UNIVERSIDAD PERUANA CAYETANO HEREDIA
FACULTAD DE CIENCIAS Y FILOSOFÍA ALBERTO CAZORLA TALLERI

Determination of potentially novel compensatory mutations in rpoC associated with rifampin resistance and rpoB mutations in *Mycobacterium tuberculosis* clinical isolates from Peru

Ana Paula Vargas Ruiz

Tesis para optar el Título Profesional de Licenciada en Biología

LIMA - PERÚ

2020

JURADO EXAMINADOR

Presidente:	Carla M. Gallo López-Aliaga, MSc.
Vocal:	Jonathan Gómez-Raja, PhD.
Secretario:	Daniel Clark Leza, PhD.

Fecha de sustentación:	06 de agosto de 2020
Calificación:	Aprobado con excelencia

ASESOR

Mirko Zimic Peralta, PhD.

DEDICATORIA

El presente trabajo va dedicado a mis padres, por su apoyo incondicional durante mi formación académica, su guía, consejos, permanente interés, y por soportar mis discusiones en los almuerzos;

a mis abuelitas y nuestros vecinos, por enseñarme a valorar la historia sin perder el norte, a navegar de buena fe por una vida de altibajos, y a darlo todo de mí sin olvidarme de pasar un buen rato;

a mi madrina, por su constante e incondicional apoyo y emoción para cada uno de mis logros, por darme el mejor ejemplo de lo que es un verdadero maestro e inspirarme a convertirme en uno.

AGRADECIMIENTOS

A Dios, por estar conmigo y mi familia y bendecirnos a diario;

a mis padres, familia y amigos, por su apoyo incondicional;

a mi asesor, el Dr. Mirko, por aceptarme en su laboratorio desde muy temprano en mi carrera, enseñarme siempre desde una perspectiva muy humana, e inspirarme y guiarme por el mundo de la ciencia;

al Laboratorio de Bioinformática y Biología Molecular, una comunidad increíble de científicos, siempre prestos a servir a los demás;

a Alexandra E., por permitirnos hacer ciencia en Perú;

y finalmente a Angela Rios, compañera de clase y laboratorio, quien fue un pilar fundamental para este trabajo.

DECLARACIÓN DE CONFLICTO DE INTERESES

Los autores declaran no tener ningún conflicto de interés.

ÍNDICE

ABREVIATURAS

1. RIF: Rifampicina.
2. RRDR: Región Determinante de Resistencia a Rifampicina.
3. RNApol: RNA polimerasa de *M. tuberculosis*.
4. TB: Tuberculosis.
5. MDR-TB: Tuberculosis multidrogoresistente.

RESUMEN

Antecedentes: La resistencia a rifampicina (RIF) en *Mycobacterium tuberculosis* es frecuentemente causada por mutaciones en el gen rpoB. Estas mutaciones están asociadas a un costo de *fitness*, que puede superarse mediante mutaciones compensatorias en otros genes, entre cuales rpoC puede ser el más importante.

Objetivos: Analizar 469 aislados clínicos peruanos de *M. tuberculosis* para identificar mutaciones compensatorias en rpoC/rpoA asociadas a la resistencia RIF.

Materiales y métodos: Los aislados de *M. tuberculosis* fueron recolectados y probados para susceptibilidad a RIF y *spoligotyping*. Las muestras se secuenciaron y alinearon al genoma de referencia para identificar mutaciones puntuales. Analizando las secuencias y los metadatos, se identificó una lista de mutaciones en rpoC exclusivamente asociadas a la resistencia a RIF y a las mutaciones en rpoB. A continuación, evaluamos la distribución de estas mutaciones a lo largo de la secuencia de la proteína y la estructura tridimensional.

Resultados: Ciento veinticinco cepas fueron susceptibles a RIF y 346 fueron resistentes. Se identificó 35 nuevas mutaciones potencialmente compensatorias, algunas de las cuales se encontraban distribuidas en la superficie de la interfaz rpoB-rpoC, surgiendo en grupos y sugiriendo la presencia de *hotspots* para las mutaciones compensatorias.

Conclusiones: En este estudio se identifican 35 posibles nuevas mutaciones compensatorias en la subunidad β' del ARNpol de *M. tuberculosis*. Se considera que seis de ellas (S428T, L507V, A734V, I997V, y V1252LM) son las que tienen más probabilidades de tener un papel compensatorio, ya que se encuentran en la zona de interacción de las dos subunidades y la mutación no produjo ningún cambio en las propiedades fisicoquímicas de la proteína.

Palabras clave: tuberculosis, resistencia a rifampicina, mutaciones compensatorias, rpoB, rpoC, evolución.

ABSTRACT

Background: Rifampicin (RIF) resistance in *Mycobacterium tuberculosis* is frequently caused by mutations in the rpoB gene. These mutations are associated with a fitness cost, which can be overcome by compensatory mutations in other genes, among which rpoC may be the most important.

Aims and objectives: We analyzed 469 Peruvian *M. tuberculosis* clinical isolates to identify compensatory mutations in rpoC/rpoA associated with RIF resistance.

Materials and methods: The *M. tuberculosis* isolates were collected and tested for RIF susceptibility and spoligotyping. Samples were sequenced and aligned to the reference genome to identify mutations. By analyzing the sequences and the metadata, we identified a list of rpoC mutations exclusively associated with RIF resistance and mutations in rpoB. We then evaluated the distribution of these mutations along the protein sequence and tridimensional structure.

Results: One hundred and twenty‑five strains were RIF susceptible and 346 were resistant. We identified 35 potential new compensatory mutations, some of which were distributed on the interface surface between rpoB and rpoC, arising in clusters and suggesting the presence of hotspots for compensatory mutations.

Conclusions: This study identifies 35 putative novel compensatory mutations in the β' subunit of *M. tuberculosis* RNApol. Six of these (S428T, L507V, A734V, I997V, and V1252LM) are considered most likely to have a compensatory role, as they fall in the interaction zone of the two subunits and the mutation did not lead to any change in the protein's physical–chemical properties.

Keywords: tuberculosis, rifampicin resistance, compensatory mutations, rpoB, rpoC, evolution.

Original Article

Determination of Potentially Novel Compensatory Mutations in rpoC Associated with Rifampin Resistance and rpoB Mutations in *Mycobacterium tuberculosis* Clinical Isolates from Peru

Ana Paula Vargas[1], Angela A. Rios[1], Louis Grandjean[2], Daniela E. Kirwan[3], Robert H. Gilman[4], Patricia Sheen[1], Mirko J. Zimic[1]

[1]Department of Cellular and Molecular Sciences, Laboratory of Bioinformatics and Molecular Biology, Faculty of Science and Philosophy, Universidad Peruana Cayetano Heredia, Lima, Peru, [2]Department of Paediatric Infectious Diseases, Imperial College, [3]Infection and Immunity Research Institute, St. George's, University of London, London, England, [4]Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

Abstract

Background: Rifampicin (RIF) resistance in *Mycobacterium tuberculosis* is frequently caused by mutations in the rpoB gene. These mutations are associated with a fitness cost, which can be overcome by compensatory mutations in other genes, among which rpoC may be the most important. We analyzed 469 Peruvian *M. tuberculosis* clinical isolates to identify compensatory mutations in rpoC/rpoA associated with RIF resistance. **Methods:** The *M. tuberculosis* isolates were collected and tested for RIF susceptibility and spoligotyping. Samples were sequenced and aligned to the reference genome to identify mutations. By analyzing the sequences and the metadata, we identified a list of rpoC mutations exclusively associated with RIF resistance and mutations in rpoB. We then evaluated the distribution of these mutations along the protein sequence and tridimensional structure. **Results:** One hundred and twenty-five strains were RIF susceptible and 346 were resistant. We identified 35 potential new compensatory mutations, some of which were distributed on the interface surface between rpoB and rpoC, arising in clusters and suggesting the presence of hotspots for compensatory mutations. **Conclusion:** This study identifies 35 putative novel compensatory mutations in the β' subunit of *M. tuberculosis* RNApol. Six of these (S428T, L507V, A734V, I997V, and V1252LM) are considered most likely to have a compensatory role, as they fall in the interaction zone of the two subunits and the mutation did not lead to any change in the protein's physical–chemical properties.

Keywords: Compensatory mutations, evolution, rifampicin resistance, rpoB, rpoC, tuberculosi

Submitted: 13-Feb-2020 **Revised:** 15-Feb-2020 **Accepted:** 04-Apr-2020 **Published:** 29-May-2020

Introduction

Tuberculosis (TB) is the leading cause of death by a single infectious agent worldwide.[1] With an average of 27,000 new cases per year, Peru has the second highest rate of TB in the Americas; thanks to a national TB control program,[2] the incidence of TB has decreased since 2000 by 1.6% per year.[3]

Rifampicin (RIF) is one of the most important drugs in the treatment of TB, and resistance to RIF threatens the success of TB control programs globally.[4] In 2017, the WHO reported more than 159,000 cases of RIF-resistant TB including 1,508 in Peru.[5] In the same year, about 390,000 new cases of multidrug-resistant TB (MDR-TB) were reported worldwide.[3] In the preceding decade, the average number of reported cases per year in Peru exceeded 1,100, with a trend toward an increase in the later years.[6]

Access this article online

Quick Response Code:

Website: www.ijmyco.org

DOI: 10.4103/ijmy.ijmy_27_20

RIF causes bacterial cell death by binding to the β-subunit of the bacteria's RNA polymerase (RNApol), blocking RNA transcription, and inhibiting gene expression.[7-10] Mutations in rpoB, the gene encoding the *Mycobacterium tuberculosis* RNApol β-subunit, have been shown to be the main cause of RIF-resistant microorganisms.[11] Moreover, the RIF resistance-determining

Address for correspondence: Prof. Mirko J. Zimic, Laboratory of Bioinformatics and Molecular Biology, Faculty of Science and Philosophy, Universidad Peruana Cayetano Heredia, Av. Honorio Delgado 430, San Martín de Porres, Lima, Peru.
E-mail: mirko.zimic@upch.pe

ORCID:
https://orcid.org/0000-0002-7203-8847

How to cite this article: Vargas AP, Rios AA, Grandjean L, Kirwan DE, Gilman RH, Sheen P, *et al*. Determination of potentially novel compensatory mutations in rpoC associated with rifampin resistance and rpoB mutations in *Mycobacterium tuberculosis* clinical isolates from Peru. Int J Mycobacteriol 2020;9:121-37.

region (RRDR) is an 81 bp region in rpoB within which 95% of all known RIF-resistant mutations are found.[11] Many studies have examined the structural variations of rpoB mutants and their association with drug binding.[12] *Escherichia coli*, distinct species of *Salmonella*, *Thermus,* and *Mycobacterium* have been the most studied models for RIF resistance to date.

