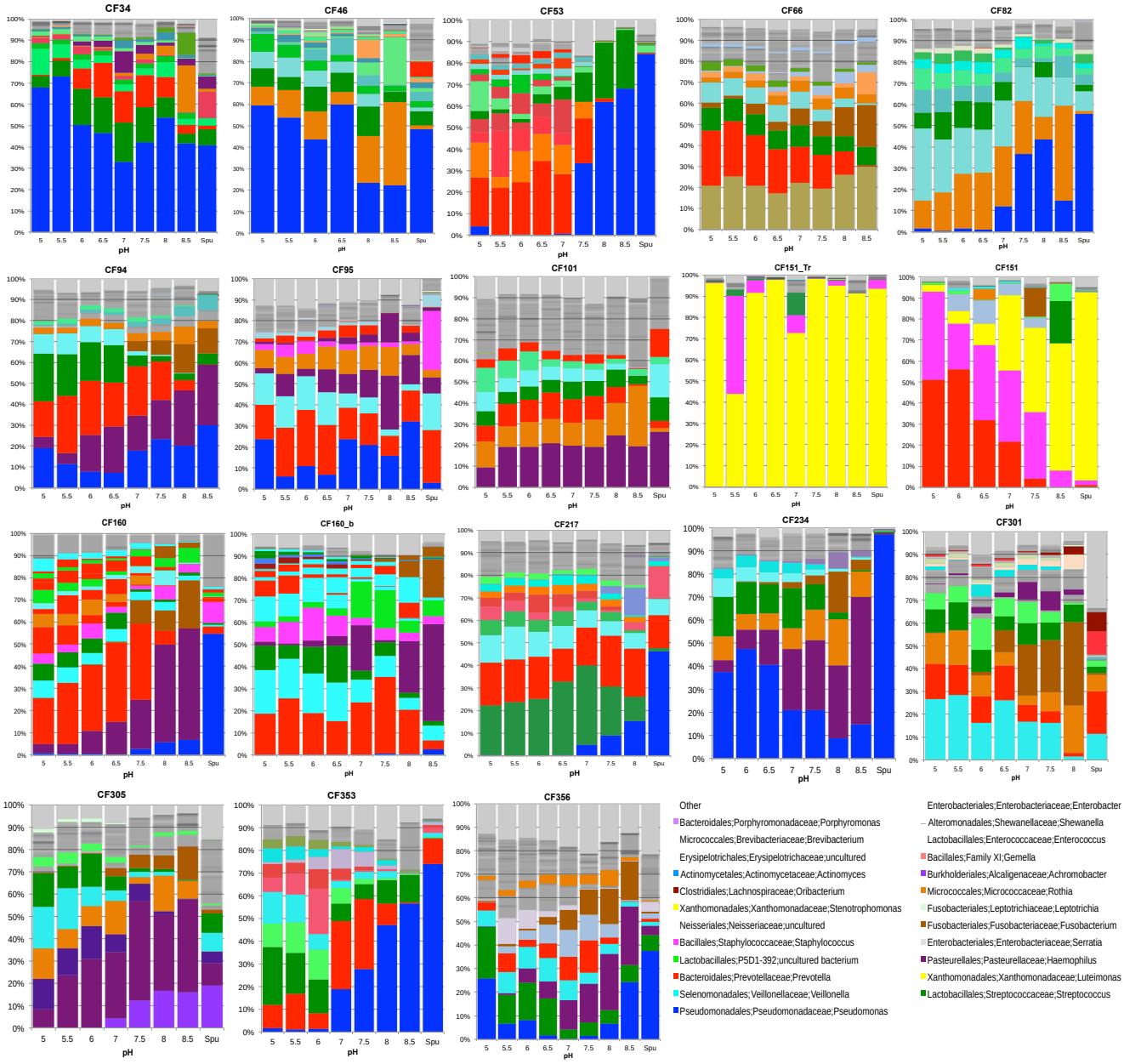
Fig. S4. O_2 microenvironment (% air saturation) through the WinCF vertical depth gradient after incubation with sputum from two patients compared to a noninoculated control. Solid lines indicate the mean while dotted lines indicate the duplicate runs for the two patients.

Microbes Associated with Gas production. Gas production explained 57.99% of the variation in the sequencing data (fig. S2). The GLMM revealed that *Pseudomonas* ($p = 0.002$), *Fusobacterium* ($p < 0.0001$) and *Haemophilus* ($p = 0.0002$) were negatively correlated with gas production and the anaerobes *Veillonella* ($p < 0.0001$), *Prevotella* ($p < 0.0001$), *Streptococcus* ($p = 0.0003$), *Atopobium* ($p < 0.0001$) and *Bulleidia* ($p < 0.0001$) were positively correlated. A random forests variable importance plot from the regression above with the gas production variable identified the microbes most positively or negatively associated with gas production (table S5).

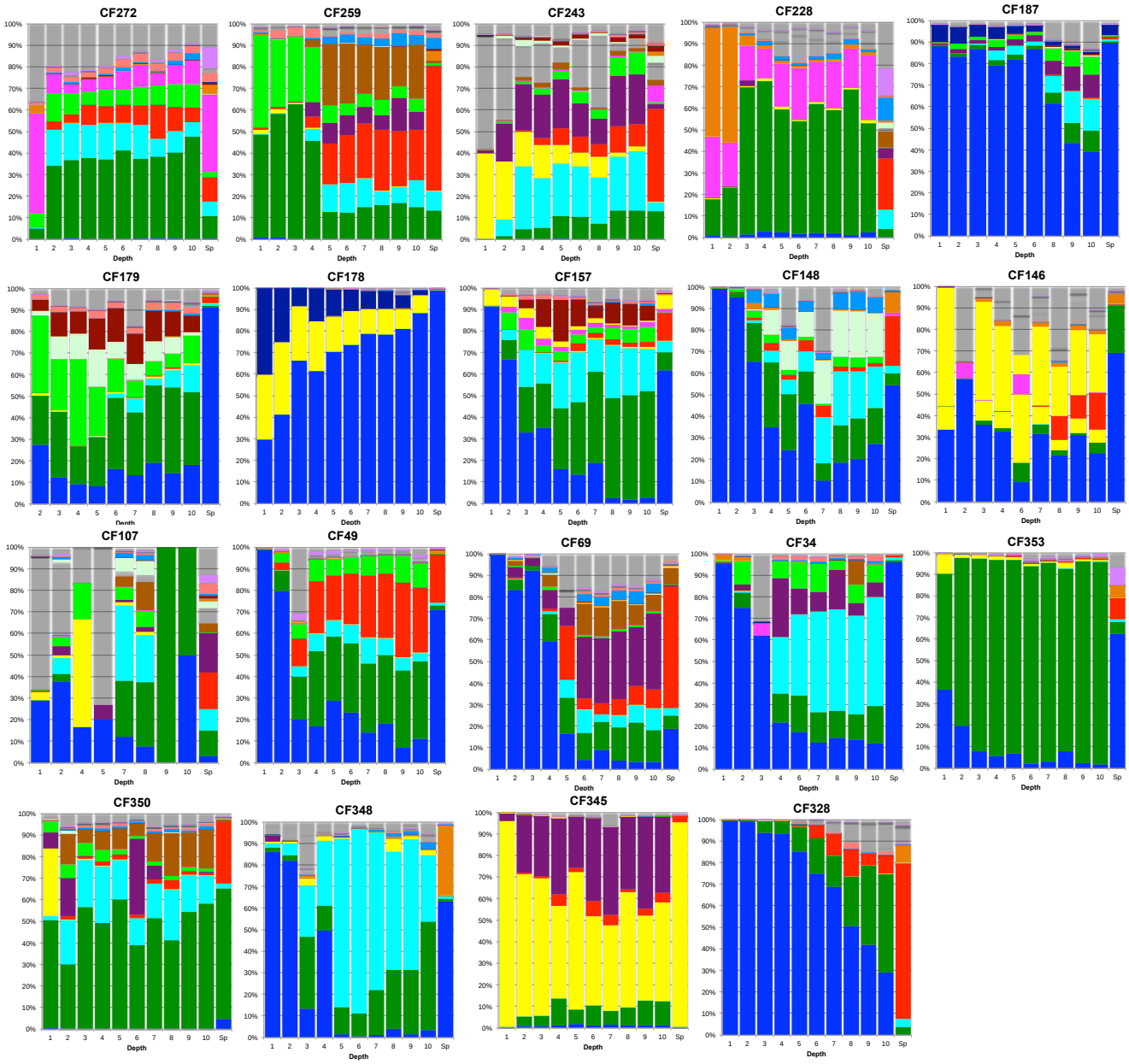
Microbes Correlated with SCFAs. Propionic acid was most strongly correlated to an *Actinomyces* sp. (*rho*=0.673), two *Prevotella* spp. (*rho* = 0.596, 0.573), *Prevotella melaninogenica* (*rho* = 0.591) and *Veillonella dispar* (*rho*=0.550). Butanoic acid was correlated with *Prevotella melaninogenica* (*rho*=0.560), a *Megasphaera* sp. (*rho*=0.519), a *Veillonella* sp. (*rho*=0.507), an *Atopobium* sp. (*rho*=0.492) and two *Fusobacterium* spp. (*rho*=0.486, 0.463).

