



**Manchester
Metropolitan
University**

Van Der Gast, CJ and Cuthbertson, L and Rogers, GB and Pope, C and Marsh, RL and Redding, GJ and Bruce, KD and Chang, AB and Hoffman, LR (2014) Three clinically distinct chronic pediatric airway infections share a common core microbiota. *Annals of the American Thoracic Society*, 11 (7). pp. 1039-1048. ISSN 2325-6621

Downloaded from: <http://e-space.mmu.ac.uk/621056/>

Publisher: American Thoracic Society

DOI: <https://doi.org/10.1513/AnnalsATS.201312-456OC>

Please cite the published version

<https://e-space.mmu.ac.uk>

1 **Three clinically distinct chronic pediatric airway infections share a common core microbiota**

2 **Running head:** Common pediatric core respiratory microbiota

3

4 Christopher J. van der Gast¹, Leah Cuthbertson¹, Geraint B. Rogers², Christopher Pope^{3,4}, Robyn L. Marsh⁵,
5 Gregory J. Redding³, Kenneth D. Bruce⁶, Anne B. Chang^{5,7}, Lucas R. Hoffman^{3,4*}

6

7 ¹ *NERC Centre for Ecology & Hydrology, Wallingford, OX10 8BB, UK*

8 ² *Immunity, Infection and Inflammation Program, Mater Research, Translational Research Institute, South
9 Brisbane, QLD, 4101, Australia*

10 *Departments of ³Pediatrics and ⁴Microbiology, University of Washington, Box 356320, HSB K328A, Seattle, WA
11 98105, USA*

12 ⁵ *Menzies School of Health Research, Charles Darwin University, Darwin, NT, 0810, Australia*

13 ⁶ *Institute of Pharmaceutical Science, Molecular Microbiology Research Laboratory, King's College London,
14 London, SE1 9NH, UK*

15 ⁷ *Queensland Children's Respiratory Centre, Royal Children's Hospital, Herston, Brisbane, QLD, 4029, Australia*

16

17 * For correspondence: Address: 4800 Sand Point Way NE, Box A-5937, Seattle Children's Hospital, Seattle,
18 WA, 98105. E-mail: lhoffm@u.washington.edu.

19

20 Word count: Abstract, 328 words. Text, 4,028 words.

21

22 Portions of this work were presented at the 2013 European Cystic Fibrosis Society Conference in Lisbon,
23 Portugal, June, 2013.

24

25 **Descriptor:** 10.11 Pediatrics: Respiratory Infections

26

27 Author contributions: CJvdG designed and performed the study and wrote the manuscript. LC, GBR, CP, and
28 RLM performed analyses and wrote the manuscript. KDB and GJR designed the study and wrote the

29 manuscript. ABC and LRH designed the study, contributed specimens, performed the study, and wrote the
30 manuscript.

31

32 This study was supported by grants from the Cystic Fibrosis Foundation (HOFFMA07P0) and the American
33 Thoracic Society (CF-07-003), the NIH (K02HL105543), the UK Natural Environment Research Council
34 (NE/H019456/1), and Australia's National Health and Medical Research Council grants (1042601, 1040830,
35 1019834 and 1034703).

36

37 This article has an online data supplement, which is accessible from this issue's table of content online at
38 www.atsjournals.org.

39

40

41 **Abstract:**

42 **Rationale:** DNA-based microbiological studies are moving beyond studying healthy human microbiota to
43 investigate diverse infectious diseases, including chronic respiratory infections such as those in the airways of
44 people with cystic fibrosis (CF) and non-CF bronchiectasis (BE). The species identified in the respiratory
45 secretion microbiota from such patients can be classified into those that are common and abundant among
46 similar subjects (core) versus those that are infrequent and rare (satellite). This categorization provides a vital
47 foundation for investigating disease pathogenesis and improving therapy. However, whether the core microbiota
48 of people with different respiratory diseases, which are traditionally associated with specific culturable
49 pathogens, are unique or shared with other chronic infections of the lower airways is not well studied. Little is
50 also known about how these chronic infection microbiota change from childhood into adulthood.

51

52 **Objectives:** We sought to compare the core microbiota in respiratory specimens from children and adults with
53 different chronic lung infections.

54

55 **Methods:** We used bacterial 16S rRNA gene pyrosequencing, phylogenetic analysis, and ecological statistical
56 tools to compare the core microbiota in respiratory samples from three cohorts of symptomatic children with
57 clinically distinct airway diseases (protracted bacterial bronchitis, BE, CF), and four healthy children. We then
58 compared the core pediatric respiratory microbiota with those in samples from adults with BE and CF.

59

60 **Measurements and Main Results:** All three pediatric disease cohorts shared strikingly similar core respiratory
61 microbiota that differed from adult CF and BE microbiota. The most common species in pediatric disease cohort
62 samples were also detected in those from healthy children. The adult CF and BE microbiota also differed from
63 each other, suggesting common early infection airway microbiota that diverge by adulthood. The shared core
64 pediatric microbiota included both traditional pathogens and many species not routinely identified by standard
65 culture.

66

67 **Conclusions:** Our results indicate that these clinically distinct chronic airway infections share common early
68 core microbiota, which are likely shaped by natural aspiration and impaired clearance of the same airway
69 microbes, but that disease-specific characteristics select for divergent microbiota by adulthood. Longitudinal and

70 interventional studies will be required to define the relationships between microbiota, treatments, and disease
71 progression.

72

73 Number of words in this abstract: 349

74

75

76 **Introduction**

77 A growing body of evidence indicates that the airways of even healthy people contain detectable
78 microbes (1–4), and that these microbiota are altered in many chronic lung diseases. For example, culture-
79 independent analyses of cystic fibrosis (CF) respiratory samples have revealed remarkably diverse microbial
80 communities compared with culture results. Among the bacteria detected most frequently, and at highest
81 abundance, in those CF samples were anaerobic species not usually detected by standard laboratory methods
82 (5). Both cross-sectional and longitudinal studies indicate that CF airway microbiota generally decrease in
83 diversity over time, concurrent with a decrease in average lung function (6–9). Whether these ecological
84 changes are the cause of lung function decline, or whether they are the result of increasing antibiotic exposure
85 or changing properties of the infected airways, are not understood (9). Furthermore, exactly where the species
86 that comprise CF secretion microbiota come from, and whether features unique to the CF airway or its
87 secretions select for CF-specific microbiota, are unknown.

88 Several other childhood lung diseases also involve chronic airway infection, including non-CF
89 bronchiectasis (BE) and protracted bacterial bronchitis (PBB) (10). At least some underlying factors that
90 predispose to infection in CF, BE, and PBB may be shared (e.g., mucus hypersecretion or stasis, impaired
91 airway clearance), while others vary (e.g., inflammatory response, intensity of antibiotic use, physicochemical
92 mucus characteristics), and the associated infectious microbiota could therefore differ at the beginning or
93 throughout these diseases. In support of this concept, specific microbes have been associated with the
94 pathogenesis of each of these diseases in culture-based studies. For example, *Staphylococcus aureus* and
95 *Pseudomonas aeruginosa* are traditionally associated with CF lung disease (11). In contrast, *Streptococcus*
96 *pneumoniae*, *Moraxella catarrhalis*, and *Haemophilus influenzae* are traditionally associated with both PBB and
97 early BE (10), with *P. aeruginosa* associated with more advanced BE disease (12). Treatment regimens
98 therefore tend to differ for each disease, designed to target specific pathogens. Moreover, it has been suggested
99 that these bacteria play primary roles in initiating or driving associated symptoms (13), which could imply
100 disease-specific infection pathogenesis. However, it is now known that children with CF have lung disease
101 without detectable infection by *P. aeruginosa* or *S. aureus* (14, 15). The identification of diverse microbiota in
102 even healthy airways further complicates the task of delineating “infected” and/or “pathogenic” from “normal”
103 conditions. Therefore, comparing the earliest stages of chronic infection for diverse airways diseases could be
104 helpful in understanding whether infection begins (and drives disease) in distinct or similar ways in different
105 conditions.