Mutations in rpoB have been found to be associated with reduced fitness of the affected strain. Several studies have demonstrated that the fitness loss due to rpoB mutations can be compensated for by other mutations, which may consequently be positively selected. These mutations have been mainly identified in the rpoC and/or rpoA genes which are responsible for coding the β' and α-subunits of the RNApol complex, respectively, and which interact closely with rpoB.[8,10,13-15] Several compensatory[13,14,16-19] and putative compensatory[20-23] mutations in these two genes have been described, mostly in rpoC, and including the V483A and L516P mutations that have been extensively described in *M. tuberculosis*.[16,19,20,23]

In this study, we analyzed rpoA, rpoB, and rpoC gene sequences from 469 *M. tuberculosis* clinical strains, recovered from unrelated TB patients from Peru, to identify compensatory mutations in rpoC/rpoA genes associated with RIF resistance. Characterization of these mutations will improve our understanding of the mycobacterial compensatory mechanisms that allow them to overcome the fitness cost associated with rpoB mutations and thereby enable the survival and ongoing transmission of RIF-resistant strains. Furthermore, this study could provide insight into evolutionary processes, leading to RIF resistance in TB.

Methods

Selection of strains and rifampicin-susceptibility testing

We selected 469 well-characterized MTB strains, isolated from unrelated TB patients in hospitals from Callao and South Lima, Peru, between 2009 and 2013. These were collected and selected as part of two large studies that have been described previously:[24] one was a population-level study and the other a household follow-up study. The MODS assay was used to confirm TB infection and determine RIF susceptibility. MODS is a highly sensitive (98%) and specific (98.6%) tool that can both diagnose TB and detect drug resistance by determining the presence or absence of bacterial growth under a fixed drug concentration, and is approved for routine clinical use.[24-26]

The genotypes of these isolates have previously been studied using spoligotyping,[24] which characterizes mycobacterial strains according to their hybridization patterns.[24,27] Thus, in this study, we also evaluated the association between RIF resistance and the spoligotype.

Sequences and mutations

The rpoA, rpoB, and rpoC gene sequences, and mutations in these genes compared to the wild-type reference strain H37Rv (GeneBank: NC_000962) were obtained from the complete genome sequences attained previously.[24] Briefly, samples were sequenced on the Illumina MiSeq platform (150 bp paired-end reads) and aligned to the reference H37Rv genome using standard methodology.

Mutations in rpoA and rpoC associated with rifampicin-resistant strains

The 469 strains were classified as RIF susceptible or RIF resistant according to their MODS result. Each group was subclassified based on the presence or absence of at least one mutation in the rpoB gene. Each of these four groups was further subclassified based on the presence or absence one or more mutations in rpoA or rpoC.

To identify possible compensatory mutations in rpoC, we selected the RIF-susceptible strains with mutations in rpoC and a wild-type rpoB gene, regardless of the presence or absence of mutations in rpoA (set A). We also selected the RIF-resistant strains harboring mutations in rpoC and any mutation in rpoB, regardless of rpoA status (set B). After removing all rpoC mutations from set B that were also present in set A, we obtained a list of rpoC mutations exclusively associated with RIF resistance and mutations in rpoB. The same approach was applied to mutations in rpoA.

rpoA and rpoC compensatory mutations reported elsewhere

A literature search was performed to identify any mutations in rpoC or rpoA that have previously been associated with RIF resistance and rpoB mutations and have been shown experimentally to be compensating for a loss of fitness.[13,14,16-23] These were compiled into a list, based on which we were able to classify the mutations identified in this study as either previously reported or as potentially novel.

Distribution of potentially novel compensatory rpoC mutations along the protein sequence

The distribution of rpoC mutations was represented in a linear histogram, which indicated an equitable amino acid spread along the complete sequence of the protein. The β' subunit (coded by rpoC) sequence comprises 1,300 amino acids, and the histogram was built on blocks of 100 amino acids. Previously reported rpoC mutations and mutations identified in this study were mapped separately. Peruvian mutations that had already been reported elsewhere were marked. This analysis was not performed for mutations in rpoA because of insufficient mutations, either previously reported or identified in this study.

Distribution of potentially compensatory rpoC mutations on the three-dimensional structure of *Mycobacterium tuberculosis* RNA polymerase

To better understand the mechanism of action of the potentially compensatory rpoC mutations, these were mapped onto the recently published crystal structure of the *M. tuberculosis* RNApol transcription initiation complex[28] (PDB ID: 5UHD). The structures were visualized and analyzed using VMD.[29] First, all of the putative compensatory rpoC mutations were mapped onto the structure. Then, the mutations unique to this study were differentiated from those reported previously.

We visualized the distribution of the mutations in the β' subunit (coded by rpoC) that fell within the region of the surface that interacts with the β subunit (coded by rpoB). To identify the residues of physical importance in the β subunit, we identified those that arise within 5.0–6.0 Å of these rpoC mutations.

Results

Determination of mutations in rpoB, rpoA, and rpoC

Among the 469 genomes analyzed [Table 1], we found 125 (27%) to be RIF sensitive and 346 (74%) to be RIF resistant according to the MODS assay [Figure 1]. We schematically distributed the strains according to their sensitivity to RIF, mutations in rpoB, and mutations in rpoC/rpoA [Figure 1]. 237 (50.5%) strains harbored mutation (s) in rpoC and/or rpoA. 322 (93.1%) RIF-resistant strains harbored at least one mutation in rpoB.

175 (54%) strains that were RIF resistant and harbored mutations in rpoB had at least one mutation in rpoA or rpoC, suggesting that these may be compensatory mutations. 41 (41.8%) RIF-susceptible strains with no mutations in rpoB had at least one mutation in rpoC or rpoA. The presence of these 41 mutations may suggest the occurrence of random mutations, not necessarily associated with bacterial fitness.

Strikingly, 11 RIF-susceptible strains of different spoligotypes harbored rpoB mutations in the critical region (RRDR) [Table 1]. These mutations represent 3.4% of the 327 strains harboring a mutation in the RRDR. Five of these 11 strains also harbored the G594E mutation in rpoC [Table 2]. Notably, G594E appears only in the Haarlem genotype among these strains. This suggests that a very small percentage (3.4%) of mutations in the RRDR may not be associated with RIF resistance.

Rifampicin susceptibility and genotyping

The distribution of RIF-resistant strains (i.e., the RIF-resistant/RIF-susceptible strains ratio) was similar (1.4–4.6) within the different spoligotype clades [Figure 2]. This suggests that our investigation is not biased toward a specific genotype. LAM was the predominant genotype (54.4%), followed by Haarlem (17.3%), other small clades and T (10.03%), Beijing (7.3%), Caprae (0.64%), and finally Bovis and EAI2-Manilla (0.21%).

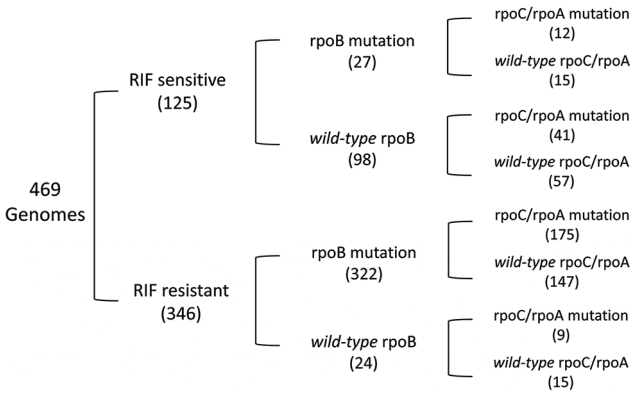

Figure 1: Representation of our 469-strain dataset [detailed in Table 1], classified according to resistance/susceptibility to rifampicin, presence/absence of mutations in rpoB, and presence/absence of mutations in rpoC or rpoA. The number in brackets indicates the number of strains with certain characteristics (i.e., there are 9 rifampicin-resistant strains that have a mutation in rpoC/rpoA but do not harbor mutations in rpoB)

Mutations in rpoA and rpoC associated with rifampicin resistant strains

The first group generated for analysis [set A, Supplementary Table 1] comprised a total of 43 strains, while the second group [set B, Supplementary Table 2] comprised 176 strains. Both groups had a variety of rpoC mutations, as depicted in Figure 3.

Interestingly, G594E appeared in both groups, being related to strains associated with RIF resistance and mutations in rpoB, but also to strains with a wild-type rpoB and RIF susceptibility. Other studies also found this mutation and excluded it from their analyses.[14,16,17,20,22,23]

Forty-four mutations found exclusively in set B [bold in Figure 3] were considered to be potentially compensatory. Nine (22%) of these mutations have been previously described [Supplementary Table 3]: N698K,[17] I491V,[17] H525Q,[17] P1040R,[17] L516P,[16,19] V1252M,[17] V483A,[16,19] V483G,[16,17,19] and V1252 L.[17] The remaining 35 (78%) were novel.

Only three mutations in the rpoA gene were found (T187A, G31S, and V183G), all of which have been reported elsewhere.[16,17,19] Our literature search identified a total of 29 compensatory mutations in rpoA.

Distribution of the potentially compensatory mutations along the rpoC protein sequence

A high frequency of mutations occurred between amino acids 400–500, 500–600, and 1000–1100 (a total of 21 mutations) [Figure 4], suggesting the existence of hotspots.

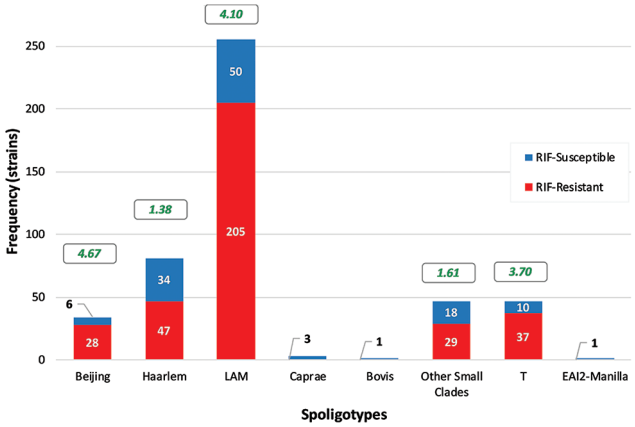

Figure 2: Frequency histogram of the 469 strains divided by spoligotype and sorted according to susceptibility to rifampicin. Red represents resistance while blue represents susceptibility. The white numbers represent the number of strains (frequency) that belong to certain spoligotypes and are susceptible/resistant to rifampicin. Black numbers have been added for contrast purposes but serve the same role. Green numbers in boxes represent the ratio of rifampicin resistant/rifampicin susceptible strains of each spoligotype