The Patient Specific Response. A random forests classification was able to accurately predict the patient that each WinCF sample came from despite the strong effects of the chemical gradients (fig. S5, table S7). This indicated that there were strong patient specific signatures in both the microbiome and metabolomic data. Thus, although the gradients strongly shaped the chemistry and community structure of the CF microbiome, each individual has unique metabolites and bacterial species. However, no matter the dominant pathogen an individual is infected with the effects of pH and oxygen were similar. The pathogens dominate in the aerobic and high pH niche, and anaerobes dominate the opposite niche (fig. S5). For example, in the pH experiments patients dominated by *Pseudomonas* (CF53) see a bloom of this bacterium in the high pH niche space, similarly, one dominated by *Stenotrophomonas* (CF 151) does as well. Thus, there was evidence for niche replacement in these experiments because the patients had different microbial profiles in their sputum but the gradient response was consistent. Patients without *P. aeruginosa*, but infected with another CF pathogen, also had their community partition into anaerobic species in anoxic layers, with the alternate pathogen such as *S. maltophilia* and *S. aureus* in the aerobic layers (fig. S5).

a)


CF34
CF46
CF53
CF66
CF82
CF94
CF95
CF101
CF151_Tr
CF151
CF160
CF160_b
CF217
CF234
CF301
CF305
CF353
CF356
pH
Other
Bacteroidales;Porphyromonadaceae;Porphyromonas
Micrococcales;Brevibacteriaceae;Brevibacterium
Erysipelotrichales;Erysipelotrichaceae;uncultured
Actinomycetales;Actinomycetaceae;Actinomyces
Clostridiales;Lachnospiraceae;Oribacterium
Xanthomonadales;Xanthomonadaceae;Stenotrophomonas
Neisseriales;Neisseriaceae;uncultured
Bacillales;Staphylococcaceae;Staphylococcus
Lactobacillales;P5D1-392;uncultured bacterium
Bacteroidales;Prevotellaceae;Prevotella
Selenomonadales;Veillonellaceae;Veillonella
Pseudomonadales;Pseudomonadaceae;Pseudomonas
Enterobacteriales;Enterobacteriaceae;Enterobacter
Alteromonadales;Shewanellaceae;Shewanella
Lactobacillales;Enterococcaceae;Enterococcus
Bacillales;Family XI;Gemella
Burkholderiales;Alcaligenaceae;Achromobacter
Micrococcales;Micrococcaceae;Rothia
Fusobacteriales;Leptotrichiaceae;Leptotrichia
Fusobacteriales;Fusobacteriaceae;Fusobacterium
Enterobacteriales;Enterobacteriaceae;Serratia
Pasteurellales;Pasteurellaceae;Haemophilus
Xanthomonadales;Xanthomonadaceae;Luteimonas
Lactobacillales;Streptococcaceae;Streptococcus

b)



Other
Bacteroidales;Porphyromonadaceae;Porphyromonas
Micrococcales;Brevibacteriaceae;Brevibacterium
Erysipelotrichales;Erysipelotrichaceae;uncultured
Actinomycetales;Actinomycetaceae;Actinomyces
Clostridiales;Lachnospiraceae;Oribacterium
Xanthomonadales;Xanthomonadaceae;Stenotrophomonas
Neisseriales;Neisseriaceae;uncultured
Bacillales;Staphylococcaceae;Staphylococcus
Lactobacillales;P5D1-392;uncultured bacterium
Bacteroidales;Prevotellaceae;Prevotella
Selenomonadales;Veillonellaceae;Veillonella
Pseudomonadales;Pseudomonadaceae;Pseudomonas
Enterobacteriales;Enterobacteriaceae;Enterobacter
Alteromonadales;Shewanellaceae;Shewanella
Lactobacillales;Enterococcaceae;Enterococcus
Bacillales;Family XI;Gemella
Burkholderiales;Alcaligenaceae;Achromobacter
Micrococcales;Micrococcaceae;Rothia
Fusobacteriales;Leptotrichiaceae;Leptotrichia
Fusobacteriales;Fusobacteriaceae;Fusobacterium
Enterobacteriales;Enterobacteriaceae;Serratia
Pasteurellales;Pasteurellaceae;Haemophilus
Xanthomonadales;Xanthomonadaceae;Luteimonas
Lactobacillales;Streptococcaceae;Streptococcus

Fig. S5. Microbiome profiles of individual patients in the (a) pH and (b) oxygen experiments. Each patient contains a different principle pathogen, either *Pseudomonas*, *Stenotrophomonas*, *Achromobacter* or *Staphylococcus*. The plots demonstrate the effects of pH on the community of pathogens and anaerobes despite which principle pathogen the patients are dominated by. The data is shown at the genus level and the sputum sample prior to incubation in the WinCF tubes is highlighted with an asterisk.

Assessment of Biases in WinCF. To assess bias in the WinCF system, the number of OTUs detected in a sputum sample but not in the culture experiments (indicating their growth was inhibited in the WinCF environment) was quantified. The incidence of bias was extremely low (less than 10 OTUs) and not consistent (table S6). There was however, some evidence for some bias against *Rothia* spp. in these experiments which had been previously reported in WinCF (*12*).

P. aeruginosa metabolites in the WinCF Samples. P. aeruginosa metabolites detected through GNPS included quinolones, phenazines, rhamnolipids and the siderophore pyochelin. There were 91 nodes representing the quinolone cluster, of these, only 2-heptyl-4-quinolone (HHQ, *m/z* 244.171), 2-nonyl-4-quinolone (NHQ, *m/z* 272.202) and its related metabolite with a desaturated fatty acyl chain NHQ-db (*m/z* 270.202), 2-nonyl-3-hydroxyquinolone (NQNO, *m/z* 288.197), and the longer fatty acid chain quinolones 2-tridecyl-4-quinolone with one unsaturated carbon (C13:db-THQ, *m/z* 326.235) and with two unsaturated carbon chains (C13:db-db-THQ, *m/z* 324.235), were detected in both the sputum and the WinCF tubes (fig. S6). Pyochelin was also detected in the CF sputum and capillary tube samples. There were not any *P. aeruginosa* metabolites unique to sputum.

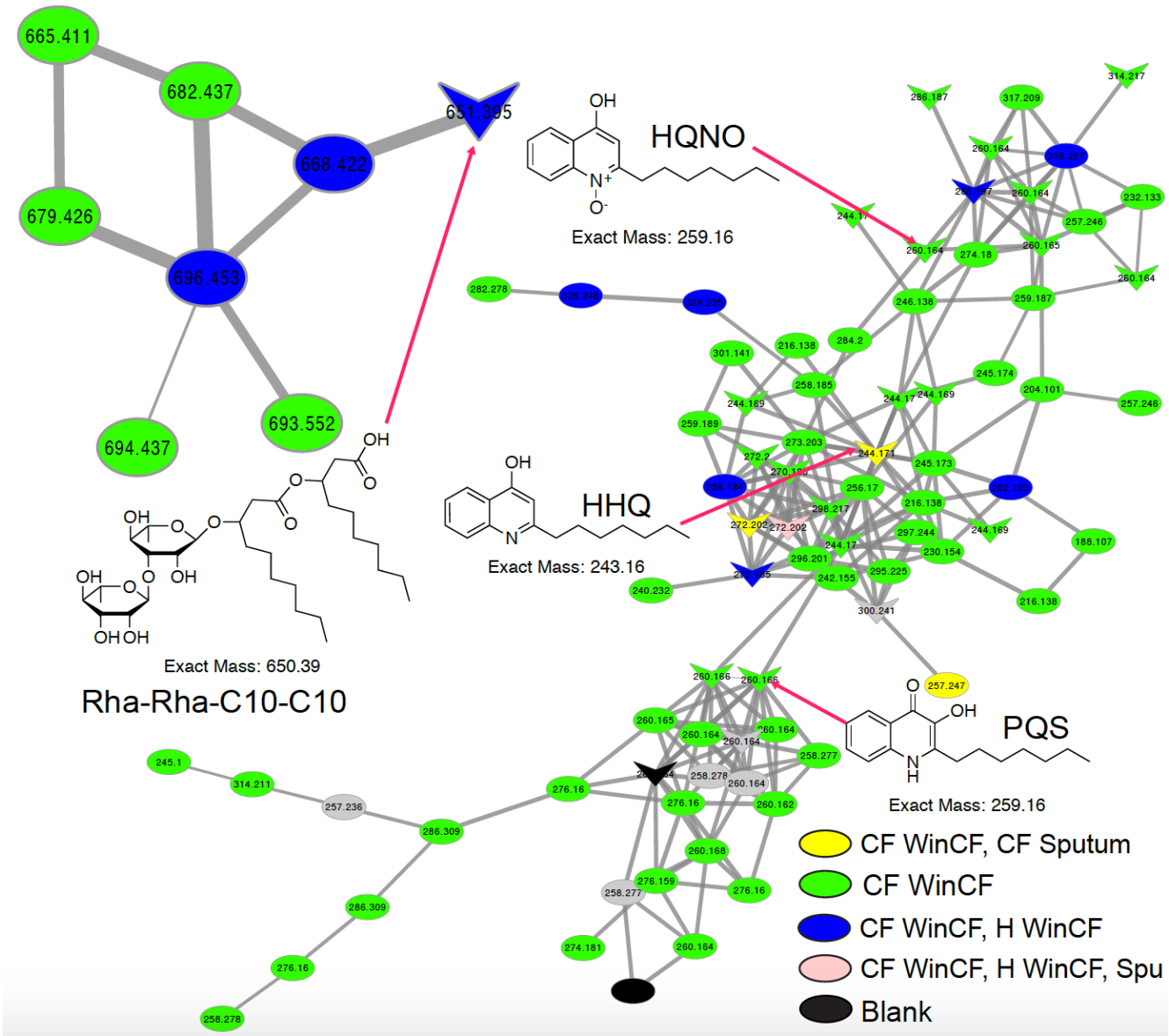
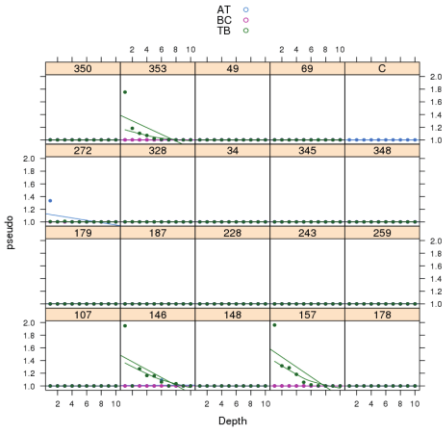
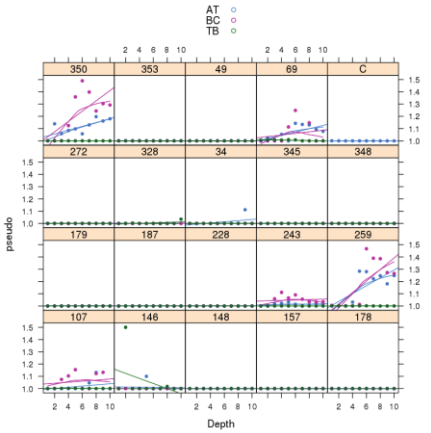


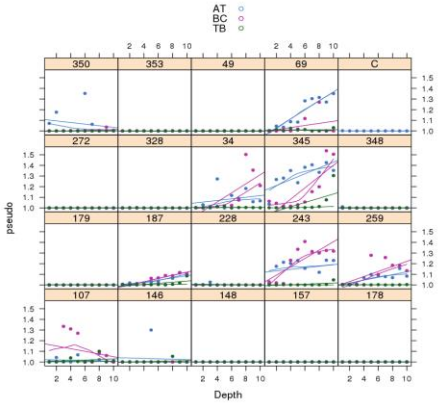
Fig. S6. Molecular network of rhamnolipids and quinolones detected in the LC-MS/MS data. Each node represents a single MS/MS clustered spectrum connections between the nodes indicate a cosine score between them of above 0.7. Arrowhead shaped nodes had hits in the GNPS database and oval nodes are unknown relatives. The nodes are colored by their sample source.



