Population-level genomics identifies the emergence and global spread of a human transmissible multidrug-resistant nontuberculous mycobacterium.

Josephine M Bryant[1,2]*, Dorothy M Grogono[2,3]*, Daniela Rodriguez-Rincon[2], Isobel Everall[1], Karen P Brown[2,3], Pablo Moreno[4], Deepshikha Verma[5], Emily Hill[5], Judith Drijkoningen[2], Peter Gilligan[6], Charles R Esther[6], Peadar G Noone[6], Olivia Giddings[6], Scott C. Bell[7,8,9], Rachel Thomson[10], Claire E. Wainwright[7,11], Chris Coulter[12], Sushil Pandey[8], Michelle E Wood[13], Rebecca E Stockwell[14], Kay A Ramsay[15], Laura J Sherrard[12], Timothy J Kidd[14,16], Nassib Jabbour[15], Graham R Johnson[17], Luke D Knibbs[18], Lidia Morawska[17], Peter D Sly[13], Andrew Jones[19], Diana Bilton[19], Ian Laurenson[20], Michael Ruddy[21], Stephen Bourke[22], Ian CJW Bowler[23], Stephen J Chapman[23], Andrew Clayton[24], Mairi Cullen[25], Thomas Daniels[25], Owen Dempsey[26], Miles Denton[27], Maya Desai[28], Richard Drew[29], Frank Edenborough[30], Jason Evans[21], Patricia Fenton[30], Jonathan Folb[31], Helen Humphrey[32], Barbara Isalska[25], Søren Jensen-Fangel[33], Bodil Jönsson[34], Andrew M. Jones[25], Terese L Katzenstein[35], Troels Lillebaek[36], Gordon MacGregor[37], Sarah Mayell[30], Michael Millar[38], Deborah Modha[39], Edward F Nash[40], Christopher O'Brien[22], Deirdre O'Brien[41], Chandra Ohri[39], Caroline S Pao[38], Daniel Peckham[28], Felicity Perrin[42], Audrey Perry[22], Tania Pressler[35], Laura Prtak[31], Tavs Qvist[35], Ali Robb[22], Helen Rodgers[43], Kirsten Schaffer[41], Nadia Shafi[3], Jakko van Ingen[44], Martin Walshaw[45], Danie Watson[38], Noreen West[46], Joanna Whitehouse[40], Charles S Haworth[3], Simon R Harris[1], Diane Ordway[5], Julian Parkhill[1]** & R. Andres Floto[2,3]**.

* JMB and DMG contributed equally to this work

1. Wellcome Trust Sanger Institute, Hinxton, UK
2. University of Cambridge Department of Medicine, MRC-Laboratory of Molecular Biology, Cambridge, UK
3. Cambridge Centre for Lung Infection, Papworth Hospital, Cambridge, UK
4. EMBL European Bioinformatics Institute, Hinxton, UK
5. Mycobacteria Research Laboratory, Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins CO, USA.
6. University of North Carolina School of Medicine, NC, USA.
7. School of Medicine, The University of Queensland, Australia
8. QIMR Berghofer Medical Research Institute, Brisbane, Australia
9. The Prince Charles Hospital, Brisbane, Australia
10. Gallipoli Medical Research Centre, University of Queensland, Brisbane, Australia
11. Lady Cilento Children's Hospital, Brisbane
12. Queensland Mycobacterial Reference Laboratory, Brisbane, Australia
13. Child Health Research Centre, The University of Queensland, Brisbane, Australia
14. Centre for Experimental Medicine, Queen's University Belfast, UK

15. Queensland University of Technology, Brisbane, Australia
16. School of Chemistry and Biomolecular sciences, The University of Queensland, Australia
17. International Laboratory for Air Quality and Health, Queensland University of Technology, Brisbane, Australia
18. School of Population Health, The University of Queensland, Brisbane, Australia
19. Royal Brompton and Harefield NHS Foundation Trust, UK
20. Scottish Mycobacteria Reference Laboratory, UK
21. Wales Centre for Mycobacteria, UK
22. The Newcastle upon Tyne Hospitals NHS Foundation Trust, UK
23. Oxford University Hospitals NHS Foundation Trust, UK
24. Nottingham University Hospitals NHS Trust, UK
25. University Hospital of South Manchester NHS Foundation Trust, UK
26. Aberdeen Royal Infirmary, NHS Grampian, Scotland, UK
27. The Leeds Teaching Hospitals NHS Trust, UK
28. Birmingham Children's Hospital NHS Foundation Trust, UK
29. Alder Hey Children's NHS Foundation Trust, UK
30. Sheffield Teaching Hospitals NHS Foundation Trust, UK
31. The Royal Liverpool and Broadgreen University Hospitals NHS Trust, UK
32. University Hospital Southampton NHS Foundation Trust, UK
33. Department of Infectious Diseases, Aarhus University Hospital, Denmark
34. Department of Infectious Medicine, Institute of Biomedicine, University of Gothenburg, Sweden
35. Copenhagen Cystic Fibrosis Center, Department of Infectious Diseases, Rigshospitalet, Copenhagen, Denmark
36. International reference Laboratory of Mycobacteriology, Statens Serum Institut, Copenhagen, Denmark
37. Gartnavel Hospital, Glasgow, NHS Greater Glasgow and Clyde, Scotland, UK
38. Barts Health NHS Trust, London, UK
39. University Hospitals of Leicester NHS Trust, UK
40. Heart of England NHS Foundation Trust, Birmingham, UK
41. St Vincent's University Hospital, Dublin, Ireland
42. King's College Hospital NHS Foundation Trust, London, UK
43. Western General Hospital, NHS Lothian, Scotland, UK
44. Department of Medical Microbiology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands
45. Liverpool Heart and Chest Hospital NHS Foundation Trust, UK
46. Sheffield Children's NHS Foundation Trust, UK