Table 1: General database

Strain number	rpoA	rpoC	rpoB	Spoligotyping	RRDR	RIF resistance
14722_6#1_H37Rv		V483A	S450L	T	1	1
14722_6#10_H37Rv				LAM	0	1
14722_6#11_H37Rv			H445D	Beijing	1	1
14722_6#12_H37Rv		P1040R	S450L	Beijing	1	1
14722_6#13_H37Rv		P1040A	S450L	LAM	1	1
14722_6#14_H37Rv			D435V	LAM	1	1
14722_6#15_H37Rv		G594E		Haarlem	0	0
14722_6#16_H37Rv		I491V	S450L, A692T	LAM	1	1
14722_6#17_H37Rv		G594E	L452P	Haarlem	1	0
14722_6#18_H37Rv				LAM	0	0
14722_6#19_H37Rv				LAM	0	0
14722_6#2_H37Rv		I491V	S450L, A692T	LAM	1	1
14722_6#20_H37Rv				Bovis	0	0
14722_6#21_H37Rv		G594E	E250G	Haarlem	0	0
14722_6#22_H37Rv			I965V	Caprae	0	0
14722_6#23_H37Rv			Q409R, S450L	LAM	1	1
14722_6#24_H37Rv			D435V	LAM	1	1
14722_6#25_H37Rv				LAM	0	1
14722_6#26_H37Rv				LAM	0	0
14722_6#27_H37Rv				Beijing	0	0
14722_6#28_H37Rv		P1040A	S450L	LAM	1	1
14722_6#29_H37Rv		W484S	S450L	Beijing	1	1
14722_6#3_H37Rv		G594E	S450L	Other small clades	1	1
14722_6#30_H37Rv		G594E		Other small clades	0	0
14722_6#31_H37Rv		G594E		Other small clades	0	0
14722_6#32_H37Rv				T	0	0
14722_6#33_H37Rv		R741S	H445C	LAM	1	1
14722_6#34_H37Rv		G594E	S450L, M920V	Haarlem	1	1
14722_6#35_H37Rv				LAM	0	0
14722_6#36_H37Rv			D435F	Beijing	1	0
14722_6#37_H37Rv		G594E	E250G	Haarlem	0	0
14722_6#38_H37Rv		P1040A	S450L	LAM	1	1
14722_6#39_H37Rv		G594E		Other small clades	0	0
14722_6#4_H37Rv			S450L, V695L	T	1	1
14722_6#40_H37Rv			D435V	LAM	1	1
14722_6#41_H37Rv			D435F	Beijing	1	0
14722_6#42_H37Rv		G594E		Haarlem	0	0
14722_6#43_H37Rv		G594E	H445Y	Other small clades	1	1
14722_6#44_H37Rv				Haarlem	0	1
14722_6#45_H37Rv		V517L	S450L	LAM	1	1
14722_6#46_H37Rv			S450L	Beijing	1	1
14722_6#47_H37Rv				LAM	0	1
14722_6#48_H37Rv			I965V	Caprae	0	0
14722_6#49_H37Rv				T	0	0
14722_6#5_H37Rv			S450L	LAM	1	1
14722_6#50_H37Rv			D435V	LAM	1	1
14722_6#51_H37Rv		V517L	S450L	LAM	1	1
14722_6#52_H37Rv			S428G, Q432E, S672F	LAM	1	1
14722_6#53_H37Rv			D435V	LAM	1	0
14722_6#54_H37Rv			D435V	LAM	1	1
14722_6#55_H37Rv		I491V, G594E	S450L	Other small clades	1	1
14722_6#56_H37Rv		G594E	D435V	Other small clades	1	1
14722_6#57_H37Rv		P1040A	S450L	LAM	1	1
14722_6#58_H37Rv		G594E		Other small clades	0	0

Contd...

Table 1: Contd...

Strain number	rpoA	rpoC	rpoB	Spoligotyping	RRDR	RIF resistance
14722_6#59_H37Rv		L527V	S450L, V695L	T	1	1
14722_6#6_H37Rv				LAM	0	0
14722_6#60_H37Rv			D435V	LAM	1	1
14722_6#61_H37Rv		G594E, E1113D	I491F	Other small clades	0	0
14722_6#62_H37Rv			S450L, V695L	T	1	1
14722_6#63_H37Rv		G594E	S450L	Haarlem	1	1
14722_6#64_H37Rv		G594E, E1113D	I491F	Other small clades	0	0
14722_6#65_H37Rv			H445P, V695L	T	1	1
14722_6#66_H37Rv				LAM	0	0
14722_6#67_H37Rv		T812I	S450L, A692T	LAM	1	1
14722_6#68_H37Rv				LAM	0	0
14722_6#69_H37Rv				LAM	0	0
14722_6#7_H37Rv			V695L	T	1	1
14722_6#70_H37Rv				Beijing	0	0
14722_6#71_H37Rv		G594E		Other small clades	0	0
14722_6#72_H37Rv		G594E	D435V	Haarlem	1	1
14722_6#73_H37Rv		G594E	S450F	Haarlem	1	1
14722_6#74_H37Rv			Q409R, S450L	LAM	1	1
14722_6#75_H37Rv			D435V	LAM	1	1
14722_6#76_H37Rv				LAM	0	0
14722_6#77_H37Rv			P45L, S450L	LAM	1	1
14722_6#78_H37Rv			S450L	LAM	1	1
14722_6#79_H37Rv		G594E		Haarlem	0	0
14722_6#8_H37Rv			D435V	LAM	0	0
14722_6#80_H37Rv			D435V	LAM	1	1
14722_6#81_H37Rv				Beijing	0	0
14722_6#82_H37Rv			D435V	LAM	1	1
14722_6#83_H37Rv				T	0	0
14722_6#84_H37Rv			V695L	T	0	0
14722_6#85_H37Rv			Q980R	LAM	1	1
14722_6#86_H37Rv		P1040A	S450L	LAM	1	1
14722_6#87_H37Rv		I491V	S450L, A692T	LAM	1	1
14722_6#88_H37Rv		I491V	S450L, A692T	LAM	1	1
14722_6#89_H37Rv		G594E	D435V	Haarlem	1	1
14722_6#9_H37Rv		L527V, G594E	S450L	Other small clades	1	1
14722_6#90_H37Rv			S450L	LAM	1	1
14722_6#91_H37Rv		G594E	S450L	Haarlem	1	1
14722_6#92_H37Rv				T	0	0
14722_6#93_H37Rv				LAM	0	0
14722_6#94_H37Rv		G594E	S450F	Haarlem	1	1
14722_6#95_H37Rv			D435V	LAM	1	1
14722_7#1_H37Rv			Q409R, S450L	LAM	1	1
14722_7#10_H37Rv		H525Q	S450L	Beijing	1	1
14722_7#11_H37Rv			D435V	LAM	1	1
14722_7#12_H37Rv		R741S	S450L, D574E	LAM	1	1
14722_7#13_H37Rv			S450L	LAM	1	1
14722_7#14_H37Rv			E207K, H445Y	LAM	1	1
14722_7#15_H37Rv			D435V	LAM	1	1
14722_7#16_H37Rv			D435V	LAM	1	1
14722_7#17_H37Rv		W484S	S450L	Beijing	1	1
14722_7#18_H37Rv			D435V	LAM	1	1
14722_7#19_H37Rv		V483A	S450L	LAM	1	1
14722_7#2_H37Rv			D435V	LAM	1	1
14722_7#20_H37Rv		L527V	S450L	T	1	1

Contd...

Table 1: Contd...

Strain number	rpoA	rpoC	rpoB	Spoligotyping	RRDR	RIF resistance
14722_7#21_H37Rv		G332R	S450L, V695L	T	1	1
14722_7#22_H37Rv		G594E	D435V	Other small clades	1	1
14722_7#23_H37Rv		E488G, G594E	S450L	Haarlem	1	1
14722_7#24_H37Rv		P1040R	S450L	LAM	1	1
14722_7#25_H37Rv			D435V	LAM	1	1
14722_7#26_H37Rv		G594E, E744G	S450L	Other small clades	1	1
14722_7#27_H37Rv				Beijing	0	1
14722_7#28_H37Rv			D435V	LAM	1	1
14722_7#29_H37Rv		G594E		Haarlem	0	0
14722_7#3_H37Rv			D435Y, N437H, V695L	T	1	1
14722_7#30_H37Rv		V517L	S450L	LAM	1	1
14722_7#31_H37Rv		G594E		Haarlem	0	0
14722_7#32_H37Rv			H445L	LAM	1	1
14722_7#33_H37Rv				LAM	0	0
14722_7#34_H37Rv		P1040Q	S450L, V695L	T	1	1
14722_7#35_H37Rv		G594E		Haarlem	0	0
14722_7#36_H37Rv		H445Y		LAM	1	1
14722_7#37_H37Rv		G594E	P45S, S450L	Haarlem	1	1
14722_7#38_H37Rv			S450L, E550G	LAM	1	1
14722_7#39_H37Rv	T187A		S450L	Beijing	1	1
14722_7#4_H37Rv				LAM	0	1
14722_7#40_H37Rv			S450W	T	1	1
14722_7#41_H37Rv			S450L	LAM	1	1
14722_7#42_H37Rv		G594E	A286V, S450L	Haarlem	1	1
14722_7#43_H37Rv				LAM	0	0
14722_7#44_H37Rv		V1252L	S450L, V695L	T	1	1
14722_7#45_H37Rv			D435V	LAM	1	1
14722_7#46_H37Rv		G594E		Haarlem	0	1
14722_7#47_H37Rv		G594E		Haarlem	0	0
14722_7#48_H37Rv			D435V	LAM	1	1
14722_7#49_H37Rv		G594E		Other small clades	0	0
14722_7#5_H37Rv				LAM	0	1
14722_7#50_H37Rv		V517L	S450L	LAM	1	1
14722_7#51_H37Rv		G594E		Haarlem	0	0
14722_7#52_H37Rv			S450L	LAM	1	1
14722_7#53_H37Rv		G594E, V1147A	S450L	Other small clades	1	1
14722_7#54_H37Rv				LAM	0	0
14722_7#55_H37Rv				LAM	0	0
14722_7#56_H37Rv				LAM	0	0
14722_7#57_H37Rv				LAM	0	0
14722_7#58_H37Rv	G31S	G594E	S450L	Other small clades	1	1
14722_7#59_H37Rv			D435V	LAM	1	1
14722_7#6_H37Rv			D435V	LAM	1	1
14722_7#60_H37Rv			D435V	LAM	1	1
14722_7#61_H37Rv		G594E		Other small clades	0	0
14722_7#62_H37Rv		Q523E	S450L	Beijing	1	1
14722_7#63_H37Rv		L516Q	S450L	LAM	1	1
14722_7#64_H37Rv				LAM	0	0
14722_7#65_H37Rv		T812I	S450L, A692T	LAM	1	1
14722_7#66_H37Rv	G31S	G594E	S450L	Other small clades	1	1
14722_7#67_H37Rv		R741S		LAM	0	1
14722_7#68_H37Rv				LAM	0	0
14722_7#69_H37Rv				LAM	0	0
14722_7#7_H37Rv		G571R, G594E	V170F	Other small clades	0	1

Contd...

Table 1: Contd...

