bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder.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 **Title**

2 Taxonomic variability over functional stability in the microbiome of 3 Cystic Fibrosis patients chronically infected by *Pseudomonas* 4 *aeruginosa*

5

6 Giovanni Bacci[1], Giovanni Taccetti[2], Daniela Dolce[2], Federica Armanini[3], Nicola Segata[3], 7 Francesca Di Cesare[1], Vincenzina Lucidi[4], Ersilia Fiscarelli[4], Patrizia Morelli[5], Rosaria Casciaro[5], 8 Anna Negroni[6], Alessio Mengoni[1], and Annamaria Bevivino[6*]

9

10 [1]Department of Biology, University of Florence, Sesto Fiorentino, 50019, Florence, Italy,

11 [2]Cystic Fibrosis Center, Anna Meyer Children's University Hospital, Department of Pediatrics 12 Medicine, Florence,

13 50139, Italy

14 [3]Centre for Integrative Biology, University of Trento, Trento, 38122, Italy

15 [4]Children's Hospital and Research Institute Bambino Gesù, Rome, 00165, Italy

16 [5]Cystic Fibrosis Center, IRCCS G. Gaslini Institute, Department of Pediatrics, Genoa, 16147, Italy

17

18 [6]Department for Sustainability, Italian National Agency for New Technologies, Energy and 19 Sustainable Economic Development, ENEA, 00123 Rome, Italy

20

21 * Corresponding author

22 Annamaria Bevivino, Department for Sustainability, ENEA, Via Anguillarese 301 – 00123 Rome, 23 Italy. Email: annamaria.bevivino@enea.it

24

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder.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.

Abstract

Although the cystic fibrosis (CF) lung microbiome has been characterized in several studies, little is still known about the functions harboured by those bacteria, and how they change with disease status and antibiotic treatment. The aim of this study was to investigate the taxonomic and functional temporal dynamics of airways microbiome in a cohort of CF patients. Multiple sputum samples were collected over 15 months from 22 patients with chronic *P. aeruginosa* infection, for a total of 79 samples. DNA extracted from samples was subjected to shotgun metagenomic sequencing allowing either strain-level taxonomic profiling and assessment of the functional metagenomic repertoire. High inter-patient taxonomic heterogeneity was found with short-term compositional changes during exacerbations and following antibiotic treatment. Each patient exhibited distinct sputum microbial communities at the taxonomic level, and strain-specific colonization of traditional CF pathogens, including *P. aeruginosa*, and emerging pathogens. Sputum microbiome was found to be extraordinarily resilient following antibiotic treatment, with rapid recovery of taxa and metagenome-associated gene functions. In particular, a large core set of genes, including antibiotic resistance genes, were shared across patients despite observed differences in clinical status or antibiotic treatment, and constantly detected in the lung microbiome of all subjects independently from known antibiotic exposure, suggesting an overall microbiome-associated functions stability despite taxonomic fluctuations of the communities.

IMPORTANCE While the dynamics of CF sputum microbial composition were highly patient-specific, the overall sputum metagenome composition was stable, showing a high resilience along time and antibiotic exposure. The high degree of redundancy in the CF lung microbiome could testifies ecological aspects connected to the disease that were never considered so far, as the large core-set of genes shared between patients despite observed differences in clinical status or antibiotic

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

48 treatment. Investigations on the actual functionality (e.g. by metatranscriptomics) of the identified
49 core-set of genes could provide clues on genetic function of the microbiome to be targeted in future
50 therapeutic treatments.
51

52 **Key words:** cystic fibrosis; lung microbiome; longitudinal studies; metagenome composition;
53 antibiotic resistance genes
54

55 Introduction

56 Bacterial lung infections reduce life expectancy in most individuals with cystic fibrosis (CF) (1).
57 Sputum bacterial loads remain equally high both during periods of clinical stability and during
58 pulmonary exacerbations (2), the latter of which contribute to the irreversible decline of lung
59 function. Though much is known about microbes that cause respiratory infections in CF (3), how
60 microbes contribute to exacerbations is still poorly understood. In the past years, studies employing
61 DNA-based analyses of the airway microbiota of CF patients have reported somewhat discordant
62 results. Indeed, while some studies showed a largely stable airway microbiota during clinical change
63 and antibiotic treatment (4), other did not, suggesting changes the involvement of some microbial
64 taxa in exacerbation (5, 6). Most of these used 16S rRNA gene sequencing, yielding the identities
65 and relative abundances of the taxa present (i.e., the microbiota), but without providing any strain-
66 level or functional (meaning based on functional genes) information (i.e., the metagenome) (7).
67 These latter characteristics are particularly relevant for studying host-microbiome interactions.
68 Indeed, defining the dynamics of individual microbial strains provides important information
69 regarding how specific sub-lineages of pathogens persist and relate to clinical change. On the other
70 hand, studying the microbial genetic repertoire, e.g. antibiotic resistance and virulence-related

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

71 genes, with respect to clinical status or treatment can identify mechanisms of microbial persistence
72 and pathogenesis (8–10). Until now, few longitudinal studies, with a limited number of patients,
73 focusing only on CF airway microbiota have been performed (11–13). Moreover, studies on the
74 complete CF microbiome (microbiota and metagenome) are few and on a limited number of patients
75 (13–16) or focused on specific metabolic functions (17). In this work, we studied the temporal
76 dynamics of CF sputum microbiomes, focusing on patients with moderate-severe lung disease,
77 chronically infected by *Pseudomonas aeruginosa*. In a previous work (18), chronic infection with *P.*
78 *aeruginosa* has been found to be associated with dysbiosis in the lungs of patients with CF. The
79 authors suggested that the dominance of one species remodels the lung microbiota and may promote
80 severity of CF lung disease. A more detailed taxonomic and functional analysis could help
81 elucidating the mechanisms leading to chronic infection with *P. aeruginosa* and the microbial
82 factors that contribute to the global changes of their lung microbiome. In the present study, a
83 shotgun metagenomic approach was used (19) to detect the entire sputum microbial genomic
84 repertoire down to the strain level (20, 21). A cohort of 22 patients with moderate-severe lung
85 disease, grouped according to homozygosity versus heterozygosity for ΔF508 (also known as
86 F508del) in the CFTR gene and chronically infected with *P. aeruginosa*, was selected and followed
87 over 15 months during which 8 patients underwent exacerbation events. We aimed to determine the
88 composition of sputum microbiomes for these patients when longitudinally sampled during periods
89 of stability and exacerbation, defining the relationship between clinical status, sputum microbial
90 metabolic gene repertoire, and the antibiotic-resistance (AR) gene composition of sputum bacterial
91 community, providing a previously unknown, high-resolution view of CF sputum microbiome
92 dynamics.
93