106 An analytical approach that has been particularly useful in understanding environmental and infectious
107 microbiota involves partitioning the constituent microbes into core and satellite species (7). In cross-sectional
108 studies, core species are those that are detected in the majority of samples from subjects with a specific
109 disease, while satellite species are those that are infrequent. While this categorization does not necessarily
110 identify which species are pathogens, it provides a useful framework within which to compare microbiota from
111 different subject groups, and the resulting similarities and differences have important implications for infection
112 pathogenesis. Here, we compared the core and satellite microbiota in three very different cohorts of children with
113 PBB, BE, and CF using next-generation sequencing and ecological analytical approaches. Because these study
114 subjects were children, and therefore at relatively early stages of disease, we hypothesized that these cohorts
115 would have similar, rather than disease-specific, early core respiratory sample microbiota. We then compared
116 pediatric and adult CF and BE respiratory metacommunities, using data from two of our recent adult studies (16,
117 17) to identify evidence of disease-specific changes in microbiota as patients with each disease progress to
118 adulthood.

119

120

121 **Materials and methods**

122 *Patients and clinical samples*

123 All children in this study were participants in previous studies of CF, PBB, and BE, respectively (18–20).
124 Inclusion and exclusion criteria and study details are listed in the Online Supplement. Sputum and
125 bronchoalveolar lavage (BAL) fluid, as well as clinical data, were collected during those studies as described
126 (18–20). The study was approved by the institutional review boards at Seattle Children’s Hospital (SCH) for the
127 CF samples, and at the Royal Children’s Hospital (RCH, Brisbane, Australia) for the PBB, BE and control
128 pediatric samples. Data from two adult BE ($n = 38$ samples, age 37 - 74 years) and CF ($n = 30$, 18 - 55 years)
129 cohorts from previous studies were used for metacommunity comparisons with the microbiota from the pediatric
130 cohorts (16, 17).

131

132 *DNA extraction and Q-PCR*

133 Full details of DNA extraction from BAL and sputum samples, and of the use of quantitative PCR to
134 determine total bacterial abundance using universal eubacterial primers, are in the Online Supplement.

135

136 *Bacterial Tag-Encoded FLX Amplicon Pyrosequencing (bTEFAP) and sequence analyses*

137 The bacterial taxonomy of each sample was evaluated using bacterial tag-encoded FLX-titanium
138 pyrosequencing (bTEFAP) of a fragment of the 16S rRNA gene covering the V1-V3 variable regions, as
139 described previously (21). Details of amplification, sequencing, and sequence processing and analysis are in the
140 Online Supplement.

141

142 *Statistical analyses*

143 To avoid potential biases in comparisons of diversity between local communities due to varying number
144 of sequences per sample, a randomised re-sampling method, using three indices of diversity (S^* , H , and $1-D$),
145 were employed as previously described (7). Bacterial species within each metacommunity were partitioned into
146 core and satellite species groups using the Poisson distribution test as previously described (7) (Online
147 Supplement).

148

149 **Results**

150 *Pediatric study cohorts*

151 We analysed the microbiota in respiratory samples (sputum and BAL fluid) from 3 cohorts of children
152 with BE, PBB, and CF, respectively. All of these children either underwent bronchoscopy for chronic wet cough
153 or were expectorating, indicative of symptomatic disease. As shown in Table 1, the cohorts were substantially
154 different in terms of demographic and treatment features that could play important roles in determining the
155 airway microbiota, including age, antibiotic exposure, and *P. aeruginosa* culture-positivity, which have each been
156 shown to correlate inversely with airway microbiota diversity in CF, and FEV₁% predicted, which correlates
157 directly with diversity (9). The children with CF also had higher rates of inhaled corticosteroid use, a treatment
158 that could conceivably alter airway microbiota constituency (2). The CF cohort originated entirely from a single
159 hospital in the US, whereas all children with BE and PBB were from a single hospital in Australia. Each of these
160 differences, if they were to influence airway microbiota, would be unlikely to bias our analyses towards finding
161 microbial similarities between the three cohorts.

162

163 *Pediatric community diversity and composition*

164 Bacterial pyrosequencing data were used to assess and compare diversity and composition of
165 microbiota from respiratory samples within and across the 3 pediatric cohorts. A total of 408,335 bacterial
166 sequence reads (mean \pm SD per sample, 5041 \pm 3557), identifying 104 genera and 225 distinct operational
167 taxonomic units (OTUs) classified to species level (Table S1), were generated from all samples combined. The
168 average numbers of bacterial sequence reads per sample were similar among the three cohorts (mean \pm SD:
169 BE, 4284 \pm 3476; PBB, 4471 \pm 3718; and CF, 4785 \pm 2702). Estimates of total bacterial abundances were also
170 similar when quantified by qPCR (for all samples, including sputum and BAL cell pellets, from which sufficient
171 DNA remained after sequencing analysis, mean cfu/mL equivalents/sample \pm SD: BE, 1.47 $\times 10^7 \pm 2.32 \times 10^7$
172 $N=18$ samples; PBB, 9.89 $\times 10^6 \pm 1.38 \times 10^7$, $N=10$; and CF, 5.82 $\times 10^7 \pm 6.28 \times 10^7$, $N=21$).

173 Bacterial diversity within and between cohorts was compared using three indices of diversity: species
174 richness (S^* , the total number of species), Shannon-Wiener (H' , a metric that accounts for both number and
175 relative abundance of species), and Simpson's (1- D , a measure of the probability that two species randomly
176 selected from a sample will differ). Bacterial diversity between BE sputum and BAL samples was compared
177 using all three measures and found not to be significantly different ($P > 0.05$, Fig. S1), and hence those samples

178 were combined for subsequent analyses. S^* was found to vary substantially within cohorts, with means of $30.6 \pm$
179 10.5 , 30.9 ± 11.5 , and 19.5 ± 10.0 for the BE, PBB, and CF cohorts, respectively (Fig. 1); similar results were
180 obtained for H' and $1-D$ (Fig. 1 and Table S2). Diversity between pediatric cohorts was not significantly different
181 ($P > 0.05$) in all instances except for one measure (H') between the PBB and CF cohorts (Table S2).

182 Similarities and differences in community membership and structure for microbiota samples both within
183 and between disease cohorts were assayed with two different indices: Sørensen (S_{SOR} , which accounts for the
184 number of species present in each community and those that are shared) and Bray-Curtis quantitative (S_{BC} ,
185 similar to S_{SOR} but also accounts for abundance of each species). Each index varies in value from 0 to 1, with
186 higher values indicating greater similarity. It has been previously observed that community composition is highly
187 variable between CF patients (7, 22). Here, we found that observation held regardless of underlying pediatric
188 lung disease, with mean similarities in bacterial community membership taken pair-wise for samples within
189 pediatric cohorts of $S_{SOR} = 0.44 \pm 0.14$ for BE ($n = 171$ pair wise comparisons); $S_{SOR} = 0.51 \pm 0.12$ for PBB ($n =$
190 66); and $S_{SOR} = 0.50 \pm 0.12$ for CF ($n = 300$), with a mean similarity of the entire pool of samples of $S_{SOR} = 0.46 \pm$
191 0.13 ($n = 1540$) (Fig. 2). For S_{BC} , the mean between all samples was $S_{BC} = 0.19 \pm 0.15$ ($n = 1540$), while that
192 within cohorts was $S_{BC} = 0.18 \pm 0.14$ for BE ($n = 171$); $S_{BC} = 0.27 \pm 0.12$ for PBB ($n = 66$); and $S_{BC} = 0.27 \pm 0.17$
193 for CF ($n = 300$). Therefore, individual sample microbiota differed within each pediatric disease group to a similar
194 degree, and the microbiota identified within each pediatric cohort's samples were pooled together into disease-
195 specific metacommunities (defined as the set of all microbiota identified in samples from a given cohort) for
196 subsequent comparisons.

197

198 *Microbiota within a metacommunity framework*

199 Previously, we established that the categorization of component species in CF microbiota into core and
200 satellite species revealed important aspects of species-abundance distributions within a metacommunity that
201 would be neglected without such a distinction (7). A coherent metacommunity could be expected to exhibit a
202 direct relationship between prevalence and abundance of individual species within the constituent communities.
203 Consistent with this prediction, the abundance of species in each pediatric study cohort was significantly
204 correlated with the number of individual respiratory sample communities those species occupied (Fig. 3A).

