1   **Title:** The Molecular and Microbial Microenvironments in Chronically Diseased Lungs

2

3   **Authors:** Alexey V. Melnik[1], Yoshiki Vázquez-Baeza[2*], Alexander A. Aksenov[1*], Embriette
4   Hyde[3], Andrew C McAvoy[4], Mingxun Wang[5], Ricardo R. da Silva[1], Ivan Protsyuk[6], Jason V.
5   Wu[1], Amina Bouslimani[1], Yan Wei Lim[7], Tal Luzzatto-Knaan[1], William Comstock[1], Robert A.
6   Quinn[1], Richard Wong[8], Greg Humphrey[3], Gail Ackermann[3], Timothy Spivey[9], Sharon S.
7   Brouha[10], Nuno Bandeira[5], Grace Y. Lin[8], Forest Rohwer[7], Douglas J. Conrad[10], Theodore
8   Alexandrov[1,6,10], Rob Knight[3,5,11,12], Pieter C. Dorrestein[1], Neha Garg[1,4,13,14, #]

9

10   **Author Affiliations:**
11   [1] Collaborative Mass Spectrometry Innovation Center, Skaggs School of Pharmacy and
12   Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 92093, USA.
13   [2] Jacobs School of Engineering, University of California, San Diego, La Jolla, CA 92093, USA.
14   [3] Department of Pediatrics, University of California, San Diego, La Jolla, CA 92093, USA.
15   Currently, Freelance Science Writer and Research Consultant.
16   [4] School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332,
17   USA.
18   [5] Department of Computer Science & Engineering, University of California, San Diego, La Jolla,
19   CA 92093, USA.
20   [6] Structural and Computational Biology Unit, European Molecular Biology Laboratory,
21   Heidelberg 69117, Germany.
22   [7] Biology Department, San Diego State University, San Diego, CA, USA.
23   [8] Department of Pathology, University of California, San Diego, La Jolla, CA 92093, USA.
24   [9] Department of Radiology, University of California, San Diego, La Jolla, CA 92093, USA.
25   [10] Department of Medicine, University of California at San Diego, La Jolla, CA 92093, USA.
26   [11] UC San Diego Center for Microbiome Innovation, University of California, San Diego, La Jolla,
27   CA 92093, USA
28   [12] Department of Bioengineering, University of California at San Diego, La Jolla, CA 92093,
29   USA.
30   [13] Emory-Children's Center for Cystic Fibrosis and Airways Disease Research, Atlanta, GA
31   30322, USA.
32   [14] Center for Microbial Dynamics and Infection, Georgia Institute of Technology, Atlanta, GA
33   30332, USA.
34   [#] To whom correspondence should be addressed: ngarg42@gatech.edu
35   [*] Equal contribution authors

36

37   **Author contributions:**
38   NG, DJC, FR, PCD created the idea for the work.
39   NG, RW, AB, TLK, YW, RAQ, GYL, DJC collected lung tissue
40   AM, NG, TS, SSB created 3D models
41   AM, NG picked coordinates on lung models
42   AM, RS, YVB performed statistical analysis
43   AM, EH, NG, GH, JG, GA, RK, AA, YVB performed 16S rRNA sequencing and analysis
44   MW, IP, RS, TA modified the tool `ili
45   NG, JW extracted lung tissue for acquisition of mass spectrometry data
46   NG acquired mass spectrometry data
47   NG, WC performed microbiology experiments
48   AM, ACM, NG, IP, PCD analyzed mass spectrometry data
49   AM, NG, AAA, PCD wrote the paper

Abstract

To visualize the personalized distributions of pathogens, chemical environments including microbial metabolites, pharmaceuticals, and their metabolic products within and between human lungs afflicted with cystic fibrosis, we generated 3D microbiome and metabolome maps of six explanted lungs from three cystic fibrosis patients. These 3D spatial maps revealed that the chemical environments are variable between patients and within the lungs of each patient. Although the patients' microbial ecosystems were defined by the dominant pathogen, their chemical diversity was not. Additionally, the chemical diversity between locales in lungs of the same individual sometimes exceeded inter-individual variation. Thus, the chemistry and microbiome of the explanted lungs appear to be not only personalized but also regiospecific. Previously undescribed analogs of microbial quinolones and antibiotic metabolites were also detected. Furthermore, mapping the chemical and microbial distributions allowed visualization of microbial community interactions, such as increased production of quorum sensing quinolones in locations where *Pseudomonas* was in contact with *Staphylococcus* and *Granulicatella*, consistent with *in vitro* observations of bacteria isolated from these patients. Visualization of microbe-metabolite associations within a host organ in early-stage CF disease in animal models will help elucidate a complex interplay between the presence of a given microbial structure, antibiotics, metabolism of antibiotics, microbial virulence factors, and host responses.

Importance

Microbial infections are now recognized to be polymicrobial and personalized in nature. A comprehensive analysis and understanding of the factors underlying the polymicrobial and personalized nature of infections remains limited, especially in the context of the host. By visualizing microbiomes and metabolomes of diseased human lungs, we describe how different the chemical environments are between hosts that are dominated by the same pathogen and

76   how community interactions shape the chemical environment, or vice versa. We highlight that

77   three-dimensional organ mapping are hypothesis building tools that allow us to design

78   mechanistic studies aimed at addressing microbial responses to other microbes, the host, and

79   pharmaceutical drugs.

80   **Introduction**

81   An increasing rate of infection from multi-drug resistant opportunistic pathogens has

82   become a significant burden in recent years. Proliferation of these pathogens due to overuse of

83   antibiotics, including those of last resort (1-4), is a threat to human health and is already

84   associated with increased mortality (5, 6). One reason for indiscriminate use of broad-spectrum

85   antibiotics and combination therapy in complex polymicrobial infections is the lack of knowledge

86   with regards to how microorganisms interact with each other, the host, and their chemical

87   environment, leading to strategies that target bacterial pathogens broadly. Herein, specific

88   microbial pathways that are involved in detrimental microbe-microbe interactions (7), microbe-

89   host interactions (8) and microbe-drug interactions (9) can serve as new targets for targeted

90   drug discovery. Thus, knowledge of such interaction-mediating microbial pathways and their

91   prevalence will shape the future of drug discovery. In this regard, even though we have begun

92   to appreciate the presence of multiple subpopulations by imaging community structures (10-12)

93   and by genome sequencing (13-15), information about the specific microbial pathways in

94   mediating the above mentioned interactions, molecular distribution of xenobiotic compounds,

95   and how such distributions are associated with specific microbial structures within the context of

96   a host is largely lacking.

97   We developed a methodology to map microbial and metabolite distributions in a human

98   lung in three dimensions to identify pathways that may be mediating microbial interactions and

99   to visualize the distribution of antibiotics in relation to microbial community structure (16). These

100   three dimensional organ maps allow visualization of chemical and microbial microenvironments

101   and consequently, may provide better insights into complex processes that take place within a

102   host. Here we apply this methodology to elucidate spatial variation within and between the lungs

103   of three individuals afflicted with cystic fibrosis (CF).

104   CF is genetic disease caused by a mutation in the cystic fibrosis transmembrane

105   conductance regulator (CFTR) gene that results in defects of the encoded CFTR protein. The

106   primary function of CFTR protein is an ion channel that regulates liquid volume (mucus) on the

107   epithelial cells through secretion of chloride ions and inhibition of sodium absorption. Sticky

108   mucus accumulates in the upper airways and lungs of CF patients, and serves as growth

109   medium for various microbes, including opportunistic pathogens, resulting in chronic and

110   recurrent polymicrobial infections. In the 1930s, the life expectancy for someone diagnosed with

111   CF was only several years (17). Due to advances in modern medicine, including the use of

112   antibiotics and better clinical management of the disease, individuals with CF can now expect to

113   live on an average into their forties even though most patients are waitlisted for organ transplant

114   by the time they reach adulthood (18). Improved clinical management is partly made possible by

115   better understanding of the polymicrobial nature of the infections of the lung and development of

116   antibiotic-based clearance of chronic infections targeting the polymicrobial community (19).

117   However, the virulence of pathogens in microbial lung diseases such as CF, pneumonia,

118   tuberculosis, and chronic obstructive pulmonary disease is mostly studied in cultures derived

119   from pulmonary secretions and by genome sequencing, which does not represent complex *in*

120   *vivo* conditions. The failure in treating an infection in a complex organ such as human lung may

121   simply stem from the inability to treat a localized infection foci which can then spread to the

122   entire organ or become systemic as, for example, in case of infections caused by *Burkholderia*

123   (13, 14, 20, 21). Understanding how the production of microbial small molecules involved in

124   pathogenicity and community interactions varies with lung biogeography leading to infection hot-

125   spots will enable the development of targeted antimicrobials and improved drug delivery

126   vehicles (14, 20, 22). Thus, CF presents an important test case for improving disease

127   management strategies for polymicrobial infections, given better understanding of community

128   structures and chemical environments within the host.

129   In this study, with patient consent, we mapped the chemical and microbial makeup of six

130   explanted lungs, removed during surgery from three CF patients, by using 3D volume

131   cartography to understand how microbes, microbial molecules, and medications are distributed

132   and metabolized throughout the entire organ, providing insights into microbe-microbe

133   interactions.

134
135   **Results and discussion**

136   The explanted lungs of three patients afflicted with CF were sectioned to inventory and

137   map the associated microbiome and metabolome in three dimensions onto lung models built

138   from CT-scans acquired prior to surgery (Methods) (16). To perform 16S rRNA gene analysis,

139   the tissue sections were swabbed, enabling detailed inventory of bacterial DNA present within

140   the patients' lungs (Table 1). We refer to our analysis of 16S rRNA gene as inventory of the

141   bacterial DNA and not the bacteria themselves, since lungs associated with CF are known to

142   contain a significant amount of DNA from dead cells as well as extracellular DNA (23). In total,

143   six lungs from three patients contained bacteria that spanned 40 genera (Supplementary Table

144   1). Bar plots of the most frequently amplified genera and their relative abundances pooled for all

145   anatomical locations are illustrated for each patient in Supplementary Figure 1. The relative

146   abundances of these genera in individual sections of each patient is available through 3D maps

147   (see below). The DNA of the most commonly occurring pathogenic organism in CF,

148   *Pseudomonas aeruginosa*, was detected at highest frequency throughout the lungs of patients 1

149   and 3, whereas patient 2 was dominated by DNA from the emerging pathogen

150   *Stenotrophomonas*. Even though the microbial population within CF-associated lungs can be

151   heterogeneous (24), dominance of a single pathogen at end-stage CF disease has been

152   described extensively in previous studies (25-27).

153   The PCoA analysis of microbiome data with weighted UniFrac distance showed

154   clustering between both lungs of patient 1 and the right lung of patient 3 along the first two

155   principal axes (Supplementary Figure 2a). Samples from the left lung of patient 3 clustered

156   separately and the comparison of 10-most abundant OTUs further highlighted the differences

157   between the microbial communities present in the left lung of patient 3 (Supplementary Figure

158   2b and Supplementary Figure 3). Apart from the dominant pathogen, the overall microbiome

159   between and within patients was different along the second and third axis (Supplementary

160   Figure 2 c,d). The Canberra distance metric yielded less apparent but visible separation of

161   patients' microbiome data in PCoA space (Figure 1 a,b); in some cases this variation within

162   patient's own lungs was found to be greater than between patients (Figure 1b). A more

163   homogeneous clustering patterns is observed in the 16S data when the unweighted UniFrac

164   metric, a phylogenetically informed metric, was applied (Supplementary Figure 4). The

165   differences arise due to the qualitative nature of unweighted UniFrac and the quantitative

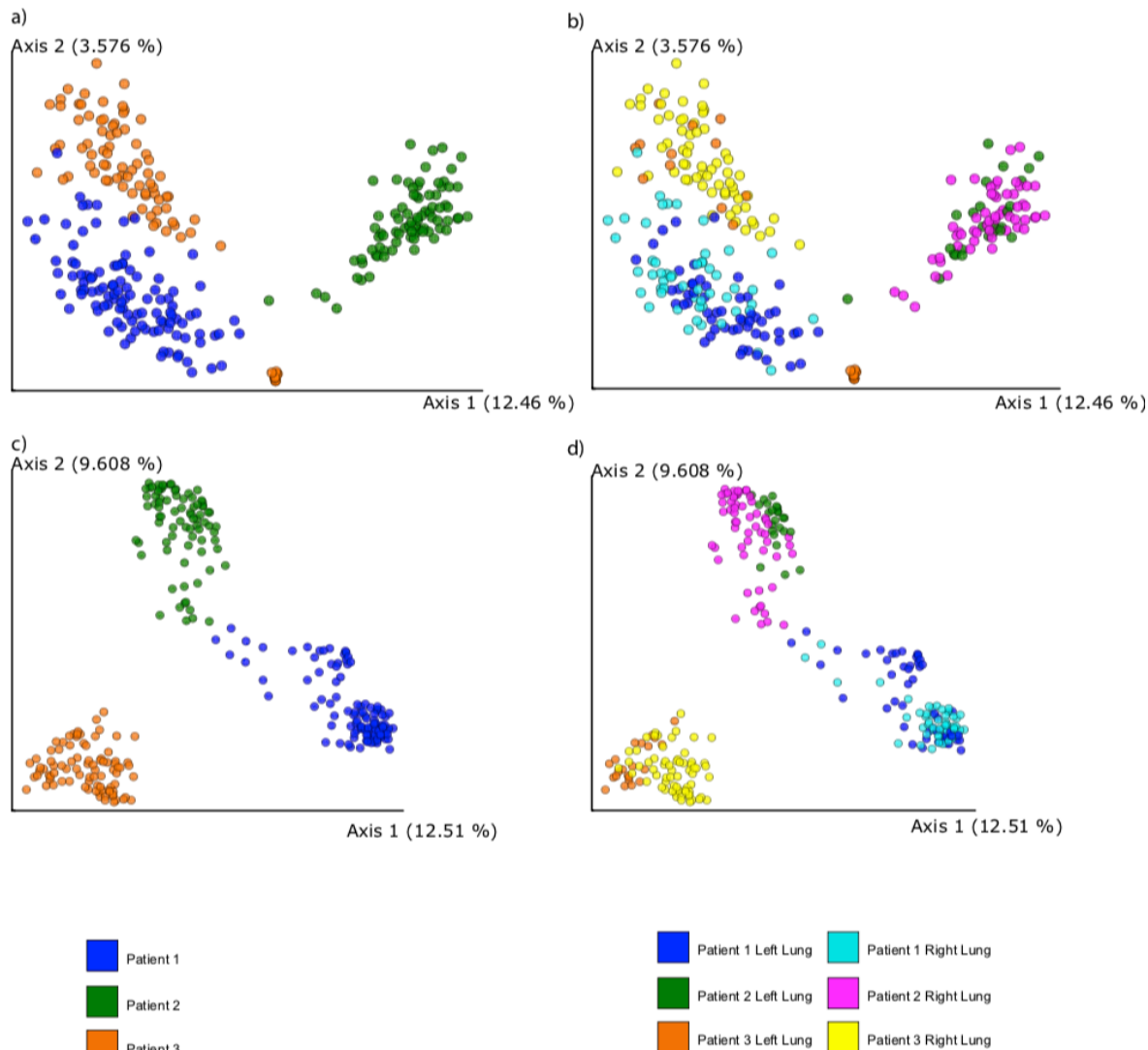166   comparisons possible with the Canberra metric.

