Genetic Recombination of the *rpoB* Gene as a Mechanism of the *Mycobacterium tuberculosis* Resistance to Rifampin

Levani Lomidze, Mamuka Kotetishvili*

Faculty of Natural Sciences and Engineering, Ilia State University
Tbilisi, Georgia

*Corresponding author's email: mamuka_kotetishvili [AT] iliauni.edu.ge

ABSTRACT--- ***The molecular genetic mechanisms whereby Mycobacterim tuberculosis develops resistance to rifampin (RIF) are poorly understood. Using the tests for selection and genetic recombination, we analyzed rpoB loci of the M. tuberculosis strains harboring the synonymous single nucleotide polymorphisms (sSNPs) to detect the horizontal gene transfer (HGT) of this genetic locus. The rpoB codons*** GGC_{507}***,*** GGG_{534}***,*** GGC_{556} ***and*** GGG_{566} ***were revealed to be the hotspots for the multiple sSNPs. The strains carrying multiple nucleotide polymorphisms across these codons were from either Iran or the Republic of Belarus, exhibiting regional patterns of the rpoB sSNPs distribution. The*** d_N/d_S ***ratios were*** ≤ 0.7363 ***for the rpoB loci that harbored these sSNPs, showing no significant selective pressures on the amino acid changes. The Index of Association (*** I_A***) being 1.110, for the DNA substitutions of the rpoB loci with the sSNPs, revealed that this genetic locus is not clonal in M. tuberculosis. Moreover, the splitstree analysis resulted in the parallelograms with the strong fit and bootstrap values (95-100 and 85-100 respectively), exhibiting the HGT events in some of these rpoB loci. Thus, it is strongly suggested that the HGT of rpoB of M. tuberculosis contributes to the generation of this species resistance to RIF.***

Keywords – *Mycobacterium tuberculosis*, resistance to rifampin, genetic recombination

1. INTRODUCTION

Mycobacterium tuberculosis is a causative agent of tuberculosis (TB), posing a significant threat to global public health. One-third of the total world's population is infected with this bacterium being responsible for ~ 8 million cases of active TB and ~ 3 million human deaths annually [1, 2]. TB infections have become particularly challenging to the prevention and treatment strategies of this disease due to a continuing emergence of the multidrug-resistant and extensively drug-resistant *M. tuberculosis* strains [3, 4].

M. tuberculosis exhibits a low mutation rate [5, 6] and a slow replication rate [7]. In this light, it is not entirely clear how the bacterium acquires resistance to antibiotics, especially in the face of concurrent multiple drug treatment [7]. Patient noncompliance [3, 4, 8], inappropriate drug regimens and dosing [5, 6, 9, 10], as well as primary infections with drug resistant strains [11, 12] have been well defined to be the most common risk factors for treatment failure due to antibiotic resistance [7]. However, the genetic mechanisms that contribute to a development of the drug resistant TB infections have not been yet comprehensively determined.

The *M. tuberculosis* resistance to RIF is a surrogate marker for the multidrug-resistant TB, because, ~ 90% of RIF-resistant phenotypes of this species demonstrate their resistance also at least to isoniazid [1, 13, 14]. The known molecular mechanisms underlying the *M. tuberculosis* resistance to RIF lead to the specific missense mutations of the *rpoB* gene encoding the RNA polymerase β subunit [14-18]. The Efflux pump induction appears to be a general step in the evolution of mycobacterial drug resistance. Most frequently, this takes place with the strains that have increased risk for high rates of drug resistance mutations [19-21]. Under these scenarios, it remains unclear whether the HGT plays any role in a development of the drug resistant TB. A genome of *M. tuberculosis* exhibits a lack of genetic recombination [22, 23]. It was suggested that genetic exchange between cells could be rare in strains of the *M. tuberculosis* complex [24]. However, a possible role of the HGT in the generation of the drug resistant TB was still hypothesized [25], although, relevant evidence was still missing. In this study, using the statistical analysis *of* the *rpoB* sSNPs, we attempt to delineate events of the HGT of the loci of this gene of *M. tuberculosis*.

2. MATERIALS AND METHODS

DNA sequence records and BLAST analyses. We analyzed the DNA sequence records for the *rpoB* gene of *M. tuberculosis* available in the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/). The DNA sequence of *rpoB* of the avirulent *M. tuberculosis* strain H37Ra (NCBI accession no.: CP000611) was used to

retrieve the *rpoB* DNA sequences of this species, utilizing the Basic Local Alignment Search Tool (BLAST) in the NCBI nucleotide database. The nucleotide search was performed using the megablast and blastn algorithms as recommended [26]. The following parameters were utilized in the BLAST analysis: Expect threshold 10; Word size 28, with automatically adjusting of the parameters for short input DNA sequences; The scoring parameters were 1,-2 for match/mismatch, and linear for gap costs; Low complexity regions were respectively filtered, and segments of the query sequences with a low compositional complexity were masked off by the DUST program as implemented in the NCBI database.