205 The species identified in each pediatric cohort metacommunity were then classified as core or satellite
206 based upon their distributions (Fig. 3B). Of the 177 species that comprised the BE metacommunity, 88 were core

207 and 89 were satellite species. The PBB metacommunity (130 species) was comprised of 63 core and 67 satellite
208 species, and the CF metacommunity (143 species) comprised of 66 and 77 species, respectively.

209 Similarities between the three pediatric metacommunities were assayed using S_{SOR} and S_{BC} indices for
210 all species, and for the core and satellite groups. The resulting cluster analysis revealed community membership
211 to be relatively well-conserved between cohorts for all species ($S_{SOR} = 0.74 \pm 0.02$ [$n = 3$ pair-wise
212 comparisons]), but much more highly conserved between the core groups ($S_{SOR} = 0.93 \pm 0.02$) (Fig. 4). The
213 satellite groups, comprised of randomly distributed species, were highly divergent ($S_{SOR} = 0.35 \pm 0.01$). Results
214 from Bray-Curtis quantitative analyses were similar: All $S_{BC} = 0.50 \pm 0.10$; core $S_{BC} = 0.51 \pm 0.10$; and satellite
215 $S_{BC} = 0.10 \pm 0.02$. These results indicate that the high level of similarity between the three pediatric cohort
216 metacommunities is attributable to core species.

217 To determine whether disease-specific microbiota emerge with disease progression into adulthood for
218 the two chronic diseases, BE and CF (PBB usually improves long-term with antibiotic treatment (10)), we
219 compared the microbiota from our previous studies of adult CF (16) and BE (17) cohorts with the pediatric cohort
220 microbiota. The adult BE cohort metacommunity ($n = 38$ patients) was comprised of 86 core and 54 satellite
221 species, while the adult CF cohort ($n = 30$) metacommunity was comprised of 68 core and 81 satellite species.
222 The resulting analyses revealed adult community membership to be highly divergent in all cases (whole
223 microbiota $S_{SOR} = 0.36 \pm 0.03$; core $S_{SOR} = 0.46 \pm 0.06$; and satellite $S_{SOR} = 0.09 \pm 0.03$ [$n = 6$ pairwise
224 comparisons in all cases]) (Fig. 4). When relative abundances of constituent taxa were included to examine
225 differences in community structure, an even more pronounced divergence was observed than for membership
226 (whole $S_{BC} = 0.07 \pm 0.03$; core $S_{BC} = 0.07 \pm 0.04$; and satellite $S_{BC} = 0.01 \pm 0.01$ [$n = 6$ in all cases]). In addition,
227 comparisons between the adult CF and BE metacommunities also revealed a high degree of divergence both in
228 terms of membership (whole $S_{SOR} = 0.38$; core $S_{SOR} = 0.48$; and satellite $S_{SOR} = 0.16$) (Fig. 4) and structure
229 (whole $S_{BC} = 0.05$; core $S_{BC} = 0.05$; and satellite $S_{BC} = 0.02$). Therefore, the adult CF and BE microbiota differed
230 substantially from each other, contrasting with the similarities among the pediatric cohort microbiota.

231 To test whether the adult and pediatric metacommunities were significantly similar or dissimilar, the
232 whole, core, and satellite microbiota were compared between cohorts using the Raup and Crick probability-
233 based similarity index (S_{RC}) to determine whether compositional similarities in the microbiota were more or less
234 significantly similar than expected by chance (Fig. 5). The resulting cluster analysis revealed the pediatric whole,
235 core, and satellite microbiota to be significantly dissimilar ($S_{RC} < 0.05$) from the corresponding adult CF and BE

236 microbiota. Likewise, the whole and core microbiota from the pediatric cohorts were significantly similar ($S_{RC} >$
237 0.95), adding further weight to the observations of pediatric community membership (S_{SOR}) and structure (S_{BC})
238 (Fig. 5). In contrast, the pediatric satellite microbiota were not significantly similar or dissimilar to each other
239 ($S_{RC} > 0.05$ and < 0.95), which would be expected given that satellite microbiota in a given cohort represent rare
240 and randomly distributed species. Furthermore, the adult CF and BE microbiota were significantly dissimilar to
241 each other ($S_{RC} < 0.05$) (Fig. 5).

242 Similarity of percentages (SIMPER) analysis of the metacommunities was used to identify those species
243 that contributed most to the observed similarity between the three pediatric cohorts. These species, all core
244 group members, are listed in decreasing order of contribution in Table 2. *H. influenzae* contributed the greatest
245 amount to the observed similarity between pediatric samples. Most of the major contributors have previously
246 been associated with respiratory tract infections and other opportunistic infections, and many are known
247 inhabitants of the oral cavity, including both aerobic and anaerobic species (Table S1). These findings
248 contrasted with the SIMPER analysis of the adult CF cohort, wherein *P. aeruginosa* contributed over 83% to the
249 observed similarity between samples (Table S3a) in addition to 3 other species: *Streptococcus pneumoniae*,
250 *Prevotella melaninogenica*, and *Veillonella parvula*. Within the adult BE cohort, the main contributors to the
251 overall similarity between samples were *H. influenzae* (50.4%), *P. aeruginosa* (8.62 %) and *V. dispar* (8.11%)
252 (Table S3b) along with *S. pneumoniae* and, as for the pediatric cohorts, other species associated with
253 respiratory infections, many of which are known oral microbiota members. As also observed for the pediatric
254 cohorts, the species identified from the SIMPER analyses for both adult cohorts were all core species in their
255 respective metacommunities. Although there were some species identified in common for the adult and pediatric
256 cohorts by the SIMPER analyses, these similarities were greatly outweighed by the differences in species
257 content, accounting for the observed differences in whole and core microbiota within each adult-pediatric
258 disease cohort pair.

259 To investigate whether the core microbiota shared between the three pediatric disease groups were
260 unique to children with airways disease, we identified the microbiota in BALF from four children who underwent
261 bronchoscopy for reasons other than suspected infection (described in the Online Supplement and ¹⁶). Of the 10
262 shared core species that contributed the most to similarity between disease cohorts from Table 2, six were
263 detected in all four “control” samples, and the other four species were detected in three of the four samples,
264 suggesting that the shared core microbiota are commonly found in healthy, as well as diseased, children’s

265 airways (Table S4), consistent with prior airway microbiota studies that included samples from healthy
266 subjects^{3,21}.
267

268 **Discussion**

269 We found that cohorts of symptomatic children with three clinically distinct airways diseases- CF, BE,
270 and PBB- shared remarkably similar respiratory sample core microbiota, the most common of which were also
271 found in airway samples from children without clinical evidence of infection. The shared core microbiota included
272 both traditional pathogens and many bacteria that are either not identified by routine clinical laboratory methods
273 or are generally classified as “oral flora” and are routinely ignored with respect to pathophysiology or treatment.
274 In contrast, the respiratory sample microbiota from adults with CF and BE differed significantly from each other
275 and from those of children with the same disease. These results indicate that the core microbes in different
276 chronic respiratory infection types likely begin similarly, yet are divergent by adulthood.

277 While these observations provide valuable insight into the microbiota in diseased airways, they also
278 identify next steps to take in understanding the pathogenesis of diverse airway infections. Given the high
279 abundance and prevalence of the core species among pediatric patients, one interpretation of these results
280 could be that the core species represent those that contribute the most to airway inflammation and disease
281 progression. Alternatively, it could be that the satellite species play key pathogenic roles in specific contexts or
282 stages of infection; for example, infection at any stage with specific low-abundance or infrequent bacteria could
283 lead to increased inflammation, airway damage, mucus hypersecretion, and/or clinical deterioration. While these
284 questions are not fully answered by the current study, the relative rarity of any specific species in each satellite
285 group may indicate a consistent, general role for the core groups in pathogenesis. Future studies will be required
286 to longitudinally compare the changes over time in airway microbiota in health and disease to better understand
287 the microbial contribution to pathogenesis, and the best therapeutic approach to infection.

