

## Article

# Horizontal Gene Transfer and Drug Resistance Involving *Mycobacterium tuberculosis*

Xuhua Xia <sup>1,2</sup> 

<sup>1</sup> Department of Biology, University of Ottawa, Ottawa, ON K1N 9A7, Canada; xxia@uottawa.ca; Tel.: +1-613-562-5718

<sup>2</sup> Ottawa Institute of Systems Biology, University of Ottawa, Ottawa, ON K1H 8M5, Canada

**Abstract:** *Mycobacterium tuberculosis* (Mtb) acquires drug resistance at a rate comparable to that of bacterial pathogens that replicate much faster and have a higher mutation rate. One explanation for this rapid acquisition of drug resistance in Mtb is that drug resistance may evolve in other fast-replicating mycobacteria and then be transferred to Mtb through horizontal gene transfer (HGT). This paper aims to address three questions. First, does HGT occur between Mtb and other mycobacterial species? Second, what genes after HGT tend to survive in the recipient genome? Third, does HGT contribute to antibiotic resistance in Mtb? I present a conceptual framework for detecting HGT and analyze 39 ribosomal protein genes, 23S and 16S ribosomal RNA genes, as well as several genes targeted by antibiotics against Mtb, from 43 genomes representing all major groups within *Mycobacterium*. I also included *mgtC* and the insertion sequence IS6110 that were previously reported to be involved in HGT. The insertion sequence IS6110 shows clearly that the Mtb complex participates in HGT. However, the horizontal transferability of genes depends on gene function, as was previously hypothesized. HGT is not observed in functionally important genes such as ribosomal protein genes, rRNA genes, and other genes chosen as drug targets. This pattern can be explained by differential selection against functionally important and unimportant genes after HGT. Functionally unimportant genes such as IS6110 are not strongly selected against, so HGT events involving such genes are visible. For functionally important genes, a horizontally transferred diverged homologue from a different species may not work as well as the native counterpart, so the HGT event involving such genes is strongly selected against and eliminated, rendering them invisible to us. In short, while HGT involving the Mtb complex occurs, antibiotic resistance in the Mtb complex arose from mutations in those drug-targeted genes within the Mtb complex and was not gained through HGT.



**Citation:** Xia, X. Horizontal Gene Transfer and Drug Resistance Involving *Mycobacterium tuberculosis*. *Antibiotics* **2023**, *12*, 1367. <https://doi.org/10.3390/antibiotics12091367>

Academic Editor: Antony T. Vincent

Received: 14 July 2023

Revised: 13 August 2023

Accepted: 21 August 2023

Published: 25 August 2023



**Copyright:** © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** antibiotic resistance; phylogenetic incongruence; horizontal gene transfer; rifampin; isoniazid

## 1. Introduction

*Mycobacterium tuberculosis* (Mtb) is enigmatic for having two seemingly incompatible traits. Mtb grows slowly, with a generation time of about 24 h [1–3]. It also features a low mutation rate of about  $2.5 \times 10^{-10}$ /bp/day [4]. Such an organism is not expected to acquire drug resistance rapidly in response to antibiotic challenges, yet Mtb has readily evolved extensive drug resistance [5–7] or even total drug resistance [5,8,9], leading to untreatable tuberculosis [8]. In a study monitoring 141 tuberculosis (TB) patients suffering from infection with Mtb strains that are susceptible to second-line drugs (SLDs), acquired resistance in Mtb was observed in 19 patients (14%) [10]. This rate of acquiring drug resistance is comparable to that in *Streptococcus pneumoniae* [11], which has a much shorter generation time [12] and a higher mutation rate [13,14].

One explanation for the rapid acquisition of drug resistance in Mtb is that drug resistance may evolve in other fast-replicating mycobacteria and then be transferred to Mtb through horizontal gene transfer (HGT). However, there has been little evidence of recent

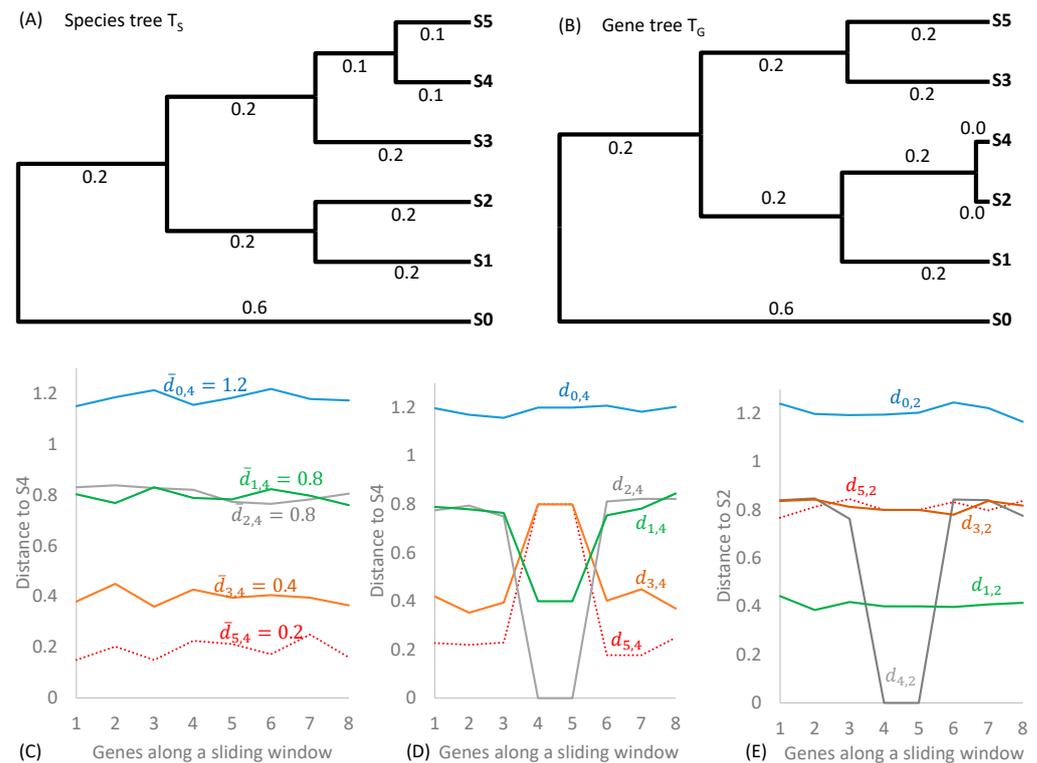
HGT. The lack or rarity of HGT has been attributed to three factors: (1) the thick cell wall that prevents exogenous DNA from entering the cell; (2) the lack of plasmids to transmit DNA across species; and (3) the limited repertoire of restriction enzymes and secretion systems to facilitate HGT. For these reasons, HGT is often discounted as a main mechanism for the acquisition of drug resistance [7,15–18]. Barring the possibility that drug resistance may arise in rapidly replicating mycobacteria and then transferred to slowly replicating mycobacteria, drug resistance in Mtb is often attributed to mutations within the Mtb complex, enhanced by genetic variation generated through insertion sequences [19–21] and fusion proteins [16], although genes responsible for antibiotic resistance are not associated with insertion sequences in the genomes of mycobacteria species [22].

The possibility that HGT contributes to drug resistance in Mtb, however, has neither been strongly rejected nor conclusively supported. Any HGT event involves a donor genome and a receptor genome. It is possible that previous studies may have missed either the donor genome or the receptor genome, or both, and consequently did not have a sufficient chance to detect rare evolutionary events such as HGT.

Two complementary phylogenetic methods have often been used for detecting HGT (and subsequent recombination) in viral and bacterial genomes. The first is based on phylogenetic incongruence (Figure 1A), illustrated numerically in Xia [23] (pp. 36–38). It addresses the question of which gene might be involved in horizontal gene transfer. The method needs a species tree (Figure 1A) and a gene tree (Figure 1B). The species tree is typically approximated by using a set of informational genes [24], i.e., those involved in genome replication, transcription, and translation that are less frequently involved in HGT and contain stronger phylogenetic signals than other genes. Single-copy genes, including most ribosomal protein genes are rarely observed in HGT [25,26], and the concatenated ribosome genes should provide an excellent approximation to the species tree. For cross-validation, one may also construct a tree from concatenated large ribosomal protein genes and another tree from concatenated small ribosomal protein genes. The gene tree is based on the gene suspected to be involved in HGT. If a gene is frequently involved in HGT, then its gene tree will tend to be incongruent with the species tree. In the gene tree in Figure 1B, species S4 gained gene 4 and 5 from species S2, so the phylogenetic position of S4 in the gene tree (Figure 1B) is incongruent with that in the species tree (Figure 1A).

I should emphasize that, although a gene frequently involved in HGT tends to have a gene tree incongruent with the species tree, the phylogenetic incongruence itself is not evidence of HGT because it can also result from factors other than HGT, e.g., gene duplication followed by lineage sorting [27–30]. Suppose Mtb variants *microti*, *bovis*, and *africanum* share an identical version of gene X, but Mtb variant *bovis* also has a highly diverged homologue of gene X. The two versions of gene X in Mtb variant *bovis* could be either paralogues (through gene duplication) or xenologues (through HGT). Further evidence is needed to discriminate between the two.

The second phylogenetic method for detecting HGT and recombination is the SimPlot/DistPlot (similarity or distance plot) method, numerically illustrated in Xia ([23], pp. 49–53). The DistPlot method (Figure 1C–E) is more informative than the SimPlot method because the latter does not correct for multiple hits, whereas the former can use commonly used substitution models to correct for multiple hits [29]. DistPlot aims to accomplish two tasks. The first is to identify species that may be the donor or recipient in an HGT event (optionally with subsequent recombination). The second is to characterize which genes are involved in HGT. Figure 1C–E illustrates the scenario of six species (S0 to S5) with eight genes (1 to 8). Without HGT (and other factors distorting phylogenetic signals), the evolutionary distance between any species and other species is expected to be relatively constant. For example, given  $T_5$  in Figure 1A, the distance between species S5 and S4 ( $d_{4,5}$ ) should be smaller than  $d_{3,4}$ , and this relative magnitude should be true for any genes, as illustrated in Figure 1C.



**Figure 1.** Phylogenetic methods for detecting horizontal gene transfer and recombination in viral and bacterial genomes. **(A)** A species tree  $T_S$ . **(B)** A gene tree  $T_G$  resulting from species  $S_4$  horizontally acquiring the gene from  $S_2$ . **(C)** Expected evolutionary distance ( $\bar{d}_{ij}$ ) between  $S_4$  and other species for eight genes (1 to 8), given  $T_S$ . The two subscripted  $i$  and  $j$  in  $\bar{d}_{ij}$  are species indicators. The observed rRNA distance may fluctuate stochastically above and below  $\bar{d}$ . **(D)** Distance  $d_{i,4}$  between  $S_4$  and other species for the eight genes, given that species  $S_4$  gained genes 4 and 5 from species  $S_2$ . For all species descended from the common ancestor of  $S_2$  and  $S_4$ ,  $d_{i,4}$  would differ among genes because  $S_4$  gained genes 4 and 5 from  $S_2$  so  $T_G$  from these two genes would differ from  $T_S$ . **(E)** Distance  $d_{i,2}$  between  $S_2$  and other species for the eight genes. Only  $d_{4,2}$  differs among genes, but other  $d_{i,2}$  distances do not, because  $S_2$  does not change its phylogenetic positions.

When HGT occurs with  $S_4$  gaining genes 4 and 5 from  $S_2$ , then  $d_{i,4}$  for genes 4 and 5 would differ from  $d_{i,4}$  from other genes for all species descending from the common ancestor of the recipient  $S_4$  and donor  $S_2$  (Figure 1D). This change of  $d_{i,4}$  suggests  $S_4$  as the recipient species. The HGT event has little effect on  $d_{0,4}$  (Figure 1D) because  $S_0$  diverged before the common ancestor of  $S_2$  and  $S_4$ . In contrast to the signature of the recipient ( $S_4$ ) in Figure 1D, the donor species, such as  $S_2$ , does not change its distance to other species ( $d_{i,2}$ ) except for  $d_{2,4}$  (Figure 1E). This is because the donor's phylogenetic position does not change except for its relationship to the recipient ( $S_4$ ). These patterns of pairwise  $d_{i,j}$  values are important for inferring the donor and recipient species involved in HGT, assuming that the donor and recipient species, or at least their close relatives, have been included in the study. In most cases, the intermediate donor is a plasmid, while the bacterial donor could have already gone extinct and is therefore unidentifiable.

The DistPlot method is often applied with a sliding window along aligned gene sequences or aligned viral genomic sequences, with the objective of testing whether the gene or the genome might have participated in recombination and consequent incorporation of a diverged gene segment from sources other than ancestral inheritance. This approach is frequently used in identifying recombination events in viral genomes ([23] pp. 49–53).