s**SNPs analyses and test for selection**. The DNA sequences of *rpoB* of the *M. tuberculosis* strains retrieved from the NCBI nucleotide database were aligned using the alignment program ClustalX (v. 2.0) [27]. The open reading frame, for the *rpoB* DNA sequences of *M. tuberculosis*, was determined using the ORF finder of the NCBI (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The sSNPs were identified within the aligned *rpoB* nucleotide sequences, by determining the silent mutations across the codons of this gene. In the analysis, we applied the codon numbering system as described previously [15]. We used the web-based program of the nucleic acid and protein sequence analyses Molecular Toolkit using the Java applet (1.0) (http://www.vivo.colostate.edu/molkit/) for a restriction mapping of the *rpoB* DNA sequences that contained two or more sSNPs. The restriction mapping analysis was conducted to determine whether there were the close linkages between the sSNPs positions and the DNA cleavage sites of the DNA restriction endonucleases of the different microbial agents in the *rpoB* loci of the *M. tuberculosis* populations. To estimate probability of matching the DNA cleavage sites of each DNA restriction endonuclease either the 5 prime ends or the 3 prime ends of the sSNPs in the *rpoB* gene, we used the binomial probability distribution:

$$P(r; p, n) = p^r (1-p)^{n-r} \frac{n!}{r!\,(n-r)!}$$

where, *r* is a number of the DNA cleavage sites of the DNA restriction endonucleases matching either the 5 prime ends or the 3 prime ends of the sSNPs of the *rpoB* loci of the *M. tuberculosis* strains; *n* is a total number of the DNA cleavage sites of each DNA restriction endonuclease in the *rpoB* gene of *M. tuberculosis*; *p* is probability of matching the DNA cleavage sites of the DNA restriction endonucleases the 5 prime end or the 3 prime end of any sSNP in the *rpoB* gene loci of *M. tuberculosis*.

The d_N/d_S (the number of non-synonymous changes per non-synonymous site/the number of synonymous changes per synonymous site) ratios were determined in the *rpoB* loci of *M. tuberculosis* strains, using the test for selection as described previously [28]. Subsequently, we calculated an average GC content of the *rpoB* loci of these strains.

Statistical tests for genetic recombination. The I_A [29] was determined for the DNA substitutions of *rpoB* of the *M. tuberculosis* strains as follows: $I_A=V_O/V_E$-1, where, V_O is the observed variance of K, and V_E is the expected variance of K, where K represents a number of the DNA polymorphisms at which two or more strains differ in the *rpoB* gene of *M. tuberculosis*. The splits-decomposition method [30], as implemented in the program SplitsTree4 (v. 4.12.8) [31], was applied to identify HGT patterns in the *rpoB* DNA sequences of *M. tuberculosis* carrying two or more sSNPs. A bootstrapping of the parallelograms constructed by SplitsTree4, was performed using 1,000 resampling of the data. A bootstrap value being ≥80 for a node, and a fit value being ≥90 for a parallelogram were considered statistically significant in these SplitsTree analyses.

3. RESULTS

BLAST and *rpoB* sSNPs analyses. Using the BLAST analysis, we could identify 23 different codons of the *rpoB* gene of 60 *M. tuberculosis* strains that carried sSNPs. A codon analysis of the clustalX-generated alignments of the above *rpoB* DNA sequences allowed us to select 33 *M. tuberculosis* strains (Table 1) with the *rpoB* alleles that contained two or more sSNPs.

Table 1. The NCBI DNA sequence records for the *rpoB* gene of *M. tuberculosis* selected in this study.

NCBI accession no.	Strain designation	Reference
CP000611	H37Ra	55
EF628311	MDR-Belarus-407	55
EF628312	MDR-Belarus-2548 M	55
EF628313	MDR-Belarus	55
EF628315	Belarus-7285	55
EF628316	Belarus-139	55
EF628318	MDR-Belarus	55

EF628319	MDR-85-vitebsk	55
EF628324	MDR-Belarus-442 Magilov	55
EF628326	Belarus-894 vitebsk	55
EF628339	Iran-441	Unpublished
EF628340	MDR-Iran-163 Zabol	Unpublished
EF628341	Iran-163 Tehran	Unpublished
EF628342	Iran-3708	Unpublished
EF628343	Iran-3548	Unpublished
EF628344	Iran-90 Tehran	Unpublished
EF628346	Iran-29 Tehran	Unpublished
EF628347	Iran-290 Mashad	Unpublished
EF628351	MDR-Iran-161 Zabol	Unpublished
EF628352	633 Tehran	Unpublished
EF628354	19 Kerman	Unpublished
EF628355	10-2 Tehran	Unpublished
EF628356	108 Tehran	Unpublished
EF628357	Iran-167 Tehran	Unpublished
EF628359	Iran-103 Zabol	Unpublished
EF628360	MDR-Iran-36-Teh	Unpublished
EF628361	MDR-Iran-663 kerman	Unpublished
EF628362	MDR-Iran-600 Tehran	Unpublished
EF628363	Iran-173 Zabol	Unpublished
EF628364	Iran-3542	Unpublished
EF628365	MDR-Iran-159 Zabol	Unpublished
EF628366	MDR-Iran-165 Zabol	Unpublished
EF628367	MDR-Iran-303-281 Mashad	Unpublished
EF628368	MDR-Iran-33 PII	Unpublished

The NCBI nucleotide sequence records, for these strains, were limited to ≤ 410-bp internal DNA fragments of *rpoB* including the 81-bp rifampin resistance-determining region (RRDR). Table 2 displays the *rpoB* codons predominantly affected by the sSNPs in the *M. tuberculosis* strains.

Table 2. The *rpoB* codons of the *M. tuberculosis* strains predominantly affected by the sSNPs, representing the adjacent regions of the DNA cleavage sites of the DNA restriction endonucleases NlaIV, AsuI, HaeIII, and CviJI

NCBI accession no.	The DNA cleavage sties of the DNA restriction endonucleases, and the *rpoB* codons affected by the sSNPs			
	NlaIV [GGN^NC]	AsuI [G^GNCC] HaeIII [GG^CC] NlaIV [GGN^NC] CviJI [RG^CY]	HaeIII [GG^CC] CviJI [RG^CY]	AsuI [G^GNCC] HaeIII [GG^CC] NlaIV [GGN^NCC] CviJI [RG^CY]
CP000611*	GGC_{507}	GGG_{534}	GGC_{556}	GGG_{566}
EF628311	GGC→GGC	GGG→GGC	GGC→GGC	GGG→GGC
EF628312	GGC→GGC	GGG→GGC	GGC→GGC	GGG→GGC
EF628313	GGC→GGC	GGG→GGC	GGC→GGC	GGG→GGC
EF628315	GGC→GGC	GGG→GGG	GGC→GGC	GGG→GGC
EF628316	GGC→GGC	GGG→GGG	GGC→GGC	GGG→GGC
EF628318	GGC→GGC	GGG→GGG	GGC→GGC	GGG→GGC
EF628319	GGC→GGC	GGG→GGC	GGC→GGC	GGG→GGG