Strain number	rpoA	rpoC	rpoB	Spoligotyping	RRDR	RIF resistance
14722_7#70_H37Rv		G594E		Other small clades	0	0
14722_7#71_H37Rv				LAM	0	0
14722_7#72_H37Rv			H445D, V695L	T	1	1
14722_7#73_H37Rv		G594E		Haarlem	0	0
14722_7#74_H37Rv		G594E	E250G, S450L, F971L	Haarlem	1	1
14722_7#75_H37Rv		G594E	L452P	Haarlem	1	0
14722_7#76_H37Rv		G594E, F831L	S450L	Other small clades	1	1
14722_7#77_H37Rv			S450L, A599V, V695L	T	0	1
14722_7#78_H37Rv		L1245R	S450L	Beijing	1	1
14722_7#79_H37Rv				LAM	0	0
14722_7#8_H37Rv			S450L, V695L	T	1	1
14722_7#80_H37Rv				LAM	0	0
14722_7#81_H37Rv		G594E	H445Y	Haarlem	1	1
14722_7#82_H37Rv		G332R	S450L, V695L	T	1	1
14722_7#83_H37Rv		G594E	S450L, M920V	Haarlem	1	1
14722_7#84_H37Rv			D435V	LAM	1	1
14722_7#85_H37Rv		G594E		Other small clades	0	0
14722_7#86_H37Rv			S450L, Q980R	LAM	1	1
14722_7#87_H37Rv		G332R	S450L, V695L	T	1	1
14722_7#88_H37Rv		G594E		Haarlem	0	0
14722_7#89_H37Rv				LAM	0	0
14722_7#9_H37Rv			S450L	Beijing	1	1
14722_7#90_H37Rv		G594E	S450L	Other small clades	1	1
14722_7#91_H37Rv				LAM	0	0
14722_7#92_H37Rv				LAM	0	0
14722_7#93_H37Rv		G594E		Haarlem	0	0
14722_7#94_H37Rv			S450L	LAM	1	1
14722_7#95_H37Rv		V483G, G594E	S450L	Other small clades	1	1
14722_8#1_H37Rv			T400N, S450L, V695L	T	1	1
14722_8#10_H37Rv			S450L, E550G	LAM	1	1
14722_8#11_H37Rv		N698K	S450L	LAM	1	1
14722_8#12_H37Rv			Q409R, S450L	LAM	1	1
14722_8#13_H37Rv	V183G		S450L	Beijing	1	1
14722_8#14_H37Rv		I885V	S450L	LAM	1	1
14722_8#15_H37Rv		A1047P	S450L	LAM	1	1
14722_8#16_H37Rv		L516P, G594E	S450L	Haarlem	1	1
14722_8#17_H37Rv		V1252M	S450L	LAM	1	1
14722_8#18_H37Rv		V1252M	S450L	LAM	1	1
14722_8#19_H37Rv		G594E	D120G, H445Y	Haarlem	1	1
14722_8#2_H37Rv		R741S	H445D	LAM	1	1
14722_8#20_H37Rv		P1040L	S450L	LAM	1	1
14722_8#21_H37Rv		V483G	S450L	LAM	1	1
14722_8#22_H37Rv		N698K	S450L	LAM	1	1
14722_8#23_H37Rv		P1040A	S450L	LAM	1	1
14722_8#24_H37Rv		N698S	S450L	LAM	1	1
14722_8#25_H37Rv			D435V	LAM	1	1
14722_8#26_H37Rv			S450L, E550G	LAM	1	1
14722_8#27_H37Rv		P1040A	S450L	LAM	1	1
14722_8#28_H37Rv			Q409R, S450L	LAM	1	1
14722_8#29_H37Rv			E207K, H445Y	LAM	1	1
14722_8#3_H37Rv		G945V	S450L	LAM	1	1
14722_8#30_H37Rv			D435V	LAM	1	1
14722_8#31_H37Rv		N698S	S450L	LAM	1	1
14722_8#32_H37Rv		T812I	S450L, A692T	LAM	1	1

Contd...

Table 1: Contd...

Strain number	rpoA	rpoC	rpoB	Spoligotyping	RRDR	RIF resistance
14722_8#33_H37Rv			S450F	LAM	1	1
14722_8#34_H37Rv			D435V	LAM	1	1
14722_8#35_H37Rv			S450L	LAM	1	1
14722_8#36_H37Rv		G594E	H445L	Haarlem	1	1
14722_8#37_H37Rv		G571R, G594E	S450L	Haarlem	1	1
14722_8#38_H37Rv		G594E, A734V	S450L	Haarlem	1	1
14722_8#39_H37Rv		P1040A	S450L	LAM	1	1
14722_8#4_H37Rv		T812I	A692T	LAM	1	1
14722_8#40_H37Rv			S450L, V695L	T	1	1
14722_8#41_H37Rv		V517L	S450L	LAM	1	1
14722_8#42_H37Rv			D435V	LAM	1	1
14722_8#43_H37Rv			S450L, E550G	LAM	1	1
14722_8#44_H37Rv		G594E	H445D	Haarlem	1	1
14722_8#45_H37Rv		G594E	S450L	Other small clades	1	1
14722_8#46_H37Rv			Q409R, S450L	LAM	1	1
14722_8#47_H37Rv			Q429H, D435Y	LAM	1	1
14722_8#48_H37Rv			S450L	LAM	1	1
14722_8#49_H37Rv				LAM	0	1
14722_8#5_H37Rv			D435V	LAM	1	1
14722_8#50_H37Rv		V1252M	S450L	LAM	1	1
14722_8#51_H37Rv				Other small clades	0	1
14722_8#52_H37Rv				LAM	0	1
14722_8#53_H37Rv		V517L	S450L	LAM	1	1
14722_8#54_H37Rv			S450L, L731P	LAM	1	1
14722_8#55_H37Rv		W484S	S450L	Beijing	1	1
14722_8#56_H37Rv		I491V	S450L, A692T	LAM	1	1
14722_8#57_H37Rv		V483G	S450L	LAM	1	1
14722_8#58_H37Rv		P1040R	S450L	LAM	1	1
14722_8#59_H37Rv		V1252M	S450L	LAM	1	1
14722_8#6_H37Rv		V1124E	S450L	LAM	1	1
14722_8#60_H37Rv			D435V	LAM	1	1
14722_8#61_H37Rv			Q409R, S450L	LAM	1	1
14722_8#62_H37Rv			D435V, V695L	T	1	1
14722_8#63_H37Rv			S450L,V695L	T	1	1
14722_8#64_H37Rv			D435V	Beijing	1	1
14722_8#65_H37Rv			D435V	LAM	1	1
14722_8#66_H37Rv		E49A	D435V	LAM	1	1
14722_8#67_H37Rv				LAM	0	1
14722_8#68_H37Rv			H445C	LAM	1	1
14722_8#69_H37Rv		N698S	S450L	LAM	1	1
14722_8#7_H37Rv			D435V	Beijing	1	1
14722_8#70_H37Rv		V1252L	S450L, A692T	LAM	1	1
14722_8#71_H37Rv		G594E		Haarlem	0	1
14722_8#72_H37Rv		G594E		Haarlem	0	1
14722_8#73_H37Rv			Q409R, S450L	LAM	1	1
14722_8#74_H37Rv			S450L, K891E	T	1	1
14722_8#75_H37Rv		V1252M	S450L	LAM	1	1
14722_8#76_H37Rv		G594E	D435V	Haarlem	1	1
14722_8#77_H37Rv		T812I	S450L, A692T	LAM	1	1
14722_8#78_H37Rv			D435V	LAM	1	1
14722_8#79_H37Rv		G594E	D435V	Haarlem	1	1
14722_8#8_H37Rv		N698S	S450L	LAM	1	1
14722_8#80_H37Rv		P1040A	S450L	LAM	1	1
14722_8#81_H37Rv		P1040A	S450L, V695L	T	1	1

Contd...

Table 1: Contd...

Strain number	rpoA	rpoC	rpoB	Spoligotyping	RRDR	RIF resistance
14722_8#82_H37Rv			Q409R, S450L	LAM	1	1
14722_8#83_H37Rv			D435V	LAM	1	1
14722_8#84_H37Rv			K312E, S450L	LAM	1	1
14722_8#85_H37Rv			Q409R, S450L	LAM	1	1
14722_8#86_H37Rv		V483G, G594E	S450L	Other small clades	1	1
14722_8#87_H37Rv		R741S	S450L	LAM	1	1
14722_8#88_H37Rv		G594E	S450F	Other small clades	1	1
14722_8#89_H37Rv			E207K, H445Y	LAM	1	1
14722_8#9_H37Rv		T812I	S450L, A692T	LAM	1	1
14722_8#90_H37Rv		E757A	S450L, A692T	LAM	1	1
14722_8#91_H37Rv			S450L	LAM	1	1
14722_8#92_H37Rv			Q409R, S450L	LAM	1	1
14722_8#93_H37Rv		G594E, A734V	S450L	Haarlem	1	1
14722_8#94_H37Rv			S450L	LAM	1	1
14722_8#95_H37Rv		G594E	Y308D, S450L	Other small clades	1	1
14892_2#1_H37Rv		V517L	S450L	LAM	1	1
14892_2#10_H37Rv		G594E		Haarlem	0	0
14892_2#11_H37Rv			L452P	LAM	1	0
14892_2#12_H37Rv			D435V	Beijing	1	1
14892_2#15_H37Rv		G594E		Other small clades	0	0
14892_2#16_H37Rv			H445D	LAM	1	1
14892_2#17_H37Rv				LAM	0	0
14892_2#18_H37Rv		G594E		Haarlem	0	0
14892_2#2_H37Rv				Beijing	0	0
14892_2#20_H37Rv				T	0	0
14892_2#21_H37Rv				LAM	0	0
14892_2#22_H37Rv				Other small clades	0	0
14892_2#23_H37Rv				LAM	0	0
14892_2#25_H37Rv		R1163H		LAM	0	0
14892_2#26_H37Rv		G594E	S450M	Haarlem	1	0
14892_2#27_H37Rv		G594E	S450M	Haarlem	1	1
14892_2#28_H37Rv			D435V	LAM	1	1
14892_2#29_H37Rv		N698S	S450L	LAM	1	1
14892_2#30_H37Rv		G594E		Haarlem	0	0
14892_2#31_H37Rv				LAM	0	0
14892_2#32_H37Rv			D435V	LAM	1	1
14892_2#33_H37Rv				LAM	0	0
14892_2#34_H37Rv			P45L, S450L	LAM	1	1
14892_2#35_H37Rv		V483A	Q432P	LAM	1	1
14892_2#36_H37Rv		G594E, G388A	E250G	Haarlem	0	0
14892_2#37_H37Rv			D435V	LAM	1	0
14892_2#39_H37Rv		G388A, R741S	H445D	LAM	1	1
14892_2#4_H37Rv				LAM	0	0
14892_2#40_H37Rv				LAM	0	0
14892_2#41_H37Rv		G594E	D435V	Haarlem	1	1
14892_2#42_H37Rv				LAM	0	0
14892_2#43_H37Rv		G388A, R741S	H445D	LAM	1	1
14892_2#44_H37Rv		G594E		Other small clades	0	0
14892_2#45_H37Rv		I851V, I997V	S450L	LAM	1	1
14892_2#46_H37Rv		G594E		Other small clades	0	0
14892_2#48_H37Rv		G594E	S450M	Haarlem	1	0
14892_2#49_H37Rv			D435Y, V695L	T	1	0
14892_2#5_H37Rv			H445Y	Beijing	1	1
14892_2#50_H37Rv			D435V	LAM	1	1

Contd...

