bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1 **Taxonomic and functional dynamics of lung microbiome in cystic**
2 **fibrosis patients chronically infected with *Pseudomonas aeruginosa***

3 Giovanni Bacci[1], Giovanni Taccetti[2], Daniela Dolce[2], Federica Armanini[3], Nicola Segata[3],
4 Francesca Di Cesare[1], Vincenzina Lucidi[4], Ersilia Fiscarelli[4], Patrizia Morelli[5], Rosaria Casciaro[5],
5 Anna Negroni[6], Alessio Mengoni[1], and Annamaria Bevivino[6*]
6
7 [1]Department of Biology, University of Florence, Sesto Fiorentino, 50019, Florence, Italy,
8 [2]Cystic Fibrosis Center, Anna Meyer Children's University Hospital, Department of Pediatrics
9 Medicine, Florence,
10 50139, Italy
11 [3]Centre for Integrative Biology, University of Trento, Trento, 38122, Italy
12 [4]Children's Hospital and Research Institute Bambino Gesù, Rome, 00165, Italy
13 [5]Cystic Fibrosis Center, IRCCS G. Gaslini Institute, Department of Pediatrics, Genoa, 16147, Italy
14 [6]Department for Sustainability, Italian National Agency for New Technologies, Energy and
15 Sustainable Economic Development, ENEA, 00123 Rome, Italy
16
17 * Corresponding author
18 Annamaria Bevivino, Department for Sustainability, ENEA, Rome, Italy. Email:
19 annamaria.bevivino@enea.it

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

20 # Abstract

21 **Background**

22 Although the cystic fibrosis (CF) lung microbiome has been taxonomically defined, little is still
23 known about the overall gene content and functional profiles of the resident microbiome and how it
24 changes in relation with patient's disease status. The aim of this study was to investigate the
25 taxonomic and functional temporal dynamics of airways microbiome in CF patients with chronic
26 *Pseudomonas aeruginosa* infection. A shotgun metagenomic approach was used, to target the entire
27 genomic repertoire of the microbial community and allow strain-level profiling.

28 **Results**

29 A cohort of 22 patients with moderate-severe lung disease carrying the ΔF_{508} mutation was enrolled
30 and followed over 15 months. The taxonomic composition, metabolic repertoire and antibiotic
31 resistance gene content of lung microbiome was investigated by time-resolved shotgun
32 metagenomics. A high inter-patient taxonomic heterogeneity was found with short-term
33 compositional changes during period of exacerbation and following antibiotic treatment. Each
34 patient represented a unique environment showing distinct communities at the taxonomic level and
35 high patient-specific colonization by both main and emerging CF pathogens. The extraordinary
36 resilience of the CF microbiome was found even from a functional perspective. All samples showed
37 a core set of antibiotic resistance genes with a limited influence of antibiotic intake.

38 **Conclusions**

39 The main outcomes from this study indicate a patient-specific temporal dynamic of the microbiome
40 and a high functional resilience following exacerbation and antibiotic treatment.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

41 # Background

42 Cystic Fibrosis (CF) is the most common lethal autosomal recessive disease in Caucasians, caused
43 by mutations in the gene coding for cystic fibrosis transmembrane conductance regulator (CFTR)
44 channel [1]. Disruption of chloride anion transport, one of the key underlying features of CF, leads
45 to altered physiological conditions at epithelial surfaces. In the airways, CFTR mutations result in a
46 dehydrated viscous mucus that compromises mucociliary clearance and predisposes CF patients to
47 repeated cycles of airway infection, mucous impaction, and bronchiectasis resulting in the majority
48 of morbidity and mortality in the patient population [2]. In particular, bacterial lung infections
49 reduce life expectancy in most CF patients [3]. The affected individuals consistently maintain high
50 bacterial loads in their airways also during periods of clinical stability that are punctuated by
51 episodes of pulmonary exacerbation [4]. Such periodic episodes of acute pulmonary exacerbation
52 strongly contribute to the irreversible decline of lung function. Though much is known about the
53 composition of the microbial infections in CF (for a recent review see [5]), the factors leading to
54 such exacerbations are still poorly understood. In the past years, studies employing DNA-based
55 analyses of the airway microbiota of CF patients have shown somewhat discordant results. Indeed,
56 some authors report a largely stable airway microbiota through periods of exacerbation and
57 antibiotic treatment [6], while other indicate of a high inter-patients variability [5,7–9], but also
58 suggested the possibility to identify some microbial taxa as biomarker of exacerbation [10], as well
59 as a role of rare species in exacerbation [11]. Most of these works are targeted metagenomic surveys
60 performed on a variable number of patients and focusing on the 16S rRNA gene sequence.
61 However, this approach offers limited possibilities to infer strain-level and functional (meaning
62 based on functional genes) insights [12]. These two last points are particularly relevant when host-
63 microbiome interactions are studied. Indeed, the overall genetic repertoire of the microbiome (i.e.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

64 the entire set of genes in all the genomes of the community members) is the main responsible of the
65 interaction with the host [13]. Recently, the functional interactions among members of a bacterial
66 community have stirred the attention of investigators for relating microbiome functionality to
67 human-microbe interaction [14] and as a perspective for understanding the airway microbiome
68 dynamics in CF [15]. In several human diseases where the microbial infection is an important factor,
69 such as CF, single patients harbor genomically different strains, which ultimately may lead to
70 explain individual differences in clinical outcomes [16–18]. Until now, few longitudinal studies,
71 with a limited number of patients, on CF airway microbiota have been performed [19,20].
72 Moreover, studies on CF microbiome are few and on a limited number of patients [9,21–23] or
73 specific metabolic functions [24]. Moving away from taxonomic inventories towards a better
74 understanding of the CF microbiome genes opens a new avenue for the identification of the
75 microbial gene repertoire associated with CF lung disease. An ecological perspective on
76 multispecies and multi-strain colonization of CF airways will permit to understand the role of
77 polymicrobial dynamics in lung disease progression [25] and provide the clinicians with new
78 biomarkers of CF progression and targets for antibiotic therapy.
79 In this work. we tried to fill the gap of knowledge about the temporal dynamics of the airway
80 microbiome in CF, paying special attention to the episodes of exacerbation, by using a shotgun
81 metagenomic approach [26], that is targeting the entire genomic repertoire of the microbial
82 community, down to the strain level [27,28]. A cohort of 22 patients with moderate-severe lung
83 disease, grouped according to different genotypes (homozygote and heterozygote for ΔF_{508}
84 mutation), and chronically infected with *Pseudomonas aeruginosa*, was selected and followed over
85 15 months during which 8 patients underwent exacerbation events. This offered the opportunity to
86 investigate the taxonomic and functional dynamics of the overall microbiome.
87

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

88 # Results

89 **Population and sampling**

90 Twenty-two patients with CF were enrolled for a total of 15 females and seven males. The patients
91 were chosen from a larger cohort of patients with moderate-severe lung disease ($30 < \%FEV_1 < 70$)
92 and chronically infected by *Pseudomonas aeruginosa*. During the study period, they were treated
93 with maintenance antibiotics (aerosol) and only a subset (n = 8) received clinical intervention in
94 form of supplementary antibiotics (oral or/and intravenous) for a pulmonary exacerbation (CFPE)
95 (Table 1 and supplementary materials Table S1). The bacterial microbiome was investigated on
96 sputum samples obtained every 3-4 months from 22 individuals along a survey of 15 months.
97 Within the 22 subjects monitored, 8 underwent episodes of exacerbations, which provided the
98 opportunity to explore the microbiomes composition along the events. In total, 79 samples from
99 these 22 subjects were collected and analyzed by a whole metagenomic sequencing approach.

100 **Airway microbiomes are taxonomically distinct and show patient-specific strain colonization**

101 The overall taxonomic representation of the microbiomes from the 79 samples is reported in Fig. 1a
102 and 1b, whereas a summary of obtained reads per sample was reported in supplementary materials
103 Table S2. Firmicutes, Proteobacteria, Bacteroidetes, and Actinobacteria were the most represented
104 phyla. A massive presence of the "classical" CF bacterial signatures (taxa), such as *Staphylococcus*
105 *aureus* and *Pseudomonas aeruginosa*, and non-traditional CF taxa, such as *Rothia mucilaginosa*,
106 and *Prevotella melaninogenica* (all present in the top-10 species within each phylum, Fig. 1b), was
107 found. These species, indeed, represent the 49% of all detected taxa as reported in supplementary
108 materials Table S3.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