EF628324	GGC→GGC	GGG→GGC	GGC→GGC	GGG→GGG
EF628326	GGC→GGC	GGG→GGG	GGC→GGC	GGG→GGC
EF628339	GGC→GGC	GGG→GGG	GGC→GGC	GGG→GGT
EF628340	GGC→GGC	GGG→GGG	GGC→GGC	GGG→GGT
EF628341	GGC→GGT	GGG→GGG	GGC→GGC	GGG→GGT
EF628342	GGC→GGC	GGG→GGG	GGC→GGC	GGG→GGT
EF628343	GGC→GGC	GGG→GGC	GGC→GGC	GGG→GGT
EF628344	GGC→GGC	GGG→GGG	GGC→GGC	GGG→GGC
EF628346	GGC→GGC	GGG→GGC	GGC→GGC	GGG→GGC
EF628347	GGC→GGC	GGG→GGC	GGC→GGC	GGG→GGC
EF628351	GGC→GGC	GGG→GGA	GGC→GGC	GGG→GGG
EF628352	GGC→GGC	GGG→GGG	GGC→GGC	GGG→GGC
EF628354	GGC→GGT	GGG→GGA	GGC→GGC	GGG→GGG
EF628355	GGC→AGT	GGG→GGA	GGC→GGC	GGG→GGG
EF628356	GGC→GGC	GGG→GGA	GGC→GGC	GGG→GGG
EF628357	GGC→GGC	GGG→GGA	GGC→GGC	GGG→GGG
EF628359	GGC→GGC	GGG→GGG	GGC→GGC	GGG→GGC
EF628360	GGC→GGC	GGG→GGG	GGC→GGC	GGG→GGC
EF628361	GGC→GGC	GGG→GGC	GGC→GGC	GGG→GGG
EF628362	GGC→GGC	GGG→GGT	GGC→GGC	GGG→GGG
EF628363	GGC→GGC	GGG→GGT	GGC→GGC	GGG→GGC
EF628364	GGC→GGT	GGG→GGC	GGC→GGG	GGG→GGC
EF628365	GGC→GGT	GGG→GGA	GGC→GGC	GGG→GGG
EF628366	GGC→GGC	GGG→GGA	GGC→GGC	GGG→GGG
EF628367	GGC→GGC	GGG→GGA	GGC→GGC	GGG→GGC
EF628368	GGC→GAT	GGG→GGA	GGC→GGC	GGG→GGT

*the NCBI accession no. for the *M. tuberculosis* strain H37Ra

As shown in Table 2, among the identified sSNPs, the G→C change was most abundant (52.0%), which was followed by the G→A substitutions (18.0%). The other DNA substitutions reflecting the sSNPs were G→T (16.0%), C→T (12.0%), and C→G (2.0%) changes. In the *rpoB* loci of the selected *M. tuberculosis* strains, the multiple sSNPs were detected at the following codons: GGC_{507}, GGG_{534}, GGC_{556} and GGG_{566}. Using the program Molecular Toolkit, we found that the 3 prime ends of these codons were followed by the DNA cleavage sites of one or more DNA restriction endonucleases (Table 2). In particular, the DNA cleavage site of the DNA restriction endonuclease NlaIV was adjacent to the 3 prime ends of the codons GGC_{507}, GGG_{534}, and GGG_{566} in a wild-type allele of the *rpoB* gene of *M. tuberculosis* (the NCBI accession no. CP000611). In addition, the DNA cleavage sites of the DNA restriction endonucleases HaeIII and CviJI were adjacent to the 3 prime ends of the codons GGG_{534}, GGC_{556} and GGG_{566}. Besides, the DNA cleavage sites of the DNA restriction endonuclease AsuI were found to be adjacent to the 3 prime ends of the *rpoB* codons GGG_{534} and GGG_{566}. Table 3 displays the probability values for matching the DNA cleavage sites of each DNA restriction endonuclease the adjacent sites of the 3 prime ends of the above four codons across the *rpoB* loci of the *M. tuberculosis* strains.

Table 3. The probability (*P*) estimates for matching the DNA cleavage sites of the DNA restriction endonucleases the 3 prime ends of the *rpoB* codons GGC_{507}, GGG_{534}, GGC_{556}, and GGG_{566}, and the d_N/d_S and I_A values for the *rpoB* DNA sequences of the *M. tuberculosis* strains carrying the multiple sSNPs.

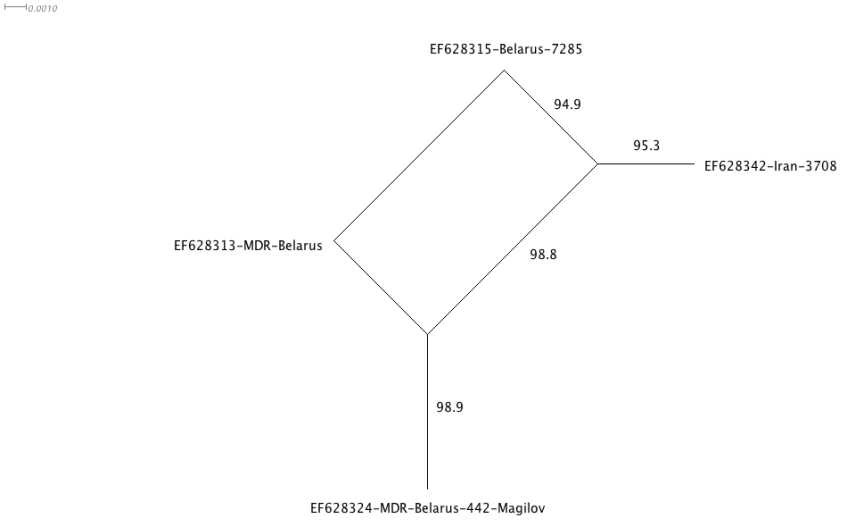
P value				d_N/d_S	I_A
CviJI [RG^CY]	NlaIV [GGN^N]	AsuI [G^GNCC]	HaeIII [GG^CC]		
0.0485	0.0026	0.0375	0.0063	≤ 0.7363	1.110