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

94 # Results

95 ## Patients and sampling

96 Twenty-two patients with CF were enrolled (15 females and seven males) who had moderate-severe
97 lung disease (30 < %FEV_1 < 70) and were chronically infected by *P. aeruginosa*, according to the
98 Leeds criteria (22). During the study period, patients were treated with maintenance antibiotics
99 (aerosol) and only a subset (n = 8) received clinical additional antibiotics (oral or/and intravenous)
100 for a pulmonary exacerbation (CFPE) (Table 1 and supplementary materials Table S1). Among the
101 22 subjects, 8 were diagnosed with exacerbations during the study period. In total, 79 sputum
102 samples were collected and analyzed using shotgun metagenomic sequencing.
103

104 ## Airway microbiomes are taxonomically distinct and show patient-specific strain colonization

105 The overall taxonomic representation of the microbiomes from the 79 samples is reported in Fig. 1a
106 and 1b, whereas a summary of obtained reads per sample was reported in supplementary materials
107 Table S2. Firmicutes, Proteobacteria, Bacteroidetes, and Actinobacteria were the most represented
108 phyla. A high relative abundance of the "classical" CF bacterial signatures (taxa), such as
109 *Staphylococcus aureus* and *Pseudomonas aeruginosa*, and non-traditional CF taxa, such as *Rothia*
110 *mucilaginosa*, and *Prevotella melaninogenica* (all present in the top-10 species within each phylum,
111 Fig. 1b), was found. These species, indeed, represent the 49% of all detected taxa as reported in
112 supplementary materials Table S3.
113 Although principal coordinates analysis showed that subjects did not cluster based on treatment
114 events and/or genotype (Fig. 2a), the PERMANOVA analysis (Table 2) reported a significant effect
115 of both factors. The R^2 values, namely the proportion of variance explained by the factor considered,
116 were very low (Table 2, R^2 = 0.03 for both factors, p-values < 0.05) probably due to intra-patient

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

117 heterogeneity. CFTR genotype did not influence the effect of antibiotic treatment (viz. exacerbation)
118 on sputum microbiota, nor vice versa (p-value > 0.05, treatment-genotype interaction effect).
119 Subject effect was predominant with an R^2 value of 0.52. A high fraction (more than 50%) of the
120 total variance can be thus explained by inter-subject variation. Neither FEV_1 nor time showed any
121 significant relationship with taxonomy or functional profile (Table 2).
122 We then performed a strain-level analysis of the sputum microbiomes. This analysis demonstrated,
123 in samples from the same patient but at different time points, that bacterial lineages were in general,
124 closely related and tightly clustered together, confirming a patient-specific bacterial colonization
125 (Fig. 3 and supplementary materials Fig. S1).
126 Alpha diversity analyses were consistent with the results above. Bacterial diversity measures
127 (Shannon and inverse Simpson indices) varied according to clinical status, genotype, and subject
128 (Fig. 4a, supplementary materials Fig. S2 and Table S4). Samples collected during clinical
129 treatments exhibited lower microbial diversity than samples collected at either baseline or recovery
130 visits, highlighting the role of clinical treatments in perturbing CF lung communities (supplementary
131 materials Table S5).
132

133 **Airway microbiomes are functionally consistent and show subject-specific distribution**
134 **patterns**

135 The results of functional metagenomic analyses were consistent with the taxonomic findings
136 described above. Exacerbation events and patient genotype significantly impacted pathway
137 distribution (Table 2, R^2 values of 0.04 and 0.03 respectively, p-values < 0.05), though with less an
138 effect than that of subject ($R^2 = 0.48$). The sample distribution according to the ordination analysis
139 (PCoA) was very heterogeneous with no sharp differences according to genotypes or exacerbation
140 events. Alpha diversity dropped significantly in samples collected during exacerbation events, but

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

141 the drop was significant only considering the inverse Simpson index (p-value = 0.036, Fig. 4b and
142 supplementary materials Table S4). Overall, the pathway distribution was more consistent with
143 respect to the taxonomic one, with biosynthetic pathways being the most represented functional
144 category (Fig. 5, supplementary materials Fig. S2 and Table S5). Pathways were mainly detected in
145 members of the phyla Firmicutes and Proteobacteria, followed by Bacteroidetes and Actinobacteria.
146 Even if these results confirmed the results from the analysis of the taxonomic distribution, metabolic
147 pathways showed a more consistent distribution across samples. Indeed, the beta-diversity analysis
148 on both taxonomic and functional distribution showed a lower similarity based on taxonomy in
149 respect with pathways (Supplementary materials Table S6, Fig. 6a and 6b). These results were
150 additionally confirmed by the differential abundance analysis. For contrasts made within each
151 genotype, only 40 pathways reported significant differences across exacerbation statuses (p-values <
152 0.05 and |log(fold-change)| > 5) all in the homozygote group (Supplementary materials Fig. S3 and
153 Table S7), whereas, considering all samples together, no pathway was found to be more abundant in
154 one condition in respect with another (data not shown). These results confirmed the extraordinary
155 resilience of the CF microbiome evident at the taxonomic level is also exhibited from a functional
156 perspective, indicating that neither clinical change nor antibiotic treatments are accompanied by
157 changes in sputum microbial functions; for example, antibiotics do not appear to select for or against
158 specific functions. To test this specifically, we focused on known antibiotic resistance genes.

159

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

161 Similar to the pathway analysis reported above, antibiotic resistance genes (ARG) were inspected in
162 relation to treatment events. Only six genes were found to be affected by an exacerbation condition,
163 all regarding samples from patients heterozygous for ΔF508 whereas, as found for metabolic

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

164 pathways, no gene was significantly impacted in terms of abundance by antibiotic treatment when
165 considering all samples at once (supplementary materials Fig. S4 and Table S8). A similar approach
166 was used to inspect the effect of antibiotic treatment on ARG distribution. ARG were inspected in
167 relation to the antibiotic treatments reported in supplementary materials Table S1. The class of each
168 antibiotic was correlated to the presence (and the abundance) of genes that may, in principle, confer
169 resistance to antibiotics from the corresponding class. Differential abundance analyses were
170 performed for each classes of antibiotics that was used in this study and results obtained were
171 reported in supplementary materials Fig. S5 and Table S9. Only 11 genes were found to be affected
172 by antibiotic intake in different ways. Indeed, 8 out of 11 reported a reduction of abundance during
173 the treatment whereas the remaining 3 reported an increased abundance in respect with antibiotic
174 intake. Results obtained confirmed the high resilience of the gene composition of CF lung
175 microbiome. The relationship between presence of ARGs with antibiotic treatment was also
176 explored. Results showed a large group of ARGs present in most of the samples and that the
177 antibiotic treatment used in each sample was mirrored by the representation of the ARG classes
178 (Supplementary materials Fig. S6 and Fig. S7).