**Correspondence:

R. Andres Floto, University of Cambridge Department of Medicine, MRC Laboratory of Molecular Biology, Cambridge, CB2 0QH, United Kingdom (+44 1223 768801), arf27@cam.ac.uk

Or

Julian Parkhill, Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, United Kingdom (+44 1223 494975), parkhill@sanger.ac.uk

ABSTRACT (117 words)

Lung infections with *Mycobacterium abscessus,* a species of multidrug resistant nontuberculous mycobacteria, are emerging as an important global threat to individuals with cystic fibrosis (CF) where they accelerate inflammatory lung damage leading to increased morbidity and mortality. Previously, *M. abscessus* was thought to be independently acquired by susceptible individuals from the environment. However, using whole genome analysis of a global collection of clinical isolates, we show that the majority of *M. abscessus* infections are acquired through transmission, potentially *via* fomites and aerosols, of recently emerged dominant circulating clones that have spread globally. These clones are associated with worse clinical outcomes, show increased virulence in cell-based and mouse infection models, and thus represent an urgent international infection challenge.

MAIN TEXT

Nontuberculous mycobacteria (NTM; referring to mycobacterial species other than *M. tuberculosis* complex and *M. leprae*) are ubiquitous environmental organisms that can cause chronic pulmonary infections in susceptible individuals *[1, 2]*, particularly those with pre-existing inflammatory lung diseases such as cystic fibrosis (CF) *[3]*. The major NTM infecting CF individuals around the world is *Mycobacterium abscessus;* a rapidly growing, intrinsically multidrug-resistant species, which can be impossible to treat despite prolonged combination antibiotic therapy *[1, 3-5]*, leads to accelerated decline in lung function *[6,7]*, and remains a contraindication to lung transplantation in many centers *[3,8,9]*.

Until recently, NTM infections were thought to be independently acquired by individuals through exposure to soil or water *[10-12].* As expected, previous analyses from the 1990s and 2000s *[13-16]* showed that CF patients were infected with unique, genetically diverse strains of *M. abscessus,* presumably from environmental sources. We used whole genome sequencing at a single UK CF center and identified two clusters of patients (11 individuals in total) infected with identical or near-identical *M. abscessus* isolates, which social network analysis suggested were acquired within hospital *via* indirect person-to-person transmission *[17]*; a possibility further supported by genomic sequencing *[18]* of a separate *M. abscessus* outbreak in a Seattle CF center *[19]*.

Given the increasing incidence of *M. abscessus* infections in CF and non-CF populations reported globally *[3, 20, 21]*, we investigated whether cross-infection, rather than independent environmental acquisition, might be the major source of infection for this organism and therefore undertook population-level, multinational, whole genome sequencing of *M. abscessus* isolates from infected CF patients, correlating results with clinical metadata and phenotypic functional analysis of isolates.

We generated whole genome sequences for 1080 clinical isolates of *M. abscessus* from 517 patients, obtained from UK CF clinics and their associated regional reference laboratories, as well as CF Centres in the US (UNC Chapel Hill), the Republic of Ireland (Dublin), mainland Europe (Denmark, Sweden, The Netherlands), and Australia (Queensland). We identified 730 isolates as *M. a. abscessus*, 256 isolates as *M. a. massiliense*, 91 isolates as *M. a. bolletii,* with three isolates (from 3 different patients) containing more than one subspecies.