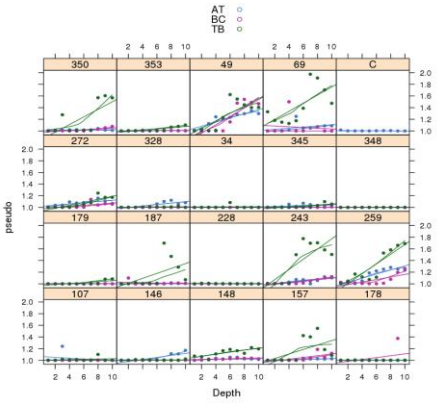
Aspergillus Mitochondrion



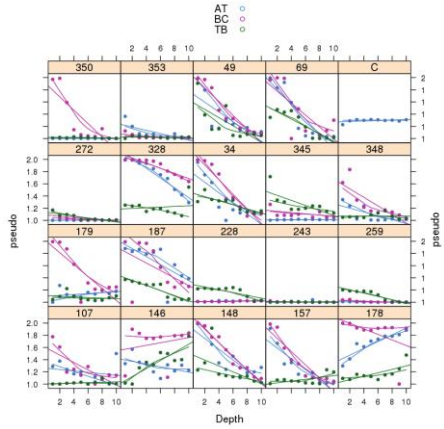
Fusobacterium



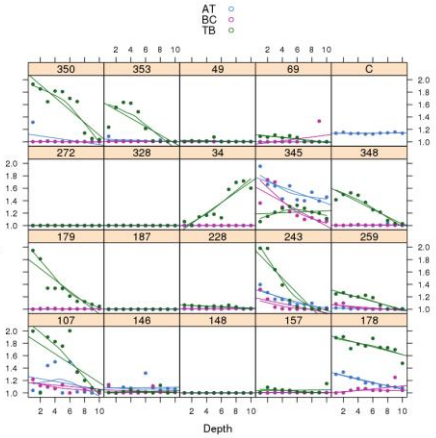
Haemophilus



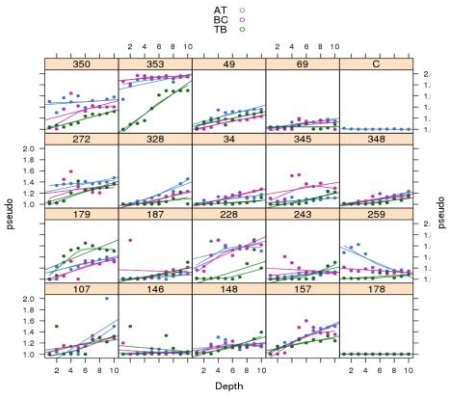
Prevotella



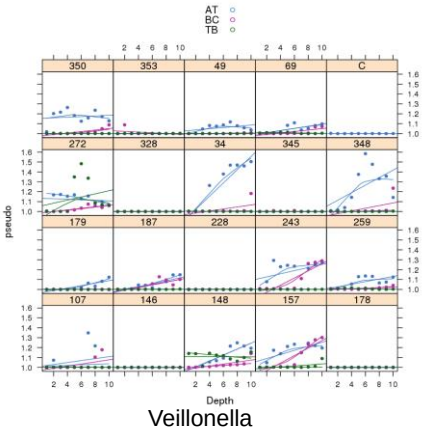
Pseudomonas



Stenotrophomonas



Streptococcus



Veillonella

Fig. S7. GLMM results for different bacterial genera on a per patient basis. The untreated, bicarbonate and tobramycin treated distributions through the WinCF column depth is plotted. AT = No treatment, BC = Bicarbonate, TB = Tobramycin

Tobramycin Penetration and Metabolism in the WinCF Columns. The antibiotic tobramycin was detected in the LC-MS/MS data and quantified using the MS_2 ion abundance in all samples. Tobramycin was consistently detected in the control samples with no sputum added and penetrated through 9 of the 10 mm layers with a decreasing abundance through the depth indicating that the compound was also present in a gradient of concentration without bacterial inoculation (fig. S8). Interestingly, while tobramycin was similarly present in a gradient when incubated with sputum, it was only sparsely detected and of low abundance in the WinCF columns. Tobramycin was detected in only 4 of the 19 WinCF columns after it was added and incubated with a sputum sample. The only samples that tobramycin was consistently detected in were those that had abundant reads mapping to the *Aspergillus* mitochondrion and this was in the same depth layers. Furthermore, there was a strong positive relationship between the abundance of *Aspergillus* reads in the column layers and tobramycin ion count (Pearson's $r = 0.577$, $p<0.001$). *Stenotrophomonas* was also positively correlated with tobramycin abundance (Pearson's $r = 0.352$, $p<0.001$). Tobramycin is known to be acetylated, phosphorylated or adenylylated as a mechanism of antimicrobial resistance (*23*). However, none of these tobramycin metabolites were detected in the LC-MS/MS dataset using the C-18 column method. Thus, we investigated samples from 6 patients, the 3 with abundant *Aspergillus* and 3 without with HILIC polar chromatography. Tobramycin was recovered in these samples and there was the presence of propionylation of the central sugar of tobramycin (fig. S9). This demonstrates that aminoglycoside modifying enzymes were modifying tobramycin in these experiments.

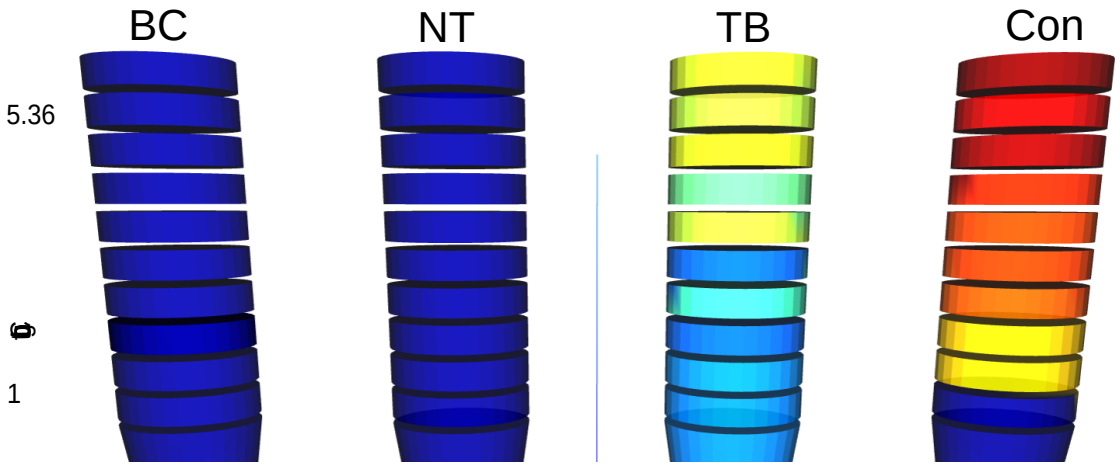
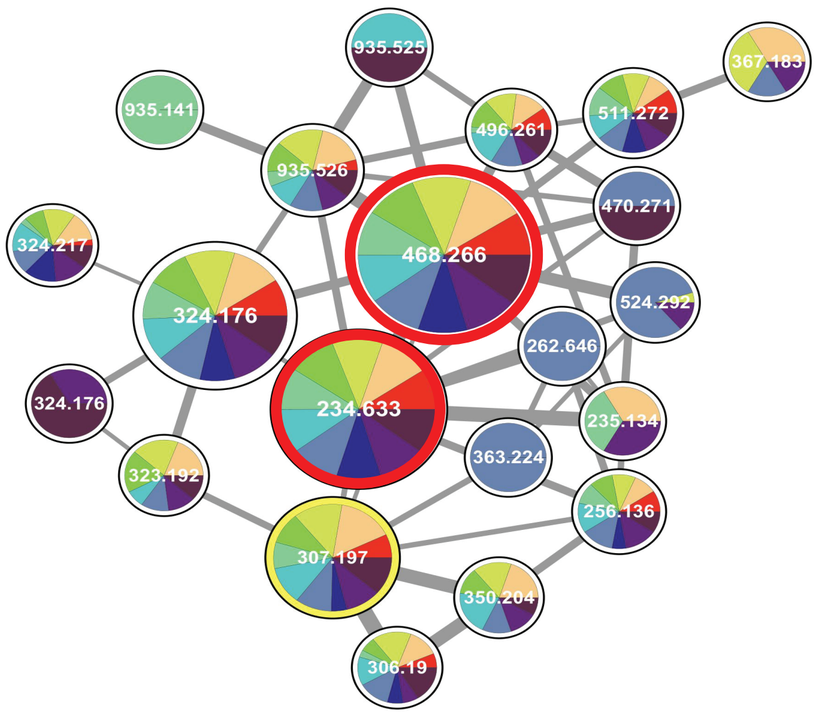
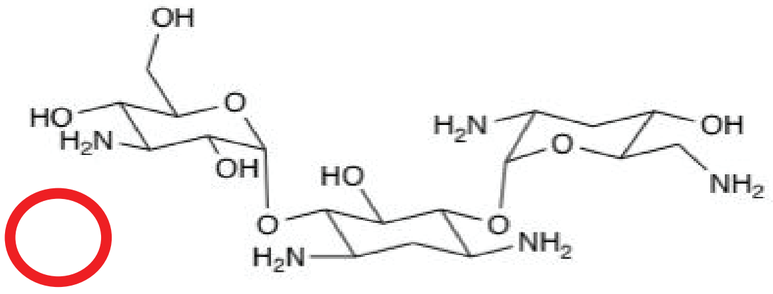


Fig. S8. Mean abundance of tobramycin by ion count in the WinCF columns after incubation. Tobramycin was added to the top of the tubes for the TB (tobramycin) and Con (uninoculated control) samples prior to incubation as described in the methods.

a)