Table 1: Contd...

Strain number	rpoA	rpoC	rpoB	Spoligotyping	RRDR	RIF resistance
14892_2#51_H37Rv			S450L, E550G	LAM	1	1
14892_2#52_H37Rv				LAM	0	0
14892_2#53_H37Rv				LAM	0	0
14892_2#54_H37Rv		G594E		Haarlem	0	0
14892_2#57_H37Rv		G594E		Haarlem	0	0
14892_2#58_H37Rv		I491V	S450L, A692T	LAM	1	1
14892_2#59_H37Rv		V483G	S450L, V695L	T	1	1
14892_2#60_H37Rv				LAM	0	0
14892_2#61_H37Rv		G594E		Other small clades	0	0
14892_2#63_H37Rv		W484S	S450L	Beijing	1	1
14892_2#64_H37Rv			D435V	LAM	1	1
14892_2#65_H37Rv		G594E	S450F	Other small clades	1	1
14892_2#66_H37Rv		P1040A	S450L	LAM	1	1
14892_2#67_H37Rv		V517L	S450L	LAM	1	1
14892_2#68_H37Rv		G594E		Other small clades	0	1
14892_2#69_H37Rv		P615L	S450L	LAM	1	1
14892_2#7_H37Rv		G594E		Haarlem	0	0
14892_2#70_H37Rv		V517L	S450L	LAM	1	1
14892_2#71_H37Rv		V1252M	S450L	LAM	1	1
14892_2#72_H37Rv				Other small clades	0	1
14892_2#73_H37Rv		G594E		Haarlem	0	1
14892_2#74_H37Rv		G594E	P45S, S450L	Haarlem	1	1
14892_2#75_H37Rv			D435V	LAM	1	1
14892_2#76_H37Rv		E1033K	F424L, I491F	LAM	0	1
14892_2#77_H37Rv			D435V	LAM	1	1
14892_2#78_H37Rv			D435V	LAM	1	1
14892_2#79_H37Rv			H445Y	Beijing	1	1
14892_2#8_H37Rv			D435V	LAM	1	1
14892_2#80_H37Rv		V1252M	S450L	LAM	1	1
14892_2#81_H37Rv			S450L	LAM	1	1
14892_2#82_H37Rv		G594E	S450L	Haarlem	1	1
14892_2#83_H37Rv			D435F	Beijing	1	1
14892_2#84_H37Rv			E207K, H445Y	LAM	1	1
14892_2#85_H37Rv		L516P, G594E	S450L	Haarlem	1	1
14892_2#86_H37Rv			D435V	LAM	1	1
14892_2#88_H37Rv		V483G	S450L, V695L	T	1	1
14892_2#89_H37Rv		S428T, G594E	V170F, I488L	Other small clades	0	1
14892_2#9_H37Rv		R741S	S450F	LAM	1	1
14892_2#90_H37Rv		G594E, A734V	S450L	Haarlem	1	1
14892_2#91_H37Rv		V483G	S450L	LAM	1	1
14892_2#92_H37Rv			E207K, H445Y	LAM	1	1
14892_2#93_H37Rv			S450L, V496M, V695L	T	1	1
14892_2#94_H37Rv		G594E, A734V	S450L	Haarlem	1	1
14892_2#95_H37Rv			D435V	LAM	1	1
14893_2#1_H37Rv				LAM	0	0
14893_2#10_H37Rv			D435V	LAM	1	1
14893_2#11_H37Rv		G594E		Haarlem	0	0
14893_2#12_H37Rv		G594E	S450L	Haarlem	1	1
14893_2#13_H37Rv				LAM	0	0
14893_2#14_H37Rv		G594E		Haarlem	0	1
14893_2#15_H37Rv			V695L	T	0	0
14893_2#16_H37Rv				Haarlem	0	0
14893_2#17_H37Rv		G594E	L452P	Haarlem	1	0
14893_2#18_H37Rv				LAM	0	1

Contd...

Table 1: Contd...

Strain number	rpoA	rpoC	rpoB	Spoligotyping	RRDR	RIF resistance
14893_2#19_H37Rv			D435V	LAM	1	1
14893_2#2_H37Rv				LAM	0	0
14893_2#20_H37Rv		G594E		Other small clades	0	0
14893_2#21_H37Rv			S450L, V695L	T	1	1
14893_2#22_H37Rv		G594E	I491F	Other small clades	0	1
14893_2#23_H37Rv		P1040A	S450L	LAM	1	1
14893_2#24_H37Rv				LAM	0	0
14893_2#25_H37Rv			S450L	LAM	1	1
14893_2#26_H37Rv			D435V	LAM	1	1
14893_2#27_H37Rv			V695L	T	0	0
14893_2#28_H37Rv				T	0	1
14893_2#29_H37Rv			H445D, V695L	T	1	1
14893_2#3_H37Rv		G594E		Haarlem	0	0
14893_2#30_H37Rv		G594E	V170F	Other small clades	0	1
14893_2#31_H37Rv			I965V	Caprae	0	0
14893_2#32_H37Rv			V695L	T	0	0
14893_2#33_H37Rv		G594E		Haarlem	0	0
14893_2#34_H37Rv			S450L, L731P	LAM	1	1
14893_2#35_H37Rv				LAM	0	0
14893_2#36_H37Rv		T812I	S450L, A692T	LAM	1	1
14893_2#37_H37Rv		N698S, R741S	S450L	LAM	1	1
14893_2#38_H37Rv		V517L	S450L	LAM	1	1
14893_2#39_H37Rv				LAM	0	0
14893_2#4_H37Rv			A286V, S450L	LAM	1	1
14893_2#40_H37Rv		P1040A	S450L	LAM	1	1
14893_2#41_H37Rv		G594E		Haarlem	0	0
14893_2#42_H37Rv		G594E		Haarlem	0	0
14893_2#43_H37Rv		G594E, P1040A	S450L, R827C	Haarlem	1	1
14893_2#44_H37Rv				LAM	0	0
14893_2#45_H37Rv		P1040A	S450L	LAM	1	1
14893_2#46_H37Rv		I491V	S450L, A692T	LAM	1	1
14893_2#47_H37Rv		G594E	S450L, G973D	Haarlem	1	1
14893_2#48_H37Rv			S450L, L731P	LAM	1	1
14893_2#49_H37Rv		G594E	H445D	Haarlem	1	1
14893_2#5_H37Rv		A172V		EAI2-Manilla	0	0
14893_2#50_H37Rv			S450L, E550G	LAM	1	1
14893_2#51_H37Rv		G594E		Other small clades	0	0
14893_2#52_H37Rv		V483G	S450L, V695L	T	1	1
14893_2#53_H37Rv				LAM	0	0
14893_2#54_H37Rv		A788V	D435F	Beijing	1	1
14893_2#55_H37Rv			D435F	Beijing	1	1
14893_2#56_H37Rv			D435V	LAM	1	1
14893_2#57_H37Rv		G594E		Haarlem	0	0
14893_2#58_H37Rv			S450L, E550G	LAM	1	1
14893_2#59_H37Rv			S450L, E550G	LAM	1	1
14893_2#6_H37Rv		G594E, V1252L	S450L	Haarlem	1	1
14893_2#60_H37Rv		G594E		Haarlem	0	0
14893_2#61_H37Rv			S450L, Q980R	LAM	1	1
14893_2#62_H37Rv			D435V	LAM	1	1
14893_2#63_H37Rv		G594E	D435V	Haarlem	1	1
14893_2#64_H37Rv		G594E	E250G	Haarlem	0	1
14893_2#65_H37Rv				LAM	0	0
14893_2#66_H37Rv			S450L, I480V, A692T	LAM	1	1
14893_2#67_H37Rv		G594E		Haarlem	0	0

Contd...

Table 1: Contd...

Strain number	rpoA	rpoC	rpoB	Spoligotyping	RRDR	RIF resistance
14893_2#68_H37Rv			D435V	LAM	1	1
14893_2#69_H37Rv			D435V	LAM	1	1
14893_2#7_H37Rv			D435V	LAM	1	1
14893_2#70_H37Rv			L452P, V695L	T	1	1
14893_2#71_H37Rv		G594E	S450L	Haarlem	1	1
14893_2#72_H37Rv		V483G	S450L, V695L	T	1	1
14893_2#73_H37Rv			D435V	LAM	1	1
14893_2#74_H37Rv		G594E, V1206G, D1218A	S450W	Haarlem	1	1
14893_2#75_H37Rv		V1252M	S450L	LAM	1	1
14893_2#76_H37Rv		G594E		Haarlem	0	1
14893_2#77_H37Rv			D435V	LAM	1	1
14893_2#78_H37Rv				Beijing	0	1
14893_2#79_H37Rv		V517L	S450L	LAM	1	1
14893_2#8_H37Rv		P1040A	S450L	LAM	1	1
14893_2#80_H37Rv		G332R	S450L	LAM	1	1
14893_2#81_H37Rv			H445Y	Beijing	1	1
14893_2#82_H37Rv			E207K, H445Y	LAM	1	1
14893_2#83_H37Rv			D435V	Beijing	1	1
14893_2#84_H37Rv		G594E	D435V	Haarlem	1	1
14893_2#85_H37Rv		G594E	S450L	Other Small Clades	1	1
14893_2#86_H37Rv		W484S	S450L	Beijing	1	1
14893_2#87_H37Rv		V483A, G594E	S450L	Other small clades	1	1
14893_2#88_H37Rv		V483G	S450L, V695L	T	1	1
14893_2#89_H37Rv		W484S	S450L	Beijing	1	1
14893_2#9_H37Rv		L507V	S450L	T	1	1
14893_2#90_H37Rv		V483A	S450L	Beijing	1	1
14893_2#91_H37Rv			D435V	LAM	1	1
14893_2#92_H37Rv		G594E	S428R, H445P	Other small clades	1	1
14893_2#93_H37Rv		V483A, G594E	S450L	Other small clades	1	1
14893_2#94_H37Rv		G571R, G594E	S450L	Haarlem	1	1
14893_2#95_H37Rv		V483G	S450L	LAM	1	1
14893_3#65_H37Rv			S450L, A692T	LAM	1	1
14893_3#66_H37Rv			S450L, V695L	T	1	1
15277_3#50_H37Rv			S450L, V695L	T	1	1
15277_3#51_H37Rv		R741S, V1039A	S450L	NA	0	0
15277_3#52_H37Rv			S450L, V695L	T	1	1
15277_3#53_H37Rv			S450L, E550G	LAM	1	1
15277_3#55_H37Rv		V517L	S450L	NA	0	0
15277_3#57_H37Rv			S450L	LAM	1	1

RIF: Rifampicin, RRDR: RIF resistance-determining region

47% of all the potentially compensatory mutations proposed in this study arose in these regions. E49A and P54 L were isolated from the rest, located within the first 100 amino acids of the protein. The distribution of the mutations found here and those reported previously was found to be remarkably similar (P = 0.2318, Kolmogorov–Smirnov nonparametric test).

