**The core microbiota of non-CF bronchiectasis airways - Online Data Supplement**

Geraint B Rogers, Christopher J van der Gast, Leah Cuthbertson, Serena K Thomson, Kenneth D Bruce, Megan L Martin, David J Serisier

**In addition to their description below, details of recruitment and inclusion are provided as part of a separate publication, [1].**

**Non-CF bronchiectasis subject inclusion criteria:**

1. Able to provide written informed consent.

2. Confirmed diagnosis of bronchiectasis by HRCT within 3 years.

3. Airways obstruction on spirometry (ratio $FEV_1$/ FVC <0.7) and $FEV_1$ ≥25% predicted.

4. Chronic productive cough with at least 5 mLs sputum production per day.

5. At least two exacerbations of bronchiectasis requiring either oral or intravenous supplemental antibiotic therapy (of at least 7 days on each occasion) in the prior 12 months.

6. Aged 20-85 inclusive.

7. Clinically stable for at least four weeks (defined as no symptoms of exacerbation, no requirement for supplemental antibiotic therapy, and FEV1 within 10% of best recently recorded value where available).

**Exclusion criteria**

1. Bronchiectasis as a result of CF or focal endobronchial obstruction.

2. Currently active tuberculosis or non-tuberculous mycobacterial (NTM) infection. Subjects with evidence of prior pulmonary NTM infection could be included only if they have completed a course of therapy that is deemed successful on the basis of negative NTM cultures following cessation of therapy. All subjects required a negative NTM culture prior to screening.

3. Any symptoms or signs to suggest recent deterioration in respiratory disease, including exacerbation of pulmonary disease (as previously defined) in the preceding 4 weeks.

4. Any change to medications in the preceding 4 weeks.

5. Prescription of either oral or intravenous antibiotic therapy in the preceding 4 weeks.

6. Cigarette smoking within the preceding 6 months.

7. Any history of malignant arrhythmia (unless in the immediate post-myocardial infarction period and not requiring any regular therapy) or QTc prolongation on baseline ECG.

8. Any of the following within the three (3) months prior to enrolment:

- Acute MI
- Acute CVA
- Major surgery

9. History of any of the following:

- Active malignancy (excepting non-melanoma skin malignancies that have been treated and considered cured)
- Listed for transplantation
- Any other significant active illness likely to affect the patient's survival within 12 months
- Receiving long-term domiciliary oxygen therapy

10. Allergy to macrolide antibiotics, other than minor, dose-related gastrointestinal intolerance that would not be anticipated to recur with low-dose erythromycin.

11. Any prescription or receipt of long-term macrolide antibiotics, or receipt of a treatment course within 4 weeks.

12. Predominant diagnosis of emphysema (rather than bronchiectasis) on HRCT scan of the chest.

13. Requirement for supplemental oxygen therapy.

14. Inability to complete required study procedures for whatever reason (including 6 minute walk test, hypertonic saline sputum induction).

15. Respiratory symptoms (including cough, sputum production, recurrent exacerbations) not predominantly the result of bronchiectasis in the opinion of the PI; where treatable causes for exacerbations existed, these were treated before considering trial enrolment.

**Excluded medications**

1. Macrolide antibiotics – long-term macrolide use was an absolute exclusion, however subjects who had received a short duration (less than 6 weeks) treatment course were eligible provided they had at least 4 weeks washout.
2. Long term oral antibiotic administration for infection prophylaxis (eg doxycycline).
3. Any other intravenous or oral antibiotic within 4 weeks.
4. While erythromycin in the current study was administered in a low dose, possible drug interactions in all patients entering the study were considered. Subjects using the following medications were not eligible for the study:
    - ergotamine or dihydroergotamine
    - triazolam/ alprazolam
    - sildenafil
    - azole antifungals (ketoconazole, itraconazole, fluconazole)
    - disopyramide
    - quinidine

Concomitant HMG-CoA reductase inhibitor ('statin') use was permitted, however subjects on high-dose statins (equivalent to 80 mgs daily of simvastatin) required dose reduction by half at study entry.

Subjects prescribed diltiazem or verapamil were screened to ensure no evidence of clinically relevant increases in levels of these medications.

The following medications were permitted, provided they had been a regular medication for at least 6 months (with the requisite number of exacerbations whilst on this therapy):

1. Inhaled antibiotics in chronic, daily, stable dose.
2. Inhaled mucolytic therapies (hypertonic saline, mannitol, dornase alpha, N-acetylcysteine).

3. Oral mucolytics or expectorants.

**Additional inclusion/ exclusion criteria for bronchoscopy subjects (with non-CF bronchiectasis):**

1. Never smokers.
2. No hypogammaglobulinaemia (levels performed at screening) or ciliary dysfunction.
3. Not asthmatic and without evidence of atopy on bloods (RAST/ s.IgE).
4. FEV1>40% predicted.
5. Not warfarinised or have other issues in relation to anticoagulation or bleeding diatheses that make them unsuitable for endobronchial biopsy due to bleeding risks.
6. No clinically significant ischaemic or myocardial disease that increases the risk in relation to bronchoscopy.
7. No hypercapnoea or evidence of overlap syndrome.

**Assignment of aetiology of non-CF bronchiectasis:**

Aetiology was assigned based on a careful consideration of patient history by the BLESS PI (DJS), who is also the CF centre director at site. This included review of all prior investigations performed (eg ciliary studies, CF testing, etc). Routine CF screening was not performed, however in subjects with phenotypic features raising the possibility of this diagnosis (eg <40 yo, malabsorption, male infertility, underweight, family history of bronchiectasis, coexistent diabetes mellitus, predominant upper lobe bronchiectasis), sweat testing +/- CFTR analysis was performed. At enrolment, all subjects also had testing for atopy (including s.IgE/ RAST testing, RAST aspergillus and aspergillus precipitins) and immunoglobulin levels. These tests were not used to assign aetiology, but rather to ensure that subjects who had any evidence of asthma or immunoglobulin deficiency were excluded from the current (bronchoscopy) substudy.

**Inclusion/ exclusion criteria for the normal control subjects:**

1. Age 18 – 85 years.

2. Normal subjects without any identifiable airway disease.

3. Capable of providing written, informed consent to participate.

4. Lifelong non-smokers (< or = 2 pack year history of smoking).

5. Normal spirometry (defined as values for $FEV_1$, FVC and ratio of $FEV_1$:FVC all lying within the normal predicted range according to age and height, no significant acute response to inhaled bronchodilator and no evidence of small airways obstruction on flow-volume loops).

Exclusion criteria:

1. Smoking history >2 pack year history of tobacco smoking.
2. Any smoking history of other substances.
3. History of any chronic respiratory disease including asthma, chronic obstructive pulmonary disease/ COPD, bronchiectasis, etc, or pneumonia within 8 weeks of bronchoscopy.
4. Any respiratory tract infection (including upper respiratory tract infection) within 4 weeks of bronchoscopy.
5. Medications: prescription of any antibiotic, antihistamine, corticosteroid, mast cell stabilising therapy, theophylline or any inhaled therapy within 4 weeks of bronchoscopy.
6. Any condition with the potential to increase the risks of bronchoscopy/ bronchoalveolar lavage or endobronchial biopsy, including (but not limited to) bleeding diathesis (platelet count < 150, abnormal coagulation profile, warfarin therapy), any unstable medical condition (eg acute cardiac or cerebral ischaemic event within 3 months, active malignancy), any significant illness likely to impact upon survival of the patient within 12 months, requirement for domiciliary oxygen.