Our review of the existing literature and that of the internet resources showed that a commensal airbborne bacterium *Neisseria lactamica* [32], airborn bacteria *Arthrobacter* species S9 [33], *Paramecium bursaria Chlorella* virus

IL3A (PBCV-IL3A) [34], and airborne human pathogen *Haemophilus influenzae* biogroup aegyptius [35] were the genetic sources for the DNA restriction endonucleases NlaIV, AsuI, CviJI, and HaeIII respectively.

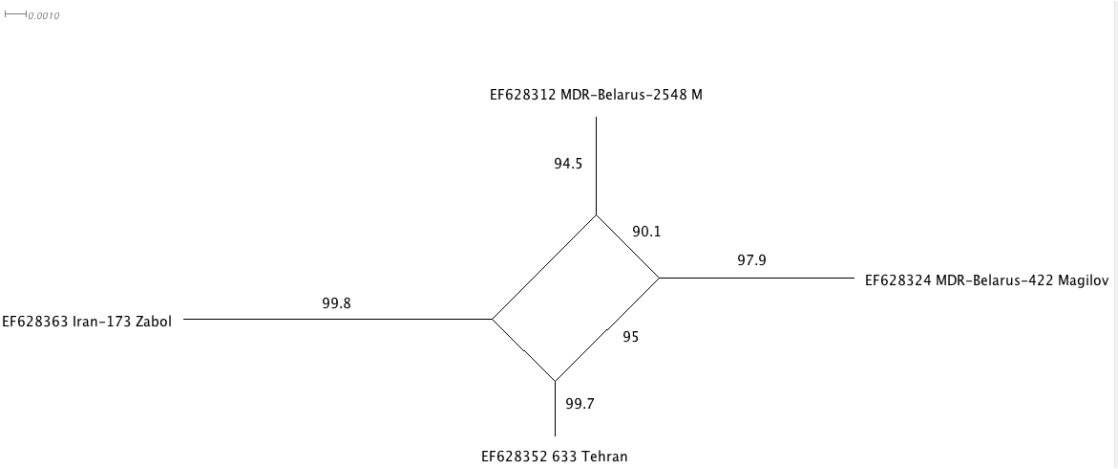
Tests for selection and genetic recombination. The average GC content was 65.94% for the DNA sequences of *rpoB* of *M. tuberculosis* strains that carried the sSNPs. The d_N/d_S ratios, for the *rpoB* loci, resulted in the values being ≤0.7363, while I_A = 1.110 (Table 3). Multiple parallelograms with the different fit and bootstrap estimates could be constructed in the *rpoB* DNA sequence analyses using the program SplitsTree4 (v. 4.12.8) [31]. Fig. 1A-D exhibits the parallelograms demonstrating evidence for the HGT of *rpoB* loci of *M. tuberculosis* strains. These parallelograms were strongly supported by the fit and bootstrap estimates varying from the high values to the highest value (being 95-100 and 85-100 respectively).