Figure 1. Principal coordinate plot of metabolome and microbiome from lungs of three patients in the study. Visualization was performed in the Emperor software using Canberra distance metric (28). a), b) PCoA plots of 16S rRNA sequencing and c), d) mass spectrometry data, respectively.

To map the relative frequency of microbes onto the 3D lung models, we used our previously described methodology (16). The distribution of prevalent (*Pseudomonas* and *Staphylococcus*) and emerging (*Stenotrophomonas* and *Achromobacter*) microbes in the CF-associated lungs is displayed in Figure 2. Although nearly uniform distribution of dominant pathogens (*Pseudomonas* in patients 1 and 3, *Stenotrophomonas* in patient 2) was observed, all other microbes were distributed unevenly, often relegated to niche spots. For example, *Achromobacter* was mainly localized in the apex of the right lung of patient 3 while

7

180   *Staphylococcus* was present in the lower lobe of both lungs of patient 1, at the apex of the lungs

181   of patient 2 and in the middle and lower lobes of the lungs of patient 3. The dominant pathogen,

182   *Stenotrophomonas*, showed uniform distribution in the lungs of the patient 2 and differential

183   distributions in the lungs of patients 1 and 3 (Figure 2). A degree of stratification is expected

184   based on the availability of oxygen: *Achromobacter* and *Stenotrophomonas* are strict aerobes

185   whereas *Staphylococcus* and *Pseudomonas* are facultative anaerobes residing as biofilms in

186   airway mucus of CF patients with the potential of undergoing anaerobic metabolism (29).

187   Furthermore, patients 1 and 3 not only share the dominant pathogen, their microbial

188   communities are also more similar to each other than either is to that of the patient 2

189   (Supplementary Figure 3 and 5). Selection pressures from competing microbes and chemical

190   microenvironments including antibiotic distributions, further leads to stratification of niches

191   occupied by specific organisms. To compare the microbial and chemical environments, we next

192   annotated the mass spectrometry data acquired on the tissue sections and mapped it onto the

193   3D models of the lungs of these patients (see below).
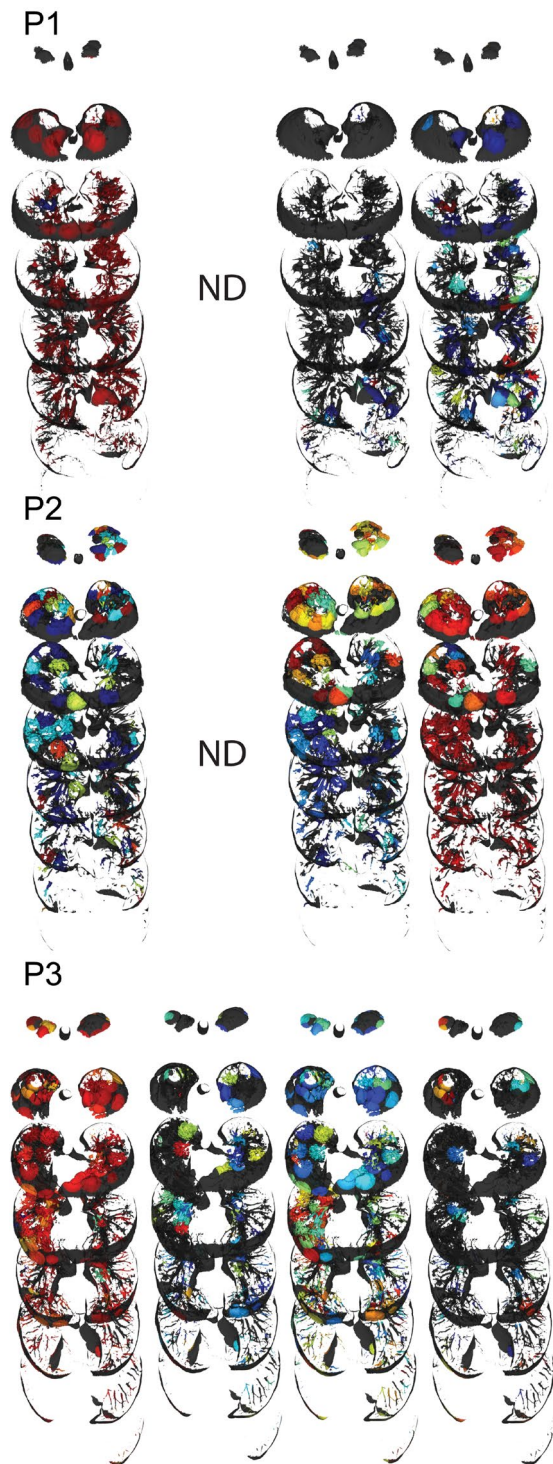
Figure 2. Distribution of microorganisms. Distribution of *Pseudomonas, Achromobacter*, *Staphylococcus*, and *Stenotrophomonas* (right to left) are shown for all three patients. P1 - patient one, P2 - patient two and P3 - patient three, ND - not detected. Intensity scale is provided at the bottom right. Full visualizations of microbial maps can be accessed via links: patient 1, patient 2 and patient 3.

200   To annotate molecular ions detected using high resolution mass spectrometry (MS)-

201   based untargeted approach, molecular network analysis was performed using the Global

202   Natural Product Social Molecular Networking (GNPS) infrastructure (30). Molecular networking

203   allows for reduction and organization of the overwhelming amount of chemical information

204   generated (in terms of mass spectra) in a high-resolution untargeted MS approach. The data

205   reduction is performed by combining and displaying the identical MS/MS spectra as a single

206   node and by displaying similar spectra as connected nodes (30, 31). Similarities in MS/MS

207   spectra relate to similarities in chemical structures, so oftentimes such connected nodes

208   represent chemical and biological transformations of a molecule. In this study, 676,451 MS/MS

209   spectra were filtered and merged into consensus spectra, producing 9,874 nodes (Figure 3a).

210   The patient-specific molecules were displayed by assigning a specific color to each patient in

211   the molecular network analysis (Figure 3a). In addition to annotating known compounds,

212   molecular networking revealed related molecules that differ by oxidation, methylation,

213   acetylation, hydroxylation, glycosylation, chain length, and saturation of alkyl chains, which

214   enabled identification of previously undescribed metabolites of administered pharmaceuticals

215   and microbial quinolones, as described below for azithromycin and *Pseudomonas aeruginosa*

216   quinolones. The frequency of detection of the antibiotics across patients' samples is shown in

217   Figure 3b, with corresponding clusters from the full network displayed for each antibiotic. The

218   nodes in antibiotic cluster represent the metabolic transformations of the antibiotic. Thus,

219   molecular networking provides a glimpse into metabolic processes. The resulting molecular

220   network revealed that among three patients, remarkably, only about 27.6% of detected

221   molecular features were shared, highlighting the diversity of chemistry present in diseased

222   human lungs (Figure 3c). All three patients in this study had different mutations in the CFTR

223   gene (see Materials and Methods) and patients 2 and 3 were also diagnosed with CF-related

224   diabetes. Two of the three patients suffered from chronic infections by *Pseudomonas*

225   *aeruginosa*. Thus, various factors may play a role leading to the observed chemical diversity,

10

226   which may arise from microbial (e.g., virulence and quorum sensing metabolites such as

227   quinolones), host (e.g., bile acids, amino acids, sugars, eukaryotic lipids, fatty acids, sterols,

228   peptides, immune-related molecules), and xenobiotic molecules. The diversity of these

229   metabolites in CF sputum has been previously characterized and many of the same compounds

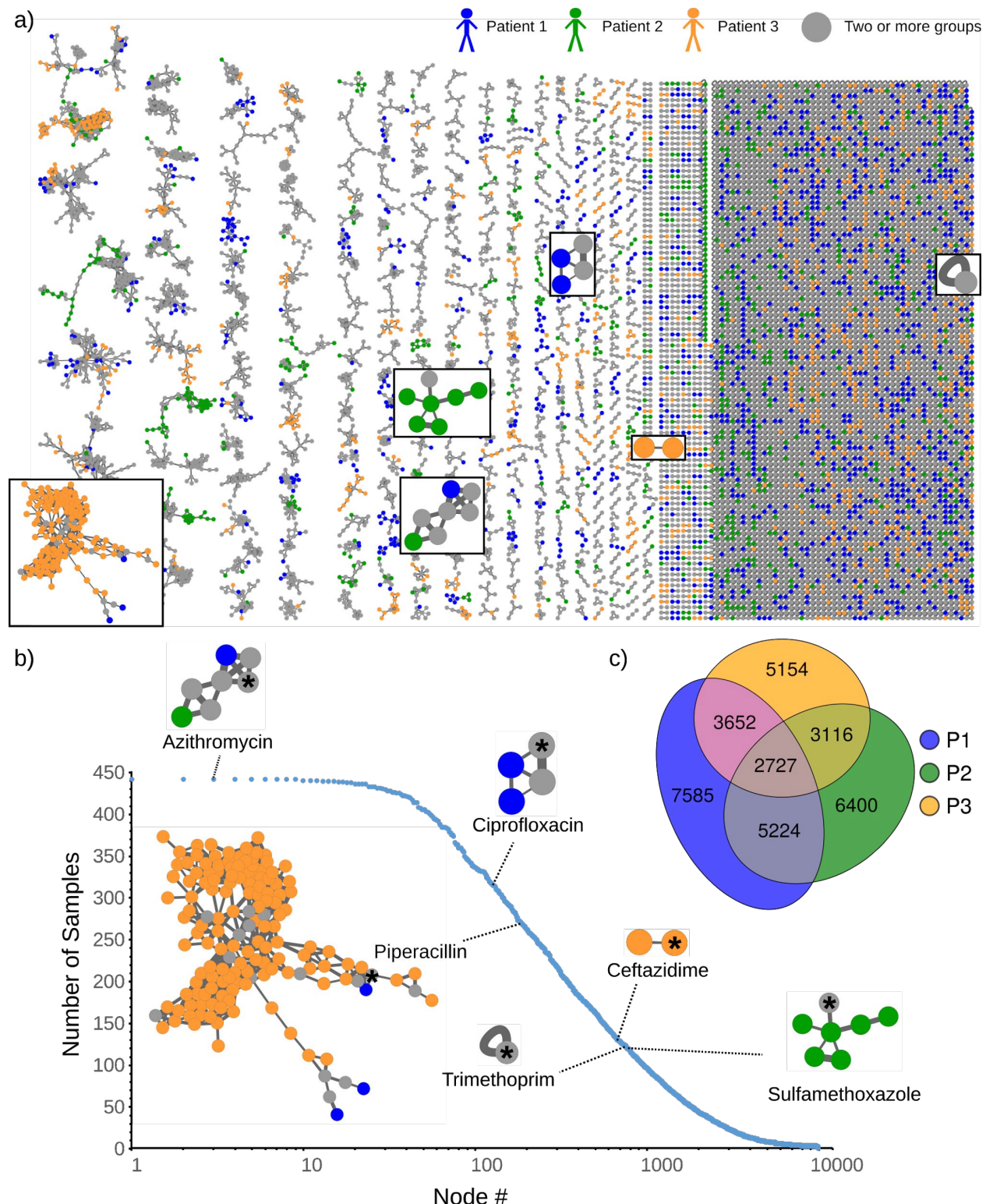230   were also found in the lung tissue in this study (32).

Figure 3. Molecular network analysis of all six lungs from three patients afflicted with CF.
a) The molecular network is color coded by patients (patient 1 (blue), patient 2 (green), and patient 3 (orange). The network clusters corresponding to antibiotics are highlighted in boxes.
b) The number of samples that contain a given consensus MS/MS spectra (represented as

12

237 node in Figure 3a) are plotted. The frequency of occurrence of antibiotics detected in this
238 dataset is highlighted on the plot. The number of nodes in a cluster is reflective of detected
239 transformations of the parent compound. The node of parent compound is highlighted by
240 asterisks. The fragmentation patterns of the most frequently observed drug - azithromycin and
241 its analogs are described in Supplementary Figure 7; the large number of nodes in Piperacillin
242 cluster stems from its structural similarity to small peptidic compounds abundant in biological
243 samples and its inherent chemical reactivity with biological molecules (33). c) Venn diagram of
244 the overlap of consensus fragmentation spectra between three patients is shown.
245
246 A Procrustes analysis of metabolomics data and 16S rRNA data with closed-reference

247 OTU picking revealed a close association between the microbiome and metabolome in the lung

248 samples (Mantel test r statistic = 0.45, P < 0.001, n = 277) (Supplementary Figure 6 (a,b)). This

249 analysis suggests that the microbial composition of each sample in large part is associated with

250 the corresponding chemical diversity. Additionally, Procrustes analysis performed on

251 metabolomics and 16S rRNA with deblurred sOTUs (34) resulted in the same trend

252 (Supplementary Figure 6 (c,d), Mantel test r statistic = 0.38, P < 0.001, n = 263). A PCoA plot of

253 the metabolome data with Canberra distance showed that a vast chemical diversity exists not

254 only between the patients (Figure 1c), but also within a patient's own lungs (Supplementary

255 table 2). This suggests that the chemical makeup of the patients with CF disease is highly

256 personalized and that a single CF lung contains unique chemical microenvironments that

257 provide different niches for microbial pathogens to live in. While metabolic diversity between

258 patients in relation to disease state is previously described (32, 35), mechanisms leading to

259 such diversity within the lungs remain poorly understood.

260 One of the additional benefits of an untargeted metabolomics analysis approach is the

261 ability to track the medications that are taken by the patient, as medical records can oftentimes

262 be incomplete and/or inaccurate due to lack of patient compliance, as well as to identify

263 metabolic transformations of the medications. For example, in the present study, in addition to

264 the prescribed medications from the clinical records (different antibiotics, bronchodilators, two

265 medications for digestive health, medications given during surgery and over-the-counter

266 medications that are used as cough suppressants), antihistamines and multiple over the counter

267   medications have been detected (Table 1). Detailed knowledge of extant exogenous

268   compounds in tissues of interest is important, among other things, to evaluate their effect on the

269   microbiome and microbial interactions for better understanding disease etiology. Another

270   advantage of a molecular networking approach for untargeted metabolomics data analysis is

271   that it allows for postulating structures for unknown compounds, nodes of which are connected

272   to nodes of known compounds (annotation propagation), and is therefore very useful for

273   identifying drug metabolites (Figure 3b). The distributions of drugs and the metabolites can then

274   be evaluated by 3D cartography even in the absence of a stable isotope tracer. Using a

275   molecular networking approach in this study, unknown metabolites that have never been

276   reported before in blood or tissue of humans and animals were detected (Supplementary Figure

277   7). The unknown metabolite of Azithromycin (*m/z* 382.26) is annotated as methylated-

278   azithromycin where the methylation, based on the analysis of the fragmentation data, occurs in

279   the core macrolide ring structure of azithromycin and another unknown metabolite is proposed

280   to have oxidation in the macrolide ring (Supplementary Figure 7). These modifications of core

281   macrolide structure of azithromycin have not been described previously and their biological

282   activities are unknown. Although these metabolites were not detected in *in vitro* cultures of

283   microbes isolated from these patients in the presence of azithromycin, the possibility that these

284   are microbially-derived warrants further investigation and cannot be ruled out. Specific *in vivo*

285   conditions may be necessary for regulation of microbial genes involved in antimicrobial

286   metabolism.