179 Discussion

180 Longitudinal studies provide important information on the stability and dynamics of microbial
181 ecosystems (23). Here, we investigated the temporal dynamics of the CF sputum microbiome using
182 shotgun metagenomics, including both periods of stability and respiratory exacerbations. The
183 sputum microbiomes of CF patients were highly patient-specific and were substantially impacted
184 both by the host and its lifestyle, suggesting the host has one of the most important determinants of
185 sputum microbiome composition. Indeed, there was less variation within the same individual at
186 different time points than between different individuals at the same time point, proving some degree

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

187 of temporal stability of an individual's sputum microbiome, as indicated by the lack of a time effect
188 on the taxonomic distribution of microbiomes. The predominant taxa detected in sputa of CF
189 patients exhibited extraordinary resilience, as demonstrated by the presence of the same strains of
190 several species during the entire study period. While similar conclusions have been drawn from
191 previous studies using both culture and 16S rRNA gene profiling, these studies failed to report a
192 comprehensive, taxonomy-wide view of strain dynamics due to the limitations of these approaches
193 (4, 6, 24). Carmody and colleagues showed a relatively stable sputum community that was often
194 altered during period of exacerbation even in the absence of viral infection or antibiotic only in a
195 small group of patients (5). A similar result was shown in the work from Fodor and colleagues (4)
196 where, though occasional short-term compositional changes in the airway microbiota were found,
197 the main taxonomic signatures of CF disease were highly stable. Even in other pulmonary diseases,
198 such as non-cystic fibrosis bronchiectasis, respiratory sample bacterial communities showed a
199 conserved structure for long period of time, as showed in the work by Cox and colleagues where
200 patients were followed for a six-month period (24). In our study a notable exception was found for
201 *Rothia mucilaginosa*. In fact, in contrast with other studies where this species was rarely identified
202 (25–27), in our samples it was detected in high relative abundance. This finding may suggest a
203 potential involvement of *R. mucilaginosa* in CF microbiome dynamics and pathogenicity, which
204 deserves further attention.
205 Antibiotic exposure did not result in durable, persistent changes in sputum microbiota; the main taxa
206 linked to CF infection were still present even after aggressive antibiotic treatment. From a
207 taxonomic perspective, samples coming from the same patient clustered together, highlighting the
208 role of the host in bacterial strain selection during the baseline but even during (and after)
209 exacerbation events. Strain selection can be indeed influenced by a number of environmental factors
210 (such as pH level and/or availability of nutrients) that are specific of the lung of a given host and

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

211 cannot be determined *a-priori*. Despite this patient-specific colonization, sputum taxonomic
212 composition differed significantly from one subject to another even when sampled at the same time.
213 Conversely, microbial functional genetic pathways were more homogeneous across patients. This
214 high conservation could indicate common features of the CF lung environment itself, such as mucus
215 composition, nutrient availability, and oxygen levels, which can be broadly similar across patients
216 with a similar clinical status. Such interpretation is consistent with the concept that the function of a
217 biotic community is more conserved than the presence of single members due to functional
218 redundancy of different microbial taxa (28). In fact, though the overall sputum microbiome in our
219 study population included large set of microorganisms, the main functions detected are similar
220 across all patients. From this point of view the airway microbiome can be considered as performing
221 a similar "ecosystem service", irrespective of the taxonomy present as pointed out by various
222 authors in other environments (28, 29). The finding that CFTR genotypes relate with different
223 representation in some pathways, may suggest that the airways microbiome is influenced by the type
224 of CFTR alteration. However, this hypothesis deserves further attention to clarify a putative role of
225 microbial pathways with respect to CFTR genotype and viceversa. Despite a clear effect of
226 antibiotic treatment during (and after) exacerbation periods, the community structure is always
227 recovered with the main pathogenic taxa emerging again. This effect is confirmed by the correlation
228 of ARG distribution and antibiotic intake. Patients subjected to a given antibiotic treatment did not
229 seem to select bacteria resistant to the antibiotic used but the detection of a particular mechanism
230 seems to be distributed in almost all patients regardless of the treatment. An evidence of functional
231 stability of the lung microbiota was previously reported in other works not concerning CF disease
232 (30, 31). Both works focused their attention on the gut microbiome of obese and healthy individuals
233 (human and mouse) reporting a considerable metabolic redundancy. This high degree of redundancy
234 in the gut microbiome supported a more ecological view where subjects can be considered to some

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

235 extent as different ecological niches, inhabited by unique collections of microbial phylotypes, but
236 sharing the same set of genes. This concept can be applied to our cohort of patients, whose lung
237 microbiome was taxonomically variable over time and among individuals (tough all chronically
238 infected by the same species, *P. aeruginosa*), but where it was possible to identify a core set of
239 metabolic-related gene features. This functional conservation may thus be needed by the whole
240 community and patients can be seen as multiple micro-environments inhabited by a peculiar set of
241 strains, which share the same functions. Investigations on the actual functionality (e.g. by
242 metatranscriptomics) of the identified core-set of genes could provide clues on genetic function of
243 the microbiome to be targeted in future therapeutic treatments (10). Additionally, the observed
244 relations of pathway representation with CFTR genotype, though needing to be validated in larger
245 studies, could offer possible opportunities for treating patients by targeting some CFTR genotype-
246 related microbial metabolism. In conclusion, the temporal dynamics of the sputum microbiome in
247 the largest cohort of patients with CF revealed analysed so far, showed patient-specific signatures of
248 the airway microbiome at strain-level, lack of variation in the microbiome across pulmonary
249 exacerbations, and a core set of antibiotic resistance genes that did not vary by antibiotic intake.

250 Materials and Methods

251 Ethics Statement

252 The study was approved by the Ethics Committees of Children's Hospital and Research Institute
253 Bambino Gesù (Rome, Italy), Cystic Fibrosis Center, Anna Meyer Children's University Hospital
254 (Florence, Italy) and G. Gaslini Institute (University of Genoa, Genoa, Italy) [Prot. N. 681 CM of
255 November 2, 2012; Prot. N. 85 of February 27, 2014; Prot. N. FCC 2012 Partner 4-IGG of
256 September 18, 2012]. All participants provided written informed consent before the enrollment in
257 the study. All sputum specimens were produced voluntarily. All procedures were performed in

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

258 agreement with the "Guidelines of the European Convention on Human Rights and Biomedicine for
259 Research in Children" and the Ethics Committee of the three CF Centers involved. All measures
260 were obtained and processed ensuring patient data protection and confidentiality.
261