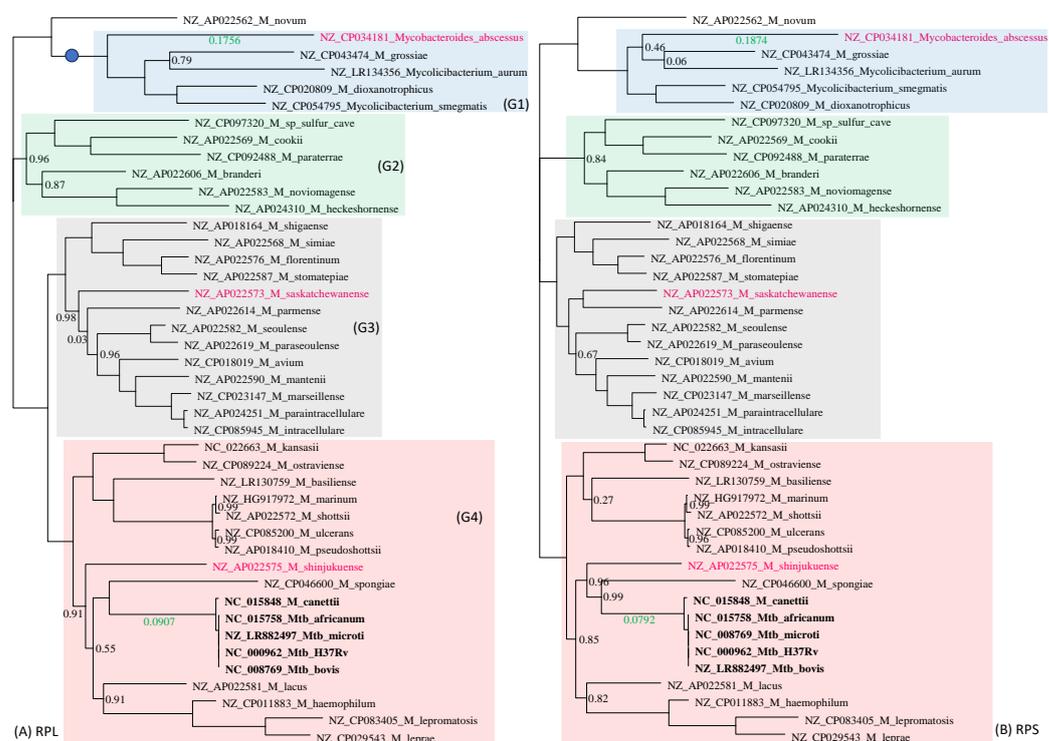
The conceptual framework of inference has not been fully applied to the study of HGT in mycobacteria. In this paper, my main focus is on whether any species in the Mtb complex has served as a recipient species for a subset of genes that have been targeted by antibiotics.

I approximated a species tree by a phylogeny based on ribosomal protein genes, and used phylogenetic incongruence and the DistPlot to detect the presence of HGT in the evolution of *Mycobacterium* lineages.

## 2. Results

### 2.1. Species Tree: Approximation from Ribosomal Protein Genes

There is excellent concordance between the tree derived from the 21 RPL proteins (Figure 2A) and that from the 18 RPS proteins (Figure 2B). Only three species (colored in red in Figure 2) differ slightly in their phylogenetic positions. Four distinct phylogenetic clades were labeled groups 1 to 4 (G1 to G4), shaded in blue, green, gray, and pink, respectively (Figure 2). The most striking feature of the two trees is the genetic homogeneity in the Mtb complex (*M. canettii* and *M. tuberculosis* variants *africanum*, *microti*, *bovis*, and H37Rv) and their clear separation from the rest of the *Mycobacterium* species.



**Figure 2.** Approximation of the species tree with genes encoding ribosomal proteins. Species names are in the format of “GenBank accession”\_“Species Name”. (A) Phylogeny from 21 aligned and concatenated RPL (large ribosome protein) genes. (B) Phylogeny from 18 aligned and concatenated RPS (small ribosomal protein) genes. The two trees are highly concordant, with only three species (colored red) differing slightly in their phylogenetic positions. The trees are unrooted. Bootstrap values equal to 1.00 are not shown, but those smaller than 1.00 are next to the node. The green-colored branch lengths, one leading to *Mycobacteroides abscessus*, and the other separating the Mtb complex from the rest, serve the function of a scale bar. Corresponding clades in (A,B) are shaded in the same color.

G1 (Figure 2) consists of culturable, relatively fast-growing, but phenotypically and genetically diverse *Mycobacterium* species. Because of this diversity, two new genera, *Mycolicibacterium* and *Mycobacteroides*, have recently been proposed to accommodate them [31]. A subsequent morphological study [32,33] and a large-scale genomic comparison [34] supported the division. However, controversies with this taxonomic change remain [35]. While genetically diverse, as manifested by the long branches, these species are consistently grouped in the same clade in both the RPL tree (Figure 2A) and the RPS tree (Figure 2B). I will continue to refer to all of them as *Mycobacterium* species.

G1 includes three nonpathogenic species. *M. dioxanotrophicus* was isolated from river sediment and capable of using 1,4-dioxane as a sole source of carbon and energy [36]. *M. aurum* was first isolated from a wastewater treatment plant in Prague and has the potential to degrade morpholine, which is an industrial waste [37,38]. *M. smegmatis* is best known for being used as a model organism for the Mtb complex [39,40]. G1 also includes *M. grossiae*, which was occasionally found in the human respiratory tract [41], and *M. abscessus*, which can cause severe diseases in immunocompromised patients and prompted the development and refinement of phage therapies because of its multi-drug resistance [42,43].

The five species in G1 also share a number of other unique features. For example, they all share a 3-codon insertion close to the 3' end of the *rpoA* gene (encoding the  $\alpha$  subunit of the RNA polymerase). However, I am interested only in features that lead to phylogenetic abnormalities implicating HGT and will not elaborate on features that are consistent with the putative species tree.

The rest of the species are slow-growing, with *M. leprae* having the longest generation time [2,44]. One would have expected the fast-growing G1 group to have much longer branches than the rest of the *Mycobacterium* species, and that the branches leading to the slow-growing *M. leprae* would be particularly short. However, this expectation is not supported (Figure 2).

G2 species are also phenotypically and genetically diverse, including a methanotrophic species found in sulfur caves (*M. sp\_sulfer\_cave* in Figure 2) with the tentative name of *Candidatus Mycobacterium methanotrophicum* [45], *M. cookie* isolated from sphagnum vegetation and surface water of moors in New Zealand [46] and occasional human pathogens such as *M. paraterrae* [47]. G3 is the *Mycobacterium avium* complex responsible for most of the nontuberculosis mycobacterial infections [48,49]. They can grow up to a temperature of 38.5 °C, which enables them to infect humans. Moreover, while a high fever can arrest the growth of these bacteria, it is insufficient to kill them because bacteria in the *Mycobacterium avium* complex can often survive up to 49 °C. G4 species includes the Mtb complex and the *M. leprae* complex, as well as many slow-growing *Mycobacterium* species causing nontuberculosis infections. Some species, such as *M. shinjukuense*, can cause severe pulmonary disease [50].

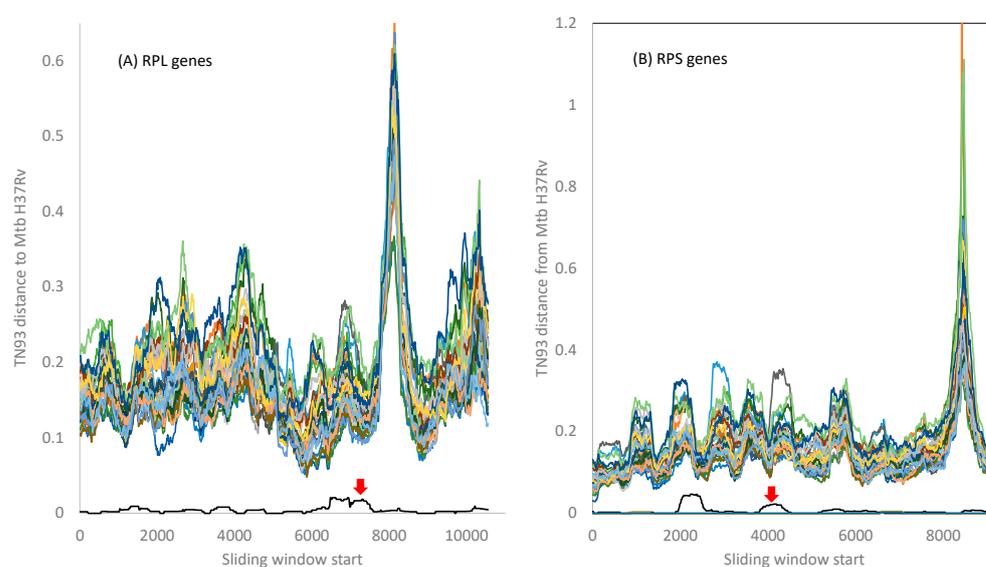
Given that many nontuberculosis mycobacterial species are also found in human respiratory tracts and lungs, and are often isolated from tuberculosis patients, chances appear to exist for them to exchange genetic materials with the Mtb complex. However, the Mtb complex exhibits little genetic variation within the complex, but is well separated from the rest of the *Mycobacterium* species, with a branch length of 0.0907 in the RPL tree and 0.0792 in the RPS tree (Figure 2). This suggests that ribosomal protein genes are unlikely to be involved in HGT, confirming ribosomal protein genes as excellent phylogenetic signals for reconstructing the species tree.

Ribosomal proteins in mycobacteria have not been used as drug targets in developing antibacterials against the Mtb complex. Oxazolidinones bind to the large subunit ribosome and inhibit translation initiation by preventing the proper positioning of the initiator tRNA at the A-site [51,52]. However, the effect is more likely mediated by rRNA than ribosomal proteins for two reasons [51]. First, although binding studies showed interaction with the 50S ribosomal subunit, crosslinking studies revealed interaction partners to be the 23S and 16S rRNAs. Second, oxazolidinone-resistant isolates were found to have specific mutations (G2447T and G2576T) on the 23S rRNA [51,52]. Thus, the drug target is rRNA instead of ribosomal protein. However, ribosomal proteins have been targeted in other pathogenic bacteria, and nonsynonymous mutations at ribosomal protein genes can confer drug resistance in other bacterial species. For example, amino acid replacements at site G70 of the large ribosomal L4 protein reduce the susceptibility of *Neisseria gonorrhoeae* to azithromycin [53]. Similarly, nonsynonymous mutations in ribosomal proteins L22 and L4 can alter the three-dimensional structure of the large ribosome subunit, leading to changes in *E. coli* susceptibility to erythromycin [54]. This later study highlights the potential cost of drug resistance to bacterial species—mutations leading to an alteration

of the three-dimensional structure of the ribosome are almost always deleterious. Thus, a foreign ribosomal protein or rRNA gained through HGT is most likely deleterious, leading to reduced fitness and eventual elimination of the HGT recipient.

### 2.2. Genomic Integrity in the Mtb Complex as Revealed by Ribosomal Proteins

I computed the evolutionary distance based on the TN93 model [29,55] over a sliding window of 500 nt between Mtb H37Rv and the other 42 *Mycobacterium* species included in this study. The 21 RPL genes and the 18 RPS genes from Mtb variants *microti*, *bovis*, and *africanum* are essentially identical to those in Mtb H37Rv. *Mycobacterium canettii* exhibits very small distances from Mtb H37Rv variant, as indicated by the red arrow in Figure 3. There is no indication of the Mtb complex gaining any RPL or RPS genes from the other *Mycobacterium* species included in this study, although it is possible that species not in the Mtb complex may participate in HGT.



**Figure 3.** TN93 distances between the Mtb H37Rv strain and the other 42 species in Figure 2. The 21 RPL genes (and their aligned lengths) are *rplA*(720), *rplB*(843), *rplC*(702), *rplD*(735), *rplE*(600), *rplF*(540), *rplI*(462), *rplJ*(657), *rplK*(435), *rplL*(405), *rplM*(444), *rplN*(369), *rplO*(447), *rplP*(417), *rplQ*(702), *rplR*(414), *rplS*(342), *rplT*(396), *rplU*(324), *rplV*(780), and *rplX*(327). The 18 genes (and their aligned lengths) are *rpsA*(1467), *rpsB*(906), *rpsC*(873), *rpsD*(606), *rpsE*(777), *rpsF*(291), *rpsG*(471), *rpsH*(399), *rpsI*(543), *rpsJ*(306), *rpsK*(444), *rpsL*(375), *rpsM*(375), *rpsO*(270), *rpsP*(633), *rpsQ*(456), *rpsS*(282), *rpsT*(267). The red arrow points to the distance curve between Mtb H37Rv and *Mycobacterium canettii*. The distance between Mtb H37Rv and other *Mycobacterium* variants is effectively zero for all ribosomal proteins.

