Supplemental Data

Appendix to: Genomic Analysis of a Hospital-Associated Outbreak of *Mycobacterium abscessus* Complex: Implications on Transmission

Rebecca M. Davidson[1], Sophie Nick[2], Sara M. Kammlade[1], Sruthi Vasireddy[3], Natalia Weakly[1], Nabeeh A. Hasan[1], L. Elaine Epperson[1], Michael Strong[1], Jerry A. Nick[4], Barbara A. Brown-Elliott[3], Jason E. Stout[5], Sarah S. Lewis[5,6], Richard J. Wallace, Jr.[3], Arthur W. Baker[5,6]

1. Center for Genes, Environment and Health, National Jewish Health, Denver, CO
2. Duke University School of Medicine, Durham, NC
3. Department of Microbiology, The University of Texas Health Science Center, Tyler, TX
4. Department of Medicine, National Jewish Health, Denver, CO
5. Division of Infectious Diseases, Duke University School of Medicine, Durham, NC
6. Duke Center for Antimicrobial Stewardship and Infection Prevention, Durham, NC

Contact Information:

Corresponding author:

Rebecca M. Davidson, PhD

Center for Genes, Environment and Health,

National Jewish Health,

1400 Jackson St., Denver CO 80206

DavidsonR@NJHealth.org; 303-270-2592 (w)

Additional Methods

Whole Genome Sequencing

Bacterial strains were grown in liquid media, and genomic DNA was isolated from bacterial pellets using a previously described method (1). Sequencing libraries were generated with 5ng of starting genomic DNA template using the Nextera XT Library Prep Kit (Illumina Inc., San Diego CA) and were sequenced on the Illumina MiSeq using the standard protocol for 2x300bp paired-end sequencing chemistry.

SNP Genotyping

Core genomes of *M. abscessus* subsp. *abscessus* isolates in the study cohort (n=26) and publicly available genomes (n=25) were compared using a reference-based mapping approach described previously (2). CO-RDP WGS data in the study cohort included the following samples: CF01818-01781.MAB (OL2-C1), CF00068-00219.MAB (OL2-C2), CF00794-00695.MAB (OL2-C3), CF00894-00919.MAB (OL2-C4), CF00924-01712.MAB (OL2-C5), CF01051-01232.MAB (OL2-C6), and CF00280-00677.MAB (OL2-C7) (3). Other publicly available genomes included the type strain for *M. abscessus* subsp. *abscessus*, ATCC19977^T (4), 21 *M. abscessus* subsp. *abscessus* isolates from previous studies (3, 5-9), two isolates of *M. abscessus* susbp. *massiliense* (10, 11), and one isolate of *M. abscessus* subsp. *bolletii* (12).

Briefly, raw sequence reads were trimmed using Skewer (13), and trimmed reads were mapped to the reference *M. abscessus* subsp. *abscessus* genome, ATCC19977^T (4) using Bowtie2 (14). SNPs relative to ATCC19977^T were called with samtools mpileup v1.5 and bcftools v1.3.1 (15). Genotype calls were filtered based on a mapping quality ≥ 20, a minimum read depth of 4x, and a minimum of 75% of reads supporting the base call. Only core genomic positions with complete genotype information for all isolates in the study were included in downstream analyses.

Genetic fingerprints in the *rpoβ* and *erm*(41) genes were analyzed in the genomic data for each isolate in the study cohort and among publicly available genomes of *M. abscessus* subsp. *abscessus*. First, the *rpoβ* mutation at position 207 in region V was confirmed by analyzing genomic coordinates corresponding to the gene in ATCC19977^T, i.e., chromosomal positions 3,916,650 to 3,916,770. Sequences of *erm*(41) genes were extracted from draft genomes and compared to the ten *erm*(41) sequevars described in Brown-Elliot et al. (16) by multiple gene sequence alignments performed with Seaview (17). A phylogenetic tree was created from concatenated sequences of partial *rpoβ* genes

and the full length *erm*(41) genes from the study cohort (n=26) and publicly available *M. abscessus* subsp. *abscessus* genomes (n=22) with Seaview (17) to confirm sample clustering based on genetic inclusion criteria (**Figure S1**).

Pan Genome Analysis

Trimmed reads were assembled into draft genomes using Unicycler (18). Contigs for each genome were ordered and oriented against the *M. abscessus* subsp. *abscessus* ATCC19977^T complete reference genome (4) with progressiveMauve (19). Genome sizes were on average 5,273,425bp (±71K bp). Genes were annotated using Prokka software (20), and pan genome analyses were performed with Roary analysis software (21). Pairwise pan genome comparisons were generated from standard Roary output with a custom python script.