14

287 **Table 1.** The medications detected in the MS data and time of administration prior to the day of
288 lung explantation surgery. C refers to: "continuously administered"
289

Medications	Patient 1	Patient 2	Patient 3	Time before surgery (days)		
Antibiotics						
1. Trimethoprim		●	●	-	8	-
2. Sulfamethoxazole		●		-	8	-
Acetylsulfamethoxazole		●				
Sulamethoxazole-glucuronide		●				
Sulfamethoxazole metabolite m/z 502.26		●				
3. Azithromycin	●	●	●	C*	C	10
Descladinose-azithromycin	●	●	●			
3''-*O*-desmethyl-azithromycin	●	●	●			
3'-*N*-desmethyldescladinose-azithromycin		●	●			
Methyl-azithromycin unknown	●	●	●			
Descladinose-azithromycin (oxidation -2Da, unknown)	●	●	●			
4. Piperacillin	●		●	6	-	1
Desethylpiperacillin	●		●			
5. Ciprofloxacin	●		●	6	-	1
6. Ceftazidime			●		-	1
7. Minocycline	●			2	-	-
8. Linezolid		●		-	X	-
Bronchodilator						
1. Pseudoephedrine		●		-	-	-
2. Albuterol (bronchodilator)	●			6		
Digestive Health						
1. Famotidine	●			6	-	-
2. Metoclopramide			●	-	-	1
Surgery (sedation, anxiolytic, paralytic)						
1. Fentanyl (for surgical pain)	●		●			
2. Haloperidol (ICU)	●					
3. Dihydrohaloperidol	●					
4. Mycophenolate (immune suppressant)		●				
5. Midazolam		●	●			
Hydroxymidazolam		●				
Hydroxymidazolam glucuronide		●				
6. Rocuronium bromide	●	●	●			
7. Vecuronium bromide			●			
8. Ondansetron			●			
Others						
1. Diphenhydramine		●	●	-	8	10
2. Amitryptiline (antidepressant)	●			X	-	-
3. Loratadine	●			6	-	-
Desloratadine (active ingredient of loratadine, antihistamine)	●					
4. Chlorhexidine (disinfectant)	●	●	●	-	-	-

290
291

15

292   As with the microbial heterogeneity, we have observed differences in metabolome

293   distributions. Molecular networking and 3D volume cartography of the antibiotics revealed

294   patient-specific metabolism and drug distributions (Figure 4 and Supplementary Figure 8). The

295   distribution of antibiotics was also found to be different between the left and right lungs of the

296   same patient. For example, the antibiotic Piperacillin and its metabolites were abundant in the

297   upper lobes of the right lung of patient 3 but present in relatively lower abundance in the right

298   lung of this patient (Figure 4). In patient 3, there was higher penetration of piperacillin in the

299   upper and middle lobes and poor penetration in the lower lobes of both lungs. Similarly, the

300   antibiotic Linezolid detected in patient 2 had lower relative abundance in the lower lobe of the

301   right lung (Supplementary Figure 8). Overall, the drug metabolites largely follow the same

302   distribution as parent drug except for the glucuronidated metabolite of sulfamethoxazole

303   (Supplementary Figure 8), indicating that metabolism may not be a significant contributing factor

304   for observed uneven distribution of detected antibiotics. Differential vascularization and tissue

305   necrosis also contributes to non-uniform drug penetration in severe end-stage CF disease.
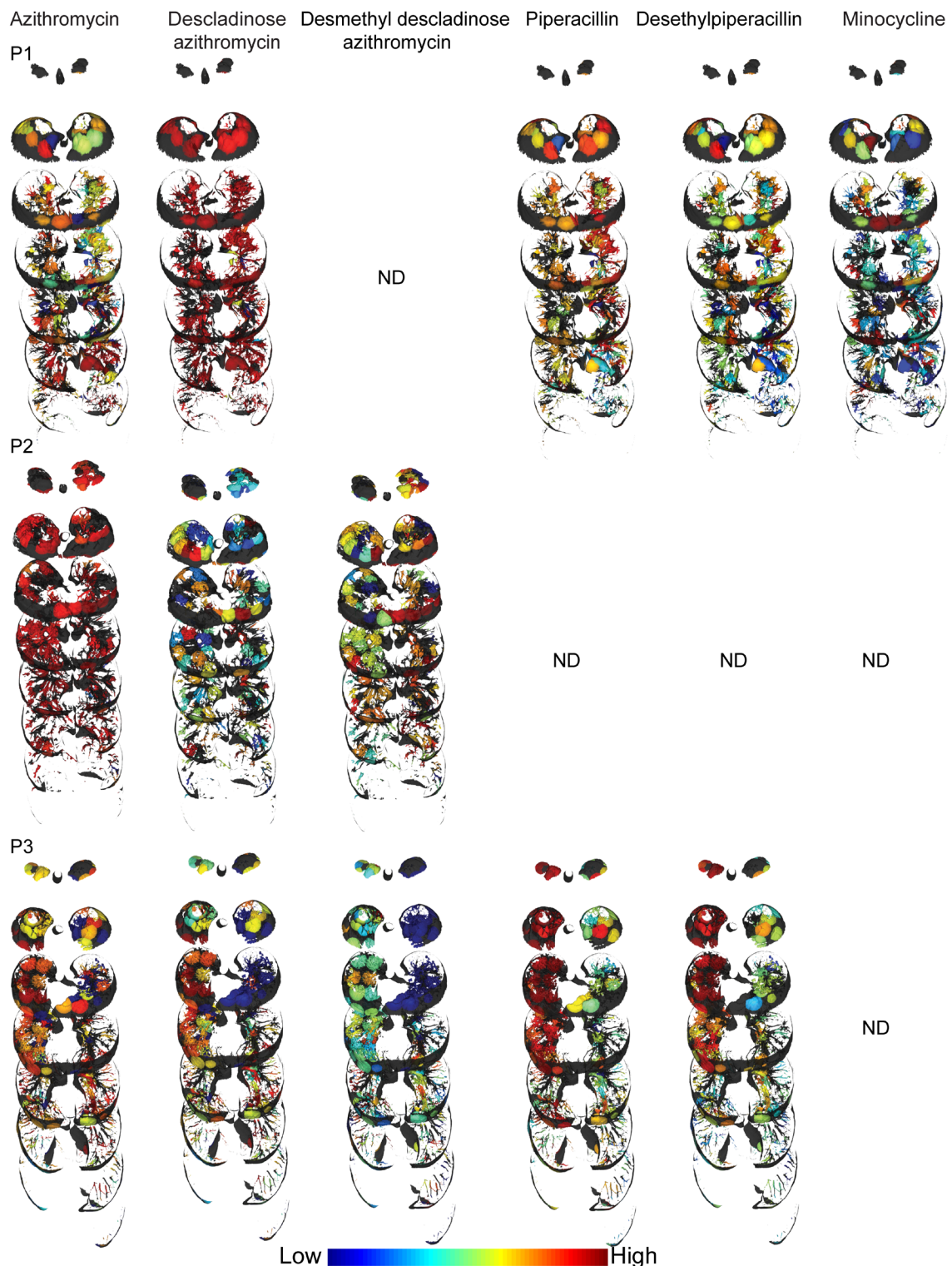
306

Figure 4. Distribution of selected antibiotics and the metabolites. P1 - patient 1, P2 - patient 2 and P3 - patient 3, ND - not detected. Intensity scale is provided at the bottom (distribution of additional antibiotics and their metabolites is provided in Supplementary Figure 8). Full visualizations of metabolite maps can be accessed via links: patient 1, patient 2 and patient 3.

312   To directly link microbiome and metabolome information and identify associations of

313   compounds detected in the lung tissue with specific microbes, isolates from lungs of all patients

314   were obtained from the clinical lab, cultured directly from swabs of the lung tissue, and MS data

315   was acquired on the organic extracts of the *in vitro* cultures using the same protocol as

316   employed for the tissue extracts. Molecular networking of MS/MS data from culture extracts and

317   tissue extracts provided insights into which molecules are shared between microbes and the

318   human host (Supplementary Figure 9). These molecules include microbe-specific virulence

319   factors, as well as various other molecules such as lipids, fatty acids, amino acid metabolites,

320   dipeptides and tripeptides. Similar to our previously reported observation for one CF lung of a

321   single CF patient (16, 32), a larger diversity of quinolones was detected in cultured isolates as

322   opposed to a smaller diversity of quinolones detected in the lung tissue of all patients dominated

323   by *Pseudomonas* in this study, including a quinolone at *m/z* 268.170 that has never been

324   reported before (Figure 5).   Based on MS[1] and MS[2] data, the structure of this quinolone is

325   proposed to contain two double bonds in the alkyl side chain as opposed to single double bond

326   found in unsaturated quinolones described in the literature (36) (Supplementary Figure 10). To

327   gain further insight into the variation in the distribution of quinolones in the patients dominated

328   with *Pseudomonas*, we investigated the distribution of quinolones directly within the lungs of

329   these patients (Figure 5b and Supplementary Figure 11). Previously, we reported that the

330   quinolones were prevalent at the upper lobe of the left lung of a single patient (16). In this study

331   quinolones were found to be exclusively present at the upper lobe of lungs of patient 1 and only

332   in the middle of the lungs of patient 3. This indicates that patients dominated by *Pseudomonas*

333   show individualized phenotypes with respect to the expression of these quorum sensing

334   molecules. Furthermore, rhamnolipids, the *Pseudomonas'* biosurfactant, were not detected in

335   the lungs of patient 1 and 3 in this study but were detected in our previous study (16). Patient-

336   specific production of rhamnolipids has been reported previously by culturing isolates in the

337   laboratory but not directly from infected tissue (13). Such compartmentalization of microbial

18

338   activity within patients, as well as variation between patients is a hallmark of complexity that is

339   inherent to polymicrobial infection in a complex organ; in the present case, a CF lung. Direct

340   visualization of the individual phenotypes in diseased organs enables informed understanding of

341   divergent evolution as well as the spatial molecular environment within a host.

a) Isolates (magenta)

Shared between isolates and Patients 1,3 (cyan)
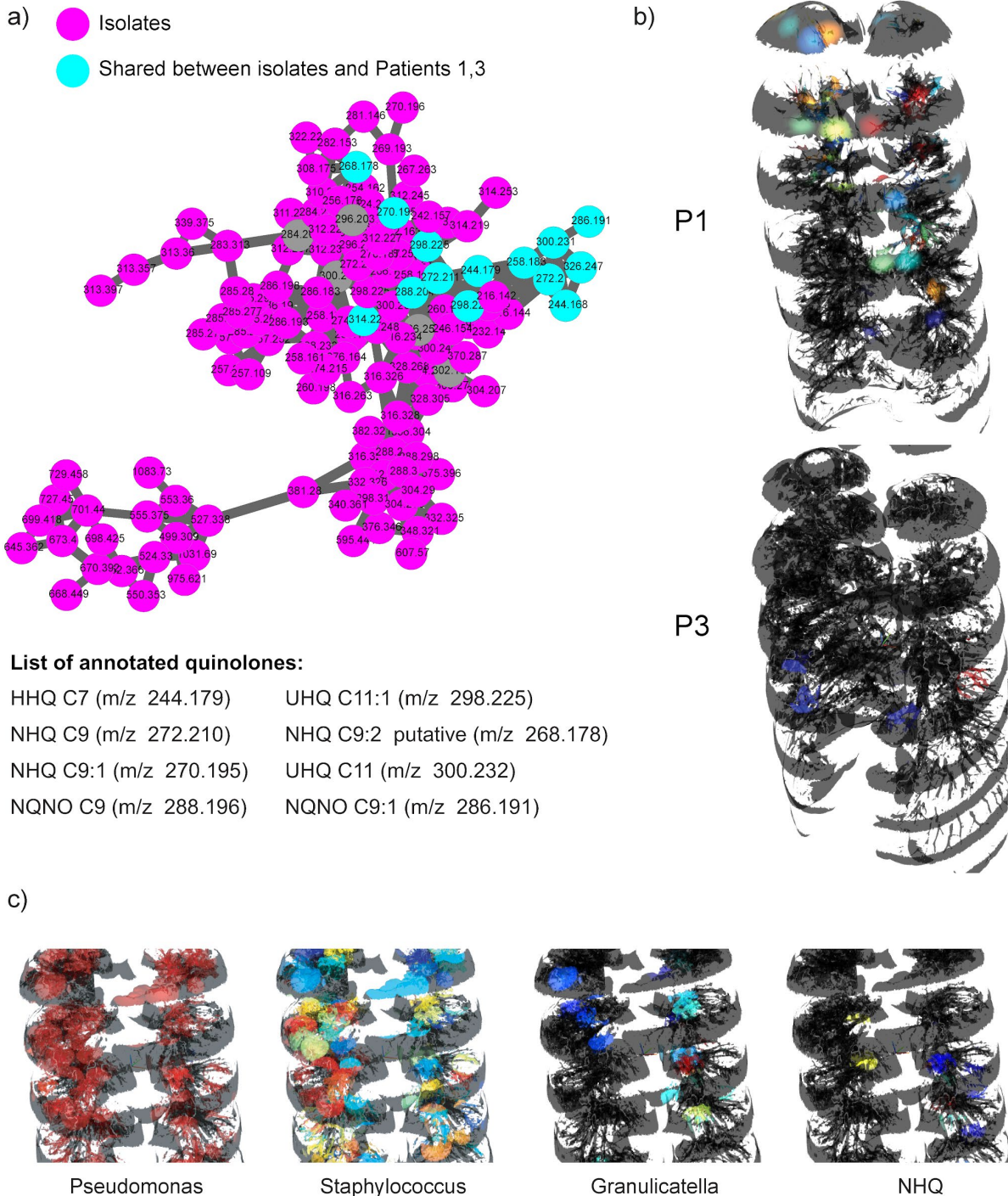
List of annotated quinolones:

HHQ C7 (m/z 244.179)   UHQ C11:1 (m/z 298.225)

NHQ C9 (m/z 272.210)   NHQ C9:2 putative (m/z 268.178)

NHQ C9:1 (m/z 270.195)   UHQ C11 (m/z 300.232)

NQNO C9 (m/z 288.196)   NQNO C9:1 (m/z 286.191)

b) P1

P3

c) Pseudomonas   Staphylococcus   Granulicatella   NHQ

Figure 5. Molecules produced by *P. aeruginosa* in patients 1 and 3. a) The molecular network cluster of quinolones detected in the lung tissue of patient 1 and 3 and *in vitro* microbial cultures of *Pseudomonas* isolated from sputum and the swabs collected from lung sections is shown. b) The distribution of the quinolone HHQ is shown in patients 1 and 3. All the other quinolones showed similar distributions in these patients (Supplementary Figure 11). c) Inset of the distribution of *Pseudomonas*, *Staphylococcus*, *Granulicatella* and the *Pseudomonas* quinolone NHQ in patient 3 suggestive of upregulation in quinolone production by