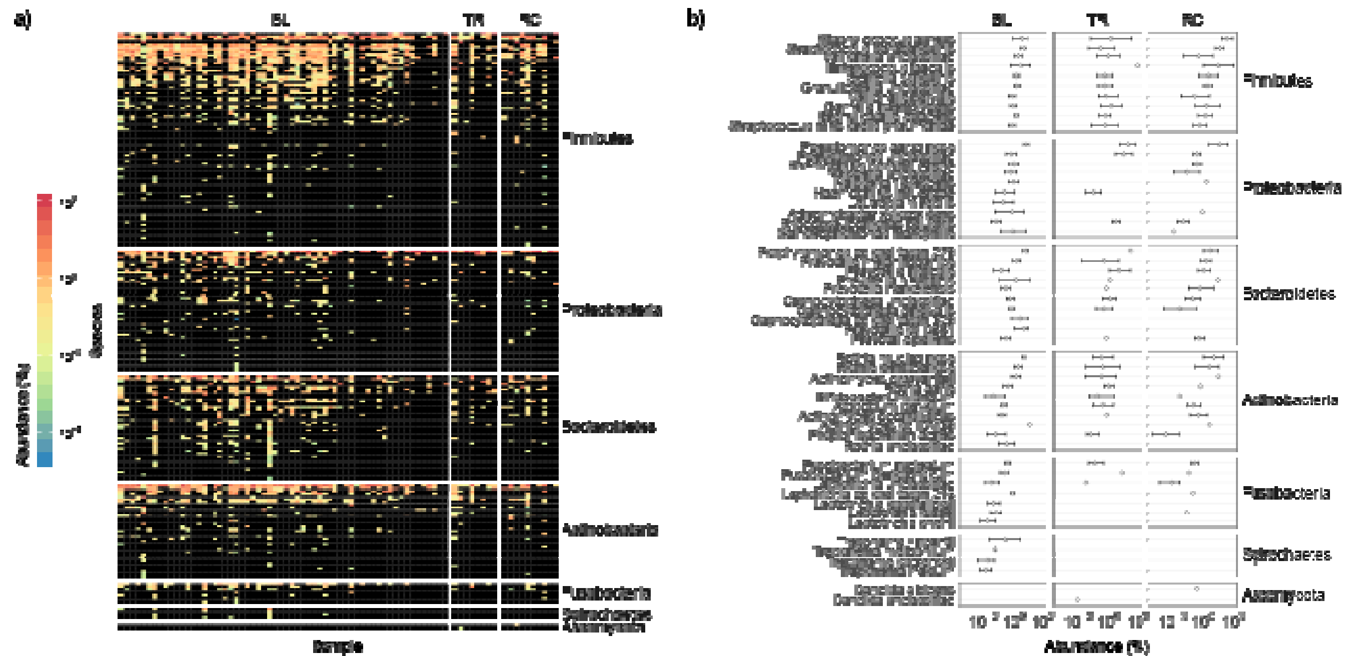
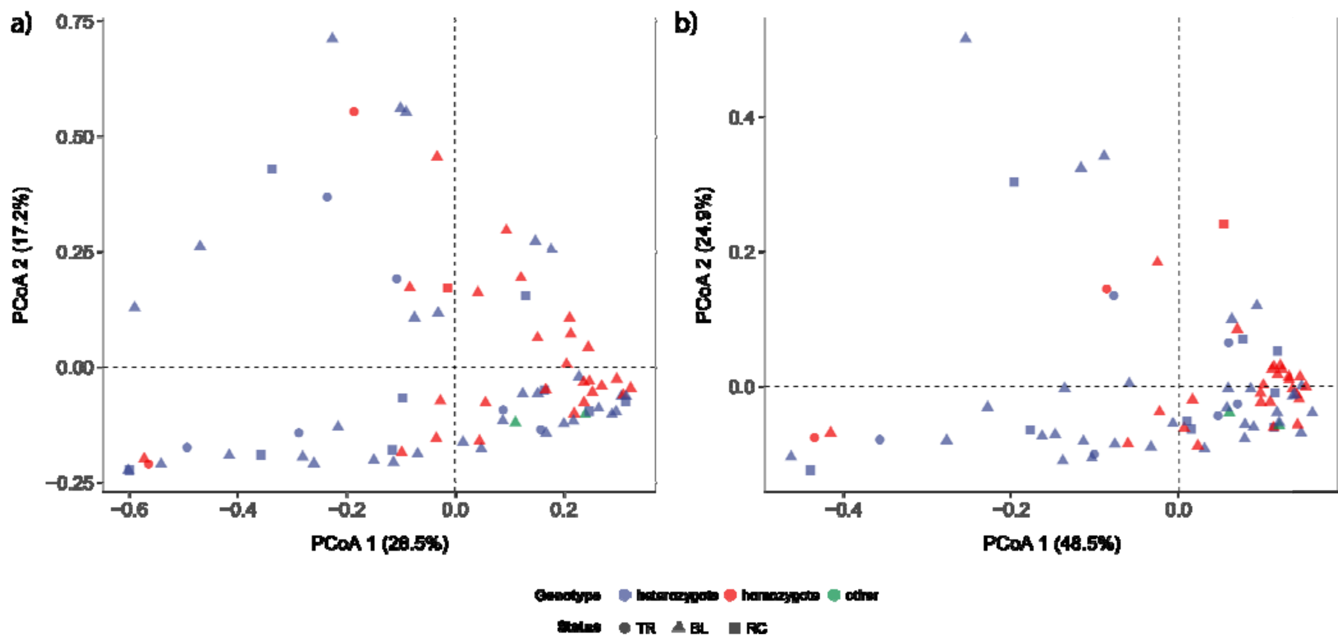


Figure 1: Taxonomic distribution in patients enrolled in the study. a) The taxonomic distribution of all species detected using MetaPhlAn2 was reported in each row of the matrix whereas columns represent samples collected during the study. Colors from dark blue to red were used to report "copies per million" (CPM) values as obtained from HUMAnN2 with black reporting a CPM value of zero. The plot was divided according to patient status: BL, baseline; TR, treatment; RC, recovery. Species were ordered according to their mean abundance and grouped according to their Phylum. b) The mean abundance value of the top-ten species (if available) detected within each Phylum was reported together with the standard error.

Although samples can be hardly clustered based on treatment events and/or genotype (Fig. 2a), the PERMANOVA analysis (Table 2) reported a significant effect of both factors. However, the R^2 values, namely the proportion of variance explained by the factor considered, were very low (0.03 for both factors). We cannot exclude that the heterogeneity of patients (especially concerning age) could have limited the statistical power of the analysis. The interaction effect between treatment events and genotype was not significant (p-value > 0.05), meaning that different genotypes did not influence the lung microbiome during treatment (viz. exacerbation) events and vice versa. The predominant effect observed was the subject effect, reporting a R^2 value of 0.52, indicating that a

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

126 high fraction (more than 50%) of the total variance can be explained by subject (patient)
127 individuality. Both FEV_1 and time did not show any significant effect (Table 2).



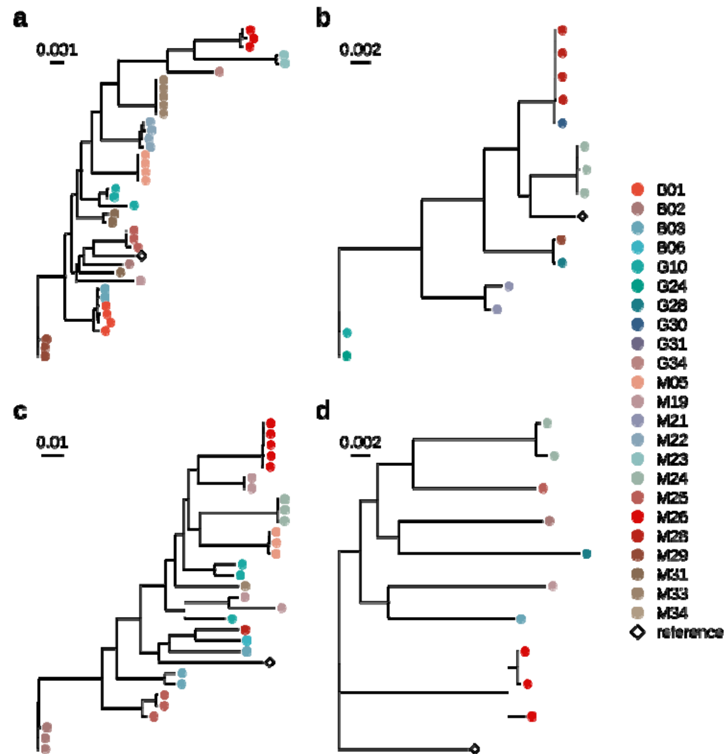
128

129 **Figure 2: Ordination analyses based on a) taxonomic assignments and b) pathway distribution**
130 **detected with MetaPhlAn2 and HUMAnN2, respectively.** Ordination analyses were conducted
131 using the Bray-Curtis dissimilarity index and ordered following the principle coordinate
132 decomposition method (PCoA). The percentage of variance explained by each coordinate was
133 reported between round brackets. Homozygote and heterozygote refer to ΔF_{508} mutation of CFTR
134 gene. BL, baseline; TR, treatment; RC, recovery.

135 With the aim to evaluate the relevance of subject (patient) individuality, a strain-level analysis was
136 conducted. Results obtained revealed that strain genotypes inferred from metagenomic data are
137 highly patient-specific for many bacterial species. Indeed, samples (i.e. strains) from the same
138 patients are very closely related and tightly clustered, confirming a high patient-specific colonization
139 by strains of the above-mentioned species (Fig. 3 and supplementary materials Fig. S1).

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

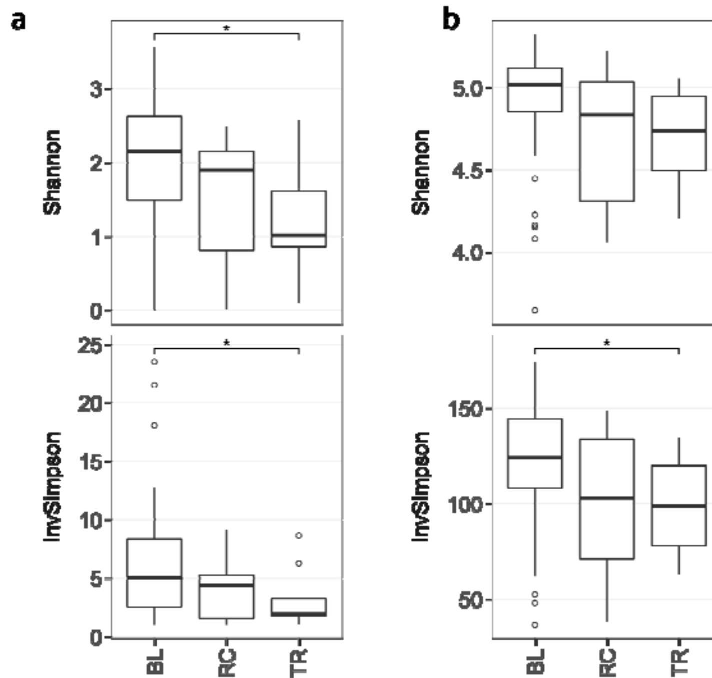
140



141 **Figure 3: Strain-level phylogenetic trees of the main CF pathogens detected in the study.**
142 Phylogenetic trees obtained through StrainPhlAn pipeline were reported for the main pathogenic
143 signatures of CF disease: a) *Pseudomonas aeruginosa*; b) *Staphylococcus aureus*; c) *Rothia*
144 *mucilaginosa*; d) *Prevotella melaninogenica*. Points at the end of each clade are colored according
145 to patients so that two points with the same color, in the same tree, represent the same species in two
146 different time points, for the same patient.