The peak distance in RPL genes (Figure 3A) is due to indels in the 3' half of the *rplQ* gene and the start of the *rplR* gene that render alignment difficult. Similarly, the peak distance in the RPS genes (Figure 3B) is caused by many indels in the *rpsP* gene, creating alignment ambiguities. However, all fluctuations in evolutionary distances among ribosomal proteins do not obscure the pattern that the genomes in the Mtb complex are highly homogenous, with little indication of HGT between the Mtb complex and other *Mycobacterium* species.

### 2.3. Genomic Integrity in the Mtb Complex as Revealed by Ribosomal RNA Genes

Small subunit (ssu) ribosomes and the initiation tRNAs decode translation initiation signals in bacteria such as the Shine-Dalgarno sequence [56–60] and the start codon. Slow-growing *Mycobacterium tuberculosis* and *M. leprae* have only one 16S rRNA gene in their genomes, whereas fast-growing *M. smegmatis* features six. The still-faster-growing

*Escherichia coli* (NC\_000913) has not only seven 16S rRNA genes, but these genes also feature the strongest transcription signals at the  $-10$  and  $-35$  promoter sites, e.g., the *rrnB* gene in *E. coli* [61]. Ribosomal protein concentration is nearly perfectly correlated with growth rate in bacteria [62,63], highlighting the importance of rRNA availability in bacterial cells.

Because translation machineries are essential for any form of life, they are often targeted by antibacterial drugs. Aminoglycoside, which binds to the A-site of the ssu ribosome to inhibit translation, is one such drug against pathogens in the Mtb complex [64]. However, drug resistance has been reported in many Mtb isolates [65]. Similarly, oxazolidinones bind to the large subunit ribosome and inhibit translation initiation by preventing the proper positioning of the initiator tRNA at the A-site [51,52]. However, specific mutations (G2447T and G2576T) on the 23S rRNA [51,52] lead to drug resistance. It is interesting to learn if rRNA genes in the Mtb complex exhibit signatures of HGT. Point mutations at rRNA genes leading to antibiotic resistance have also been observed in other bacterial pathogens, such as *Neisseria gonorrhoeae* [66,67].

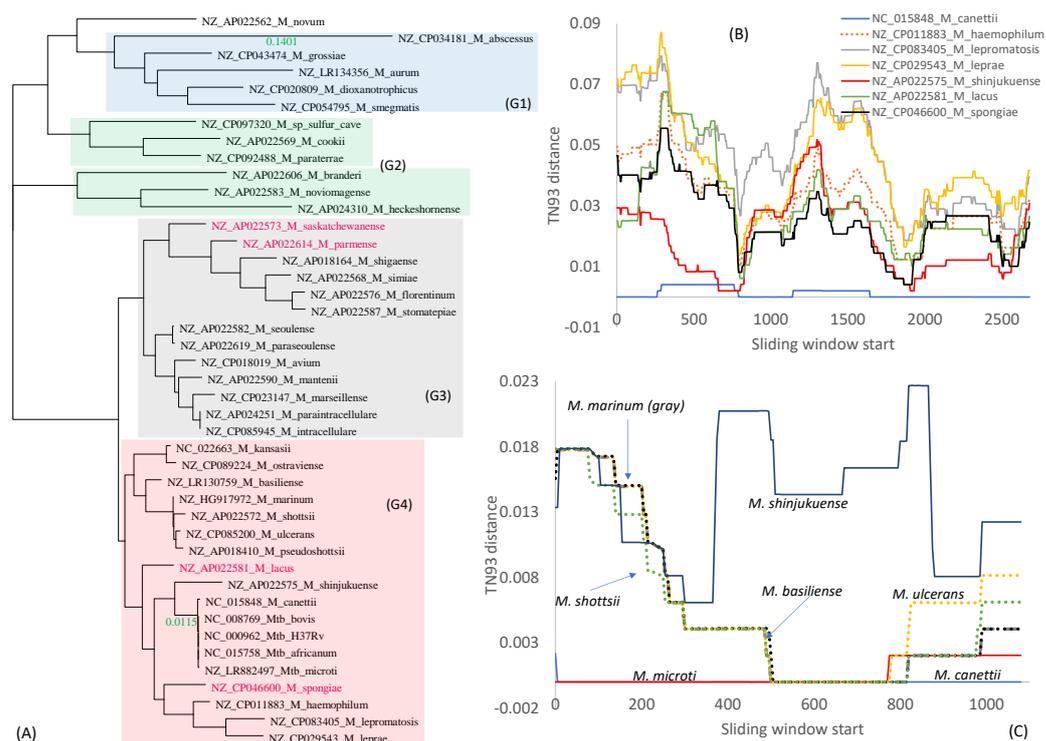
Multiple copies of rRNA genes within each *Mycobacterium* species are nearly identical and are always tightly clustered together in a phylogenetic tree. Therefore, only one copy of the gene was used to represent each species in Figure 4A. The 23S and 16S rRNA genes were individually aligned and then concatenated with DAMBE. The phylogeny from this concatenated alignment (Figure 4) is largely concordant with the RPL tree (Figure 2A) and the RPS tree (Figure 2B). However, several differences are notable. First, the G2 clade is not monophyletic. Second, four species (colored red in Figure 4A) have changed their phylogenetic position relative to trees in Figure 2. Third, while the branch length leading to *M. abscessus* is comparable to those in Figure 2, the branch length separating the Mtb complex from the other species is only 0.0115, much smaller than those in Figure 2. This means that the genomes of the Mtb complex are not as genetically unique from the perspective of rRNA genes as they are from the ribosomal protein genes. In other words, rRNA genes in the Mtb complex may be involved in HGT.

The DistPlot output for the 23S rRNA gene (Figure 4B) shows TN93 distances along a sliding window of 500 nt between the Mtb H37Rv and other species. All members of the Mtb complex have nearly identical 23S rRNA genes. Only *Mycobacterium canettii* exhibits a slight variation (Figure 4B). If any member of the Mtb complex had gained an rRNA gene (or a segment of rRNA genes) from genetically diverged species, then its distance from the Mtb H37Rv would increase. However, this increase is small (Figure 4), suggesting that, if HGT occurred, the donor and the receptor should be very closely related.

The only species with a segment of 23 rRNA identical to Mtb H37Rv is *M. shinjukuense* (red line in Figure 4B). However, the pattern suggests that it is *M. shinjukuense* that may have gained a segment from Mtb rather than the other way around, i.e., *M. shinjukuense* is a recipient rather than a donor in the putative HGT. *M. shinjukuense* is a recently identified pathogen that can cause severe pulmonary diseases [50]. Given the discovery of a plasmid that can shuttle information among slow-growing *Mycobacterium* species [68], it is possible for *M. shinjukuense* to gain genes from the Mtb complex.

Figure 4B omitted all fast-growing species because the distances between them and the Mtb complex never came close. This is consistent with the finding of the only plasmid in the slow-growing *Mycobacterium* species [68]. This plasmid is limited to the slow-growing mycobacteria species and cannot cross the boundary to fast-growing relatives.

The pattern in 16S rRNA is similar (Figure 4C). Several *Mycobacterium* species (*M. basiliense*, *M. ulcerans*, *M. marinum*, and *M. shottsii*) have a long segment of 16S rRNA identical to that in the Mtb H37Rv. This is not because the rRNA segment is particularly conserved because it differs much from the homologous segment in *M. shinjukuense* (Figure 4C). The similarity of the 16S rRNA segment is more likely due to a plasmid donating the segment to the Mtb complex and the four *Mycobacterium* species. However, the sequence difference is small overall, as one can see from the scale on the vertical axis, so the data set has limited power in detecting HGT events.

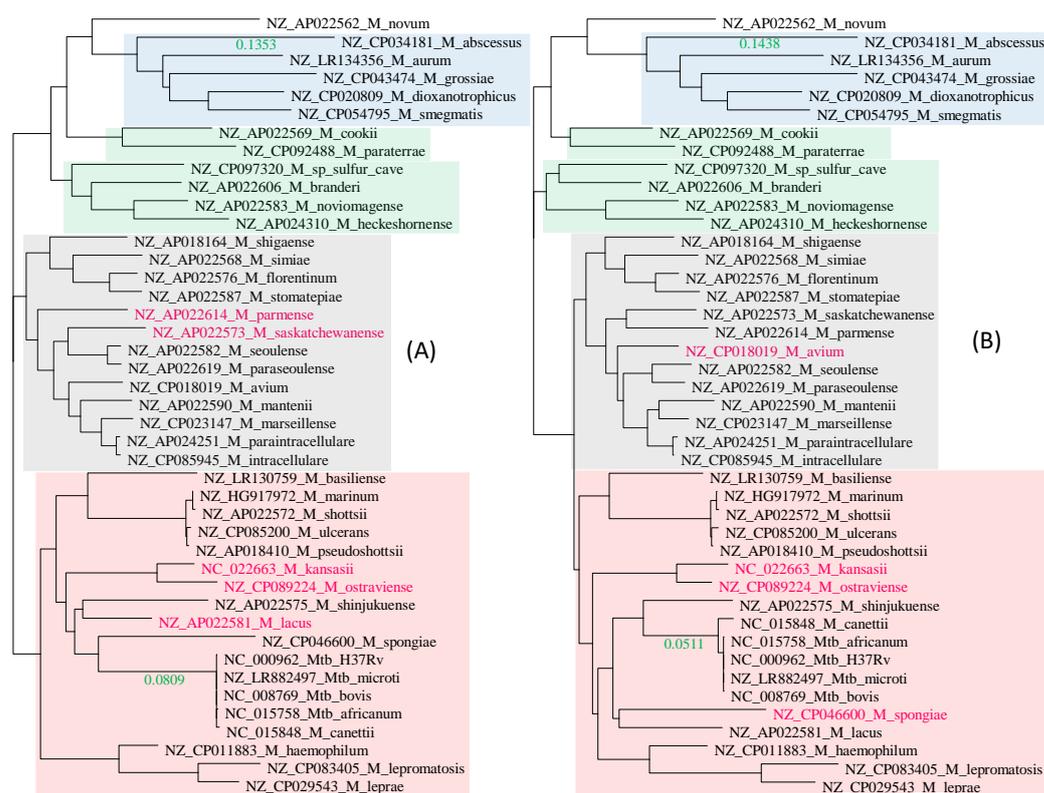


**Figure 4.** Genomic integrity as seen from rRNA genes. (A) Phylogeny from concatenated 23S + 16S rRNA genes. DistPlot of the 23S rRNA (B) and 16S rRNA (C) genes between Mtb H37Rv and its close relatives within the G4 clade. The G2 clade is not monophyletic as seen in Figure 2. Species that changed their phylogenetic position relative to tree in Figure 2 are colored red. The branch length to *M. abscessus* (colored green) is comparable to that in Figure 2, but the branch length (colored green) separating the Mtb complex from the rest is much shorter than that in Figure 2. TN93 distances were calculated over a sliding window of 500 nt.

#### 2.4. RNA Polymerase $\beta$ and $\beta'$ Subunit (*rpoB* and *rpoC*)

Transcription is essential for life, so the RNA polymerase  $\beta$  subunit (*rpoB*) gene has been targeted by antibiotics such as rifamycin including rifampin (rifampicin), rifapentine, and rifabutin. The mechanism of action is through the binding of rifampin to the *rpoB* protein, leading to the inhibition of transcription initiation and elongation. This is substantiated by five lines of evidence. First, the  $\beta$  subunit is involved in the resistance of several antibiotics, including rifampicin [69,70]. Second, rifampicin binds tightly to *E. coli* RNA polymerase [71]. Third, mutations in the  $\beta$  subunit often change rifampicin resistance [70,72]. Fourth, when the  $\beta$  subunit of RNA polymerase from a rifampicin-sensitive strain of *E. coli* is replaced by the  $\beta$  subunit from a rifampicin-resistant strain, the resulting RNA polymerase is rifampicin-resistant; in contrast, when the replacement is in the opposite direction, the resulting RNA polymerase is rifampicin-sensitive [69]. Fifth, when a particular mutation in the  $\beta$  subunit in a rifampin-resistant *Mycobacterium smegmatis* LR223 strain was introduced into the rifampin-sensitive LR223 strain of *M. smegmatis*, the latter became rifampin-resistant [73], with MIC increasing from 25 micrograms/mL to 200 micrograms/mL.

*Mycobacterial* RNA polymerases contain  $\beta$  and  $\beta'$  subunits in tandem *rpoBC* configuration. Both were extracted, aligned, and analyzed. The  $\beta$  and  $\beta'$  phylogenies (Figures 5A and 5B, respectively) are concordant with each other and similar to the trees in Figure 2. The Mtb complex is well separated from the rest of *Mycobacterium* species, with the branch length being 0.0809 for  $\beta$  and 0.0511 for  $\beta'$  (Figure 5). The Mtb complex is homogeneous in *rpoB*, but *M. canettii* differs slightly from the rest of the Mtb complex in *rpoC* (Figure 5B).