20

350   *Pseudomonas* in the regions where interactions of *Pseudomonas* with *Staphylococcus*,
351   *Granulicatella* and possibly other microbes take place. In agreement with this observation, the
352   production of HHQ and NHQ was also found to increase in co-cultures of *Pseudomonas* and
353   *Staphylococcus* compared to Pseudomonas grown alone under identical conditions
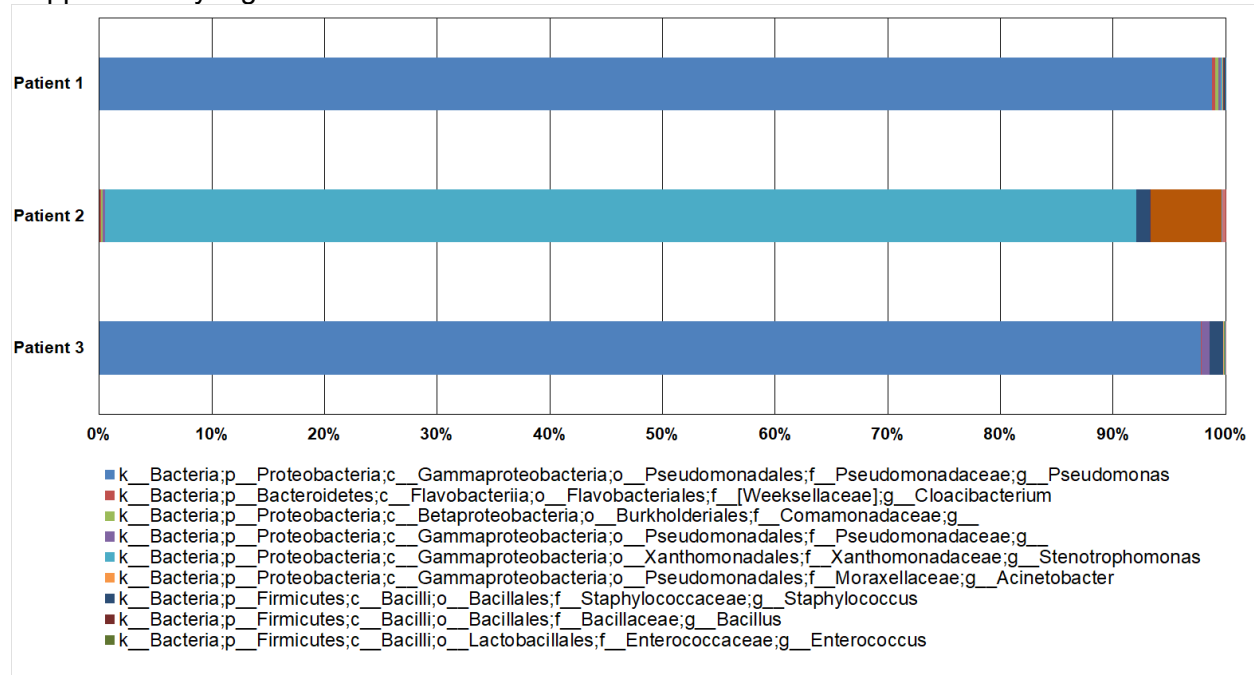354   (Supplementary Figure 12).
355
356
357   The spatial co-distribution of microorganisms, antibiotics and microbial molecules were

358   investigated to establish microbe-metabolite interactions. Although the presence of a single

359   dominant pathogen renders correlation analysis rather uninformative, several trends have been

360   observed. In particular, the distribution of certain microorganisms such as *Staphylococcus* and

361   *Granulicatella* were found to be associated with the distribution of quinolones produced by

362   *Pseudomonas* in patient 3, as shown for NHQ in Figure 5c. We have recently shown that

363   *Staphylococcus aureus* isolated from a CF patient increases quinolone and biofilm production

364   by co-isolated *Pseudomonas in vitro* (37). Similarly, mixing cultures of the *Pseudomonas* and

365   *Staphylococcus* isolated from patient 3 in this study showed increased production of HHQ and

366   NHQ when compared to *Pseudomonas* grown alone under identical conditions (Supplementary

367   Figure 12a). This observation indicates that the production of quinolone molecules is in part also

368   modulated by microbial interactions present in a polymicrobial infection. The complexity of these

369   microbial interactions is further increased as antibiotics cause perturbations of microbial

370   communities reflected by suppression of the virulence factors. Variation in production of

371   quinolones by patient isolates of *Pseudomonas* was observed upon exposure to sub-MIC

372   concentrations (Supplementary Figures 12b). This, and the other observations reported here

373   support the hypothesis that not only are genetic changes responsible for changing metabolism,

374   but microbial interactions in conjunction with multiple other factors including sub-MIC

375   concentrations of antibiotics and perhaps other xenobiotics may also play a role and call for the

376   design of specific studies investigating these phenomena in multiple patient isolates. Thus, it is

377   reasonable to hypothesize that both specific microbial interactions in the lungs and differential

378   abundances of antibiotics could result in metabolic divergence, creating isolated regions of

379 enhanced biofilm formation and tissue damage that is often observed in CF patients by chest X-

380 rays and CT-scans. Application of advanced techniques such as Ultra-High-Resolution

381 Computed Tomography in conjunction with the approach presented here could be a focus of

382 future studies (38).

383

384 **Conclusion**

385   Cystic fibrosis is a devastating genetic disease which affects tens of thousands of people

386 worldwide. In this work, we presented the findings of spatial distributions of microbes,

387 medications, and their metabolites throughout lungs of three patients afflicted with CF. We have

388 found that although the microbiome is predominantly patient-specific, the chemical differences

389 between locations within patient's own lungs may be greater than inter-patient variations. In-

390 depth analyses revealed differential drug penetration, metabolism of prescribed medications,

391 and microbial compartmentalization resulting in metabolic divergence governed by local

392 microbial interactions. Mapping of microbial communities and localized chemistries allowed for

393 visualization of interactions among community members, such as production of quinolones by

394 *Pseudomonas* when present in a community structure with other microbes such as

395 *Staphylococcus* or *Granulicatella.* Visualization of such local infection loci highlight the

396 importance of development of effective drug delivery approaches. Considering recent advances

397 in the development of small-scale robots (as small as few micrometers) that can non-invasively

398 access confined spaces (39), targeted access of internal tissues as well as precision delivery of

399 drug payloads may become feasible in the near future. In general, a paradigm shift of

400 considering localized regions of divergent microbial and chemical distributions is an important

401 next step for effective disease management of polymicrobial infections.

402

403   Supplementary Figures and Tables:

- k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas
- k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__[Weeksellaceae];g__Cloacibacterium
- k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__
- k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__
- k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Stenotrophomonas
- k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter
- k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;g__Staphylococcus
- k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus
- k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae;g__Enterococcus

404
405   Supplementary Figure 1. The bar charts of taxonomic summary of all samples for each patient
406   is shown. The legend for 9 most abundant taxa is shown. The complete taxonomic summary is
407   available in Supplementary table 1. The complete description of all OTU's for each individual
408   sections are available through spatial maps: patient 1, patient 2 and patient 3.
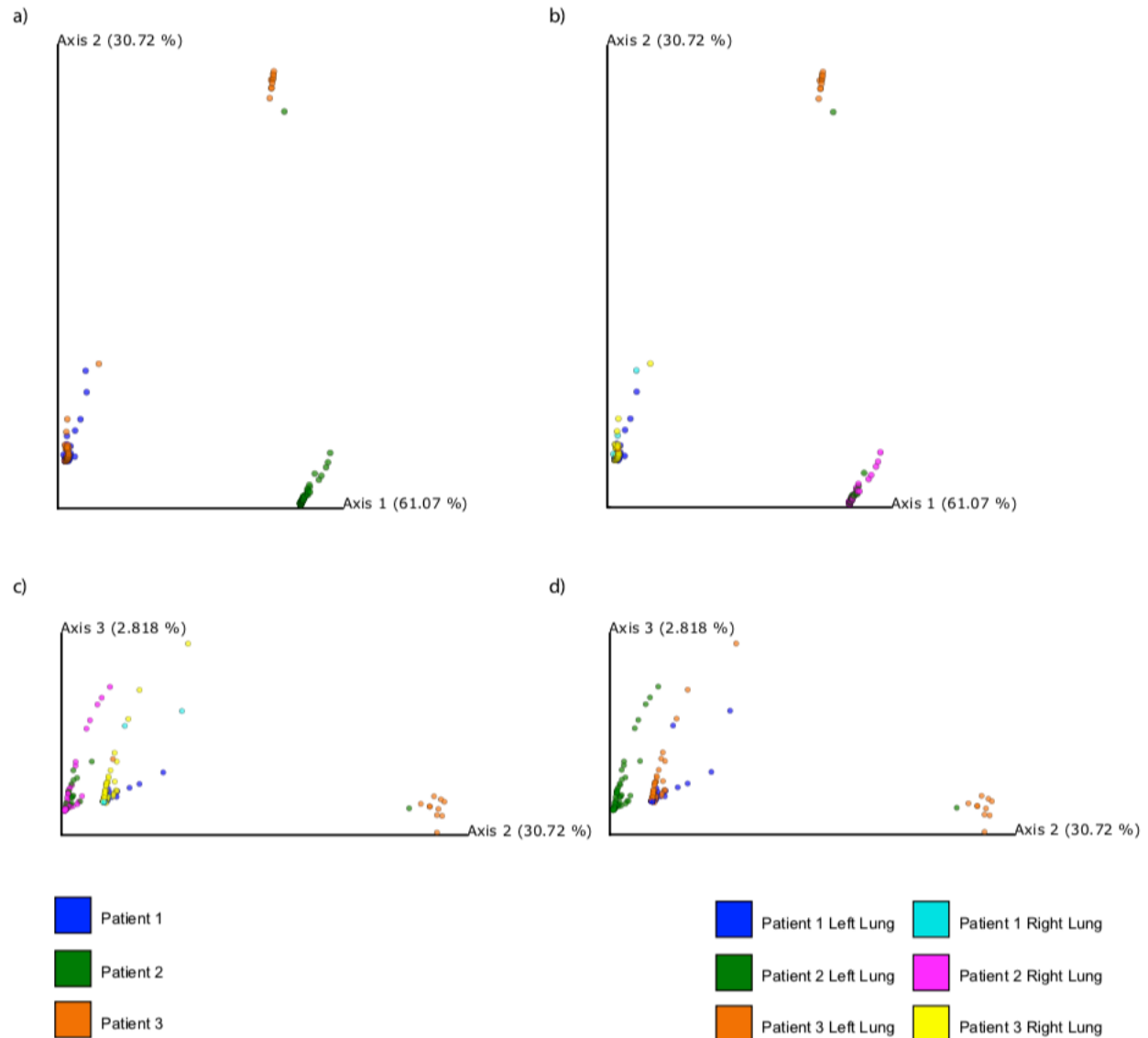
Supplementary Figure 2. PCoA plot on Weighted Unifrac distance of 16S rRNA data from three patients is shown (40). a) and b) first and second principal component is shown. c) and d) second and third principal component is shown.

24

414
415
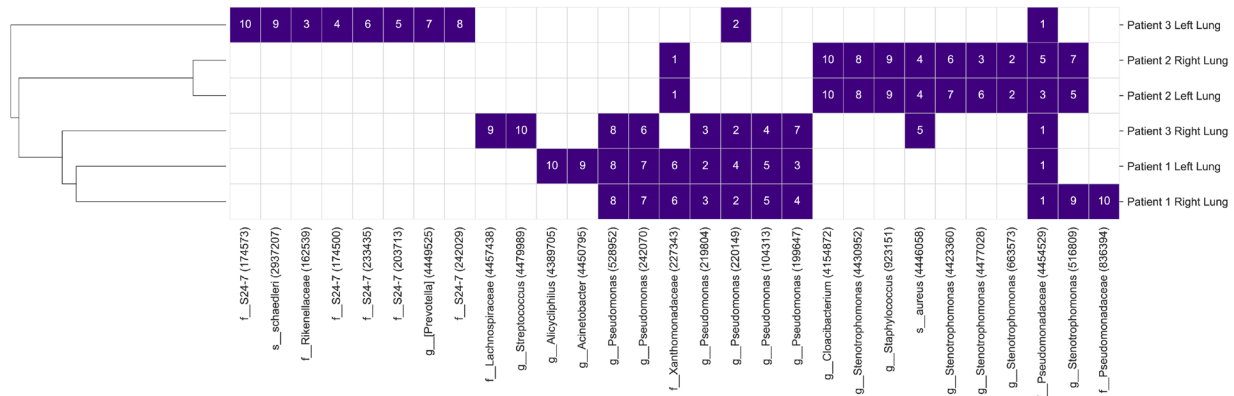
Supplementary Figure 3. Top 10 most abundant OTUs and their ranks (1 most abundant; 10 least abundant) according to host and sampling site. Columns represent OTUs and are labeled according to their taxonomic classification and their Greengenes identifier. Rows represent the sites and are hierarchically clustered based on the rank vectors.

421
422
423
424

425
426
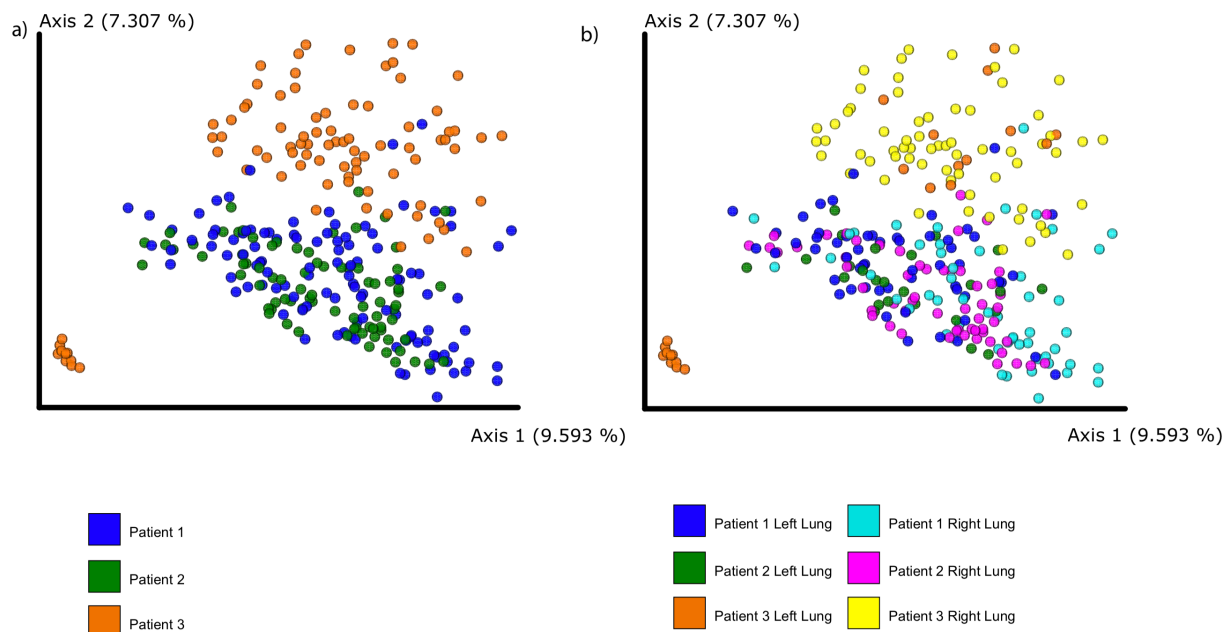
Supplementary Figure 4. Unweighted UniFrac PCoA plot of the 16S data. a) Ordination colored by subject identifier. b) Ordination colored by subject identifier and sampling site.

