The Journal of Infectious Diseases

MAJOR ARTICLE

Genomic Analyses of Longitudinal *Mycobacterium abscessus* Isolates in a Multicenter Cohort Reveal Parallel Signatures of In-Host Adaptation

JooHee Choi,[1] Eric C. Keen,[1] Meghan A. Wallace,[2] Skye Fishbein,[2] Jerome Prusa,[2] Madsen Zimbric,[3] Carlos R. Mejia-Chew,[4] Shail B. Mehta,[4] Thomas C. Bailey,[4] Lindsay J. Caverly,[3] Carey-Ann D. Burnham,[2,4,5,6,e] and Gautam Dantas[1,2,5,6,7,e]

[1]The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine in St Louis, St Louis, Missouri, USA; [2]Department of Pathology and Immunology, Division of Laboratory and Genomic Medicine, Washington University School of Medicine in St Louis, St Louis, Missouri, USA; [3]Department of Pediatrics, University of Michigan Medical School, Ann Arbor, Michigan, USA; [4]Department of Medicine, Washington University School of Medicine in St Louis, St Louis, Missouri, USA; [5]Department of Pediatrics, Washington University School of Medicine in St Louis, St Louis, Missouri, USA; [6]Department of Molecular Microbiology, Washington University School of Medicine in St Louis, St Louis, Missouri, USA; and [7]Department of Biomedical Engineering, Washington University in St Louis, St Louis, Missouri, USA

Background. Nontuberculous mycobacteria (NTM) are ubiquitous in the environment and an increasingly frequent cause of opportunistic infections. *Mycobacterium abscessus* complex (MABC) is one of the major NTM lung pathogens that disproportionately colonize and infect the lungs of individuals with cystic fibrosis (CF). MABC infection can persist for years, and antimicrobial treatment is frequently ineffective.

Methods. We sequenced the genomes of 175 isolates longitudinally collected from 30 patients with MABC lung infection. We contextualized our cohort amidst the broader MABC phylogeny and investigated genes undergoing parallel adaptation across patients. Finally, we tested the phenotypic consequences of parallel mutations by conducting antimicrobial resistance and mercury-resistance assays.

Results. We identified highly related isolate pairs across hospital centers with low likelihood of transmission. We further annotated nonrandom parallel mutations in 22 genes and demonstrated altered macrolide susceptibility co-occurring with a nonsynonymous *whiB1* mutation. Finally, we highlighted a 23-kb mercury-resistance plasmid whose loss during chronic infection conferred phenotypic susceptibility to organic and nonorganic mercury compounds.

Conclusions. We characterized parallel genomic processes through which MABC is adapting to promote survival within the host. The within-lineage polymorphisms we observed have phenotypic effects, potentially benefiting fitness in the host at the putative detriment of environmental survival.

Keywords. comparative genomics; in-host adaptation; *Mycobacterium abscessus* complex; nontuberculous mycobacteria.

Nontuberculous *Mycobacterium* spp. (NTM) are a diverse group of mycobacteria outside the *Mycobacterium tuberculosis* and *Mycobacterium leprae* complexes [1]. Commonly found in soil and water [2], NTM are mostly considered environmental saprophytes. However, NTM can cause opportunistic infection, particularly in humans who are immunocompromised or have preexisting lung conditions such as chronic obstructive pulmonary disease, cystic fibrosis (CF), or non-CF bronchiectasis [3]. Within the NTM, *Mycobacterium abscessus* complex (MABC) disproportionately colonize and infect patients with CF [4–7]. Colonization and infection can persist for years, and even with prolonged multidrug therapy [8] only approximately 30% of patients experience treatment success [9]. With infection rates increasing worldwide [5], MABC pose an imminent public health challenge.

The *M. abscessus* complex comprises 3 subspecies: *Mycobacterium* subsp. *bolletii*, *Mycobacterium* subsp. *massiliense*, and *Mycobacterium* subsp. *abscessus* [10]. Of these, over 50% of global clinical isolates have been shown to belong to 3 clusters of dominant circulating clones (DCCs) [11, 12]. However, it is unclear how frequently infection is caused by transmission between patients relative to independent environmental acquisition. Recent studies demonstrating the intercontinental presence of highly related pairs of clinical isolates (less than 20 single nucleotide polymorphisms; SNPs) suggest transmission between patients may occur more often than previously thought [13, 14], with some studies demonstrating epidemiological linkage such as overlapping clinic visit dates or outpatient visits by the same team of clinicians [15]. DCCs have also been found to exhibit elevated drug resistance and

Received 07 November 2022; editorial decision 21 May 2023; accepted 30 May 2023; published online 31 May 2023

[e]C. A. B. and G. D. jointly supervised this work.

Correspondence: Carey-Ann D. Burnham, PhD, Washington University in St. Louis, School of Medicine, 660. S Euclid Ave., St. Louis, MO 63110 (cburnham@wustl.edu); Gautam Dantas, PhD, School of Medicine, Washington University in St Louis, 4515 McKinley Avenue, St Louis, MO 63110 (dantas@wustl.edu).

The Journal of Infectious Diseases® 2023;228:321–31

https://doi.org/10.1093/infdis/jiad187

intracellular survival [11], suggestive of adaptation to a pathogenic lifestyle. However, epidemiological evidence of transmission is often lacking [16–19]. It remains unclear whether the prevalence of DCCs is a result of recent transmission, widespread environmental presence, sampling bias, a slower mutation rate among DCCs maintaining fitness benefits in the lung milieu—or a combination of these factors. Thus, understanding how this environmental species survives and persists in the host is critical for elucidating MABC transmission and evolution.

Despite their ability to cause chronic infection, little is known of the in-host adaptive behavior of MABC. MABC is intrinsically resistant to many antimicrobials [20, 21] but can also acquire antimicrobial resistance through SNPs in ribosomal rRNA genes *rrl* and *rrs* [20]. A smooth-to-rough morphotype switch via loss of glycopeptidolipid (GPL) expression on the cell surface has been shown to correspond with heightened virulence [22], but the mechanisms regulating this process are unknown [23]. To date, most studies investigating in-host adaptation of MABC have been limited to isolates from a single patient [24–26]. One study querying longitudinal samples from 18 patients found evidence for convergent evolution in 30 genes, including the GPL locus and virulence regulators [13]. Some of these mutations impaired survival on fomites [13], suggesting fitness trade-offs between the environment and host. Thus, utilizing high-resolution sequencing to explicitly investigate MABC in-host adaptation may inform prevention strategies and effective treatment development.

Here, we leverage a multicenter cohort of 175 isolates, longitudinally collected from 30 patients with MABC infection who were treated at academic medical centers in the United States between 2002 and 2020. Through whole-genome sequencing we first contextualized the genomic relatedness of isolates, incorporating an additional 1455 published MABC genomes for comparative genomic analyses. Next, we investigated within-lineage genomic diversity and characterized parallel in-host adaptation. Finally, we probed differences in antimicrobial resistance and mercury compound resistance correlated with specific genomic variants. This work provides high-resolution insight into MABC adaptive mechanisms and identifies novel candidate genes for parallel adaptation. We highlight how MABC adapts to the host milieu while shedding obsolete functions, furthering our understanding of MABC evolution during chronic infection.