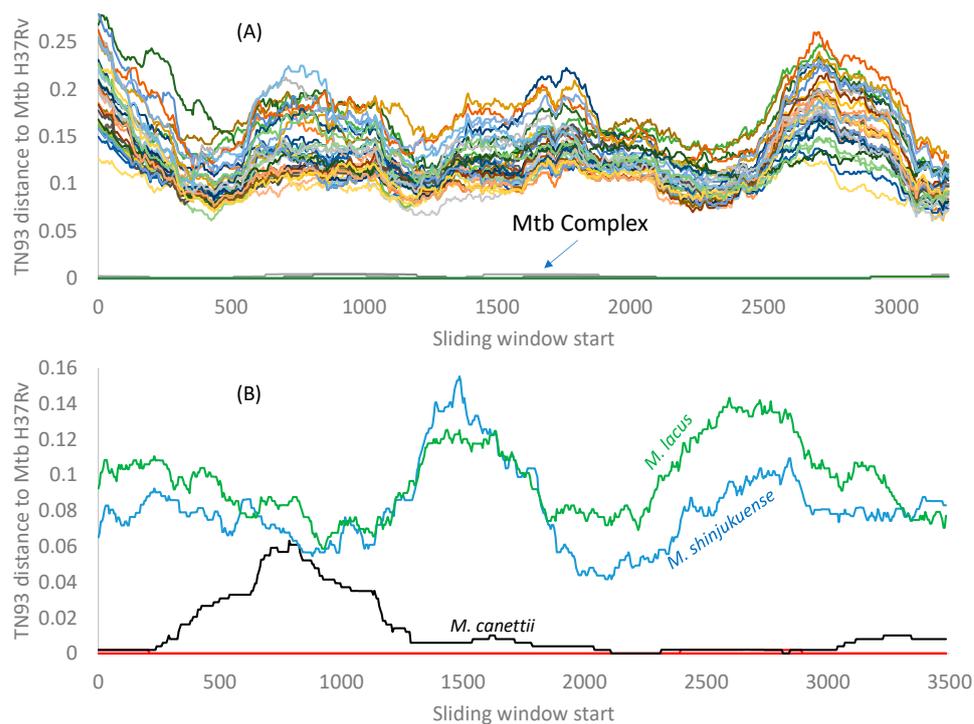
**Figure 5.** Phylogenetics of  $\beta$  (A) and  $\beta'$  (B) subunits of RNA polymerase. Species with phylogenetic position different from those in Figure 2 are colored in red. The branch lengths (colored green) leading to *M. abscessus* and separating the Mtb complex from the rest were shown.

The *rpoB* sequences within the Mtb complex are nearly identical to each other (Figure 6A), with TN93 distances all nearly 0 between the Mtb H37Rv and the other members of the Mtb complex. There is no indication that any member of the Mtb complex has gained either a *rpoB* or a segment of it from any other *Mycobacterium* species (Figure 6A).

The pattern for the  $\beta'$  subunit (Figure 6B) shows a difference between *M. canettii* and the rest of the Mtb complex that have nearly identical  $\beta'$  subunit sequences with their  $d_{i,MtbH37Rv}$  essentially all equal to 0 (the red line nearly overlapping the horizontal axis (Figure 6B)). The sites 738–1284 in *M. canettii* contain 33 nucleotide differences (with two being nonsynonymous) from the other members of the Mtb complex. It is likely that *M. canettii* gained this segment from other species through HGT. Previous experimental assays have shown *M. canettii* to be an efficient DNA recipient in HGT [74]. However, searching this segment against GenBank returned only those within the Mtb complex, so the putative donor of this segment remains unknown.

### 2.5. Genomic Integrity in the Mtb Complex as Revealed by *inhA* and *katG* Genes

The success of Mtb depends heavily on its mycolic acids, which form a thick shell that prevents antibiotics from reaching the cell membrane in the Mtb complex [75]. The mycolic acids serve two key functions in Mtb survival [76]: (1) they contribute to bacterial resistance to antibiotics, and (2) they contribute to bacterial evasion of the host immune system [77]. Once reaching the lungs, Mtb is phagocytosed by pulmonary macrophages [2,3,78]. The infected macrophage dies, and the clumps of Mtb are phagocytosed by other macrophages. Macrophages usually trigger the apoptosis pathway after engulfing parasites, killing themselves as well as the parasites they have engulfed. Mtb can not only survive macrophage apoptosis but also use macrophages (and some other immune cells) to build a granuloma that serves as a protected niche for Mtb to survive and proliferate inside the cell. The defensive fortress of Mtb is constructed mainly with mycolic acids as the key components [76,79], so genes essential for the synthesis of mycolic acids represent the Achilles' heel of Mtb [75].



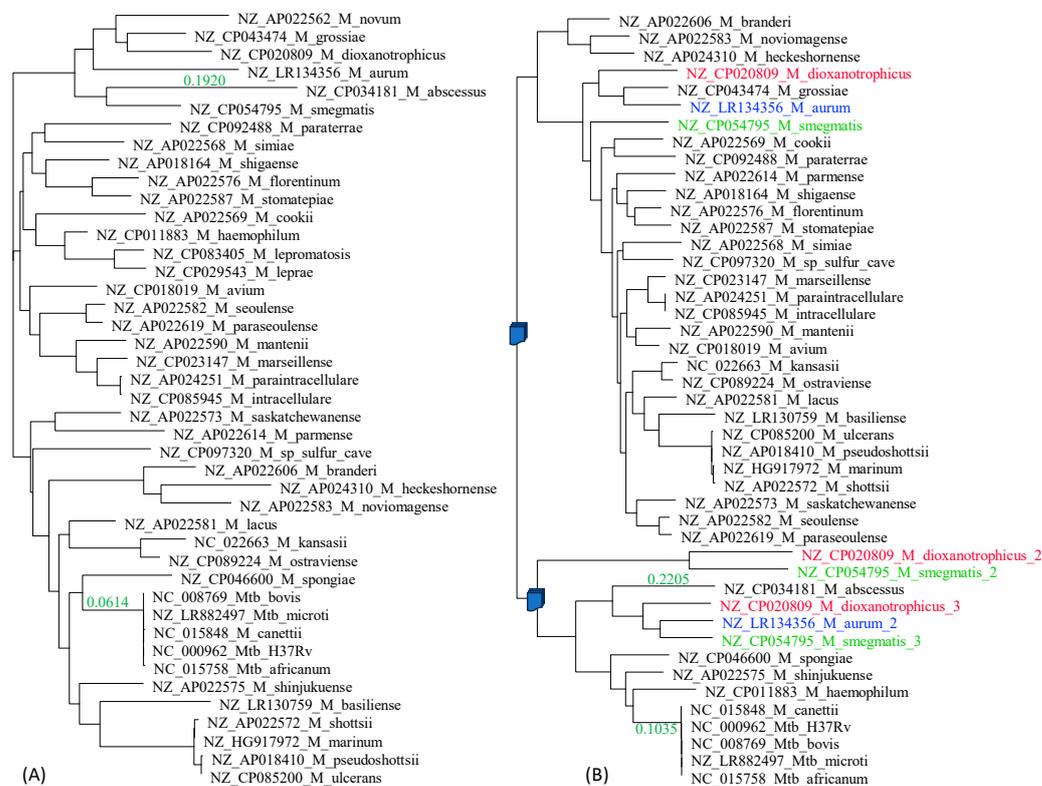
**Figure 6.** DistPlots for  $\beta$  (A) and  $\beta'$  (B) subunits of the RNA polymerase in *Mycobacterium* species. TN93 distances were calculated between Mtb H37Rv and each of the other species over a sliding window of 500 nt. All members in the Mtb complex were colored red in (B) except for *M. canettii*, which was colored black.

Isoniazid inhibits the synthesis of mycolic acids by inhibiting NADH-dependent enoyl-ACP reductase encoded by *InhA* that is essential for mycolic acid synthesis [80]. The drug is specific against *Mycobacterium* species because it requires activation by a catalase-peroxidase encoded by the *katG* gene in *Mycobacterium* genomes. The normal function of KatG is to catabolize peroxides and protect mycobacteria from the harmful effects of peroxides [81]. However, KatG also oxidizes isoniazid to form isonicotinoyl radicals that inhibit InhA activity essential for mycolic acid biosynthesis [82]. Resistance to isoniazid in *Mycobacterium* species can arise from mutations in *InhA* [80,83], mutations in *katG* [84–86], or the deletion of *katG* [85].

*InhA* and *katG* differ from previous genes in that they are not informational genes and are therefore more likely to participate in HGT. The two trees (Figure 7) indeed differ much from the topologies in Figure 2 reconstructed from ribosomal protein genes. However, the genetic homogeneity within the Mtb complex in these two genes is also obvious from the trees. The branch lengths within the Mtb complex are essentially zero, but the branch length separating the Mtb complex from the rest of the species is large, being 0.0614 for the *inhA* gene (Figure 7A) and 0.1035 for the *katG* gene (Figure 7B). Thus, there is no indication that these two genes in the Mtb complex have participated in recent HGT, although HGT involving these two genes might occur beyond the Mtb complex.

Three species have multiple versions of *katG* (colored in Figure 7B). Are they paralogues from gene duplication or xenologues from HGT? The distribution of *katG* copies on the tree (Figure 7B) can be explained feasibly by two gene duplication events followed by gene loss. The first gene duplication gave rise to paralogues P<sub>1</sub> and P<sub>2</sub>. *M. smegmatis*, *M. dioxanotrophicus* and *M. aurum* retained both paralogues in their genomes, but one paralogue, say P<sub>2</sub>, was lost in the ancestor of the blue and green clades (Figure 7B). Similarly, paralogue P<sub>1</sub> was lost in the ancestor of the pink clade (Figure 7B). The second gene duplication generated paralogues P<sub>21</sub> and P<sub>22</sub> (Figure 7B). However, only *M. smegmatis* and *M. dioxanotrophicus* retained both P<sub>21</sub> and P<sub>22</sub>, whereas three gene loss events occurred independently in the ancestor of the pink clade, in *M. abscessus*, and in *M. aurum*. This

gene-loss event appears to be ongoing. For example, among the three *katG* genes in *M. smegmatis*, one gene (corresponding to *M. smegmatis\_2* in Figure 7B) does not produce a functional protein, and the other two catalyze the decomposition of peroxides during the exponential and stationary phases, respectively [81]. Thus, the first gene is already a pseudogene and on its way to gene loss.



**Figure 7.** Phylogenies from *inhA* (A) and *katG* (B) genes. The branch lengths (colored green) leading to *M. abscessus* and separating the Mtb complex from the rest were shown next to the branch. Species with multiple versions of *katG* are highlighted with different colors, with an appended “\_1”, “\_2” and “\_3” to distinguish different versions of the gene within the same genome. The numbering does not imply relationship.