429
430
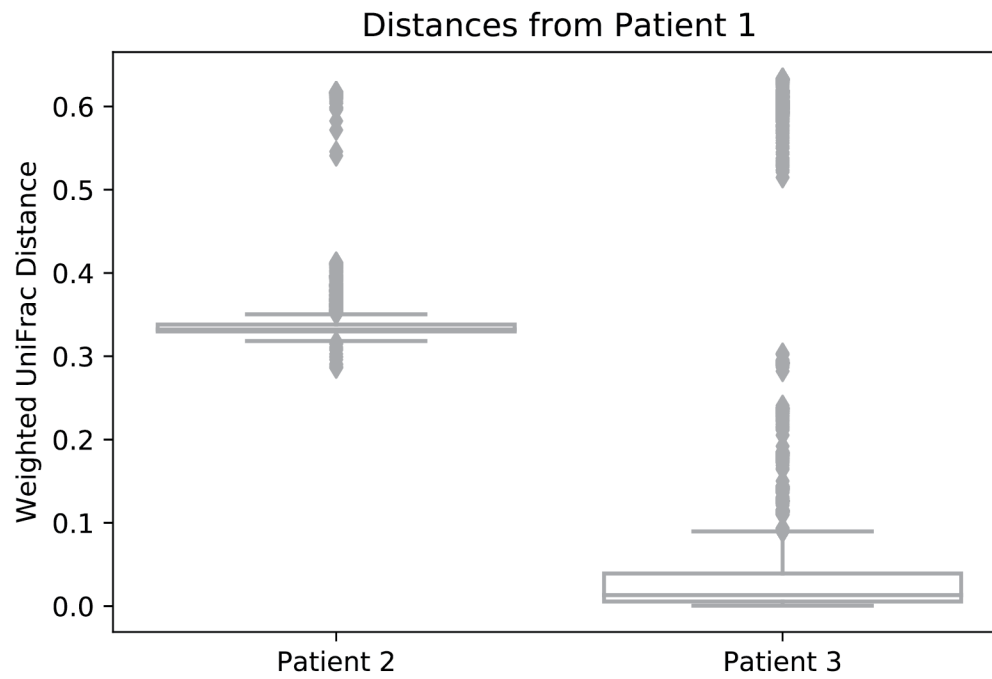
Supplementary Figure 5. Weighted UniFrac distances from Patient 1 to Patients 2 and 3.

26
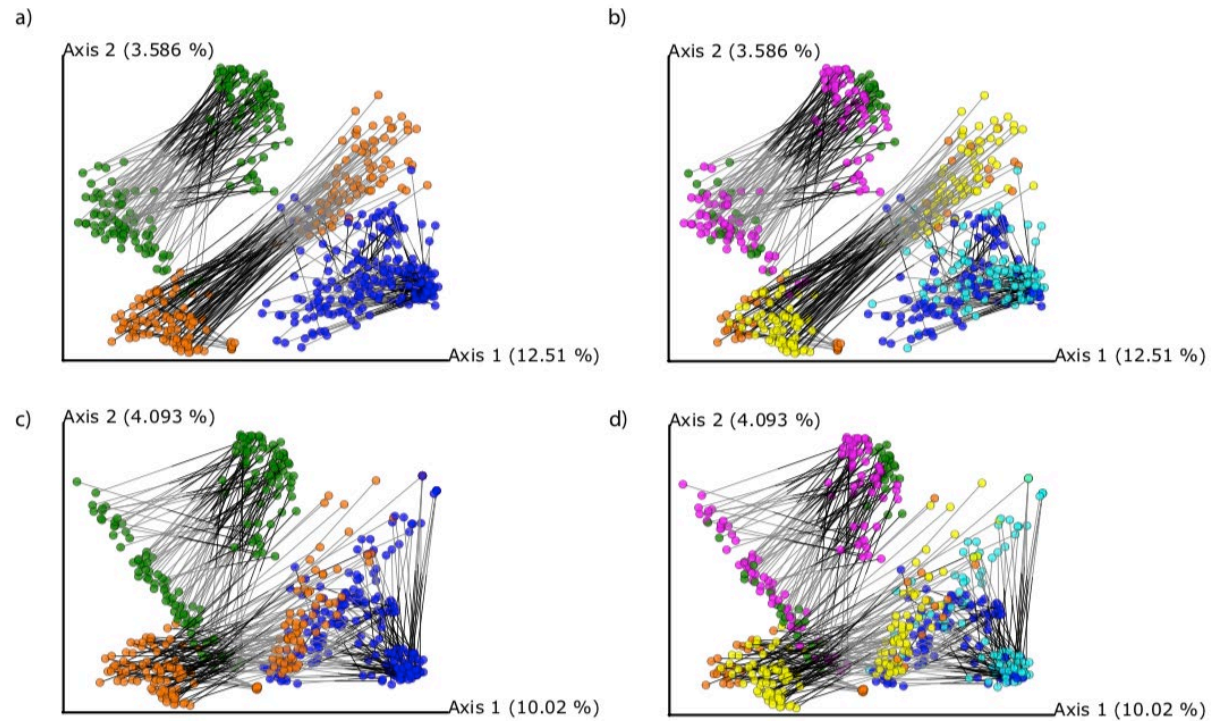
Supplementary Figure 6. Procrustes plots generated using principal component analysis of Canberra distance metric for untargeted metabolomics data and 16s rRNA sequencing data. a),b) - Emperor visualization plots of metabolomics and closed-reference picked OTUs. c),d) - Emperor visualization plots of metabolomics and deblurred sOTUs.
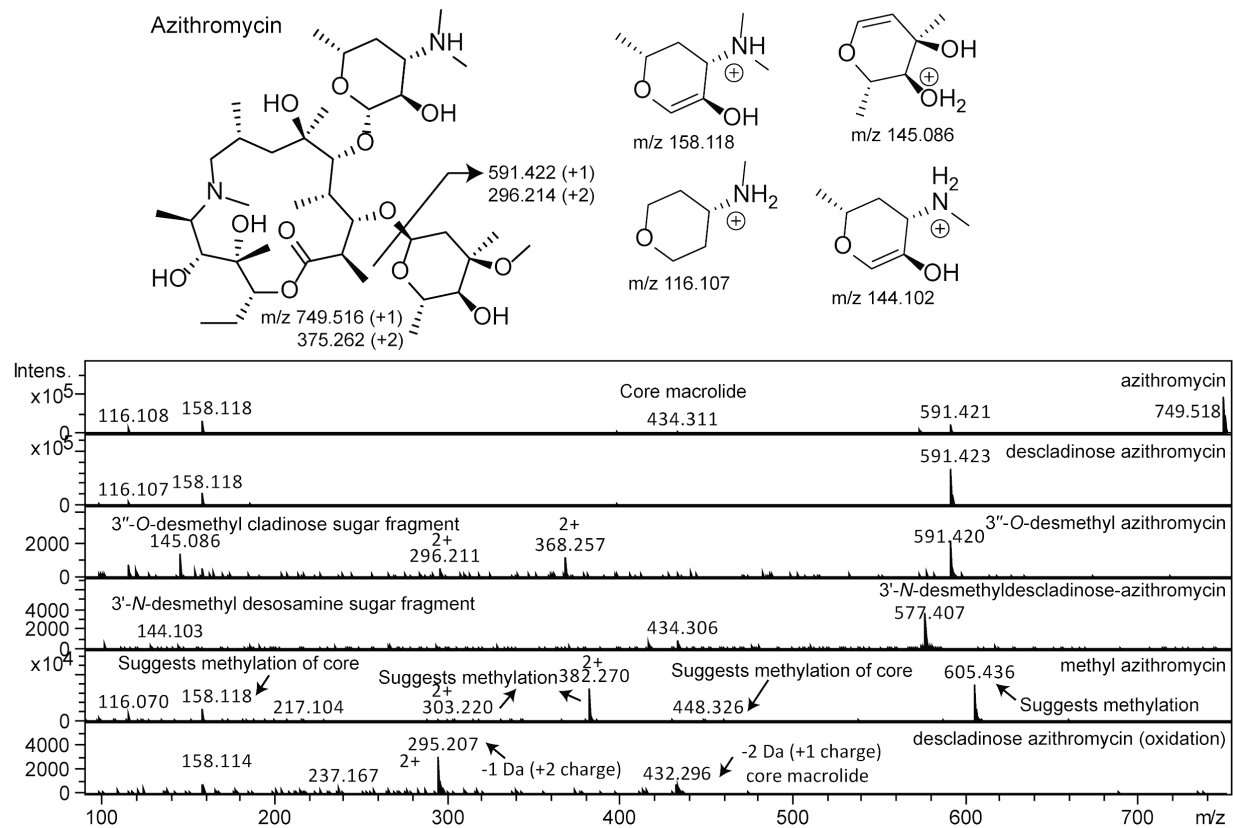
Supplementary Figure 7. The mass spectral analysis for azithromycin and its proposed metabolites is shown.
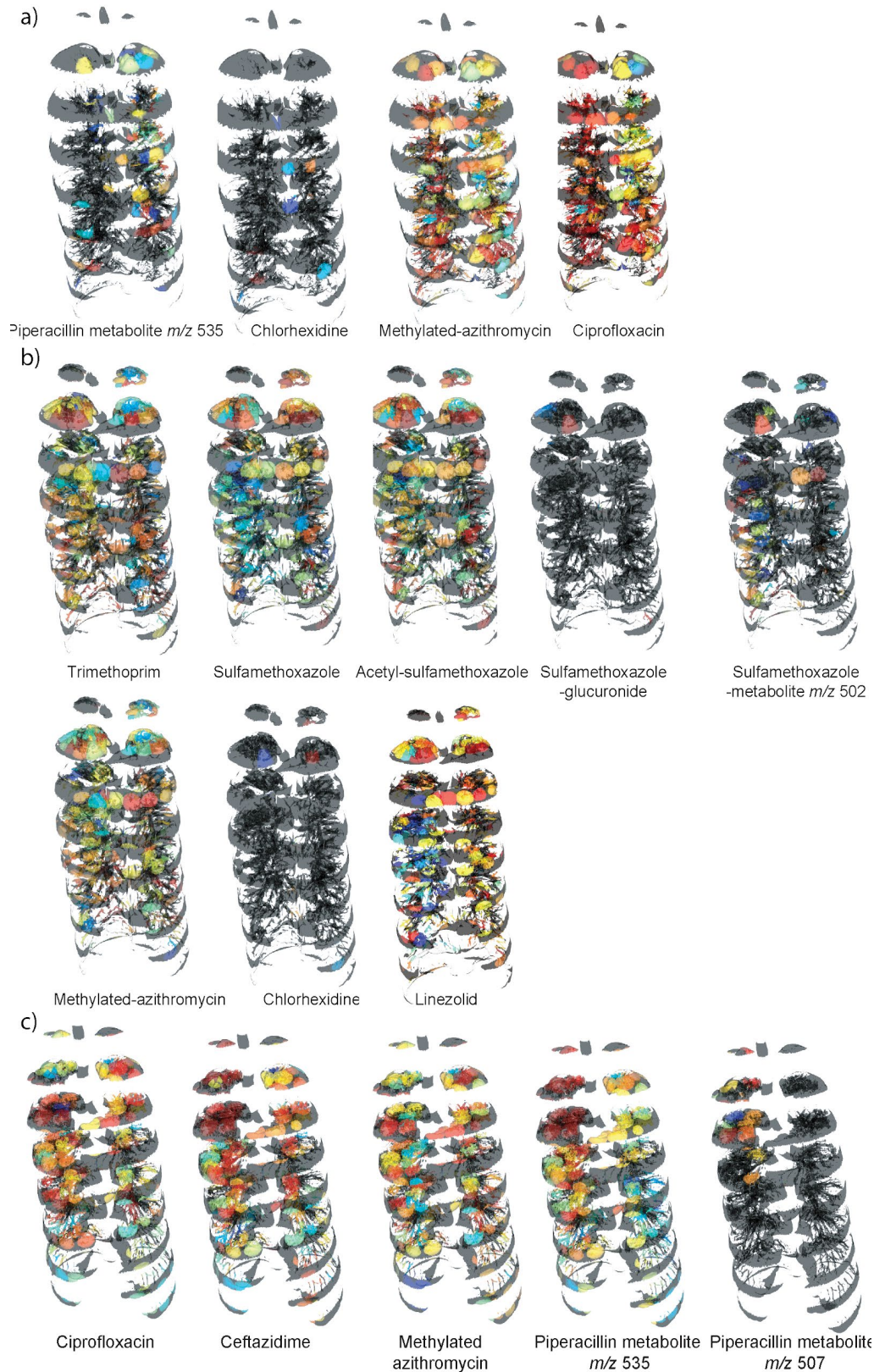
Supplementary Figure 8. Distribution of antibiotics and its metabolites in patients. a) patient 1, b) patient 2 and c) patient 3.
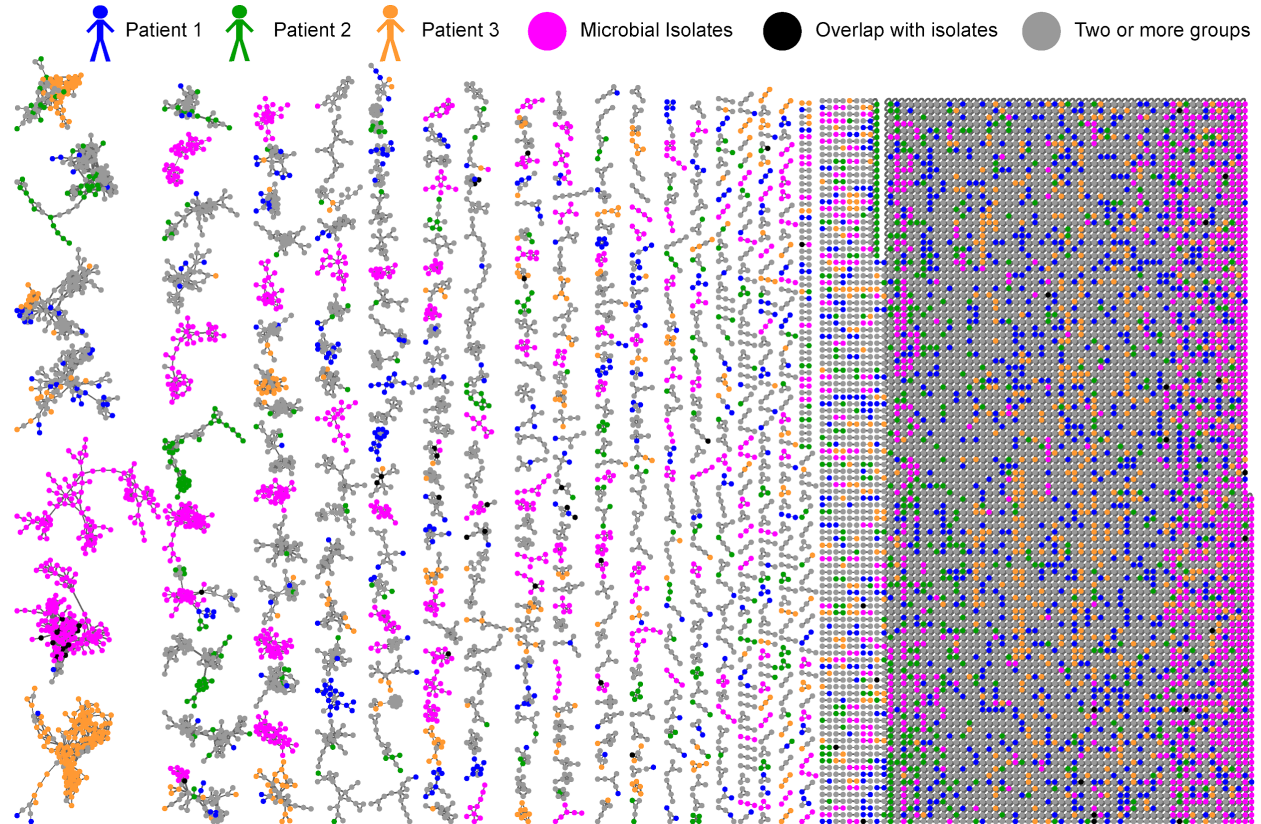
446
447
448 Supplementary Figure 9. Molecular network analysis of all six lungs from three patients and
449 microbial isolates. The molecular network is color coded by sample source.
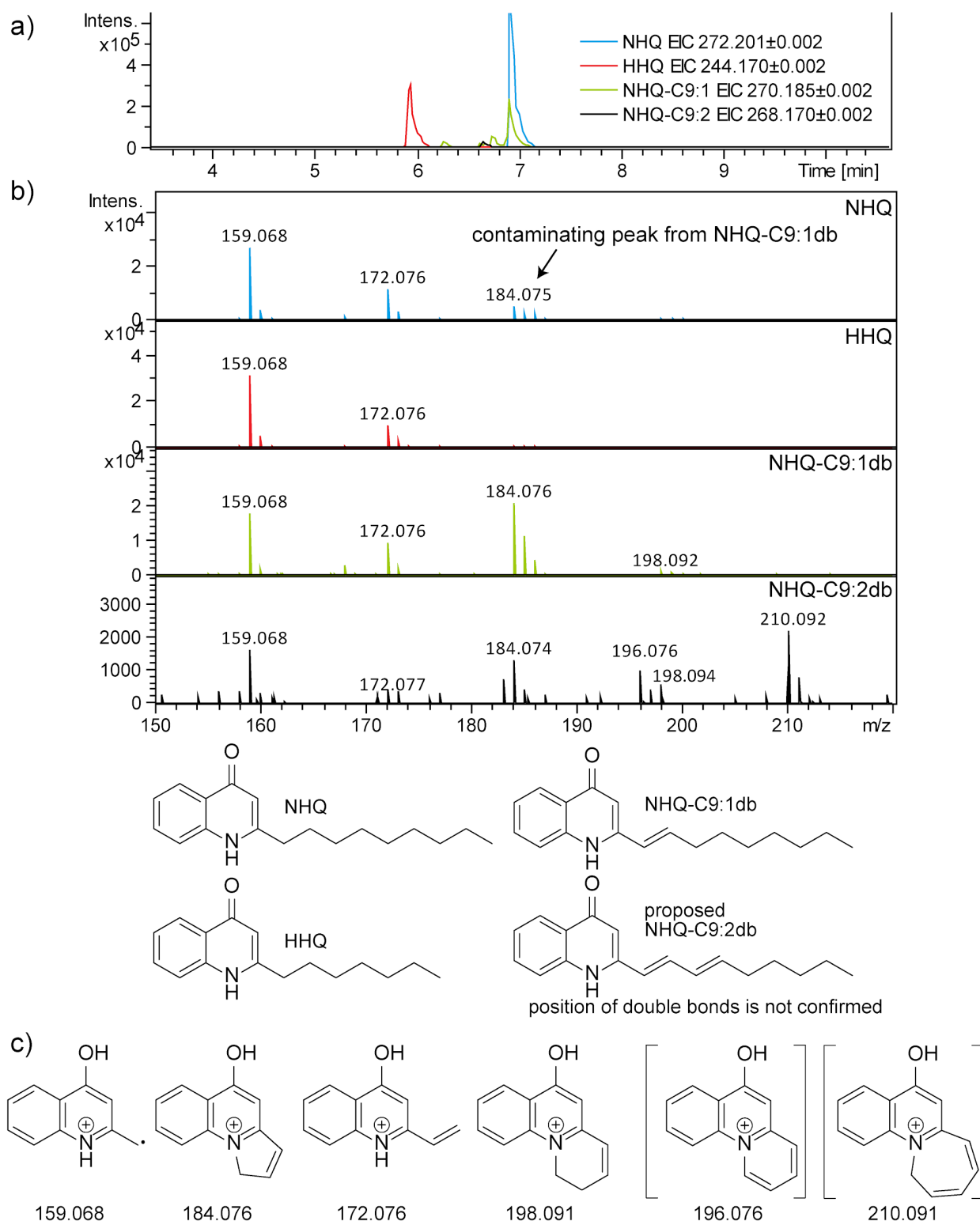
Supplementary Figure 10. a) The extracted ion chromatograms (EIC), b) mass spectrum and structures of quinolones NHQ, HHQ, NHQ-C9:1db, NHQ-C9:2db (proposed) are shown. The position of additional double bond for NHQ-C9:2db is not known and is putatively drawn. c) The structures of fragments described previously in the literature for NHQ, HHQ and NHQ-C9:1db[40]