Phylogenetic analysis of these sequences (using one isolate per patient), supplemented by published genomes from US, France, Brazil, Malaysia, China, and South Korea (**Table S1**), was performed and analysed in the context of the geographical provenance of isolates (**Figure 1; Figure S1**). As done previously *[17]*, we obtained maximum likelihood phylogenetic trees demonstrating separation of *M. abscessus* into three clearly divergent subspecies (*M. a. abscessus, M. a. bolletii,* and *M. a. massiliense*), challenging recent reclassifications of *M. abscessus* into only two subspecies *[22]*.

Within each subspecies, we found multiple examples of deep branches (indicating large genetic differences) between isolates from different individuals, consistent with independent acquisition of unrelated environmental bacteria. However, we also identified multiple clusters of near-identical isolates from geographically diverse locations (**Figure 1**), suggesting widespread transmission of circulating clones within the global CF patient community.

To investigate further the relatedness of isolates from different individuals, we analysed each subspecies for the presence of high density phylogenetic clades (see Supplementary Methods *[23]*). We identified multiple dense clusters of isolates, predominantly within the *M. a. abscessus* and *M. a. massiliense* subspecies (**Figure 2A**), indicating the presence of dominant circulating clones. We next excluded clusters found in only one CF centre from further analysis to remove related isolates that might have been acquired from a local environmental point source. We found that most patients (74%) were infected with clustered, rather than unclustered, isolates, principally from *M. a. abscessus* Cluster 1 and 2, and *M. a. massiliense* Cluster 1 (**Figure 2B**). The median branch lengths of almost all clusters found in two or more CF centers was less than 20 SNPs (range 1-175 SNPs), indicating a high frequency of identical or near identical isolates infecting geographically separate individuals.

To determine how much of the genetic relatedness found within clusters was attributable to recent person-to-person transmission, we first examined the within-patient genetic diversity of *M. abscessus* isolates from single individuals. In keeping with our previously published results *[17]*, we found that 90% of same-patient isolates differed by less than 20 SNPs, while 99% of same-patient isolates differed by less than 38 SNPs (**Figure S2**). We therefore classified isolates from different individuals varying by less than 20 SNPs as indicating 'probable*',* and those varying by 20-38 SNPs as indicating 'possible'*,* recent direct or indirect transmission. We thereby

identified multiple episodes of likely recent transmission in virtually all multi-site clusters of *M. abscessus* (**Figure 2B),** and across the majority of CF centers (**Figure S3**).

We next examined the global distribution of clustered isolates and found that, in all countries, the majority of patients were infected with clustered rather than unclustered isolates (**Figure 2C**), suggesting frequent and widespread infection of patients with closely related isolates. Moreover, the three dominant circulating clones, *M. abscessus* Clusters 1 and 2, and *M. massiliense* Cluster 1, were all represented in the USA, European, and Australian collections of clinical isolates, indicating trans-continental dissemination of these clades.

We then compared the genetic differences between isolates (measured by pairwise SNP distance) as a function of geography. As expected from our previous detection of hospital-based transmission of *M. abscessus [17]*, average genetic distances were significantly shorter for *M. abscessus* isolates from the same CF center than those from different CF centers within the same country or from different countries (**Figure 2D**). However, we also detected numerous examples of identical or near-identical isolates infecting groups of patients in different CF centers and, indeed, across different countries (**Figure 2D**), indicating the recent global spread of *M. abscessus* clones throughout the international CF patient community.

We applied Bayesian analysis *[24]* to date the establishment and spread of dominant circulating clones, focusing on *M. a. massiliense* Cluster 1, which includes isolates from both the Seattle *[19]* and Papworth *[17]* CF Center outbreaks, as well as isolates from CF centres across England (Birmingham, London, Leicester), Scotland (Lothian, Glasgow), Ireland (Dublin), Denmark (Copenhagen), Australia (Queensland), and the USA (Chapel Hill, NC) (**Figure 3A**). We estimate that the most recent common ancestor of isolates infecting patients from all these locations emerged around 1978 (95% CI: 1955-1995), clearly indicating recent global dissemination of this dominant circulating clone amongst individuals with CF (**Figure 3A**).

Furthermore we were able to resolve individual transmission events between patients infected with dominant circulating clones through two orthogonal approaches. Firstly, using high-depth genomic sequencing of colony sweeps, we were able to track changes in within-patient bacterial diversity in sputum cultures of a single individual over time. By linking the frequency of occurrence of minority variants in longitudinal samples, we were able to define the presence of particular subclones within infected

individuals, assign their likely evolutionary development (involving the successive acquisition of non-synonymous mutations in likely virulence genes; **Figure 3B**), monitor their relative frequencies over time, and demonstrate their transmission between patients (**Figure 3B**). Secondly, through longitudinal whole genome sequencing of isolates collected over time from individuals, we were able to find multiple examples of the complete nesting of one patient's sampled diversity within another's (**Figure S4**). Such paraphyletic relationships are strongly indicative of recent person-to-person transmission *[25]*.