To better analyze this distribution, we classified the amino acids into the following groups: basic (R, K, and H), acidic (D and E), polar (G, S, Y, C, Q, T, and N), and nonpolar (F, L, W, P, I, M, V, and A). Based on this, we found that 54% of the potentially compensatory mutations found in rpoC preserve their physical–chemical properties despite the amino acid change.

Distribution of the potential compensatory rpoC mutations on the three-dimensional structure of *Mycobacterium tuberculosis* RNApol

The potentially compensatory mutations that we described above [Supplementary Table 3] were mainly distributed on the surface of the three-dimensional (3D)-crystal structure of the β' subunit (82%), and 9 were buried below the surface [Figure 5]. Only one of the buried mutated sites, H525Q, has been previously reported to be compensatory.[17] E49A, which was visibly separated from the rest in the primary structure of rpoC,

can be seen on the side of the subunit 3D-structure, distant from the main cluster of mutations [Figure 5].

β' RNApol has a domain that resembles the shape of an "arm" that hosts a particularly large number of mutations. The "arm" is best visible on the right side of the structure [Figure 6]. Our data show that some of the putative compensatory mutations fall in similar spots as those that have been previously reported, in particular those located in this "arm". Interestingly, the reported mutation P54 L appears hidden below the surface along with E49A [Figure 6]. We also found that two of the proposed mutations, G332R and W484S, occur in regions where other mutations have already been reported.

The 3D-crystal structure of β' interacting with the β subunit confirms that several mutations fall in the regions of interaction (S428T, L507V, A734V, I997V, and V1252 LM) [Figure 7]. These represent 13.5% of the surface mutations. All of the putative compensatory mutations that we identified in the rpoC/rpoB contact surface conserved their physical–chemical properties during the amino acid change [Figure 7].

Discussion

This study describes 35 novel mutations in rpoC, associated with RIF resistance and rpoB mutations, detected in Peruvian clinical isolates of *M. tuberculosis*. These mutations may serve to compensate for any loss of fitness caused by a mutation in rpoB, as reported in other studies.[8,10,13-15] Of note, some were found to be associated with controversial mutations in rpoB that have been identified in RIF-susceptible MTB strains.

The similarity in the distribution of the potentially compensatory rpoC mutations found in Peru compared to those reported

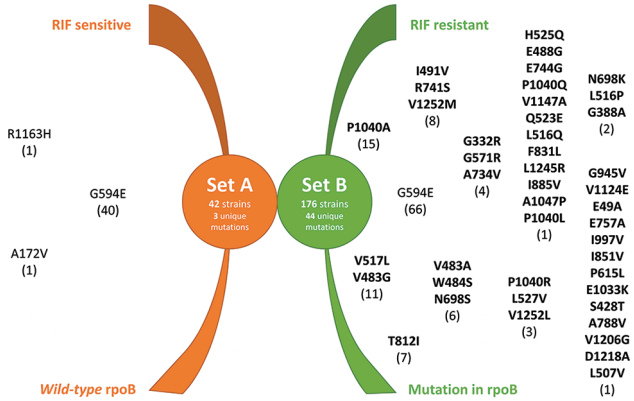

Figure 3: Graphical representation of rpoC mutations in Set A (mutations in rifampicin susceptible strains with wild-type rpoB, shown in orange on the left) and Set B (mutations in rifampicin-resistant strains harboring rpoB mutations, shown in green on the right). Mutations in bold are those found exclusively in Set B and thus possible compensatory mutations

Table 2: Mutations in the rpoB critical region (rifampicin resistance-determining region) not associated with rifampicin resistance

Strain number	Mutation in rpoC	Mutation in RRDR	Spoligotyping
14722_6#36_H37Rv	0	D435F	Beijing
14722_6#41_H37Rv	0	D435F	Beijing
14722_6#53_H37Rv	0	D435V	LAM
14892_2#11_H37Rv	0	L452P	LAM
14892_2#37_H37Rv	0	D435V	LAM
14892_2#49_H37Rv	0	D435Y, V695L	T
14722_6#17_H37Rv	G594E	L452P	Haarlem
14722_7#75_H37Rv	G594E	L452P	Haarlem
14892_2#26_H37Rv	G594E	S450M	Haarlem
14892_2#48_H37Rv	G594E	S450M	Haarlem
14893_2#17_H37Rv	G594E	L452P	Haarlem

Trait present: Mutation in amino acid code, Trait absent: 0.
RIF: Rifampicin, RRDR: RIF resistance-determining region

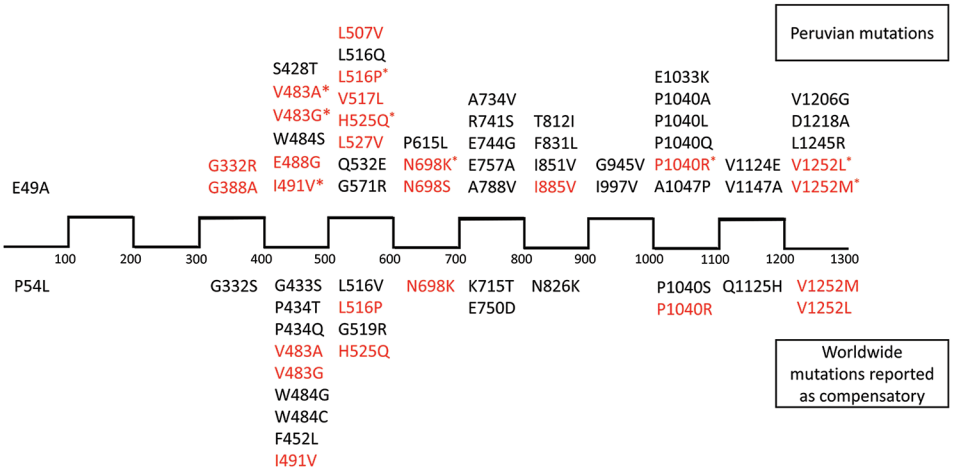

Figure 4: Linear distribution of the identified putative rpoC compensatory mutations throughout the protein sequence (scale: 100 amino acids per space). The upper part of the image lists the mutations found in our analysis, while the lower part shows mutations reported in the scientific literature. Mutations in red are those already reported elsewhere, but not necessarily underwent experimental confirmation of compensatory behavior. Upper mutations marked with an asterisk (*) are those already reported as compensatory by experimental assays

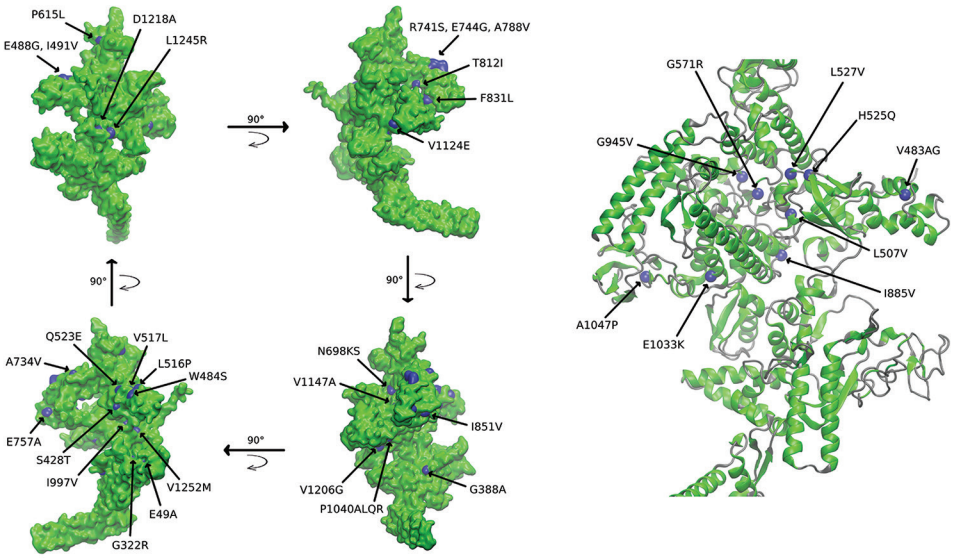

Figure 5: Crystal of the *Mycobacterium tuberculosis* β' subunit (rpoC), represented according to its surface and its secondary structure. The surface, β sheets, and α helices are shown in green, while loops are colored gray. The mutations found in our study are either pictured in blue patches (surface) or as Van der Waals spheres (alpha carbon of buried residues only, radius scale = 1.4 Å). The arrows in the left panel indicate the direction of the rotation of the protein along the Y axis

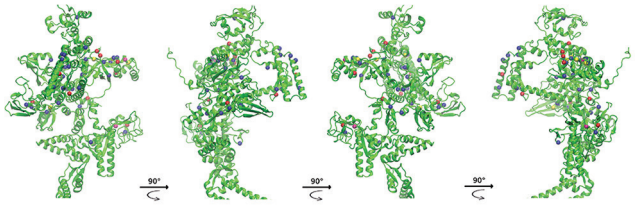

Figure 6: Crystal of the *Mycobacterium tuberculosis* β' subunit (rpoC), represented by its secondary structure. The β sheets and α helices are shown in green, while loops are colored gray. Different types of mutations are shown in different colors, as follows: Blue for the putative compensatory mutations found in our study, red for previously reported compensatory mutations, yellow for mutations reported in the literature and also found in our study, and magenta for mutations found in this study that had another amino acid change (for the same position) reported in the literature. All the mutations are represented as Van der Waals spheres (alpha carbon only, radius scale = 1.4 Å). The arrows indicate the direction of the rotation of the protein along the Y axis

elsewhere is striking. The spatial distribution along the protein sequence reveals a similar, almost mirrored pattern, suggesting that compensatory mutations in rpoC preferentially occur in certain positions. These cluster-like regions could be related to functional aspects of the protein or may be structurally critical. According to the Pfam protein domain family database,[30-32] these high-frequency mutation zones correspond to specific domains. The first zone (amino acids 400–500, second domain) contains the protein's active site, the second zone (amino acids 500–600, third domain) contains the pore thought to act as a channel for nucleotides, and the third zone (amino acids 1000–1100, fifth domain) includes the discontinuous cleft required to form the channel where DNA is bound. Mutations occurring in these critical regions may cause significant changes in function.