31

456 and the proposed structures for fragments corresponding to NHQ-C9:2db are enclosed in
457 brackets.
458
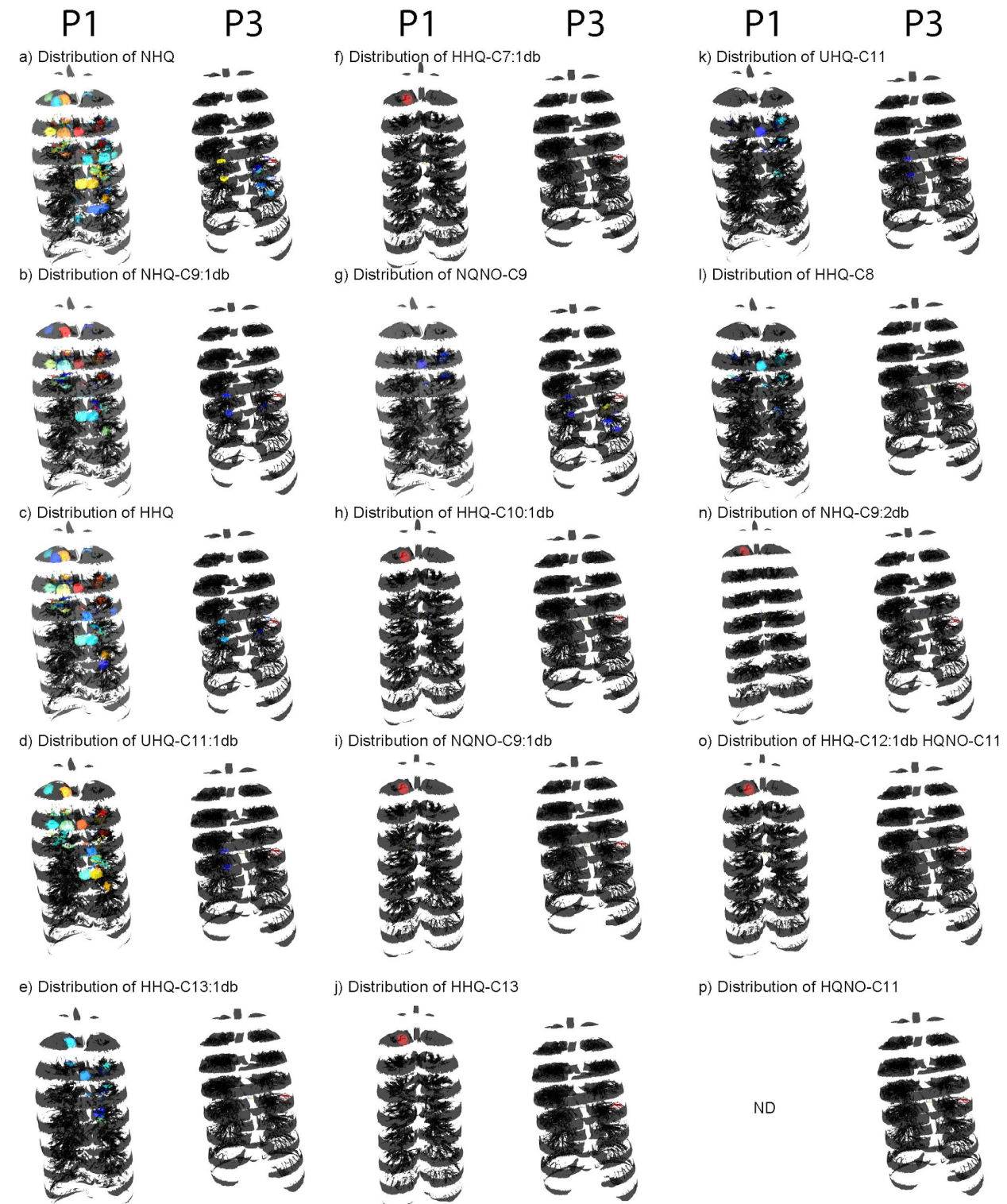
459
460 Supplementary Figure 11. Distribution of quinolones that are common to patients 1 and 3.
461
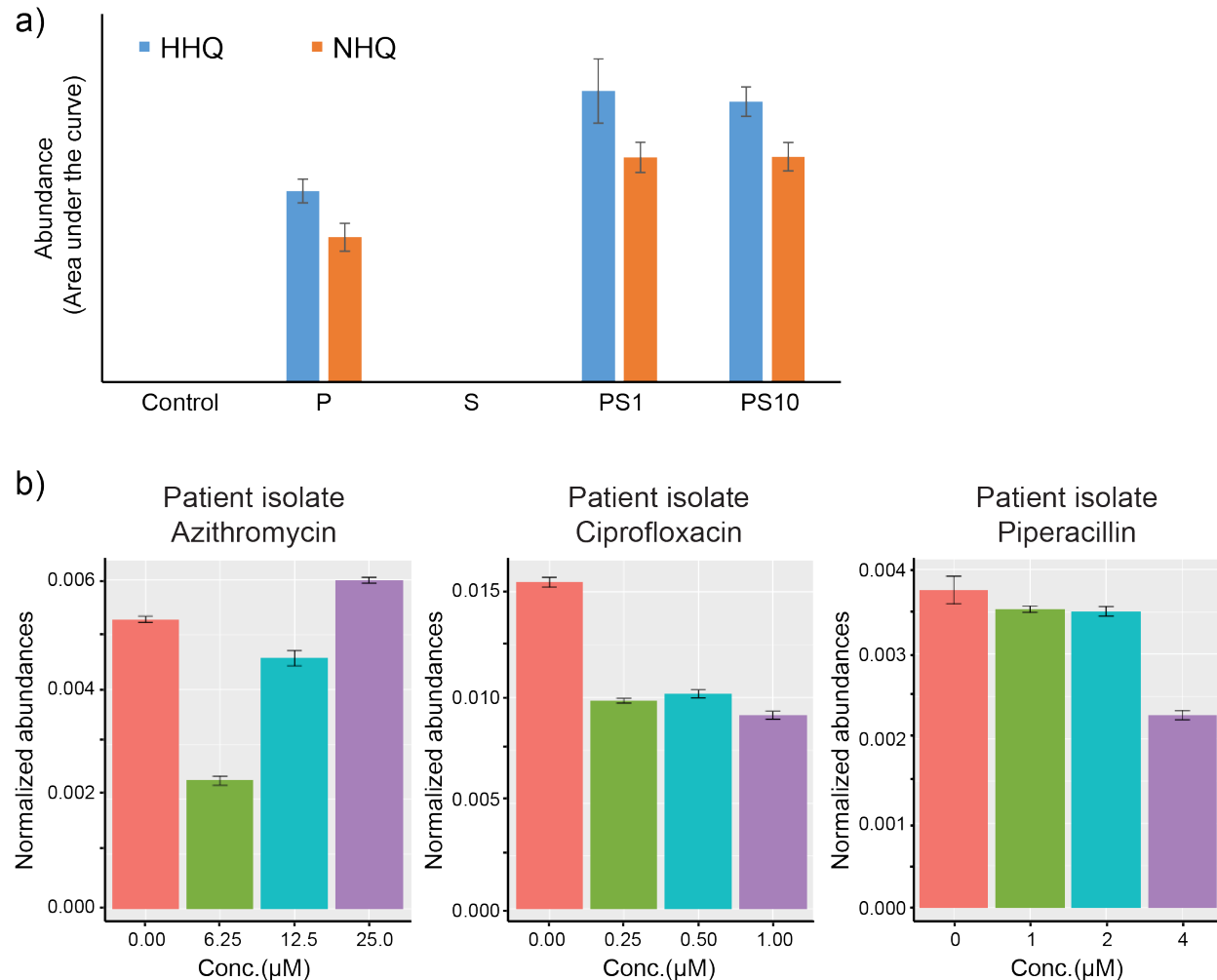
462
463 Supplementary Figure 12. Bar graphs of plots of area under the curves normalized to total ion
464 current (error bars indicate standard deviation from mean values of three independent
465 experiments (n = 3)). a) Quinolone production by *Pseudomonas aeruginosa* in single culture
466 and in mixed culture with *Staphylococcus aureus* is shown. PS1 represents cultures were mixed
467 in the ratio 1:1 and PS10 represents a ratio of 1:10, where *S. aureus* culture was diluted 10 fold.
468 b) Production of quinolone NHQ in response to antibiotic exposure by *Pseudomonas*
469 *aeruginosa* isolated from patient 3. Each color represent different concentrations of antibiotics.
470