A



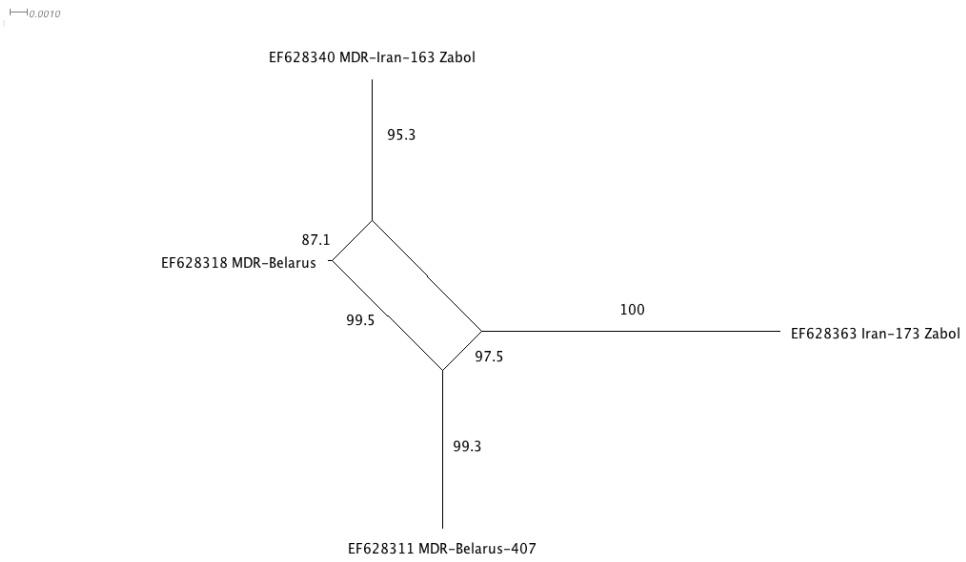
Fit 95.338

B



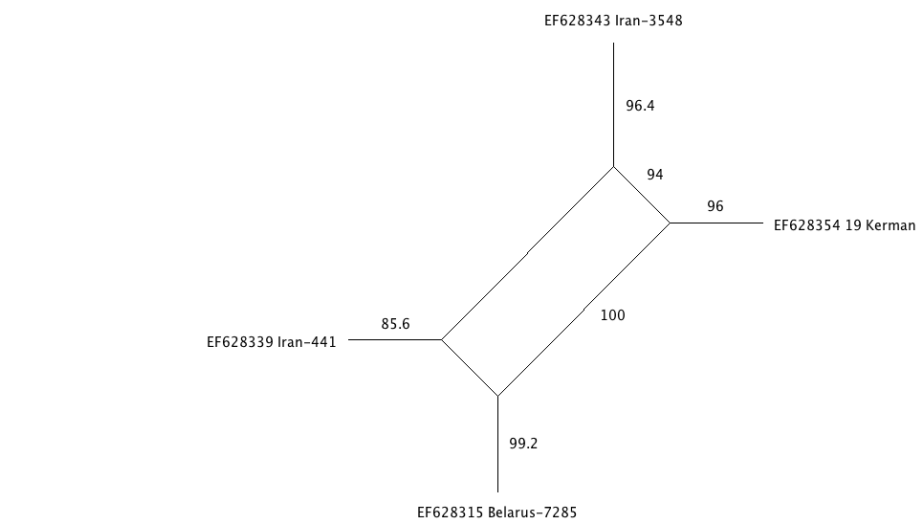
Fit 100

C



Fit 100

D



Fit 100

Fig. 1A-D. The parrarelograms (A-D) constructed by the SplitsTree analyses, exhibiting the genetic recombination in the *rpoB* gene of the *M. tuberculosis* isolates.

4. DISCUSSION

The silent DNA polymorphisms have been used as the phylogenetic markers to assess the evolution of *M. tuberculosis* [23, 36]. In this study, we examined the *rpoB* loci of the different *M. tuberculosis* carrying the sSNPs to determine whether this genetic locus is involved in the HGT.

In a genome of *M. tuberculosis*, the sSNPs appear across multiple loci including *rpoB* [37], the gene that strictly undergoes the selective pressures of the antimicrobial agent rifampin. It is important to indicate that, unlike the sSNPs reflecting neutral genomic variations, specific missense mutations of the *rpoB* RRDR were found to be a major driving force of this species resistance to RIF [14-18, 37]. In our study, the d_N/d_S ratios (≤ 0.7363), for the *M. tuberculosis rpoB* loci that carried both the drug resistance mutations and the multiple sSNPs, showed no significant selective pressures on the amino acid changes. It is noteworthy that these values were in a range of the d_N/d_S ratios (01.109-0.848) determined earlier for the antibiotic resistance-associated loci of the sensitive strains of *M. tuberculosis* [23]. Hence, it can be thought that the *rpoB* loci of these strains shared some similar evolutionary trends.

It was previously reported that oxidative damage of the DNA, in the *rpoB* gene of *M. tuberculosis*, can lead to C⟶T base substitutions [37], while the circumstances under which this process takes place remain unknown. Here we show that C⟶T base changes constituted 12.0% of the sSNPs in *rpoB* of the *M. tuberculosis* strains (Table 2), suggesting that they could occur during the HGT of the *rpoB* gene loci of the *M. tuberculosis* strains. In particular, in our dataset, the C⟶T base substitutions were detected at the *rpoB* codon GGC_{507}, which, along with the other sSNPs-affected *rpoB* codons GGG_{534} and $GGG_{566,}$ was adjacent to the DNA cleavage site of NlaIV produced by *N. lactamica*. This organism is a commensal airbborne bacterium that occupies the human nasopharynx being one of the habitats of *M. tuberculosis* as well during the TB infection. It is very likely that *N. lactamica* excretes NlaIV in abundance and releases it extracellularly in a surrounding environment, where the enzyme could interact with the naked genomic DNAs of the degrading cells of the *M. tuberculosis* strains when these two species share the same living compartment. Such an enzymatic interaction of NlaIV with the naked genomic DNAs would possibly induce the DNA excision of the *rpoB* loci from a genome of the donor strain of *M. tuberculosis*. Subsequently, the excised DNA fragments could be acquired by a recipient strain of this species during the mixed TB infections. We suggest that the DNA insertion of the *rpoB* loci into the genome of the recipient strain via the HGT could lead sometimes to the DNA errors in a capacity of the sSNPs in the adjacent sites of the targeted genomic region. The statistical analysis of the *rpoB* sSNPs (Table 3) using the binomial probability distribution showed that there was very slim probability ($P \leq 0.0485$) for these polymorphisms to emerge in parallel by chance at the adjacent regions of the DNA cleavage sites of the same DNA restriction endonucleases in this gene. Very interestingly, GGC_{507} harboring the sSNPs is the first codon of the *rpoB* RRDR, the region that from its 3 prime end is flanked by the codon GGG_{534} that also carried the sSNPs. Therefore, it is possible that enzymatic activities of NlaIV of *N. lactamica* could selectively affect the *rpoB* RRDR, resulting in the flanking sSNPs upon the insertion of this region into the *M. tuberculosis* genome. The similar suggestions can be made in respect to the *rpoB* codons GGG_{534} and GGG_{566} encompassing the sSNPs. In the wild type allele of *rpoB*, the adjacent regions of these codons from their 3 prime ends represent a potential target of the DNA restriction endonuclease AsuI produced by the airborn bacteria *Arthrobacter* species S9. Besides, from the 3 prime end, the above two codons and the codon GGC_{556}, in this gene, are adjacent to the DNA cleavage sites of HaeIII from the airborn human pathogen *H. influenzae* biogroup aegyptius. Thus, based on these observations, at least NlaIV, AsuI, and HaeIII are thought to play a role in the DNA excision of the *rpoB* loci of *M. tuberculosis*, later leading to the sSNPs through the HGT. It can be also thought that the sSNPS remained fixed in some populations of *M. tuberculosis*.