7. Positive bronchial challenge test or chronic abnormalities detected on CT scan of the chest.

### Leicester Cough Questionnaire

The Leicester Cough Questionnaire (LCQ) is a valid, repeatable 19 item self-completed quality of life measure of chronic cough which is responsive to change. LCQ was employed as described previously. [2]

### St George's Respiratory Questionnaire

The St. George's Respiratory Questionnaire (SGRQ) is a disease-specific measure used to assess patients with mild to severe airway disease. This measure is a disease-specific instrument designed to measure impact on overall health, daily life, and perceived well-being. It was developed for use by patients with fixed and reversible airway obstruction. The measure consists of 50 (76 responses) items that produce three domain scores and one overall score measuring: Symptom (frequency and severity); Activity (activities that cause or are limited by breathlessness); and Impacts (social functioning, psychological disturbances resulting from airways disease. The SGRQ has been validated and approved as a self-administered instrument, and was applied here as described previously. [3]

### CRP measurement

Subjects were asked to provide a blood sample to be analysed at Mater Pathology laboratory. CRP measurement was performed using an Architect ci16200 latex-enhanced immunoturbidimetric assay (Abbott Diagnostics, Abbott Park, IL).

## Procedures

*Sputum microbiology*

Sputum processing for culture and sensitivity testing was performed in the Division of Microbiology, Mater Pathology. Sputum was transported to the laboratory within 60 min of collection and processed within 3 h (refrigerated at 4°C in the interim). The most purulent portion of the specimen was selected and streaked directly onto horse blood agar (HBA), MacConkey agar, chocolate agar supplemented with bacitracin (CHOC-B), Sabouraud agar, and mannitol salt agar. Plates were incubated at 37°C (HBA and CHOC-B in $CO_2$ and CHOC-B anaerobically and the rest in $O_2$) for at least 48 h and examined daily. Organism identification incorporated a combination of typical morphology, species specific manual tests (eg. catalase testing, coagulase testing, Gram stain, pigment production etc.), API, and Vitek identification.

*Sputum induction procedure*

Subjects were instructed to perform their usual chest physiotherapy regime on the morning of the sputum induction procedure. Prior to commencement of hypertonic saline inhalation, any spontaneous sputum expectorated was collected for standard culture. Sputum induction (SI) was performed after inhalation of 400 ug of albuterol, using 4.5% hypertonic saline nebulised from an ultrasonic nebuliser (output >1 mL/ min) for 20 minutes in 4 periods of 5 minutes each, according to the standardised protocol recommended by the European Respiratory Society taskforce [4] Following mouth-rinsing and expectoration, sputum was collected following each nebulisation period, on each occasion preceded immediately by spirometry. The first sputum sample was refrigerated immediately following collection and frozen at -80 °C within an hour. A cold chain was maintained up until the point of DNA extraction.

Ten, 15 and 20 minute samples were pooled and an aliquot from this placed on ice immediately and transferred for inflammatory cell count processing within 60 minutes. Sputum was processed according to the methods of the US Cystic Fibrosis Therapeutics Development Network Standard

Operating Procedure.[5] Briefly, an equal volume of sterile 10% dithiothreitol (DTT) (Sputolysin; Calbiochem-Novabiochem Corp., San Diego, CA), was added to the sputum, then incubated in a shaking water bath at 37° C for 5-10 min, and mixed using a transfer pipette at 5-min intervals. A further three times the volume of both DTT and phosphate-buffered saline (Dulbecco's; Gibco BRL, Grand Island, NY) was added and the mixture incubated again in the 37° C shaking water bath for another 5-10 min. 10 μl of homogenized sputum samples, mixed with Trypan Blue, was used to calculate total cell counts using a standard hemacytometer. A further 0.25-0.50 ml of both samples was used to prepare cytospin slides for differential cell counts. After staining the slides with Wright's stain, 300 non-squamous cells were counted and cell differentials calculated.

*Bronchoscopy procedure*

Subjects were fully informed about the potential risks of the procedure and provide written consent. Bronchoscopy was performed as an outpatient procedure in the endoscopy unit of the operating theatres of the Mater Adult Hospital, using an Olympus flexible fibre-optic bronchoscope according to the safety standards of the Thoracic Society of Australia and New Zealand,[6] with details of the research bronchoscopy procedure adapted from prior methods.[7-9] Subjects fasted for 6 hours before the procedure. The procedure was performed transorally, under light sedation using intravenous midazolam and fentanyl to ensure patient comfort. Topical lignocaine was applied to the vocal cords and bronchi by instillation through the bronchoscope. Where possible, no suction was employed prior to performing BAL in the target lobe. After wedging in a right middle lobe bronchus, BAL was performed. In non-CF bronchiectasis subjects without significant right middle lobe bronchiectasis (on HRCT), an alternative lobe with bronchiectatic change was selected according to a hierarchy –lingula, upper lobe (right or left), apical segment lower lobe.

One hundred and forty (140) mLs of warmed sterile saline was gently instilled, followed by gentle aspiration through the bronchoscope's suction channel.

Following bronchoalveolar lavage, all subjects also had endobronchial biopsies collected before completion of the procedure, although this does not form part of the current report. Subjects were observed for 2 hours after the bronchoscopy before being allowed home.

Samples were refrigerated immediately following collection and frozen at -80 °C within an hour. A cold chain was maintained up until the point of DNA extraction.

*DNA extraction protocol*

Nucleic acid extractions were performed on 100 µL portions of sputum or BAL cell pellets. Guanidinium thiocyanate–EDTA–sarkosyl (500 µL) and PBS (500 µL), pH 8.0, were added to samples. Cell disruption was achieved using a Fastprep-24 Instrument (MP Biomedicals Europe, Illkirch, France) 6.5 m/s, 60 s, followed by incubation at 90 °C for 1 min and −20 °C for 5 min. Cell debris was pelleted by centrifugation at 12 000 × *g* for 2 min at 4 °C. Supernatant was transferred to a fresh microfuge tube. NaCl (to a final concentration of 0.5 mol/L and polyethylene glycol (to a final concentration of 15%) were added and DNA precipitated at 4 °C for 30 min. DNA was pelleted by centrifugation at 12 000 × *g* for 2min at 4 °C and resuspended in 300 µL of sterile distilled water. Samples were heated at 90 °C for 30 s and vortexed. Phenol/chloroform (1:1) (300 µL) was added, and samples were vortexed for 20 s before centrifugation at 12 000 × *g* at 4 °C for 3min. The upper phase was then transferred to a fresh microfuge tube. Total DNA was then precipitated by the addition of an equal volume of isopropanol, a 0.1-volume 10 mol/L ammonium acetate, and 1 µL of GenElute linear polyacrylamide (Sigma-Aldrich, Gillingham, UK) and incubated at −20 °C for 25 min. DNA was pelleted by centrifugation at 12 000 × g at 4 °C for 5 min. Pelleted DNA was then washed 3 times in 70% ethanol, dried, and resuspended in 50 µL of sterile distilled water. DNA extracts were quantified using the Picodrop Microlitre Spectrophotometer (GRI, Braintree, UK).

*Pyrosequencing*

Bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP) was performed as described previously using Gray28F 5'-TTTGATCNTGGCTCAG-3' and Gray519r 5'-GTNTTACNGCGGCKGCTG-3'). Initial generation of the sequencing library involved a one-step PCR of 30 cycles, using a mixture of Hot Start and HotStar high fidelity Taq DNA polymerase, as described previously.[10] Tag-encoded FLX amplicon pyrosequencing analyses utilized Roche 454 FLX instrument with Titanium reagents, titanium procedures performed at the Research and Testing Laboratory (Lubbock, TX) using RTL protocols (www.researchandtesting.com).