262 **Characteristics of enrolled patients**

263 Twenty-two adolescents and adults with moderate-severe lung disease and carrying the ΔF508
264 mutation were enrolled in the study between October 2014 and March 2015 (Table 1). The inclusion
265 criteria are described in detail in the supplementary methods. Clinical status at the time of collection
266 was designated as *baseline* (BL), when clinically stable and at their clinical and physiological
267 baseline, *on treatment* (TR), at exacerbation-associated antibiotic treatments, and *at recovery* (RC),
268 upon completion of antibiotic treatment. Subjects were treated according to current standards of care
269 with periodical microbiological controls (32) with at least four microbiological controls per year (2).
270 At each visit, clinical data collection and microbiological status (colonizing pathogens with
271 available cultivation protocols) were performed according to the European CF Society standards of
272 care (33). Forced expiratory volume in 1 second as a percentage of predicted (%FEV_1) is a key
273 outcome of monitoring lung function in CF (34). FEV_1 values were measured according to the
274 American Thoracic Society and European Respiratory Society standards (32). CFTR genotype, sex,
275 age, and antibiotic treatment for each patient were reported in (Table 1 and supplementary materials
276 Table S1). During serial sampling, data (antibiotic usage and spirometry) were collected. A total of
277 79 sputum sample were collected and DNA extraction were performed as reported in supplementary
278 methods.
279

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

280 **Bioinformatic analyses**

281 Sequence quality was ensured by trimming reads using StreamingTrim 1.0 (35), with a quality
282 cutoff of 20. Bowtie2 (36) was used to screen out human-derived sequences from metagenomic data
283 with the latest version of the human genome available in the NCBI database (GRCh38) as reference.
284 Sequences displaying a concordant alignment (mate pair that aligns with the expected relative mate
285 orientation and with the expected range of distances between mates) against the human genomes
286 were then removed from all subsequent analyses. Metabolic and regulatory patterns were estimated
287 using HUMAnN2 (37) and considering only those pathways with a coverage value ≥ 80%, whereas
288 the taxonomic microbial community composition was assessed using MetaPhlAn2 (38). Reads were
289 assembled into contigs using the metaSPAdes microbial assembler (39) with automatic k-mer length
290 selection. To establish an airway microbiome gene catalog (7) we first removed contigs smaller than
291 500bp and then used prodigal in Anonymous mode (40), as suggested by the author of the tool, to
292 predict open reading frames (ORFs). Translated protein sequences obtained from assembled contigs
293 were classified using eggNOG mapper against the bactNOG database (41). Each protein was
294 classified according to its best hit with an e-value lower than 0.001 as suggested in (42). The CARD
295 database (43) was used in combination to the Resistance Gene Identifier (RGI, version 4.0.3) to
296 inspect the distribution of antibiotic resistance gene (AR genes). Genes predicted within each
297 metagenome were quantified using the number of reads that mapped against metagenomic contigs
298 obtained for each sample. Reads were mapped back to contigs using Bowtie2 (36) and the number
299 of reads mapping each ORF was obtained with the bedtools command "multicov" (version 2.26.0).
300 To quantify gene content across different samples, genes were collapsed using the bestOG given by
301 eggNOG mapper by summing together the number of reads that mapped genes with the same
302 annotation. The same approach was used to quantify AR genes predicted with RGI but this time the
303 unique identifier provided by CARD was used to collapse counts.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

304 Strain characterization was performed using StrainPhlAn (20). Sequence variants for each organism
305 detected were assessed against the MetaPhlAn2 (38) marker genes and a tree has been generated
306 including all samples in which the organism was found at least in one time point. Since all
307 organisms detected had at least one reference genome available in the RefSeq database, the most
308 recent version of their genome was downloaded and added to the tree.

309 Acknowledgments

310 We greatly acknowledge Professor Luke Hoffman (University of Washington, School of Medicine,
311 Seattle, WA, USA) for critically reading and reviewing the earlier drafts of the manuscript. We also
312 thank the Italian Cystic Fibrosis Research Foundation (FCC) (https://www.fibrosicisticaricerca.it)
313 for financial support and administrative tasks, and the medical student Marta Maggisano
314 (Sant'Andrea Hospital, Sapienza University of Rome, Rome, Italy) for the stimulating discussion.
315 We would like to express their gratitude to Ricciotti Gabriella (Children's Hospital and Research
316 Institute Bambino Gesù, Rome), Dr. Tuccio Vanessa (Children's Hospital and Research Institute
317 Bambino Gesù, Rome) and Dr. Campana Silvia (Cystic Fibrosis Center, Anna Meyer Children's
318 University Hospital, Department of Pediatrics Medicine, Florence) for their technical support.

319

320 References

321 1. Knapp EA, Fink AK, Goss CH, Sewall A, Ostrenga J, Dowd C, Elbert A, Petren KM,
322 Marshall BC. 2016. The Cystic Fibrosis Foundation Patient Registry. Design and methods of
323 a national observational disease registry. Ann Am Thorac Soc 13:1173–1179.

324 2. Flume PA, Mogayzel PJ, Robinson KA, Goss CH, Rosenblatt RL, Kuhn RJ, Marshall BC,
325 Bujan J, Downs A, Finder J, Goss C, Gutierrez H, Hazle L, Lester M, Quittell L, Sabadosa K,
326 Vender RL, White TB, Willey-Courand DB, Saldanha I, Oyegunle M, Shankar MB, Mckoy
327 N, Sengupta S, Odelola OA, Waybright S. 2009. Cystic fibrosis pulmonary guidelines:
328 Treatment of pulmonary exacerbations. Am J Respir Crit Care Med 180:802–808.

329 3. Huang YJ, LiPuma JJ. 2016. The Microbiome in Cystic Fibrosis. Clin Chest Med 37:59–67.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

330 4. Fodor AA, Klem ER, Gilpin DF, Elborn JS, Boucher RC, Tunney MM, Wolfgang MC. 2012.
331 The Adult Cystic Fibrosis Airway Microbiota Is Stable over Time and Infection Type, and
332 Highly Resilient to Antibiotic Treatment of Exacerbations. PLoS One 7:e45001.

333 5. Carmody LA, Zhao J, Schloss PD, Petrosino JF, Murray S, Young VB, Li JZ, LiPuma JJ.
334 2013. Changes in Cystic Fibrosis Airway Microbiota at Pulmonary Exacerbation. Ann Am
335 Thorac Soc 10:179–187.

336 6. Carmody LA, Zhao J, Kalikin LM, LeBar W, Simon RH, Venkataraman A, Schmidt TM,
337 Abdo Z, Schloss PD, LiPuma JJ. 2015. The daily dynamics of cystic fibrosis airway
338 microbiota during clinical stability and at exacerbation. Microbiome 3:12.

339 7. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez
340 F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y,
341 Tap J, Lepage P, Bertalan M, Batto J-M, Hansen T, Le Paslier D, Linneberg A, Nielsen HB,
342 Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li
343 Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Doré J, Guarner F, Kristiansen K,
344 Pedersen O, Parkhill J, Weissenbach J, Bork P, Ehrlich SD, Wang J. 2010. A human gut
345 microbial gene catalogue established by metagenomic sequencing. Nature 464:59–65.