Supplemental Figure

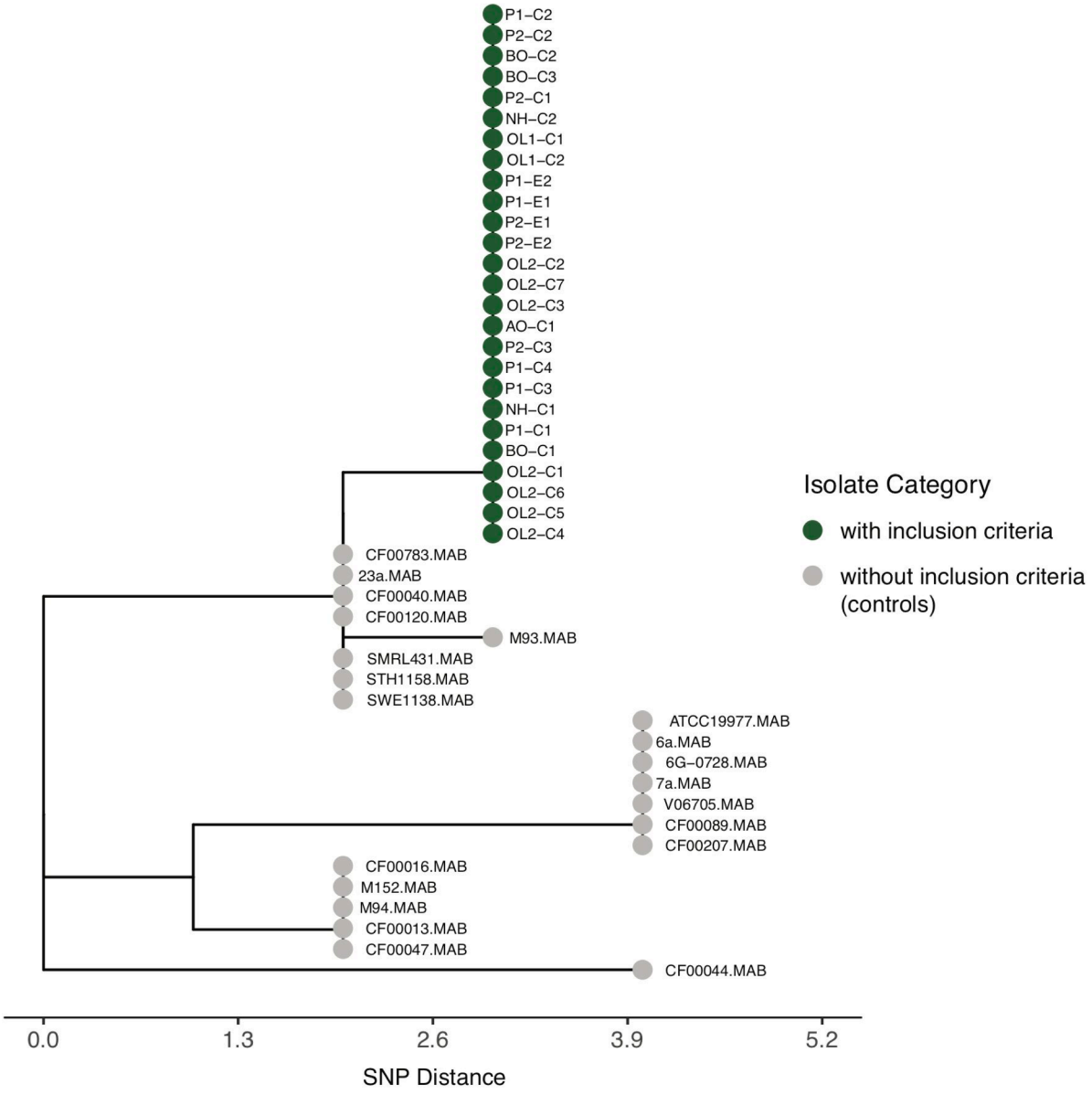

Figure S1. Phylogenetic analysis of *M. abscessus* subsp. *abscessus* isolates with genetic inclusion criteria in the study cohort (n=26) and isolates without inclusion criteria (n=22) based on concatenated sequence alignments of partial *rpoB* gene and full length *erm*(41) gene sequences.

Table S1. Genome assembly metrics for *M. abscessus* subsp. *abscessus* isolates in the study cohort

Isolate Name	Genome Size	# Contigs	# Genes
P1-E1	5271202	75	5246
P1-E2	5262971	79	5237
P1-C1	5269364	62	5235
P1-C2	5179189	78	5146
P1-C3	5272279	60	5240
P1-C4	5269813	64	5235
P2-E1	5266768	87	5236
P2-E2	5159744	77	5131
P2-C1	5338954	74	5361
P2-C2	5197591	70	5143
P2-C3	5290843	67	5260
BO-C1	5352265	71	5338
BO-C2	5363150	80	5393
BO-C3	5368843	85	5360
AO-C1	5149850	66	5091
NH-C1	5407916	70	5449
NH-C2	5290453	95	5265
OL1-C1	5196950	81	5166
OL1-C2	5248295	75	5206
OL2-C1	5413840	57	5380
OL2-C2	5218114	66	5136
OL2-C3	5304922	72	5236
OL2-C4	5228325	48	5169
OL2-C5	5210619	38	5151
OL2-C6	5318779	117	5344
OL2-C7	5258012	64	5183

