

RESEARCH ARTICLE

# Profiling mycobacterial communities in pulmonary nontuberculous mycobacterial disease

Steven A. Cowman<sup>1,2</sup>, Phillip James<sup>1</sup>, Robert Wilson<sup>1,2</sup>, William O. C. Cookson<sup>1</sup>, Miriam F. Moffatt<sup>1</sup>, Michael R. Loebinger<sup>1,2</sup>\*

**1** National Heart and Lung Institute, Imperial College London, London, United Kingdom, **2** Host Defence Unit, Royal Brompton Hospital, London, United Kingdom

✉ These authors contributed equally to this work.

\* [M.Loebinger@rbht.nhs.uk](mailto:M.Loebinger@rbht.nhs.uk)



**OPEN ACCESS**

**Citation:** Cowman SA, James P, Wilson R, Cookson WOC, Moffatt MF, Loebinger MR (2018) Profiling mycobacterial communities in pulmonary nontuberculous mycobacterial disease. PLoS ONE 13(12): e0208018. <https://doi.org/10.1371/journal.pone.0208018>

**Editor:** Yoshihiko Hoshino, National Institute of Infectious Diseases, JAPAN

**Received:** May 24, 2018

**Accepted:** November 10, 2018

**Published:** December 11, 2018

**Copyright:** © 2018 Cowman et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The raw sequence data used in this study are available from the European Nucleotide Archive (<https://www.ebi.ac.uk/ena>) under accession number PRJEB22153.

**Funding:** S.A.C. and M.R.L. were funded by a grant from the Welton Foundation. W.O.C.C. and M.F.M. were funded by a joint investigator award from the Wellcome Trust (<https://wellcome.ac.uk/>) reference numbers WT097117 and WT096964. Additional funding for W.O.C.C. and M.F.M. was from the Asmarley Charitable Trust. Support for all authors

## Abstract

The diagnosis of pulmonary non-tuberculous mycobacterial disease (pNTM) is dependent on the isolation of NTM in culture, which is prone to overgrowth and contamination and may not capture the diversity of mycobacteria present, including rare or unidentified species.

This study aimed to develop a culture independent method of detecting and identifying mycobacteria from sputum samples using partial sequencing of the hsp65 gene. DNA was extracted from sputum samples from subjects with pNTM and disease controls. Multiplexed partial sequencing of the hsp65 gene was performed using the Illumina MiSeq and custom primers. A reference database of hsp65 sequences was created for taxonomy assignment. Sequencing results were obtained from 42 subjects (31 cases, 11 controls). Mycobacterial sequences were identified in all subjects. In 90.5% of samples more than one species was found (median 5.5). The species isolated in culture was detected by sequencing in 81% of subjects and was the most abundant species in 62%. The sequencing of NTM from clinical samples reveals a far greater diversity than conventional culture and suggests NTM are present as communities rather than a single species. NTM were found to be present even in the absence of isolation in culture or clinical disease.

## Introduction

Pulmonary infection with non-tuberculous mycobacteria (NTM) is a challenging disease which is becoming increasingly prevalent. It is associated with a high mortality, and treatment is prolonged, expensive and frequently poorly tolerated [1,2]. The diagnosis of NTM infection relies upon the isolation of NTM in culture which is challenging due to the overgrowth of other co-pathogens, such as *Pseudomonas*, contamination from environmental mycobacteria and failure to grow due to fastidious requirements, as well as the time and labour involved.

Molecular techniques utilising nucleotide amplification have been used directly on clinical samples in TB infection and have become established in clinical practice [3]. Their use in NTM disease, however, is currently confined to the identification of species previously

was also provided by the NIHR Respiratory Disease Biomedical Research Unit at the Royal Brompton and Harefield NHS Foundation Trust and Imperial College London. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** M.R.L reports receiving personal fees for consultancy/advisory board work from Insmid Inc. (New Jersey, USA) and Savara Inc. (Texas, USA) outside the submitted work. The other authors have declared that no competing interests exist. This does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

cultured from clinical samples. As well as overcoming the limitations of culture based techniques, the sequencing of mycobacteria directly from a sample has the potential to capture the diversity of mycobacterial species present including rare and as yet unidentified species. This may be valuable, as the isolation of multiple species is not uncommon and may pose a therapeutic challenge [4,5]. Molecular techniques have been successfully employed in both environmental [6–8] and more recently in human samples [9].

The aim of this study was to employ next generation sequencing of the *hsp65* gene to characterise mycobacterial communities in the lung directly from sputum samples.

## Materials and methods

Written consent was gained from all participants and the study was approved by the London - Chelsea Research Ethics Committee (reference 12/LO/1034). Subjects were recruited from the outpatient department of the Royal Brompton Hospital between September 2012 and November 2013. Cases were enrolled if they met ATS/IDSA 2007 criteria [10] and had non-cystic fibrosis bronchiectasis, chronic obstructive pulmonary disease (COPD) or no underlying lung disease prior to their diagnosis of NTM disease. Controls had non-cystic fibrosis bronchiectasis or COPD with no clinical or radiological evidence of NTM infection, previous negative sputum mycobacterial cultures and a negative sputum mycobacterial culture on enrolment. Exclusion criteria were: pregnancy, immunosuppression (other than oral corticosteroids), and current malignancy or tuberculosis.

Participants underwent clinical assessment and pulmonary function testing. If not performed within the previous six months high resolution computed tomography (HRCT) was performed. Spontaneously expectorated sputum was stored at -80 degrees Celsius for DNA extraction as well as undergoing phenol auramine microscopy and culture for bacteria, mycobacteria and fungi.