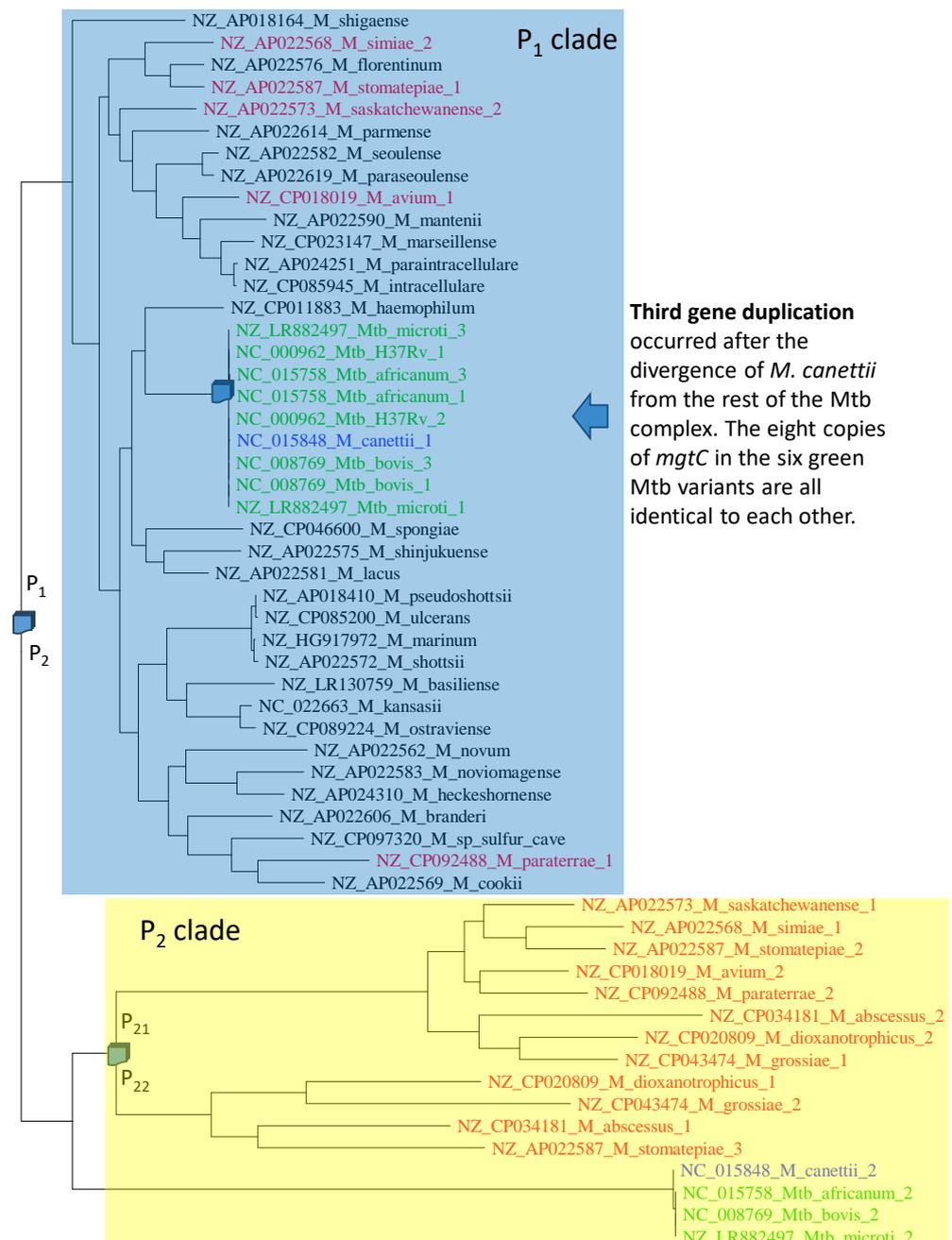
Note that one could also concoct a scheme of multiple HGT events to explain the distribution of gene copies. However, these HGT events would need to occur at different times between donor and recipient species in such a way that phylogenetic relationships are largely maintained.

## 2.6. Is *mgtC* Involved in HGT?

The *mgtC* is a virulence factor important for maintaining  $Mg^{2+}$  homeostasis [87–89]. It has been suggested previously to be involved in HGT [90]. However, it has not been examined critically in a phylogenetic framework, so it is not clear if multiple versions of the gene in the Mtb complex represent paralogues or xenologues.

The phylogeny of aligned *mgtC* sequences (Figure 8) is compatible with three gene duplication events. The first gene duplication generated paralogues  $P_1$  and  $P_2$ . The colored species retained both  $P_1$  and  $P_2$ , but the uncolored (black) species shaded in blue have lost  $P_2$ . The second gene duplication event gave rise to  $P_{21}$  and  $P_{22}$  (Figure 8), but only four species (*M. dioxanotrophicus*, *M. grossiae*, *M. abscessus* and *M. stomatopiae*) have retained both copies. *M. stomatopiae\_2* has become a pseudogene with a frameshifting mutation and is therefore on its way to complete loss. The third duplication must have occurred very recently, after the divergence of *M. canettii* from the rest of the Mtb complex (Figure 8). Two observations substantiate this recent duplication. First, not only are the two copies

of *mgtC* identical within *Mtb\_H37Rv*, *Mtb\_microti*, *M\_bovis*, and *M\_africanum*, but all eight of them are also identical to each other. Second, *M. canettii* has only one copy of  $P_1$ , suggesting that the gene duplication occurred after the divergence of *M. canettii* from the rest of the *Mtb* complex.



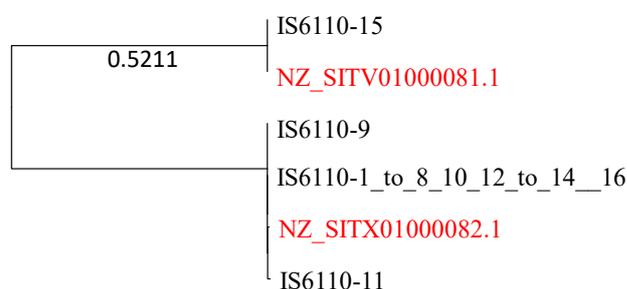
**Figure 8.** The phylogeny of *mgtC* is compatible with three gene duplication events. Species with multiple versions of *mgtC* are colored and appended with “\_1”, “\_2” and “\_3” to distinguish different versions of the gene within the same genome. The numbering does not imply relationship. The *Mtb* complex is colored in green and blue. Other species with multiple versions of *mgtC* are colored red.

### 2.7. Strong Evidence of HGT Involving Insertion Sequence IS6110

Insertion sequences occur frequently in mycobacteria, and their restricted distribution is often used for typing mycobacterial species or subspecies [91–94]. Insertion sequence IS6110 was previously thought to be specific and consequently was recognized from early on as a diagnostic signal for the *Mtb* complex in epidemiological studies [95–97].

Because IS6110 exists in multiple copies in the Mtb complex, e.g., 16 copies in the reference genome of *Mycobacterium tuberculosis* H37Rv (NC\_000962), their locations and the length of intervening sequences can be used to distinguish among variants within the Mtb complex until a homologue of IS6110 was found in *M. smegmatis* [19]. HGT was suggested for the presence of an IS6100 homologue in *M. smegmatis* [19]. However, because the IS6110 homologue in *M. smegmatis* is only 67% identical to the IS6110 in the Mtb complex [19], both ancient gene duplication and HGT could be invoked as valid hypotheses.

Here, I present strong evidence for HGT involving IS6110. The genomes of the Mtb complex feature 16 IS6110 sequences that vary slightly in sequence length (1355–1375 nt). They are numbered from 1 to 16, of which numbers 1–8, 10, 12–14 and 16 are identical. I used this sequence to query the 67 *M. smegmatis* genomes for highly similar sequences. Two *M. smegmatis* sequences are sufficient to make a strong case for HGT because they are essentially identical to the IS6110 sequences from the Mtb complex (Figure 9).



**Figure 9.** Phylogeny of the 16 IS6110 sequences from Mtb H37Rv (numbered from 1 to 16) and two homologues from *M. smegmatis* (colored red). The Mtb H37Rv sequences numbered 1–8, 10, 12–14 and 16 are identical and represented by one sequence. The tree is midpoint-rooted, with the number 0.5211 indicating the branch length.

There must be two very recent HGT events that generate the phylogenetic pattern in Figure 9. The Mtb complex and *M. smegmatis* are highly diverged from each other, as one can see from all previous phylogenies. The pattern in Figure 9 cannot be explained by ancient gene duplication because the divergence between NZ\_SITV01000081 and IS6110-15 and, or between NZ\_SITX01000082.1 and the other three Mtb H37Rv sequences, would correspond to the high divergence between *M. smegmatis* and Mtb H37Rv. However, the sequences within each of the two clades are essentially identical, with hardly any divergence.

### 3. Discussion

To facilitate discussion, I should first define the concept of horizontal transferability that was previously proposed but not explicitly defined [98,99]. Let  $e$  represent an HGT event involving a gene transferred from a donor to a recipient,  $p_e$  be the probability of such an HGT event in a given period, and  $p_f$  be the probability of the transferred gene establishing its function in the recipient. Horizontal transferability (HT) would be equal to  $HT = p_e p_f$ .

Both  $p_e$  and  $p_f$  depend on a variety of factors. In species within the family Pasteurellaceae and the genus *Neisseria* [100–102], gene sequences featuring an uptake signal sequence (or a DNA uptake sequence) have higher  $p_e$  than other sequences. For mycobacterial species, no such uptake sequences have been reported, and  $p_e$  may be assumed to be similar among genes. Thus, HT will depend mainly on  $p_f$ . Previous studies [98,99] suggest that functionally important genes, especially those forming protein complexes with many other partners such as ribosomal proteins and rRNA in ribosomes, would have a smaller  $p_f$  than functionally unimportant genes such as insertion sequences. While this study does not measure  $p_e$  or  $p_f$ , the framework is useful for discussion.

The phylogenetic pattern of the insertion sequence IS6110 suggests frequent HGT even between highly diverged mycobacterial species, such as between *M. smegmatis* and the Mtb complex. Whether these HGT events would be visible to us depends on  $p_f$ , i.e., how

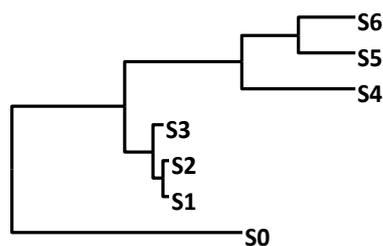
strong natural selection is against genes acquired through HGT. The Mtb complex appears to be highly adapted [103] and cannot tolerate mutations in functionally important genes. Many drug-resistant strains involve only one or a few mutations in specific genes, yet such a slightly modified gene would result in a significant reduction in fitness. For example, drug-sensitive strains of the Mtb complex have a generation time of about 20 h, but the corresponding drug-resistant strains have a generation time of about 30 h [104]. Thus, both functionally important and unimportant genes may be involved in HGT, but those strains acquiring a gene homologous to a functionally important gene are likely eliminated efficiently by selection, so we do not observe such HGT events. This is consistent with the functionally important genes studied in this paper, such as ribosomal protein genes, rRNA genes, and polymerase. However, for functionally unimportant genes such as insertion sequences, natural selection is expected to be weaker, leading to a relatively higher  $p_f$ , i.e., a higher chance for us to observe the HGT events.

The detection of frequent HGT events through insertion sequences between the Mtb complex and *M. smegmatis* is highly relevant to the battle against the drug-resistant Mtb variants. Genomic comparisons show that many virulence factors are shared between the Mtb complex and environmental mycobacteria [105], suggesting the potential of their exchange through HGT. The process could also generate new pathogens. For example, PPE25 and PPE26 from Mtb, which are not present in *M. smegmatis*, can significantly enhance the survival of *M. smegmatis* inside mouse macrophages when the two genes were expressed in *M. smegmatis* [106]. Similar results have been reported for other Mtb proteins, such as SA-5K [107], PE17 [108], Erp [109], Rv0431 [110], and isocitrate lyase [111] that enhance the survival of *M. smegmatis* in macrophages or other host cells. Multiple iterations of HGT, recombination, and natural selection could potentially generate new pathogens.

The results of this study do show the rarity of HGT involving functionally important genes, presumably because of a low  $p_f$ , and this could eliminate certain risks of HGT. For example, Mtb features both rifampicin-susceptible and rifampicin-resistant strains, and *M. smegmatis* also features rifampicin-susceptible and rifampicin-resistant strains, e.g., the LR222 and L223 strains, respectively [73]. Given that rifampicin-resistant mutations typically occur in the *rpoB* gene, there are at least two possibilities. First, rifampicin-resistant mutations occur independently at *rpoB* in different lineages of Mtb and *M. smegmatis*. Second, rifampicin-resistant mutations occur in the *rpoB* gene only once in either an Mtb lineage or an *M. smegmatis* lineage and are acquired by the other lineage through HGT. This study, however, does not favor the second hypothesis because *rpoB* is a functionally important gene with its protein interacting with multiple partners. Incorporating a diverged variant into the Mtb genome is likely deleterious, so such a recipient is likely eliminated efficiently by natural selection.

I wish to emphasize the point that phylogenetics can shed light on many biopharmaceutical problems. Here is one additional example. Suppose we have a gene tree with a fast-evolving clade and a slow-evolving clade (Figure 10). If this gene or its product is a drug target, then the drug target is good if the pathogen is within the conserved clade (e.g., S1, S2, or S3), but not good if the pathogen is in the fast-evolving clade (S4, S5, or S6). A fast-evolving clade implies a relatively rapid accumulation of mutations, which facilitate the origin of drug resistance. If the pathogen is in the slow-evolving clade, then it has a reduced chance of evolving drug resistance. Such a simple analysis is often not done with the selection of drug targets.

Comparison of phylogeny and DistPlot patterns among genes can also help refine the selection of drug targets. For example, given the DistPlot for ribosomal protein genes and that for rRNA genes, one would choose ribosomal protein genes over rRNA genes as drug targets, everything else being equal. Similarly, given the DistPlot for *rpoB* and *rpoC* genes, one would choose *rpoB* over *rpoC* as a drug target.



**Figure 10.** A tree with a fast-evolving clade (S1 to S3) and a slow-evolving clade (S4 to S6).

#### 4. Materials and Methods

There are 101 *Mycobacterium* species with sequenced genomes in GenBank. However, only 41 are annotated as “complete”. These do not include species that were traditionally in *Mycobacterium* but have recently been assigned new generic names, such as *Mycolicibacterium smegmatis*, *M. aurum*, and *Mycobacteroides abscessus*. In this 41 genomes, *Mycobacterium tuberculosis* was represented only by the strain H37Rv. After including *Mycolicibacterium smegmatis*, *M. aurum* and *Mycobacteroides abscessus*, as well as four additional members within the Mtb complex (i.e., *M. microti*, *M. bovis*, *M. africanum*, and *M. canettii*), there are a total of 48 genomes. However, five genomes exhibit phylogenetic incongruence between the tree built with individually aligned and concatenated large ribosomal protein genes and that built with aligned and concatenated small ribosomal protein genes. I cannot decide if this is due to sequencing errors or genomic fluidity. For simplicity, I just removed these five genomes, leaving 43 genomes included in this study.