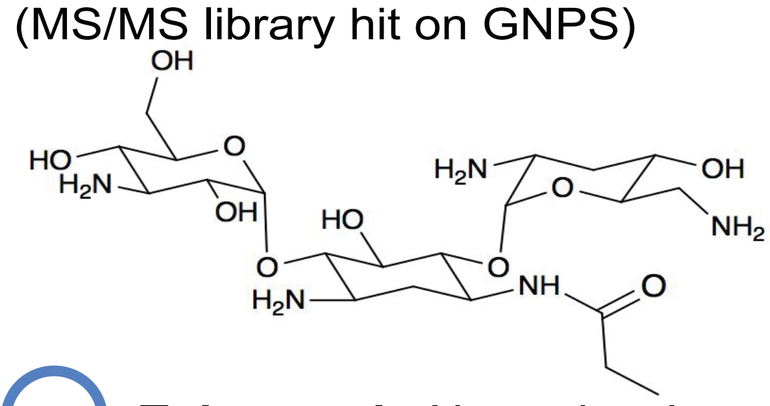
Molecular networking 1.0
[GNPS with MS-Cluster]





Tobramycin *m/z* 468.266 $[M+H]^+$
(MS/MS library hit on GNPS)

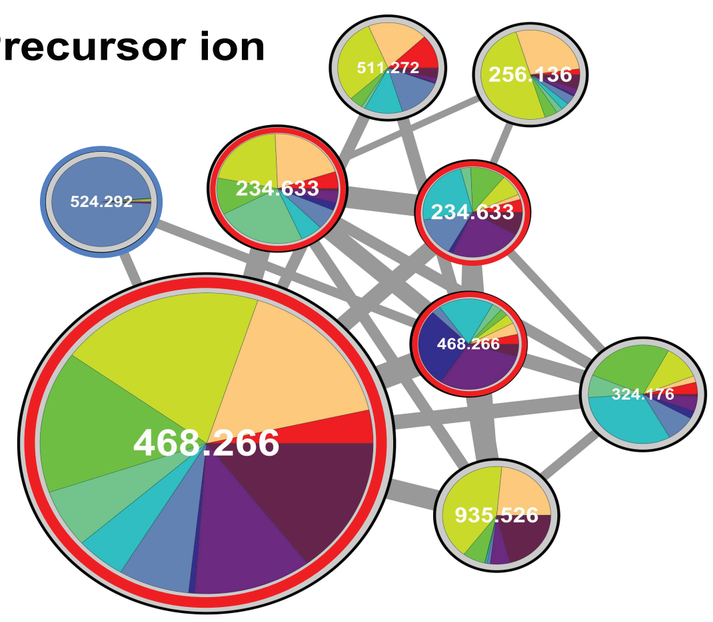
3(N-propionylation)



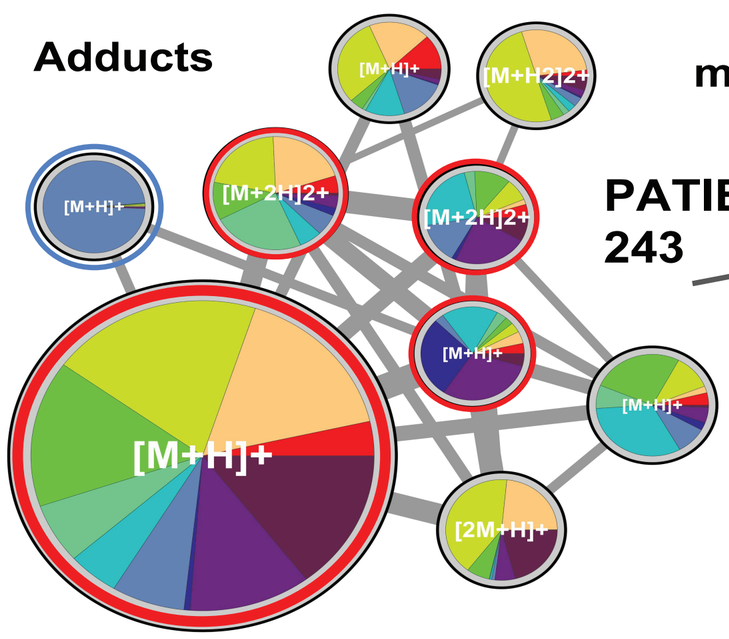
Tobramycin N-propionyl
m/z 524.292 $[M+H]^+$

Molecular networking 2.0 [MZmine2 and CAMERA]

Precursor ion

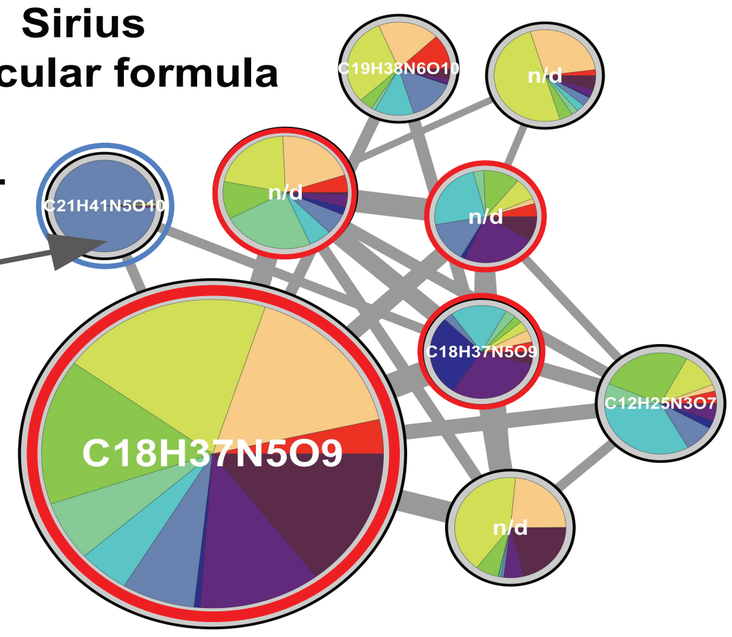


Adducts



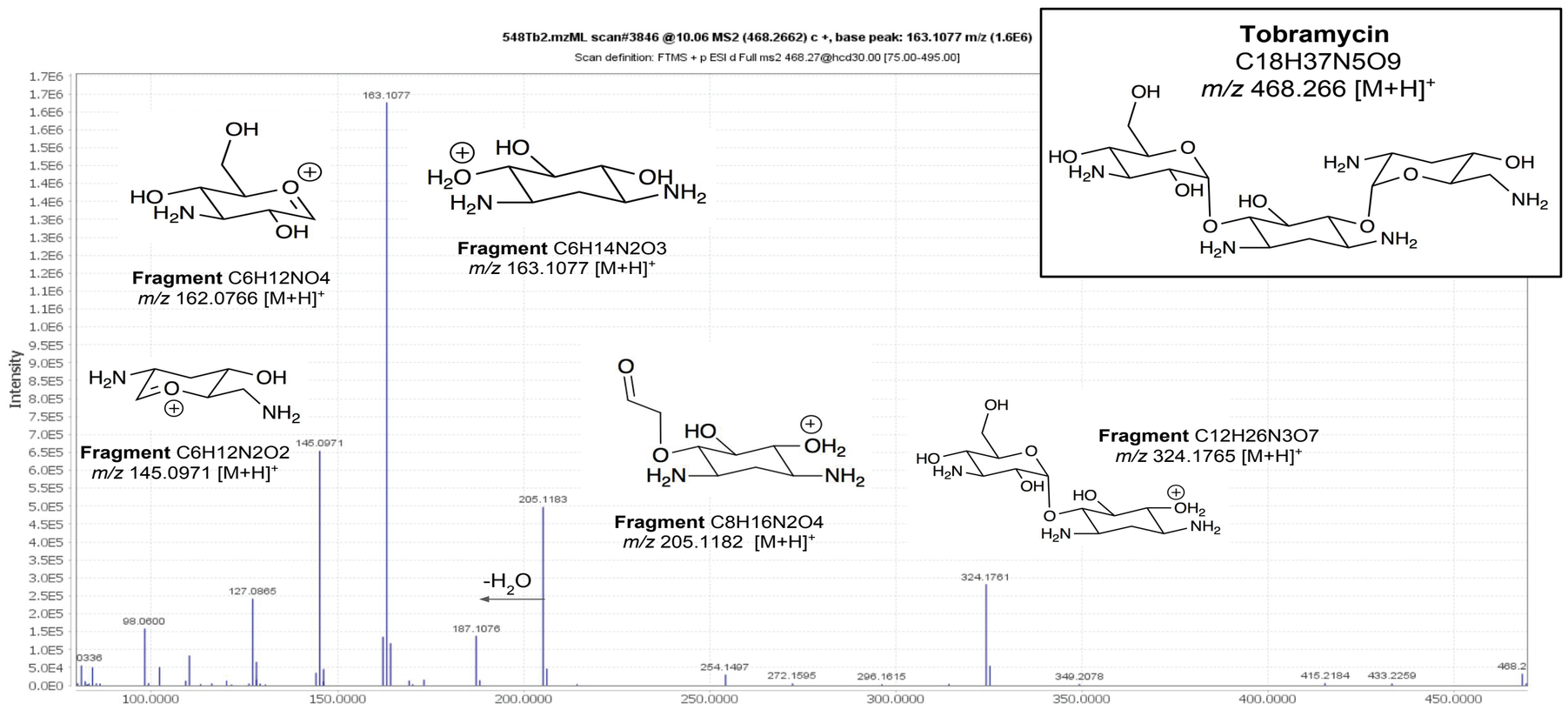
Sirius molecular formula

PATIENT 243



b)

Fragmentation spectrum of tobramycin (MS/MS of m/z468.266)



Fragmentation spectrum of the N-propionyl tobramycin (MS/MS of *m/z* 524.292)

c)