We next examined potential mechanisms of transmission of *M. abscessus* between individuals (which our previous epidemiological analysis had suggested was indirect rather than *via* direct contact between patients *[17]*). We found support for fomite spread of *M. abscessus* (detecting three separate transmission events associated with surface contamination of an inpatient room by an individual infected with a dominant circulating clone; **Figure S5**), and also for potential airborne transmission (by experimentally demonstrating the generation of long-lived, potentially infectious cough aerosols by an infected CF patient; **Figure S6**).

A potential explanation for the emergence of dominant clones of *M. abscessus* is that they are more efficient at infection and/or transmission. We therefore analysed clinical metadata to establish whether outcomes were different for patients infected with clustered rather than unclustered isolates. We correlated clinical outcomes with bacterial phylogeny and the presence of constitutive resistance to two key NTM antibiotics, amikacin and macrolides *[26, 27]*, acquired through point mutations in the 16S and 23S ribosomal RNA respectively (**Figure 4A**). We found no differences in the proportions of *M.* abscessus-positive individuals diagnosed with ATS-defined NTM pulmonary disease *[1]*, (namely the presence of two or more culture-positive sputum samples with NTM-associated symptoms and radiological changes), but did observe increased rates of chronic infection in individuals infected with clustered rather than unclustered isolates (**Figure 4B**). As anticipated for transmissible clones exposed to multiple rounds of antibiotic therapy, we also found high rates of constitutive amikacin and/or macrolide resistance in clustered isolates (**Figure 4B**). Of note, resistance to these two antibiotics did not itself correlate with poor clinical outcomes, suggesting that additional bacterial factors might contribute to worse responses in patients infected with clustered isolates (**Figure S7**).

To explore differences in intrinsic virulence between clustered and unclustered *M. abscessus*, we subjected a panel of representative isolates (27 clustered and 17

unclustered *M. a. abscessus*; 25 clustered and 13 unclustered *M. a. massiliense*) to a series of *in vitro* phenotypic assays. While we found no or only minor differences between groups in their colony morphotype, biofilm formation, ability to trigger cytokine release from macrophages (**Figure S8**) and their overall phenotypic profile (by multifactorial analysis; **Figure S9**), we detected significantly increased phagocytic uptake (**Figure 4C**) and intracellular survival in macrophages (**Figure 4D**) of clustered isolates of both *M. a. abscessus* and *M. a. massiliense* compared to unclustered controls, indicating clear differences in pathogenic potential. Moreover, infection of *SCID* mice revealed significantly greater bacterial burden (**Figure 4E**) and granulomatous inflammation (**Figure 4F**) following inoculation with clustered rather than unclustered isolates of *M. a. abscessus* and *M. a. massiliense*, confirming differences in virulence between these groups.

In summary, our results reveal that the majority of *M. abscessus* infections of individuals with CF worldwide are caused by genetically-clustered isolates, suggesting recent person-to-person transmission, rather than through independent acquisition of genetically-unrelated environmental organisms. Given the widespread implementation of individual and cohort segregation of patients in CF centres in Europe *[28]*, the USA *[29]*, and Australia *[30]* (which have led to falling levels of MRSA, *Burkholderia*, and transmissible *Pseudomonas* infections *[31-33]*), we believe that the likely mechanism of local spread of *M. abscessus* is through the generation of long-lived infectious aerosols (as identified for other CF pathogens *[34-36]*) or via fomite spread; both routes are plausible given our findings (**Figures S5, S6**), and would be potentially enhanced by the intrinsic desiccation resistance of *M. abscessus*. Such indirect transmission, involving environmental contamination by patients, is supported by our previous social network analysis of a UK outbreak of *M. abscessus [17]* in CF patients, which revealed hospital-based cross-infection without direct person-to-person contact, and by the termination of a Seattle *M. abscessus* outbreak associated with the introduction of clinic room negative pressure ventilation and double room cleaning *[19]*. The long-distance spread of circulating clones is more difficult to explain; with no supporting evidence for the movement of potentially infected CF patients or contaminated equipment between countries, or for zoonotic spread of *M. abscessus*, one possibility is that this may be driven by healthy individuals acting as vectors of transmission.

Our study illustrates the power of population-level genomics to uncover modes of transmission of emerging pathogens and has revealed the recent emergence of global dominant circulating clones of *M. abscessus* that have spread between

continents. These clones are better able to survive within macrophages, cause more virulent infection in mice, and are associated with worse clinical outcomes, suggesting that person-to-person transmission may have permitted multiple rounds of within-host genetic adaptation to allow *M. abscessus* to evolve from an environmental organism to a true lung pathogen.

Figure legends

Figure 1. Global phylogeny of clinical isolates of M. abscessus.