To approximate the species tree, I extracted 21 large ribosomal protein (RPL) genes and 18 small ribosomal protein (RPS) genes that are shared among all 43 species using DAMBE [112]. These genes were individually aligned with MAFFT [113], with the slow but accurate LINSI option. The 21 RPL genes were then concatenated to create an RPL43.FAS file. The 18 RPS genes were similarly concatenated to create an RPS43.FAS file. These two supermatrix files were included as Supplemental Files. Aligned and concatenated 23S and 16S rRNA genes from the 43 genomes are also included as a Supplemental File 23S\_16S.FAS. Maximum likelihood trees were reconstructed with PhyML [114] with the GTR +  $\Gamma$  substitution model and simultaneous optimization of tree topology, branch lengths, and model parameters.

Drug development typically targets genes with essential functions for pathogen survival and reproduction. Such genes include 23S and 16S rRNA genes, *rpoB*, *rpoC*, *inhA*, and *katG*. Two genes were previously reported to be involved in HGT: the *mgtC* gene [90] and the insertion sequence IS6110 [19], which were included in this study. Gene trees were reconstructed from these individual genes to check if they might participate in HGT. All these genes are present in all 43 *Mycobacterium* species included in this study.

For the DistPlot method, simultaneous estimated TN93 distances [29] were computed with a sliding window of 500 nt. The simultaneous estimation is more robust and accurate than the commonly used independent estimation [29,115]. The use of the TN93 distance instead of the more general GTR distance is because the window size of 500 nt often does not include all six different nucleotide replacements (two transitions and four transversions) needed for parameter estimation in the GTR model.

For identifying HGT between the Mtb complex and *M. smegmatis* involving the insertion sequence IS6110, I downloaded the 67 genomes available for *M. smegmatis*, created a local BLAST database, and searched the IS6110 sequences from the Mtb complex against *M. smegmatis* genomes. Matches could arise from HGT or ancient gene duplication followed by subsequent lineage sorting. However, if the IS6110 sequences in *M. smegmatis* and the Mtb complex are identical, then HGT is favored.

#### 5. Conclusions

As revealed by insertion sequences, horizontal gene transfer (HGT) occurs frequently between *Mycobacterium tuberculosis* (Mtb) and other mycobacterial species, even between

Mtb and the highly diverged *M. smegmatis*. This raises the possibility that drug resistance may evolve in fast-replicating and large populations of nontuberculosis mycobacteria and get imported into Mtb through HGT. Fortunately, the horizontal transferability of genes depends on the functional importance of the gene. Functionally important genes are typically well adapted with their interacting partners, so a recipient of a diverged homologue through HGT is most likely deleterious and would be eliminated efficiently by natural selection. Phylogenetic patterns characterized by genomic data suggest that functionally important genes are rarely involved in HGT, but functionally unimportant sequences, such as insertion sequences, are readily observed in HGT. Because genes that have been chosen as drug targets are typically functionally important genes, they do not seem to participate in HGT, so there is no evidence suggesting that drug resistance in Mtb is gained through HGT. The analysis also shows that *mgcC*, which was previously suggested to be horizontally transferred, has phylogenetic patterns more consistent with gene duplication and subsequent lineage sorting than with HGT.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/antibiotics12091367/s1>, File S1: the 21 aligned and concatenated RPL genes, File S2: the 18 aligned and concatenated RPS genes, and File S3: the concatenated 23S\_16S rRNA genes from 43 *Mycobacterium tuberculosis* genomes.

**Funding:** This research was funded by Discovery Grant from Natural Science and Engineering Research Council (NSERC) of Canada, grant number RGPIN/2018-03878.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** I thank P. Aris and A. Kruglikov for discussion. J. Y. Xia contributed significantly to data compilation and analysis.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Cole, S.T. Comparative and functional genomics of the *Mycobacterium tuberculosis* complex. *Microbiology* **2002**, *148*, 2919–2928. [[CrossRef](#)] [[PubMed](#)]
2. Zhu, M.; Dai, X. On the intrinsic constraint of bacterial growth rate: *M. tuberculosis*'s view of the protein translation capacity. *Crit. Rev. Microbiol.* **2018**, *44*, 455–464. [[CrossRef](#)] [[PubMed](#)]
3. Gengenbacher, M.; Kaufmann, S.H. *Mycobacterium tuberculosis*: Success through dormancy. *FEMS Microbiol. Rev.* **2012**, *36*, 514–532. [[CrossRef](#)] [[PubMed](#)]
4. Ford, C.B.; Lin, P.L.; Chase, M.R.; Shah, R.R.; Iartchouk, O.; Galagan, J.; Mohaideen, N.; Ioerger, T.R.; Sacchetti, J.C.; Lipsitch, M.; et al. Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. *Nat. Genet.* **2011**, *43*, 482–486. [[CrossRef](#)] [[PubMed](#)]
5. Centers for Disease Control and Prevention. Emergence of *Mycobacterium tuberculosis* with extensive resistance to second-line drugs—worldwide, 2000–2004. *MMWR Morb. Mortal. Wkly. Rep.* **2006**, *55*, 301–305.
6. Gandhi, N.R.; Moll, A.; Sturm, A.W.; Pawinski, R.; Govender, T.; Lalloo, U.; Zeller, K.; Andrews, J.; Friedland, G. Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet* **2006**, *368*, 1575–1580. [[CrossRef](#)] [[PubMed](#)]
7. Oppong, Y.E.A.; Phelan, J.; Perdigão, J.; Machado, D.; Miranda, A.; Portugal, I.; Viveiros, M.; Clark, T.G.; Hibberd, M.L. Genome-wide analysis of *Mycobacterium tuberculosis* polymorphisms reveals lineage-specific associations with drug resistance. *BMC Genom.* **2019**, *20*, 252. [[CrossRef](#)]
8. Udawadia, Z.F.; Amale, R.A.; Ajbani, K.K.; Rodrigues, C. Totally drug-resistant tuberculosis in India. *Clin. Infect. Dis.* **2012**, *54*, 579–581. [[CrossRef](#)]
9. Velayati, A.A.; Masjedi, M.R.; Farnia, P.; Tabarsi, P.; Ghanavi, J.; ZiaZarifi, A.H.; Hoffner, S.E. Emergence of new forms of totally drug-resistant tuberculosis bacilli: Super extensively drug-resistant tuberculosis or totally drug-resistant strains in Iran. *Chest* **2009**, *136*, 420–425. [[CrossRef](#)]
10. Kempker, R.R.; Kipiani, M.; Mirtskhulava, V.; Tukvadze, N.; Magee, M.J.; Blumberg, H.M. Acquired Drug Resistance in *Mycobacterium tuberculosis* and Poor Outcomes among Patients with Multidrug-Resistant Tuberculosis. *Emerg. Infect. Dis.* **2015**, *21*, 992–1001. [[CrossRef](#)]

11. Cillóniz, C.; Garcia-Vidal, C.; Ceccato, A.; Torres, A. Antimicrobial Resistance among *Streptococcus pneumoniae*. In *Antimicrobial Resistance in the 21st Century*; Fong, I.W., Shlaes, D., Drlica, K., Eds.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 13–38.
12. Ercoli, G.; Fernandes, V.E.; Chung, W.Y.; Wanford, J.J.; Thomson, S.; Bayliss, C.D.; Straatman, K.; Crocker, P.R.; Dennison, A.; Martinez-Pomares, L.; et al. Intracellular replication of *Streptococcus pneumoniae* inside splenic macrophages serves as a reservoir for septicemia. *Nat. Microbiol.* **2018**, *3*, 600–610. [[CrossRef](#)]
13. Tonkin-Hill, G.; Ling, C.; Chaguza, C.; Salter, S.J.; Hinfonthong, P.; Nikolaou, E.; Tate, N.; Pastusiak, A.; Turner, C.; Chewapreecha, C.; et al. Pneumococcal within-host diversity during colonization, transmission and treatment. *Nat. Microbiol.* **2022**, *7*, 1791–1804. [[CrossRef](#)] [[PubMed](#)]
14. Stevens, K.E.; Sebert, M.E. Frequent beneficial mutations during single-colony serial transfer of *Streptococcus pneumoniae*. *PLoS Genet.* **2011**, *7*, e1002232. [[CrossRef](#)] [[PubMed](#)]
15. Eldholm, V.; Balloux, F. Antimicrobial Resistance in *Mycobacterium tuberculosis*: The Odd One Out. *Trends Microbiol.* **2016**, *24*, 637–648. [[CrossRef](#)] [[PubMed](#)]
16. Gallant, J.; Mouton, J.; Ummels, R.; Ten Hagen-Jongman, C.; Kriel, N.; Pain, A.; Warren, R.M.; Bitter, W.; Heunis, T.; Sampson, S.L. Identification of gene fusion events in *Mycobacterium tuberculosis* that encode chimeric proteins. *NAR Genom. Bioinform.* **2020**, *2*, lqaa033. [[CrossRef](#)] [[PubMed](#)]
17. Hirsh, A.E.; Tsolaki, A.G.; DeRiemer, K.; Feldman, M.W.; Small, P.M. Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4871–4876. [[CrossRef](#)] [[PubMed](#)]
18. Warner, D.F.; Rock, J.M.; Fortune, S.M.; Mizrahi, V. DNA Replication Fidelity in the *Mycobacterium tuberculosis* Complex. *Adv. Exp. Med. Biol.* **2017**, *1019*, 247–262. [[PubMed](#)]
19. Coros, A.; DeConno, E.; Derbyshire, K.M. IS6110, a *Mycobacterium tuberculosis* complex-specific insertion sequence, is also present in the genome of *Mycobacterium smegmatis*, suggestive of lateral gene transfer among mycobacterial species. *J. Bacteriol.* **2008**, *190*, 3408–3410. [[CrossRef](#)]
20. Mariani, F.; Piccolella, E.; Colizzi, V.; Rappuoli, R.; Gross, R. Characterization of an IS-like element from *Mycobacterium tuberculosis*. *J. Gen. Microbiol.* **1993**, *139*, 1767–1772. [[CrossRef](#)] [[PubMed](#)]
21. Singh, A.; Gaur, M.; Sharma, V.; Khanna, P.; Bothra, A.; Bhaduri, A.; Mondal, A.K.; Dash, D.; Singh, Y.; Misra, R. Comparative Genomic Analysis of *Mycobacteriaceae* Reveals Horizontal Gene Transfer-Mediated Evolution of the CRISPR-Cas System in the *Mycobacterium tuberculosis* Complex. *mSystems* **2021**, *6*, e00934-20. [[CrossRef](#)]
22. Razavi, M.; Kristiansson, E.; Flach, C.F.; Larsson, D.G.J. The Association between Insertion Sequences and Antibiotic Resistance Genes. *mSphere* **2020**, *5*, e00418-20. [[CrossRef](#)] [[PubMed](#)]
23. Xia, X. *Comparative Genomics*; Springer: Berlin/Heidelberg, Germany, 2013; p. VIII. 67p.
24. Rivera, M.C.; Jain, R.; Moore, J.E.; Lake, J.A. Genomic evidence for two functionally distinct gene classes. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6239–6244. [[CrossRef](#)]
25. Sorek, R.; Zhu, Y.; Creevey, C.J.; Francino, M.P.; Bork, P.; Rubin, E.M. Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* **2007**, *318*, 1449–1452. [[CrossRef](#)] [[PubMed](#)]
26. Bull, J.J.; Springman, R.; Molineux, I.J. Compensatory evolution in response to a novel RNA polymerase: Orthologous replacement of a central network gene. *Mol. Biol. Evol.* **2007**, *24*, 900–908. [[CrossRef](#)]
27. Hudson, R.R. Gene trees, species trees and the segregation of ancestral alleles. *Genetics* **1992**, *131*, 509–513. [[CrossRef](#)] [[PubMed](#)]
28. Page, R.D.M. Introduction. In *Tangled Trees: Phylogeny, Cospeciation and Coevolution*; Page, R.D.M., Ed.; University of Chicago Press: Chicago, IL, USA, 2003; pp. 1–21.
29. Xia, X. *A Mathematical Primer of Molecular Phylogenetics*; CRC Press: New York, NY, USA, 2020; p. 380.
30. Kinsella, R.J.; Fitzpatrick, D.A.; Creevey, C.J.; McInerney, J.O. Fatty acid biosynthesis in *Mycobacterium tuberculosis*: Lateral gene transfer, adaptive evolution, and gene duplication. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10320–10325. [[CrossRef](#)]
31. Gupta, R.S.; Lo, B.; Son, J. Phylogenomics and Comparative Genomic Studies Robustly Support Division of the Genus *Mycobacterium* into an Emended Genus *Mycobacterium* and Four Novel Genera. *Front. Microbiol.* **2018**, *9*, 67. [[CrossRef](#)]
32. Yamada, H.; Chikamatsu, K.; Aono, A.; Murata, K.; Miyazaki, N.; Kayama, Y.; Bhatt, A.; Fujiwara, N.; Maeda, S.; Mitarai, S. Fundamental Cell Morphologies Examined With Cryo-TEM of the Species in the Novel Five Genera Robustly Correlate with New Classification in Family *Mycobacteriaceae*. *Front. Microbiol.* **2020**, *11*, 562395. [[CrossRef](#)]
33. Yamada, H.; Yamaguchi, M.; Igarashi, Y.; Chikamatsu, K.; Aono, A.; Murase, Y.; Morishige, Y.; Takaki, A.; Chibana, H.; Mitarai, S. *Mycolicobacterium smegmatis*, Basonym *Mycobacterium smegmatis*, Expresses Morphological Phenotypes Much more Similar to *Escherichia coli* than *Mycobacterium tuberculosis* in Quantitative Structome Analysis and CryoTEM Examination. *Front. Microbiol.* **2018**, *9*, 1992. [[CrossRef](#)]
34. Matsumoto, Y.; Kinjo, T.; Motooka, D.; Nabeya, D.; Jung, N.; Uechi, K.; Horii, T.; Iida, T.; Fujita, J.; Nakamura, S. Comprehensive subspecies identification of 175 nontuberculous mycobacteria species based on 7547 genomic profiles. *Emerg. Microbes Infect.* **2019**, *8*, 1043–1053. [[CrossRef](#)]
35. Nouioui, I.; Carro, L.; García-López, M.; Meier-Kolthoff, J.P.; Woyke, T.; Kyrpides, N.C.; Pukall, R.; Klenk, H.-P.; Goodfellow, M.; Göker, M. Genome-Based Taxonomic Classification of the Phylum *Actinobacteria*. *Front. Microbiol.* **2018**, *9*, 2007. [[CrossRef](#)]
36. He, Y.; Wei, K.; Si, K.; Mathieu, J.; Li, M.; Alvarez, P.J.J. Whole-Genome Sequence of the 1,4-Dioxane-Degrading Bacterium *Mycobacterium dioxanotrophicus* PH-06. *Genome Announc.* **2017**, *5*, e00625-17. [[CrossRef](#)] [[PubMed](#)]