147 Alpha diversity analysis confirmed the overall picture of results mentioned above. Different values
148 of bacterial diversity were found according to clinical status, genotypes, and subjects (Fig. 4a,
149 supplementary materials Fig. S2 and Table S4). Samples collected during clinical treatments
150 reported a lower biodiversity than samples collected during normal visits, highlighting the role of
151 clinical treatments in perturbing CF lung communities as confirmed by the Tukey's post hoc test
152 (supplementary materials Table S5).

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



153

154 **Figure 4: Differences across exacerbation events.** The effect of an exacerbation event on alpha
155 diversity was inspected using both the Shannon index and the inverse Simpson index. Diversity
156 indexes were computed for both a) taxonomic signature and b) metabolic pathways. BL, baseline;
157 TR, treatment; RC, recovery. Each box shows the "interquartile range" (IQR) that is the differences
158 between the third and the first quartile of data (the 75th and the 25th percentile). Horizontal bars are
159 medians whereas whiskers represent the minimum and maximum values defined as Q1 – (1.5 x
160 IQR) and Q3 + (1.5 x IQR), respectively. Observations that fell outside minimum and maximum
161 values were defined as outliers and reported using white points.

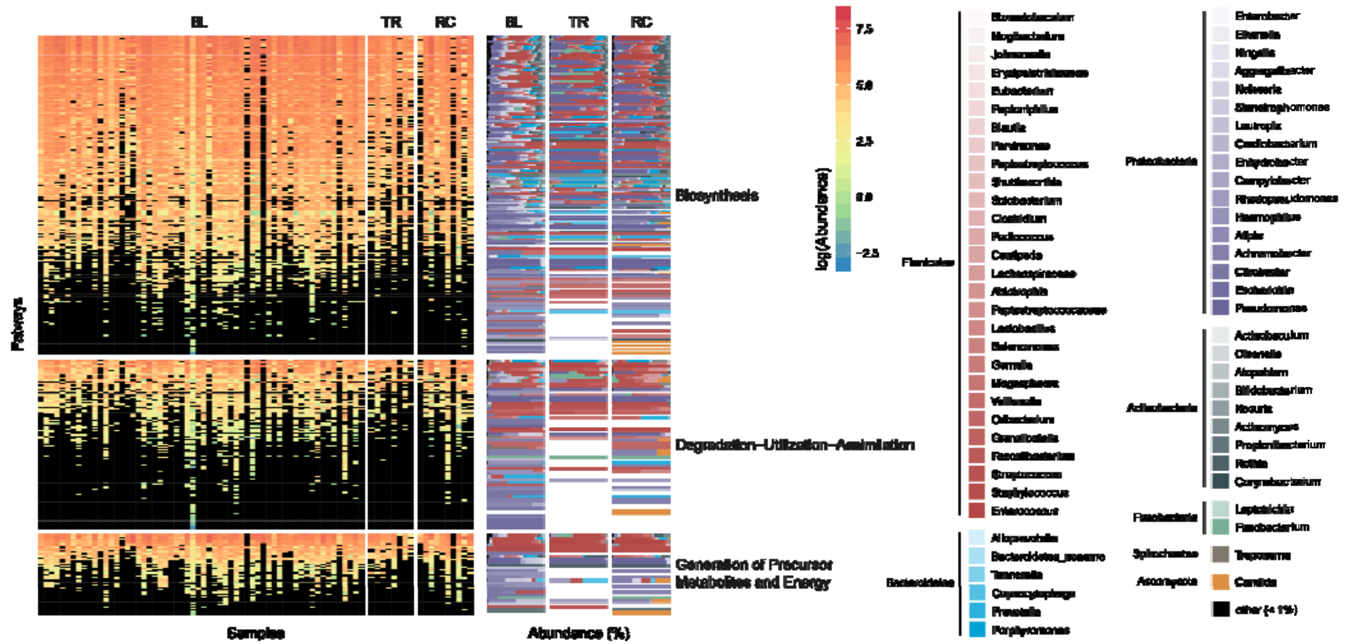
162 **Airway microbiomes are functionally consistent and show subject-specific distribution**
163 **patterns**

164 Similar results as those reported above were obtained considering the pathway distribution. Indeed,
165 the PERMANOVA analysis (Table 2) confirmed the effect of exacerbation events and genotypes in
166 shaping the pathway distribution of CF lung microbiome (R^2 values of 0.04 and 0.03 respectively),
167 though less marked than the subject-specific effect (R^2 = 0.48). The sample distribution according to
168 the ordination analysis (PCoA) was very heterogeneous with no sharp differences according to

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

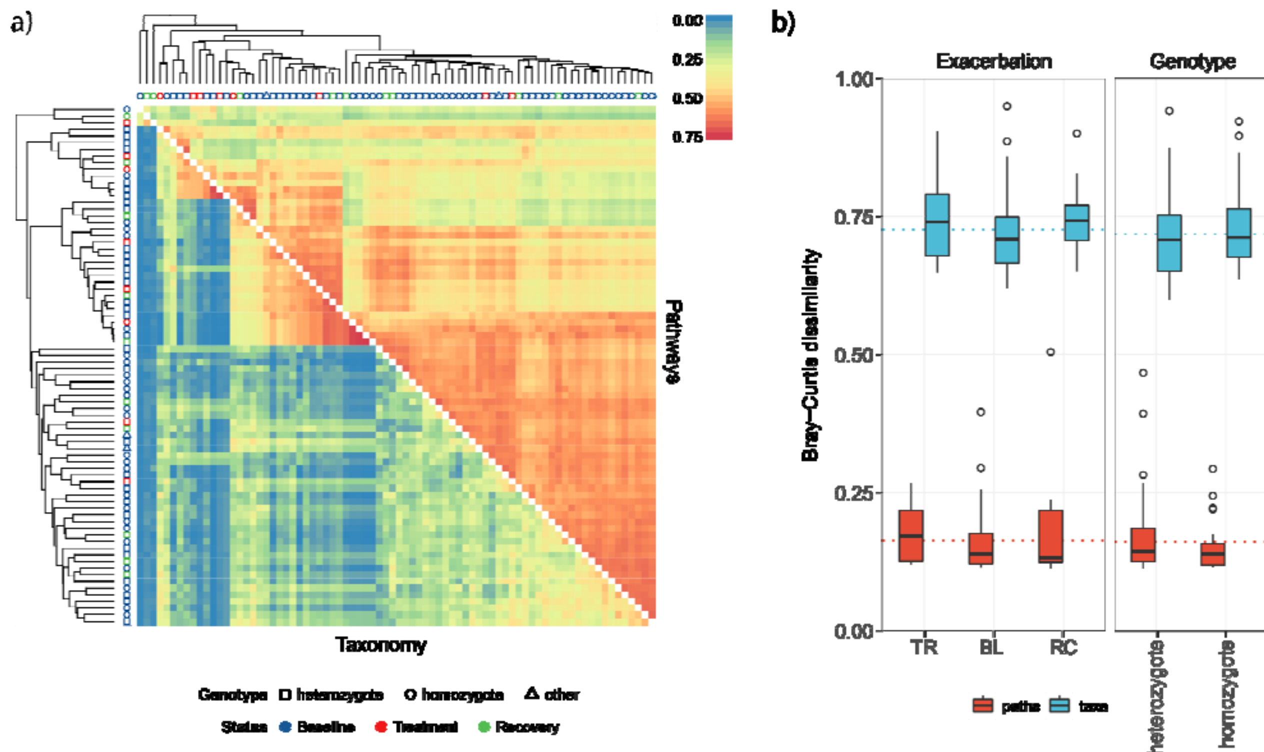
169 genotypes or exacerbation events. Even here, alpha diversity analyses reported a significant drop of
170 diversity in samples collected during exacerbation events, but the drop was significant only
171 considering the inverse Simpson index (Fig. 4b and supplementary materials Table S4). Overall, the
172 pathway distribution was more consistent with respect to the taxonomic one, with biosynthetic
173 pathways being the most represented functional category (Fig. 5, supplementary materials Fig. S2
174 and Table S5). Pathways were mainly detected in members of Firmicutes and Proteobacteria phyla,
175 though Bacteroidetes and Actinobacteria were quite well represented. Even if these results
176 confirmed the results from the analysis of the taxonomic distribution, metabolic pathways showed a
177 more consistent distribution across samples. Indeed, the beta-diversity analysis on both taxonomic
178 and functional distribution showed a lower similarity based on taxonomy in respect with pathways
179 (Supplementary materials Table S6, Fig. 6a and 6b). These results were additionally confirmed by
180 the differential abundance analysis. For contrasts made within each genotype, only 40 pathways
181 reported significant differences across exacerbation statuses (p-values < 0.05 and $|\log(\text{fold-change})|$
182 > 5) all in the homozygote group (Supplementary materials Fig. S3 and Table S7), whereas,
183 considering all samples together, no pathway was found to be more abundant in one condition in
184 respect with another (data not shown). These results confirmed the extraordinary resilience of the
185 CF microbiome even from a functional perspective.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

186


187 **Figure 5: Pathway distribution according to exacerbation events.** The pathway distribution was
188 reported for each sample (columns) and for each pathway detected (rows). Colors from dark blue to
189 red were used to report "copies per million" (CPM) values as obtained from HUMAnN2 with black
190 reporting a CPM value of zero. On the left, the percentage of taxa in which each pathway was
191 detected was reported using different colors. The main colors correspond to the Phylum whereas the
192 different shades correspond to the genus detected (if available). BL, baseline; TR, treatment; RC,
193 recovery.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



194

195 **Figure 6: Beta diversity analysis on both taxonomic and functional distribution.** a) Hierarchical
196 clustering based on UPGMA method. Clustering was performed on both pathway distribution (the
197 upped triangle) and taxonomic composition of samples (lower triangle). The Bray-Curtis distance
198 was used to compute distances between samples, but it was transformed into similarity value by
199 subtracting 1 before plotting. Thus, red colors represent high similarity values whereas blue colors
200 represent low similarity values. The shape of the points on each tip of trees refers to the hospital
201 whereas the colors refer to the exacerbation events. b) Results of Tukey's post hoc test on beta
202 diversity values across patient genotypes and exacerbation events. Contrasts were computed even to
203 test differences between taxonomic distribution and pathways with taxa reporting higher level of
204 beta diversity. Homozygote and heterozygote refer to ΔF_{508} mutation of CFTR gene. BL, baseline;
205 TR, treatment; RC, recovery. Boxplot were computed as described in Figure 4 legend.