*Sequence processing pipeline*

The following information is as described in protocol documentation provided by Molecular Research DNA, Texas, USA (www.mrdnalab.com). Custom software written in C# within a Microsoft*H*.NET (Microsoft Corp, Seattle, WA) development environment was used for all post sequencing processing. Quality trimmed sequences obtained from the FLX sequencing run were derived directly from FLX sequencing run output files. Tags were extracted from the multi-FASTA file into individual sample-specific files based upon the tag sequence. Tags which did not have 100% homology to the sample designation were not considered. Sequences which were less than 150 bp after quality trimming were not considered. All failed sequence reads, low quality sequence ends and tags and primers were removed. Sequences with ambiguous base calls, sequences with homopolymers > 6bp were removed. Further, any non-bacterial ribosomal sequences and chimeras using B2C2 [11] as described previously.[10]

Individual samples were assembled using CAP3 after parsing the tags into individual FASTA files.[12] The ace files generated by CAP3 were then processed to generate a secondary FASTA

file containing the tentative consensus (TC) sequences of the assembly along with the number of reads integrated into each consensus. TC were required to have at least 2-fold coverage.

To determine the identity of bacterial species in the remaining sequences, sequences were de-noised, assembled into OUT clusters at 97% identity, and queried using a distributed .N*Et al*gorithm that utilizes Blastn+ (KrakenBLAST www.krakenblast.com) against a database of high quality 16S rRNA gene bacterial sequences. Using a .NET and C# analysis pipeline the resulting BLASTn+ outputs were compiled, data reduction analysis performed, and sequence identity classification carried out, as described previously.[10]

Based on the described protocols, identification at the species level should only be considered putative.

| Patient | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 73 | 56 | 65 | 62 | 62 | 67 | 56 | 53 | 58 | 63 | 69 | 48 | 57 | 65 | 62 | 60 | 37 | 68 | 69 | 71 | 67 |
| Gender | F | F | F | F | F | F | M | F | M | F | F | F | M | M | M | M | F | F | M | M | F |
| FEV1 % pre-BD | 66.5 | 71.9 | 49.8 | 76.1 | 94.6 | 83.9 | 70.3 | 73.9 | 76.2 | 54.8 | 100.0 | 75.3 | 34.9 | 102.0 | 69.0 | 76.9 | 101.8 | 54.4 | 61.0 | 72.2 | 67.2 |
| FEV1% post-BD | 75.2 | 69.3 | 62.3 | 81.0 | 98.1 | 80.5 | 75.6 | 72.3 | 81.1 | 66.1 | 94.6 | 77.9 | 40.9 | 101.4 | 74.4 | 88.0 | 103.2 | 65.2 | 59.3 | 72.6 | 66.2 |
| Sputum weight (g) | 15 | 24 | 12 | 5 | 12 | 5 | 11 | 3 | 12 | 18 | 32 | 1.3 | 44.9 | 7 | 0.7 | 23.7 | 6.5 | 22.8 | 17.8 | 21.9 | 23.6 |
| Leicester cough score | 10.8 | 11.9 | 16.6 | 13.7 | 8.3 | 12.3 | 16.8 | 10.3 | 19.7 | 18.1 | 9.3 | 18.8 | 19.2 | 18.8 | 19.8 | 13.7 | 15.4 | 19.2 | 16.5 | 8.8 | 17.7 |
| SGRQtotal | 62.8 | 46.2 | 34.8 | 46.4 | 63.9 | 25.3 | 22.3 | 59.7 | 20.3 | 31.9 | 29.2 | 24.1 | 22.0 | 30.4 | 18.6 | 30.4 | 20.7 | 19.2 | 53.7 | 53.3 | 20.4 |
| SGRQ symptoms | 68.7 | 77.0 | 74.8 | 63.4 | 92.8 | 53.8 | 60.7 | 79.5 | 26.1 | 58.6 | 57.3 | 23.8 | 54.0 | 49.9 | 36.2 | 39.6 | 39.6 | 33.0 | 73.4 | 90.5 | 41.9 |
| CRP | 0 | 2.4 | 15 | 8.8 | 3.5 | 0 | 3.1 | 3.9 |  | 0.8 | 4.2 |  | 10 | 0 | 1.4 | 0 | 0.8 | 8.7 | 0.8 | 10 | 5.6 |
| Neuts - induced | 236 | 403 | 479 | 590 | 472 | 314 | 446 | 515 | 177 | 401 | 286 | 476 | 86 | 515 | 442 | 396 |  | 390 | 408 | 254 | 410 |
| Neuts % - induced | 83.4 | 96.9 | 91.4 | 98.5 | 89.2 | 73.4 | 91.2 | 96.8 | 39.2 | 89.7 | 69.9 | 91.7 | 97.7 | 82.9 | 87.4 | 84.3 |  | 93.8 | 76.0 | 58.3 | 85.4 |
| Aetiology - idiopathic | ✓ | ✓ |  |  | ✓ | ✓ | ✓ | ✓ |  | ✓ | ✓ | ✓ |  |  | ✓ | ✓ | ✓ | ✓ |  | ✓ |  |
| - post infective |  |  | ✓ | ✓ |  |  |  |  | ✓ |  |  |  | ✓ |  |  |  |  |  | ✓ |  | ✓ |
| - pink disease |  |  |  |  |  |  |  |  |  |  |  |  |  | ✓ |  |  |  |  |  |  |  |
| - 'burnt-out' ABPA |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *_P. aeruginosa_ (standard culture) |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | ✓ |  |  |  |  |  |
| *_H. influenzae_ |  |  |  |  |  |  |  |  |  | ✓ |  |  |  |  | ✓ |  | ✓ | ✓ | ✓ | ✓ | ✓ |
| 454 pyrosequencing analysis | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |  | ✓ | ✓ | ✓ | ✓ |  | ✓ | ✓ | ✓ | ✓ |