Barcoded primers were designed for use with the MiSeq benchtop sequencer (Illumina, San Diego, CA, USA) as previously described [11] using the TB11/TB12 primer pair, which amplify a 401bp region of the mycobacterial *hsp65* gene [12]. This target was chosen as the amplicon length is suitable for the MiSeq platform, and enables clear differentiation between *M. abscessus* and *M. chelonae* [13] as well as differentiation between *M. abscessus* spp. *abscessus* and ssp. *bolletii/massiliense* which are identical by partial sequencing of the 16S rRNA gene [14–16].

As a positive control, a mock community was created consisting of *hsp65* amplicons from 16 NTM typestrains cloned, sequenced and pooled in equal abundance. A reference database was created consisting of all mycobacterial *hsp65* sequences present in the National Center for Biotechnology Information (NCBI) GenBank (<http://www.ncbi.nlm.nih.gov/nucleotide/>) in addition to the sequences of the NTM typestrains of the mock community and non-mycobacterial sequences which also may be amplified by the TB11/TB12 primer pair, identified using NCBI primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>).

For each patient a 300µl aliquot of stored sputum underwent bead-beating and DNA was extracted using hexadecyltrimethylammonium bromide and phenol:chloroform:isoamylalcohol (see supporting information for further information). As a control, molecular-grade water was collected in the same receptacles as used for sputum and subjected to the same storage and DNA extraction protocol. Barcoded *hsp65* sequences were amplified from each DNA sample in quadruplicate and libraries prepared and sequenced using the MiSeq benchtop sequencer and MiSeq v3 600 cycle reagents (Illumina).

Sequence processing was performed using the QIIME pipeline 1.8.0 [17] applying a 99% similarity for Operational taxonomic unit (OTU) picking. The resulting data was imported

into the R environment version 3.1.2 [18] where all downstream analysis was performed. The relative abundance of taxa between groups were compared using the Wilcoxon rank sum test using the Benjamini-Hochberg correction for multiple testing. Diversity metrics were calculated using the phyloseq and vegan packages [19,20] and compared between groups using the t-test and correlated with clinical variables using Pearson correlation. Differences in community structure were tested using permuted analysis of variance (PERMANOVA) in the vegan package. A *P* value of < 0.05 was considered significant.

Full details of the *hsp65* database and mock community creation, DNA extraction, library preparation, sequencing and the processing of sequence data can be found in the supporting information. Sequence data are available from the European Nucleotide Archive under accession number PRJEB22153.

## Results

Sputum was collected for DNA extraction from 56 eligible subjects. A total of 4.84 million joined sequences were obtained, leading to 4.16 million sequences available for analysis following processing and quality control (for more detail see the supporting information). Samples with a low depth of sequencing (under 5,000 sequences) were removed (N = 14) leaving 42 subjects for analysis with a median depth of sequencing of 80,960 sequences per sample. The clinical characteristics of the final study population are shown in Table 1. All controls were smear and culture negative and had provided at least one prior negative mycobacterial culture (median 4 samples, range 1 to 15).

Continuous variables are expressed as mean +/- standard deviation, categorical variables as count and percentage. NTM species refers to the species isolated in culture leading to the original diagnosis of pulmonary NTM disease.

**Table 1. Clinical characteristics of the study population.**

	Cases N = 31	Controls N = 11
Age (years)	65.77 (±11.32)	65.58 (±4.78)
Male sex	12 (38.7%)	6 (54.5%)
Ex- or current smoker	15 (48.4%)	6 (54.5%)
<b>Underlying diagnosis</b>		
COPD	5 (16.1%)	0
Bronchiectasis	23 (74.2%)	11 (100%)
No underlying lung disease	3 (9.7%)	0
<b>FEV1% predicted</b>	52.85 (±29.30)	79.19 (±16.67)
<b>FVC % predicted</b>	89.96 (±32.01)	92.73 (±15.29)
<b>Prophylactic antibiotics</b>	7 (22.6%)	5 (45.5%)
<b>Oral steroid</b>	1 (3.2%)	2 (18.2%)
<b>NTM treatment</b>	7 (22.6%)	-
<b>NTM species</b>		
<i>M. abscessus</i>	4 (12.9%)	-
<i>M. avium</i> complex (MAC)	18 (58.1%)	-
<i>M. fortuitum</i>	1 (3.2%)	-
<i>M. kansasii</i>	5 (16.1%)	-
<i>M. malmoense</i>	1 (3.2%)	-
<i>M. simiae</i>	1 (3.2%)	-
<i>M. xenopi</i>	1 (3.2%)	-

<https://doi.org/10.1371/journal.pone.0208018.t001>

All species within the mock community were detected. The median difference between the observed proportional abundance of each species and the expected proportion was 2.7% (range 0.2% to 5.5%) with *M. xenopi* and *M. tuberculosis* complex being the most overrepresented and *M. abscessus* subsp. *abscessus* and *M. peregrinum* being the most underrepresented (S2 Fig). There was no evidence of contamination in the negative extraction controls or the negative PCR control.

### Genus level

Mycobacterial sequences were detected in all samples, although only 6.23% of all sequences were from mycobacteria and the median proportion of mycobacterial sequences in each sample was 0.15% (S3 Table and S4 Fig). The most abundant genus was Actinomyces, accounting for 51.5% of all sequences. Samples tended to be dominated by one genus and in 38 samples one genus represented over 50% of all sequences. In two subjects Mycobacterium was the most abundant genus; both were culture positive for *Mycobacterium avium* complex (MAC) and *M. abscessus* respectively. There was no significant difference in the median relative abundance of mycobacteria in cases compared with controls (0.19% vs 0.02%,  $P = 0.65$ ). There was no significant correlation between mycobacterial abundance and *hsp65* amplicon concentration (Spearman rho -0.18  $P = 0.255$ , see S3 Fig in the supporting information).