Maximum likelihood phylogenetic tree of clinical isolates of *M. abscessus* collected with relevant local and/or national Ethical Board approval from 517 patients (using one isolate per patient), obtained from UK CF clinics and their associated regional reference laboratories, CF Centres in the US (UNC Chapel Hill), the Republic of Ireland (Dublin), mainland Europe (Denmark, Sweden, The Netherlands), and Australia (Queensland), supplemented by published genomes from US, France, Brazil, Malaysia, China, and South Korea (listed in **Table S1**).

Figure 2. Transcontinental spread of dominant circulating clones.

(**A**). Hierarchical branch density analysis of phylogenetic trees for each subspecies of *M. abscessus* identifies multiple clusters of closely related isolates predominantly within the *M. a. abscessus* and *M. a. massiliense* subspecies (numbered, and spectrally coloured red to blue, from most densely clustered to least; black indicating no significant clustering). (**B**). Analysis of *M. abscessus* clusters found in two or more CF centers showing (*top*) numbers of patients infected with each cluster (*grey bars*) or unclustered isolates (*green)* and median branch length (SNPs) of different patients' isolates within each cluster (*blue circles*); (bottom) numbers of recent transmission events classified as probable (*red*; < 20 SNPs difference between isolates) or possible (*yellow*; 20-38 SNPs difference between isolates). (**C**) Global distribution of clustered *M. abscessus* isolates showing *M. a. abscessus* Cluster 1 (*red*) and Cluster 2 (*green*), *M. a. massiliense* Cluster 1 (*blue*), other clustered isolates (*yellow*) and unclustered isolates (*black*) with numbers of patients (*n*) sampled per location. (**D**) Genetic differences between isolates (measured by pairwise SNP distance) from different patients attending the same CF center, different CF centers within the same country, or CF centers in different countries (boxes indicate median and interquartile range; *p* values obtained from Mann Whitney Rank Sum tests). To exclude multiple highly distant comparisons, for each isolate only the smallest pairwise distance with an isolate from another patient is included. Colour coding indicates whether recent transmission was probable (*red*; < 20 SNPs difference between isolates), possible (*yellow*; 20-38 SNPs difference between isolates) or unlikely (> 38 SNPs difference between isolates).

Figure 3. Dating the emergence of dominant circulating clones.

(**A**). Dating the emergence of the *M. a. massiliense* Cluster 1 (responsible for the Papworth and Seattle CF center outbreaks), using Bayesian analysis, with

geographical annotation of isolates within the cluster. (**B**). Predicted evolution of subclones (identified through minority variant linkage; ***see Supplementary Methods [23]***) within a single patient with CF (Patient 2 from *Ref. 19*) chronically infected with the dominant circulating clone *Massiliense* Cluster 1 (representative of a total of 11 patients studied). (i) Analysis revealed successive acquisition of non-synonymous polymorphisms (NS) by the most common recent ancestral clone (MRCA; white) in potential virulence genes (UBiA, MAB_0173; Crp/Fnr, MAB_0416c; mmpS, MAB_0477; PhoR, MAB_0674) and then transmission of a single subclone to another patient from the same CF center (Patient 28 from *Ref. 19*). (ii) Frequency of each subclone within longitudinal sputum isolates analysed during the course of Patient 2's infection and the subsequent transmission of a subclone to Patient 28. We observed considerable heterogeneity in the detected repertoire of subclones within each sputum sample (vertical rectangles coloured to illustrate the proportion of detected subclones coded as for (i) in each sputum sample), reflecting either temporal fluctuations in dominant sub-lineages or variable sampling of geographical diversity of subclones within the lung (as previously described for *P. aeruginosa [37]*). Previously determined opportunities for hospital-based cross-infection between the two patients (using social network and epidemiologic analysis *[17]*), are shown in grey vertical bars.

Figure 4. Comparison of clinical outcomes and functional phenotyping of clustered and unclustered *M. abscessus* isolates.

(**A, B**). Relationship of phylogeny with clinical metadata. Phylogenetic tree of *M. abscessus* isolates (one isolate per patient) with dominant circulating clones *M. a. abscessus* 1 (*Absc 1*), *abscessus* 2 (*Absc 2*), and *M. a. massiliense* (*Mass 1*) highlighted (*grey*). For each isolate, clinical data (where available) was used to determine whether (*column 1*) the infected patient fulfilled the ATS/IDSA criteria for NTM pulmonary disease, namely the presence of two or more culture-positive sputum samples with NTM-associated symptoms and radiological changes *[1]* (yes: *blue;* no: *orange*); whether (*column 2*) patients culture converted (*green*) or remained chronically infected (*red*) with *M. abscessus* ; and whether (*column 3*) isolates have acquired amikacin resistance (through 16S rRNA mutations; *red*), macrolide resistance (through 23S rRNA mutations; *yellow*), or both (*orange*). (**C, D**). *In vitro* phenotyping of representative isolates of clustered (*blue*) and unclustered (*green*) *M. a. abscessus* and clustered (*red*) and unclustered (*yellow*) *M. a. massiliense* comparing phagocytosis by (C) and intracellular survival (normalised for uptake)