| Patient | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 64 | 65 | 67 | 64 | 63 | 62 | 73 | 65 | 57 | 59 | 67 | 67 | 66 | 74 | 57 | 67 | 65 | 66 | 61 | 60 |
| Sex | F | F | F | F | M | F | F | F | M | M | F | F | M | F | F | F | F | F | F | M |
| FEV1 % pre-BD | 59.8 | 88.4 | 91.6 | 96.5 | 74.6 | 58.2 | 52.6 | 92.8 | 65.7 | 75.2 | 72.0 | 57.4 | 93.4 | 67.6 | 57.3 | 67.8 | 73.7 | 74.4 | 72.2 | 65.3 |
| FEV1% post-BD | 57.8 | 94.2 | 89.2 | 99.6 | 80.3 | 58.7 | 59.4 | 98.6 | 78.2 | 81.6 | 72.5 | 61.5 | 104.3 | 72.6 | 69.4 | 68.3 | 76.0 | 81.0 | 74.9 | 65.0 |
| Sputum weight (g) | 37.7 | 27.1 | 6 | 15.1 | 24.1 | 14.3 | 25.4 | 0.9 | 16.4 | 17.1 | 44.9 | 22.1 | 19.8 | 20.5 | 10.3 | 15.5 | 6.2 | 3 | 18.5 | 25.1 |
| Leicester cough score | 15.6 | 8.7 | 7.7 | 15.2 | 12.2 | 17.4 | 13.5 | 20.3 | 14.8 | 15.4 | 9.9 | 5.5 | 7.5 | 13.9 | 10.5 | 15.7 | 18.7 | 17.8 | 15.7 | 11.2 |
| SGRQtotal | 45.8 | 64.4 | 55.4 | 25.8 | 43.7 | 31.5 | 58.9 | 23.8 | 53.3 | 31.0 | 49.3 | 79.6 | 48.5 | 59.6 | 52.6 | 19.6 | 31.8 | 21.0 | 31.6 | 51.9 |
| SGRQ symptoms | 32.7 | 77.1 | 100.0 | 23.8 | 93.5 | 55.9 | 57.9 | 42.2 | 74.3 | 66.7 | 48.5 | 64.0 | 79.1 | 85.7 | 85.8 | 26.1 | 58.3 | 42.0 | 68.3 | 75.4 |
| CRP | 18 | 1.5 | 9.4 | 1.4 | 5 | 7.6 | 14 | 1.5 | 0 | 1.6 |  | 6.5 | 0.6 | 2.3 | 0.6 | 0 | 8.2 | 13 | 5.3 | 2.8 |
| Neuts - induced | 402 | 56 |  | 399 | 399 |  | 498 | 331 | 231 | 441 | 439 | 611 | 477 | 437 | 373 | 319 |  | 246 | 375 | 432 |
| Neuts % - induced | 92.4 | 41.8 |  | 93.2 | 90.1 |  | 92.2 | 77.2 | 53.1 | 95.7 | 91.1 | 95.8 | 90.0 | 91.0 | 93.0 | 74.7 |  | 60.6 | 93.5 | 94.7 |
| Aetiology - idiopathic | ✓ | ✓ |  | ✓ | ✓ |  | ✓ | ✓ | ✓ |  |  | ✓ | ✓ |  | ✓ | ✓ |  |  |  | ✓ |
| - post infective |  |  | ✓ |  |  | ✓ |  |  |  |  |  |  |  | ✓ |  |  | ✓ |  |  |  |
| - pink disease |  |  |  |  |  |  |  |  |  | ✓ |  |  |  |  |  |  |  | ✓ | ✓ |  |
| - 'burnt-out' ABPA |  |  |  |  |  |  |  |  |  |  | ✓ |  |  |  |  |  |  |  |  |  |
| **P. aeruginosa* (standard culture) |  |  |  |  |  |  |  |  |  |  | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| **H. influenzae* | ✓ |  | ✓ |  |  |  | ✓ |  |  | ✓ |  |  |  |  |  |  | ✓ |  |  |  |
| 454 pyrosequencing analysis | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |  | ✓ | ✓ | ✓ |

**Table S1.** Patient data. $FEV_1$% – forced expiratory volume in 1 second as a percentage of the predicted value; SGRQ – St George's Respiratory Questionnaire – range 1-100, lower scores indicate better quality of life, symptoms component measures the effect of respiratory symptoms; CRP – C reactive protein, sputum % neutrophils – neutrophils in induced sputum as a percent of total non-squamous cells; ICS – inhaled corticosteroid; LABA – long-acting β-agonist, BD – bronchodilator.

| **Patient** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Induced sputum*** | | | | | | | | | | | | | | | | | | | | |
| *P. aeruginosa* | | | | | | | | | | | | | | | | ✓ | | | | |
| *H. influenzae* | | | | | | | | | | | | | | | ✓ | | | ✓ | ✓ | ✓ |
| *S. aureus* | | | | | | | | | | | | | | | | | | | | |
| *P. multocida* | | | | | | | | | | | | | | | | | | | | |
| *S. pneumoniae* | | | | | | | | | | | | | | | | | | | | |
| *A. xylosoxidans* | | | | | | | | | | | | | | | | | | | | |
| *S. maltophilia* | | | | | | | | | | | | | | | | | | | | |
| *M. catarrhalis* | | | | | | | | | | | | | | | | | | | | |
| *Aspergillus spp.* | | | | | | | | | | | | | | | | | | | | |
| *Candida albicans* | | | | | | | | | | | | | | | | | | | | |
| *Penicillium spp* | | | | | | | | | | | | | | | | | | | | |
| **BAL**** | | | | | | | | | | | | | | | | | | | | |
| *P. aeruginosa* | | | | | | | | | | | ✓ | | | | | ✓ | | | | |
| *H. influenzae* | | | | | | | | | ✓ | | | | | | ✓ | | | ✓ | ✓ | ✓ |
| *S. aureus* | | | | | ✓ | | | | | | | | | | | | | | | |
| *P. multocida* | | | | ✓ | | | | | | | | | ✓ | | | | | | | |
| *S. pneumoniae* | | | | | | | | | | | | | | | | | | | | ✓ |
| *A. xylosoxidans* | | | | | | | | | | | | | | | | | | | | |
| *S. maltophilia* | | | | | | | | | | | | | | | | | | | | |
| *M. catarrhalis* | | | | | | | | | | | | | | | | | | | | |
| *Aspergillus spp.* | | | | | | | | | | | | | | | | | | | | |
| *Candida albicans* | | | | | | | | | | | | | | | | | | | | |
| *Penicillium spp* | | | | | | | | | | | | | | | | | | | | |

| Patient | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Induced sputum*** | | | | | | | | | | | | | | | | | | | | | |
| *P. aeruginosa* | | | | | | | | | | | | ✓ | ✓ | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| *H. influenzae* | ✓ | | | ✓ | | | | ✓ | | | ✓ | | | | | | | ✓ | | | |
| *S. aureus* | | | | | | | | | | | | | | | | | | | | | |
| *P. multocida* | | | | | | | | | | | | | | | | | | | | | |
| *S. pneumoniae* | | | | | | | | | | | | | | | | | | | | | |
| *A. xylosoxidans* | | | | | | | | | | | | | | ✓ | | | | | | | |
| *S. maltophilia* | | | | | | | | | | | | | | | | | | | | | |
| *M. catarrhalis* | | | | | | | | | | | | | ✓ | | | ✓ | | | | | |
| *Aspergillus spp.* | | | | | | | | | | | | | | | | | | | | | |
| *Candida albicans* | | | | | | | | | | | | | | | | | | | | | |
| *Penicillium spp* | | | | | | | | | | | | | | | | | | | | | |
| **BAL**** | | | | | | | | | | | | | | | | | | | | | |
| *P. aeruginosa* | | | | | | | | | | | | ✓ | ✓ | | | ✓ | ✓ | | ✓ | ✓ | ✓ |
| *H. influenzae* | ✓ | | | ✓ | | | | ✓ | | ✓ | ✓ | | | | | | | ✓ | | | |
| *S. aureus* | | | | | ✓ | | | ✓ | | | | | | | | | | | | | |
| *P. multocida* | | | | | | | | | | | | | | | | | | | | | |
| *S. pneumoniae* | | | | | | ✓ | | | | | | | | | | | | | | | |
| *A. xylosoxidans* | | | | | | | ✓ | | | | | | | | | | | | | | |
| *S. maltophilia* | | | | | | | | | | | | | | | | | | | | | |
| *M. catarrhalis* | | | | | | | | | | | | | ✓ | | | | | | | | |
| *Aspergillus spp.* | | | | | ✓ | | | | | | | | | | | | | | | | |
| *Candida albicans* | | | | | | | | | | | ✓ | | | | | | | | | | |
| *Penicillium spp* | | | | | | | | | | | ✓ | | | | | | | | | | |

**Table S2.** Diagnostic culture analysis of induced sputum and BAL samples. *-Samples collected directly concurrently with those analysed by pyrosequencing. **-Samples from the same sample as analysed by pyrosequencing.