METHODS

Isolate Collection

In total, 122 isolates from 22 patients were recovered from clinical specimens collected as a part of routine clinical care at the Barnes-Jewish Hospital microbiology laboratory between 2002 and 2019. Another 53 isolates from 8 patients were obtained from clinical samples at Michigan Medicine (University of Michigan) between 2014 and 2020 (Supplementary Table 1). Culture collection dates, clinic appointment dates, and pulmonary provider were reviewed for isolates with high genomic relatedness to assess potential epidemiological links. This research was conducted under Washington University Institutional Review Board No. 201904074.

Characterization of Within-Lineage Diversity

Lineage-specific whole-genome alignments were generated and annotated for SNPs and insertions/deletions (Supplementary Material). For the permutation analysis, the total number of observed mutations per lineage was noted (based on the lineage-specific whole-genome alignment). Next, the same number of mutations per lineage were randomly generated across each lineage's initial isolate's genome, and then annotated for which genes they landed on by mapping back to the prokka-generated gff file. For each round of permutation, the total number of lineages randomly mutated in parallel for a given gene was noted. This process was repeated for a total of 10 000 rounds to generate a hypothetical distribution (script: s_randommuts.py). *P* values were calculated by calculating the percentile of the actual number of lineages within the hypothetical distribution (s_pval.R). Bubble plots were visualized in R version 3.6.3 [27] using ggplot2 [28]. Genes encodig PE/PPE proteins were removed from the final list of significant hits due to their highly repetitive nature [29]. The presence of a 23-kb mercury plasmid was validated through BLASTN [30] (100% identity, 100% coverage).

Antimicrobial Resistance Assays

Isolates were inoculated into 7H9 broth supplemented with oleic albumin dextrose catalase. After aerobic growth to mid-log phase in 37°C, suspensions were diluted to 0.05 optical density $(OD)_{600}$ in 7H9, approximately 10^8 colony-forming units. In a 96-well plate, serial dilutions of each drug were prepared by adding 100 µL antimicrobial solution to 100 µL 7H9 broth. Then, 100 µL of the diluted sample was added to each well and mixed by pipetting up and down. The plate was placed in a sealed container to grow with shaking for 4 days in 37°C in air, after which 10 uL 0.02% resazurin was added to each well. After 24 hours (shaking in 37°C in air), the minimum inhibitory concentration was recorded as the minimum concentration observed to inhibit cell growth (determined from a color change from blue to pink). American Type Culture Collection (ATCC19977) was used as a control strain, along with a positive control row (containing just sample and 7H9 broth), and negative control row (containing just 7H9 broth) for each sample.

Mercury-Resistance Assays

Isolates were grown out to mid-log phase and diluted to 0.05 OD_{600} as described above. Mercury-resistance was tested in

the methodology described by Steingrube et al [31]: 600 µL of each isolate suspension was spread on a 7H10 plate supplemented with Albumin Dextrose Catalase (ADC), and 6-mm discs were loaded with 20 µL $HgCl_2$ or phenylmercury acetate (PMA) at 10-fold diluted concentrations: 10^{-2}M, 10^{-3}M, 10^{-4}M, and 10^{-5}M. A blank disc was added as a negative control. The plates were set to grow for 72 hours at 37°C in air, after which the zone of inhibition was measured.

RESULTS

Cohort Comprises 2 Major Subspecies of MABC Spanning the Global Phylogeny

To understand the phylogenetic relationships in the cohort, we annotated the genomes of all 175 clinical isolates, as well as 3 National Center for Biotechnology Information (NCBI) reference genomes for each MABC subspecies (subsp. *abscessus*, subsp. *massiliense*, subsp. *bolletii*). We aligned 2693 core genes to generate a maximum likelihood phylogenetic tree (Figure 1*A* and Supplementary Table 2). Most isolates (74.9%, 131/175) belonged to subsp. *abscessus*, and a smaller group of subsp. *massiliense* genomes (25.1%, 44/175) was also identified. No subsp. *bolletii* were present in the cohort. Therefore, the tree was rerooted with the subsp. *bolletii* reference genome as the outgroup.

We then calculated pairwise average nucleotide identity (ANI) for all isolate pairs (Supplementary Figure 1). All isolate pairs showed at least 97% ANI, fulfilling the genomic gold standard for microbial species [32]. Again, 2 major clusters were identified corresponding to each of the major subspecies present in the cohort. Pairwise comparisons within the same subspecies were at least 98.49% ANI (Figure 1*B* and Supplementary Table 3).

To contextualize our cohort within the global phylogeny, we conducted another core genome alignment (1423 core genes), incorporating 1452 additional published MABC genomes downloaded from NCBI and the European Nucleotide Archive for a total of 1630 genomes included in the analysis (genomes downloaded 28 December 2021; Supplementary Figure 2A and Supplementary Table 2). Most genomes in this larger cohort belonged to subsp. *abscessus* (72.1%, 1175/1630), followed by subsp. *massiliense* (27.2%, 444/1630). There were 8 subsp. *bolletii* genomes (0.49%, 8/1630), all contributed from downloaded genomes. The 175 isolates in our study were broadly distributed throughout the species phylogeny, suggesting a large range of genomic diversity is represented in the study cohort. ANI was also measured against each of the 3 subspecies reference genomes, and again the same subspecies were determined to be at least 98.5% ANI against a reference genome (Supplementary Figure 2B and Supplementary Table 3).

Genomic Relatedness Indicates Within-Patient Diversity of Subspecies and Lineages

To define lineages, we measured the relatedness of intrapatient isolate pairs in the 175 isolate core genome alignment. When looking at the overall distribution of pairwise core genome SNP distances, isolate pairs > 40 000 SNPs apart belonged to different subspecies (Figure 1*C* and Supplementary Figure 3). While same-patient isolate pairs generally exhibited low core genome SNP distance (median, 134; mean, 1671; range, 1–62 011), pairs of isolates belonging to the same subspecies but corresponding to large SNP distances (10 000–14 000 SNPs) were also identified, pointing to diverse MABC populations within some hosts (Figure 1*C*).

To further quantify within-host diversity, we applied pairwise ANI as an orthogonal approach and found that at a 99.99% ANI cutoff, we could distinguish clusters of highly related isolates (Figure 1*C* and Figure 2*A*). This cutoff corresponded to <400 whole-genome SNPs and <150 core genome Multilocus sequence type (cgMLST) distance (Supplementary Table 4). While most clusters comprised isolates from a single patient, there were also groups of closely related isolates found across multiple patients (L1-L4; Figure 2*A*). Interestingly, these multipatient groups were closely related (<200 core genome SNPs) to published DCC genomes (Supplementary Table 4). Whole-genome alignment of multipatient clusters further revealed SNP distances ranging from 7 to 490 SNPs (mean, 177; median, 135; Supplementary Table 4). Two patient pairs (MAB_04 and MAB_26, MAB_04 and MAB_30) exhibited DCC1 isolates less than 10 SNPs apart but were treated at different sites. Three other patient pairs exhibited DCC1 and DCC3 isolates less than 38 SNPs apart (suggested as an indicator of possible transmission [11]), with one pair (MAB_25 and MAB_05) treated at different sites in different years, one pair (MAB_26 and MAB_30) treated at the same site in different years, and just one pair (MAB_06 and MAB_17) at the same site with overlapping treatment years, with first positive sputum occurring 4 months apart. A review of these patients' culture collection dates, clinic appointment dates, and pulmonary provider did not suggest an epidemiological link between the two. Overall, the observation of highly related isolate pairs from distinct locations and years suggests presence of nearly clonal DCCs on a broad geographic scale.