288 These findings provide valuable perspective for results from studies of many respiratory diseases,
289 usually at more advanced stages (2). A study using a relatively low-depth sequencing technique, 16S rRNA
290 gene clone sequencing, comparing microbiota in explanted lungs from adults with end-stage BE and CF
291 identified diverse communities often dominated by *Pseudomonas* in both sample sets (23). Likewise, two studies
292 using the same technique applied to sputa from patients with CF and BE found some similarities and some
293 differences between the subject groups, but no formal comparisons were made (24, 25). In support of our results
294 (17), a recent pyrosequencing study of 21 older adults with non-CF bronchiectasis showed their sputum to
295 contain diverse microbiota (26). Studies of sputum, BAL fluid, and explanted lung tissue from patients at various
296 stages of disease due to chronic obstructive pulmonary disease (COPD) and asthma, sometimes comparing with

297 healthy subjects, have yielded varying and often conflicting results, in some cases demonstrating disease-
298 specific microbiota and in others identifying substantial overlap (1, 2, 27–29). In all of these studies, the effects
299 of therapies that could conceivably impact airway microbiota, such as intubation, steroids and antibiotics, were
300 difficult to control for. Our results suggest that the microbiota observed in those older patients, at later stages of
301 disease, likely had similar (or identical) beginnings and diverged over time. Whether those changes resulted
302 from features of the underlying disease or from differences in treatment are not yet clear. While the differences in
303 antibiotic use among study cohorts support the idea that antibiotics did not have a profound, lasting impact on
304 airway microbiota (similar to observations from CF (8, 30)), and there was no evidence that differences in steroid
305 use resulted in differences in microbiota in these children, chronic or repetitive drug exposure could have
306 gradual effects (as indicated by longitudinal studies in CF (9)). Regardless, these findings suggest that factors
307 intrinsic to airways predominate in establishing the early microbiota in these chronic respiratory infections.

308 Several observations support these conclusions. Culture-based studies of children with CF
309 demonstrated that radiographic lung disease, respiratory symptoms, and inflammation all precede the detection
310 of “standard CF pathogens” (14, 15, 31, 32). All of the shared core microbes identified in the current study, many
311 of which would traditionally be considered “oral flora”, were also commonly identified in oropharyngeal swabs
312 from infants with CF by pyrosequencing (33). The identification of such “oral flora” has been associated with both
313 inflammation (32) and structural lung disease (15) in children with CF. Similarly, piglets with engineered CFTR
314 mutations (34) concurrently develop diverse airway microbiota and lung disease similar to those observed in
315 children with CF. The cultured microbiota of CF pigs were similar in constituency to those of wild-type pigs, but
316 more persistent and higher in abundance. Therefore, the requirement for traditional pathogens for CF lung
317 disease or symptom development is questionable.

318 Both cross-sectional and longitudinal studies of CF respiratory samples using molecular methods have
319 demonstrated correlations between decreasing microbiota diversity, age, and disease progression (3, 8, 9, 30).
320 Interestingly, some organisms traditionally associated with chronic airways disease were classified as core
321 species in some disease cohorts but not others in our study (for example, *P. aeruginosa*, *S. aureus*, and several
322 species each of *Veillonella* and *Prevotella* were core members in both the CF and BE metacommunities, but
323 satellite for PBB, Table S1). Collectively, these results could indicate that particular microbes are associated with
324 and/or accelerate specific lung diseases. Alternatively, changes over time in airway microbiota, including the
325 gradual decrease in community diversity observed in CF and in COPD (28), and the gradual identification of *P.*

326 *aeruginosa* in people with CF and non-CF bronchiectasis (23), could all result from intensified antibiotic
327 treatment and/or underlying disease, but this concept would not require the microbes that eventually dominate to
328 be particularly pathogenic. Similarly, changes in the respiratory microbiota could result from evolution (therapy-
329 driven or otherwise) in another reservoir that seeds the airways, such as the oropharynx (as suggested by
330 abundance in our study of microbes traditionally associated with this space (35)) or the GI tract (which a recent
331 longitudinal study of infants with CF suggested could be a source for respiratory microbiota (33)).

332 Together with the above earlier studies, our results support a unified model of early airway infections, in
333 which a defect in the clearance of otherwise normal airway microbiota (for example, due to early airway injury,
334 altered mucus, defective ciliary motility, mechanical obstruction, or immunodeficiency) contributes to progressive
335 inflammation and airway damage. In this model, an extension of the “vicious cycle hypothesis” (13), either
336 medical treatment or the physical effects of underlying disease (such as diseased tissue or accumulated
337 secretions) could lead to progressive change in the microbiota. These microbiological changes (including
338 eventual selection for *P. aeruginosa* in CF or BE) could in turn accelerate lung disease, or they could be
339 consequences rather than causes of this progression. Future studies, either interventional or longitudinal, may
340 be able to refine this model.

341 Antibiotics play important roles in the management of all three of these symptomatic chronic infections.
342 For example, inhaled and oral antibiotics are commonly used in CF and BE both for maintenance (11, 36) and
343 exacerbation treatments (11, 13). For PBB, a prolonged course of antibiotics usually leads to resolution of
344 symptoms (18). These observations underscore the importance of bacteria in the pathogenesis of each disease;
345 the current findings provide a rationale for reexamining and comparing the culture-based antibiotic approaches
346 to each of these airway infections, particularly during early stages.

347 Our study was limited by the small sizes of each subject group. The groups were not matched for age,
348 treatment, disease severity, or demographics, limiting our ability to determine relationships between microbiota
349 and these factors. However, the similarity in microbiota among demographically and clinically divergent pediatric
350 groups could be interpreted to further support the concept of a common early pediatric respiratory microbiota;
351 additional study with demographically well-matched cohorts could strengthen this argument. We did not
352 distinguish between viable and nonviable cells in the current analysis due to the small volumes of many of the
353 samples analyzed, and thus the detection of a microbe cannot definitively be said to indicate its persistence
354 (although even transient or nonviable bacterial cells could potentially contribute to airway inflammation).

355 Similarly, it could be proposed that the microbial diversity found in our study and others is due to contamination
356 by upper airway bacteria during sample collection. However, the significant similarity in microbiota among
357 sputum and BAL samples from the children with BE in our study, as observed in studies of children with CF (5,
358 22), as well as the high abundance of microbes in our samples, argue against this issue significantly altering our
359 findings.

360 In conclusion, children with early chronic lung infections due to three distinct diseases shared strikingly
361 similar core airway microbiota that included many bacteria not generally identified by clinical laboratory culture.
362 These similarities were independent of underlying diagnosis or disease severity, geography, antibiotic and
363 steroid use, and age. These results indicate that different chronic airway infections begin similarly, remaining
364 similar even after symptoms have begun, but diverging by adulthood. Our findings provide a baseline to
365 compare with in future, longitudinal microbiota studies of diverse chronic airway infections, which would provide
366 vital insight into the microbial determinants of chronic lung disease progression and could lead to improved
367 treatments.

368

369 **Acknowledgements**

370 We thank the participating patients and their families, B. van Yserloo and W. Osborne for technical
371 assistance, and R. Gibson and H. Smith-Vaughan for helpful discussions. This study was supported by grants
372 from the Cystic Fibrosis Foundation (HOFFMA07P0), the NIH (K02HL105543) and the American Thoracic
373 Society (CF-07-003) to LRH, UK Natural Environment Research Council (NE/H019456/1) to CJvdG. The studies
374 in Brisbane were supported by Australia's National Health and Medical Research Council grants (1042601,
375 1040830, 1019834 and 1034703).

376

377 Data deposition: The sequence data reported in this paper have been submitted to the NCBI Short Read
378 Archive database (Bioproject accession number: PRJNA200702).