| Subject | FEV1 % pre-BD | FEV1% post-BD | Age | Gender |
|---|---|---|---|---|
| 1 | 93.4 | 96.8 | 25 | F |
| 2 | 85.5 | 88.0 | 37 | F |
| 3* | 96.1 | 96.8 | 40 | M |
| 4 | 96.6 | 97.9 | 51 | M |
| 5 | 97.1 | 96.1 | 40 | M |
| 6 | 124.6 | 125.9 | 32 | F |
| 7 | 93.5 | 99.3 | 42 | F |
| 8** | 85.8 | 89.4 | 23 | F |
| 9** | 85.0 | 90.9 | 59 | F |
| 10 | 108.2 | 109.8 | 29 | F |
| 11* | 129.5 | 129.1 | 51 | F |
| 12 | 103.6 | 102.2 | 22 | M |
| 13 | 107.9 | 110.8 | 26 | M |
| 14 | 97.9 | 104.2 | 49 | M |
| 15 | 100.9 | 101.2 | 26 | F |
| 16 | 92.5 | 100.0 | 26 | M |
| **mean** | **99.9** | **102.4** | **36.1** | **F:M ratio 9:7** |
| std. dev. | 12.8 | 11.7 | 11.7 | |

**Table S3**. Control subject data. $FEV_1$% – forced expiratory volume in 1 second as a percentage of the predicted value. BD – bronchodilator.

| Class | Order | Family | Taxon name | Code |
|---|---|---|---|---|
| Actinobacteria | Actinomycetales | Actinomycetaceae | *Actinomyces odontolyticus* | C, Ae, O |
| | | | *Actinomyces oris* | C, Ae |
| | | | *Actinomyces urogenitalis* | C, Ae |
| | | Corynebacteriaceae | *Corynebacterium durum* | S, Ae, O |
| | | | *Corynebacterium freneyi* | S, Ae |
| | | | *Corynebacterium macginleyi* | S, Ae |
| | | | *Corynebacterium matruchotii* | S, Ae, O |
| | | | *Corynebacterium riegelii* | S, Ae |
| | | Geodermatophilaceae | *Blastococcus aggregatus* | S, Ae |
| | | Gordoniaceae | *Gordonia terrae* | S, Ae |
| | | Micrococcaceae | *Rothia aeria* | C, Ae, O |
| | | | *Rothia dentocariosa* | S, Ae, O |
| | | | *Rothia mucilaginosa* | C, Ae, O |
| | | Nocardioidaceae | *Propionibacterium acnes* | C, Ae, O |
| | | Bifidobacteriaceae | *Scardovia inopinata* | S, An, O |
| | Coriobacteriales | Coriobacteriaceae | *Atopobium parvulum* | C, An, O |
| | | | *Collinsella aerofaciens* | S, An |
| Bacteroidia | Bacteroidales | Bacteroidaceae | *Bacteroides acidifaciens* | C, An |
| | | | *Bacteroides coprophilus* | C, An |
| | | | *Bacteroides plebeius* | C, An |
| | | Porphyromonadaceae | *Dysgonomonas mossii* | S, Ae |
| | | | *Parabacteroides gordonii* | C, An |
| | | | *Porphyromonas endodontalis* | C, An, O |
| | | | *Porphyromonas gingivalis* | C, An, O |
| | | | *Tannerella forsythia* | C, An, O |
| | | | *Prevotella baroniae* | C, An, O |
| | | | *Prevotella copri* | C, An |
| | | | *Prevotella histicola* | C, An, O |
| | | | *Prevotella intermedia* | S, An, O |
| | | | *Prevotella melaninogenica* | C, An, O |
| | | | *Prevotella nanceiensis* | C, An |
| | | | *Prevotella nigrescens* | C, An, O |
| | | | *Prevotella oris* | C, An, O |
| | | | *Prevotella oulorum* | C, An, O |
| | | | *Prevotella pallens* | C, An, O |
| | | | *Prevotella ruminicola* | C, An |
| | | | *Prevotella tannerae* | C, An, O |
| | | | *Prevotella veroralis* | C, An, O |
| Flavobacteria | Flavobacteriales | Flavobacteriaceae | *Capnocytophaga ochracea* | S, Ae, O |
| | | | *Capnocytophaga sputigena* | C, Ae, O |
| | | | *Flavobacterium frigidarium* | C, Ae |
| | | | *Riemerella anatipestifer* | S, Ae, O |
| | | | *Zhouia amylolytica* | S, Ae |
| Bacilli | Bacillales | Bacillaceae | *Bacillus cereus* | C, Ae |
| | | Paenibacillaceae | *Aneurinibacillus thermoaerophilus* | S, Ae |
| | | Staphylococcaceae | *Staphylococcus aureus* | C, Ae, O, P |
| | | | *Staphylococcus epidermidis* | C, Ae, O |
| | Gemellales | Gemellaceae | *Gemella sanguinis* | C, Ae, O |
| | Lactobacillales | Aerococcaceae | *Abiotrophia defectiva* | S, Ae, O |
| | | Carnobacteriaceae | *Granulicatella adiacens* | C, Ae, O |
| | | | *Granulicatella elegans* | S, Ae, O |
| | | Lactobacillaceae | *Lactobacillus iners* | S, Ae, O |
| | | Streptococcaceae | *Streptococcus anginosus* | S, Ae, O, P |

*Table S4 Continued*

| Class | Order | Family | Taxon name | Code |
|---|---|---|---|---|
| | | | *Streptococcus constellatus* | S, Ae, O, P |
| | | | *Streptococcus cristatus* | C, Ae, O |
| | | | *Streptococcus gordonii* | C, Ae, O |
| | | | *Streptococcus infantis* | C, Ae, O |
| | | | *Streptococcus mutans* | S, Ae, O |
| | | | *Streptococcus oralis* | C, Ae, O |
| | | | *Streptococcus parasanguinis* | C, Ae, O |
| | | | *Streptococcus pneumoniae* | C, Ae, O, P |
| | | | *Streptococcus pseudopneumoniae* | C, Ae |
| | | | *Streptococcus salivarius* | C, Ae, O |
| | | | *Streptococcus sanguinis* | C, Ae, O |
| Clostridia | Clostridiales | Clostridiaceae | *Finegoldia magna* | S, An, O |
| | | | *Peptoniphilus asaccharolyticus* | S, An, O |
| | | Lachnospiraceae | *Blautia producta* | S, An |
| | | | *Catonella sp.* | C, An, O |
| | | | *Moryella indoligenes* | C, An |
| | | | *Oribacterium sinus* | C, An, O |
| | | | *Pseudobutyrivibrio ruminis* | S, An |
| | | | *Ruminococcus gnavus* | S, An |
| | | Peptostreptococcaceae | *Peptostreptococcus anaerobius* | S, An, O |
| | | Ruminococcaceae | *Ruminococcus flavefaciens* | C, An |
| | | Veillonellaceae | *Acidaminococcus intestini* | S, An |
| | | | *Dialister invisus* | C, An, O |
| | | | *Dialister micraerophilus* | C, An, O |
| | | | *Dialister pneumosintes* | S, An, O |
| | | | *Mitsuokella multacida* | C, An, O |
| | | | *Selenomonas noxia* | C, An, O |
| | | | *Selenomonas sputigena* | C, An, O |
| | | | *Veillonella dispar* | C, An, O |
| | | | *Veillonella parvula* | C, An, O |
| | | | *Veillonella ratti* | S, An |
| Fusobacteria | Fusobacteriales | Fusobacteriaceae | *Fusobacterium canifelinum* | C, An |
| | | | *Leptotrichia buccalis* | C, An, O |
| | | | *Leptotrichia hofstadii* | C, An, O |
| | | | *Leptotrichia trevisanii* | C, An |
| | | | *Sebaldella termitidis* | S, An |
| Alphaproteobacteria | Rhizobiales | Bradyrhizobiaceae | *Afipia genosp. 1* | C, Ae |
| | | | *Bradyrhizobium elkanii* | C, Ae, O |
| | | | *Nitrobacter vulgaris* | C, Ae |
| | Rhodobacterales | Rhodobacteraceae | *Paracoccus yeei* | S, Ae |
| | Sphingomonadales | Sphingomonadaceae | *Sphingobium yanoikuyae* | S, Ae |
| | | | *Sphingomonas asaccharolytica* | S, Ae |
| Betaproteobacteria | Burkholderiales | Alcaligenaceae | *Bordetella hinzii* | S, Ae |
| | | | *Bordetella petrii* | C, Ae |
| | | Burkholderiaceae | *Burkholderia vietnamiensis* | S, Ae, P |
| | | | *Lautropia mirabilis* | C, Ae, O |
| | | | *Methylibium petroleiphilum* | S, Ae |
| | | | *Roseateles depolymerans* | C, Ae |
| | | Oxalobacteraceae | *Massilia timonae* | S, Ae |
| | Neisseriales | Neisseriaceae | *Bergeriella denitrificans* | C, Ae |
| | | | *Conchiformibius kuhniae* | S, Ae |
| | | | *Eikenella corrodens* | C, Ae, O |