To contextualize the evolutionary history of MABC lineages, we generated maximum parsimony trees using PHYLIP [33]. From these trees' ancestral nodes, we were able to infer the average distance from the most recent common ancestor (dMRCA) for each lineage (mean, 4.82 SNPs; 95% confidence interval [CI], 2.41–7.23). We then applied the estimated molecular clock for each corresponding subspecies [10] (1.8 SNPs/year for subsp. *abscessus*, 0.46 SNPs/year for subsp. *massiliense*) to each lineage and calculated the average time from most recent common ancestor (tMRCA; mean, 3.76 years; 95% CI, 2.18–5.34 years). Comparison of these estimates with collection dates revealed that despite variation among lineages (Supplementary Figure 4), average estimated tMRCA corresponded closely to the actual average time since initial positive NTM isolate culture

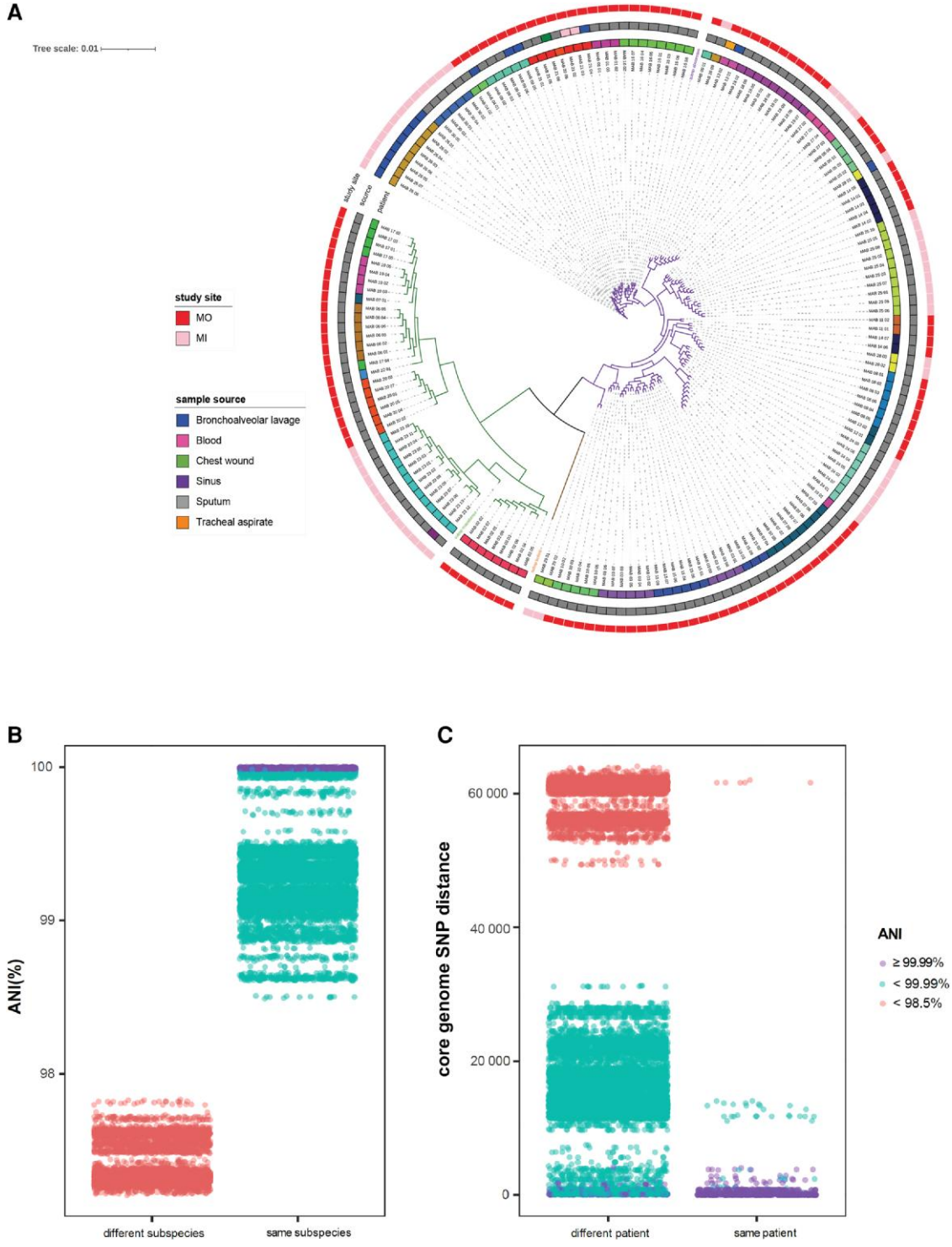

Figure 1. Genomic comparisons of entire cohort. *A*, Core genome alignment of 175 isolate genomes with 3 reference genomes for each *Mycobacterium abscessus* subspecies. Aligned with roary, converted with FastTree, visualized with iTOL. 2693 core genes. Tree is rerooted at the *Mycobacterium bolletii* reference genome. Subspecies were determined by clade: subspecies *massiliense* (n = 44, left, green); subspecies *abscessus* (n = 131, right, purple). Outer rings (from inner to outer) denote patient ID, sample source, and study site. *B*, ANI across isolate pairs. X-axis shows comparisons between different subspecies and same subspecies. Y-axis denotes ANI value. Isolate pairs belonging to the same subspecies have pairwise ANI values above 98.5%. *C*, Core genome SNP distance across isolate pairs. Each point represents a pairwise comparison. Comparisons are grouped as distances between isolates from different patients or the same patient. Points are colored by subspecies comparison as well as corresponding ANI: different subspecies pairs, purple; highly related pairs of at least 99.99% ANI, turquoise; and less related pairs less than 99.99% ANI, salmon. Abbreviations: ANI, average nucleotide identity; MI, Michigan; MO, Missouri; SNP, single-nucleotide polymorphism.