within (D) differentiated THP1 cells. Data points represent averages of at least three independent replicates. (**E, F**) Using SCID mice, infection with clustered *M. a. abscessus* (*blue*) and *M. a. massiliense* (*red*) led to (E) greater intracellular survival within (i) bone marrow-derived macrophages *in vitro* and (ii) higher bacterial burdens in lung and spleen with (F) worse granulomatous lung inflammation (*arrowheads*), than unclustered controls (*M. a. abscessus green; M. a. massiliense yellow*) *scale bar*. CFU data is shown as mean ± *sem*; * $p < 0.05$; ** $p < 0.005$ (two-tailed unpaired Student's t-test).

References

1. D. E. Griffith *et al.,* An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. *Am J Respir Crit Care Med.* **175,** 367-416 (2007).
2. K. L. Winthrop *et al.,* Pulmonary nontuberculous mycobacterial disease prevalence and clinical features: an emerging public health disease. *Am J Respir Crit Care Med.* **82,** 977-82 (2010).
3. R. A. Floto *et al.,* US Cystic Fibrosis Foundation and European Cystic Fibrosis Society consensus recommendations for the management of non-tuberculous mycobacteria in individuals with cystic fibrosis. *Thorax* **71** Suppl 1**,** i1-i22 (2016).
4. K. Jeon *et al.,* Antibiotic treatment of Mycobacterium abscessus lung disease: a retrospective analysis of 65 patients. *Am J Respir Crit Care Med* **180**, 896-902 (2009).
5. J. Jarand *et al.,* Clinical and microbiologic outcomes in patients receiving treatment for Mycobacterium abscessus pulmonary disease. *Clin Infect Dis.* **52,** 565-71 (2011).
6. C. R. Esther Jr *et al.,* Chronic Mycobacterium abscessus infection and lung function decline in cystic fibrosis. *J Cyst Fibros*. **9**, 117-23 (2010).
7. C. R. Esther Jr *et al.,* Nontuberculous mycobacterial infection in young children with cystic fibrosis. *Pediatr Pulmonol*. **40**, 39-44 (2005).
8. M. Sanguinetti *et al.,* Fatal pulmonary infection due to multidrug-resistant Mycobacterium abscessus in a patient with cystic fibrosis. *J Clin Microbiol*. **39,** 816-9 (2001).
9. J. L. Taylor, S.M. Palmer. Mycobacterium abscessus chest wall and pulmonary infection in a cystic fibrosis lung transplant recipient. *J Heart Lung Transplant* **25**, 985-8 (2006).

10. J. O. Falkinham 3rd. Environmental sources of nontuberculous mycobacteria. *Clin Chest Med.* **36**, 35-41 (2015).
11. J. van Ingen *et al.,* Environmental sources of rapid growing nontuberculous mycobacteria causing disease in humans. *Clin Microbiol Infect*. **15**, 888-93 (2009).
12. R. Thomson *et al.,* Isolation of nontuberculous mycobacteria (NTM) from household water and shower aerosols in patients with pulmonary disease caused by NTM. *J. Clin Microbiol* **51**, 3006-11.
13. K. N. Olivier *et al.,* Nontuberculous mycobacteria. I: multicenter prevalence study in cystic fibrosis. *Am J Respir Crit Care Med.* **167**, 828-34 (2003).
14. I. Sermet-Gaudelus *et al.,* Mycobacterium abscessus and children with cystic fibrosis. *Emerg Infect Dis*. **9**, 1587-91 (2003).
15. F.C. Bange *et al.,* Lack of transmission of mycobacterium abscessus among patients with cystic fibrosis attending a single clinic. *Clin Infect Dis.* **32**, 1648-50 (2001).
16. B. E. Jonsson *et al.,* Molecular epidemiology of Mycobacterium abscessus, with focus on cystic fibrosis. *J Clin Microbiol* **45**, 1497-594 (2007).
17. J. M. Bryant *et al.,* Whole-genome sequencing to identify transmission of Mycobacterium abscessus between patients with cystic fibrosis: a retrospective cohort study. *Lancet* **381**, 1551-60 (2013).
18. H. Tettelin *et al.,* High-level relatedness among Mycobacterium abscessus subsp. massiliense strains from widely separated outbreaks. *Emerg. Infect. Dis.* **20**, 364-71 (2014).
19. M. L. Aitken *et al.,* Respiratory outbreak of Mycobacterium abscessus subspecies massiliense in a lung transplant and cystic fibrosis center. *Am J Respir Crit Care Med.* **185**, 231-2 (2012).
20. D. R. Prevots *et al.,* Nontuberculous mycobacterial lung disease prevalence at four integrated health care delivery systems. *Am J Respir Crit Care Med*. **182**, 970-6 (2010).
21. C. Pierre-Audigier *et al.,* Age-related prevalence and distribution of nontuberculous mycobacterial species among patients with cystic fibrosis. *J Clin Microbiol.* **43**, 3467-70 (2005).
22. S.C. Leao *et al.,* Proposal that Mycobacterium massiliense and Mycobacterium bolletii be united and reclassified as Mycobacterium abscessus subsp. bolletii comb. nov., designation of Mycobacterium abscessus subsp. abscessus subsp. nov. and emended description of Mycobacterium abscessus. *Int J Syst Evol Microbiol*. **61**, 2311-3 (2011).