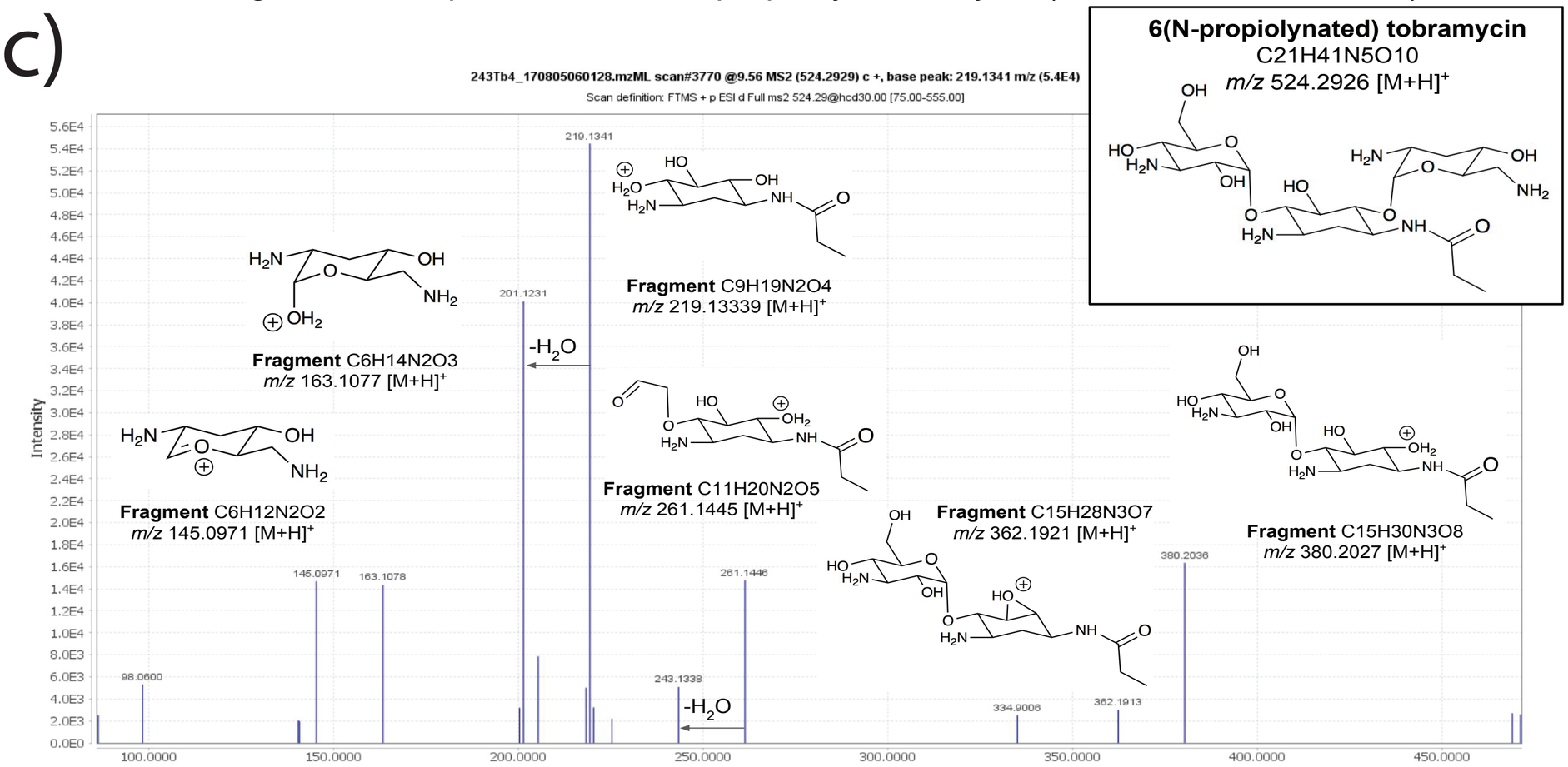
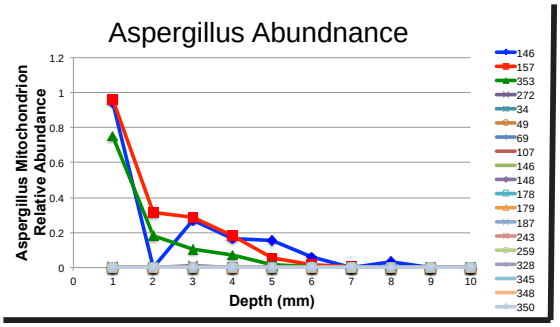


Fig. S9. Tobramycin and *N*-propionyl tobramycin identification from polar LC-MS/MS data. a) Molecular network of Tobramycin from CF patient 234 revealing annotated GNPS nodes and related metabolites. Structures of tobramycin and its propionylated from are shown. The result of the molecular networking 2.0 analysis showing adducts and the calculated Sirius molecular formula (see supplementary methods is also shown). MS/MS fragmentation spectra and annotated molecular structures of fragments from b) tobramycin and c) N-propionyl tobramycin. The location of the propionylation is shown on the middle saccharide of tobramycin, but the exact location cannot be determined from our mass spectrometry data. However, the current location is the most common location for acetylation.

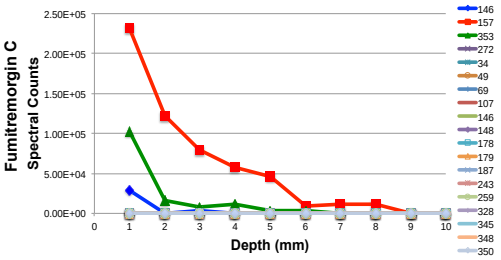


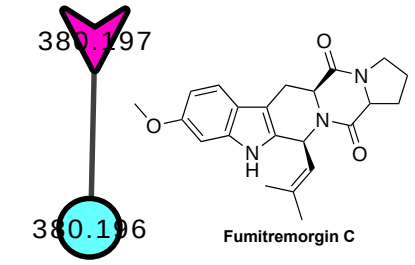
= Metabolite Found in WinCF and Aspergillus cultures

= Metabolite Found in WinCF only

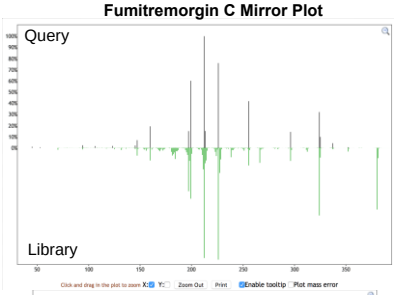
= Annotated metabolite through GNPS

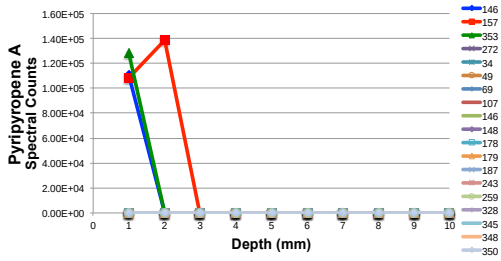
Aspergillus Metabolite Abundnance

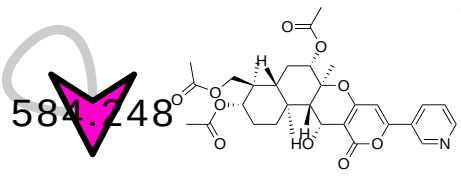


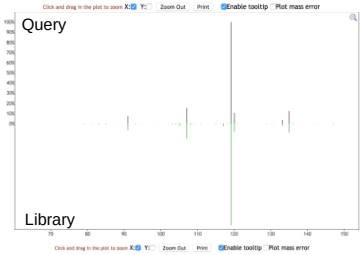


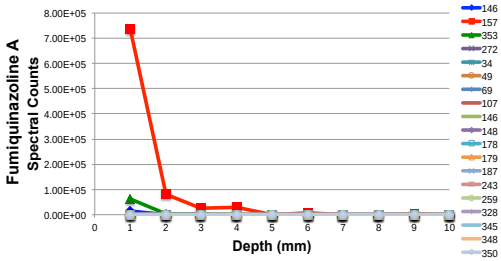
Fumitremorgin C Mirror Plot

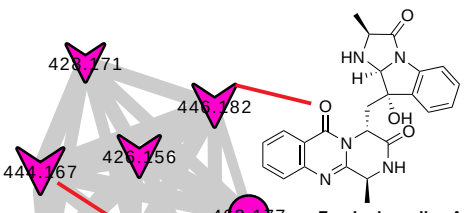




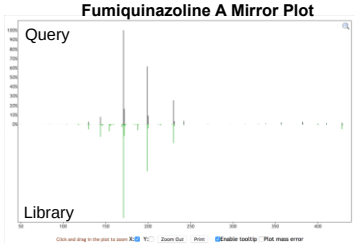


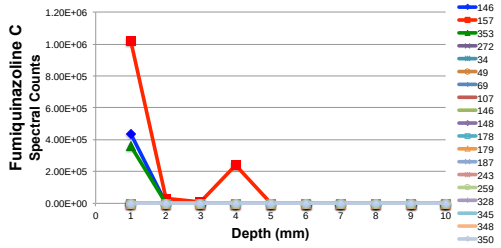


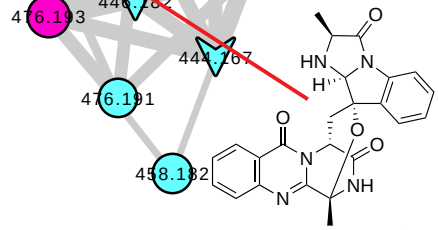




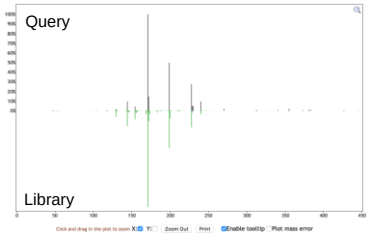
Fumiquinazoline A Mirror Plot

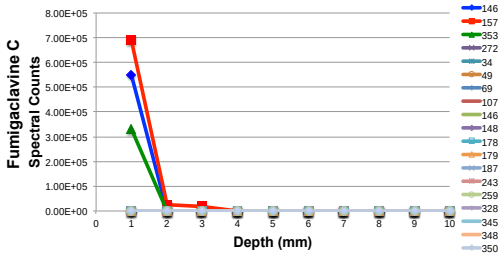


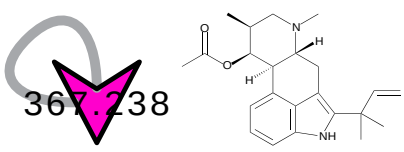




Fumiquinazoline F Mirror Plot







Fumigaclavin C Mirror Plot

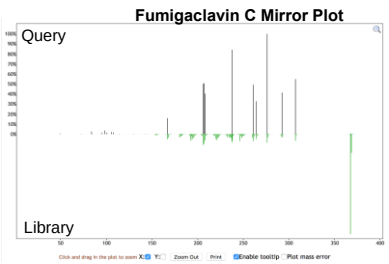


Fig. S10. *A. fumigatus* metabolites in tobramycin-treated WinCF columns. The abundance of reads to the *Aspergillus* mitochondrion and the spectral counts of four *A. fumigatus* metabolites are shown colored by the patient the sputum sample was collected from. The same patients that have the metabolites also have abundant reads to *Aspergillus*. Molecular clusters of *A. fumigatus* metabolites and the chemical structures of those detected in the WinCF columns.