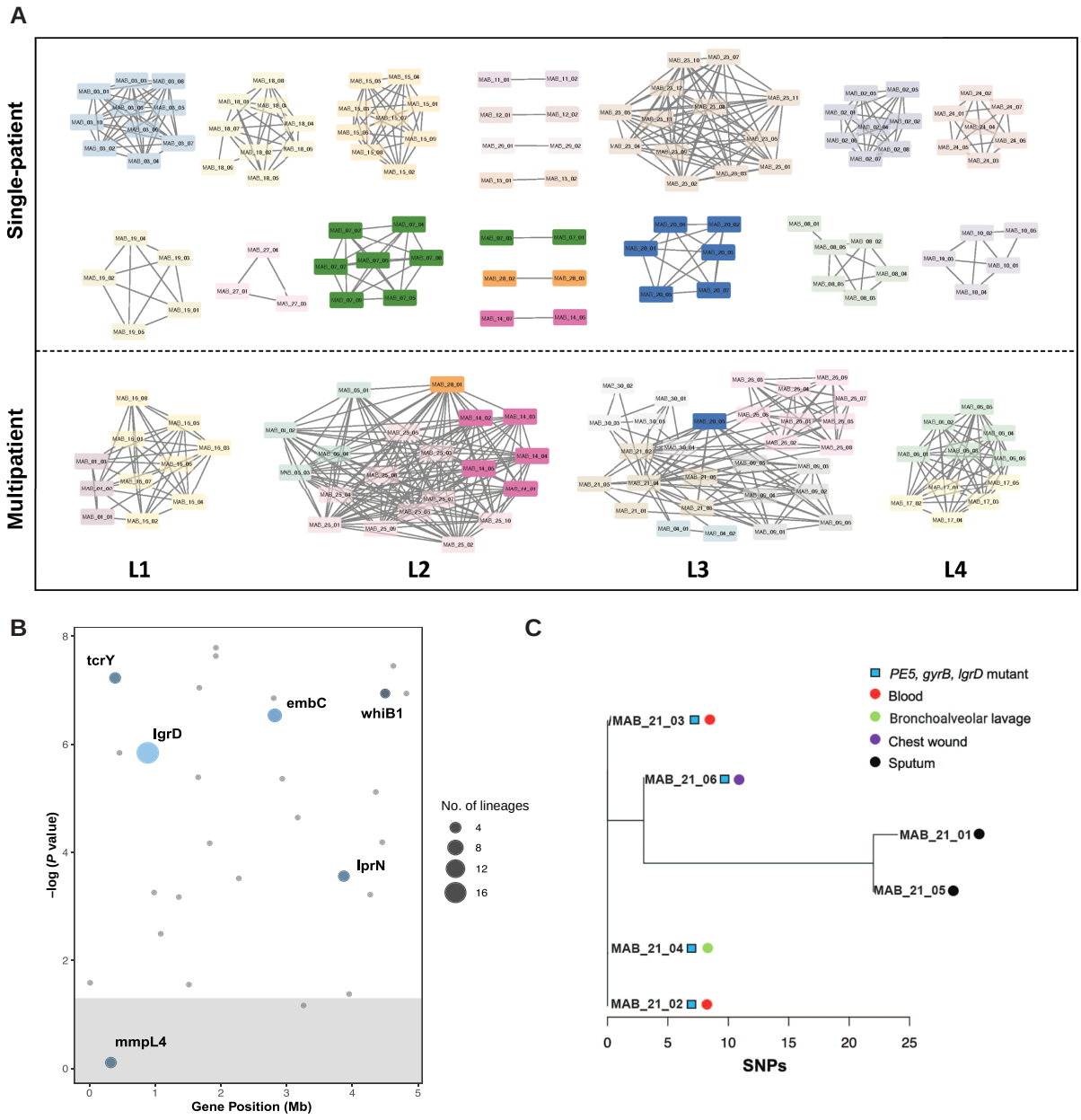

Figure 2. Multiple subspecies or lineages coexist within 4 patients. *A*, Network visualization of isolate genomes at least 99.99% ANI. Nodes indicate genomes and edges indicate a pairwise ANI value of at least 99.99%. Colors indicate patient ID; nodes of the same color are isolates from the same patient. Clusters of isolates from a single patient are shown above, and multipatient clusters are shown below. Multipatient clusters are labeled L1–L4. Nodes highlighted with higher opacity are isolates from patients with multiple subspecies present (MAB_20, blue) or multiple lineages (MAB_07, green; MAB_14, pink; MAB_28, orange). Two isolates without any 99.99% ANI matches are not shown: MAB_22_01 and MAB_27_02. ANI was measured with dnadiff, and clusters visualized on Cytoscape. *B*, Bubble plot displaying results of permutation analysis. Mutations were randomly distributed across representative isolate genomes to generate a neutral (expected) distribution for parallel mutations across lineages. This distribution was then compared with the observed number of lineages with mutations in each gene. X-axis denotes position in the ATCC19977 reference genome, Y-axis denotes negative log *P* value. Bubble size corresponds to number of lineages the gene was found to be mutated in. Genes mutated in at least 3 lineages are colored and named. Area with grey background indicates *P* value < .05. *C*, Unrooted tree showing whole-genome SNP distances between isolates from MAB_21, which were collected within 9 days of each other. X-axis is SNP distance, and tree nodes contain isolate IDs. Nodes are also annotated for site of collection and key mutations observed. SNPs are annotated by aligning reads against initial isolate genome MAB_21_01. Abbreviations: ANI, average nucleotide identity; SNP, single-nucleotide polymorphism.