431 23. Materials and methods are available as Supplementary Online Materials
432 24. A. J. Drummond, A. Rambaut. BEAST: Bayesian evolutionary analysis by
433 sampling trees. *BMC Evol Biol*. **7**, 214 (2007).
434 25. E. O. Romero-Sevenson, I Bulla, T Leitner Phylogenetically resolving
435 epidemiologic linkage. *PNAS* **113**, 2690-5 (2016).
436 26. T. Prammananan *et al.,* A single 16S ribosomal RNA substitution is
437 responsible for resistance to amikacin and other 2-deoxystreptamine
438 aminoglycosides in Mycobacterium abscessus and Mycobacterium chelonae.
439 *J Infect Dis.* **177**, 1573-81 **(**1998).
440 27. R. J. Wallace Jr. et al., Genetic basis for clarithromycin resistance among
441 isolates of Mycobacterium chelonae and Mycobacterium abscessus.
442 *Antimicrob Agents Chemother*. **40**, 1676-81 (1996).
443 28. S. Conway *et al.,* European Cystic Fibrosis Society Standards of Care:
444 Framework for the Cystic Fibrosis Centre. *J Cyst Fibros.* **13** Suppl 1, S3-22
445 (2014).
446 29. L. Saiman *et al.,* Infection prevention and control guideline for cystic fibrosis:
447 2013 update. *Infect Control Hosp Epidemiol*. **35** Suppl 1, S1-67 (2014).
448 30. S. C. Bell, P. J. Robinson. Cystic Fibosis Standards of Care, Australia.
449 https://www.rah.sa.gov.au/thoracic/whats_new/documents/CFA_Standards_o
450 f_Care_journal_31_Mar_08.pdf (2008).
451 31. S. J. Doe *et al.,* Patient segregation and aggressive antibiotic eradication
452 therapy can control methicillin-resistant Staphylococcus aureus at large cystic
453 fibrosis centres. *J Cyst Fibros*. **9**, 104-9 (2010).
454 32. M.W. France *et al.,* The changing epidemiology of Burkholderia species
455 infection at an adult cystic fibrosis centre. *J Cyst Fibros*. **7**, 368-72 (2008).
456 33. A.M. Jones *et al.,* Prospective surveillance for Pseudomonas aeruginosa
457 cross-infection at a cystic fibrosis center. *Am J Respir Crit Care Med.* **171**,
458 257-260 (2005).
459 34. C. E. Wainwright *et al.,* Cough-generated aerosols of Pseudomonas
460 aeruginosa and other Gram-negative bacteria from patients with cystic
461 fibrosis. *Thorax* **64**, 926-31 (2009).
462 35. L. D. Knibbs *et al.,* Viability of *Pseudomonas aeruginosa* in cough aerosols
463 generated by persons with cystic fibrosis. *Thorax* **69**, 740-5 (2014).
464 36. S. Panagea *et al.,* Environmental contamination with an epidemic strain of
465 Pseudomonas aeruginosa in a Liverpool cystic fibrosis centre, and study of its
466 survival on dry surfaces. *J Hosp Infect*. **59**, 102-7 (2005).