It is important to note that clonal populations, in general, are identified by the I_A value that differs significantly from zero [29]. In our analysis, the I_A estimate 1.110 showed that *rpoB* of *M. tuberculosis* is not clonal. While the average GC content of the *rpoB* gene (65.94%) was only slightly higher as compared with the average GC content of the whole genome of *M. tuberculosis* being 65.60% [38], evidence, for the presence of HGT in *rpoB*, was supported further by the results of the Splitstree analyses. The method of the splits decomposition utilized by the program SplitsTree [30, 31] has been an extremely powerful tool for detecting the parallel nucleotide substitutions of the microbial genomes involved in the genetic recombination [39, 40-44]. In the SplitsTree analysis, the highest fit being 100 for the parallelograms (Fig. 1B-D), and very strong bootstrap values being 90-100, for the nodes of the parallelograms, provide a clear evidence for the HGT events in the *rpoB* gene of *M. tuberculosis*.

The relative importance of bacterial determinants of TB treatment failure has been unclear [7]. Taken together, our results show that the parallel sSNPs of *rpoB* were limited to the subsets of *M. tuberculosis* strains (Table 1) from two distinct geographic regions, Iran and the Republic of Belarus. These results can be in agreement with the observation that the *rpoB* allelic diversity and DNA polymorphisms reflect sometimes the geographical distribution of *M. tuberculosis* strains [16, 45]. The experimental and clinical evidence demonstrates that *M. tuberculosis* strains from different lineages can vary in their capacity to cause disease [46-48] and acquire drug resistance [11-12, 46, 49, 50]. Besides, it was found that strains from the East Asian lineage and the Beijing sublineage acquire drug resistance in vitro more rapidly than *M. tuberculosis* strains from the Euro-American lineage [7]. Here, we suggest that certain strains of *M. tuberculosis* from the Eastern Europe and the Asia have unique abilities to recombine *rpoB* loci, and which can easily result in the development of resistance or in increased resistance to RIF. The mixed TB infections are essential for the HGT of *rpoB* of *M. tuberculosis* to occur, although, it is noteworthy that the infections [51-54] driven by more than one *M. tuberculosis* strain during the same episode are relatively infrequent. Nevertheless, the mixed infections, in a presence of the air pollution with the above indicated airborne microbes (e.g., *N. lactamica*, *H. influenzae* and *Arthrobacter* species), seem to be a risk factor at least in the development of RIF resistant TB. Thus, while the HGT of *rpoB* can be a rare event, this phenomenon is strongly suggested to be an additional genetic mechanism that contributes to *M. tuberculosis* resistance to RIF. The results of this study open a broad avenue for the further research on the *M. tuberculosis* interactions with other airborne microbes to better understand the causes of the reemerging drug resistant and severe TB cases.

5. ACKNOWLEDGMENTS

We thank Dr. Cort Anderson from the Faculty of Natural Sciences and Engineering, Ilia State University, for his useful comments that helped us in a progression of this study. This work was supported by the intramural fund of the Ilia State University, Tbilisi, Georgia.

6. REFERENCES

[1] Herrera L, Jiménez S, Valverde A, García-Aranda MA, Sáez-Nieto JA. 2003. Molecular analysis of rifampicin-resistant *Mycobacterium tuberculosis* isolated in Spain (1996-2001). Description of new mutations in the *rpoB* gene and review of the literature. Int J Antimicrob Agents 21(5): 403-8.

[2] Raviglione MC, Snider DE Jr, Kochi A. 1995. Global epidemiology of tuberculosis. Morbidity and mortality of a worldwide epidemic. JAMA 273(3):220-6.

[3] Velayati AA, Masjedi MR, Farnia P, Tabarsi P, Ghanavi J, Ziazarifi AH, Hoffner SE. 2009. Emergence of new forms of totally drug-resistant tuberculosis bacilli: super extensively drug-resistant tuberculosis or totally drug-resistant strains in iran. Chest 136(2):420-5.

[4] Gandhi NR, Nunn P, Dheda K, Schaaf HS, Zignol M, van Soolingen D, Jensen P, Bayona J. 2010. Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. Lancet 375(9728):1830-43.

[5] David HL. 1970. Probability distribution of drug-resistant mutants in unselected populations of *Mycobacterium tuberculosis*. Applied microbiology 20:810–814.

[6] Ford CB, Lin PL, Chase MR, Shah RR, Iartchouk O, Galagan J, Mohaideen N, Ioerger TR, Sacchettini JC, Lipsitch M, Flynn JL, Fortune SM. 2011. Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. Nat Genet 43(5):482-6.

[7] Ford CB, Shah RR, Maeda MK, Gagneux S, Murray MB, Cohen T, Johnston JC, Gardy J, Lipsitch M, Fortune SM. 2013. *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. Nat Genet 45(7):784-90.

[8] Migliori GB, De Iaco G, Besozzi G, Centis R, Cirillo DM. 2007. First tuberculosis cases in Italy resistant to all tested drugs. Euro Surveill 12:E070517.1.

[9] Pablos-Mendez A, Knirsch CA, Barr RG, Lerner BH, Frieden TR. 1997. Nonadherence in tuberculosis treatment: predictors and consequences in New York City. Am J Med 102:164–170.

[10] Udwadia ZF, Pinto LM, Uplekar MW. 2010. Tuberculosis management by private practitioners in Mumbai, India: has anything changed in two decades? PLoS ONE 5:e12023.