471   Supplementary Table 1. Taxa summary table

Constructed from biom file

#OTU ID	patient_1	patient_2	patient_3
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas	0.987999611	0.0000134	0.978101853
k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__[Weeksellaceae];g__Cloacibacterium	0.00275499	0.001671197	0.000108784
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__	0.002739072	0.001235305	0.0000858
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__	0.001915169	0.001952885	0.007403893
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Stenotrophomonas	0.001218315	0.915415264	0.000047
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter	0.001097751	0.000449253	0.0000796
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;g__Staphylococcus	0.000685652	0.012052438	0.011663445
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus	0.00037761	0.0000507	0.0000301
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae;g__Enterococcus	0.000318949	0.00000668	0
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Hylemonella	0.000273553	0.000110226	0
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Rhodocyclales;f__Rhodocyclaceae;g__Dechloromonas	0.000243486	0.000118019	0
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__	0.000239949	0.062973295	0.0000707
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium	0.0000607	0.000868444	0.0000801
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides	0.0000528	0.003076851	0.0000879
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Micrococcaceae;g__Rothia	0.00000973	0.00000111	0.000120101
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Actinomycetaceae;g__Actinomyces	0.00000855	0.00000111	0.0000271
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__	0.00000265	0.00000334	0.0000351
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Bradyrhizobiaceae;g__	0.000000884	0	0.000337207
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Alicycliphilus	0.00000059	0	0.0000825
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]	0	0.000000557	0.00000508
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus	0	0	0.000698204
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__	0	0	0.000280506
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Neisseriales;f__Neisseriaceae;g__Neisseria	0	0	0.000117676
k__Bacteria;p__Firmicutes;c__Bacilli;o__Gemellales;f__Gemellaceae;g__	0	0	0.0000984
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella	0	0	0.0000903
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Carnobacteriaceae;g__Granulicatella	0	0	0.0000762
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus	0	0	0.0000654
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Ralstonia	0	0	0.0000619
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella	0	0	0.0000462
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Porphyromonas	0	0	0.0000385
k__Bacteria;p__Fusobacteria;c__Fusobacteriia;o__Fusobacteriales;f__Fusobacteriaceae;g__Fusobacterium	0	0	0.0000182
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Achromobacter	0	0	0.0000117
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Barnesiellaceae];g__	0	0	0.00000901
k__Bacteria;p__Fusobacteria;c__Fusobacteriia;o__Fusobacteriales;f__Leptotrichiaceae;g__Leptotrichia	0	0	0.00000658
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Sphingobium	0	0	0.00000612
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Neisseriales;f__Neisseriaceae;g__Conchiformibius	0	0	0.00000404
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Campylobacteraceae;g__Campylobacter	0	0	0.00000381
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiaceae;g__	0	0	0.00000104
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Enterobacter	0	0	0
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Yersinia	0	0	0

472

473

474

475 Supplementary Table 2. Percentage of shared mass spectrometry features between left and
476 right lungs of all patients. This percentage was calculated by taking the features that overlap in
477 comparison to the total number of features associated with each pair of lungs.
478

		P1	P1	P2	P2	P3	P3
		Left	Right	Left	Right	Left	Right
P1	Left	100	88.02	75.71	76.92	70.46	72.7
P1	Right	88.02	100	71.67	72.39	74.01	76.73
P2	Left	75.71	71.67	100	89.46	67.91	70.54
P2	Right	76.92	72.39	89.46	100	70.35	72.74
P3	Left	70.46	74.01	67.91	70.35	100	91.56
P3	Right	72.7	76.73	70.54	72.74	91.56	100

479
480

481 **Materials and Methods.**

482 **Tissue collection and processing.** To map the microbiome and metabolome of explanted

483 lungs in 3D, the lungs of three patients were obtained in close coordination with the patient's

484 physician and the surgical team. This work was approved by the University of California

485 Institutional Review Board (project #081500) and informed consents were obtained prior to

486 tissue collection. The CFTR mutation in Patient 1 was dF508/G551D with no clinical diabetes, in

487 patient 2 was dF508/3120+1G>A with observed clinical diabetes and in patient 3 was

488 dF508/dF508 with observed clinical diabetes. The general workflow for tissue sectioning is

489 described previously (16). Briefly, both the right and left lungs for subject 1, 2, and 3 were

490 collected. The tissue sectioning was performed at the hospital under the guidance of a

491 pathologist. The lungs were first sliced horizontally. The anatomical orientation of each slice was

492 recorded. Every alternate slice starting from apex of the lung was further sub-sectioned into

493 small sections 1-2 cm^3 in size maintaining the recorded orientation. Each of the sub-sectioned

494 tissue pieces were swabbed with sterile soft foam swabs moistened with Tris-EDTA, pH 7.4.

495 The swabs were stored in 96-well bead plate provided in the PowerSoil®-htp 96 Well Soil DNA

496 Isolation Kit. The plate was placed on dry ice prior and during the collection. The individual

497 tissue pieces were stored in glass jars placed on dry ice. The samples were kept frozen at -80

498 °C until further processing. Bacterial DNA was isolated from the swabs using the PowerSoil®-

499 htp 96 Well Soil DNA Isolation Kit following the manufacturer's instructions and was subjected to

500 prokaryotic ribosomal 16S rRNA-based sequencing using the standardized Earth Microbiome

501 Protocol (http://www.earthmicrobiome.org/emp-standard-protocols/). Amplicons were cleaned,

502 pooled and then sequenced on an Illumina MiSeq. Because lungs were obtained at different

503 times, the sequences analyzed for this study were obtained from a total of two individual

504 sequencing runs (sequencing run 1: patients 1-2, sequencing run 2: patient 3). The sequencing

505 runs were performed at the University of California San Diego Institute for Genomic Medicine

506 Center. For untargeted metabolomics analysis, the tissue sections were weighed and extracted

507   with 1 mL/g of tissue with a 2:2:1 mixture of ethyl acetate, methanol and water. An aliquot of 150

508   µL of the extract was dried for each tissue section and analyzed by MS.

509

510   **MS data acquisition.** The tissue extracts and extracts of bacterial isolates from the subjects

511   cultured on sheep blood agar and MacConkey agar were resuspended in 80% methanol

512   containing 1µM sulfadimethoxine and analyzed with a UltiMate 3000 UHPLC system (Thermo

513   Scientific) using a Kinetex™ 1.7 µm C18 reversed phase UHPLC column (50 X 2.1 mm) and

514   Maxis Q-TOF mass spectrometer (Bruker Daltonics) equipped with ESI source. The column was

515   equilibrated with 2% solvent B (98% acetonitrile, 0.1% formic acid in LC-MS grade water with

516   solvent A as 0.1% formic acid in water) for 1 min, followed by a linear gradient from 2% B to

517   100% B in 10 min, held at 100% B for 2.5 min. A small wash segment was employed to wash

518   the column (100% B for 0.5 min, 100%-10 % B in 0.5 min) following which the column was kept

519   at 2% B for A min at a flow rate of 0.5 mL/min throughout the run. MS spectra were acquired in

520   positive ion mode in the range of 50-2000 *m/z.* A mixture of 10 µg/mL of each sulfamethazine,

521   sulfamethizole, sulfachloropyridazine, sulfadimethoxine, amitriptyline, and coumarin-314 was

522   run after every eight injections for quality control. An external calibration with ESI-L Low

523   Concentration Tuning Mix (Agilent technologies) was performed prior to data collection and

524   internal calibrant Hexakis(1H,1H,3H-tetrafluoropropoxy)phosphazene was used throughout the

525   runs. The capillary voltage of 4500 V, nebulizer gas pressure (nitrogen) of 2 bar, ion source

526   temperature of 200 °C, dry gas flow of 9 L/min source temperature, spectral rate of 3 Hz for MS[1]

527   and 10 Hz for MS[2] was used. For acquiring MS/MS fragmentation, the 10 most intense ions per

528   MS[1] were selected and collision induced dissociation energy given in Table 1 was used. Basic

529   stepping function was used to fragment ions at 50% and 125% of the CID calculated for each

530   *m/z* from Table 1 with timing of 50% for each step. Similarly, basic stepping of collision RF of

531   550 and 800 Vpp with a timing of 50% for each step and transfer time stepping of 57 and 90 µs

532   with a timing of 50% for each step was employed. MS/MS active exclusion parameter was set to

533   3 and released after 30 seconds. The mass of internal calibrant was excluded from the MS/MS

534   list using a mass range of *m/z* 921.5–924.5. The data was deposited in the online repository

535   namely MassIVE and is available under the id MSV000079652 and MSV000079398.

536   The microbial isolates of *Pseudomonas aeruginosa, Staphylococcus aureus* and

537   *Stenotrophomonas maltophila* from the patients were obtained from the Center of Advanced

538   Clinical Medicine, UC San Diego. The culturing of the isolates and the extractions were

539   performed as described previously (16). The MS data was collected using the same conditions

540   as described above for lung tissue.

541

542   **LC-MS/MS data analysis.** All mzXML files were cropped with an *m/z* range of 50.00 to

543   2,000.00 Da and RT range of 0.5 - 18.5 min. Feature extraction was performed using MZmine2

544   (http://mzmine.sourceforge.net/) with a signal height threshold of 5.0e3 (41). The mass

545   tolerance was set to 10 ppm, and the maximum allowed retention time deviation was set to 0.01

546   min. For chromatographic deconvolution, the local minimum search algorithm was used with a

547   minimum relative peak height of 1% and a minimum retention time range of 0.01 min. The

548   maximum peak width was set to 1 min. After isotope peak removal, the peak lists of all samples

549   were aligned with the above-mentioned retention time and mass tolerances. After the creation of

550   a feature matrix containing the feature retention times and the exact mass and peak areas of

551   the corresponding extracted ion chromatograms, the metadata of the samples were added. The

552   signal intensities of the features were normalized (probabilistic quotient normalization [PQN])

553   (42).

554

555   Statistical analysis was carried out as follows: QIIME 1.9.1 was used to perform principal-

556   coordinate analysis (PCoA) (beta_diversity.py, a Canberra distance metric in Adkins form). The

557   PCoA plots were visualized in EMPeror (28).

558

559   **Molecular networking.** The molecular network was created using the online workflow at GNPS

560   platform. The data was then clustered with MS-Cluster with a parent mass tolerance of 0.1 Da

561   and a MS/MS fragment ion tolerance of 0.1 Da to create consensus spectra. Further, consensus

562   spectra that contained less than 3 spectra were discarded. A network was then created where

563   edges were filtered to have a cosine score above 0.7 and more than 4 matched peaks. The

564   edges between two nodes were kept in the network if and only if each of the nodes appeared in

565   each other's respective top 10 most similar nodes. The spectra in the network were then

566   searched against GNPS's spectral libraries. All matches kept between network spectra and

567   library spectra were required to have a score above 0.7 and at least 4 matched peaks. The

568   molecular networks and the parameters used are available at the links below:

569   The molecular network and parameters for the patient data are available at:

570   https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=6f92a21af31d4569bcdb3cce803c600c

571   The molecular network and parameters for the patient data and data acquired on cultured

572   microbial isolates are available at:

573   https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=45d70e56faae4081bbba1f7a9ce38019

574   In total, 1776 of the nodes were annotated (7.8%) which is higher than the typical rate of

575   annotations of 1.8% annotated in an untargeted metabolomics experiment (30). This is likely

576   because many of the reference MS/MS libraries in the public domain are populated from studies

577   of human samples and also contain most of therapeutics used in the clinic. The error rate of

578   these annotations have been assessed by the GNPS community; with the scoring settings used

579   to obtain the annotations, 1% of which are deemed incorrect, for 4% not enough information is

580   available and 4% could be an isomer or correct, while 91% is presumed correct (30).

581

582   **16S rRNA gene analysis.** As described above, sequences were obtained over the course of

583   two months through two independent sequencing runs. The samples for patient 1 and patient 2

584   were sequenced in one batch and samples for patient 3 were sequenced separately. Each set

39

585   of sequences were processed and analyzed using Qiita (43). First, the sequencing runs were

586   quality trimmed and filtered using default parameters, resulting in 15,629,914 sequences with a

587   mean length of 150 nucleotides. Next, trimmed sequences (at 150 nucleotides) were clustered

588   into operational taxonomic units (OTUs) using the closed reference OTU picking method at 97%

589   sequence similarity. UCLUST was the underlying clustering algorithm and Greengenes (August

590   2013 release) was the reference database used (44). This resulted in 340 samples with a mean

591   of 25,078 sequences per sample. After rarefaction at 3,369 sequences per sample 277 samples

592   were used for downstream analyses, including the creation of taxonomy summaries and in the

593   calculation of the UniFrac and Canberra (Adkins form) distances. The most abundant OTU for

594   patient 2 was identified as unclassified genus in the family Xanthomonadaceae. BLAST analysis

595   of the sequence corresponding to this OTU from patient 2 revealed that it belongs to the genus

596   *Stenotrophomonas*. As controls, a total of 49 wells (either containing blank swab or empty well)

597   were interleaved between each of the two sampling sites (left and right lung) of the three

598   subjects. The vast majority of the samples (72%) yielded zero sequences. The remaining 14

599   samples had a non-zero amount of sequences. Of these, 7 samples were represented by under

600   4 sequences, a negligible amount compared to the 3,500 sequences per sample used for

601   analysis. And the last 7 samples were represented by over 6,000 sequences each. Although the

602   last sample set was processed without any DNA, the well-to-well contamination during the DNA

603   extraction step yielded these sequences. We removed these samples since the DNA is

604   biological and not representative of a type of actionable contamination (45, 46).

605

606   For statistical analysis, QIIME2 (47) was used to perform PCoA (Canberra (in Adkins form),

607   weighted, and unweighted UniFrac distances (weighted UniFrac distances (48)) and Procrustes

608   analysis with metabolomics data. The PCoA and Procrustes plots were visualized in EMPeror

609   (28). The Mantel test was used to calculate r^2 scores between mass spectrometry and for both

610   closed-reference and deblur 16S rRNA gene analysis data using scikit-bio's 0.5.5 Mantel's test

611   implementation.

612   **3D lung model generation and visualization.** The procedure for creation and visualization of

613   3D models has been previously described (16). Briefly, the CT- scan images obtained from the

614   radiology department at the Hillcrest hospital in San Diego were combined to create a 3D lung

615   model and exported in the .stl format using InVesalius 3.0. The extraneous pixels corresponding

616   to the chest and back of each model were manually deleted using the 3D modelling software

617   Geomagic Wrap. The relative abundances of detected microbes and molecules were plotted on

618   to   these   models   using   a   modified   version   of   the   `ili   software   available   at

619   http://mingwangbeta.ucsd.edu/public/ili/ (16, 49).

620

621   **Data availability statement.** All data presented in this manuscript is publically available. The

622   metabolomics data was deposited in the online repository namely MassIVE and is available

623   under the id MSV000079652 and MSV000079398. The molecular network analysis and

624   parameters for the patient data are available at:

625   https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=6f92a21af31d4569bcdb3cce803c600c

626   The molecular network analysis and parameters for the patient data and data acquired on

627   cultured microbial isolates are available at:

628   https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=45d70e56faae4081bbba1f7a9ce38019

629   All raw and processed 16S amplicon sequencing data and metadata are available with Qiita

630   study identification 10169 and as EBI study with accession no. ERP110498.

631   All figures in this manuscript have associated raw data, which is available through above

632   described accession numbers.

633

41

634   **Code availability Statement.** The code for 3D mapping via the browser tool github `ili is

635   available at `https://github.com/mwang87/ili and the tool is accessed via the link

636   http://mingwangbeta.ucsd.edu/public/ili/

References

639

640   1.   Laxminarayan R, Duse A, Wattal C, Zaidi AK, Wertheim HF, Sumpradit N, Vlieghe E, Hara
641   GL, Gould IM, Goossens H, Greko C, So AD, Bigdeli M, Tomson G, Woodhouse W,
642   Ombaka E, Peralta AQ, Qamar FN, Mir F, Kariuki S, Bhutta ZA, Coates A, Bergstrom R,
643   Wright GD, Brown ED, Cars O. 2013. Antibiotic resistance-the need for global solutions.
644   Lancet Infect Dis 13:1057-98.