206

207 **Antibiotic resistance genes through exacerbation events and treatments**

208 Similar to the pathway analysis reported above, antibiotic resistance genes (ARG) were inspected in
209 relation to treatment events. Only six genes were found to be affected by an exacerbation condition,

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

210 all regarding samples form ΔF_{508} heterozygote patients whereas, as found for metabolic pathways,
211 no gene was significantly impacted by antibiotic treatment when considering all samples at once
212 (supplementary materials Fig. S4 and Table S8). A similar approach was used to inspect the effect
213 of antibiotic treatment on ARG distribution. ARG were inspected in relation to the antibiotic
214 treatments reported in supplementary materials Table S1. The class of each antibiotic was correlated
215 to the presence (and the abundance) of genes that may, in principle, confer resistance to antibiotics
216 from the corresponding class. Differential abundance analyses were performed for each classes of
217 antibiotics that was used in this study and results obtained were reported in supplementary materials
218 Fig. S5 and Table S9. Only 11 genes were found to be affected by antibiotic intake in different
219 ways. Indeed, 8 out of 11 reported a reduction of abundance during the treatment whereas the
220 remaining 3 reported an increased abundance in respect with antibiotic intake. Results obtained
221 confirmed the high resilience of the gene composition of CF lung microbiome. The presence of
222 ARGs coupled with antibiotic intake was also explored. Results showed that the antibiotic resistance
223 classes of each gene corresponded to the antibiotic treatment used in each sample reporting a big
224 block of ARGs that were present in most of the sample considered (Supplementary materials Fig. S6
225 and Fig. S7).

226 Discussions

227 Longitudinal studies allow to provide important clues on stability and dynamics of microbial
228 ecosystems [29]. As all biotic communities, microbial communities tend to evolve towards a stable
229 composition, either in natural environment and in association with host (as human-associated
230 microbiomes). Changes in the community can be triggered by external conditions, as changes in
231 host physiology (e.g. inflammation status) and/or other perturbations (e.g. antibiotic treatment).
232 Indeed, perturbation studies help to probe community dynamics and resilience and possibly discover

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

233 new findings for accessing ways for modifying the microbiome [30,31]. Although patients with CF
234 experiments repeated episodes of pulmonary exacerbations during their lives, a broadly accepted
235 definition of these events is still missing [4]. Here, we have investigated the temporal dynamics of
236 CF airway microbiome by using shotgun metagenomics posing attention on exacerbation events
237 which usually bring to an acute decrease in lung function and an increase in respiratory symptoms
238 (such as: increased cough, sputum production, and shortness of breath). Key questions were i) what
239 was the composition and stability of the lung microbiome in patients with CF when longitudinally
240 sampled at stable and exacerbation events; and ii) if the clinical status influenced the metabolic
241 repertoire and the AR gene composition of lung bacterial community. Our results describe a unique
242 examination of the dynamic of the lung microbiome in patients with moderate-severe lung disease
243 carrying the ΔF_{508} mutation of CFTR gene and containing clinical measurements over a 15-month
244 period.,
245 The lung microbiome of CF patients seems to be a highly patient-specific environment which can be
246 directly conditioned by the host and its habits. Indeed, there was less variation within the same
247 individual at different time points than between different individuals at the same time point, proving
248 some degree of temporal stability of an individual's lung microbiome. This last point agrees with the
249 lack of a time effect on the taxonomic distribution of microbiome. The predominant taxa that
250 colonized the lung of CF patients showed an extraordinary resilience, as witnessed by the presence
251 of the same strains during the whole period of infection. These results agree with previous
252 observations based on 16S rRNA gene profiling, though these studies failed to report a strain-
253 specific overview of the whole dynamic due to the limitations intrinsic to the approach [6,8,11].
254 Carmody and colleagues showed a relatively stable lung community that may be altered during
255 period of exacerbation even in the absence of viral infection or antibiotic only in a small group of
256 patients [10]. Even in other pulmonary diseases, such as non-cystic fibrosis bronchiectasis, lung

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

257 bacterial communities showed a conserved structure for long period of time, as showed in the work
258 by Cox and colleagues where patients were followed for a six-month period [8]. A similar result was
259 shown in the work from Fodor and colleagues [6] where, though occasional short-term
260 compositional changes in the airway microbiota were found, the main taxonomic signatures of CF
261 disease were highly stable. A notable exception was found for *Rothia mucilaginosa*. In fact, in
262 contrast with other studies where it was rarely identified [15,32,33], in our samples it was detected
263 in high relative abundance. This finding may suggest a potential involvement of *Rothia*
264 *mucilaginosa* in CF microbiome dynamics and pathogenicity, which deserves further attention.
265 The antibiotic treatment used did not seem to alter this micro-environment for long period of time
266 since most of the main taxa linked to CF infection are still present even after exacerbation events
267 that are usually handled by a massive amount of antibiotic. From a taxonomic perspective, samples
268 coming from the same patient clustered together highlighting the role of the host in bacterial strain
269 selection during the baseline but even during (and after) exacerbation events. Despite this patient-
270 specific colonization, the taxonomic composition was very different from one subject to another
271 event if sampled at the same time point.
272 On the other hand, pathways reported a more homogeneous distribution across patients. This high
273 conservation could be related to the characteristic of the lung environment itself, such as mucus
274 compositions, nutrient availability, and oxygen levels, which can be broadly similar across patients
275 with a similar clinical status. This, is in line with the finding that the function of a biotic community
276 is more conserved than the presence of single members [34]. In fact, though the lung microbiome in
277 our study was populated with a relatively large set of microorganisms, the main functions detected
278 are similar across all patients. From this point of view the airway microbiome can be considered as
279 performing a similar "ecosystem service", irrespective of the taxonomy present as pointed out by
280 various authors in other environments [34–36]. The finding that CFTR genotypes a different

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

281 representation in some pathways, may suggest that the airways microbiome is influenced by the type
282 of CFTR alteration. However, this hypothesis deserves further attention to clarify the specific role of
283 microbial pathways with respect to CFTR genotype. Pathogenic bacteria, such as *Pseudomonas*
284 *aeruginosa*, need to colonize human tissues to grow and in this sense, even pathway that could be
285 related to a worsening of clinical conditions or that could be targeted by antibiotic molecules will be
286 part of this core set of functions. Despite a clear effect of antibiotic treatment during (and after)
287 exacerbation periods, the community structure is always recovered with the main pathogenic taxa
288 emerging again. This effect is confirmed by the correlation of ARG distribution and antibiotic
289 intake. Patients subjected to a given antibiotic treatment did not seem to select bacteria resistant to
290 the antibiotic used but the detection of a particular mechanism seems to be distributed in almost all
291 patients regardless of the treatment. An evidence of functional stability of the lung microbiota was
292 previously reported in other works not concerning CF disease [37,38]. Both works focused their
293 attention on the gut microbiome of obese and healthy individuals (human and mouse) reporting a
294 considerable metabolic redundancy. This high degree of redundancy in the gut microbiome supports
295 a more ecological view where subjects can be considered as different ecological niches all inhabited
296 by unique collections of microbial phylotypes but all sharing the same set of genes. This concept can
297 be extended to the lung microbiome where it is possible to define a core set of features only at the
298 level of metabolic functions. This functional conservation may thus be needed by the whole
299 community and patients can be seen as multiple micro-environments inhabited by a peculiar set of
300 strains, which share the same functions. This work represents a step forward toward a patient-
301 specific interpretation of CF microbiology [39].

302 **Conclusions**

303 In summary, the temporal dynamics of the airway microbiome in a large cohort of patients with CF
304 revealed patient-specific signatures of the airway microbiome at strain-level profiling, the lack of

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

305 variation in the microbiome across pulmonary exacerbations, and a core set of antibiotic resistance
306 genes that did not vary by antibiotic intake. The main conclusion of the present study is that the
307 management of chronic CF infection may be improved by a more patient-specific personalization of
308 clinical care and treatment. Longitudinal studies of CF airway microbiota will permit to tailor
309 therapeutic interventions and select antibiotic therapies based on the composition and relative
310 abundance of antibiotic resistance genes within the respiratory microbiome.

311 Methods

312 Demographic and clinical characteristics of enrolled patients

313 Twenty-two adolescents and adults with CF were enrolled in the study between October 2014 and
314 March 2015 (Table 1). The study subjects were selected based on eligibility criteria that included all
315 of the following: (i) a diagnosis of CF, i.e., a sweat test showing sweat Cl > 60 mmol/l and two
316 known CFTR mutations causing the disease with pancreatic insufficiency (elastase< 5 µg/g/feces)
317 [40], (ii) aged more than six years, i.e., between 11 and 55 years, (iii) chronically infected with
318 *Pseudomonas aeruginosa* according to the Leeds criteria [41] and iv) decline in %FEV_1 in the
319 previous three years before enrollment by measuring the difference between the best %FEV_1
320 registered within the previous year and the best %FEV_1 registered two-years before specimen
321 collection, following the criteria previously reported [42]. Patients were excluded if they were
322 chronically infected with *Burkholderia cepacia* complex. Using these criteria, 22 patients with
323 moderate-severe lung disease and carrying the ΔF_{508} mutation were included in the study for a total
324 of 79 shotgun metagenomic samples. The cohort was enrolled in three Italian Hospital, namely:
325 Bambino Gesù Children's Hospital (Rome, Italy), G. Gaslini Children's Hospital (Genoa, Italy) and
326 Meyer Children's Hospital (Florence, Italy). Clinical status at the time of collection was
327 designated as *baseline* (BL), when clinically stable and at their clinical and physiological baseline,

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

328 *on treatment* (TR), at exacerbation-associated antibiotic treatments, and *at recovery* (RC), upon
329 completion of antibiotic treatment. Subjects were treated according to current standards of care with
330 periodical microbiological controls [43] with at least four microbiological controls per year [4]. At
331 each visit, clinical data collection and microbiological status (colonizing pathogens with available
332 cultivation protocols) were performed according to the European CF Society standards of care [44].
333 Forced expiratory volume in 1 second as a percentage of predicted (%FEV_1) is a key outcome of
334 monitoring lung function in CF [45]. FEV_1 values were measured according to the American
335 Thoracic Society and European Respiratory Society standards [43]. CFTR genotype, sex, age, and
336 antibiotic treatment for each patient were reported in (Table 1 and supplementary materials Table
337 S1). During serial sampling, data (antibiotic usage and spirometry) were collected.