379

380

381

382
383

References

- 384 1. Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, Bushman FD, Collman RG.
385 Topographical continuity of bacterial populations in the healthy human respiratory tract. *Am J Respir Crit*
386 *Care Med* 2011;184:957–963.
- 387 2. Han MK, Huang YJ, Lipuma JJ, Boushey HA, Boucher RC, Cookson WO, Curtis JL, Erb-Downward J, Lynch
388 SV, Sethi S, Toews GB, Young VB, Wolfgang MC, Huffnagle GB, Martinez FJ. Significance of the
389 microbiome in obstructive lung disease. *Thorax* 2012;67:456–463.
- 390 3. Blainey PC, Milla CE, Cornfield DN, Quake SR. Quantitative analysis of the human airway microbial ecology
391 reveals a pervasive signature for cystic fibrosis. *Sci Transl Med* 2012;4:153ra130.
- 392 4. Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, Flores SC, Fontenot AP, Ghedin E,
393 Huang L, Jablonski K, Kleerup E, Lynch SV, Sodergren E, Twigg H, Young VB, Bassis CM, Venkataraman
394 A, Schmidt TM, Weinstock GM, Lung HIV Microbiome Project. Comparison of the respiratory microbiome in
395 healthy nonsmokers and smokers. *Am J Respir Crit Care Med* 2013;187:1067–1075.
- 396 5. Rogers GB, Carroll MP, Serisier DJ, Hockey PM, Jones G, Bruce KD. characterization of bacterial community
397 diversity in cystic fibrosis lung infections by use of 16s ribosomal DNA terminal restriction fragment length
398 polymorphism profiling. *J Clin Microbiol* 2004;42:5176–5183.
- 399 6. Cox MJ, Allgaier M, Taylor B, Baek MS, Huang YJ, Daly RA, Karaoz U, Andersen GL, Brown R, Fujimura KE,
400 Wu B, Tran D, Koff J, Kleinhenz ME, Nielson D, Brodie EL, Lynch SV. Airway microbiota and pathogen
401 abundance in age-stratified cystic fibrosis patients. *PLoS ONE* 2010;5:e11044.
- 402 7. Van der Gast CJ, Walker AW, Stressmann FA, Rogers GB, Scott P, Daniels TW, Carroll MP, Parkhill J, Bruce
403 KD. Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *ISME J*
404 2011;5:780–791.
- 405 8. Stressmann FA, Rogers GB, van der Gast CJ, Marsh P, Vermeer LS, Carroll MP, Hoffman L, Daniels TWV,
406 Patel N, Forbes B, Bruce KD. Long-term cultivation-independent microbial diversity analysis demonstrates
407 that bacterial communities infecting the adult cystic fibrosis lung show stability and resilience. *Thorax*
408 2012;doi:10.1136/thoraxjnl-2011-200932.
- 409 9. Zhao J, Schloss PD, Kalikin LM, Carmody LA, Foster BK, Petrosino JF, Cavalcoli JD, VanDevanter DR,
410 Murray S, Li JZ, Young VB, LiPuma JJ. Decade-long bacterial community dynamics in cystic fibrosis
411 airways. *Proc Natl Acad Sci USA* 2012;109:5809–5814.

- 412 10. Chang AB, Redding GJ, Everard ML. Chronic wet cough: Protracted bronchitis, chronic suppurative lung
413 disease and bronchiectasis. *Pediatr Pulmonol* 2008;43:519–531.
- 414 11. Gibson RL, Burns JL, Ramsey BW. Pathophysiology and management of pulmonary infections in cystic
415 fibrosis. *Am J Respir Crit Care Med* 2003;168:918–951.
- 416 12. Redding GJ. Bronchiectasis in children. *Pediatr Clin North Am* 2009;56:157–171, xi.
- 417 13. Chalmers JD, Smith MP, McHugh BJ, Doherty C, Govan JR, Hill AT. Short- and long-term antibiotic
418 treatment reduces airway and systemic inflammation in non-cystic fibrosis bronchiectasis. *Am J Respir Crit*
419 *Care Med* 2012;186:657–665.
- 420 14. Sagel SD, Gibson RL, Emerson J, McNamara S, Burns JL, Wagener JS, Ramsey BW. Impact of
421 *Pseudomonas* and *Staphylococcus* infection on inflammation and clinical status in young children with cystic
422 fibrosis. *J Pediatr* 2009;154:183–188.
- 423 15. Mott LS, Park J, Murray CP, Gangell CL, de Klerk NH, Robinson PJ, Robertson CF, Ranganathan SC, Sly
424 PD, Stick SM. Progression of early structural lung disease in young children with cystic fibrosis assessed
425 using CT. *Thorax* 2012;67:509–516.
- 426 16. Rogers GB, Cuthbertson L, Hoffman LR, Wing PAC, Pope C, Hoofman DAP, Lilley AK, Oliver A, Carroll
427 MP, Bruce KD, van der Gast CJ. Reducing bias in bacterial community analysis of lower respiratory
428 infections. *ISME J* 2013;7:697–706.
- 429 17. Rogers GB, van der Gast CJ, Cuthbertson L, Thomson SK, Bruce KD, Martin ML, Serisier DJ. Clinical
430 measures of disease in adult non-CF bronchiectasis correlate with airway microbiota composition. *Thorax*
431 2013;68:731–737.
- 432 18. Chang AB, Yerkovich ST, Gibson PG, Anderson-James S, Petsky HL, Carroll ML, Masters IB, Marchant JM,
433 Wurzel D, Upham JW. Pulmonary innate immunity in children with protracted bacterial bronchitis. *J Pediatr*
434 2012;161:621–625.e1.
- 435 19. Kapur N, Grimwood K, Masters IB, Morris PS, Chang AB. Lower airway microbiology and cellularity in
436 children with newly diagnosed non-CF bronchiectasis. *Pediatr Pulmonol* 2012;47:300–307.
- 437 20. Wolter DJ, Emerson JC, McNamara S, Buccat AM, Qin X, Cochrane E, Houston LS, Rogers GB, Marsh P,
438 Prehar K, Pope CE, Blackledge M, Déziel E, Bruce KD, Ramsey BW, Gibson RL, Burns JL, Hoffman LR.
439 *Staphylococcus aureus* small-colony variants are independently associated with worse lung disease in
440 children with cystic fibrosis. *Clin Infect Dis* 2013;doi:10.1093/cid/cit270.

- 441 21. Dowd SE, Wolcott RD, Sun Y, McKeehan T, Smith E, Rhoads D. Polymicrobial nature of chronic diabetic
442 foot ulcer biofilm infections determined using bacterial tag encoded FLX amplicon pyrosequencing
443 (bTEFAP). *PLoS ONE* 2008;3:e3326.
- 444 22. Harris JK, De Groote MA, Sagel SD, Zemanick ET, Kapsner R, Penvari C, Kaess H, Deterding RR, Accurso
445 FJ, Pace NR. Molecular identification of bacteria in bronchoalveolar lavage fluid from children with cystic
446 fibrosis. *Proc Natl Acad Sci USA* 2007;104:20529–20533.
- 447 23. Maughan H, Cunningham KS, Wang PW, Zhang Y, Cypel M, Chaparro C, Tullis DE, Waddell TK, Keshavjee
448 S, Liu M, Guttman DS, Hwang DM. Pulmonary bacterial communities in surgically resected noncystic
449 fibrosis bronchiectasis lungs are similar to those in cystic fibrosis. *Pulm Med* 2012;2012:746358.
- 450 24. Bittar F, Richet H, Dubus J-C, Reynaud-Gaubert M, Stremier N, Sarles J, Raoult D, Rolain J-M. Molecular
451 detection of multiple emerging pathogens in sputa from cystic fibrosis patients. *PLoS ONE* 2008;3:e2908.
- 452 25. Duff RM, Simmonds NJ, Davies JC, Wilson R, Alton EW, Pantelidis P, Cox MJ, Cookson WO, Bilton D,
453 Moffatt MF. A molecular comparison of microbial communities in bronchiectasis and cystic fibrosis. *Eur*
454 *Respir J* 2013;41:991–993.
- 455 26. Tunney MM, Einarsson GG, Wei L, Drain M, Klem ER, Cardwell C, Ennis M, Boucher RC, Wolfgang MC,
456 Elborn JS. The Lung Microbiota and Bacterial Abundance in Patients With Bronchiectasis When Clinically
457 Stable and During Exacerbation. *Am J Respir Crit Care Med* 2013;doi:10.1164/rccm.201210-1937OC.
- 458 27. Huang YJ, Nelson CE, Brodie EL, Desantis TZ, Baek MS, Liu J, Woyke T, Allgaier M, Bristow J, Wiener-
459 Kronish JP, Sutherland ER, King TS, Icitovic N, Martin RJ, Calhoun WJ, Castro M, Denlinger LC, Dimango
460 E, Kraft M, Peters SP, Wasserman SI, Wechsler ME, Boushey HA, Lynch SV. Airway microbiota and
461 bronchial hyperresponsiveness in patients with suboptimally controlled asthma. *J Allergy Clin Immunol*
462 2011;127:372–381.e1–3.
- 463 28. Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, Young VB, Toews
464 GB, Curtis JL, Sundaram B, Martinez FJ, Huffnagle GB. Analysis of the lung microbiome in the “healthy”
465 smoker and in COPD. *PLoS ONE* 2011;6:e16384.
- 466 29. Sze MA, Dimitriu PA, Hayashi S, Elliott WM, McDonough JE, Gosselink JV, Cooper J, Sin DD, Mohn WW,
467 Hogg JC. The lung tissue microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*
468 2012;185:1073–1080.