Surface mutations are more likely to interact with other subunits at a quaternary structure level or with other proteins. Furthermore, a surface mutation would not significantly affect the tertiary structure of the protein, whereas a deep mutation may destabilize the whole subunit. Thus, it would be favorable for compensatory mutations to be located on the surface of the protein, even more on the interaction sites. In our study, the position of the novel mutations on the crystal structure revealed that 82.2% fall on the surface of the β' subunit.

Moreover, only six mutations were found in the region of interaction between the surfaces of β and β', which constitute our most likely candidates for compensation, as a mutation at this site would produce a conformational change sufficient to affect the binding between the subunits but not significant enough to disrupt the inner core. Two of these six mutations have already been reported, and although W484S was not found in the literature, W484G/C has in fact been previously described. Moreover, as noted before and according to our expectations that mutations would result in minimal (but significant) structural changes, every mutation that we reported as a candidate conserves the physical–chemical properties of the original amino acid, favoring the maintenance of the protein's stability.[33] Several mutations found in this study are very close to mutations reported elsewhere, and in some cases, even at the same position. This may also indicate the preference of certain locations for the emergence of compensatory mutations.

Only 20.5% of the mutations that we identified coincided with previously reported compensatory mutations.[16,17,19] This is likely to be due to the large genotypic diversity among Peruvian *M. tuberculosis* strains. Of note, in our study, most of the isolates (250) corresponded to the "LAM" spoligotype.

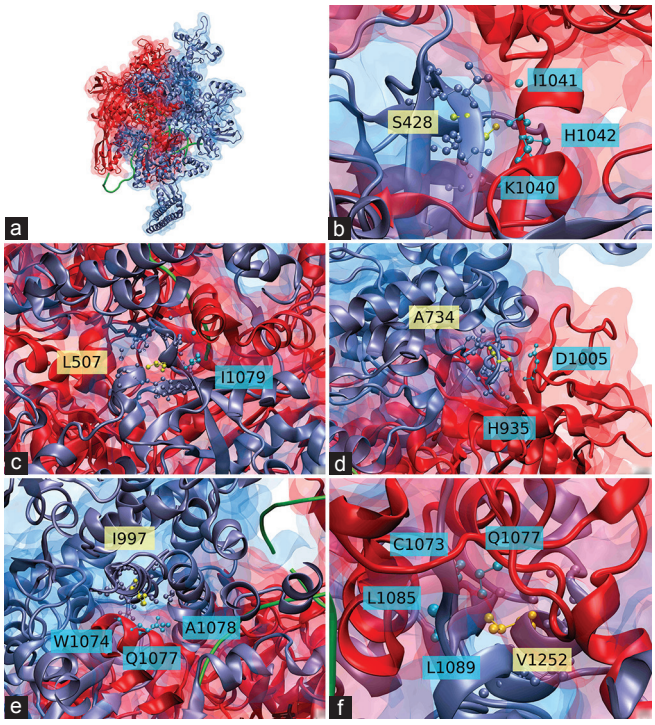

Figure 7: Crystal structure of *Mycobacterium tuberculosis* transcription complex showing interaction of β' (rpoC) and β (rpoB) subunits. Proteins are represented by their secondary structure and their surface; rpoC is shown in iceblue and rpoB is shown in red. Two DNA strands are shown in green as a ribbon, along with a rifampin molecule in light blue (panel a) or black (panels b– f). rpoC residues whose sites would carry compensatory mutations are marked in yellow in Corey-Pauling-Koltun model (CPK) representation (side chain included). rpoB residues within 5.00 Å (6.00 Å for panel c) of the marked rpoC residues are shown in cyan in CPK representation. rpoC residues within that distance are shown in iceblue in CPK representation. (a) Provides a full, general view of the interaction between the subunits, while the rest of the panels show the putative rpoC compensatory mutations interacting with close residues in rpoB: S428T (b), L507V (c), A734V (d), I997V (e), V1252 LM (f). The amino acids are indicated according to standard code, according to the subunit they belong (yellow for rpoC, blue for rpoB)

We identified just three mutations in rpoA, which is thought to be, along with rpoC, the most important site of RIF-resistant compensatory mutations. All three have been previously reported. This is in keeping with previous findings that the majority of compensatory mutations arise in rpoC,[16,17,19] suggesting that these may be more effective at recovering loss of fitness arising from mutations in rpoB.[16,21-23]

Furthermore, we found mutations in rpoC that were not associated with any mutations in rpoB but were present in RIF-resistant strains. It is likely that the appearance of these mutations is random, but further studies are necessary to clarify their true nature.

We reported 11 RIF-susceptible strains harboring mutations in the RRDR, with 5 harboring rpoC G594E [Table 2]. This is unlikely to be due to incorrect DST results, as MODS is a highly sensitive (98%) and specific (98.6%) method.[26] Rather, this finding led us to further analyze our data in search of patterns in the distribution of rpoC G594E. We found it to be present in 17.3% of RIF-resistant and 39.8% of RIF-susceptible strains, being by far the most common mutation. G594E has been described in strains with no rpoB mutations and is a SNP present in all strains from a group within the Haarlem genotype, which would mean that it arises independently of rpoB mutations and RIF resistance. Taken together, these features make it unlikely that G594E is associated with drug resistance, as broadly stated before.[14,16,17,20,22,23]

In view of this, we found G594E in both of our analysis groups: set A (rpoC mutations in RIF-susceptible strains without rpoB mutations) and set B (rpoC mutations in RIF-resistant strains harboring rpoB mutations). Furthermore, we found it in the five previously mentioned RIF-susceptible strains with mutations in the RRDR [Table 2]. Three of these harbor rpoB L452P, described elsewhere.[19,34] The remaining two harbor rpoB S450M, which is similar to the widely-known S450 L[10,17-19,34] and their described variants S450F/Q/Y/W.[19,21,34-38] The detection of rpoB S450M is puzzling because a Ser to Met mutation would require changes in two nucleotides (TCG → ATG), making it unlikely to occur spontaneously. It is possible that S450F occurred first (TCG → TTG), followed by a change from TTG → ATG, resulting in the end of a Ser to Met change. It is notable that both amino acids are nonpolar; that is, they have similar physicochemical properties. We could not find any other study describing S450M, so we encourage further investigation of the role and evolution of this mutation, as our theory of its development is mere speculation.

Nonetheless, we cannot avoid discussing the dichotomy regarding the appearance of rpoC G594E. Considering all of the above, we propose another possibility toward its appearance, which is that under certain conditions some mutations in rpoB do not cause RIF resistance but may still decrease the fitness of the bacteria. This would subdue the bacteria to selective pressure, leading to the positive selection of a compensatory mutation in rpoC. Until now, most investigations have attributed the unlikeliness of G594E being associated with drug resistance to its uncharacteristic appearance, but believe that there may be other ways for this to happen and thus we cannot draw any conclusions about the true nature of this mutation until more is known.

Our study identifies mutations that may have a compensatory role, but there is a need for experimental confirmation of the effect of each mutation as some of these are likely to arise at random. Previous studies have used methods such as the measurement of the catalytic activity and rate of bacterial growth.[14-18] We suggest that the use of site-directed mutagenesis techniques, such as the emerging CRISPR-Cas9 technology,[39] would be useful to assess the fitness of mutant strains.

Conclusion

This study identifies 35 putative novel compensatory mutations in the β' subunit of *M. tuberculosis* RNApol. Six

of these (S428T, L507V, A734V, I997V, and V1252 LM) are considered most likely to have a compensatory role, as they fall in the interaction zone of the two subunits and the mutation did not lead to any change in the protein's physical–chemical properties. Further studies assessing the fitness change in strains harboring these mutations are needed to confirm their potential compensatory role. In addition, it would be interesting to evaluate the evolutionary relationship of the various SNPs found in this study.

Financial support and sponsorship

This research was funded by the Wellcome Trust (Ref: 099805/Z/12/Z), the Grand Challenge Canada (GCC Number 0687-01-10), and the LOREAL-UNESCO-CONCYTEC 2014 award.

PS was supported by a Wellcome Trust Intermediate fellowship.

Conflicts of interest

There are no conflicts of interest.