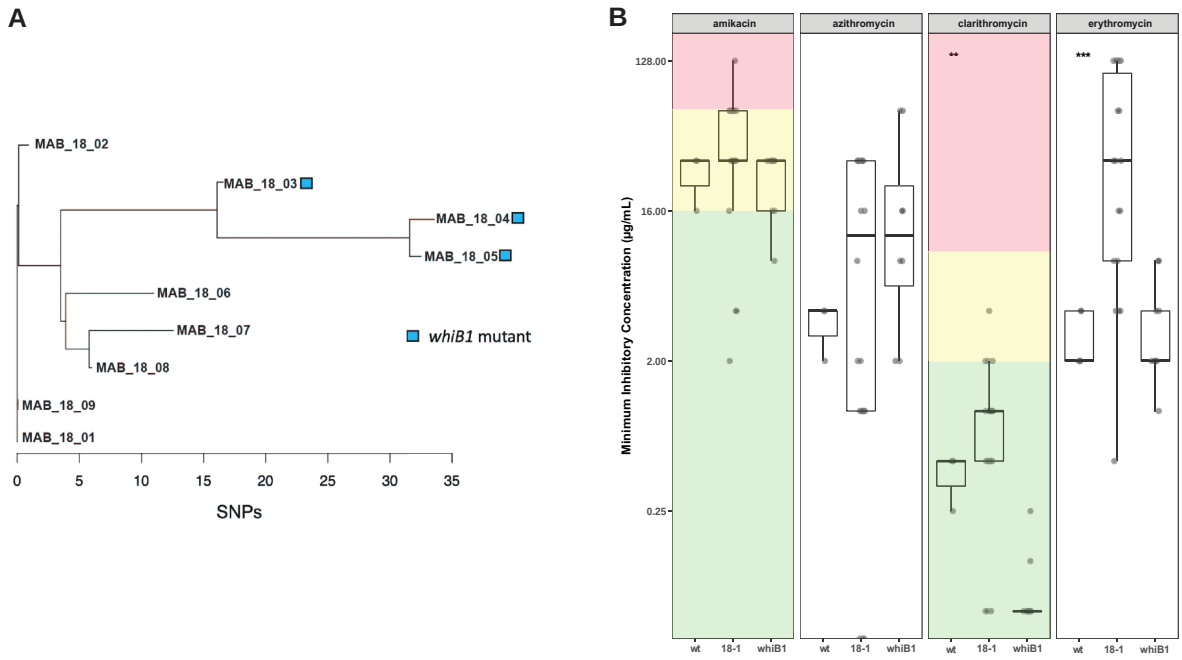

Figure 3. MAB_18 isolates show varied macrolide susceptibility. *A*, Unrooted tree showing whole-genome single-nucleotide polymorphism (SNP) distances between isolates from MAB_18. X-axis is SNP distance, and tree nodes contain isolate IDs. Nodes are also annotated for key observed mutations. *B*, Results of antimicrobial resistance assay. Each panel indicates a tested drug. Isolates (n = 9) are grouped by observed mutation: MAB_18_03, MAB_18_04, MAB_18_05 are categorized as "whiB1" and the remaining MAB_18 isolates are "18-1". ATCC19977 was included as a control ("wt"). Significant differences among the groups were observed for clarithromycin ($P = .001053$) and erythromycin ($P = .000789$, Kruskal-Wallace rank sum test). Background colors in the panel represent clinical interpretation according to CLSI M24-A2 guidelines: resistant, red; intermediate, yellow; or susceptible, green. Interpretations for erythromycin or azithromycin are not available and thus left blank.Whiskers extend 1.5 times the interquartile range while the box indicates first and third quartiles. Line in box indicates median.

of 3.57 years (95% CI, 2.53–4.62 years), lending credence to the accuracy of our genomic measurements.

Antimicrobial Resistance Genes Are Prevalent and Conserved Within Lineages

Next, we sought to characterize the antimicrobial resistance genes (ARGs) that may influence survival in vivo (Supplementary Table 5). We found that 100% (175/175) of isolates carried *bla*, *arr*, and *cmx_cmrA* genes, which confer resistance to β-lactams, rifamycin, and chloramphenicol, respectively; 92.6% (162/175) of isolates carried the *aph(3′′)* gene conferring resistance to streptomycin; and 74.3% (130/175) of isolates, all subsp. *abscessus*, carried the *erm(41)* gene for macrolide (induced clarithromycin) resistance. We did not observe substantial within-lineage variance for ARGs, and in 97% of cases (33/34 lineages) whole lineages were identical in predicted ARG profile (Supplementary Table 5). The single instance of within-lineage ARG profile variability was observed in patient MAB_14, in which the initial isolate MAB_14_01 carried mercury-resistance genes *merA* and *merB*, but subsequent isolates did not.

Lineages Undergo Parallel Within-Host Adaptation in Mycobacterial Virulence Genes

We sought to identify adaptive mutations that may have conferred fitness advantages for long-term survival in the host. For each lineage, we identified polymorphisms by aligning isolate reads against the genome from the lineage's earliest collection time and annotating whole-genome SNPs and insertions/deletions (indels). At this higher resolution, isolate pairs from the same lineage and patient were on average 10 SNPs apart (95% CI, 8.74–11.3; Supplementary Table 4).

In total we found 28 genes mutated in parallel across multiple patients and lineages (Figure 2*B* and Supplementary Table 6). We then applied a permutation test with 10 000 iterations to assess the significance (nonrandomness) of these parallel findings ("Methods" section; Figure 2*B* and Supplementary Table 6). Of the genes, 78.6% (22/28) were also significant by the permutation test (P value $< .05$, Benjamini-Hochberg). Some of these genes were also found to be variable in isolates collected just 9 days apart from multiple body sites (Figure 2*C*). Four genes (*crp*, *embC*, *whiB1*, and *espR*) were previously reported in the literature as evidence of within-host parallel evolution [13], suggesting diverse populations of MABC undergo similar adaptive trajectories during chronic infection.

Within-Lineage Diversity Affects Phenotypic Antimicrobial Resistance

We sought to test the phenotypic effects of parallel mutated genes by examining a lineage from patient MAB_18, which