645   2.   Shallcross LJ, Davies DS. 2014. Antibiotic overuse: a key driver of antimicrobial
646   resistance. Br J Gen Pract 64:604-5.

647   3.   Ventola CL. 2015. The antibiotic resistance crisis: part 1: causes and threats. P T 40:277-
648   83.

649   4.   Mollenkopf DF, Stull JW, Mathys DA, Bowman AS, Feicht SM, Grooters SV, Daniels JB,
650   Wittum TE. 2017. Carbapenemase-Producing Enterobacteriaceae Recovered from the
651   Environment of a Swine Farrow-to-Finish Operation in the United States. Antimicrob
652   Agents Chemother 61:e01298-16.

653   5.   de Kraker ME, Stewardson AJ, Harbarth S. 2016. Will 10 Million People Die a Year due to
654   Antimicrobial Resistance by 2050? PLoS Med 13:e1002184.

655   6.   Woolhouse M, Waugh C, Perry MR, Nair H. 2016. Global disease burden due to
656   antibiotic resistance - state of the evidence. J Glob Health 6:010306.

657   7.   Venkataraman A, Rosenbaum MA, Werner JJ, Winans SC, Angenent LT. 2014. Metabolite
658   transfer with the fermentation product 2,3-butanediol enhances virulence by
659   Pseudomonas aeruginosa. ISME J 8:1210-20.

660   8.   Nougayrede JP, Homburg S, Taieb F, Boury M, Brzuszkiewicz E, Gottschalk G, Buchrieser
661   C, Hacker J, Dobrindt U, Oswald E. 2006. *Escherichia coli* induces DNA double-strand
662   breaks in eukaryotic cells. Science 313:848-51.

663   9.   Chankhamjon P, Javdan B, Lopez J, Hull R, Chatterjee S, Donia MS. 2019. Systematic
664   mapping of drug metabolism by the human gut microbiome. doi:10.1101/538215 %J
665   bioRxiv:538215.

666   10. Valm AM, Mark Welch JL, Rieken CW, Hasegawa Y, Sogin ML, Oldenbourg R, Dewhirst
667   FE, Borisy GG. 2011. Systems-level analysis of microbial community organization
668   through combinatorial labeling and spectral imaging. Proc Natl Acad Sci U S A 108:4152-
669   7.

670   11. Stacy A, Everett J, Jorth P, Trivedi U, Rumbaugh KP, Whiteley M. 2014. Bacterial fight-
671   and-flight responses enhance virulence in a polymicrobial infection. Proc Natl Acad Sci U
672   S A 111:7819-24.

673   12. Mark Welch JL, Rossetti BJ, Rieken CW, Dewhirst FE, Borisy GG. 2016. Biogeography of a
674   human oral microbiome at the micron scale. Proc Natl Acad Sci U S A 113:E791-800.

675   13. Jorth P, Staudinger BJ, Wu X, Hisert KB, Hayden H, Garudathri J, Harding CL, Radey MC,
676   Rezayat A, Bautista G, Berrington WR, Goddard AF, Zheng C, Angermeyer A, Brittnacher
677   MJ, Kitzman J, Shendure J, Fligner CL, Mittler J, Aitken ML, Manoil C, Bruce JE, Yahr TL,
678   Singh PK. 2015. Regional Isolation Drives Bacterial Diversification within Cystic Fibrosis
679   Lungs. Cell Host Microbe 18:307-19.

680   14.   Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Beck JM, Huffnagle GB, Curtis
681   JL. 2015. Spatial Variation in the Healthy Human Lung Microbiome and the Adapted
682   Island Model of Lung Biogeography. Ann Am Thorac Soc 12:821-30.
683   15.   Chung H, Lieberman TD, Vargas SO, Flett KB, McAdam AJ, Priebe GP, Kishony R. 2017.
684   Global and local selection acting on the pathogen Stenotrophomonas maltophilia in the
685   human lung. Nat Commun 8:14078.
686   16.   Garg N, Wang M, Hyde E, da Silva RR, Melnik AV, Protsyuk I, Bouslimani A, Lim YW,
687   Wong R, Humphrey G, Ackermann G, Spivey T, Brouha SS, Bandeira N, Lin GY, Rohwer F,
688   Conrad DJ, Alexandrov T, Knight R, Dorrestein PC. 2017. Three-Dimensional Microbiome
689   and Metabolome Cartography of a Diseased Human Lung. Cell host & microbe 22:705-
690   716 e4.
691   17.   Aurora P, Wade A, Whitmore P, Whitehead B. 2000. A model for predicting life
692   expectancy of children with cystic fibrosis. Eur Respir J 16:1056-60.
693   18.   Szwed A, John A, Gozdzik-Spychalska J, Czainski W, Czerniak W, Ratajczak J, Batura-
694   Gabryel H. 2018. Survival of Patients with Cystic Fibrosis Depending on Mutation Type
695   and Nutritional Status. Adv Exp Med Biol 1023:65-72.
696   19.   Caverly LJ, Zhao J, LiPuma JJ. 2015. Cystic fibrosis lung microbiome: opportunities to
697   reconsider management of airway infection. Pediatr Pulmonol 50 Suppl 40:S31-8.
698   20.   Bumann D. 2015. Heterogeneous host-pathogen encounters: act locally, think globally.
699   Cell Host Microbe 17:13-9.
700   21.   Zlosnik JE, Zhou G, Brant R, Henry DA, Hird TJ, Mahenthiralingam E, Chilvers MA, Wilcox
701   P, Speert DP. 2015. Burkholderia species infections in patients with cystic fibrosis in
702   British Columbia, Canada. 30 years' experience. Ann Am Thorac Soc 12:70-8.
703   22.   Whiteson KL, Bailey B, Bergkessel M, Conrad D, Delhaes L, Felts B, Harris JK, Hunter R,
704   Lim YW, Maughan H, Quinn R, Salamon P, Sullivan J, Wagner BD, Rainey PB. 2014. The
705   upper respiratory tract as a microbial source for pulmonary infections in cystic fibrosis.
706   Parallels from island biogeography. Am J Respir Crit Care Med 189:1309-15.
707   23.   Rogers GB, Marsh P, Stressmann AF, Allen CE, Daniels TV, Carroll MP, Bruce KD. 2010.
708   The exclusion of dead bacterial cells is essential for accurate molecular analysis of
709   clinical samples. Clin Microbiol Infect 16:1656-8.
710   24.   Willner D, Haynes MR, Furlan M, Schmieder R, Lim YW, Rainey PB, Rohwer F, Conrad D.
711   2012. Spatial distribution of microbial communities in the cystic fibrosis lung. ISME J
712   6:471-4.
713   25.   Goddard AF, Staudinger BJ, Dowd SE, Joshi-Datar A, Wolcott RD, Aitken ML, Fligner CL,
714   Singh PK. 2012. Direct sampling of cystic fibrosis lungs indicates that DNA-based
715   analyses of upper-airway specimens can misrepresent lung microbiota. Proc Natl Acad
716   Sci U S A 109:13769-74.
717   26.   Coburn B, Wang PW, Diaz Caballero J, Clark ST, Brahma V, Donaldson S, Zhang Y,
718   Surendra A, Gong Y, Elizabeth Tullis D, Yau YC, Waters VJ, Hwang DM, Guttman DS.
719   2015. Lung microbiota across age and disease stage in cystic fibrosis. Sci Rep 5:10241.
720   27.   Quinn RA, Whiteson K, Lim YW, Zhao J, Conrad D, LiPuma JJ, Rohwer F, Widder S. 2016.
721   Ecological networking of cystic fibrosis lung infections. NPJ Biofilms Microbiomes 2:4.
722   28.   Vazquez-Baeza Y, Pirrung M, Gonzalez A, Knight R. 2013. EMPeror: a tool for visualizing
723   high-throughput microbial community data. Gigascience 2:16.

724 29. Hassett DJ, Cuppoletti J, Trapnell B, Lymar SV, Rowe JJ, Yoon SS, Hilliard GM, Parvatiyar
725   K, Kamani MC, Wozniak DJ, Hwang SH, McDermott TR, Ochsner UA. 2002. Anaerobic
726   metabolism and quorum sensing by Pseudomonas aeruginosa biofilms in chronically
727   infected cystic fibrosis airways: rethinking antibiotic treatment strategies and drug
728   targets. Adv Drug Deliv Rev 54:1425-43.
729 30. Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, Nguyen DD, Watrous J,
730   Kapono CA, Luzzatto-Knaan T, Porto C, Bouslimani A, Melnik AV, Meehan MJ, Liu WT,
731   Crusemann M, Boudreau PD, Esquenazi E, Sandoval-Calderon M, Kersten RD, Pace LA,
732   Quinn RA, Duncan KR, Hsu CC, Floros DJ, Gavilan RG, Kleigrewe K, Northen T, Dutton RJ,
733   Parrot D, Carlson EE, Aigle B, Michelsen CF, Jelsbak L, Sohlenkamp C, Pevzner P, Edlund
734   A, McLean J, Piel J, Murphy BT, Gerwick L, Liaw CC, Yang YL, Humpf HU, Maansson M,
735   Keyzers RA, Sims AC, Johnson AR, Sidebottom AM, Sedio BE, et al. 2016. Sharing and
736   community curation of mass spectrometry data with Global Natural Products Social
737   Molecular Networking. Nat Biotechnol 34:828-837.
738 31. Watrous J, Roach P, Alexandrov T, Heath BS, Yang JY, Kersten RD, van der Voort M,
739   Pogliano K, Gross H, Raaijmakers JM, Moore BS, Laskin J, Bandeira N, Dorrestein PC.
740   2012. Mass spectral molecular networking of living microbial colonies. Proc Natl Acad
741   Sci U S A 109:E1743-52.
742 32. Quinn RA, Phelan VV, Whiteson KL, Garg N, Bailey BA, Lim YW, Conrad DJ, Dorrestein PC,
743   Rohwer FL. 2016. Microbial, host and xenobiotic diversity in the cystic fibrosis sputum
744   metabolome. ISME J 10:1483-98.
745 33. Whitaker P, Meng X, Lavergne SN, El-Ghaiesh S, Monshi M, Earnshaw C, Peckham D,
746   Gooi J, Conway S, Pirmohamed M, Jenkins RE, Naisbitt DJ, Park BK. 2011. Mass
747   spectrometric characterization of circulating and functional antigens derived from
748   piperacillin in patients with cystic fibrosis. J Immunol 187:200-11.
749 34. Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, Kightley EP,
750   Thompson LR, Hyde ER, Gonzalez A, Knight R. 2017. Deblur Rapidly Resolves Single-
751   Nucleotide Community Sequence Patterns. mSystems 2:e00191-16.
752 35. Quinn RA, Lim YW, Mak TD, Whiteson K, Furlan M, Conrad D, Rohwer F, Dorrestein P.
753   2016. Metabolomics of pulmonary exacerbations reveals the personalized nature of
754   cystic fibrosis disease. PeerJ 4:e2174.
755 36. Lepine F, Milot S, Deziel E, He J, Rahme LG. 2004. Electrospray/mass spectrometric
756   identification and analysis of 4-hydroxy-2-alkylquinolines (HAQs) produced by
757   Pseudomonas aeruginosa. J Am Soc Mass Spectrom 15:862-9.
758 37. Frydenlund Michelsen C, Hossein Khademi SM, Krogh Johansen H, Ingmer H, Dorrestein
759   PC, Jelsbak L. 2016. Evolution of metabolic divergence in Pseudomonas aeruginosa
760   during long-term infection facilitates a proto-cooperative interspecies interaction. ISME
761   J 10:1323-36.
762 38. Kakinuma R, Moriyama N, Muramatsu Y, Gomi S, Suzuki M, Nagasawa H, Kusumoto M,
763   Aso T, Muramatsu Y, Tsuchida T, Tsuta K, Maeshima AM, Tochigi N, Watanabe S,
764   Sugihara N, Tsukagoshi S, Saito Y, Kazama M, Ashizawa K, Awai K, Honda O, Ishikawa H,
765   Koizumi N, Komoto D, Moriya H, Oda S, Oshiro Y, Yanagawa M, Tomiyama N, Asamura
766   H. 2015. Ultra-High-Resolution Computed Tomography of the Lung: Image Quality of a
767   Prototype Scanner. PLoS One 10:e0137165.

768   39.   Hu W, Lum GZ, Mastrangeli M, Sitti M. 2018. Small-scale soft-bodied robot with
769   multimodal locomotion. Nature 554:81-85.
770   40.   Lozupone CA, Hamady M, Kelley ST, Knight R. 2007. Quantitative and qualitative beta
771   diversity measures lead to different insights into factors that structure microbial
772   communities. Appl Environ Microbiol 73:1576-85.
773   41.   Pluskal T, Castillo S, Villar-Briones A, Oresic M. 2010. MZmine 2: modular framework for
774   processing, visualizing, and analyzing mass spectrometry-based molecular profile data.
775   BMC Bioinformatics 11:395.
776   42.   Dieterle F, Ross A, Schlotterbeck G, Senn H. 2006. Probabilistic quotient normalization as
777   robust method to account for dilution of complex biological mixtures. Application in 1H
778   NMR metabonomics. Anal Chem 78:4281-90.
779   43.   Gonzalez A, Navas-Molina JA, Kosciolek T, McDonald D, Vazquez-Baeza Y, Ackermann G,
780   DeReus J, Janssen S, Swafford AD, Orchanian SB, Sanders JG, Shorenstein J, Holste H,
781   Petrus S, Robbins-Pianka A, Brislawn CJ, Wang M, Rideout JR, Bolyen E, Dillon M,
782   Caporaso JG, Dorrestein PC, Knight R. 2018. Qiita: rapid, web-enabled microbiome
783   meta-analysis. Nat Methods 15:796-798.
784   44.   Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST.
785   Bioinformatics 26:2460-1.
786   45.   Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS. 2019. Contamination
787   in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. Trends
788   Microbiol 27:105-117.
789   46.   Minich JJ, Sanders JG, Amir A, Humphrey G, Gilbert J, Knight R. 2019. Quantifying and
790   understanding well-to-well contamination in microbiome research. doi:10.1101/577718
791   %J bioRxiv:577718.
792   47.   Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N,
793   Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE,
794   Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ,
795   Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows
796   analysis of high-throughput community sequencing data. Nat Methods 7:335-6.
797   48.   Lozupone C, Knight R. 2005. UniFrac: a new phylogenetic method for comparing
798   microbial communities. Appl Environ Microbiol 71:8228-35.
799   49.   Protsyuk I, Melnik AV, Nothias LF, Rappez L, Phapale P, Aksenov AA, Bouslimani A,
800   Ryazanov S, Dorrestein PC, Alexandrov T. 2018. 3D molecular cartography using LC-MS
801   facilitated by Optimus and 'ili software. Nat Protoc 13:134-154.
802