[11] Johnson R, Warren RM, van der Spuy GD, Gey van Pittius NC, Theron D, Streicher EM, Bosman M, Coetzee GJ, van Helden PD, Victor TC. 2010. Drug-resistant tuberculosis epidemic in the Western Cape driven by a virulent Beijing genotype strain. Int J Tuberc Lung Dis 14(1):119-21.

[12] Streicher EM, Warren RM, Kewley C, Simpson J, Rastogi N, Sola C, van der Spuy GD, van Helden PD, Victor TC. 2004. Genotypic and phenotypic characterization of drug-resistant *Mycobacterium tuberculosis* isolates from rural districts of the Western Cape Province of South Africa. J Clin Microbiol 42(2):891-4.

[13] Heep M, Brandstätter B, Rieger U, Lehn N, Richter E, Rüsch-Gerdes S, Niemann S. 2001. Frequency of *rpoB* mutations inside and outside the cluster I region in rifampin-resistant clinical *Mycobacterium tuberculosis* isolates. J Clin Microbiol 39(1):107-10.

[14] Ramaswamy S, Musser JM. 1998. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. Tuber Lung Dis 79(1):3-29. Review.

[15] Telenti, A, Imboden P, Marchesi F, Lowrie D, Cole S, Colston MJ, Matter L, Schopfer K, and Bodmer T. 1993. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. Lancet 341:647–650.

[16] Musser JM. 1995. Antimicrobial agent resistance in mycobacteria: molecular genetic insights. Clin Microbiol Rev 8(4):496-514. Review.

[17] Miller LP, Crawford JT, Shinnick TM. 1994. The *rpoB* gene of *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 38(4):805-11.

[18] Heep M, Rieger U, Beck D, Lehn N. 2000. Mutations in the beginning of the *rpoB* gene can induce resistance to rifamycins in both Helicobacter pylori and *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 44(4):1075-7.

[19] Schmalstieg AM, Srivastava S, Belkaya S, Deshpande D, Meek C, Leff R, van Oers NS, Gumbo T. 2012. The antibiotic resistance arrow of time: efflux pump induction is a general first step in the evolution of mycobacterial drug resistance. Antimicrob Agents Chemother 56(9):4806-15.

[20] Louw GE, Warren RM, Gey van Pittius NC, Leon R, Jimenez A, Hernandez-Pando R, McEvoy CR, Grobbelaar M, Murray M, van Helden PD, Victor TC. 2011. Rifampicin reduces susceptibility to ofloxacin in rifampicin-resistant *Mycobacterium tuberculosis* through efflux. Am J Respir Crit Care Med. 184(2):269-76.

[21] Adams KN, Takaki K, Connolly LE, Wiedenhoft H, Winglee K, Humbert O, Edelstein PH, Cosma CL, Ramakrishnan L. 2011. Drug tolerance in replicating mycobacteria mediated by a macrophage-induced efflux mechanism. Cell 145(1):39-53.

[22] Eldholm V, Norheim G, von der Lippe B, Kinander W, Dahle UR, Caugant DA, Mannsåker T, Mengshoel AT, Dyrhol-Riise AM, Balloux F. 2014. Evolution of extensively drug-resistant *Mycobacterium tuberculosis* from a susceptible ancestor in a single patient. Genome Biol 15(11):490.

[23] Baker L, Brown T, Maiden MC, Drobniewski F. 2004. Silent nucleotide polymorphisms and a phylogeny for

Mycobacterium tuberculosis. Emerg Infect Dis. 10(9):1568-77.

[24] Smith NH, Dale J, Inwald J, Palmer S, Gordon SV, Hewinson RG, Smith JM. 2003. The population structure of *Mycobacterium bovis* in Great Britain: clonal expansion. Proc Natl Acad Sci U S A 100(25):15271-5.

[25] Krzywinska E, Krzywinski J, Schorey JS. 2004. Naturally occurring horizontal gene transfer and homologous recombination in *Mycobacterium*. Microbiology 150(Pt 6):1707-12.

[26] Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421.

[27] Jeanmougin F, Thompson J. D, Gouy M, Higgins DG, and Gibson TJ. 1998. Multiple sequence alignment with Clustal X. Trends Biochem Sci 23:403–405.

[28] Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3(5):418-26.

[29] Smith JM, Smith NH, O'Rourke M, Spratt BG. 1993. How clonal are bacteria? Proc Natl Acad Sci U S A 15;90(10):4384-8.

[30] Bandelt HJ, Dress AW. 1992. Split decomposition: a new and useful approach to phylogenetic analysis of distance data. Mol Phylogenet Evol 1(3):242-52.

[31] Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23(2):254-67.

[32] Lau PC, Forghani F, Labbé D, Bergeron H, Brousseau R, Höltke HJ. 1994. The NlaIV restriction and modification genes of *Neisseria lactamica* are flanked by leucine biosynthesis genes. Mol Gen Genet 243(1):24-31.

[33] Nagano Y, Walker J, Loughrey A, Millar C, Goldsmith C, Rooney P, Elborn S, Moore J. 2009. Identification of airborne bacterial and fungal species in the clinical microbiology laboratory of a university teaching hospital employing ribosomal DNA (rDNA) PCR and gene sequencing techniques. Int J Environ Health Res 19(3):187-99.

[34] Blanc G, Mozar M, Agarkova IV, Gurnon JR, Yanai-Balser G, Rowe JM, Xia Y, Riethoven JJ, Dunigan DD, Van Etten JL. 2014. Deep RNA sequencing reveals hidden features and dynamics of early gene transcription in *Paramecium bursaria chlorella* virus 1. PLoS One 7;9(3).

[35] Strouts FR, Power P, Croucher NJ, Corton N, van Tonder A, Quail MA, Langford PR, Hudson MJ, Parkhill J, Kroll JS, Bentley SD. 2012. Lineage-specific virulence determinants of *Haemophilus influenzae* biogroup aegyptius. Emerg Infect Dis 18(3):449-57.