*Table S4 Continued*

| Class | Order | Family | Taxon name | Code |
|---|---|---|---|---|
| | | | *Kingella oralis* | C, Ae, O |
| | | | *Neisseria bacilliformis* | S, Ae, O |
| | | | *Neisseria cinerea* | C, Ae |
| | | | *Neisseria meningitidis* | C, Ae, O |
| | | | *Neisseria subflava* | C, Ae, O |
| Epsilonproteobacteria | Campylobacterales | Campylobacteraceae | *Campylobacter concisus* | C, Ae, O |
| | | | *Campylobacter gracilis* | S, Ae, O |
| | | | *Campylobacter mucosalis* | C, Ae |
| | | | *Campylobacter rectus* | S, Ae, O |
| | | | *Campylobacter showae* | S, Ae, O |
| Gammaproteobacteria | Enterobacteriales | Enterobacteriaceae | *Enterobacter hormaechei* | S, Ae, O |
| | | | *Enterobacter sp.* | S, Ae, O |
| | | | *Escherichia fergusonii* | S, Ae |
| | | | *Klebsiella oxytoca* | C, Ae |
| | | | *Proteus mirabilis* | S, Ae, O |
| | Legionellales | Legionellaceae | *Legionella pneumophila* | S, Ae, P |
| | Pasteurellales | Pasteurellaceae | *Actinobacillus porcitonsillarum* | S, Ae |
| | | | *Aggregatibacter aphrophilus* | C, Ae, O |
| | | | *Aggregatibacter segnis* | C, Ae, O |
| | | | *Haemophilus influenzae* | C, Ae, O, P |
| | | | *Haemophilus parainfluenzae* | C, Ae, O |
| | | | *Haemophilus parahaemolyticus* | C, Ae, O |
| | | | *Pasteurella multocida* | C, Ae |
| | Pseudomonadales | Moraxellaceae | *Acinetobacter schindleri* | S, Ae |
| | | | *Moraxella catarrhalis* | C, Ae |
| | | Pseudomonadaceae | *Pseudomonas sp.* | C, Ae |
| | | | *Pseudomonas geniculata* | S, Ae |
| | | | *Pseudomonas aeruginosa* | C, Ae, O, P |
| | | | *Pseudomonas stutzeri* | S, O |
| | Xanthomonadales | Xanthomonadaceae | *Stenotrophomonas acidaminiphila* | S, Ae |
| Spirochaetes | Spirochaetales | Spirochaetaceae | *Treponema denticola* | C, An, O |
| | | | *Treponema lecithinolyticum* | C, An, O |
| | | | *Treponema medium* | C, An, O |
| | | | *Treponema socranskii* | C, An, O |
| Mollicutes | Mycoplasmatales | Mycoplasmataceae | *Mycoplasma arthritidis* | C, Ae |

**Table S4.** Bacterial species sampled across the 38 sputum samples. Ae, denotes aerobe; An, Anerobe[1]; O, taxa associated with the oral microbiome[2]; P, species with populations reported to be pathogenic in the context of respiratory infections. Also given are indications of core (C) and satellite (S) species group membership.

[1] Only strict anaerobes were classified as anaerobes, whereas aerobes, facultative anaerobes, and microaerophiles were classified as aerobes, as described previously.[13]

[2] Classification of oral taxa was according to the Human Oral Microbiome Database.[14]