338 **Sample collection, processing, DNA extraction and sequencing**

339 Sputum samples were obtained by spontaneous expectoration at baseline, exacerbation-associated
340 antibiotic treatments and recovery status. Sampled were processed according to standard methods as
341 previously described [13,46]. Bacterial respiratory pathogens were identified using the conventional
342 techniques reported in the Guidelines, as previously described [46,47]. The number of samples,
343 microbiological status at sampling and samplings following exacerbation events are reported in
344 Table 1. Sputum samples were washed in 5 ml PBS and then centrifuged (3,800 g) for 15 minutes.
345 Resulting pellets were resuspended in 5-10 ml DNAse buffer (10 mM Tris-HCl pH 7.5; 2.5 mM
346 MgCl2; 0.5 mM CaCl2, pH 6.5) with 7.5 ul of DNAse I (2000 Units/ml) per 1 ml of sample
347 (15U/ml final), incubated for 2 hours at 37C, and washed twice by pelleting at 3,800 g for 15
348 minutes and resuspending in 10 ml SE buffer (75 mM NaCl, 25 mM EDTA, pH 7.5). Pellets were
349 then resuspended in 0.5 ml lysis buffer (20 mM Tris-HCl pH 8.0; 2 mM EDTA pH 8.0; 1% (v/v)
350 Triton; 20 mg/ml Lysozyme final concentration), incubated for 30 minutes at 37C before extracting

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

351 DNA with the MoBio Powersoil DNA extraction kit as per manufacturer's instructions. Libraries
352 were prepared with Nextera XT kit (Illumina) Sequencing was performed on an Illumina HiSeq2500
353 apparatus (Illumina). Raw sequence data reported in this study have been deposited in the NCBI
354 "Sequence Read Archive" (SRA) under the project accession PRJNA516870.

355 **Bioinformatic analyses**

356 Sequence quality was ensured by trimming reads using StreamingTrim 1.0 [48], with a quality
357 cutoff of 20. Bowtie2 [49] was used to screen out human-derived sequences from metagenomic data
358 with the latest version of the human genome available in the NCBI database (GRCh38) as reference.
359 Sequences displaying a concordant alignment (mate pair that aligns with the expected relative mate
360 orientation and with the expected range of distances between mates) against the human genomes
361 were then removed from all subsequent analyses. Metabolic and regulatory patterns were estimated
362 using HUMAnN2 [50] and considering only those pathways with a coverage value $\geq$ 80%, whereas
363 the taxonomic microbial community composition was assessed using MetaPhlAn2 [51]. Reads were
364 assembled into contigs using the metaSPAdes microbial assembler [52] with automatic k-mer length
365 selection. To establish an airway microbiome gene catalog [12] we first removed contigs smaller
366 than 500bp and then used prodigal in Anonymous mode [53], as suggested by the author of the tool,
367 to predict open reading frames (ORFs). Translated protein sequences obtained from assembled
368 contigs were classified using eggNOG mapper against the bactNOG database [54]. Each protein was
369 classified according to its best hit with an e-value lower than 0.001 as suggested in [55]. The CARD
370 database [56] was used in combination to the Resistance Gene Identifier (RGI, version 4.0.3) to
371 inspect the distribution of antibiotic resistance gene (AR genes). Genes predicted within each
372 metagenome were quantified using the number of reads that mapped against metagenomic contigs
373 obtained for each sample. Reads were mapped back to contigs using Bowtie2 [49] and the number

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

374 of reads mapping each ORF was obtained with the bedtools command "multicov" (version 2.26.0).
375 To quantify gene content across different samples, genes were collapsed using the bestOG given by
376 eggNOG mapper by summing together the number of reads that mapped genes with the same
377 annotation. The same approach was used to quantify AR genes predicted with RGI but this time the
378 unique identifier provided by CARD was used to collapse counts.
379 Strain characterization was performed using StrainPhlAn [27]. Sequence variants for each organism
380 detected were assessed against the MetaPhlAn2 [51] marker genes and a tree has been generated
381 including all samples in which the organism was found at least in one time point. One reference
382 genome per organism was downloaded from the RefSeq database and added to the tree.

383 **Taxonomic classification of metagenomic contigs**

384 Assembled contigs were taxonomically classified using BLAST. First, all genomes available for
385 each species detected with MetaPhlAn2 were downloaded from NCBI and used to build a database
386 for each sample. All genomes reporting an identity higher than 90% and a coverage higher than 80%
387 were collected and used for taxonomic classification. Contigs reporting hits with genomes coming
388 from a single species were assigned to that species whereas contigs reporting hits from multiple
389 species were flagged as unknown.

390 **Statistical analyses**

391 Statistical analyses were performed in R [57] version 3.4.4. The taxonomical and functional
392 composition on lung microbiome was explored using permutational multivariate analysis of variance
393 (PERMANOVA with 1000 permutations), 'adonis2' function of vegan package version 2.5-2;
394 whereas differences in bacterial diversity were tested using analysis of covariance (ANCOVA),
395 'aov' function. The model fitted for both analyses was:

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

396 $$X \sim Status + Genotype + Subject + FEV_1 + days$$

397 where, Exacerbation is the exacerbation event, Genotype is the CFTR genotype, Subject is the
398 patient, FEV_1 was the forced expiatory volume in 1 second, and days, was the number of days from
399 the enrollment in the study. For the ANCOVA analyses Tukey's post hoc tests were performed to
400 test for mean differences within each factor used to build the full model (excluding FEV_1 value and
401 days since they were not categorical variable). Ordination analyses were conducted on both taxa and
402 pathways using the function 'ordinate' of the phyloseq package (version 1.23.1) with principle
403 coordinate decomposition method (PCoA) and the Bray-Curtis dissimilarity index. The same index
404 was used to inspect the distribution of samples and compare beta diversity level in bot taxonomic
405 composition and pathways.

406 To test for differentially distributed pathways and taxa across exacerbation events and genotypes we
407 used a moderated t-test as implemented in the limma package [58], version 3.34.9. Data obtained
408 with MetaPhlAn2 (taxonomic composition) and HUMAnN2 (pathway composition) were fitted into
409 limma's model using subjects as blocking variable. Since both software quantify biological units
410 using relative counts (HUMAnN2 uses "copies per million" and MetaPhlAn2 uses percentages) we
411 transformed this data into logarithmic values using the formula: $\log_2(x + 0.1)$, where x are the
412 relative counts. Obtained p-values were corrected using the Benjamini-Hochberg correction method.
413 A similar approach has been used for antibiotic genes detect along assembled contigs. Here the
414 number of reads that mapped onto each gene was used to estimate differentially abundant gene.
415 Since the number of reads for each sample was variable (the ratio of the largest library size to the
416 smallest was more than 10-fold) we used limma's voom method [59] to fit our model, as suggested
417 by the author of limma.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

418 # Declarations

419 **Ethics Statement**

420 The study was approved by the Ethics Committees of Children's Hospital and Research Institute
421 Bambino Gesù (Rome, Italy), Cystic Fibrosis Center, Anna Meyer Children's University Hospital
422 (Florence, Italy) and G. Gaslini Institute (University of Genoa, Genoa, Italy) [Prot. N. 681 CM of
423 November 2, 2012; Prot. N. 85 of February 27, 2014; Prot. N. FCC 2012 Partner 4-IGG of
424 September 18, 2012]. All participants provided written informed consent before the enrollment in
425 the study. All sputum specimens were produced voluntarily. All procedures were performed in
426 agreement with the "Guidelines of the European Convention on Human Rights and Biomedicine for
427 Research in Children" and the Ethics Committee of the three CF Centers involved. All measures
428 were obtained and processed ensuring patient data protection and confidentiality.

429 **Availability of data and material**

430 All data generated or analysed during this study are included in this article. Raw sequence data
431 reported in this study have been deposited in the NCBI "Sequence Read Archive" (SRA) under the
432 project accession PRJNA516870.

433 **Conflict of Interest**

434 We have no conflict of interest to declare.

435 **Funding**

436 This study was supported by Italian Grants funded by the Italian Cystic Fibrosis Research
437 Foundation (FFC) (http://www.fibrosicisticaricerca.it/) to Annamaria Bevivino: Project n.
438 FFC#14/2015 with the contribution of "Delegazione FFC di Latina", "Latteria Montello Nonno
439 Nanni", and "Gruppo di Sostegno FFC Valle Scrivia Alessandria"), and Project n. FFC#19/2017
440 with the contribution of "Delegazione FFC Lago di Garda con i Gruppi di Sostegno FFC di

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

441 Chivasso, dell'Isola Bergamasca, di Arezzo". Giovanni Bacci was supported by a postdoctoral
442 fellowship by Italian Cystic Fibrosis Research Foundation. The funders had no role in study design,
443 data collection and analysis, decision to publish, or preparation of the manuscript.