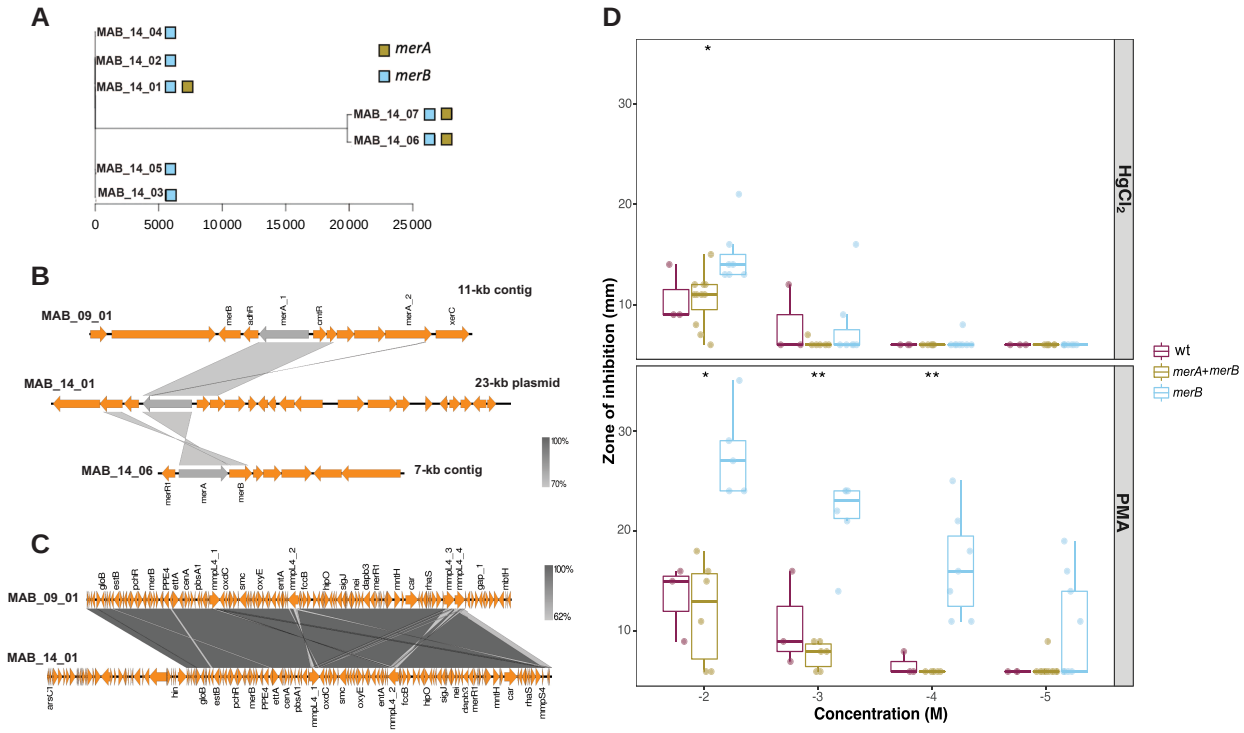

Figure 4. Loss of mercury-resistance genes affects mercury susceptibility. *A*, Unrooted tree showing whole-genome SNP distances between isolates from MAB_14. X-axis is SNP distance and tree nodes contain isolate IDs. Nodes are also annotated for presence of *merA* or *merB*. Isolates MAB_14_01 through MAB_14_05 belong to one lineage, while MAB_14_06 and MAB_14_07 belong to a second lineage. *B* and *C*, Genomic context of *merA* and *merB* genes across isolate genomes. *merA* is colored grey, while other coding sequences are orange. Each row is a visualization of an assembled contig containing *merA* or *merB*. Grey regions between contig rows indicate regions of high percent identity according to BLASTn. Visualized using easyfig. *D*, Results of mercury-resistance assays. Each clinical isolate was exposed to inorganic ($HgCl_2$) or organic (PMA) mercury compounds via disc diffusion assay. ATCC19977 was included as a control (purple). Isolates are grouped by genotype: *merA* and *merB* (MAB_14_01, MAB_14_07 and MAB_14_08, turquoise) or *merB* only (MAB_14_02 through MAB_14_05, orange). ATCC19977 (wt, purple) contains a 23-kb plasmid identical to MAB_14_01. Significant differences were observed between groups in $HgCl_2$ 10^{-2} M ($P = .011$, Kruskal-Wallace rank sum test), PMA 10^{-2} M ($P = .011$), PMA 10^{-3} M ($P = .0064$), and PMA 10^{-4} ($P = .0016$). Six mm indicates disc size and no zone of inhibition. Abbreviations: PMA, phenylmercury acetate; SNP, single-nucleotide polymorphism; wt, wild type. Whiskers extend 1.5 times the interquartile range while the box indicates first and third quartiles. Line in box indicates median (*: $P < 0.05$; **: $P < 0.01$, ***: $P < 0.001$).

featured nonsynonymous variation in the coding region of the *whiB1* gene (Figure 3*A*). *whiB1* (MAB_3539) encodes for a nitric oxide-sensitive transcriptional repressor in the WhiB family of proteins, and has been implicated in regulation of the ESX-1 secretion system in *M. tuberculosis* [34]. Deletion of the *whiB7* gene has been reported to confer sensitivity to the ribosome targeting drugs amikacin, clarithromycin, erythromycin, tetracycline, and spectinomycin [35]. In MAB_18, three isolates exhibited a nonsynonymous *whiB1* mutation resulting in a glycine (Gly) to alanine (Ala) switch at the Gly24 locus conserved between *M. tuberculosis* H37rV and *M. abscessus* ATCC 19977. We hypothesized that the nonsynonymous *whiB1* mutation at this conserved locus would affect clinical isolates' antimicrobial susceptibility. Thus, we measured their minimum inhibitory concentrations in amikacin, erythromycin, azithromycin, and clarithromycin using a resazurin microplate assay [36, 37] ("Methods" section). We observed that the 3 isolates with a *whiB1* mutation were significantly more susceptible to erythromycin and clarithromycin, but not amikacin or azithromycin (clarithromycin $P = .001053$, erythromycin $P = .000789$, Kruskal-Wallace rank sum test; Figure 3*B*), demonstrating within-lineage variability in phenotypic resistance corresponding to *whiB1* mutation.

Loss of 23-kb Mercury-Resistance Plasmid Is Linked to Mercury Susceptibility

We observed the loss of a 23-kb contig carrying mercury-resistance genes in isolates from patient MAB_14. This patient had 7 isolates belonging to 2 distinct lineages (Figure 4*A*). The initial isolate MAB_14_01 genome contained a 23-kb mercury-resistance plasmid identical to one in ATCC19977, reported to have originated in *Mycobacterium marinum* [38, 39] and predicted to encode a 472 amino acid (AA) MerA protein and 218 AA MerB protein. This plasmid was not present in subsequent isolate genomes of the same lineage (MAB_14_02 through MAB_14_05). Instead, these genomes contained a predicted 281 AA MerB in the chromosome, in a region encoding for cell wall components MmpL4 and PPE4 (Figure 4*C*). This

region was highly conserved across all subsp. *abscessus* isolates (Supplementary Table 7). In contrast, the MAB_14_06 and MAB_14_07 genomes of a separate lineage contained a 7-kb contig carrying both predicted 474 AA MerA and 219 AA MerB (Figure 4*B*).

MerA is a mercuric reductase, which reduces inorganic mercury Hg(II) to Hg(0), while MerB is an organic mercury lyase, which cleaves the Hg-C bond in organic mercury compounds to generate inorganic mercury Hg(II) (Supplementary Material). We sought to validate the phenotypic differences between isolates in MAB_14 that had different combinations and alleles of *merA* and *merB*, by conducting disc diffusion assays with an inorganic mercury compound, mercury chloride ($HgCl_2$), and an organic mercury compound, PMA ("Methods" section). We found striking differences in phenotypic resistance to both compounds between the 2 genotypes (Figure 4*D*). Isolates encoding *merA* had higher resistance to both PMA and $HgCl_2$, but this difference was significant only at the highest concentration of the inorganic compound, $HgCl_2$ (Kruskal-Wallis test, $P = .011$). In contrast, resistance to the organic compound PMA was significantly higher in the *merA* isolates across 3 different concentrations: 10^{-2}M, 10^{-3}M, and 10^{-4}M (Kruskal-Wallis test, $P = .011$, $P = .0064$, $P = .0016$), despite *merB* being present in all tested isolates.

DISCUSSION

The possibility of patient-to-patient or fomite-directed transmission of MABC has been debated [10, 15, 16, 18, 19]. Here, we found instances of highly related isolate pairs (<20 whole-genome SNPs) across different hospital centers as well as within the same center. We were unable to substantiate transmission in the current study, without formal epidemiological investigation. Importantly, we did not conduct environmental sampling of the health care centers, which have previously been linked to MABC reservoirs and outbreaks [14, 40]. Future studies will benefit from conducting both genomic and epidemiological investigation. Barring the existence of an undiscovered clinical reservoir, these occurrences may indicate widely circulating lineages of MABC acquired through separate infection events, as has been suggested by others [16–19]. The high genomic relatedness of these isolates could be explained by a slow rate of mutation among highly successful host-adapted pathogens. Whether this is a species-wide trend towards obligate pathogenicity (as the case for *M. tuberculosis* and *M. leprae*) [13, 41], or driving a divide between clinical and environmental populations of MABC, warrants exploration. Further research on the species' molecular clock and genomic comparisons with environmental isolates will provide more clarity on the transmission dynamics and evolutionary trajectory of MABC.