37. P. Jorth *et al .,* Regional isolation drives bacterial diversification within Cystic Fibrosis Lungs*. Cell Host Microbe* **18**, 307-19 (2015).
38. H. Ponstingl. SMALT, http://www.sanger.ac.uk/science/tools/smalt-0 (Wellcome Trust Sanger Institute, Hinxton), accessed. 2012.
39. F. Ripoll et al., Non mycobacterial virulence genes in the genome of the emerging pathogen Mycobacterium abscessus. *PloS one* **4**, e5660 (2009).
40. H. Li et al., The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079 (2009).
41. S. R. Harris et al., Evolution of MRSA during hospital transmission and intercontinental spread. *Science* **327**, 469-474 (2010).
42. A. Stamatakis, RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690 (2006).
43. M. Richter, R. Rosselló-Móra, Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* **106**, 19126-19131 (2009).
44. D. R. Zerbino, Using the Velvet de novo assembler for short-read sequencing technologies. *Curr Protoc Bioinformatics* Chapter 11, Unit 11.15 (2010).
45. L. Cheng, T. R. Connor, J. Siren, D. M. Aanensen, J. Corander, Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Mol Biol Evol* **30**, 1224-1228 (2013).
46. S. Harris. (2016), https://github.com/simonrharris/tree_gubbins.
47. N. J. Croucher et al., Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* **43**, e15 (2015).
48. D. A. J. Rambaut A, Tracer v1.5, Available from http://beast.bio.ed.ac.uk/Tracer. (2007).
49. C. Firth et al., Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. *Mol Biol Evol* **27**, 2038-2051 (2010).
50. L. Hepburn et al., Innate immunity. A Spaetzle-like role for nerve growth factor β in vertebrate immunity to Staphylococcus aureus. *Science* **346**, 641-646 (2014).
51. M. Schiebler et al., Functional drug screening reveals anticonvulsants as enhancers of mTOR-independent autophagic killing of Mycobacterium tuberculosis through inositol depletion. *EMBO Mol Med* **7**, 127-139 (2015).
52. S. Lê, J. Josse, F. Husson, FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software* **25**, 1-18 (2008).

53. S. Shang et al., Increased virulence of an epidemic strain of Mycobacterium massiliense in mice. *PLoS One* **6**, e24726 (2011).
54. D. Ordway et al., Animal model of Mycobacterium abscessus lung infection. *J Leukoc Biol* **83**, 1502-1511 (2008).
55. A. Obregón-Henao et al., Susceptibility of Mycobacterium abscessus to antimycobacterial drugs in preclinical models. *Antimicrob Agents Chemother* **59**, 6904-6912 (2015).
56. D. Ordway et al., Influence of Mycobacterium bovis BCG Vaccination on Cellular Immune Response of Guinea Pigs Challenged with Mycobacterium tuberculosis. *Cellul. Vaccine Immunol* **15**, 1248-1258 (2008).
57. D. Ordway et al., The hypervirulent Mycobacterium tuberculosis strain HN878 induces a potent TH1 response followed by rapid down-regulation. *J. Immunol*. **179.**, 522-531 (2007).
58. Y. Saito *et al.*, Characterization of endonuclease III (nth) and endonuclease VIII (nei) mutants of Escherichia coli K-12. *J Bacteriol* **179**, 3783-3785 (1997).

Acknowledgements

This work was supported by The Wellcome Trust grant 098051 (JMB, SH, JP) and 107032AIA (DG,RAF) The Medical Research Council (JMB), The UK Cystic Fibrosis Trust (DMG, DR-R, IE, JP, RAF), Papworth Hospital (DMG, KPB, CSH, RAF), NIHR Cambridge Biomedical Research Centre (RAF), NIHR Specialist Respiratory Biomedical Research Unit, Imperial College London (DB), The UKCRC Translational Infection Research Initiative (JP), CF Foundation Therapeutics grant (SB, TK, CW, LM, PS), and The Prince Charles Hospital Foundation (SB), and National Services Division, NHS Scotland (IL) We are grateful to the following for their assistance with microbiological culture, environmental sampling, and isolate retrieval: Kate Ball (Aintree University Hospitals NHS Foundation Trust); Malcolm Brodlie and Matt Thomas (Newcastle upon Tyne Hospitals NHS Foundation Trust, UK); Grace Smith (Regional Mycobacterial Reference Laboratory, Birmingham, UK); Keith Thickett (Sheffield Teaching Hospitals NHS Foundation Trust, UK); Rhian Williams (Wales Centre for Mycobacteria, UK); Vasanthini Athithan and Margaret Gillham (Papworth Hospital NHS Foundation Trust). Jukka Corander (University of Oslo) assisted with BAPs clustering analysis. All sequence data associated with this study is deposited in the European Nucleotide Archive under project accession ERP001039. Ethical approval for the study was obtained nationally for centres in England and Wales from the National Research Ethics Service (NRES; REC reference: 12/EE/0158) and the

National Information Governance Board (NIGB; ECC 3-03 (f)/2012), for Scottish centres from NHS Scotland Multiple Board Caldicott Guardian Approval (NHS Tayside AR/SW), and locally for other centres.

Supplementary Material

Materials and Methods

Supplementary References (*38-58*)

Figs. S1 to S10

Tables S1