### Species level

All sequences from the genus Mycobacterium (N = 258,421) were selected for further analysis. The most common species was *M. avium*, followed by *M. abscessus* subsp. *bolletii* and *M. simiae* (Table 2). In 32 samples the most abundant species constituted over 50% of all

**Table 2. The abundance of mycobacterial species discovered by sequencing.**

Species	Mean proportion reads per sample	No. subjects where present	No. subjects where most abundant	Number of OTUs
<i>M. avium</i>	35.01%	33	14	107
<i>M. abscessus</i> subsp. <i>bolletii</i>	17.44%	28	8	28
<i>M. simiae</i>	13.67%	27	8	15
<i>M. abscessus</i> subsp. <i>abscessus</i>	6.55%	23	3	35
<i>M. xenopi</i>	5.32%	13	2	12
<i>M. intracellulare chimaera</i>	4.40%	22	3	3
<i>M. intracellulare</i>	4.16%	25	1	37
<i>M. kansasii</i>	3.81%	16	2	5
<i>M. arupense</i>	1.37%	13	0	11
<i>M. chelonae</i>	1.08%	13	1	9
<i>M. rhodesiae</i>	0.99%	9	0	8
<i>M. fortuitum</i>	0.97%	11	0	13
<i>M. gordonae</i>	0.94%	7	0	4
<i>M. psychrotolerans</i>	0.89%	13	0	4
<i>M. malmoense</i>	0.71%	13	0	13
<i>M. peregrinum</i>	0.58%	9	0	7
<i>M. tuberculosis</i> complex	0.58%	11	0	20
<i>M. neoaurum</i>	0.52%	3	0	1
<i>M. szulgai</i>	0.39%	13	0	4
<i>M. kubicae</i>	0.02%	2	0	1
<i>M. abscessus</i> complex	0.01%	1	0	1
Unidentified	0.58%	10	0	17

<https://doi.org/10.1371/journal.pone.0208018.t002>

mycobacterial sequences. More than one species was identified in 38 (90.5%) of samples. The median number of species present was 5.5 (range 1 to 19). Details of sequencing for each subject are shown in [Table 3](#).

To allow comparison at the community level between subjects with varying depth of sequencing, samples were rarefied to a depth of 100 sequences leaving 25 subjects (18 cases and 7 controls) for analysis. The community structure of each subject is shown in [Fig 1](#). After rarefaction the majority of subjects (92%) had more than one species present (median 5, range 2 to 17). There were no significant differences in the abundance of any NTM species between cases and controls, and no difference in community structure by PERMANOVA.

Exploration of diversity ([S5](#), [S6](#) and [S7](#) Figs) revealed a significant reduction in diversity as measured by Shannon and Pielou's indices in NTM cases compared to controls. Within NTM cases alone, Shannon index and observed richness were higher in those receiving treatment for NTM, and Shannon and Pielou's indices were positively correlated with forced expiratory volume in 1 second.

### Comparison with sputum culture

The cultured NTM species was detected by sequencing in 17 (81%) of the 21 culture positive samples ([Table 3](#)). NTM were detected by sequencing in all of the 19 culture-negative samples. The median proportion of mycobacterial sequences was lower in the four discordant samples where the isolated species was not detected by sequencing (0.006%) compared to the concordant samples where the isolated species was also detected by sequencing (0.12%) or the culture negative samples (0.31%), although this difference did not reach significance ([S8 Fig](#), Kruskal-Wallis  $P = 0.054$ ).

In 13 of the 21 cases (62%) with culture-positive sputum the isolated species was also the most abundant species found by sequencing. Agreement was highest for *M. kansasii*, which was the most abundant species in both culture positive subjects, and lowest for *M. simiae*, which was not cultured from any of the five subjects in which it was the most abundant species by sequencing. No sputum sample isolated more than one NTM species by culture. However, in 12 cases (39%) more than one NTM species had been previously isolated at the time of study enrolment ([S4 Table](#)).

### Discussion

This study has demonstrated the utility of next generation sequencing with multiplexed bar-coded paired-end primers to sequence mycobacterial hsp65 directly from sputum samples. The majority of sequences were identified to the species level. The study is the first to provide a profile of mycobacterial communities within the lung.

The use of a mock community confirms the ability to detect a broad range of NTM species and allows assessment of amplification bias. Contamination is an important concern in microbiome studies [21] and the use of extraction controls mirroring the processing of sputum samples provides reassurance that the findings are unlikely to be the result of contamination.

Mycobacterial sequences were recovered from every sample in both cases and controls, regardless of their growth in culture or the presence of clinical disease. It has long been appreciated that mycobacteria may be cultured from subjects in the absence of disease. This has led to the development of microbiological criteria for the diagnosis of NTM disease [10] which are not met by around half of subjects in whom NTM are isolated [22,23]. This has been attributed to contamination or 'casual' isolation of transiently inhaled organisms, however the data from this present study suggest that this isolation may reflect the true recovery of NTM from the respiratory tract, and that sputum culture may in fact lack sensitivity to the presence of NTM.

Table 3. Sequencing and culture results for each subject.