- 469 30. Fodor AA, Klem ER, Gilpin DF, Elborn JS, Boucher RC, Tunney MM, Wolfgang MC. The adult cystic fibrosis
470 airway microbiota is stable over time and infection type, and highly resilient to antibiotic treatment of
471 exacerbations. *PLoS ONE* 2012;7:e45001.
- 472 31. Zemanick ET, Wagner BD, Sagel SD, Stevens MJ, Accurso FJ, Harris JK. Reliability of quantitative real-time
473 PCR for bacterial detection in cystic fibrosis airway specimens. *PLoS ONE* 2010;5:e15101.
- 474 32. Gangell C, Gard S, Douglas T, Park J, de Klerk N, Keil T, Brennan S, Ranganathan S, Robins-Browne R,
475 Sly PD. Inflammatory responses to individual microorganisms in the lungs of children with cystic fibrosis.
476 *Clin Infect Dis* 2011;53:425–432.
- 477 33. Madan JC, Koestler DC, Stanton BA, Davidson L, Moulton LA, Housman ML, Moore JH, Guill MF, Morrison
478 HG, Sogin ML, Hampton TH, Karagas MR, Palumbo PE, Foster JA, Hibberd PL, O'Toole GA. Serial analysis
479 of the gut and respiratory microbiome in cystic fibrosis in infancy: interaction between intestinal and
480 respiratory tracts and impact of nutritional exposures. *MBio* 2012;3:
- 481 34. Stoltz DA, Meyerholz DK, Pezzulo AA, Ramachandran S, Rogan MP, Davis GJ, Hanfland RA, Wohlford-
482 Lenane C, Dohrn CL, Bartlett JA, Nelson GA 4th, Chang EH, Taft PJ, Ludwig PS, Estin M, Hornick EE,
483 Launspach JL, Samuel M, Rokhlina T, Karp PH, Ostedgaard LS, Uc A, Starner TD, Horswill AR, Brogden
484 KA, Prather RS, Richter SS, Shilyansky J, McCray PB Jr, *et al.* Cystic fibrosis pigs develop lung disease and
485 exhibit defective bacterial eradication at birth. *Sci Transl Med* 2010;2:29ra31.
- 486 35. Cardenas PA, Cooper PJ, Cox MJ, Chico M, Arias C, Moffatt MF, Cookson WO. Upper airways microbiota in
487 antibiotic-naïve wheezing and healthy infants from the tropics of rural Ecuador. *PLoS ONE* 2012;7:e46803.
- 488 36. Barker AF, Couch L, Fiel SB, Gotfried MH, Ilowite J, Meyer KC, O'Donnell A, Sahn SA, Smith LJ, Stewart
489 JO, Abuan T, Tully H, Van Dalnsen J, Wells CD, Quan J. Tobramycin solution for inhalation reduces sputum
490 *Pseudomonas aeruginosa* density in bronchiectasis. *Am J Respir Crit Care Med* 2000;162:481–485.

491

492

493

494 **Figure legends**

495 **Fig. 1.** Box plot comparisons of bacterial diversity between the BE, PBB, and CF cohorts. Given are three
496 measures of diversity: species richness (S^*), Simpson's complement index ($1-D$), and Shannon-Wiener index
497 (H'). The top and bottom boundaries of each box indicate the 75th and 25th quartile values, respectively, and lines
498 within each box represent the 50th quartile (median) values. Ends of whiskers mark the lowest and highest
499 diversity values in each instance. One-way ANOVA summary statistics are given in Table S2. The only
500 significant difference found was for H' between the CF and PBB cohorts ($P=0.027$, Table S2).

501

502 **Fig. 2.** Dendrograms of bacterial community membership for all patients from the BE, PBB, and CF cohorts.
503 Black, grey, and white shaded boxes are given to show positions of BE, PBB, and CF samples, respectively.
504 Patient species profiles were compared using the Sørensen index of similarity and un-weighted pair-group
505 averages (UPGMA).

506

507 **Fig. 3.** Distribution and dispersal of bacterial species among BE, PBB, and CF pediatric cohorts. (A) The number
508 of samples for which each detected bacterial species (open circles) was observed, plotted against the
509 abundance (\log_{10} scale) of that species among all samples within each cohort (BE, $r^2 = 0.64$, $F_{1,175} = 311.5$, $P <$
510 0.0001 ; PBB, $r^2 = 0.72$, $F_{1,128} = 321.1$, $P < 0.0001$; and CF, $r^2 = 0.75$, $F_{1,141} = 418.4$, $P < 0.0001$). (B) A dispersal
511 plot to identify which bacterial species are randomly distributed within each cohort, a measure used to assign
512 core versus satellite status. Index of dispersion was calculated as the ratio of variance to mean of abundance for
513 each species within each cohort and plotted for each sample. The line depicts the 2.5 % confidence limit for the
514 χ^2 distribution. Species that fall below this line are randomly distributed and were considered satellite species,
515 whereas those that are above the line are non-randomly distributed and were considered core species. The 97.5
516 % confidence limit was not plotted, as no species fell below that line.

517

518 **Fig. 4.** Dendrograms of community membership similarity between the pediatric BE, PBB, and CF bacterial
519 metacommunities and compared with adult CF and BE metacommunities. Given are whole, core and satellite
520 microbiota. Metacommunity profiles were compared using the Sørensen index of similarity and UPGMA.
521 Similarities between the microbiota from different cohorts are read as the location of the horizontal line

522 connecting those cohorts (the “node”) on the Y-axis; for example, the similarity between pediatric core microbiota
523 was $\geq 90\%$, while that between adult core microbiota was $\geq 40\%$.

524

525 **Fig. 5.** Dendrograms of Raup and Crick (S_{RC}) probability-based index of similarity between the pediatric BE,
526 PBB, and CF bacterial metacommunities and compared with adult CF and BE metacommunities. Given are the
527 whole, core and satellite microbiota. $S_{RC} < 0.95$ and $S_{RC} > 0.05$ denote similarity no greater than expected by
528 chance. $S_{RC} < 0.05$ denotes significant dissimilarity and $S_{RC} > 0.95$ significant similarity. The 0.05 and 0.95
529 thresholds are depicted with dashed lines in each instance. Dendrograms were constructed using UPGMA.

530

531

532 **Tables:**

533 **Table 1.** Demographic characteristics of the pediatric study groups.

534

Clinical Characteristic	BE ¹	PBB	CF
Number of subjects	19	12	25
Geographic origin	Australia (Brisbane)	Australia (Brisbane)	US (Seattle)
Sample type	9 BAL & 10 Sputum	All BAL	All Sputum
Age (y; Mean ± SD, range)	8.9± 4.7, 1.8-16.3 ⁶	2.3 ± 1.7, 0.6-5.5 ^{4,5}	12.5±3.5, 2.3-17.7
Gender (% female)	36%	25%	68%
FEV ₁ % predicted (Mean ± SD, range, no. subjects with measurements)	85.7 ± 21%, 41-120%, 15 subjects	117 ± 1.4%, 116 and 118%, 2 subjects ^{4,5}	77.9 ± 24%, 22.2-116.2%, 21 subjects
Number (%) of subjects on antibiotics at sampling ²	8 of 19 (42%)	1 of 12 (8%)	7 of 25 (28%)
Number (%) of subjects on corticosteroids at sampling	6 of 19 (32%), one oral, 5 inhaled	2 of 12 (16%), all inhaled	11 of 25 (48%), all inhaled
Number (%) culture-positive for <i>P. aeruginosa</i> ³	0	0	8 (34%)
Number (%) culture-positive for <i>S. aureus</i> ³	2 (11%)	2 (16%)	22 (96%)
Number (%) culture-positive for <i>H. influenzae</i> ³	4 (19%)	6 (50%)	7 (30%)

535

536

¹ Of the 10 children with BE with additional diagnoses available, 5/10 (50%) had idiopathic bronchiectasis, one

537

had previously had an airway foreign body, one had primary ciliary dyskinesia, one had chronic aspiration due to

538

a tracheoesophageal fistula, one had Mounier-Kuhn syndrome, and one had common variable

539

immunodeficiency.