| Class | Order | Family | Taxon name | BAL1 (161) | BAL2 (10) | Sputum 1 (2049) | Sputum 2 (1474) |
|---|---|---|---|---|---|---|---|
| Rubrobacteridae | Solirubrobacterales | Solirubrobacteraceae | Solirubrobacter sp. | 14.3 | | | |
| Actinobacteria | Actinomycetales | Actinomycetaceae | Actinomyces odontolyticus | | | 1.07 | 0.61 |
| | | | Actinomyces oris | | | 0.05 | |
| | | | Actinomyces urogenitalis | | | 4.49 | 0.14 |
| | | Corynebacteriaceae | Corynebacterium tuberculostearicum | | | | 0.07 |
| | | Nocardioidaceae | Propionibacterium acnes | 1.2 | | | |
| | | Micrococcaceae | Rothia dentocariosa | | | | 0.47 |
| | | | Rothia mucilaginosa | | | | 0.14 |
| | Coriobacteriales | Coriobacteriaceae | Atopobium parvulum | | | 0.49 | 0.20 |
| Actinomycetales | Corynebacterineae | Mycobacteriaceae | Mycobacterium celatum | 6.8 | | | |
| | Frankineae | Geodermatophilaceae | Modestobacter sp. | 7.5 | | | |
| | Propionibacterineae | Nocardioidaceae | Marmoricola sp. | 1.9 | | | |
| | Pseudonocardineae | Pseudonocardiaceae | Actinomycetospora sp. | 0.6 | | | |
| Flavobacteria | Flavobacteriales | Flavobacteriaceae | Capnocytophaga sputigena | | | | 0.27 |
| | | | Flavobacterium succinicans | | | 0.05 | 0.07 |
| Mollicutes | Mycoplasmatales | Mycoplasmataceae | Mycoplasma arthritidis | | | 0.05 | 0.07 |
| Bacteroidia | Bacteroidales | Bacteroidaceae | Bacteroides coprophilus | | | 0.05 | |
| | | Porphyromonadaceae | Porphyromonas gingivalis | | | | 0.54 |
| | | | Porphyromonas levii | | | 0.05 | 0.20 |
| | | | Prevotella copri | | | 0.05 | 0.20 |
| | | | Prevotella histicola | | | 0.93 | 1.97 |
| | | | Prevotella melaninogenica | | | 37.77 | 14.45 |
| | | | Prevotella nanceiensis | | | | 0.20 |
| | | | Prevotella nigrescens | | | 0.05 | |
| | | | Prevotella oris | | | 0.05 | 0.14 |
| | | | Prevotella oulorum | | | 1.61 | |
| | | | Prevotella pallens | | | 3.27 | 0.88 |
| | | | Prevotella tannerae | | | | 0.07 |
| | | | Prevotella veroralis | | | 0.15 | 0.81 |
| Bacilli | Bacillales | Staphylococcaceae | Staphylococcus aureus | 0.6 | | | |
| | | Staphylococcaceae | Staphylococcus epidermidis | 41.6 | | | |
| | Gemellales | Gemellaceae | Gemella sanguinis | | | | 0.20 |
| | Lactobacillales | Carnobacteriaceae | Granulicatella adiacens | | | 0.10 | 0.61 |
| | | Streptococcaceae | Streptococcus anginosus | | | | 0.20 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | Streptococcus infantis | | | | 1.02 |
| | | | Streptococcus oralis | | | 0.49 | 2.10 |
| | | | Streptococcus parasanguinis | | 10 | 0.24 | 1.29 |
| | | | Streptococcus pneumoniae | | | | 0.14 |
| | | | Streptococcus pseudopneumoniae | | | 0.15 | 1.22 |
| | | | Streptococcus salivarius | | | 1.17 | 2.24 |
| | | | Streptococcus sanguinis | | | 0.05 | |
| Clostridia | Clostridiales | Clostridiaceae | Clostridium phytofermentans | | | | |
| | | | Clostridium sp. | | | 0.15 | 0.07 |
| Negativicutes | Selenomonadales | Veillonellaceae | Dialister invisus | | | | 0.07 |
| | | Lachnospiraceae | Moryella indoligenes | | | 0.24 | 0.34 |
| | | | Oribacterium sinus | | | 0.10 | 0.20 |
| | | | Ruminococcus flavefaciens | | | | 0.14 |
| | | Veillonellaceae | Selenomonas noxia | | | 1.56 | 0.14 |
| | | | Veillonella dispar | | 10 | 36.41 | 7.19 |
| | | | Veillonella parvula | | | 4.69 | 49.32 |
| | | Veillonellaceae | Anaeroglobus sp. | | | 1.76 | 0.20 |
| Fusobacteria | Fusobacteriales | Fusobacteriaceae | Fusobacterium canifelinum | 0.6 | | 0.83 | 2.71 |
| | | | Leptotrichia buccalis | | | 0.15 | |
| | | | Leptotrichia trevisanii | | | 0.20 | 0.20 |
| Alphaproteobacteria | Rhizobiales | Bradyrhizobiaceae | Afipia felis | 11.2 | | | |
| Betaproteobacteria | Burkholderiales | Burkholderiaceae | Roseateles depolymerans | 0.6 | | | |
| | | | Ralstonia mannitolilytica | 8.7 | | | |
| | | | Ralstonia solanacearum | 3.1 | | | |
| | | Comamonadaceae | Comamonas aquatica | | 20 | | |
| | | | Delftia sp. | | 30 | | |
| | Neisseriales | Neisseriaceae | Eikenella corrodens | | | 0.10 | 0.07 |
| | | | Kingella oralis | | | | 0.27 |
| | | | Neisseria subflava | | | 0.39 | 0.41 |
| Epsilonproteobacteria | Campylobacterales | Campylobacteraceae | Campylobacter concisus | | | 0.44 | 0.14 |
| Gammaproteobacteria | Pasteurellales | Pasteurellaceae | Aggregatibacter segnis | | | | 3.39 |
| | | | Haemophilus parainfluenzae | | | 0.59 | 3.87 |
| | Pseudomonadales | Pseudomonadaceae | Pseudomonas geniculata | 1.2 | 30 | | 0.14 |
| | | | Pseudomonas nitroreducens | | | 0.05 | |
| Tm7 candidate division | | | | | | | 0.61 |

**Table S5.** Bacterial taxa identified through 16S rRNA gene pyrosequencing analysis in samples from healthy control subjects. Total numbers of sequences obtained are shown in brackets. Sequences identification is expressed as a percentage of species obtained.

| | $FEV_1$% (Pre-BD) | $FEV_1$% (Post-BD) | LCS | Neutrophils | Gender |
|---|---|---|---|---|---|
| $FEV_1$% (Pre-BD) | - | 0.868 | 0.152 | 0.122 | -0.025 |
| $FEV_1$% (Post-BD) | <0.0001* | - | 0.134 | 0.150 | 0.020 |
| LCS | <0.0001* | 0.001* | - | 0.116 | -0.005 |
| Neutrophils | 0.0001* | 0.0001* | 0.002* | - | 0.003 |
| Gender | 0.507 | 0.593 | 0.897 | 0.930 | - |

**Table S6.** Mantel tests for co-correlations between clinical factors. Given for each test are the Mantel statistic (*r*) and significance (*P)* to ascertain whether the Mantel regression coefficients (Pearson's correlation coefficient) were significantly different following 9,999 Monte Carlo permutations. Values for *r* and *P* are given in the upper and lower triangles of the matrix, respectively. Asterisks denote *P*-values significant after Bonferroni correction for multiple comparisons (0.05/5 = 0.01). LCS denotes Leicester Cough Score and BD denotes bronchodilator.

| Parameter | Control for | All $r$ | $P$ | Core $r$ | $P$ |
|---|---|---|---|---|---|
| $FEV_1\%$ | Gender | -0.169 | <0.0001* | -0.170 | <0.0001* |
| $FEV_1\%$ | Leicester Cough Score | -0.151 | <0.0001* | -0.151 | <0.0001* |
| $FEV_1\%$ | Neutrophils | -0.156 | <0.0001* | -0.156 | <0.0001* |
| Gender | $FEV_1\%$ | -0.078 | 0.020 | -0.077 | 0.021 |
| Gender | Leicester Cough Score | -0.074 | 0.025 | -0.073 | 0.026 |
| Gender | Neutrophils | -0.073 | 0.027 | -0.072 | 0.028 |
| Leicester Cough Score | $FEV_1\%$ | -0.103 | 0.003* | -0.103 | 0.003* |
| Leicester Cough Score | Gender | -0.126 | 0.0001* | -0.127 | 0.0001* |
| Leicester Cough Score | Neutrophils | -0.114 | 0.001* | -0.114 | 0.001* |
| Neutrophils | $FEV_1\%$ | -0.094 | 0.004* | -0.094 | 0.0006* |
| Neutrophils | Gender | -0.112 | 0.001* | -0.113 | 0.001* |
| Neutrophils | Leicester Cough Score | -0.099 | 0.004* | -0.099 | 0.004* |

**Table S7.** Summary statistics for partial Mantel tests. The partial Mantel statistic $r(AB.C)$ estimates the correlations (Pearson's correlation coefficient) between two proximity matrices, $A$ and $B$ whilst controlling for the effects of $C$. Given are the bacterial community (Bray-Curtis) similarities for whole communities (All) and the core group members, and also differences in clinical factors. Also given is $P$ to ascertain whether the Mantel regression coefficients were significantly different from zero following 9,999 Monte Carlo permutations. Asterisks denote $P$-values significant after Bonferroni correction for multiple comparisons (0.05/12 = 0.0042). Satellite group members are excluded from these analyses. $FEV_1\%$ pre-bronchodilator is presented here.