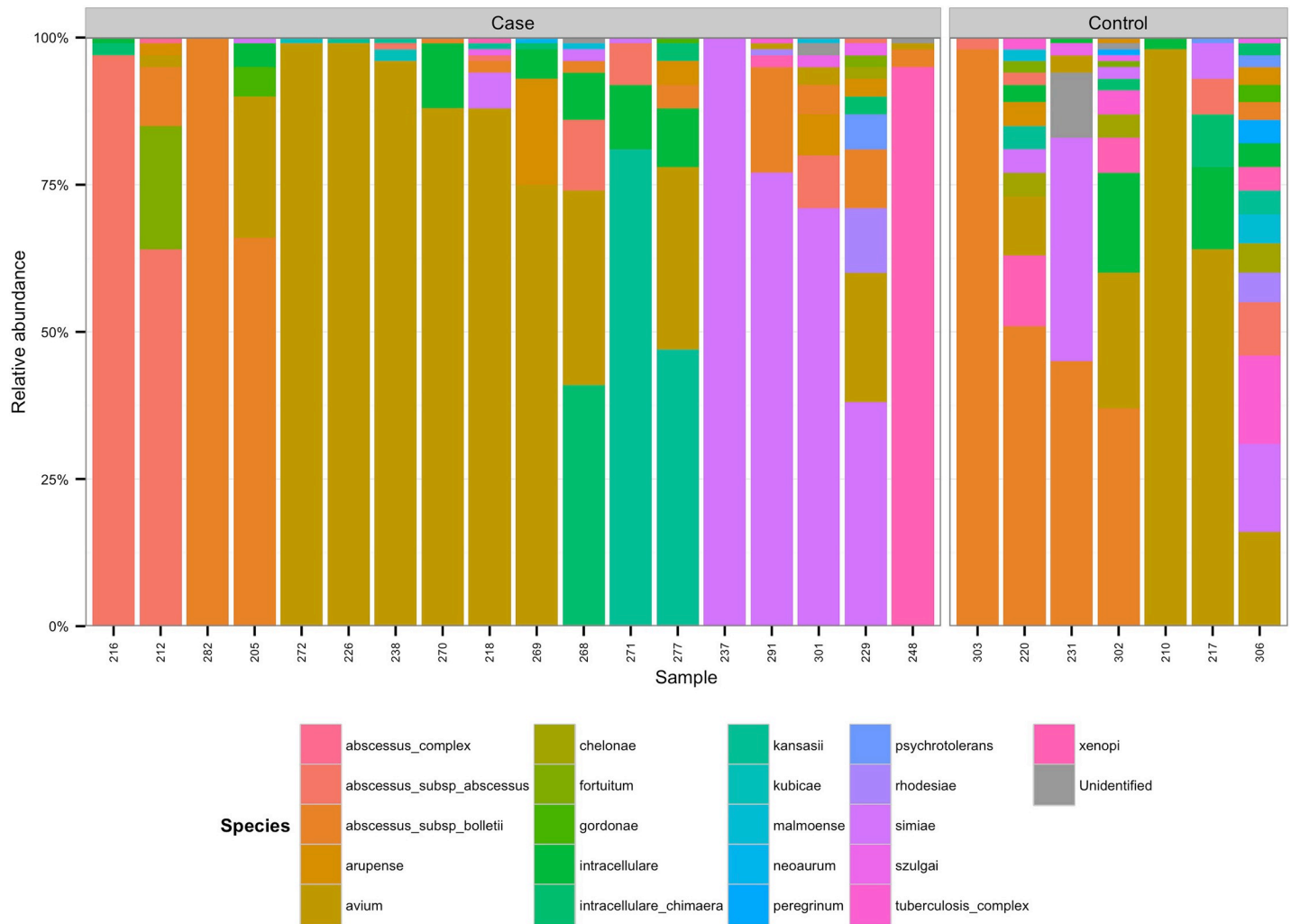
ID	Group	Diagnosed species	No. mycobacterial sequences	No. mycobacterial OTUs	Most abundant species by sequencing	Cultured species	Cultured species detected by sequencing
282	Case	<i>M. abscessus</i>	156074	57	<i>M. abscessus subsp. bolletii</i>	<i>M. abscessus</i>	Yes
272	Case	MAC	75301	173	<i>M. avium</i>	MAC	Yes
238	Case	MAC	11622	119	<i>M. avium</i>	MAC	Yes
212	Case	<i>M. abscessus</i>	4092	74	<i>M. abscessus subsp. abscessus</i>	<i>M. abscessus</i>	Yes
216	Case	<i>M. abscessus</i>	3745	39	<i>M. abscessus subsp. abscessus</i>	MAC	Yes
291	Case	<i>M. simiae</i> *	1648	55	<i>M. simiae</i>	No sample	-
248	Case	<i>M. fortuitum</i>	1410	21	<i>M. xenopi</i>	No growth	-
270	Case	MAC	557	23	<i>M. avium</i>	No growth	-
269	Case	MAC	441	25	<i>M. avium</i>	MAC	Yes
218	Case	MAC	310	27	<i>M. avium</i>	MAC	Yes
226	Case	<i>M. kansasii</i>	247	10	<i>M. avium</i>	<i>M. kansasii</i>	Yes
277	Case	<i>M. kansasii</i>	204	15	<i>M. kansasii</i>	<i>M. kansasii</i>	Yes
301	Case	MAC	199	15	<i>M. simiae</i>	No sample	-
237	Case	<i>M. malmoense</i>	165	3	<i>M. simiae</i>	<i>M. xenopi</i>	No
205	Case	MAC	133	26	<i>M. abscessus subsp. bolletii</i>	No growth	-
268	Case	MAC	119	30	<i>M. intracellulare chimaera</i>	MAC	Yes
271	Case	<i>M. kansasii</i>	102	6	<i>M. kansasii</i>	<i>M. kansasii</i>	Yes
229	Case	MAC	101	29	<i>M. simiae</i>	MAC	Yes
249	Case	<i>M. kansasii</i>	97	11	<i>M. simiae</i>	<i>M. kansasii</i>	No
266	Case	<i>M. kansasii</i>	67	5	<i>M. abscessus subsp. abscessus</i>	<i>M. kansasii</i>	Yes
294	Case	<i>M. abscessus</i>	57	9	<i>M. avium</i>	No growth	-
214	Case	MAC	54	4	<i>M. xenopi</i>	<i>M. xenopi</i>	Yes
276	Case	<i>M. xenopi</i>	47	2	<i>M. abscessus subsp. bolletii</i>	<i>M. xenopi</i>	No
250	Case	MAC	32	3	<i>M. abscessus subsp. bolletii</i>	No growth	-
245	Case	MAC	17	9	<i>M. avium</i>	MAC	Yes
235	Case	MAC	9	6	<i>M. intracellulare chimaera</i>	No growth	-
241	Case	MAC	9	5	<i>M. intracellulare chimaera</i>	MAC	Yes
299	Case	MAC	7	3	<i>M. avium</i>	No growth	-
224	Case	MAC	6	2	<i>M. abscessus subsp. bolletii</i>	No growth	-
263	Case	MAC	2	1	<i>M. avium</i>	MAC	Yes
222	Case	MAC	1	1	<i>M. avium</i>	<i>M. abscessus</i>	No
303	Control		468	15	<i>M. abscessus subsp. bolletii</i>	No growth	-
220	Control		209	50	<i>M. abscessus subsp. bolletii</i>	No growth	-
306	Control		176	47	<i>M. simiae</i>	No growth	-
217	Control		167	22	<i>M. avium</i>	No growth	-
231	Control		143	12	<i>M. simiae</i>	No growth	-
210	Control		140	15	<i>M. avium</i>	No growth	-
302	Control		111	38	<i>M. abscessus subsp. bolletii</i>	No growth	-
211	Control		68	9	<i>M. simiae</i>	No growth	-
305	Control		31	7	<i>M. avium</i>	No growth	-
225	Control		26	12	<i>M. intracellulare</i>	No growth	-
304	Control		7	6	<i>M. chelonae</i>	No growth	-