We identified candidate genes as hotspots of in-host adaptation, potentially conferring advantages for survival within the lung milieu. Although average tMRCA corresponded closely to average time difference between first and last isolate collection times, tMRCA values derived from SNP distances trended higher than actual collection times within each lineage, suggestive of great-than-expected genomic diversity. It is possible that even greater diversity of MABC was present but not captured due to limitations in study design. Here we only obtained one isolate per time point, but there may be additional coexisting lineages, or lineages that emerged prior to the dates to which we attributed them. Furthermore, CF patients exhibit polymicrobial lung infections [42], and cross-species interactions such as horizonal gene transfer or competition may occur. Future studies may capture more diversity by picking multiple colonies [43], conducting plate sweeps [11, 16], or conducting metagenomic sequencing of sputum samples [44, 45].

We demonstrated in vitro that within-lineage variations are associated with diverse phenotypic susceptibility to drugs. We found that a nonsynonymous mutation at a conserved site in *whiB1* was associated with increased susceptibility to clarithromycin and erythromycin, but not amikacin or azithromycin. The patient was on azithromycin therapy for CF at earlier time points (MAB_18_01 through MAB_18_05), while latter isolates (MAB_18_07 through MAB_18_09) were exposed to azithromycin and cefoxitin, and amikacin (Supplementary Figure 5). WhiB proteins are involved in regulation of pathogenesis, cell division, as well as oxidative stress response [20]. It is possible that *whiB1*, similar to *whiB7* [35], regulates antimicrobial resistance. Thus, the observed *whiB1* mutation may pose a trade-off between increased antimicrobial susceptibility and greater in-host survival through differences in transcriptional regulation, presenting an opportune window for treatment. A greater understanding of the regulatory pathways of *whiB1* in MABC is required to test this hypothesis. Further studies investigating treatment outcome in relation to prevalence of MABC parallel mutations may reveal additional fitness trade-offs, which may be exploited for treatment development.

Finally, we demonstrated how the loss of a 23-kb mercury-resistance plasmid is correlated with increased susceptibility to both organic and inorganic mercury compounds. Mercury exists naturally in the soil, water, and atmosphere, and cycles globally through processes such as industrial wastewater, landfills, burning fossil fuels, and processing through microorganisms [46]. Bacteria use MerA and MerB to break down organomercury into Hg(II), and subsequently to elemental mercury, Hg(0), which is highly volatile and rapidly diffuses out of the bacterial cell [47, 48]. A study of *Arabidopsis thaliana* found that insertion of bacterial *merA* and *merB* genes together conferred 10-fold higher resistance to organic methylmercury than insertion of *merB* alone [48]. These findings reflect our own observations, and potentially signify that both genes are imperative to successfully break down and remove mercury from the bacterial cell. While we cannot rule out the possibility

of plasmid loss during subcultivation, the consistent finding of the mercury plasmid among temporally initial isolates suggests loss of the mercury-resistance plasmid is likely due to the fitness trade-off of maintaining the plasmid in the low-mercury pulmonary environment [49, 50].

In this study we highlight genomic processes through which MABC adapts to promote its own survival within the host. Many of these events occur in parallel across patients and hospital sites and include DCCs circulating on a global scale. We suggest highly infectious strains of MABC may exhibit low rates of mutation to maintain a pathogen lifestyle. Furthermore, the within-lineage polymorphisms we observed have phenotypic effects, potentially benefiting fitness in the host, at the putative detriment of environmental survival. This work thus contributes to our understanding of in-host survival of MABC and may inform development of treatment strategies against these chronic infections.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Acknowledgments. We thank members of the Dantas laboratory for general discussion; Jessica Hoisington-López and MariaLynn Crosby for sequencing support; Eric Martin and Brian Koebbe for computational support; and Mark R. Sullivan for experimental guidance.

Disclaimer. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Financial support. This work was supported by the National Institute for Occupational Safety and Health of the US Centers for Disease Control and Prevention (grant number R01 OH011578 to G. D.); the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (NIH; grant number R01 AI123394 to G. D.); the Cystic Fibrosis Foundation (grant number CAVERL17A0 to L. J. C.); and the National Heart, Lung, and Blood Institute of the NIH (grant number K23HL136934 to L. J. C.). E. C. K. is supported by a Graduate Research Fellowship from the National Science Foundation (grant number DGE-1143945). J. P. is supported by the National Cancer Institute, NIH (grant numbers T32CA113275 and P30CA091842).

Potential conflicts of interest. All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

Presented in part: Lake Arrowhead Microbial Genomics Meeting (Lake Arrowhead, California, 12 September 2022); and the World Microbe Forum (virtual, 20-24 June 2021).

References

1. Faria S, Joao I, Jordao L. General overview on nontuberculous mycobacteria, biofilms, and human infection. J Pathog **2015**; 2015:e809014.
2. Falkinham JO. Nontuberculous mycobacteria in the environment. Clin Chest Med **2002**; 23:529–51.
3. Griffith DE, Aksamit T, Brown-Elliott BA, 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.
4. Levy I, Grisaru-Soen G, Lerner-Geva L, et al. Multicenter cross-sectional study of nontuberculous mycobacterial infections among cystic fibrosis patients, Israel. Emerg Infect Dis **2008**; 14:378–84.
5. Lopeman RC, Harrison J, Desai M, Cox JAG. *Mycobacterium abscessus*: environmental bacterium turned clinical nightmare. Microorganisms **2019**; 7:90.
6. Olivier KN, Weber DJ, Wallace RJ, et al. Nontuberculous mycobacteria. Am J Respir Crit Care Med **2003**; 167: 828–34.
7. Griffith DE. Emergence of nontuberculous mycobacteria as pathogens in cystic fibrosis. Am J Respir Crit Care Med **2003**; 167:810–2.
8. Martiniano SL, Nick JA, Daley CL. Nontuberculous mycobacterial infections in cystic fibrosis. Thorac Surg Clin **2019**; 29:95–108.
9. van Dorn A. Multidrug-resistant *Mycobacterium abscessus* threatens patients with cystic fibrosis. Lancet Respir Med **2017**; 5:15.
10. Bryant JM, Grogono DM, Greaves D, et al. Whole-genome sequencing to identify transmission of *Mycobacterium abscessus* between patients with cystic fibrosis: a retrospective cohort study. Lancet **2013**; 381:1551–60.
11. Bryant JM, Grogono DM, Rodriguez-Rincon D, et al. Population-level genomics identifies the emergence and global spread of a human transmissible multidrug-resistant nontuberculous mycobacterium. Science **2016**; 354:751–7.
12. Davidson RM, Hasan NA, Epperson LE, et al. Population genomics of *Mycobacterium abscessus* from U.S. cystic fibrosis care centers. Ann Am Thorac Soc **2021**; 18:1960–9.
13. Bryant JM, Brown KP, Burbaud S, et al. Stepwise pathogenic evolution of *Mycobacterium abscessus*. Science **2021**; 372:eabb8699.
14. Davidson RM, Nick SE, Kammlade SM, et al. Genomic analysis of a hospital-associated outbreak of *Mycobacterium*