540

² Antibiotics for children with CF included tobramycin (inhaled), linezolid, levofloxacin, ciprofloxacin, trimethoprim-

541

sulfamethoxazole, ticarcillin-clavulanate, and amoxicillin-clavulanate. Antibiotics for children with PBB included

542

only erythromycin. Antibiotics for children with BE included erythromycin, azithromycin, roxithromycin,

543

clarithromycin, amoxicillin-clavulanate, cefotaxime, ceftriaxone, and ticarcillin-clavulanate.

544

³ Of 23 subjects with CF with culture data available; all subjects with PBB and BE had available culture data.

545

⁴ $P < 0.001$ compared with CF; ⁵ $P < 0.001$ compared with BE; ⁶ $P < 0.01$ compared with CF.

546

547 **Table 2.** Similarity of Percentages (SIMPER) analysis of bacterial community similarity (Bray-Curtis) between
 548 pediatric metacomunities. Given is occupancy, or the number of patients for each disease group a given
 549 species was detected in. Next is mean % abundance of sequences for a species across the samples it was
 550 observed to occupy. Mean contribution represents the average contribution of a given species to the average
 551 similarity between samples (overall mean = 50.2%). Percentage contribution is the mean contribution divided by
 552 mean similarity across samples. The list of species is not exhaustive, so cumulative % value does not sum to
 553 100%. Species level identities of detected taxa are reported here. However, given the length of the ribosomal
 554 sequences analysed, these identities should be considered putative.

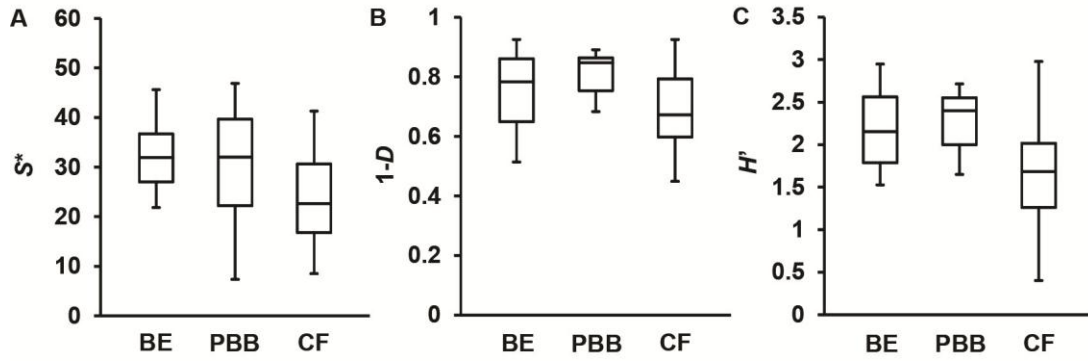
555

Species	Occupancy ¹			Mean Abund.	Mean Cont.	% Cont.	Cum. %
	BE	PBB	CF				
<i>Haemophilus influenzae</i>	12	11	19	23.22	18.79	37.44	37.44
<i>Streptococcus mitis</i>	18	12	25	14.59	8.45	16.84	54.28
<i>Prevotella melaninogenica</i>	16	11	24	9.16	4.59	9.15	63.43
<i>Veillonella dispar</i>	19	12	24	4.48	3.90	7.77	71.20
<i>Fusobacterium nucleatum</i>	17	11	21	2.63	1.99	3.97	75.17
<i>Neisseria flavescens</i>	12	11	19	5.04	1.73	3.44	78.61
<i>Porphyromonas catoniae</i>	17	11	18	1.66	1.05	2.09	80.70
<i>Haemophilus parainfluenzae</i>	15	10	16	1.30	1.04	2.06	82.76
<i>Porphyromonas gingivalis</i>	17	9	17	1.60	0.97	1.92	84.69
<i>Prevotella nanceiensis</i>	13	11	15	0.85	0.54	1.08	85.77
<i>Prevotella histicola</i>	12	7	19	0.81	0.46	0.92	86.69
<i>Granulicatella adiacens</i>	15	10	19	0.53	0.44	0.87	87.57
<i>Prevotella oris</i>	10	4	13	1.07	0.37	0.73	88.30
<i>Sphingomonas echinoides</i>	9	7	3	0.73	0.36	0.71	89.02
<i>Gemella morbillorum</i>	13	11	21	0.86	0.35	0.70	89.71
<i>Prevotella pallens</i>	16	8	13	0.63	0.32	0.63	90.35

556 ¹Total N for each pediatric disease cohort: BE, 19; PBB, 12; CF, 25.