Aspergillus fumigatus Metabolites in WinCF Columns. The LC-MS/MS data was searched through GNPS with molecular networking against a panel of *Aspergillus* cultures obtained from the UCSD CALM lab from CF patients. These included *A. flavus, A. niger, A. fumigatus, A. terreus* and unknown *Aspergillus* spp. These isolates were grown in ASM media and metabolites extracted using the same extraction protocol as the WinCF and CF sputum samples (see methods). The network was prepared such that the four patients that contained abundant *Aspergillus* in the WinCF sequencing data were separated from those without to aid in identification of metabolites unique to those samples and media blanks were included to remove background media. The molecular network is available here: http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=bc47b6a1d2f54cb6a8d2fff236a80b04
Five unique *A. fumigatus* metabolites were detected by GNPS library searching, which makes the annotation level 2 according to the metabolomics consortium minimum standards (*49*). These included fumigaclavin C, pyripyropene A, Fumitramorgin C and the fumiquinazolines A and C by MS/MS matching. These metabolites were uploaded to GNPS as part of the GNPS NIH Natural Products Library (round 2). These metabolites were also matched to MS/MS patterns available in the Metlin Database if available. The abundance of these metabolites through the WinCF gradient matches the abundance of *Aspergillus* mitochondrion reads and these metabolites are only detected in the three patients with *Aspergillus* reads. This is strong evidence for the growth of *A. fumigatus* in these WinCF columns.

WinCF pH Mathematical Modeling. We propose a simplified mathematical model to qualitatively describe the CF microbiome in the WinCF capillary tube environment. The model includes two microbial communities, distilled down to *Pseudomonas aeruginosa* and fermentative anaerobes, with their populations denoted by $\theta_p(x, t)$ and $\theta_f(x, t)$ respectively, where x and t are spatial and temporal variables. *P. aeruginosa*'s behavior in the CF lung is sourced from evidence in the literature (*18,25,50*). *P. aeruginosa* grows either aerobically or anaerobically, using oxygen (S_O, fast growth) or nitrate (S_N, slower growth) as electron acceptor respectively, consumes amino acid (S_A), and produces

ammonium (P) (a compound that has been detected in high levels in the WinCF capillary columns after growth of CF sputum (*11*)) and a chemical (I) which inhibits fermenters growth (this chemical is meant to represent its antimicrobial metabolites such as phenazines and rhamnolipids detected in the WinCF tubes). Fermenters consume sugar (S_G) and produces acid (F). pH is crudely represented by $F - P + C_0$, where C_0 describes the initial buffering pH (bigger C_0 means lower pH). Both microbial communities are assumed to have logistic growth with growth rate given by Monod-kinetics. The carrying capacity of *P. aeruginosa* is a decreasing function of $F - P + C_0$, reflecting that it grows better at higher pH. The growth rate of fermenters decreases with increasing oxygen concentration, and the carrying capacity of fermenters is a decreasing function of I, reflecting that the presence of *P. aeruginosa* inhibits its growth.

The effect of adding bicarbonate (T_B) and the anti-*P. aeruginosa* antibiotic tobramycin (T_T) treatments are also included in the model. The model assumes that bicarbonate raises the pH and tobramycin kills *P. aeruginosa*, but not fermenters. The chemical species concentrations are governed by reaction-diffusion equations with no-flux boundary conditions at all boundaries, except the oxygen has a fixed value at the top boundary. The bacteria populations satisfy ordinary differential equations (ODEs) in time describing growth and decay at a fixed spatial location. Detailed model equations are given in fig. S12. Simulation was done for a rectangular domain with width 0.1 cm and height 0.4 cm, the initial population is 0.1 for θ_f and 0.2 for θ_p everywhere. All the chemicals have constant initial values (with F and P have zero initial values since they are not produced initially). Two simulations were run with initial value of C_0 being 10 (low initial pH) and 0 (high initial pH) respectively. It must be noted that these pH values are a representative scale of acidity and not meant to reflect any actual pH measurement. Figure 4a) in the main text shows the contours of θ_p and θ_f at time t = 50 hours, when all quantities essentially reached steady states. In figure 4a, the left column is for low initial pH, and the right column is for high initial pH. It shows that in the low initial pH case, θ_f grows from 0.1 to about 0.37 in the anaerobic region and to about 0.12 in the aerobic region; in the high initial pH case, θ_f grows from 0.1 to about 0.33 in the anaerobic region and to about 0.12 in the aerobic region. In the low initial pH case, θ_p grows from 0.2 to about 0.5 in the anaerobic region and to about 0.64 in the aerobic region; in the high initial pH case, θ_p grows from 0.2 to about 0.44 in the anaerobic region and to about 1.81 in the aerobic region.

To determine what were the main drivers of the model outcome, the contours of θ_p and θ_f given by the simulation with the effect of inhibition chemical *I* turned off (set $\beta_1 = 0$ in the model), therefore *P. aeruginosa* has no effect on the growth of fermenters. In this case the fermenters grew well (to a peak

value around 0.75 which is larger than the peak value given in Fig. 4a in both low and high initial pH, and they produce enough acid to further inhibit the growth of *P. aeruginosa* in both low and high initial pH. Since pH still affects the growth of *P. aeruginosa*, θ_p has a slightly higher value in the high initial pH case, and it consumes more oxygen at the top layer of the tube, thus the growth of θ_f expands to a bigger region in the high initial pH case, which results to a smaller peak value of θ_f (at the lower part of the tube) since the total amount of sugar S_G is fixed. Actually the totally amount of θ_f are almost the same for the high and low pH case.

The contours of θ_p and θ_f given were also calculated after simulation with treatments of bicarbonate or tobramycin. Starting with homogeneous initial value 0.2, θ_p grows to a maximum value around 0.66 with no treatment (NT), a maximum value around 1.86 with bicarbonate treatment (BC) due to the pH increase, and a maximum value around 0.24 with tobramycin treatment (TB). In the model, these treatments did not affect fermenters directly. Comparing to the NT case, θ_f grows slower under bicarbonate treatment since more inhibition chemical I is produced due to better growth of θ_p, and grows even more slower under tobramycin treatment since the slow growth of θ_p allows deeper penetration of oxygen which inhibits the growth of θ_f.

To compare the results of the treatment effects in the model to the experimental data all known anaerobes in the microbiome data (minimum mean normalized abundance 0.002, which included *Streptococcus, Prevotella, Veillonella, Fusobacterium, Actinomyces, Leptotrichia, Porphyromonas, Enterococcus, Oribacterium, Gemella* and *Granulicatella*) were summed to get 'total anaerobic abundance'. This was plotted through the BC, TB and NT treatments using the *'ili* software (*48*) to visualize the effect of these treatments on the anaerobes collectively as they are treated in the model. Bicarbonate had little effect on the total anaerobic load, but tobramycin induced some reduction of anaerobic abundance in the upper layers of the WinCF column (fig. S11).

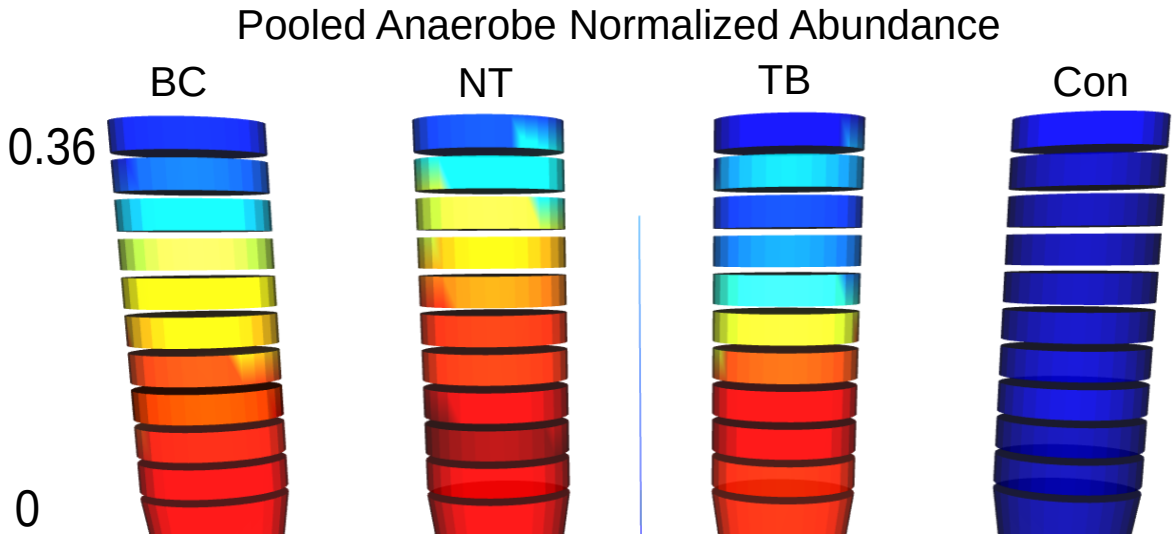


Fig. S11. Mean abundance of pooled anaerobes in the WinCF columns after the different treatments. BC = Bicarbonate, NT = No Treatment, TB = Tobramycin, Con = No Sputum