abscessus: implications on transmission. J Clin Microbiol **2022**; 60:e01547-21.
15. Yan J, Kevat A, Martinez E, et al. Investigating transmission of *Mycobacterium abscessus* amongst children in an Australian cystic fibrosis centre. J Cyst Fibros **2020**; 19: 219–24.
16. Waglechner N, Tullis E, Stephenson AL, et al. Genomic epidemiology of *Mycobacterium abscessus* in a Canadian cystic fibrosis centre. Sci Rep **2022**; 12:16116.
17. Lipworth S, Hough N, Weston N, et al. Epidemiology of *Mycobacterium abscessus* in England: an observational study. Lancet Microbe **2021**; 2:e498–507.
18. Harris KA, Underwood A, Kenna DTD, et al. Whole-genome sequencing and epidemiological analysis do not provide evidence for cross-transmission of *Mycobacterium abscessus* in a cohort of pediatric cystic fibrosis patients. Clin Infect Dis **2015**; 60:1007–16.
19. Doyle RM, Rubio M, Dixon G, et al. Cross-transmission is not the source of new *Mycobacterium abscessus* infections in a multicenter cohort of cystic fibrosis patients. Clin Infect Dis **2020**; 70:1855–64.
20. Nessar R, Cambau E, Reyrat JM, Murray A, Gicquel B. *Mycobacterium abscessus*: a new antibiotic nightmare. J Antimicrob Chemother **2012**; 67:810–8.
21. Sanguinetti M, Ardito F, Fiscarelli E, et al. Fatal pulmonary infection due to multidrug-resistant *Mycobacterium abscessus* in a patient with cystic fibrosis. J Clin Microbiol **2001**; 39:816–9.
22. Howard ST, Rhoades E, Recht J, et al. Spontaneous reversion of *Mycobacterium abscessus* from a smooth to a rough morphotype is associated with reduced expression of glycopeptidolipid and reacquisition of an invasive phenotype. Microbiology **2006**; 152:1581–90.
23. Johansen MD, Herrmann JL, Kremer L. Non-tuberculous mycobacteria and the rise of *Mycobacterium abscessus*. Nat Rev Microbiol **2020**; 18:392–407.
24. Nick JA, Dedrick RM, Gray AL, et al. Host and pathogen response to bacteriophage engineered against *Mycobacterium abscessus* lung infection. Cell **2022**; 185:1860–74.e12.
25. Kreutzfeldt KM, McAdam PR, Claxton P, et al. Molecular longitudinal tracking of *Mycobacterium abscessus* spp. during chronic infection of the human lung. PLoS One **2013**; 8: e63237.
26. Lewin A, Kamal E, Semmler T, et al. Genetic diversification of persistent *Mycobacterium abscessus* within cystic fibrosis patients. Virulence **2021**; 12:2415–29.
27. R Core Team. R: A language and environment for statistical computing, **2020**. www.R-project.org. Accessed 8 May 2022
28. Hadley W. Ggplot2: elegant graphics for data analysis. New York: Springer International Publishing, **2016**.
29. Sampson SL. Mycobacterial PE/PPE proteins at the host-pathogen interface. *Clin Dev Immunol* **2011**; 2011:e497203.
30. Standard nucleotide BLAST. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_SPEC=GeoBlast&PAGE_TYPE=BlastSearch. Accessed 20 February 2023.
31. Steingrube VA, Wallace RJ, Steele LC, Pang YJ. Mercuric reductase activity and evidence of broad-spectrum mercury resistance among clinical isolates of rapidly growing mycobacteria. Antimicrob Agents Chemother **1991**; 35:819–23.
32. Richter M, Roselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. PNAS **2009**: 106:19126–31 .
33. Felsenstein J. PHYLIP (phylogeny inference package) version 3.697, 2009. https://evolution.genetics.washington.edu/phylip.html. Accessed 4 February 2022.
34. Kudhair BK, Hounslow AM, Rolfe MD, et al. Structure of a wbl protein and implications for NO sensing by *M. tuberculosis*. Nat Commun **2017**; 8:2280.
35. Hurst-Hess K, Rudra P, Ghosh P. *Mycobacterium abscessus* WhiB7 regulates a species-specific repertoire of genes to confer extreme antibiotic resistance. Antimicrob Agents Chemother **2017**; 61:11.
36. Bich Hanh BT, Quang NT, Park Y, et al. Omadacycline potentiates clarithromycin activity against *Mycobacterium abscessus*. Front Pharmacol **2021**; 12:790767.
37. de Carvalho NFG, Sato DN, Pavan FR, Ferrazoli L, Chimara E. Resazurin microtiter assay for clarithromycin susceptibility testing of clinical isolates of *Mycobacterium abscessus* group. J Clin Lab Anal **2016**; 30:751–5.
38. Medjahed H, Gaillard JL, Reyrat JM. *Mycobacterium abscessus*: a new player in the mycobacterial field. Trends Microbiol **2010**; 18:117–23.
39. Schué M, Dover LG, GurdyalS B, Parkhill J, Brown NL. Sequence and analysis of a plasmid-encoded mercury resistance operon from *Mycobacterium marinum* identifies MerH, a new mercuric ion transporter. J Bacteriol **2009**; 191:439–44.
40. Baker AW, Lewis SS, Alexander BD, et al. Two-phase hospital-associated outbreak of *Mycobacterium abscessus*: investigation and mitigation. Clin Infect Dis **2017**; 64:902–11.
41. Chavarro-Portillo B, Soto CY, Guerrero MI. *Mycobacterium leprae*'s evolution and environmental adaptation. Acta Tropica **2019**; 197:105041.
42. Filkins LM, O'Toole GA. Cystic fibrosis lung infections: polymicrobial, complex, and hard to treat. PLoS Pathog **2015**; 11:e1005258.
43. Lieberman TD, Flett KB, Yelin I, et al. Genetic variation of a bacterial pathogen within individuals with cystic fibrosis

provides a record of selective pressures. Nat Genet **2014**; 46:82–7.
44. Raghuvanshi R, Vasco K, Vázquez-Baeza Y, et al. High-resolution longitudinal dynamics of the cystic fibrosis sputum microbiome and metabolome through antibiotic therapy. mSystems **2020**; 5:e00292-20.
45. Feigelman R, Kahlert CR, Baty F, et al. Sputum DNA sequencing in cystic fibrosis: non-invasive access to the lung microbiome and to pathogen details. Microbiome **2017**; 5:20.
46. Barkay T, Miller SM, Summers AO. Bacterial mercury resistance from atoms to ecosystems. FEMS Microbiol Rev **2003**; 27:355–84.
47. Summers AO. Organization, expression, and evolution of genes for mercury resistance. Annu Rev Microbiol **1986**; 40:607–34.
48. Bizily SP, Rugh CL, Meagher RB. Phytodetoxification of hazardous organomercurials by genetically engineered plants. Nat Biotechnol **2000**; 18:213–7.
49. Lim HE, Shim JJ, Lee SY, et al. Mercury inhalation poisoning and acute lung injury. Korean J Intern Med **1998**; 13: 127–30.
50. Reichl FX, Walther UI, Durner J, et al. Cytotoxicity of dental composite components and mercury compounds in lung cells. Dent Mater **2001**; 17:95–101.