557

558 Figure 1



559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

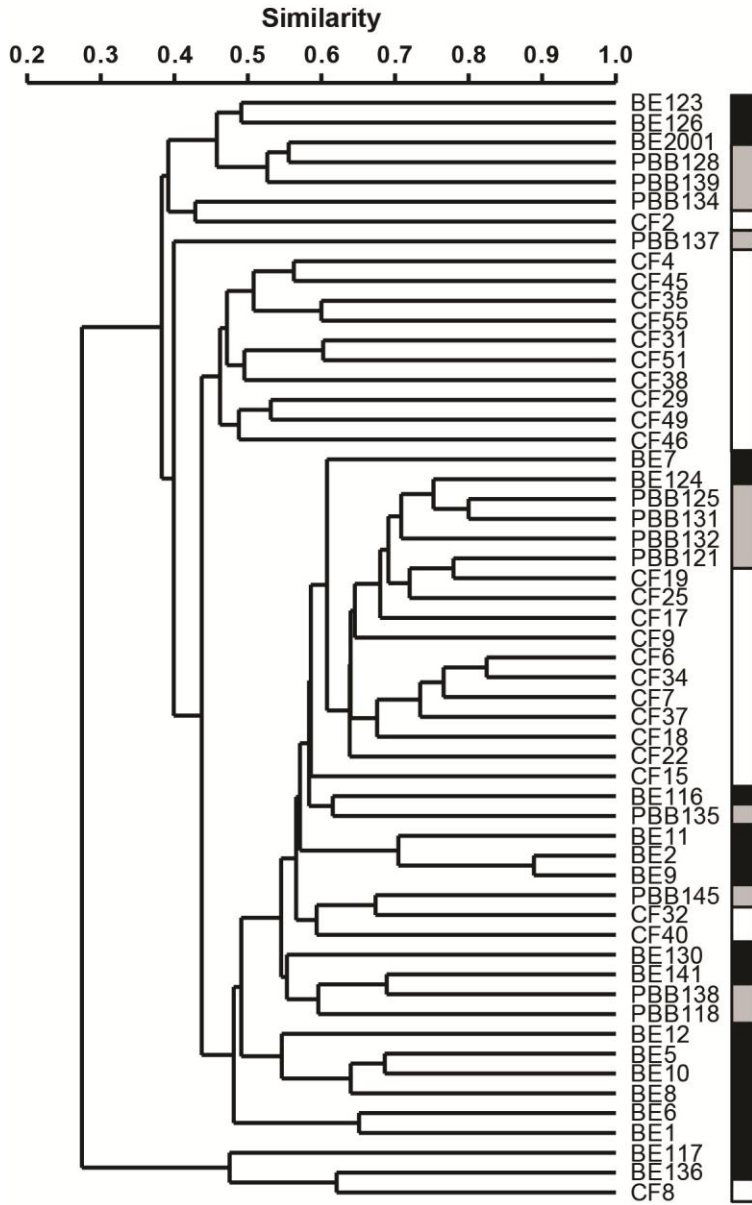
578

579

580

581

582 Figure 2



583

584

585

586

587

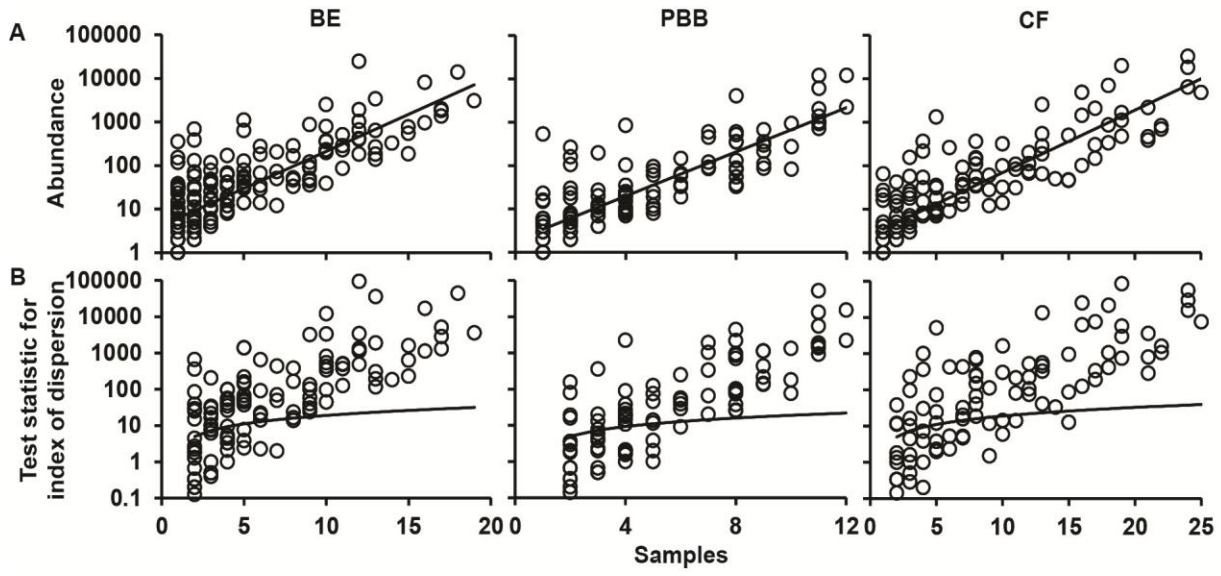
588

589

590

591

592 Figure 3



593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

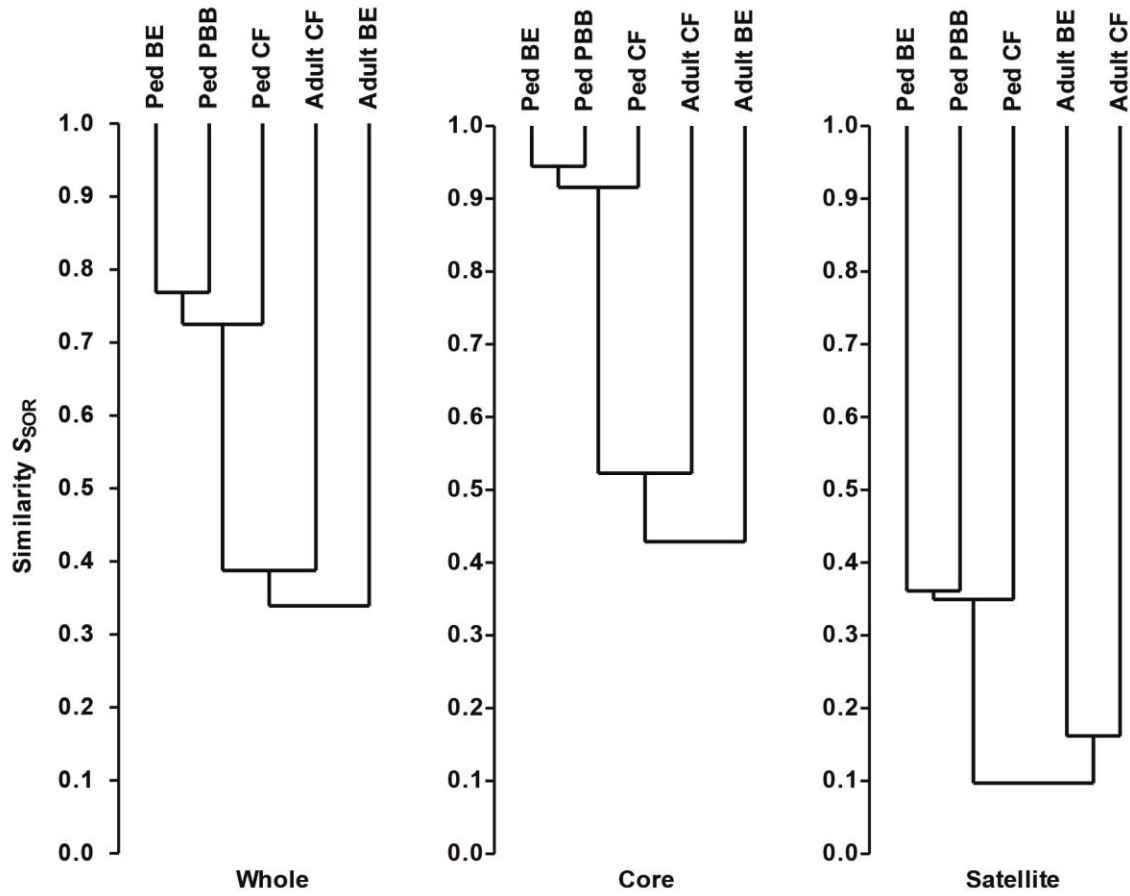
609

610

611

612

613 Figure 4



614

615

616

617

618

619

620

621

622

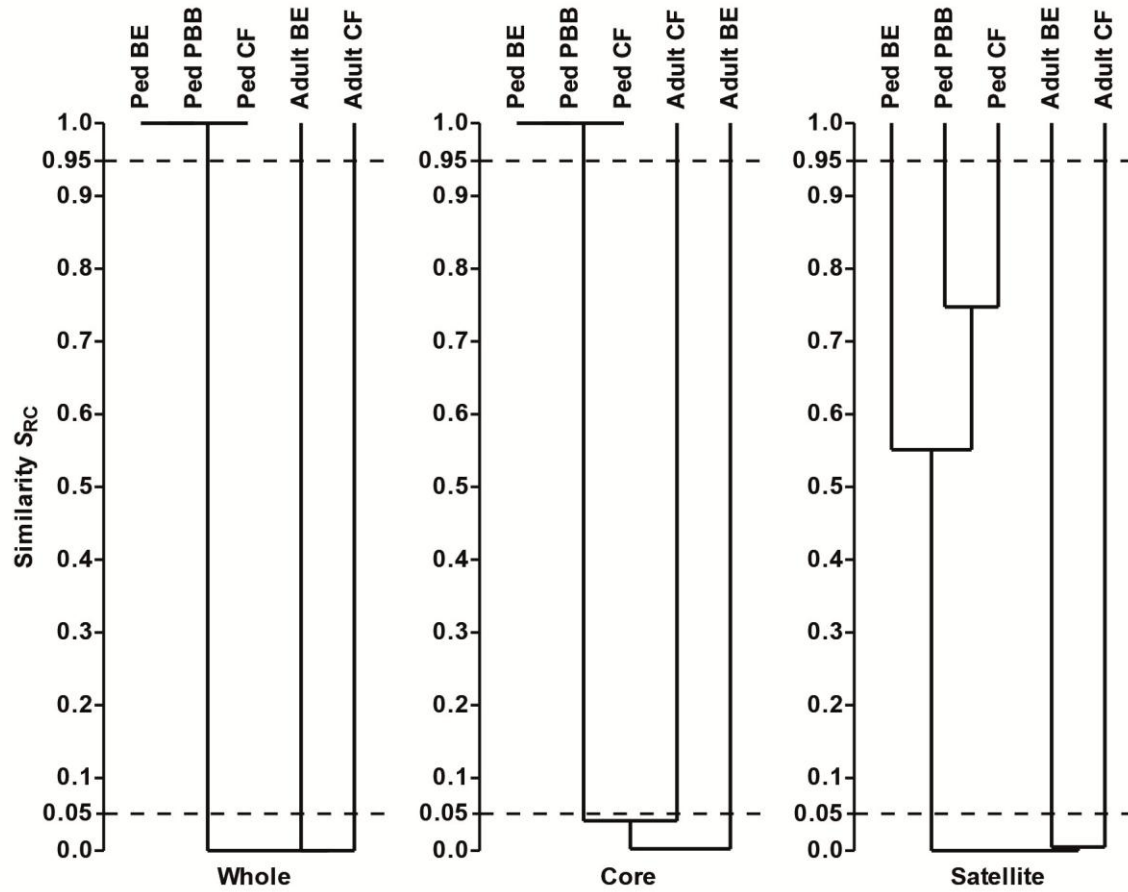
623

624

625

626

627 Figure 5



628

629