444 **Authors' contributions**

445 Conceived and designed the experiments: AB VL GT EVF AM. Performed the experiments: GB
446 FDC. Analyzed the data: GB AM AB. Contributed reagents/materials/analysis tools: DD FA PM RS
447 AN. Wrote the paper: GB AM AB. Provided comments and recommendations that improved the
448 manuscript: NS GT VL. Supervised research: AB VL GT EVF AM.

449 **Acknowledgments**

450 The authors greatly acknowledge the Italian Cystic Fibrosis Research Foundation (FCC) for its
451 support and administrative tasks, and Ricciotti Gabriella, Tuccio Vanessa and Campana Silvia for
452 their technical support.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

453 # Tables

454 **Table 1. Characteristics of patients enrolled in the study.** ID, study id; Hospital, hospital in
455 which patient has been enrolled [OPBG=Children's Hospital and Research Institute Bambino Gesù
456 (Rome, Italy); Gaslini=G. Gaslini Institute (University of Genoa, Genoa, Italy); Meyer=Cystic
457 Fibrosis Center, Anna Meyer Children's University Hospital (Florence, Italy)]; Genotype, CFTR
458 genotype; Gender, gender; Age, enrollment's age; n, number of samples collected; EX, yes if an
459 exacerbation event has occurred during the study (no otherwise) [4]; FEV1, mean value of forced
460 expiratory volume in 1 second plus/minus the standard error on the mean; heterozygote and
461 homozygote refers to ΔF508 genotype; %FEV1 status: S = with a rate decline lower than 1.5%, SD
462 = with a rate decline higher than 5% [42].

ID	Hospital	Genotype	Gender	FEV_1 status	Age	n	EX	$\%FEV_1$
B01	OPBG	ΔF_{508}/2183AA->G	M	S	27	5	yes	37.0 ± 1.70
B02	OPBG	ΔF_{508}/N1303K	F	SD	26	3	no	54.7 ± 3.48
B03	OPBG	ΔF_{508}/4016insT	F	S	30	4	no	55.0 ± 1.08
B06	OPBG	ΔF_{508}/ΔF_{508}	F	SD	21	4	no	60.2 ± 3.42
G10	Gaslini	ΔF_{508}/ΔF_{508}	M	S	51	4	no	54.0 ± 3.08
G24	Gaslini	ΔF_{508}/ΔF_{508}	F	S	49	3	yes	NA ± NA
G28	Gaslini	ΔF_{508}/ΔF_{508}	F	NA	38	2	no	42.5 ± 1.50
G30	Gaslini	ΔF_{508}/ΔF_{508}	F	S	50	1	no	54
G31	Gaslini	G1244E/G42X	F	SD	53	2	no	41.5 ± 1.50
G34	Gaslini	ΔF_{508}/ΔF_{508}	F	S	39	1	no	47
M05	Meyer	ΔF_{508}/ΔF_{508}	M	SD	32	4	no	34.8 ± 0.85
M19	Meyer	ΔF_{508}/ΔF_{508}	M	S	24	4	no	44.0 ± 2.04
M21	Meyer	ΔF_{508}/N1303K	M	SD	27	4	yes	51.5 ± 4.35
M22	Meyer	ΔF_{508}/2789+5G->A	F	S	29	5	yes	50.4 ± 1.03
M23	Meyer	ΔF_{508}/G542X	F	S	30	4	yes	37.0 ± 1.47
M24	Meyer	ΔF_{508}/ΔF_{508}	M	S	32	4	no	35.2 ± 0.85
M25	Meyer	ΔF_{508}/296+1G->T	F	SD	41	4	no	42.5 ± 2.02
M26	Meyer	ΔF_{508}/3849+10	F	SD	49	5	yes	39.6 ± 1.94
M28	Meyer	ΔF_{508}/N1303K	M	S	23	4	no	39.0 ± 1.08
M29	Meyer	ΔF_{508}/G542X	F	S	12	4	no	43.5 ± 3.75
M31	Meyer	ΔF_{508}/ΔF_{508}	F	SD	11	3	yes	32.7 ± 4.41
M33	Meyer	ΔF_{508}/G85E	F	SD	13	5	yes	35.4 ± 5.78
Total: 22	Gaslini:6 Meyer:12 OPBG:4	Heterozygote :47 Homozygote :29 Other:2	F:15 M:7	S:12 SD:9	32.1 ± 2.73	79	no:14 yes:8	43.5 ± 1.09

463

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

464 **Table 2.** Permutational multivariate analysis of variance (PERMANOVA) on both taxonomic
465 distribution and metabolic pathways. The analysis based on taxonomic distribution was reported in
466 the upper part of the table whereas the analysis based on metabolic pathways was reported at the
467 bottom. Df, degrees of freedom; SumOfSqs, sum of squares; R^2, r-squared statistic (reported as
468 proportion); F, F-statistic; Pr(>F), p-value associated to the F-statistic. Significant effects, namely
469 those reporting a p-value lower than 0.05, were reported in bold.

	Df	**SumOfSqs**	**R^2**	**F**	**Pr(>F)**
Taxonomy					
Status	**2**	**0.68**	**0.03**	**1.91**	**0.0300**
Genotype	**1**	**0.77**	**0.03**	**4.30**	**0.0020**
Sample	**18**	**11.97**	**0.52**	**3.74**	**0.0010**
FEV_1 value	1	0.27	0.01	1.53	0.1349
Days	1	0.28	0.01	1.58	0.1229
Status:Genotype	1	0.11	0.01	0.64	0.7642
Residual	49	8.72	0.38	-	-
Pathway					
Status	**2**	**0.20**	**0.04**	**2.37**	**0.0220**
Genotype	**1**	**0.14**	**0.03**	**3.42**	**0.0080**
Sample	**18**	**2.43**	**0.48**	**3.20**	**0.0010**
FEV_1 value	1	0.09	0.02	2.14	0.0989
Days	1	0.05	0.01	1.26	0.2458
Status:Genotype	1	0.08	0.02	1.96	0.1169
Residual	49	2.07	0.41	-	-

470

471

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

472 # References

473 1. Ratjen F, Döring G. Cystic fibrosis. Lancet. 2003;361:681–9.

474 2. Gibson RL, Burns JL, Ramsey BW. Pathophysiology and management of pulmonary infections in
475 cystic fibrosis. Am J Respir Crit Care Med. 2003;168:918–51.

476 3. Knapp EA, Fink AK, Goss CH, Sewall A, Ostrenga J, Dowd C, et al. The cystic fibrosis
477 Foundation patient registry. Design and methods of a national observational disease registry. Ann
478 Am Thorac Soc. Am Thoracic Soc; 2016;13:1173–9.

479 4. Flume PA, Mogayzel PJ, Robinson KA, Goss CH, Rosenblatt RL, Kuhn RJ, et al. Cystic fibrosis
480 pulmonary guidelines: Treatment of pulmonary exacerbations. Am J Respir Crit Care Med.
481 2009;180:802–8.

482 5. Huang YJ, LiPuma JJ. The microbiome in cystic fibrosis. Clin Chest Med. 2016;37:59–67.

483 6. Fodor AA, Klem ER, Gilpin DF, Elborn JS, Boucher RC, Tunney MM, et al. The adult cystic
484 fibrosis airway microbiota is stable over time and infection type, and highly resilient to antibiotic
485 treatment of exacerbations. PLoS One. 2012;7:e45001.

486 7. Price KE, Hampton TH, Gifford AH, Dolben EL, Hogan DA, Morrison HG, et al. Unique
487 microbial communities persist in individual cystic fibrosis patients throughout a clinical
488 exacerbation. Microbiome. 2013;1:27.

489 8. Cox MJ, Turek EM, Hennessy C, Mirza GK, James PL, Coleman M, et al. Longitudinal
490 assessment of sputum microbiome by sequencing of the 16S rRNA gene in non-cystic fibrosis
491 bronchiectasis patients. PLoS One. 2017;12:1–17.

492 9. Whelan FJ, Heirali AA, Rossi L, Rabin HR, Parkins MD, Surette MG. Longitudinal sampling of
493 the lung microbiota in individuals with cystic fibrosis. PLoS One. 2017;12:e0172811.

494 10. Carmody LA, Zhao J, Schloss PD, Petrosino JF, Murray S, Young VB, et al. Changes in cystic
495 fibrosis airway microbiota at pulmonary exacerbation. Ann Am Thorac Soc. 2013;10:179–87.

496 11. Carmody LA, Zhao J, Kalikin LM, LeBar W, Simon RH, Venkataraman A, et al. The daily
497 dynamics of cystic fibrosis airway microbiota during clinical stability and at exacerbation.
498 Microbiome. 2015;3:12.

499 12. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial
500 gene catalogue established by metagenomic sequencing. Nature. 2010;464:59–65.

501 13. Bacci G, Mengoni A, Fiscarelli E, Segata N, Taccetti G, Dolce D, et al. A different microbiome
502 gene repertoire in the airways of cystic fibrosis patients with severe lung disease. Int J Mol Sci.
503 2017;18:1654.

504 14. Franzosa EA, Morgan XC, Segata N, Waldron L, Reyes J, Earl AM, et al. Relating the
505 metatranscriptome and metagenome of the human gut. Proc Natl Acad Sci. 2014;111:E2329–38.

506 15. Lim YW, Schmieder R, Haynes M, Willner D, Furlan M, Youle M, et al. Metagenomics and
507 metatranscriptomics: Windows on CF-associated viral and microbial communities. J Cyst Fibros.
508 2013;12:154–64.

509 16. Segata N. On the Road to Strain-Resolved Comparative Metagenomics. mSystems.
510 2018;3:e00190-17.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

511 17. Beghini F, Pasolli E, Truong TD, Putignani L, Cacciò SM, Segata N. Large-scale comparative
512 metagenomics of *Blastocystis*, a common member of the human gut microbiome. ISME J.
513 2017;11:2848–63.

514 18. Zolfo M, Tett A, Jousson O, Donati C, Segata N. MetaMLST: Multi-locus strain-level bacterial
515 typing from metagenomic samples. Nucleic Acids Res. 2017;45:e7.