37. Cech, J.S.; Hartman, P.; Slosárek, M.; Chudoba, J. Isolation and identification of a morpholine-degrading bacterium. *Appl. Environ. Microbiol.* **1988**, *54*, 619–621. [[CrossRef](#)]
38. Poupin, P.; Mazure, N.; Truffaut, N. Morpholine degradation by strain *Mycobacterium aurum* MOI: Improvement of cells growth and morpholine degradation rate by cells immobilization. In *Progress in Biotechnology*; Wijffels, R.H., Buitelaar, R.M., Bucke, C., Tramper, J., Eds.; Elsevier: Amsterdam, The Netherlands, 1996; Volume 11, pp. 770–776.
39. Lelovic, N.; Mitachi, K.; Yang, J.; Lemieux, M.R.; Ji, Y.; Kurosu, M. Application of *Mycobacterium smegmatis* as a surrogate to evaluate drug leads against *Mycobacterium tuberculosis*. *J. Antibiot.* **2020**, *73*, 780–789. [[CrossRef](#)] [[PubMed](#)]
40. Reytrat, J.M.; Kahn, D. *Mycobacterium smegmatis*: An absurd model for tuberculosis? *Trends Microbiol.* **2001**, *9*, 472–474. [[CrossRef](#)]
41. Paniz-Mondolfi, A.E.; Greninger, A.L.; Ladutko, L.; Brown-Elliott, B.A.; Vasireddy, R.; Jakubiec, W.; Vasireddy, S.; Wallace Jr, R.J.; Simmon, K.E.; Dunn, B.E.; et al. *Mycobacterium grossiae* sp. nov., a rapidly growing, scotochromogenic species isolated from human clinical respiratory and blood culture specimens. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 4345–4351. [[CrossRef](#)] [[PubMed](#)]
42. Dedrick, R.M.; Guerrero-Bustamante, C.A.; Garlena, R.A.; Russell, D.A.; Ford, K.; Harris, K.; Gilmour, K.C.; Soothill, J.; Jacobs-Sera, D.; Schooley, R.T.; et al. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. *Nat. Med.* **2019**, *25*, 730–733. [[CrossRef](#)]
43. Xia, X. Optimizing Protein Production in Therapeutic Phages against a Bacterial Pathogen, *Mycobacterium abscessus*. *Drugs Drug Candidates* **2023**, *2*, 189–209. [[CrossRef](#)]
44. Change, Y.T.; Andersen, R.N.; Vaituzis, Z. Growth of *Mycobacterium lepraemurium* in cultures of mouse peritoneal macrophages. *J. Bacteriol.* **1967**, *93*, 1119–1131. [[CrossRef](#)]
45. van Spanning, R.J.M.; Guan, Q.; Melkonian, C.; Gallant, J.; Polerecky, L.; Flot, J.-F.; Brandt, B.W.; Braster, M.; Iturbe Espinoza, P.; Aerts, J.W.; et al. Methanotrophy by a *Mycobacterium* species that dominates a cave microbial ecosystem. *Nat. Microbiol.* **2022**, *7*, 2089–2100. [[CrossRef](#)]
46. Kazda, J.; Stackebrandt, E.; Smida, J.; Minnikin, D.E.; Daffe, M.; Parlett, J.H.; Pitulle, C. *Mycobacterium cookii* sp. nov. *Int. J. Syst. Bacteriol.* **1990**, *40*, 217–223. [[CrossRef](#)]
47. Lee, H.; Lee, S.A.; Lee, I.K.; Yu, H.K.; Park, Y.G.; Jeong, J.; Lee, S.H.; Kim, S.R.; Hyun, J.W.; Kim, K.; et al. *Mycobacterium paraterrae* sp. nov. recovered from a clinical specimen: Novel chromogenic slow growing mycobacteria related to *Mycobacterium terrae* complex. *Microbiol. Immunol.* **2010**, *54*, 46–53. [[CrossRef](#)] [[PubMed](#)]
48. van Ingen, J.; Turenne, C.Y.; Tortoli, E.; Wallace, R.J., Jr.; Brown-Elliott, B.A. A definition of the *Mycobacterium avium* complex for taxonomical and clinical purposes, a review. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 3666–3677. [[CrossRef](#)] [[PubMed](#)]
49. Goring, S.M.; Wilson, J.B.; Risebrough, N.R.; Gallagher, J.; Carroll, S.; Heap, K.J.; Obradovic, M.; Loebinger, M.R.; Diel, R. The cost of *Mycobacterium avium* complex lung disease in Canada, France, Germany, and the United Kingdom: A nationally representative observational study. *BMC Health Serv. Res.* **2018**, *18*, 700. [[CrossRef](#)] [[PubMed](#)]
50. Taoka, T.; Shinohara, T.; Hatakeyama, N.; Iwamura, S.; Murase, Y.; Mitarai, S.; Ogushi, F. *Mycobacterium Shinjukuense* Pulmonary Disease Progressed to Pleuritis after Iatrogenic Pneumothorax: A Case Report. *J. Clin. Tuberc. Other Mycobact. Dis.* **2020**, *19*, 100160. [[CrossRef](#)]
51. Shinabarger, D. Mechanism of action of the oxazolidinone antibacterial agents. *Expert Opin. Investig. Drugs* **1999**, *8*, 1195–1202. [[CrossRef](#)]
52. Foti, C.; Piperno, A.; Scala, A.; Giuffrè, O. Oxazolidinone Antibiotics: Chemical, Biological and Analytical Aspects. *Molecules* **2021**, *26*, 4280. [[CrossRef](#)]
53. Ma, K.C.; Mortimer, T.D.; Duckett, M.A.; Hicks, A.L.; Wheeler, N.E.; Sánchez-Busó, L.; Grad, Y.H. Increased power from conditional bacterial genome-wide association identifies macrolide resistance mutations in *Neisseria gonorrhoeae*. *Nat. Commun.* **2020**, *11*, 5374. [[CrossRef](#)]
54. Gregory, S.T.; Dahlberg, A.E. Erythromycin resistance mutations in ribosomal proteins L22 and L4 perturb the higher order structure of 23 S ribosomal RNA. *J. Mol. Biol.* **1999**, *289*, 827–834. [[CrossRef](#)]
55. Tamura, K.; Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **1993**, *10*, 512–526.
56. Shine, J.; Dalgarno, L. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 1342–1346. [[CrossRef](#)] [[PubMed](#)]
57. Shine, J.; Dalgarno, L. Determinant of cistron specificity in bacterial ribosomes. *Nature* **1975**, *254*, 34–38. [[CrossRef](#)]
58. Hui, A.; de Boer, H.A. Specialized ribosome system: Preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 4762–4766. [[CrossRef](#)] [[PubMed](#)]
59. Steitz, J.A.; Jakes, K. How ribosomes select initiator regions in mRNA: Base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 4734–4738. [[CrossRef](#)]
60. Taniguchi, T.; Weissmann, C. Inhibition of Qbeta RNA 70S ribosome initiation complex formation by an oligonucleotide complementary to the 3' terminal region of *E. coli* 16S ribosomal RNA. *Nature* **1978**, *275*, 770–772. [[CrossRef](#)] [[PubMed](#)]
61. Ross, W.; Gosink, K.K.; Salomon, J.; Igarashi, K.; Zou, C.; Ishihama, A.; Severinov, K.; Gourse, R.L. A Third Recognition Element in Bacterial Promoters: DNA Binding by the  $\alpha$  Subunit of RNA Polymerase. *Science* **1993**, *262*, 1407–1413. [[CrossRef](#)] [[PubMed](#)]
62. Klumpp, S.; Scott, M.; Pedersen, S.; Hwa, T. Molecular crowding limits translation and cell growth. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16754–16759. [[CrossRef](#)] [[PubMed](#)]