Diagnosed species refers to the species isolated in culture leading to the original diagnosis of pulmonary NTM disease. Cultured species refers to the NTM species isolated from the study sample at the time of sequencing. ID = sample identifier.

\* isolated MAC prior to *M. simiae* but not re-isolated for 10 years

<https://doi.org/10.1371/journal.pone.0208018.t003>





**Fig 1. Mycobacterial community structure in samples with more than 100 mycobacterial sequences.** X-axis = samples, Y-axis = relative abundance, colours = species, left panel = cases, right panel = controls.

<https://doi.org/10.1371/journal.pone.0208018.g001>

However, the detection of microbial DNA by sequencing does not necessarily indicate the presence of viable organisms and it is also possible that the recovery of some sequences, particularly those from species not known to cause disease, may originate from non-viable inhaled organisms. The control group was composed of subjects with bronchiectasis, a known risk factor for NTM disease, and it may be that the ubiquitous finding of NTM is restricted to this particular clinical phenotype. A useful avenue for further work will be to determine if NTM are present in the lungs of subjects with other chronic lung diseases, or even in healthy individuals.

Despite their ubiquity in the environment, mycobacteria have not been frequently reported in previous studies of the lung microbiome. Their abundance may however be underestimated for a variety of reasons. In contrast to many other taxa most mycobacteria possess only a single copy of the 16S rRNA gene [24], their protective cell wall is particularly resistant to lysis for DNA extraction [25] and a number of primers targeting the 16S rRNA gene do not comprehensively cover the genus [26].

Several culture-based studies have found that the isolation of more than one NTM species from the same individual is not uncommon [4,5,22] and another key finding of this study is

that in nearly all samples mycobacteria were present as a community rather than a single species. This was not detected by sputum culture which failed to detect more than one species in any individual, although the historic isolation of more than one species was seen in over one-third of cases. This is important finding especially as the current paradigm of NTM infection is that of infection with a single species, which along with the radiological pattern of disease defines that individual's diagnosis and governs their treatment.

There was a tendency for one NTM species to dominate the community and this was more marked in NTM cases. Although not statistically significant, cases had lower alpha diversity indices than controls. Diversity indices were also lower in cases with poor lung function but higher in those treated for NTM. This may represent the expansion of one pathogenic species in subjects who develop clinical NTM disease. The differences however were small and dominance by a single species was still frequently seen in controls without disease. In four samples the cultured NTM species was not detected by sequencing; this is highly likely to be due to the very low numbers of mycobacterial sequences (between 1 and 165) recovered from these samples.

The most abundant mycobacteria species sequenced was *M. avium*, in keeping with the results of conventional culture in the study cohort and other published culture-based studies [23,27–29]. Surprisingly the second most common species was *M. simiae*, despite not being detected by culture in any sample and not being a commonly encountered species in the UK or the rest of Europe. The reason for this finding is unclear. There was no clear bias seen in favour of *M. simiae* amplification in the mock community and the bacterium is readily isolated by conventional culture techniques with no special requirements.

Measures of mycobacterial diversity were lower in NTM cases compared to controls, and within NTM cases higher diversity was associated with better lung function and with treatment directed against NTM. These findings parallel those seen in studies of the bacterial microbiome in respiratory disease which have repeatedly found lower bacterial diversity to be associated with more severe disease [30–36].