516 19. Zhao J, Schloss PD, Kalikin LM, Carmody LA, Foster BK, Petrosino JF, et al. Decade-long
517 bacterial community dynamics in cystic fibrosis airways. Proc Natl Acad Sci. 2012;109:5809–14.

518 20. Delhaes L, Monchy S, Fréalle E, Hubans C, Salleron J, Leroy S, et al. The airway microbiota in
519 cystic fibrosis: A complex fungal and bacterial community-implications for therapeutic
520 management. PLoS One. 2012;7:e36313.

521 21. Willner D, Daly J, Whiley D, Grimwood K, Wainwright CE, Hugenholtz P. Comparison of
522 DNA extraction methods for microbial community profiling with an application to pediatric
523 bronchoalveolar lavage samples. PLoS One. 2012;7:e34605.

524 22. Lim YW, Evangelista JS, Schmieder R, Bailey B, Haynes M, Furlan M, et al. Clinical insights
525 from metagenomic analysis of sputum samples from patients with cystic fibrosis. J Clin Microbiol.
526 2014;52:425–37.

527 23. Willner D, Haynes MR, Furlan M, Hanson N, Kirby B, Lim YW, et al. Case studies of the
528 spatial heterogeneity of DNA viruses in the cystic fibrosis lung. Am J Respir Cell Mol Biol.
529 2012;46:127–31.

530 24. Whiteson KL, Meinardi S, Lim YW, Schmieder R, Maughan H, Quinn R, et al. Breath gas
531 metabolites and bacterial metagenomes from cystic fibrosis airways indicate active pH neutral 2,3-
532 butanedione fermentation. ISME J. 2014;8:1247–58.

533 25. Surette MG. The cystic fibrosis lung microbiome. Ann Am Thorac Soc. 2014;11:61–5.

534 26. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from
535 sampling to analysis. Nat Biotechnol. 2017;35:833–44.

536 27. Truong DT, Tett A, Pasolli E, Huttenhower C, Segata N. Microbial strain-level population
537 structure and genetic diversity from metagenomes. Genome Res. 2017;626–38.

538 28. Scholz M, Ward D V., Pasolli E, Tolio T, Zolfo M, Asnicar F, et al. Strain-level microbial
539 epidemiology and population genomics from shotgun metagenomics. Nat Methods. 2016;13:435–8.

540 29. Faust K, Lahti L, Gonze D, de Vos WM, Raes J. Metagenomics meets time series analysis:
541 unraveling microbial community dynamics. Curr Opin Microbiol. 2015;25:56–66.

542 30. Shade A, Peter H, Allison SD, Baho DL, Berga M, Bürgmann H, et al. Fundamentals of
543 microbial community resistance and resilience. Front Microbiol. 2012;3:417.

544 31. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and
545 resilience of the human gut microbiota. Nature. 2012;489:220–30.

546 32. Zemanick ET, Wagner BD, Robertson CE, Ahrens RC, Chmiel JF, Clancy JP, et al. Airway
547 microbiota across age and disease spectrum in cystic fibrosis. Eur Respir J. 2017; 50(5). pii:
548 1700832

549 33. Carmody LA, Caverly LJ, Foster BK, Rogers MAM, Kalikin LM, Simon RH, et al. Fluctuations
550 in airway bacterial communities associated with clinical states and disease stages in cystic fibrosis.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

551 PLoS One. 2018; 13(3): e0194060.

552 34. Doolittle WF, Booth A. It's the song, not the singer: an exploration of holobiosis and
553 evolutionary theory. Biol Philos. 2016;32:5–24.

554 35. Eng A, Borenstein E. Taxa-function robustness in microbial communities. Microbiome.
555 2018;6:45.

556 36. Louca S, Jacques SMS, Pires APF, Leal JS, Srivastava DS, Parfrey LW, et al. High taxonomic
557 variability despite stable functional structure across microbial communities. Nat Ecol Evol.
558 2016;1(1):15.

559 37. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-
560 associated gut microbiome with increased capacity for energy harvest. Nature. 2006;444:1027–31.

561 38. Thaiss CA, Itav S, Rothschild D, Meijer M, Levy M, Moresi C, et al. Persistent microbiome
562 alterations modulate the rate of post-dieting weight regain. Nature. 2016;540:1–27.

563 39. Bevivino A, Bacci G, Drevinek P, Nelson MT, Hoffman LR, Mengoni A. Deciphering the
564 ecology of cystic fibrosis bacterial communities: Towards systems-level integration. Trends Mol
565 Med. 2019; in press. DOI:10.1016/j.molmed.2019.07.008

566 40. Farrell PM, White TB, Ren CL, Hempstead SE, Accurso F, Derichs N, et al. Diagnosis of cystic
567 fibrosis: Consensus guidelines from the cystic fibrosis foundation. J Pediatr. 2017;181:S4–15.

568 41. Proesmans M, Balinska-Miskiewicz W, Dupont L, Bossuyt X, Verhaegen J, Høiby N, et al.
569 Evaluating the "Leeds criteria" for *Pseudomonas aeruginosa* infection in a cystic fibrosis centre.
570 Eur Respir J. 2006;27:937–43.

571 42. Bacci G, Paganin P, Lopez L, Vanni C, Dalmastri C, Cantale C, et al. Pyrosequencing unveils
572 cystic fibrosis lung microbiome differences associated with a severe lung function decline. PLoS
573 One. 2016;11:e0156807.

574 43. Kerem E, Conway S, Elborn S, Heijerman H, Committee C, others. Standards of care for
575 patients with cystic fibrosis: a European consensus. J Cyst Fibros. 2005;4:7–26.

576 44. Smyth AR, Bell SC, Bojcin S, Bryon M, Duff A, Flume P, et al. European cystic fibrosis society
577 standards of care: Best practice guidelines. J Cyst Fibros. 2014;13:S30–53.

578 45. Taylor-Robinson D, Whitehead M, Diderichsen F, Olesen H V, Pressler T, Smyth RL, et al.
579 Understanding the natural progression in %FEV1 decline in patients with cystic fibrosis: A
580 longitudinal study. Thorax. 2012;67:860–6.

581 46. Paganin P, Fiscarelli EV, Tuccio V, Chiancianesi M, Bacci G, Morelli P, et al. Changes in cystic
582 fibrosis airway microbial community associated with a severe decline in lung function. PLoS One.
583 2015;10:e0124348.

584 47. Farrell PM, Rosenstein BJ, White TB, Accurso FJ, Castellani C, Cutting GR, et al. Guidelines
585 for diagnosis of cystic fibrosis in newborns through older adults: Cystic fibrosis Foundation
586 consensus report. J Pediatr. 2008;153:S4–14.

587 48. Bacci G, Bazzicalupo M, Benedetti A, Mengoni A. StreamingTrim 1.0: A Java software for
588 dynamic trimming of 16S rRNA sequence data from metagenetic studies. Mol Ecol Resour.
589 2014;14:426–34.

590 49. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

591 2012;9:357–9.

592 50. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, et al. Metabolic
593 reconstruction for metagenomic data and its application to the human microbiome. PLoS Comput
594 Biol. 2012;8:e1002358.

595 51. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlAn2 for
596 enhanced metagenomic taxonomic profiling. Nat Methods. 2015;12:902–3.

597 52. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: A new versatile metagenomic
598 assembler. Genome Res. 2017;27:824–34.

599 53. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: Prokaryotic gene
600 recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.

601 54. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, et al. eggNOG 4.5: A
602 hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic
603 and viral sequences. Nucleic Acids Res. 2016;44:D286–93.

604 55. Pearson WR. An introduction to sequence similarity (“homology”) searching. Curr Protoc
605 Bioinformatics. 2013; Chapter 3 Unit 3.1: p. 3.1.1-3.1.8.

606 56. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, et al. CARD 2017:
607 Expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic
608 Acids Res. 2017;45:D566–73.

609 57. R Development Core Team. R: A language and environment for statistical computing. R
610 foundation for statistical computing, Vienna, Austria. http://www.R-project.org/. R Found Stat
611 Comput Vienna, Austria. 2012.

612 58. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. *limma* powers differential
613 expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43:e47.

614 59. Law CW, Chen Y, Shi W, Smith GK. voom: precision weights unlock linear model analysis
615 tools for RNA-seq read counts. Genome Biol. 2014;15(2):R29.

616

a)

BL TR RC

Firmicutes

Proteobacteria

Bacteroidetes

Actinobacteria

Fusobacteria

Spirochaetes

Ascomycota

Species

Sample

Abundance (%)

10^2

10^0

10^{-2}

10^{-4}

b)

BL TR RC

Staphylococcus aureus
Streptococcus parasanguinis
Veillonella unclassified
Enterococcus faecalis
Gemella sanguinis
Granulicatella unclassified
Veillonella parvula
Gemella haemolysans
Granulicatella adiacens
Streptococcus mitis oralis pneumoniae

Firmicutes

Pseudomonas aeruginosa
Pseudomonas unclassified
Haemophilus parainfluenzae
Neisseria unclassified
Neisseria flavescens
Haemophilus influenzae
Lautropia mirabilis
Achromobacter xylosoxidans
Rhodopseudomonas palustris
Stenotrophomonas maltophilia

Proteobacteria

Porphyromonas sp oral taxon 279
Prevotella melaninogenica
Prevotella histicola
Prevotella pleuritidis
Prevotella nanceiensis
Capnocytophaga unclassified
Capnocytophaga gingivalis
Capnocytophaga sp oral taxon 329
Prevotella sp C561
Prevotella pallens

Bacteroidetes

Rothia mucilaginosa
Rothia dentocariosa
Actinomyces graevenitzii
Rothia aeria
Bifidobacterium longum
Atopobium parvulum
Actinomyces odontolyticus
Actinomyces sp ICM39
Propionibacterium acnes
Rothia unclassified

Actinobacteria

Fusobacterium nucleatum
Fusobacterium periodonticum
Leptotrichia unclassified
Leptotrichia sp oral taxon 215
Leptotrichia goodfellowii
Leptotrichia wadei
Leptotrichia shahii