346 8. Bacci G, Mengoni A, Fiscarelli E, Segata N, Taccetti G, Dolce D, Paganin P, Morelli P,
347 Tuccio V, De Alessandri A, Lucidi V, Bevivino A. 2017. A different microbiome gene
348 repertoire in the airways of cystic fibrosis patients with severe lung disease. Int J Mol Sci
349 18:1654.

350 9. Franzosa EA, Morgan XC, Segata N, Waldron L, Reyes J, Earl AM, Giannoukos G, Boylan
351 MR, Ciulla D, Gevers D, Izard J, Garrett WS, Chan AT, Huttenhower C. 2014. Relating the
352 metatranscriptome and metagenome of the human gut. Proc Natl Acad Sci 111:E2329–
353 E2338.

354 10. Bevivino A, Bacci G, Drevinek P, Nelson MT, Hoffman LR, Mengoni A. 2019. Deciphering
355 the Ecology of Cystic Fibrosis Bacterial Communities: Towards Systems-Level Integration.
356 Trends Mol Med in press.

357 11. Zhao J, Schloss PD, Kalikin LM, Carmody LA, Foster BK, Petrosino JF, Cavalcoli JD,
358 VanDevanter DR, Murray S, Li JZ, Young VB, LiPuma JJ. 2012. Decade-long bacterial
359 community dynamics in cystic fibrosis airways. Proc Natl Acad Sci 109:5809–5814.

360 12. Delhaes L, Monchy S, Fréalle E, Hubans C, Salleron J, Leroy S, Prevotat A, Wallet F,
361 Wallaert B, Dei-Cas E, Sime-Ngando T, Chabé M, Viscogliosi E. 2012. The airway
362 microbiota in cystic fibrosis: A complex fungal and bacterial community-implications for
363 therapeutic management. PLoS One 7:e36313.

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

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

369 15. Lim YW, Evangelista JS, Schmieder R, Bailey B, Haynes M, Furlan M, Maughan H,
370 Edwards R, Rohwer F, Conrad D. 2014. Clinical insights from metagenomic analysis of
371 sputum samples from patients with cystic fibrosis. J Clin Microbiol 52:425–437.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

372 16. Willner D, Haynes MR, Furlan M, Hanson N, Kirby B, Lim YW, Rainey PB, Schmieder R,
373 Youle M, Conrad D, Rohwer F. 2012. Case studies of the spatial heterogeneity of DNA
374 viruses in the cystic fibrosis lung. Am J Respir Cell Mol Biol 46:127–31.

375 17. Whiteson KL, Meinardi S, Lim YW, Schmieder R, Maughan H, Quinn R, Blake DR, Conrad
376 D, Rohwer F. 2014. Breath gas metabolites and bacterial metagenomes from cystic fibrosis
377 airways indicate active pH neutral 2,3-butanedione fermentation. ISME J 8:1247–58.

378 18. Boutin S, Graeber SY, Stahl M, Dittrich AS, Mall MA, Dalpke AH. 2017. Chronic but not
379 intermittent infection with Pseudomonas aeruginosa is associated with global changes of the
380 lung microbiome in cystic fibrosis. Eur Respir J 50.

381 19. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. 2017. Shotgun metagenomics,
382 from sampling to analysis. Nat Biotechnol 35:833–844.

383 20. Truong DT, Tett A, Pasolli E, Huttenhower C, Segata N. 2017. Microbial strain-level
384 population structure and genetic diversity from metagenomes 626–638.

385 21. Scholz M, Ward D V., Pasolli E, Tolio T, Zolfo M, Asnicar F, Truong DT, Tett A, Morrow
386 AL, Segata N. 2016. Strain-level microbial epidemiology and population genomics from
387 shotgun metagenomics. Nat Methods 13:435–438.

388 22. Proesmans M, Balinska-Miskiewicz W, Dupont L, Bossuyt X, Verhaegen J, Høiby N, de
389 Boeck K. 2006. Evaluating the "Leeds criteria" for Pseudomonas aeruginosa infection in a
390 cystic fibrosis centre. Eur Respir J 27:937–943.

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

393 24. Cox MJ, Turek EM, Hennessy C, Mirza GK, James PL, Coleman M, Jones A, Wilson R,
394 Bilton D, Cookson WOC, Moffatt MF, Loebinger MR. 2017. Longitudinal assessment of
395 sputum microbiome by sequencing of the 16S rRNA gene in non-cystic fibrosis
396 bronchiectasis patients. PLoS One 12:1–17.

397 25. Zemanick ET, Wagner BD, Robertson CE, Ahrens RC, Chmiel JF, Clancy JP, Gibson RL,
398 Harris WT, Kurland G, Laguna TA, McColley SA, McCoy K, Retsch-Bogart G, Sobush KT,
399 Zeitlin PL, Stevens MJ, Accurso FJ, Sagel SD, Harris JK. 2017. Airway microbiota across
400 age and disease spectrum in cystic fibrosis. Eur Respir J.

401 26. Carmody LA, Caverly LJ, Foster BK, Rogers MAM, Kalikin LM, Simon RH, VanDevanter
402 DR, LiPuma JJ. 2018. Fluctuations in airway bacterial communities associated with clinical
403 states and disease stages in cystic fibrosis. PLoS One.

404 27. Lim YW, Schmieder R, Haynes M, Willner D, Furlan M, Youle M, Abbott K, Edwards R,
405 Evangelista J, Conrad D, Rohwer F. 2013. Metagenomics and metatranscriptomics: Windows
406 on CF-associated viral and microbial communities. J Cyst Fibros 12:154–164.

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

409 29. Eng A, Borenstein E. 2018. Taxa-function robustness in microbial communities. Microbiome
410 6:45.

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

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

413 1031.

414 31. Thaiss CA, Itav S, Rothschild D, Meijer M, Levy M, Moresi C, Dohnalová L, Braverman S,
415 Rozin S, Malitsky S, Dori-Bachash M, Kuperman Y, Biton I, Gertler A, Harmelin A, Shapiro
416 H, Halpern Z, Aharoni A, Segal E, Elinav E. 2016. Persistent microbiome alterations
417 modulate the rate of post-dieting weight regain. Nature 540:1–27.

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

420 33. Smyth AR, Bell SC, Bojcin S, Bryon M, Duff A, Flume P, Kashirskaya N, Munck A, Ratjen
421 F, Schwarzenberg SJ, Sermet-Gaudelus I, Southern KW, Taccetti G, Ullrich G, Wolfe S.
422 2014. European cystic fibrosis society standards of care: Best practice guidelines. J Cyst
423 Fibros 13:S30–S53.