The only comparable culture independent study employing human clinical samples was reported by Macovei and colleagues [9] who used a nested PCR targeting mycobacteria to perform partial sequencing of the 16S rRNA gene on samples from the oral cavity and upper respiratory tract of healthy subjects. In keeping with the present study multiple taxa of NTM were found to be present in every subject.

The principal limitation of the present study is the recovery of a large amount of non-mycobacterial sequences, which had the effect of significantly reducing the depth of mycobacterial sequencing. Consequently the data may underestimate the true diversity of mycobacteria present, and may fail to detect mycobacteria in some samples with low sequencing depth. The amplification of non-mycobacterial sequences also meant that the primers could not be used for quantitative PCR to estimate the absolute abundance of mycobacteria in the samples. The mock community data also suggests that there is some variation in amplification between species which may influence estimates of community structure.

Sputum samples are subject to oral contamination and may originate from anywhere within the lung, and so potentially may not reflect the community composition at the site of localised disease. However bronchoscopy would not be feasible in many of the study participants who have advanced pulmonary disease and other comorbidities, and bronchoscopic sampling is also not without risk of contamination [37]. Given the known presence of NTM in the oral cavity [9] contamination is a potential source of NTM sequences in the sputum, however in comparison to sputum oral secretions are likely to have contributed only a small fraction of the total biomass sequenced, and reassuringly there was no relationship between amplicon concentration and mycobacterial abundance, which has been shown to be a marker of contamination in 16S data [38].



In the present study subjects were sampled at a single time point and only one sputum sample was used for sequencing. Future work will analyse successive sputum samples at a single time point to assess the reproducibility of the community structure, as well as longitudinal sampling to characterise the resilience and responsiveness of the community to changes in clinical state over time.

Finally, the study group was heterogeneous and composed of individuals with different pulmonary conditions receiving different treatments, including antibiotics and therapy for NTM. There are differences between NTM cases and the control group, notably the use of NTM therapy which was by definition absent in the control group, and a lack of subjects with COPD in the control group as sufficient sputum could not be obtained for sequencing in these subjects. These differences may confound analyses between groups. Furthermore, the small sample size means the analysis of diversity metrics is underpowered. Therefore these particular findings should be treated as exploratory.

## Conclusions

This study represents the first use of multiplexed next generation partial sequencing of the *hsp65* gene to characterise mycobacterial communities in the lung. This found that NTM were present in all study subjects regardless of their isolation by conventional culture or the presence of clinical NTM disease, and that multiple species of NTM were usually present in the same individual.

## Supporting information

### S1 File. Supporting information.

(DOCX)

### S1 Table. NTM typestrains included in the mock community.

(DOCX)

### S2 Table. Primers used for *hsp65* sequencing.

(DOCX)

### S3 Table. The abundance of genera in the study population.

(DOCX)

### S4 Table. Isolation of multiple NTM species in study subjects.

(DOCX)

### S1 Fig. Selection of minimum sequencing depth.

(PDF)

### S2 Fig. Structure of the NTM mock community.

(PDF)

### S3 Fig. Mycobacterial abundance and *hsp65* amplicon concentration prior to library pooling.

(PDF)

### S4 Fig. Genus level community structure by sample.

(PDF)

### S5 Fig. Mycobacterial alpha diversity between controls (blue) and cases (orange).

(PDF)

**S6 Fig. Mycobacterial alpha diversity between untreated (blue) and treated (orange) cases.**

(PDF)

**S7 Fig. Mycobacterial alpha diversity and forced expiratory volume in 1 second (FEV<sub>1</sub>) in NTM cases.**

(PDF)

**S8 Fig. Proportion of mycobacterial sequences and discordance between sequencing and culture.**

(PDF)

## Acknowledgments

Just as the advent of culture independent techniques has revealed the unprecedented diversity of the lung microbiome, and a new appreciation of the complex relationship between bacterial communities in the lung and clinical disease, these findings suggest that the relationship between NTM present in the lung and clinical disease may also be more complex than previously thought. Whilst requiring further validation in independent cohorts, our findings potentially have important implications for our understanding of the pathogenesis and also potentially the management of this challenging disease.

## Acknowledgments

The authors would like to thank the staff of the Royal Brompton Hospital Respiratory Clinical Research Facility for their help in the collection of samples and phenotypic data, and Dr Michael Cox and Dr Leah Cuthbertson at the National Heart and Lung Institute, Imperial College London for their assistance with the sequencing.

## Author Contributions

**Conceptualization:** Steven A. Cowman, Miriam F. Moffatt, Michael R. Loebinger.

**Data curation:** Steven A. Cowman, Phillip James.

**Formal analysis:** Steven A. Cowman, Phillip James, William O. C. Cookson, Miriam F. Moffatt.

**Funding acquisition:** William O. C. Cookson, Miriam F. Moffatt, Michael R. Loebinger.

**Investigation:** Steven A. Cowman, Phillip James, Robert Wilson.

**Methodology:** Steven A. Cowman, Phillip James, Miriam F. Moffatt, Michael R. Loebinger.

**Project administration:** Steven A. Cowman, Robert Wilson, Michael R. Loebinger.

**Resources:** Robert Wilson, William O. C. Cookson, Miriam F. Moffatt.

**Software:** William O. C. Cookson, Miriam F. Moffatt.

**Supervision:** Miriam F. Moffatt, Michael R. Loebinger.

**Writing – original draft:** Steven A. Cowman, Michael R. Loebinger.

**Writing – review & editing:** Steven A. Cowman, Phillip James, Robert Wilson, William O. C. Cookson, Miriam F. Moffatt, Michael R. Loebinger.