Fusobacteria

Treponema vincentii
Treponema maltophilum
Treponema socranskii
Treponema medium

Spirochaetes

Candida albicans
Candida unclassified

Ascomycota

10^{-2} 10^0 10^2 10^{-2} 10^0 10^2 10^{-2} 10^0 10^2

Abundance (%)

a)

PCoA 2 (17.2%)

0.75
0.50
0.25
0.00
−0.25

−0.6 −0.4 −0.2 0.0 0.2

PCoA 1 (26.5%)

b)

PCoA 2 (24.9%)

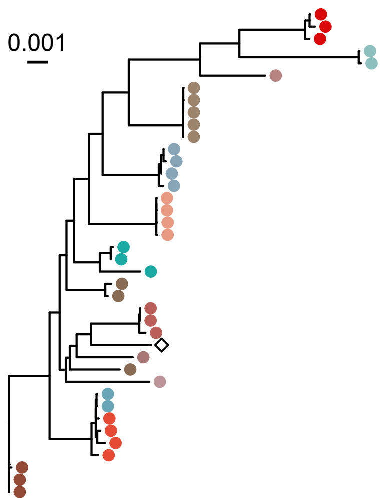
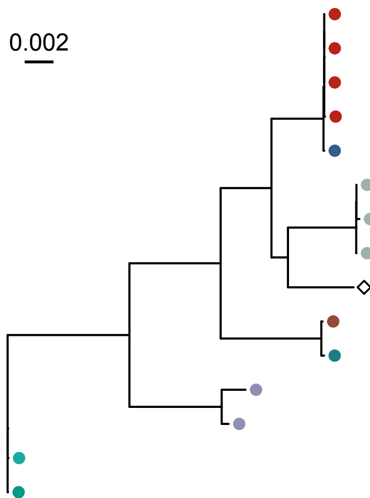
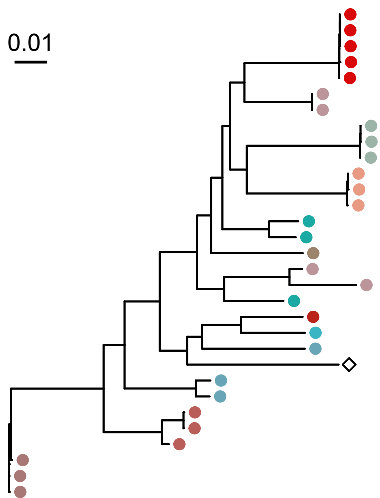
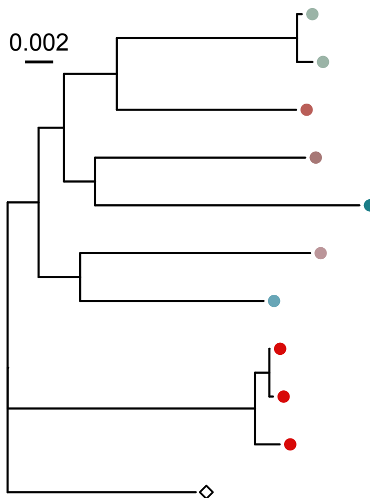
0.4
0.2
0.0

−0.4 −0.2 0.0

PCoA 1 (48.5%)

Genotype heterozygote homozygote other

Status TR BL RC


a
0.001

b
0.002

c
0.01

d
0.002

B01
B02
B03
B06
G10
G24
G28
G30
G31
G34
M05
M19
M21
M22
M23
M24
M25
M26
M28
M29
M31
M33
M34
reference

a

*

Shannon

3
2
1
0

*

InvSimpson

25
20
15
10
5
0

BL RC TR

b

Shannon

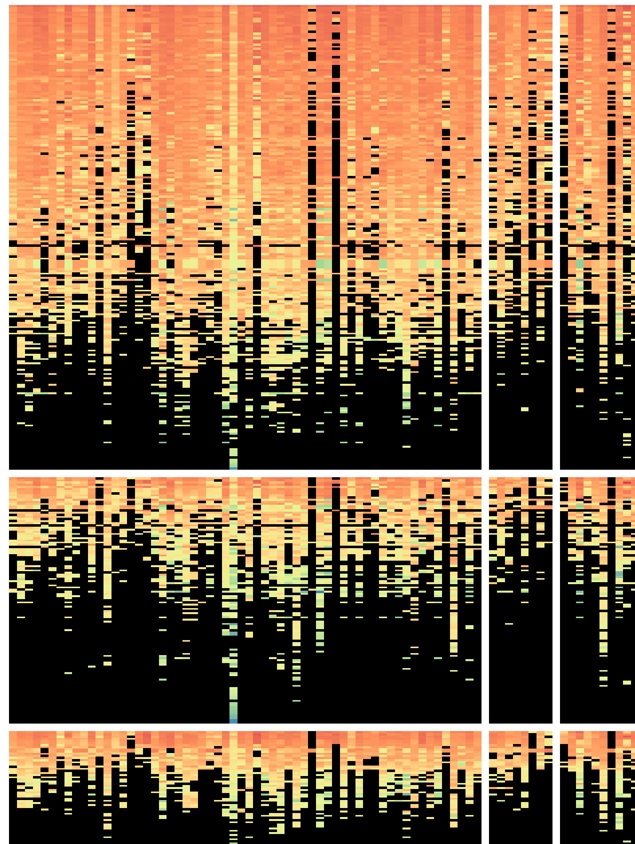
5.0
4.5
4.0

*

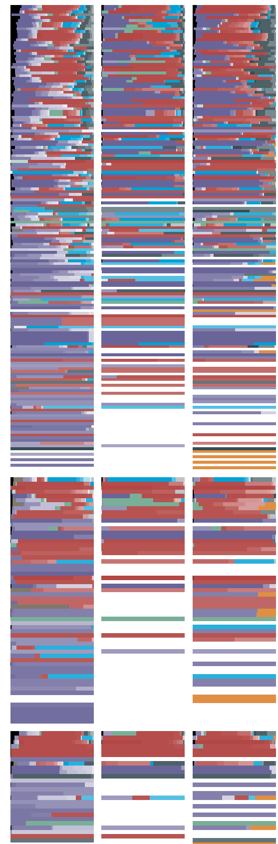
InvSimpson

150
100
50

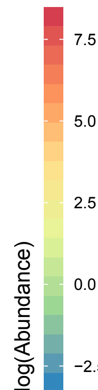
BL RC TR



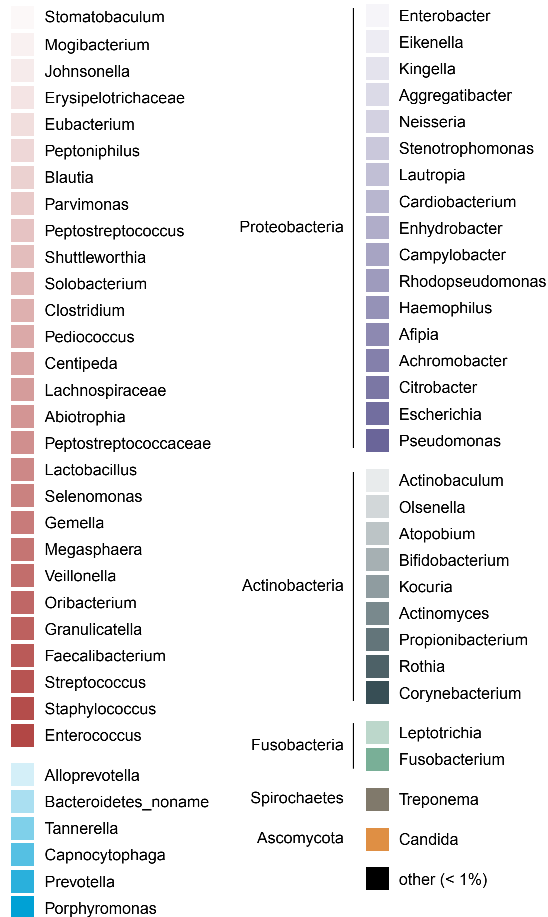
BL
TR
RC
Patways
Samples



BL
TR
RC
Biosynthesis
Degradation−Utilization−Assimilation
Generation of Precursor Metabolites and Energy
Abundance (%)



log(Abundance)
7.5
5.0
2.5
0.0
−2.5



Firmicutes
Stomatobaculum
Mogibacterium
Johnsonella
Erysipelotrichaceae
Eubacterium
Peptoniphilus
Blautia
Parvimonas
Peptostreptococcus
Shuttleworthia
Solobacterium
Clostridium
Pediococcus
Centipeda
Lachnospiraceae
Abiotrophia
Peptostreptococcaceae
Lactobacillus
Selenomonas
Gemella
Megasphaera
Veillonella
Oribacterium
Granulicatella
Faecalibacterium
Streptococcus
Staphylococcus
Enterococcus
Bacteroidetes
Alloprevotella
Bacteroidetes_noname
Tannerella
Capnocytophaga
Prevotella
Porphyromonas
Proteobacteria
Enterobacter
Eikenella
Kingella
Aggregatibacter
Neisseria
Stenotrophomonas
Lautropia
Cardiobacterium
Enhydrobacter
Campylobacter
Rhodopseudomonas
Haemophilus
Afipia
Achromobacter
Citrobacter
Escherichia
Pseudomonas
Actinobacteria
Actinobaculum
Olsenella
Atopobium
Bifidobacterium
Kocuria
Actinomyces
Propionibacterium
Rothia
Corynebacterium
Fusobacteria
Leptotrichia
Fusobacterium
Spirochaetes
Treponema
Ascomycota
Candida
other (< 1%)

a)

0.00
0.25
0.50
0.75

Pathways

Taxonomy

Genotype □ heterozygote ○ homozygote △ other

Status ● Baseline ● Treatment ● Recovery

b)

Exacerbation

Genotype

Bray−Curtis dissimilarity

1.00
0.75
0.50
0.25
0.00

TR BL RC

heterozygote homozygote

paths taxa