424 34. Taylor-Robinson D, Whitehead M, Diderichsen F, Olesen H V, Pressler T, Smyth RL, Diggle
425 P. 2012. Understanding the natural progression in %FEV1 decline in patients with cystic
426 fibrosis: a longitudinal study. Thorax 67:860–866.

427 35. Bacci G, Bazzicalupo M, Benedetti A, Mengoni A. 2014. StreamingTrim 1.0: A Java
428 software for dynamic trimming of 16S rRNA sequence data from metagenetic studies. Mol
429 Ecol Resour 14:426–434.

430 36. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods
431 9:357–9.

432 37. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B,
433 Zucker J, Thiagarajan M, Henrissat B, White O, Kelley ST, Methé B, Schloss PD, Gevers D,
434 Mitreva M, Huttenhower C. 2012. Metabolic reconstruction for metagenomic data and its
435 application to the human microbiome. PLoS Comput Biol 8:e1002358.

436 38. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, Tett A, Huttenhower
437 C, Segata N. 2015. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat
438 Methods 12:902–903.

439 39. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. MetaSPAdes: A new versatile
440 metagenomic assembler. Genome Res 27:824–834.

441 40. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal:
442 Prokaryotic gene recognition and translation initiation site identification. BMC
443 Bioinformatics 11:119.

444 41. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende
445 DR, Sunagawa S, Kuhn M, Jensen LJ, Von Mering C, Bork P. 2016. EGGNOG 4.5: A
446 hierarchical orthology framework with improved functional annotations for eukaryotic,
447 prokaryotic and viral sequences. Nucleic Acids Res 44:D286–D293.

448 42. Pearson WR. 2013. An introduction to sequence similarity (“homology”) searching., p. 3.1.1-
449 3.1.8. *In* Current protocols in bioinformatics.

450 43. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, Lago BA, Dave BM,
451 Pereira S, Sharma AN, Doshi S, Courtot M, Lo R, Williams LE, Frye JG, Elsayegh T, Sardar
452 D, Westman EL, Pawlowski AC, Johnson TA, Brinkman FSL, Wright GD, McArthur AG.
453 2017. CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic
454 resistance database. Nucleic Acids Res 45:D566–D573.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

455 44. Bacci G, Paganin P, Lopez L, Vanni C, Dalmastri C, Cantale C, Daddiego L, Perrotta G,
456 Dolce D, Morelli P, Tuccio V, De Alessandri A, Fiscarelli EV, Taccetti G, Lucidi V,
457 Bevivino A, Mengoni A. 2016. Pyrosequencing unveils cystic fibrosis lung microbiome
458 differences associated with a severe lung function decline. PLoS One 11.

459

460

461 **Availability of data and material**

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

465

466 **Conflict of Interest**

467 We have no conflict of interest to declare.

468

469 **Funding**

470 This study was supported by Italian Grants funded by the Italian Cystic Fibrosis Research
471 Foundation (FFC) (http://www.fibrosicisticaricerca.it/) to Annamaria Bevivino: Project n.
472 FFC#14/2015 with the contribution of "Delegazione FFC di Latina", "Latteria Montello Nonno
473 Nanni", and "Gruppo di Sostegno FFC Valle Scrivia Alessandria"), and Project n. FFC#19/2017
474 with the contribution of "Delegazione FFC Lago di Garda con i Gruppi di Sostegno FFC di
475 Chivasso, dell'Isola Bergamasca, di Arezzo". Giovanni Bacci was supported by a postdoctoral
476 fellowship by Italian Cystic Fibrosis Research Foundation. The funders had no role in study design,
477 data collection and analysis, decision to publish, or preparation of the manuscript.

478

479 **Authors' contributions**

480 Conceived and designed the experiments: AB VL GT EVF AM. Performed the experiments: GB

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

481 FDC. Analyzed the data: GB AM AB. Contributed reagents/materials/analysis tools: DD FA PM RS
482 AN. Wrote the paper: GB AM AB. Provided comments and recommendations that improved the
483 manuscript: NS GT VL. Supervised research: AB VL GT EVF AM.
484

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

485 # Tables

486 **TABLE 1**

487 Characteristics of patients enrolled in the study

ID	Hospital	Genotype	Gender	FEV_1 status	Age	n	EX	$\%FEV_1$
B01	OPBG	ΔF508/2183AA->G	M	S	27	5	yes	37.0 ± 1.70
B02	OPBG	ΔF508/N1303K	F	SD	26	3	no	54.7 ± 3.48
B03	OPBG	ΔF508/4016insT	F	S	30	4	no	55.0 ± 1.08
B06	OPBG	ΔF508/ΔF508	F	SD	21	4	no	60.2 ± 3.42
G10	Gaslini	ΔF508/ΔF508	M	S	51	4	no	54.0 ± 3.08
G24	Gaslini	ΔF508/ΔF508	F	S	49	3	yes	NA ± NA
G28	Gaslini	ΔF508/ΔF508	F	NA	38	2	no	42.5 ± 1.50
G30	Gaslini	ΔF508/ΔF508	F	S	50	1	no	54
G31	Gaslini	G1244E/G42X	F	SD	53	2	no	41.5 ± 1.50
G34	Gaslini	ΔF508/ΔF508	F	S	39	1	no	47
M05	Meyer	ΔF508/ΔF508	M	SD	32	4	no	34.8 ± 0.85
M19	Meyer	ΔF508/ΔF508	M	S	24	4	no	44.0 ± 2.04
M21	Meyer	ΔF508/N1303K	M	SD	27	4	yes	51.5 ± 4.35
M22	Meyer	ΔF508/2789+5G->A	F	S	29	5	yes	50.4 ± 1.03
M23	Meyer	ΔF508/G542X	F	S	30	4	yes	37.0 ± 1.47
M24	Meyer	ΔF508/ΔF508	M	S	32	4	no	35.2 ± 0.85
M25	Meyer	ΔF508/296+1G->T	F	SD	41	4	no	42.5 ± 2.02
M26	Meyer	ΔF508/3849+10	F	SD	49	5	yes	39.6 ± 1.94
M28	Meyer	ΔF508/N1303K	M	S	23	4	no	39.0 ± 1.08
M29	Meyer	ΔF508/G542X	F	S	12	4	no	43.5 ± 3.75
M31	Meyer	ΔF508/ΔF508	F	SD	11	3	yes	32.7 ± 4.41
M33	Meyer	ΔF508/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

488

489 ID, study id; Hospital, hospital in which patient has been enrolled [OPBG=Children's Hospital and Research
490 Institute Bambino Gesù (Rome, Italy); Gaslini=G. Gaslini Institute (University of Genoa, Genoa, Italy);
491 Meyer=Cystic Fibrosis Center, Anna Meyer Children's University Hospital (Florence, Italy)]; Genotype,
492 CFTR genotype; Gender, gender; Age, enrollment's age; n, number of samples collected; EX, yes if an
493 exacerbation event has occurred during the study (no otherwise) (2); FEV1, mean value of forced expiratory
494 volume in 1 second plus/minus the standard error on the mean; heterozygote and homozygote refers to ΔF508
495 genotype; %FEV1 status: S = with a rate decline lower than 1.5%, SD = with a rate decline higher than 5%
496 (44).