## References

1. Stout JE, Koh W-J, Yew WW. Update on pulmonary disease due to non-tuberculous mycobacteria. *Int J Infect Dis*. Elsevier; 2016; 45: 123–134. <https://doi.org/10.1016/j.ijid.2016.03.006> PMID: 26976549
2. Strollo SE, Adjemian J, Adjemian MK, Prevots DR. The Burden of Pulmonary Nontuberculous Mycobacterial Disease in the United States. *Ann Am Thorac Soc*. American Thoracic Society; 2015; 12: 1458–1464. <https://doi.org/10.1513/AnnalsATS.201503-173OC> PMID: 26214350
3. Marangu D, Devine B, John-Stewart G. Diagnostic accuracy of nucleic acid amplification tests in urine for pulmonary tuberculosis: a meta-analysis. *Int J Tuberc Lung Dis*. 2015; 19: 1339–1347. <https://doi.org/10.5588/ijtld.15.0209> PMID: 26467586
4. Lim H-J, Park CM, Park YS, Lee J, Lee S-M, Yang S-C, et al. Isolation of multiple nontuberculous mycobacteria species in the same patients. *Int J Infect Dis*. 2011; 15: e795–8. <https://doi.org/10.1016/j.ijid.2011.07.004> PMID: 21873093
5. Griffith DE, Phillely JV, Brown-Elliott BA, Benwill JL, Shepherd S, York D, et al. The significance of *Mycobacterium abscessus* subspecies *abscessus* isolation during *Mycobacterium avium* complex lung disease therapy. *Chest*. 2015; 147: 1369–1375. <https://doi.org/10.1378/chest.14-1297> PMID: 25357074
6. Angenent LT, Kelley ST, St Amand A, Pace NR, Hernandez MT. Molecular identification of potential pathogens in water and air of a hospital therapy pool. *Proc Natl Acad Sci USA*. 2005; 102: 4860–4865. <https://doi.org/10.1073/pnas.0501235102> PMID: 15769858
7. van der Wielen PWJJ, van der Kooij D. Nontuberculous mycobacteria, fungi, and opportunistic pathogens in unchlorinated drinking water in The Netherlands. *Appl Environ Microbiol*. 2013; 79: 825–834. <https://doi.org/10.1128/AEM.02748-12> PMID: 23160134
8. Pontiroli A, Khera TT, Oakley BB, Mason S, Dowd SE, Travis ER, et al. Prospecting environmental mycobacteria: combined molecular approaches reveal unprecedented diversity. Hoshino Y, editor. *PLoS ONE*. 2013; 8: e68648. <https://doi.org/10.1371/journal.pone.0068648> PMID: 23874704
9. Macovei L, McCafferty J, Chen T, Teles F, Hasturk H, Paster BJ, et al. The hidden “mycobacteriome” of the human healthy oral cavity and upper respiratory tract. *J Oral Microbiol*. 2015; 7: 26094. <https://doi.org/10.3402/jom.v7.26094> PMID: 25683180
10. Griffith DE, Aksamit T, Brown-Elliott BA, Catanzaro A, Daley C, Gordin F, et al. An Official ATS/IDSA Statement: Diagnosis, Treatment, and Prevention of Nontuberculous Mycobacterial Diseases. *Am J Respir Crit Care Med*. 2007; 175: 367–416. <https://doi.org/10.1164/rccm.200604-571ST> PMID: 17277290
11. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol*. 2013; 79: 5112–5120. <https://doi.org/10.1128/AEM.01043-13> PMID: 23793624
12. Telenti A, Marchesi F, Balz M, Bally F, Böttger EC, Bodmer T. Rapid identification of mycobacteria to the species level by polymerase chain reaction and restriction enzyme analysis. *Journal of Clinical Microbiology*. 1993; 31: 175–178. PMID: 8381805
13. Ringuet H, Akoua-Koffi C, Honore S, Varnerot A, Vincent V, Berche P, et al. hsp65 sequencing for identification of rapidly growing mycobacteria. *Journal of Clinical Microbiology*. 1999; 37: 852–857. PMID: 9986875
14. Adékambi T, Reynaud-Gaubert M, Greub G, Gevaudan M-J, La Scola B, Raoult D, et al. Amoebal coculture of “*Mycobacterium massiliense*” sp. nov. from the sputum of a patient with hemoptitic pneumonia. *Journal of Clinical Microbiology*. 2004; 42: 5493–5501. <https://doi.org/10.1128/JCM.42.12.5493-5501.2004> PMID: 15583272
15. Leão SC, Tortoli E, Viana-Niero C, Ueki SYM, Lima KVB, Lopes ML, et al. Characterization of mycobacteria from a major Brazilian outbreak suggests that revision of the taxonomic status of members of the *Mycobacterium chelonae*-*M. abscessus* group is needed. *Journal of Clinical Microbiology*. 2009; 47: 2691–2698. <https://doi.org/10.1128/JCM.00808-09> PMID: 19571015
16. Kim H-Y, Kook Y, Yun Y-J, Park CG, Lee NY, Shim TS, et al. Proportions of *Mycobacterium massiliense* and *Mycobacterium boletii* Strains among Korean *Mycobacterium chelonae*-*Mycobacterium abscessus* Group Isolates. *Journal of Clinical Microbiology*. 2008; 46: 3384–3390. <https://doi.org/10.1128/JCM.00319-08> PMID: 18753344
17. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010; 7: 335–336. <https://doi.org/10.1038/nmeth.f.303> PMID: 20383131
18. Core TR. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.

19. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE*. 2013; 8: e61217. <https://doi.org/10.1371/journal.pone.0061217> PMID: 23630581
20. Oksanen J, Guillaume Blanchet F, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. *vegan*: Community Ecology Package. R package version 2.2-0. 2014.
21. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol*. 2014; 12: 87. <https://doi.org/10.1186/s12915-014-0087-z> PMID: 25387460
22. Cowman S, Burns K, Benson S, Wilson R, Loebinger MR. The antimicrobial susceptibility of non-tuberculous mycobacteria. *J Infect*. 2015. <https://doi.org/10.1016/j.jinf.2015.12.007> PMID: 26723913
23. Russell CD, Claxton P, Doig C, Seagar A-L, Rayner A, Laurenson IF. Non-tuberculous mycobacteria: a retrospective review of Scottish isolates from 2000 to 2010. *Thorax*. 2014; 69: 593–595. <https://doi.org/10.1136/thoraxjnl-2013-204260> PMID: 23986391
24. Větrovský T, Baldrian P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS ONE*. 2013; 8: e57923. <https://doi.org/10.1371/journal.pone.0057923> PMID: 23460914
25. Hosek J, Svastova P, Moravkova M, Pavlik I, Bartos M. Methods of mycobacterial DNA isolation from different biological material: a review. *Veterinarni Medicina*. 2006; 51: 180–192.
26. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res*. 2013; 41: e1. <https://doi.org/10.1093/nar/gks808> PMID: 22933715
27. Hoefsloot W, van Ingen J, Andréjak C, Angeby K, Bauriaud R, Berner P, et al. The geographic diversity of nontuberculous mycobacteria isolated from pulmonary samples: an NTM-NET collaborative study. *European Respiratory Journal*. 2013; 42: 1604–1613. <https://doi.org/10.1183/09031936.00149212> PMID: 23598956
28. Moore JE, Kruijshaar ME, Ormerod LP, Drobniewski F, Abubakar I. Increasing reports of non-tuberculous mycobacteria in England, Wales and Northern Ireland, 1995-2006. *BMC Public Health*. BioMed Central Ltd; 2010; 10: 612. <https://doi.org/10.1186/1471-2458-10-612> PMID: 20950421
29. Henry MT, Inamdar L, O'riordain D, Schweiger M, Watson JP. Nontuberculous mycobacteria in non-HIV patients: epidemiology, treatment and response. *European Respiratory Journal*. 2004; 23: 741–746. <https://doi.org/10.1183/09031936.04.00114004> PMID: 15176690
30. Einarsson GG, Comer DM, McIlreavey L, Parkhill J, Ennis M, Tunney MM, et al. Community dynamics and the lower airway microbiota in stable chronic obstructive pulmonary disease, smokers and healthy non-smokers. *Thorax*. 2016. <https://doi.org/10.1136/thoraxjnl-2015-207235> PMID: 27146202
31. Garcia-Nuñez M, Millares L, Pomares X, Ferrari R, Pérez Brocal V, Gallego M, et al. Severity-related changes of bronchial microbiome in chronic obstructive pulmonary disease. *Journal of Clinical Microbiology*. 2014; 52: 4217–4223. <https://doi.org/10.1128/JCM.01967-14> PMID: 25253795
32. Flight WG, Smith A, Paisey C, Marchesi JR, Bull MJ, Norville PJ, et al. Rapid Detection of Emerging Pathogens and Loss of Microbial Diversity Associated with Severe Lung Disease in Cystic Fibrosis. McAdam AJ, editor. *Journal of Clinical Microbiology*. American Society for Microbiology; 2015; 53: 2022–2029. <https://doi.org/10.1128/JCM.00432-15> PMID: 25878338
33. van der Gast CJ, Walker AW, Stressmann FA, Rogers GB, Scott P, Daniels TW, et al. Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *ISME J*. 2011; 5: 780–791. <https://doi.org/10.1038/ismej.2010.175> PMID: 21151003
34. Delhaes L, Monchy S, Fréalle E, Hubans C, Salleron J, Leroy S, et al. The airway microbiota in cystic fibrosis: a complex fungal and bacterial community—implications for therapeutic management. Brown SP, editor. *PLoS ONE*. 2012; 7: e36313. <https://doi.org/10.1371/journal.pone.0036313> PMID: 22558432
35. Rogers GB, van der Gast CJ, Cuthbertson L, Thomson SK, Bruce KD, Martin ML, et al. Clinical measures of disease in adult non-CF bronchiectasis correlate with airway microbiota composition. *Thorax*. 2013; 68: 731–737. <https://doi.org/10.1136/thoraxjnl-2012-203105> PMID: 23564400
36. Rogers GB, Zain NMM, Bruce KD, Burr LD, Chen AC, Rivett DW, et al. A Novel Microbiota Stratification System Predicts Future Exacerbations in Bronchiectasis. *Ann Am Thorac Soc*. 2014. <https://doi.org/10.1513/AnnalsATS.201310-335OC> PMID: 24592925
37. Gubler JG, Salfinger M, Graevenitz von A. Pseudoepidemic of nontuberculous mycobacteria due to a contaminated bronchoscope cleaning machine. Report of an outbreak and review of the literature. *Chest*. 1992; 101: 1245–1249. PMID: 1582279
38. Jarvis-Bardy J, Leong LEX, Marri S, Smith RJ, Choo JM, Smith-Vaughan HC, et al. Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data. *Microbiome* 2015; 3:19. <https://doi.org/10.1186/s40168-015-0083-8> PMID: 25969736