$$
\begin{aligned}
\frac{d\,\theta_f}{dt} &= g_0(S_O)\cdot\frac{S_G}{K_G+S_G}\cdot\theta_f\left(1-\frac{\theta_f}{g_1(I)}\right)\\
\frac{d\,\theta_p}{dt} &= \left(\mu_{pa}\frac{S_O}{K_O+S_O}+e^{-d_p\cdot S_O}\cdot\mu_{pn}\frac{S_N}{K_N+S_N}\right)\cdot\frac{S_A}{K_A+S_A}\cdot\theta_p\left(1-\frac{\theta_p}{g_2(pHA)}\right)\\
&\quad -d_T\cdot T_T\cdot\theta_p\\
\frac{\partial F}{dt} &= D_F\frac{\partial^2 F}{\partial x^2}+\frac{1}{Y_{fF}}\cdot g_0(S_O)\cdot\frac{S_G}{K_G+S_G}\cdot\theta_f\left(1-\frac{\theta_f}{g_1(I)}\right)\\
\frac{\partial P}{dt} &= D_P\frac{\partial^2 P}{\partial x^2}+\left(\frac{1}{Y_{pPo}}\cdot\mu_{pa}\frac{S_O}{K_O+S_O}+\frac{1}{Y_{pPn}}\cdot e^{-d_p\cdot S_O}\cdot\mu_{pn}\frac{S_N}{K_N+S_N}\right)\cdot\frac{S_A}{K_A+S_A}\cdot\\
&\quad \theta_p\left(1-\frac{\theta_p}{g_2(pHA)}\right)\\
\frac{\partial I}{dt} &= D_I\frac{\partial^2 I}{\partial x^2}+\left(\frac{1}{Y_{pIo}}\cdot\mu_{pa}\frac{S_O}{K_O+S_O}+\frac{1}{Y_{pIn}}\cdot e^{-d_p\cdot S_O}\cdot\mu_{pn}\frac{S_N}{K_N+S_N}\right)\cdot\frac{S_A}{K_A+S_A}\cdot\\
&\quad \theta_p\left(1-\frac{\theta_p}{g_2(pHA)}\right)\\
\frac{\partial S_O}{dt} &= D_O\frac{\partial^2 S_O}{\partial x^2}-\frac{1}{Y_{pO}}\cdot\mu_{pa}\cdot\frac{S_O}{K_O+S_O}\cdot\frac{S_A}{K_A+S_A}\cdot\theta_p\left(1-\frac{\theta_p}{g_2(pHA)}\right)-\frac{d_O}{Y_{pO}}\cdot\frac{S_O}{K_O+S_O}\cdot\theta_p\\
\frac{\partial S_N}{dt} &= D_N\frac{\partial^2 S_N}{\partial x^2}-\frac{1}{Y_{pN}}\cdot e^{-d_p\cdot S_O}\cdot\mu_{pn}\cdot\frac{S_N}{K_N+S_N}\cdot\frac{S_A}{K_A+S_A}\cdot\theta_p\left(1-\frac{\theta_p}{g_2(pHA)}\right)\\
\frac{\partial S_A}{dt} &= D_A\frac{\partial^2 S_A}{\partial x^2}-\frac{1}{Y_{pA}}\cdot\left(\mu_{pa}\frac{S_O}{K_O+S_O}+e^{-d_p\cdot S_O}\cdot\mu_{pn}\frac{S_N}{K_N+S_N}\right)\cdot\frac{S_A}{K_A+S_A}\cdot\\
&\quad \theta_p\left(1-\frac{\theta_p}{g_2(pHA)}\right)\\
\frac{\partial S_G}{dt} &= D_G\frac{\partial^2 S_G}{\partial x^2}-\frac{1}{Y_{fG}}\cdot g_0(S_O)\cdot\frac{S_G}{K_G+S_G}\cdot\theta_f\left(1-\frac{\theta_f}{g_1(I)}\right)\\
\frac{\partial T_B}{dt} &= D_B\frac{\partial^2 T_B}{\partial x^2}\\
\frac{\partial T_T}{dt} &= D_T\frac{\partial^2 T_T}{\partial x^2}-\frac{1}{Y_{pT}}\cdot d_T\cdot T_T\cdot\theta_p
\end{aligned}
$$

$$
g_0(x)=\mu_f\left[1-\frac{2}{\pi}\operatorname{atan}(\beta_0 x)\right],\quad g_1(x)=K_f\left[1-0.9\frac{2}{\pi}\operatorname{atan}(\beta_1 x)\right]
$$

$$
g_2(x)=K_p\left[0.9\left(\frac{1}{2}-\frac{1}{\pi}\operatorname{atan}(\beta_2 x)\right)+0.1\right],\quad pHA=F-P+C_0-T_B
$$

Fig. S12. WinCF model equations.

Global Analysis of Microbiome and Metabolomic Data. In the pH experiments, principle coordinate analysis (PCoA) of the weighted UniFrac distance from the microbiome data revealed that the CF and non-CF microbial profiles were distinct (fig. S13). The CF WinCF and sputum samples did not separate, but the non-CF samples had separation between capillary tubes and sputum (fig. S13). This

indicated that the CF microbiome from the sputum was well represented in the WinCF tubes and distinct from non-CF. The LC-MS/MS untargeted approach showed that the sample types were distinctly different in their chemical makeup as seen in a principle co-ordinate projection of a Bray-Curtis dissimilarity matrix (fig. S13). Sputum samples from CF and non-CF were distinct from WinCF metabolomes and each other. The WinCF media metabolomes from CF and non-CF samples were also distinct, although they showed more overlap between the sample types compared to sputum (fig. S13). For the CF samples, the WinCF and sputum samples did not cluster separately, indicting similar profiles between the capillary tubes and the sputum samples, the non-CF samples however, had tight clustering for the WinCF capillary microbiomes, which was distinct from the sputum (fig. S13). This indicated that the CF microbiome from the sputum was well represented in the WinCF tubes, distinct from non-CF and highly variable through the pH gradient.

In the oxygen experiments, the sputum samples were not as clearly distinct from the cultured communities in the metabolomics data and not distinguishable at all in the microbiome data (fig. S13). Within each treatment group the metabolomics and microbiome data varied greatly as the samples dispersed through the PCoA space. The metabolomics data showed clear separation between the tobramycin and bicarbonate treated samples from the untreated. The tobramycin treated samples separated in the microbiome data, but the bicarbonate and untreated samples were not separated in PCoA space.

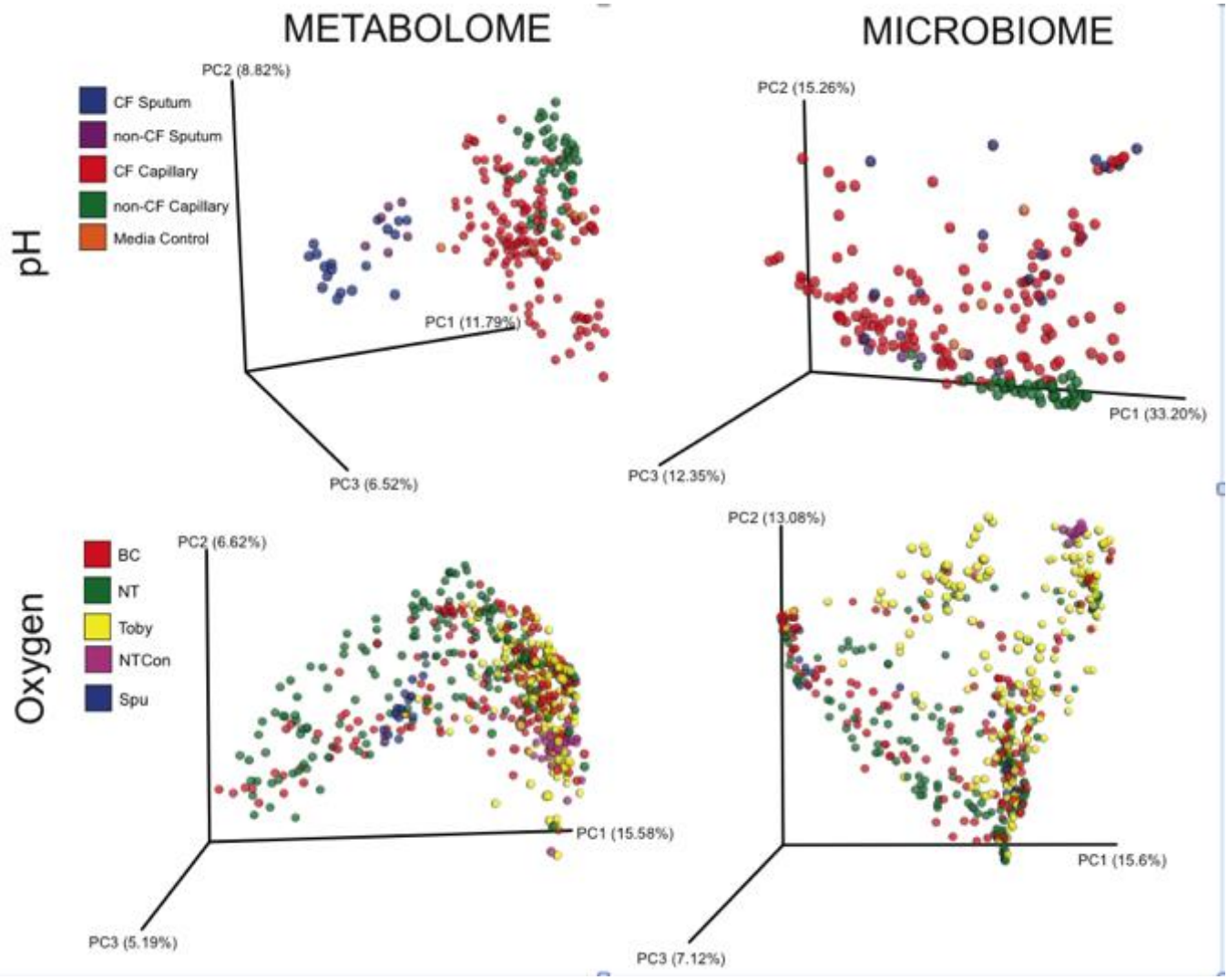


Fig. S13. PCoA plots of metabolome and microbiome data from all samples. The metabolomics plot was generated using the Bray-Curtis dissimilarity on the abundance of all metabolites in the LC-MS/MS untargeted data based on the MS^1 feature finding. The microbiome data was generated using the weighted UniFrac distance at the OTU level.

Supplementary Table Captions:

Table S1. Patient samples and information collected in this study.

Table S2. ANOVA of qualitative and quantitative variables measured during the WinCF pH gradient experiments.

Table S3. Metabolites that most changed with the WinCF gas production gradient according to an RF variable importance plot from the untreated samples. Metabolites are listed according to the mass followed by their retention time. Those that have a hit to GNPS libraries are listed next to these identifiers. MSE = Mean Squared Error

Table S4. Mean abundance through the depth gradient (1 to 10 mm) of *P. aeruginosa* virulence factor metabolites detected in the WinCF depth experiments and the corresponding Pearson's correlation (*r*).

Table S5. Deblurred OTUs and their sequences that most changed with the WinCF depth gradient according to an RF variable importance plot from the untreated samples. MSE = Mean Squared Error

Table S6. Results of the assessment of bias in WinCF system. All OTUs that had any instance of detection in a sputum sample, but not in the corresponding capillary tubes are listed for the pH and oxygen experiments. The number of patients where that specific instance of bias occurred is tabulated for each OTU in the pH and oxygen experiments.