63. Karpinets, T.V.; Greenwood, D.J.; Sams, C.E.; Ammons, J.T. RNA: Protein ratio of the unicellular organism as a characteristic of phosphorous and nitrogen stoichiometry and of the cellular requirement of ribosomes for protein synthesis. *BMC Biol.* **2006**, *4*, 30. [[CrossRef](#)]
64. Chambers, H.F.; Sande, M.A. Antimicrobial agents—The aminoglycosides. In *Goodman and Gilman's the Pharmacological Basis of Therapeutics*, 9th ed.; Goodman, L.S., Limbird, L.E., Milinoff, P.B., Gilman, A.G., Hardmann, J.G., Eds.; McGraw-Hill: New York, NY, USA, 1996; pp. 1103–1121.
65. Shcherbakov, D.; Akbergenov, R.; Matt, T.; Sander, P.; Andersson, D.I.; Böttger, E.C. Directed mutagenesis of *Mycobacterium smegmatis* 16S rRNA to reconstruct the in vivo evolution of aminoglycoside resistance in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **2010**, *77*, 830–840. [[CrossRef](#)]
66. Chisholm, S.A.; Dave, J.; Ison, C.A. High-level azithromycin resistance occurs in *Neisseria gonorrhoeae* as a result of a single point mutation in the 23S rRNA genes. *Antimicrob. Agents Chemother.* **2010**, *54*, 3812–3816. [[CrossRef](#)]
67. Pham, C.D.; Nash, E.; Liu, H.; Schmerer, M.W.; Sharpe, S.; Woods, G.; Roland, B.; Schlanger, K.; St Cyr, S.B.; Carlson, J.; et al. Atypical Mutation in *Neisseria gonorrhoeae* 23S rRNA Associated with High-Level Azithromycin Resistance. *Antimicrob. Agents Chemother.* **2021**, *65*, e00885-20. [[CrossRef](#)] [[PubMed](#)]
68. Ummels, R.; Abdallah, A.M.; Kuiper, V.; Aâjoud, A.; Sparrius, M.; Naeem, R.; Spaink, H.P.; Soolingen, D.V.; Pain, A.; Bitter, W. Identification of a Novel Conjugative Plasmid in Mycobacteria That Requires Both Type IV and Type VII Secretion. *mBio* **2014**, *5*, e01744-14. [[CrossRef](#)] [[PubMed](#)]
69. Heil, A.; Zillig, W. Reconstitution of bacterial DNA-dependent RNA-polymerase from isolated subunits as a tool for the elucidation of the role of the subunits in transcription. *FEBS Lett.* **1970**, *11*, 165–168. [[CrossRef](#)]
70. Rabussay, D.; Zillig, W. A rifampicin resistant ma-polymerase from *E. coli* altered in the  $\beta$ -subunit. *FEBS Lett.* **1969**, *5*, 104–106. [[CrossRef](#)]
71. Wehrli, W.; Knüsel, F.; Schmid, K.; Staehelin, M. Interaction of rifamycin with bacterial RNA polymerase. *Proc. Natl. Acad. Sci. USA* **1968**, *61*, 667–673. [[CrossRef](#)]
72. Boyd, D.H.; Zillig, W.; Scaife, F.J.G. Reference mutations for the  $\beta$  subunit of RNA polymerase. *Mol. Gen. Genet. MGG* **1974**, *130*, 315–320. [[CrossRef](#)] [[PubMed](#)]
73. Miller, L.P.; Crawford, J.T.; Shinnick, T.M. The rpoB gene of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **1994**, *38*, 805–811. [[CrossRef](#)]
74. Madacki, J.; Orgeur, M.; Mas Fiol, G.; Frigui, W.; Ma, L.; Brosch, R. ESX-1-Independent Horizontal Gene Transfer by *Mycobacterium tuberculosis* Complex Strains. *mBio* **2021**, *12*, e00965-21. [[CrossRef](#)]
75. Nataraj, V.; Varela, C.; Javid, A.; Singh, A.; Besra, G.S.; Bhatt, A. Mycolic acids: Deciphering and targeting the Achilles' heel of the tubercle bacillus. *Mol. Microbiol.* **2015**, *98*, 7–16. [[CrossRef](#)]
76. Lee, W.; Engels, B. The protonation state of catalytic residues in the resting state of KasA revisited: Detailed mechanism for the activation of KasA by its own substrate. *Biochemistry* **2014**, *53*, 919–931. [[CrossRef](#)]
77. Zhao, J.; Siddiqui, S.; Shang, S.; Bian, Y.; Bagchi, S.; He, Y.; Wang, C.R. Mycolic acid-specific T cells protect against *Mycobacterium tuberculosis* infection in a humanized transgenic mouse model. *eLife* **2015**, *4*, e08525. [[CrossRef](#)]
78. Khan, S.R.; Manialawy, Y.; Siraki, A.G. Isoniazid and host immune system interactions: A proposal for a novel comprehensive mode of action. *Br. J. Pharmacol.* **2019**, *176*, 4599–4608. [[CrossRef](#)]
79. Asselineau, J.; Lederer, E. Structure of the mycolic acids of *Mycobacteria*. *Nature* **1950**, *166*, 782–783. [[CrossRef](#)]
80. Quémard, A.; Sacchetti, J.C.; Dessen, A.; Vilcheze, C.; Bittman, R.; Jacobs, W.R., Jr.; Blanchard, J.S. Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. *Biochemistry* **1995**, *34*, 8235–8241. [[CrossRef](#)]
81. Iwao, Y.; Nakata, N. Roles of the three *Mycobacterium smegmatis* katG genes for peroxide detoxification and isoniazid susceptibility. *Microbiol. Immunol.* **2018**, *62*, 158–167. [[CrossRef](#)] [[PubMed](#)]
82. Timmins, G.S.; Deretic, V. Mechanisms of action of isoniazid. *Mol. Microbiol.* **2006**, *62*, 1220–1227. [[CrossRef](#)]
83. Parikh, S.L.; Xiao, G.; Tonge, P.J. Inhibition of InhA, the Enoyl Reductase from *Mycobacterium tuberculosis*, by Triclosan and Isoniazid. *Biochemistry* **2000**, *39*, 7645–7650. [[CrossRef](#)] [[PubMed](#)]
84. Zhang, Y.; Garbe, T.; Young, D. Transformation with katG restores isoniazid-sensitivity in *Mycobacterium tuberculosis* isolates resistant to a range of drug concentrations. *Mol. Microbiol.* **1993**, *8*, 521–524. [[CrossRef](#)] [[PubMed](#)]
85. Altamirano, M.; Marostenmaki, J.; Wong, A.; FitzGerald, M.; Black, W.A.; Smith, J.A. Mutations in the catalase-peroxidase gene from isoniazid-resistant *Mycobacterium tuberculosis* isolates. *J. Infect. Dis.* **1994**, *169*, 1162–1165. [[CrossRef](#)]
86. Heym, B.; Alzari, P.M.; Honoré, N.; Cole, S.T. Missense mutations in the catalase-peroxidase gene, katG, are associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **1995**, *15*, 235–245. [[CrossRef](#)]
87. Alix, E.; Blanc-Potard, A.B. MgtC: A key player in intramacrophage survival. *Trends Microbiol.* **2007**, *15*, 252–256. [[CrossRef](#)]
88. Choi, S.; Choi, E.; Cho, Y.J.; Nam, D.; Lee, J.; Lee, E.J. The *Salmonella* virulence protein MgtC promotes phosphate uptake inside macrophages. *Nat. Commun.* **2019**, *10*, 3326. [[CrossRef](#)] [[PubMed](#)]
89. Lee, J.W.; Lee, E.J. Regulation and function of the *Salmonella* MgtC virulence protein. *J. Microbiol.* **2015**, *53*, 667–672. [[CrossRef](#)]
90. Alix, E.; Godreuil, S.; Blanc-Potard, A.B. Identification of a Haarlem genotype-specific single nucleotide polymorphism in the mgtC virulence gene of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* **2006**, *44*, 2093–2098. [[CrossRef](#)]
91. Guilhot, C.; Gicquel, B.; Davies, J.; Martín, C. Isolation and analysis of IS6120, a new insertion sequence from *Mycobacterium smegmatis*. *Mol. Microbiol.* **1992**, *6*, 107–113. [[CrossRef](#)]

92. Ichikawa, K.; Yagi, T.; Moriyama, M.; Inagaki, T.; Nakagawa, T.; Uchiya, K.I.; Nikai, T.; Ogawa, K. Characterization of *Mycobacterium avium* clinical isolates in Japan using subspecies-specific insertion sequences, and identification of a new insertion sequence, ISMav6. *J. Med. Microbiol.* **2009**, *58*, 945–950. [[CrossRef](#)]
93. Park, H.T.; Park, H.E.; Jung, Y.H.; Yoo, H.S. An ISMap02-like insertion sequence in *Mycobacterium* spp. interferes with specific detection of *Mycobacterium avium* subsp. *paratuberculosis*. *Vet. Microbiol.* **2018**, *216*, 1–6. [[CrossRef](#)] [[PubMed](#)]
94. Thierry, D.; Brisson-Noël, A.; Vincent-Lévy-Frébault, V.; Nguyen, S.; Guesdon, J.L.; Gicquel, B. Characterization of a *Mycobacterium tuberculosis* insertion sequence, IS6110, and its application in diagnosis. *J. Clin. Microbiol.* **1990**, *28*, 2668–2673. [[CrossRef](#)] [[PubMed](#)]
95. McEvoy, C.R.; Falmer, A.A.; Gey van Pittius, N.C.; Victor, T.C.; van Helden, P.D.; Warren, R.M. The role of IS6110 in the evolution of *Mycobacterium tuberculosis*. *Tuberculosis* **2007**, *87*, 393–404. [[CrossRef](#)]
96. Thierry, D.; Cave, M.D.; Eisenach, K.D.; Crawford, J.T.; Bates, J.H.; Gicquel, B.; Guesdon, J.L. IS6110, an IS-like element of *Mycobacterium tuberculosis* complex. *Nucleic Acids Res.* **1990**, *18*, 188. [[CrossRef](#)]
97. van Embden, J.D.; Cave, M.D.; Crawford, J.T.; Dale, J.W.; Eisenach, K.D.; Gicquel, B.; Hermans, P.; Martin, C.; McAdam, R.; Shinnick, T.M.; et al. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: Recommendations for a standardized methodology. *J. Clin. Microbiol.* **1993**, *31*, 406–409. [[CrossRef](#)] [[PubMed](#)]
98. Nakamura, Y.; Itoh, T.; Matsuda, H.; Gojobori, T. Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nat. Genet.* **2004**, *36*, 760–766. [[CrossRef](#)] [[PubMed](#)]
99. Creevey, C.J.; Doerks, T.; Fitzpatrick, D.A.; Raes, J.; Bork, P. Universally distributed single-copy genes indicate a constant rate of horizontal transfer. *PLoS ONE* **2011**, *6*, e22099. [[CrossRef](#)]
100. Smith, H.O.; Tomb, J.F.; Dougherty, B.A.; Fleischmann, R.D.; Venter, J.C. Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome. *Science* **1995**, *269*, 538–540. [[CrossRef](#)]
101. Maughan, H.; Wilson, L.A.; Redfield, R.J. Bacterial DNA uptake sequences can accumulate by molecular drive alone. *Genetics* **2010**, *186*, 613–627. [[CrossRef](#)]
102. Frye, S.A.; Nilsen, M.; Tønjum, T.; Ambur, O.H. Dialects of the DNA uptake sequence in *Neisseriaceae*. *PLoS Genet.* **2013**, *9*, e1003458. [[CrossRef](#)] [[PubMed](#)]
103. Bishai, W. The *Mycobacterium tuberculosis* genomic sequence: Anatomy of a master adaptor. *Trends Microbiol.* **1998**, *6*, 464–465. [[CrossRef](#)]
104. Mukherjee, T.; Goswami, A.; Chakraborty, U.; Majumdar, M.; Sinha, S.; Pal, N.K. A Study on Generation Time of Sensitive and Resistant *Mycobacterium tuberculosis* Isolates. *J. Evol. Med. Dent. Sci.* **2019**, *8*, 2489–2494. [[CrossRef](#)]
105. Wang, J.; Behr, M.A. Building a better bacillus: The emergence of *Mycobacterium tuberculosis*. *Front. Microbiol.* **2014**, *5*, 139. [[CrossRef](#)]
106. Mi, Y.; Bao, L.; Gu, D.; Luo, T.; Sun, C.; Yang, G. *Mycobacterium tuberculosis* PPE25 and PPE26 proteins expressed in *Mycobacterium smegmatis* modulate cytokine secretion in mouse macrophages and enhance mycobacterial survival. *Res. Microbiol.* **2017**, *168*, 234–243. [[CrossRef](#)] [[PubMed](#)]
107. Batoni, G.; Bottai, D.; Maisetta, G.; Pardini, M.; Boschi, A.; Florio, W.; Esin, S.; Campa, M. Involvement of the *Mycobacterium tuberculosis* secreted antigen SA-5K in intracellular survival of recombinant *Mycobacterium smegmatis*. *FEMS Microbiol. Lett.* **2001**, *205*, 125–129. [[CrossRef](#)] [[PubMed](#)]
108. Li, Z.; Liu, H.; Li, H.; Dang, G.; Cui, Z.; Song, N.; Wang, Q.; Liu, S.; Chen, L. PE17 protein from *Mycobacterium tuberculosis* enhances *Mycobacterium smegmatis* survival in macrophages and pathogenicity in mice. *Microb. Pathog.* **2019**, *126*, 63–73. [[CrossRef](#)] [[PubMed](#)]
109. Ganaie, A.A.; Trivedi, G.; Kaur, A.; Jha, S.S.; Anand, S.; Rana, V.; Singh, A.; Kumar, S.; Sharma, C. Interaction of Erp Protein of *Mycobacterium tuberculosis* with Rv2212 Enhances Intracellular Survival of *Mycobacterium smegmatis*. *J. Bacteriol.* **2016**, *198*, 2841–2852. [[CrossRef](#)] [[PubMed](#)]
110. Deng, G.; Zhang, F.; Yang, S.; Kang, J.; Sha, S.; Ma, Y. *Mycobacterium tuberculosis* Rv0431 expressed in *Mycobacterium smegmatis*, a potentially mannosylated protein, mediated the immune evasion of RAW 264.7 macrophages. *Microb. Pathog.* **2016**, *100*, 285–292. [[CrossRef](#)]
111. Li, J.M.; Li, N.; Zhu, D.Y.; Wan, L.G.; He, Y.L.; Yang, C. Isocitrate lyase from *Mycobacterium tuberculosis* promotes survival of *Mycobacterium smegmatis* within macrophage by suppressing cell apoptosis. *Chin. Med. J.* **2008**, *121*, 1114–1119. [[CrossRef](#)]
112. Xia, X. DAMBE7: New and improved tools for data analysis in molecular biology and evolution. *Mol. Biol. Evol.* **2018**, *35*, 1550–1552. [[CrossRef](#)]
113. Katoh, K.; Toh, H. Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.* **2008**, *9*, 286–298. [[CrossRef](#)]
114. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)]
115. Xia, X. Nucleotide Substitution Models and Evolutionary Distances. In *Bioinformatics and the Cell: Modern Computational Approaches in Genomics, Proteomics and Transcriptomics*; Springer: Cham, Switzerland, 2018; pp. 269–314.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.