497

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

498 **TABLE 2**

499 Permutational multivariate analysis of variance on both taxonomic distribution and metabolic pathways

	Df	SumOf Sqs	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**
Subject	**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**
Subject	**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	-	-

500 The permutational multivariate analysis of variance (PERMANOVA) analysis based on taxonomic distribution
501 was reported in the upper part of the table whereas the analysis based on metabolic pathways was reported at the
502 bottom. Df, degrees of freedom; SumOfSqs, sum of squares; R^2, r-squared statistic (reported as proportion); F,
503 F-statistic; Pr(>F), p-value associated to the F-statistic. Significant effects, namely those reporting a p-value
504 lower than 0.05, were reported in bold.

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

505 # Figure legends

506 **FIGURE 1**

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

515 **FIGURE 2**

516 Ordination analyses based on a) taxonomic assignments and b) pathway distribution detected with
517 MetaPhlAn2 and HUMAnN2, respectively. Ordination analyses were conducted using the Bray-
518 Curtis dissimilarity index and ordered following the principle coordinate decomposition method
519 (PCoA). The percentage of variance explained by each coordinate was reported between round
520 brackets. Homozygote and heterozygote refer to ΔF508 mutation of CFTR gene. BL, baseline; TR,
521 treatment; RC, recovery.

522 **FIGURE 3**

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

529 **FIGURE 4**

530 Differences across exacerbation events. The effect of an exacerbation event on alpha diversity was
531 inspected using both the Shannon index and the inverse Simpson index. Diversity indexes were
532 computed for both a) taxonomic signature and b) metabolic pathways. BL, baseline; TR, treatment;
533 RC, recovery. Each box shows the "interquartile range" (IQR) that is the differences between the
534 third and the first quartile of data (the 75th and the 25th percentile). Horizontal bars are medians
535 whereas whiskers represent the minimum and maximum values defined as Q1 – (1.5 x IQR) and Q3

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

536 + (1.5 x IQR), respectively. Observations that fell outside minimum and maximum values were
537 defined as outliers and reported using white points.

538 **FIGURE 5**

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

545 **FIGURE 6**

546 Beta diversity analysis on both taxonomic and functional distribution. a) Hierarchical clustering
547 based on UPGMA method. Clustering was performed on both pathway distribution (the upped
548 triangle) and taxonomic composition of samples (lower triangle). The Bray-Curtis distance was used
549 to compute distances between samples, but it was transformed into similarity value by subtracting 1
550 before plotting. Thus, red colors represent high similarity values whereas blue colors represent low
551 similarity values. The shape of the points on each tip of trees refers to the hospital whereas the
552 colors refer to the exacerbation events. b) Results of Tukey's post hoc test on beta diversity values
553 across patient genotypes and exacerbation events. Contrasts were computed even to test differences
554 between taxonomic distribution and pathways with taxa reporting higher level of beta diversity.
555 Homozygote and heterozygote refer to ΔF508 mutation of CFTR gene. BL, baseline; TR, treatment;
556 RC, recovery. Boxplot were computed as described in Figure 4 legend.

557

bioRxiv preprint doi: https://doi.org/10.1101/609057; this version posted January 27, 2020. 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.

558 # Supplementary materials

559 Supplementary methods

560 **SUPPLEMENTARY TABLES**

561 **TABLE S1** CFTR genotype, sex, age, and antibiotic treatment for each patient
562
563 **TABLE S2** Number of reads for each sample
564
565 **TABLE S3** Summary of all species detected with a mean abundance higher than 0.2%
566
567 **TABLE S4** Analysis of variance on alpha diversity indices
568
569 **TABLE S5** Tukey post hoc test on alpha diversity indices
570
571 **TABLE S6** Tukey post hoc test on Sorensen similarity index
572
573 **TABLE S7** Metabolic pathways differentially distributed across clinical statuses
574
575 **TABLE S8** Antibiotic resistance genes differentially distributed across clinical statuses
576
577 **TABLE S9** Antibiotic resistance genes differentially distributed depending on drug intake

578

579 **SUPPLEMENTARY FIGURES**

580 **FIGURE S1** Strain-level phylogenetic trees of all detected microbes in the study
581
582 **FIGURE S2** Effect of genotypes and samples on the bacterial diversity of lung microbiome
583
584 **FIGURE S3** Differential abundant pathways
585
586 **FIGURE S4** Differential abundant antibiotic resistance genes
587
588 **FIGURE S5** Effect of the antibiotic intake on the distribution of antibiotic resistance genes
589
590 **FIGURE S6** Antibiotic resistance genes map
591
592 **FIGURE S7** Antibiotic resistance map of each sample included in the study

a)

BL TR RC

Firmicutes

Proteobacteria

Bacteroidetes

Actinobacteria

Fusobacteria

Spirochaetes

Ascomycota

Sample

Species

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%)

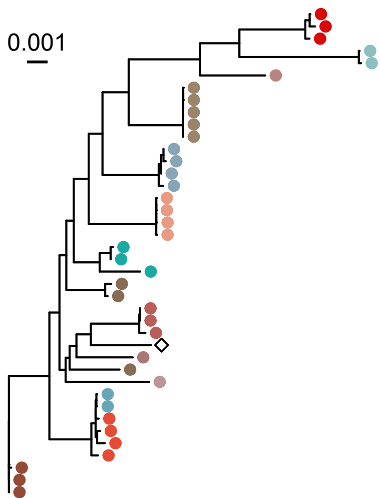
0.4
0.2
0.0

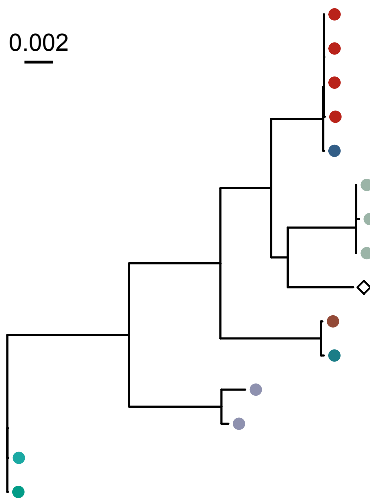
−0.4 −0.2 0.0

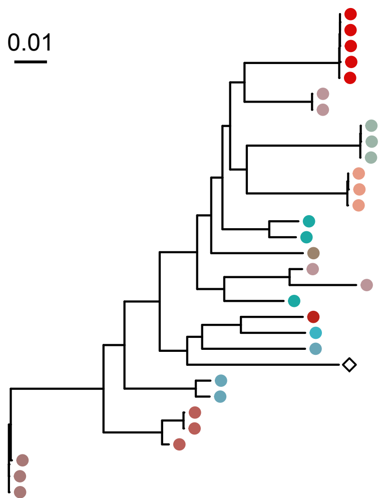
PCoA 1 (48.5%)