Table S7. Confusion matrix and out-of-bag error from an RF classification of the pH experiment metabolomics data based on patient source.

Supplementary Methods

WinCF pH Experiments. Sputum samples from both CF and non-CF subjects were homogenized with a 3ml syringe without the needle and then diluted 1:1 in PBS upon return to the laboratory. After inoculation all capillary tubes were sealed at the bottom with capillary tube sealant (Fisher Scientific, USA) and placed into an incubation chamber with a paper towel moistened with sterile water to create a humid environment. After incubation the media was removed using a sterile blunt ended syringe, by piercing the plug sealant at the bottom of the tube and injecting the media into a 0.5 ml eppendorf tube using a 200 µl pipette tip. This methodological approach resulted in all 18 sputum smaples being grown in a gradient of 8 different pH intervals in triplicate. These triplicates were then pooled and analyzed with the different omics methods.

16S rRNA Gene Amplicon Sequencing. Genomic DNA was extracted using the PowerMag DNA isolation kit (MoBio Laboratories, Carlsbad, CA) and the V4 region was PCR amplified in triplicate from each sample and combined. The PCR mixtures contained 13 µl MoBio PCR water, 10 µl 5 Prime HotMasterMix, 0.5 µl both the barcoded forward and reverse primers (515f and 806rB; 10 µM final concentration), and 1.0 µl genomic DNA. Thermocycling consisted of ramping tot 94°C for 3 min (denaturation), with amplification proceeding for 35 cycles at 94°C for 45 s, 50°C for 60 s, and 72°C for 90 s, followed by a final extension for 10 min at 72°C. After amplification, the DNA concentration was quantified using PicoGreen double-stranded DNA (dsDNA) reagent in 10 mM Tris buffer (pH 8.0). A composite sample for sequencing was created by combining equimolar ratios of amplicons from the individual samples, followed by ethanol precipitation to remove any remaining contaminants and PCR artifacts.

LC-MS/MS Reverse Phase (RP) Analysis. A volume of 200 µl of LC-MS grade ethyl acetate was first added to the samples and incubated at room temperature for two hours, the ethyl acetate layer was then removed and evaporated using a centrifugal evaporator. The same volume of LC-MS grade methanol was then added to the remaining sample and incubated for another 2 hours. The extract and sample preparation was then spun in a tabletop centrifuge at 10,000 x g for 30 s to pellet debris and the supernatant was added to the ethyl acetate extract and dried down again.
The mass spectrometer was tuned using Tuning Mix ES-TOF (Agilent Technologies) at a 3 mL min^{-1} flow rate once a day during the run. For accurate mass measurements, lock mass internal calibration used a wick saturated with hexakis (1H,1H,3H-tetrafluoropropoxy) phosphazene ions (Synquest Laboratories, *m/z* 922.0098) located within the source. Full scan MS spectra (*m/z* 50 – 2000) were acquired in the qTOF and the top ten most intense ions in a particular scan were fragmented using collision induced dissociation at 35 eV for +1 ions and 25 eV for +2 ions in the collision cell. Automatic exclusion was used such that an ion was fragmented upon its first detection, then fragmented twice more, but not again unless its intensity was 2.5x the previous fragmentation. This exclusion method was removed after 30 seconds and the mass spectrometer would repeat its ion detection and automatic exclusion cycle.

LC-MS/MS HILIC Analysis. The methanol metabolite extracts from three patients for the oxygen experiments with *Aspergillus* reads (CF157, CF146 and CF353) and three without (CF178, CF243,

CF69) were additionally used for the HILIC chromatography analysis to more efficiently detect tobramycin and its metabolites. These samples were extracted further using a water/methanol 4:1 extract solution for 5 min, followed by centrifugation to pellet debris prior to mass spectrometry analysis. The analysis was conducted on a Vanquish UHPLC (Thermo Fisher Scientific, Waltham, MA) connected to Orbitrap (Q Exactive, Thermo Fisher Scientific, Waltham, MA) mass spectrometer equipped with HESI-II probe source. The separation was conducted using a 100 x 2.1 mm Kinetex 1.7 µM, HILIC, 100Å column (Phenomenex, Torrance, CA). The column was held at 40°C during the analysis. The mobile phases used were: A 99.9 % HPLC grade water 0.1% formic acid (Thermo Fisher Scientific, Optima LC/MS), and B 99.9% HPLC grade acetonitrile 0.1% formic acid (Fisher Scientific, Optima LC/MS). Following gradient steps were used: 0-1 min 100% B, 1-4 min 100-90% B, 4-14 min 90-0% B, 14.0-14.9 min 100% A, 14.9-17.0 min 100% B. Flow rate was set at 0.350 mL/min. For the MS acquisition, following settings were used: positive ion mode, Spray voltage of 3500 V, ion source temperature of 270°C, Capillary temperature 250°C, S-lens RF level of 50 Hz, Sheath gas (N_2) pressure of 45 psi, Auxiliary gas pressure (N2) of 10 psi, and Aux gas heater temp. at 410°C. 5 µl of the sample was injected. The data were acquired in a data dependent acquisition (DDA) mode with both MS1 full scan (150-1500 *m*/*z*) and up to 5 MS^2 scans of the most abundant ions per duty cycle. The resolution of Orbitrap mass analyzer was set at 30,000. The MS^2 precursor selection window was set to 1.5 m/z with 0.5 m/z offset. The normalized collision energy was set to stepwise increase from 20 to 30 to 40 units with z = 1 as default charge state. The MS^2 acquisition was set to be automatically triggered at the apex of a peak within 2 to 15 s from their first occurrence with the dynamic exclusion time of 5 s. The data were uploaded on MassIVE (MSV000081456). http://massive.ucsd.edu/ProteoSAFe/status.jsp?task=56f6d2f53e774297a84df347175dfbb8

GC-MS Protocol. The GC protocol analysis included: cryofocusing on the head of the column at -10 °C for 1.25 min; 100 °C/min oven ramp to 40 °C (hold of 0.1 min), 15 °C/min oven ramp to 280 °C (hold of 0.1 min), and a 3 min hold period to purge the column. The helium carrier gas was set to constant 2 mL/min flow, splitless injection mode was used throughout. The scanned m/z range in a single quadrupole was 35-350 Th. The empty vial blanks were interspersed with the samples. Quality controls of natural mint oil extract were run along with samples throughout the analysis to monitor instrument performance and SPME wear. The data was normalized to the total ion abundance of each sample. The data were uploaded on MassIVE (MSV000081455).

Feature Finding for RP LC-MS/MS Data. The parameters of the Optimus run were set as default expect for the following: *m*/*z* tolerance 15.0 ppm, noise threshold 1000, retention time tolerance 20 s, intensity factor compared to blanks 2.0, and a feature observation rate 0.01. The metabolome was trimmed by removing data from the first 60 s and between 600 s and 840 s of the chromatographic run (where most contaminants eluted) and then normalized to total ion current abundance of each sample). The data were uploaded on MassIVE (MSV000081456)

Feature Finding Parameters for the HILIC LC-MS/MS Data. The parameters for MZmine2 (*56*) (v2.29) were as follow: noise level MS^1 (1,000); noise level MS^2 (100); ADAP chromatogram builder (min group size: 5; group intensity threshold 50,000; min highest intensity 20,000; *m*/*z* tolerance 7.5 ppm); Chromatogram deconvolution module (Local minimum search; chromatographic threshold: 0.05; search minimum in RT range: 0.05 min; minimum absolute height: 100,000; min ratio of peak top/edge: 1.5: peak duration: 0-4 min); Camera module CITE CAMERA (default parameters; and *m*/*z* tolerance:7.5 ppm); Join aligner (*m*/*z* tolerance of 7.5 ppm; and retention time absolute tolerance:0.5 min); Peak row filter (MS/MS filter; reset row ID); Remove duplicate filter (retention time 2.0 min; and *m*/*z* tolerance of 7.5 ppm); Peak finder module (minimum absolute intensity of 0.05; retention time tolerance 1.5 min; and *m*/*z* tolerance of 7.5 ppm).

Feature Finding for GC-MS Data. The data were processed with MZmine2 using ADAP algorithm (version ADAP-in-MZmine2, https://github.com/du-lab/ADAP-in-MZmine2) , and parameters were set as follow: Crop filter (0.35-25.0 min; m/z 20-350), noise level MS1 (10,000); ADAP chromatogram builder (min group size: 6; group intensity threshold 20,000; min highest intensity 20,000; *m*/*z* tolerance 0.45 Da); Smoothing (Filter width: 5); Chromatogram deconvolution module (Wavelets ADAP: S/N threshold: 0.5; Peak width multiplicity 1.0; abs(wavelet coeffs): true; min feature height: 100,000; coefficient area threshold 1.0; Peak duration range: 0.005-0.7; RT wavelet range: 0.0001-0.04); ADAP decomposition module (Min cluster distance: 0.001 min; Min cluster size: 3; Min cluster intensity: 100; Find shared peaks: false; Min edge-to-height ratio: 0.2; Min sharpness: 10.0; Shape-similarity tolerance: 90; Choice of Model Peak based Sharpness); ADAP aligner module (Min confidence:0.3; Retention time range 0.125 min; Score tolerance 0.3; Score weight: 0.3; EIC score: retention time); Peak finder module (minimum absolute intensity of 0.05; retention time tolerance 1.5 min; and *m*/*z* tolerance of 0.45 Da). Data were uploaded to GNPS and searched against NIST and WILEY spectral library.

Aspergillus Metabolite Detection. Seventy-one strains of *Aspergillus* fungi isolated from CF patients at the UCSD CALM laboratory were cultured in ASM media, these included *A. flavus, A. niger, A. fumigatus, A. terreus* and unknown *Aspergillus* spp. These samples were then extracted using ethyl acetate and methanol with the same procedures as the sputum samples. The extract data was then converted and uploaded to GNPS as described and a network was built where the WinCF samples with tobramycin recovery were separated from samples without tobramycin after growth and networked with this Aspergillus dataset. Nodes mapping the tobramycin recovery group with those of *Aspergillus* spp. were analyzed for GNPS annotations. This led to the identification of fumigaclavin C, pyripyropene A, Brevianamide F (detected only in WinCF sample CF157), Fumitramorgin C and the fumiquinazolines A by MS/MS matching.