[36] Gibson A, Brown T, Baker L, Drobniewski F. 2005. Can 15-locus mycobacterial interspersed repetitive unit-variable-number tandem repeat analysis provide insight into the evolution of *Mycobacterium tuberculosis*? Appl Environ Microbiol 71(12):8207-13.

[37] O'Sullivan DM, McHugh TD, Gillespie SH. 2005. Analysis of rpoB and pncA mutations in the published literature: an insight into the role of oxidative stress in *Mycobacterium tuberculosis* evolution? J Antimicrob Chemother 55(5):674-9.

[38] Miyoshi-Akiyama T, Satou K, Kato M, Shiroma A, Matsumura K, Tamotsu H, Iwai H, Teruya K, Funatogawa K, Hirano T, Kirikae T. 2015. Complete annotated genome sequence of *Mycobacterium tuberculosis* (Zopf) Lehmann and Neumann (ATCC35812) (Kurono). Tuberculosis (Edinb). 95(1):37-9. doi: 10.1016/j.tube.2014.10.007.

[39] Norberg P, Kasubi MJ, Haarr L, Bergström T, Liljeqvist JA. 2007. Divergence and recombination of clinical herpes simplex virus type 2 isolates. J Virol 81(23):13158-67.

[40] Kotetishvili M, Stine OC, Kreger A, Morris JG Jr, Sulakvelidze A. 2002. Multilocus sequence typing for characterization of clinical and environmental *salmonella* strains. J Clin Microbiol 40(5):1626-35.

[41] Kotetishvili M, Stine OC, Chen Y, Kreger A, Sulakvelidze A, Sozhamannan S, Morris JG Jr. 2003. Multilocus sequence typing has better discriminatory ability for typing *Vibrio cholerae* than does pulsed-field gel electrophoresis and provides a measure of phylogenetic relatedness. J Clin Microbiol 41(5):2191-6.

[42] Revazishvili T, Kotetishvili M, Stine OC, Kreger AS, Morris JG Jr, Sulakvelidze A. 2004. Comparative analysis of multilocus sequence typing and pulsed-field gel electrophoresis for characterizing *Listeria monocytogenes* strains isolated from environmental and clinical sources. J Clin Microbiol 42(1):276-85.

[43] Kotetishvili M, Kreger A, Wauters G, Morris JG Jr, Sulakvelidze A, Stine OC. 2005. Multilocus sequence typing for studying genetic relationships among *Yersinia* species. J Clin Microbiol 43(6):2674-84.

[44] Stalder U, Stephan R, Corti S, Bludau M, Maeschli A, Klocke P, Johler S. 2014. Short communication: *Staphylococcus aureus* isolated from colostrum of dairy heifers represent a closely related group exhibiting highly homogeneous genomic and antimicrobial resistance features. J Dairy Sci 97(8):4997-5000.

[45] Qian L, Abe C, Lin TP, Yu MC, Cho SN, Wang S, Douglas JT. 2002. *rpoB* genotypes of *Mycobacterium tuberculosis* Beijing family isolates from East Asian countries. J Clin Microbiol 40(3):1091-4.

[46] Glynn JR, Whiteley J, Bifani PJ, Kremer K, Van Soolingen D. 2002. Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. Emerging Infect Dis 8:843–849.

[47] Drobniewski F, Balabanova Y, Nikolayevsky V, Ruddy M, Kuznetzov S, Zakharova S, Melentyev A, Fedorin I.

2005. Drug-resistant tuberculosis, clinical virulence, and the dominance of the Beijing strain family in Russia. JAMA 293(22):2726-31.

[48] Coscolla M, Gagneux S. Does. 2010. *M. tuberculosis* genomic diversity explain disease diversity? Drug Discovery Today: Disease Mechanisms 7:e43–e59.

[49] European Concerted Action on New Generation Genetic Markers and Techniques for the Epidemiology and Control of Tuberculosis. 2006. Beijing/W genotype Mycobacterium tuberculosis and drug resistance. Emerging Infect Dis 12:736–743.

[50] Pang Y, Zhou Y, Zhao B, Liu G, Jiang G, Xia H, Song Y, Shang Y, Wang S, Zhao YL. 2012. Spoligotyping and drug resistance analysis of *Mycobacterium tuberculosis* strains from national survey in China. PLoS One 7(3):e32976.

[51] Braden, C. R., G. P. Morlock, C. L. Woodley, K. R. Johnson, A. C. Colombel, M. D. Cave, Z. Yang, S. E. Valway, I. M. Onorato, and J. T. Crawford 2001. Simultaneous infection with multiple strains of *Mycobacterium tuberculosis*. Clin. Infect. Dis 33:e42–e47.

[52] Chaves, F, Dronda F, Alonso-Sanz M, and Noriega AR. 1999. Evidence of exogenous reinfection and mixed infection with more than one strain of *Mycobacterium tuberculosis* among Spanish HIV-infected inmates. AIDS 13:615–620.

[53] Warren RM, Victor TC, Streicher EM, Richardson M, Beyers N, van Pittius NC, and van Helden PD. 2004. Patients with active tuberculosis often had different strains in the same sputum specimen. Am. J. Respir. Crit. Care Med 169:610–614.

[54] Yeh RW, Hopewell PC, and Daley CL. 1999. Simultaneous infection with two strains of *Mycobacterium tuberculosis* identified by restriction fragment length polymorphism analysis. Int. J. Tuberc. Lung Dis 3:537–539.

[55] Titov LP, Zakerbostanabad S, Slizen V, Surkova L, Taghikhani M, Bahrmand A. 2006. Molecular characterization of *rpoB* gene mutations in rifampicine-resistant *Mycobacterium tuberculosis* isolates from tuberculosis patients in Belarus. Biotechnol J 12:1447-52.