References

1. Epperson LE, Strong M. 2020. A scalable, efficient, and safe method to prepare high quality DNA from mycobacteria and other challenging cells. *J Clin Tuberc Other Mycobact Dis* 19:100150.
2. Davidson RM, Hasan NA, Reynolds PR, Totten S, Garcia B, Levin A, Ramamoorthy P, Heifets L, Daley CL, Strong M. 2014. Genome sequencing of *Mycobacterium abscessus* isolates from patients in the united states and comparisons to globally diverse clinical strains. *J Clin Microbiol* 52:3573-82.
3. Davidson RM, Hasan NA, Epperson LE, Benoit JB, Kammlade SM, Levin AR, Calado de Moura V, Hunkins J, Weakly N, Beagle S, Sagel SD, Martiniano SL, Salfinger M, Daley CL, Nick JA, Strong M. 2021. Population Genomics of *Mycobacterium abscessus* from United States Cystic Fibrosis Care Centers. *Ann Am Thorac Soc* doi:10.1513/AnnalsATS.202009-1214OC.
4. Ripoll F, Pasek S, Schenowitz C, Dossat C, Barbe V, Rottman M, Macheras E, Heym B, Herrmann JL, Daffé M, Brosch R, Risler JL, Gaillard JL. 2009. Non mycobacterial virulence genes in the genome of the emerging pathogen *Mycobacterium abscessus. PLoS One* 4:e5660.
5. Bryant JM, Grogono DM, Greaves D, Foweraker J, Roddick I, Inns T, Reacher M, Haworth CS, Curran MD, Harris SR, Peacock SJ, Parkhill J, Floto RA. 2013. Whole-genome sequencing to identify transmission of *Mycobacterium abscessus* between patients with cystic fibrosis: a retrospective cohort study. *Lancet* 381:1551-60.
6. Bryant JM, Grogono DM, Rodriguez-Rincon D, Everall I, Brown KP, Moreno P, Verma D, Hill E, Drijkoningen J, Gilligan P, Esther CR, Noone PG, Giddings O, Bell SC, Thomson R, Wainwright CE, Coulter C, Pandey S, Wood ME, Stockwell RE, Ramsay KA, Sherrard LJ, Kidd TJ, Jabbour N, Johnson GR, Knibbs LD, Morawska L, Sly PD, Jones A, Bilton D, Laurenson I, Ruddy M, Bourke S, Bowler IC, Chapman SJ, Clayton A, Cullen M, Daniels T, Dempsey O, Denton M, Desai M, Drew RJ, Edenborough F, Evans J, Folb J, Humphrey H, Isalska B, Jensen-Fangel S, Jonsson B, Jones AM, et al. 2016. Emergence and spread of a human-transmissible multidrug-resistant nontuberculous mycobacterium. *Science* 354:751-757.
7. Choo SW, Wee WY, Ngeow YF, Mitchell W, Tan JL, Wong GJ, Zhao Y, Xiao J. 2014. Genomic reconnaissance of clinical isolates of emerging human pathogen *Mycobacterium abscessus* reveals high evolutionary potential. *Sci Rep* 4:4061.
8. Choo SW, Wong YL, Yusoff AM, Leong ML, Wong GJ, Ong CS, Ng KP, Ngeow YF. 2012. Genome sequence of the *Mycobacterium abscessus* strain M93. *J Bacteriol* 194:3278.
9. Pang S, Renvoise A, Perret C, Guinier M, Chelghoum N, Brossier F, Capton E, Jarlier V, Sougakoff W. 2013. Whole-Genome Sequence of *Mycobacterium abscessus* Clinical Strain V06705. *Genome Announc* 1.
10. Tettelin H, Sampaio EP, Daugherty SC, Hine E, Riley DR, Sadzewicz L, Sengamalay N, Shefchek K, Su Q, Tallon LJ, Conville P, Olivier KN, Holland SM, Fraser CM, Zelazny AM. 2012. Genomic insights into the emerging human pathogen *Mycobacterium massiliense. J Bacteriol* 194:5450.
11. Kim BJ, Kim BR, Hong SH, Seok SH, Kook YH, Kim BJ. 2013. Complete Genome Sequence of *Mycobacterium massiliense* Clinical Strain Asan 50594, Belonging to the Type II Genotype. *Genome Announc* 1.
12. Choi GE, Cho YJ, Koh WJ, Chun J, Cho SN, Shin SJ. 2012. Draft genome sequence of *Mycobacterium abscessus* subsp. *bolletii* BD(T). *J Bacteriol* 194:2756-7.
13. Jiang H, Lei R, Ding SW, Zhu S. 2014. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* 15:182.
14. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357-9.

15. Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27:2987-93.
16. Brown-Elliott BA, Vasireddy S, Vasireddy R, Iakhiaeva E, Howard ST, Nash K, Parodi N, Strong A, Gee M, Smith T, Wallace RJ, Jr. 2015. Utility of sequencing the erm(41) gene in isolates of *Mycobacterium abscessus* subsp. *abscessus* with low and intermediate clarithromycin MICs. *J Clin Microbiol* 53:1211-5.
17. Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 27:221-4.
18. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595.
19. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147.
20. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068-9.
21. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691-3.