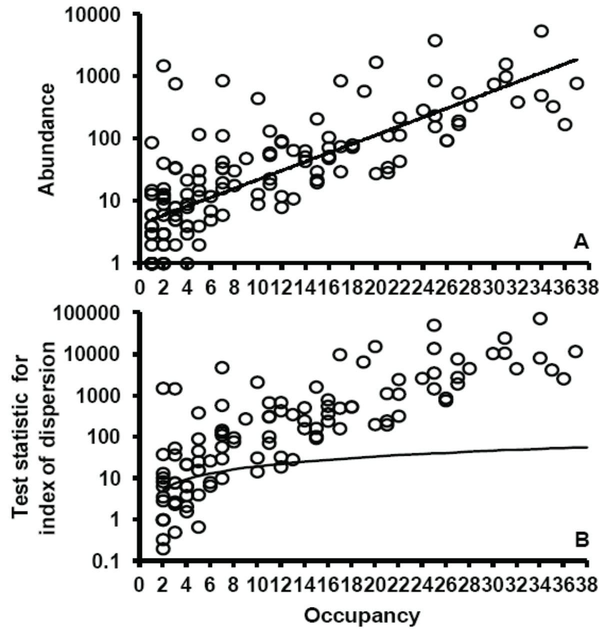


**Figure S1.** Distribution and dispersal of bacterial species across patients. (A) The number of samples for which each bacterial species was observed, plotted against its maximum abundance ($\log_{10}$ scale) across all 38 samples ($r^2$ = 060, $F_{1,138}$ = 203.12, $P$ < 0.0001). (B) Random and non-random dispersal through space visualised by decomposing the overall distribution using an index of dispersion based on the ratio of variance to the mean abundance for each bacterial species from the 38 samples. The line depicts the 2.5 % confidence limit for the $\chi^2$ distribution. The 97.5% confidence limit was not plotted, as no species fell below that line.

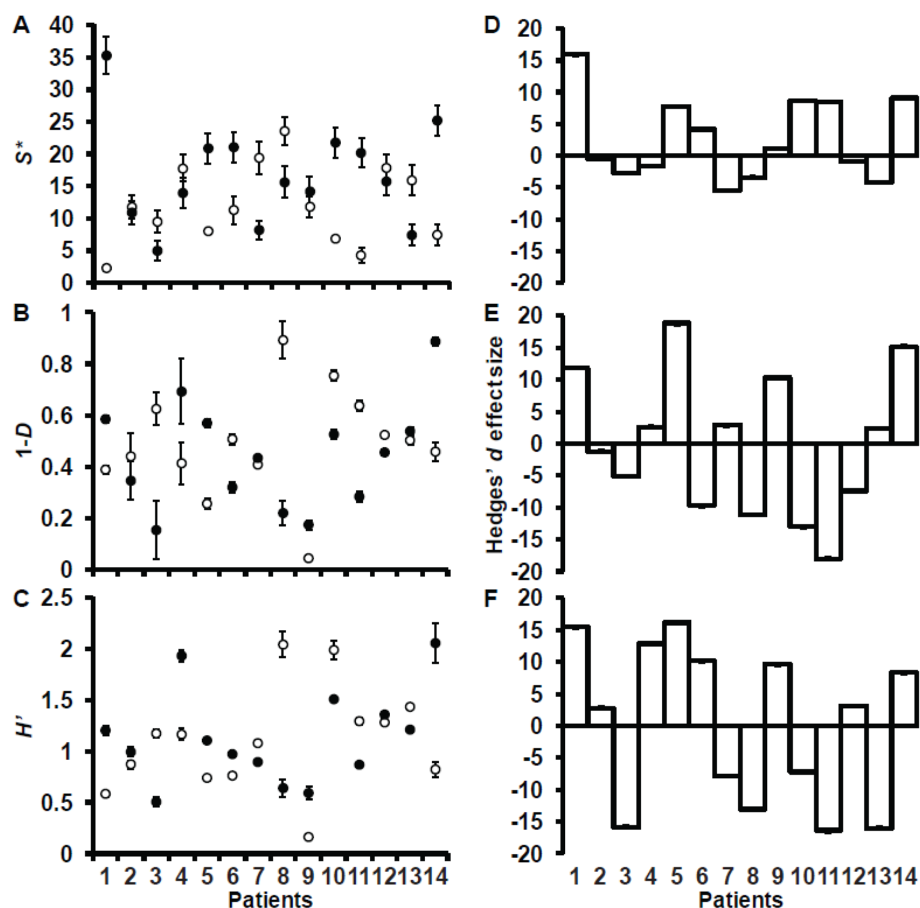


**Figure S2.** Comparisons of diversity and the effects of sample type between paired BAL and sputum samples. Given are (A) species richness (*S**), (B) Simpson's complement index (1-*D*), and (C) Shannon-Wiener index (*H*') for paired BAL (open circles) and sputum samples (closed circles). Diversity measures were calculated with a uniform re-sample size of 809 sequences following 1000 iterations in each instance. Error bars represent the SD of the mean ($n$ = 1000). Meta-analysis of (D) *S**, (E) 1-*D*, and (F) *H*' using Hedges' *d* effect size measure between paired BAL and sputum samples. Columns represent effect size and error bars represent the SE of the effect size ($n$ = 2000). The SE bars that cross zero indicate no significant effect on diversity by sample type.

## References


1. David J Serisier, Megan L Martin, Michael A McGuckin, et al. Effect of long-term, low-dose erythromycin on pulmonary exacerbations among patients with non-cystic fibrosis bronchiectasis. The BLESS randomized, controlled trial. JAMA. 2013. In press.
2. Birring SS, Prudon B, Carr AJ, *et al*. Development of a symptom specific health status meaure for patients with chronic cough: Leicester Cough Questionnaire (LCQ). *Thorax* 2003;**58**:339-343.
3. Jones P, Quirk F, Baveystock C, *et al*. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis* 1992;**145**:1321-758.
4. Paggiaro PL, Chanez P, Holz O, *et al*. Sputum induction. *Eur Respir J* 2002;**20**:Suppl 37, 3s-8s.
5. Sagel SD, Kapsner R, Osberg I, *et al*. Airway inflammation in children with cystic fibrosis and healthy children assessed by sputum induction. *Am J Respir Crit Care Med* 2011;**164**:1425–31.
6. Wood-Baker R, Burdon J, McGregor A, *et al*. Fibre-optic bronchoscopy in adults – a position paper of the Thoracic Society of Australia and New Zealand. *Intern Med J* 2001;**31**:479-87.
7. Monton C, Torres A, El-Ebiary M, *et al*. Cytokine expression in severe pneumonia: a bronchoalveolar lavage study. *Crit Care Med* 1999;**27**:1745-53.
8. Hilliard JB, Konstan MW, Davis PB. Inflammatory mediators in CF patients. In Methods in Molecular Medicine 2002, vol 70: Cystic Fibrosis Methods and protocols. WR Skach (ed), Humana Press Inc, Totowa, NJ, USA

9. Hattotuwa K, Gamble EA, O'Shaughnessy T, *et al*. Safety of bronchoscopy, biopsy and BAL in research patients with COPD. *Eur Respir J* 2002;**122**:1909-12

10. Dowd SE, Wolcott RD, Sun Y, McKeehan T, *et al*. Polymicrobial nature of chronic diabetic foot ulcer biofilm infections determined using bacterial tag encoded FLX amplicon pyrosequencing (bTEFAP). *PLOS One* 2008;**3**:e3326.

11. Gontcharova V, Youn E, Sun Y, *et al*. A comparison of bacterial composition in diabetic ulcers and contralateral intact skin. *Open Microbiol J* 2010;**4:**8-19.

12. Huang X, Madan A (1999) CAP3: A DNA sequence assembly program.Genome Res 9: 868–877.

13. van der Gast CJ, Walker AW, Stressmann FA, *et al*. Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *ISME J* 2011;**5**:780-91.

14. Dewhirst FE, Chen T, Izard J, *et al*. The human oral microbiome. *J Bacteriol* 2010;**192**:5002-17.