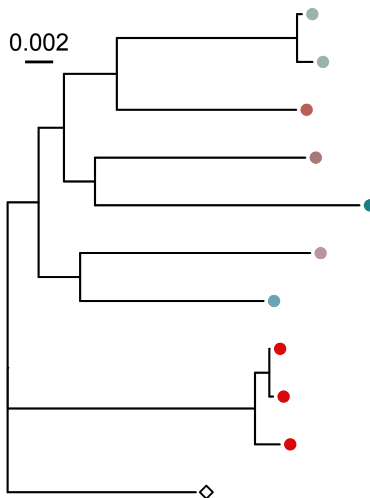
Genotype heterozygote homozygote other

Status TR BL RC


a
0.001

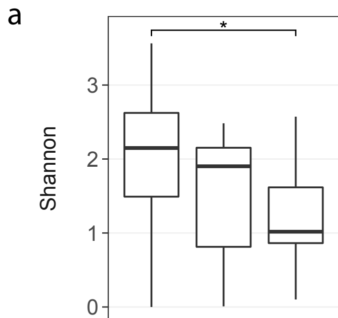

b
0.002

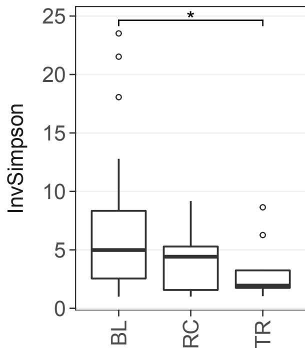

c
0.01


d
0.002

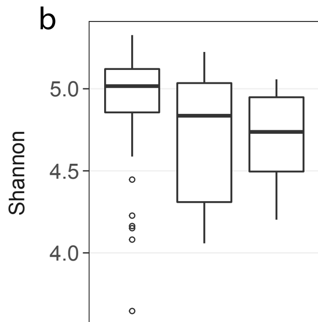
B01
B02
B03
B06
G10
G24
G28
G30
G31
G34
M05
M19
M21
M22
M23
M24
M25
M26
M28
M29
M31
M33
M34
reference

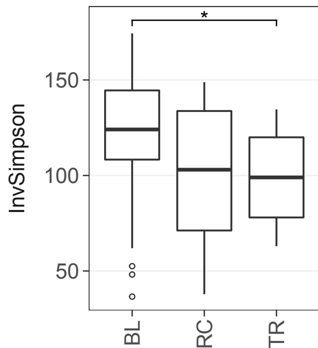
a

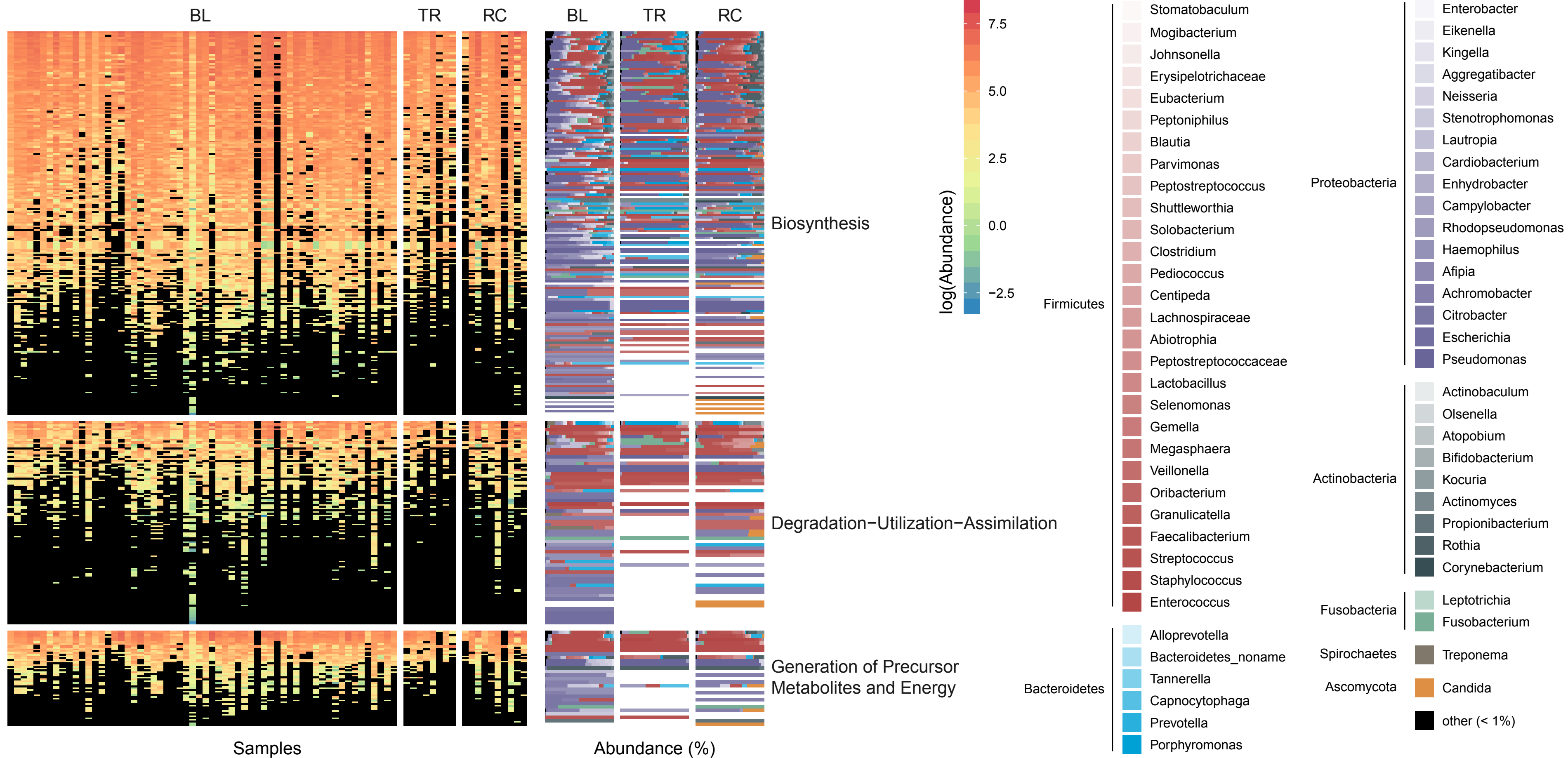




b






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