References

1. World Health Organization. World Health Statistics 2016: Monitoring Health for the SDGs, Sustainable Development Goals. Geneva: World Health Organization. Available from: https://apps.who.int/iris/bitstream/handle/10665/324835/9789241565707-eng.pdf?ua=1. [Last updated on 2020 Jan 13; Last accessed on 2019 Jan 20].
2. Silva C, Bermúdez V, Arraiz N, Bermúdez F, Rodríguez M, Valdelamar L, *et al*. Fármacos de primera línea utilizados en el tratamiento de la tuberculosis. Arch Venez Farmacol Y Ter 2007;26:21-6.
3. World Health Organization. Global Tuberculosis Report. Geneva: World Health Organization Available from: https://apps.who.int/iris/bitstream/handle/10665/329368/9789241565714-eng.pdf?ua=1. [Last updated on 2020 Jan 13; Last accessed on 2019 Jan 20].
4. World Health Organization. Tuberculosis. Available from: https://www.who.int/en/news-room/fact-sheets/detail/tuberculosis. [Last updated on 2020 Jan 13; Last accessed on 2019 Jan 20]
5. World Health Organization. Drug resistant TB-Data by WHO Region. Available from: http://apps.who.int/gho/data/view.main.MDRTBWHOREG?lang=en. [Last updated on 2020 Jan 13; Last accessed on 2019 Jan 20]
6. Ministry of Health. Lima: Analysis of the Epidemiological Situation in Peru. Available from: http://bvs.minsa.gob.pe/local/MINSA/3446.pdf. [Last updated on 2020 Jan 13; Last accessed on 2019 Jan 20]
7. Mendoza L, Rodríguez H, Rodríguez E, Tarazona C. Conocimientos básicos sobre antibacterianos. 1st ed. Maracaibo: Editorial Astro Data; 1999. p. 73-5.
8. Koch A, Mizrahi V, Warner DF. The impact of drug resistance on *Mycobacterium tuberculosis* physiology: What can we learn from rifampicin? Emerg Microbes Infect 2014;3:e17.
9. Hughes D, Brandis G. Rifampicin resistance: Fitness costs and the significance of compensatory evolution. Antibiotics (Basel) 2013;2:206-16.
10. Alifano P, Palumbo C, Pasanisi D, Talà A. Rifampicin-resistance, rpoB polymorphism and RNA polymerase genetic engineering. J Biotechnol 2015;202:60-77.
11. Telenti A, Imboden P, Marchesi F, Lowrie D, Cole S, Colston MJ, *et al*. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. Lancet 1993;341:647-50.
12. Goldstein BP. Resistance to rifampicin: A review. J Antibiot (Tokyo) 2014;67:625-30.
13. Brandis G, Hughes D. Genetic characterization of compensatory evolution in strains carrying rpoB Ser531Leu, the rifampicin resistance mutation most frequently found in clinical isolates. J Antimicrob Chemother 2013;68:2493-7.
14. Song T, Park Y, Shamputa IC, Seo S, Lee SY, Jeon HS, *et al*. Fitness costs of rifampicin resistance in *Mycobacterium tuberculosis* are amplified under conditions of nutrient starvation and compensated by mutation in the β' subunit of RNA polymerase. Mol Microbiol 2014;91:1106-19.
15. Mariam DH, Mengistu Y, Hoffner SE, Andersson DI. Effect of rpoB mutations conferring rifampin resistance on fitness of *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 2004;48:1289-94.
16. Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, *et al*. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. Nat Genet 2011;44:106-10.
17. Cohen KA, Abeel T, Manson McGuire A, Desjardins CA, Munsamy V, Shea TP, *et al*. Evolution of extensively drug-resistant tuberculosis over four decades: Whole genome sequencing and dating analysis of *Mycobacterium tuberculosis* isolates from kwazulu-natal. PLoS Med 2015;12:e1001880.
18. Brandis G, Wrande M, Liljas L, Hughes D. Fitness-compensatory mutations in rifampicin-resistant RNA polymerase. Mol Microbiol 2012;85:142-51.
19. Zhang H, Li D, Zhao L, Fleming J, Lin N, Wang T, *et al*. Genome sequencing of 161 *Mycobacterium tuberculosis* isolates from China identifies genes and intergenic regions associated with drug resistance. Nat Genet 2013;45:1255-60.
20. Ali A, Hasan Z, McNerney R, Mallard K, Hill-Cawthorne G, Coll F, *et al*. Whole genome sequencing based characterization of extensively drug-resistant *Mycobacterium tuberculosis* isolates from Pakistan. PLoS One 2015;10:e0117771.
21. Casali N, Nikolayevskyy V, Balabanova Y, Ignatyeva O, Kontsevaya I, Harris SR, *et al*. Microevolution of extensively drug-resistant tuberculosis in Russia. Genome Res 2012;22:735-45.
22. de Vos M, Müller B, Borrell S, Black PA, van Helden PD, Warren RM, *et al*. Putative compensatory mutations in the rpoC gene of rifampin-resistant *Mycobacterium tuberculosis* are associated with ongoing transmission. Antimicrob Agents Chemother 2013;57:827-32.
23. Li QJ, Jiao WW, Yin QQ, Xu F, Li JQ, Sun L, *et al*. Compensatory mutations of rifampin resistance are associated with transmission of multidrug-resistant *Mycobacterium tuberculosis* Beijing genotype strains in China. Antimicrob Agents Chemother 2016;60:2807-12.
24. Grandjean L, Gilman RH, Iwamoto T, Köser CU, Coronel J, Zimic M, *et al*. Convergent evolution and topologically disruptive polymorphisms among multidrug-resistant tuberculosis in Peru. PLoS One 2017;12:e0189838.
25. Caviedes L, Lee TS, Gilman RH, Sheen P, Spellman E, Lee EH, *et al*. Rapid, efficient detection and drug susceptibility testing of *Mycobacterium tuberculosis* in sputum by microscopic observation of broth cultures. The Tuberculosis Working Group in Peru. J Clin Microbiol 2000;38:1203-8.
26. Minion J, Leung E, Menzies D, Pai M. Microscopic-observation drug susceptibility and thin layer agar assays for the detection of drug resistant tuberculosis: A systematic review and meta-analysis. Lancet Infect Dis 2010;10:688-98.
27. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, *et al*. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol 1997;35:907-14.
28. Lin W, Mandal S, Degen D, Liu Y, Ebright YW, Li S, *et al*. Structural Basis of *Mycobacterium tuberculosis* Transcription and Transcription Inhibition. Mol Cell 2017;66:169-790.
29. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph 1996;14:33-8, 27-8.
30. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, *et al*. The Pfam protein families database: Towards a more sustainable future. Nucleic Acids Res 2016;44:D279-85.
31. Severinov K, Mustaev A, Kukarin A, Muzzin O, Bass I, Darst SA, *et al*. Structural modules of the large subunits of RNA polymerase. Introducing archaebacterial and chloroplast split sites in the beta and beta' subunits of *Escherichia coli* RNA polymerase. J Biol Chem 1996;271:27969-74.
32. Cramer P, Bushnell DA, Kornberg RD. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science

2001;292:1863-76.
33. Ladunga I, Smith RF. Amino acid substitutions preserve protein folding by conserving steric and hydrophobicity properties. Protein Eng. Des. Sel. 1997;10:187-96.
34. Casali N, Broda A, Harris SR, Parkhill J, Brown T, Drobniewski F. Whole genome sequence analysis of a large isoniazid-resistant tuberculosis outbreak in London: A Retrospective Observational Study. PLoS Med 2016;13:e1002137.
35. Salamon H, Yamaguchi KD, Cirillo DM, Miotto P, Schito M, Posey J, *et al*. Integration of published information into a resistance-associated mutation database for *Mycobacterium tuberculosis*. J Infect Dis 2015;211 Suppl 2:S50-7.
36. Yang C, Luo T, Shen X, Wu J, Gan M, Xu P, *et al*. Transmission of multidrug-resistant *Mycobacterium tuberculosis* in Shanghai, China: A retrospective observational study using whole-genome sequencing and epidemiological investigation. Lancet Infect Dis 2017;17:275-84.
37. Ullah I, Ahmad W, Shah AA, Shahzada A, Tahir Z, Qazi O, *et al*. Detection of rifampicin resistance of *Mycobacterium tuberculosis* using multiplex allele specific polymerase chain reaction (MAS-PCR) in Pakistan. Infect Genet Evol 2019;71:42-6.
38. Miotto P, Cabibbe AM, Borroni E, Degano M, Cirillo DM. Role of Disputed Mutations in the rpoB Gene in Interpretation of Automated Liquid MGIT Culture Results for Rifampin Susceptibility Testing of Mycobacterium tuberculosis. J Clin Microbiol 2018;56:e01599-17.
39. Rock JM, Hopkins FF, Chavez A, Diallo M, Chase MR, Gerrick ER, *et al*. Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. Nat Microbiol 2017;2:16274.

Supplementary Table 1: rpoC mutations in rifampicin-sensitive strains without mutations in rpoB (Set A)

Strain number	Spoligotyping	Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change
14722_6#15_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14722_6#30_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14722_6#31_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14722_6#39_H37Rv	Other small clades	g	A	ggg	2	gAg	G594E
14722_6#42_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14722_6#58_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14722_6#71_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14722_6#79_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14722_7#29_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14722_7#31_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14722_7#35_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14722_7#47_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14722_7#49_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14722_7#51_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14722_7#61_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14722_7#70_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14722_7#73_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14722_7#85_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14722_7#88_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14722_7#93_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14892_2#10_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14892_2#18_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14892_2#25_H37Rv	LAM	G	A	cgc	2	cAc	R1163H
14892_2#30_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14892_2#44_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14892_2#46_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14892_2#54_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14892_2#57_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14892_2#61_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14892_2#7_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14893_2#11_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14893_2#20_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14893_2#3_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14893_2#33_H37Rv	T	G	A	ggg	2	gAg	G594E
14893_2#41_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14893_2#42_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14893_2#5_H37Rv	EAI2-manilla	C	T	gcc	2	gTc	A172V
14893_2#51_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E
14893_2#57_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14893_2#60_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14893_2#67_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E
14892_2#15_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E

Supplementary Table 2: rpoC mutations in rifampicin-resistant strains that harbored rpoB mutations (Set B)

Strain Number	Spoligotyping	Mutations in rpoC						Mutations in rpoB					
		Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change	Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change
14722_6#1_H37Rv	T	T	C	gtg	2	gCg	V483A	C	T	tcg	2	tTg	S450L
14722_6#12_H37Rv	Beijing	C	G	ccg	2	cGg	P1040R	C	T	tcg	2	tTg	S450L
14722_6#13_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14722_6#16_H37Rv	LAM	A	G	atc	1	Gtc	I491V	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_6#2_H37Rv	LAM	A	G	atc	1	Gtc	I491V	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_6#28_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14722_6#29_H37Rv	Beijing	G	C	tgg	2	tCg	W484S	C	T	tcg	2	tTg	S450L
14722_6#3_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
14722_6#33_H37Rv	LAM	C	A	cgc	1	Agc	R741S	C,A	T,G	cac	1,2	TGT	H445C
14722_6#34_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
								A	G	atg	1	Gtg	M920V
14722_6#38_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14722_6#43_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	cac	1	Tac	H445Y
14722_6#45_H37Rv	LAM	G	T	gtg	1	Ttg	V517L	C	T	tcg	2	tTg	S450L
14722_6#51_H37Rv	LAM	G	T	gtg	1	Ttg	V517L	C	T	tcg	2	tTg	S450L
14722_6#55_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
		A	G	atc	1	Gtc	I491V						
14722_6#56_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	A	T	gac	2	gTc	D435V
14722_6#57_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14722_6#59_H37Rv	T	T	G	ttg	1	Gtg	L527V	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L
14722_6#63_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
14722_6#67_H37Rv	LAM	C	T	acc	2	aTc	T812I	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_6#72_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	A	T	gac	2	gTc	D435V
14722_6#73_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450F
14722_6#86_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14722_6#87_H37Rv	LAM	A	G	atc	1	Gtc	I491V	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_6#88_H37Rv	LAM	A	G	atc	1	Gtc	I491V	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_6#89_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	A	T	gac	2	gTc	D435V

Contd...

Supplementary Table 2: Contd...

Strain Number	Spoligotyping	Mutations in rpoC						Mutations in rpoB					
		Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change	Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change
14722_6#9_H37Rv	Other small clades	G	A	ggg	3	gAg	G594E	C	T	tcg	2	tTg	S450L
	Other small clades	T	G	ttg	1	Gtg	L527V						
14722_6#91_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
14722_6#94_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450F
14722_7#10_H37Rv	Beijing	C	G	cac	3	caG	H525Q	C	T	tcg	2	tTg	S450L
14722_7#12_H37Rv	LAM	C	A	cgc	1	Agc	R741S	C	T	tcg	2	tTg	S450L
								C	A	gac	3	gaA	D574E
14722_7#17_H37Rv	Beijing	G	C	tgg	2	tCg	W484S	C	T	tcg	2	tTg	S450L
14722_7#19_H37Rv	LAM	T	C	gtg	2	gCg	V483A	C	T	tcg	2	tTg	S450L
14722_7#20_H37Rv	T	T	G	ttg	1	Gtg	L527V	C	T	tcg	2	tTg	S450L
14722_7#21_H37Rv	T	G	C	ggc	1	Cgc	G332R	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L
14722_7#22_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	A	T	gac	2	gTc	D435V
14722_7#23_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
	Haarlem	A	G	gaa	2	gGa	E488G						
14722_7#24_H37Rv	LAM	C	G	ccg	2	cGg	P1040R	C	T	tcg	2	tTg	S450L
14722_7#26_H37Rv	Other small clades	A	G	gag	2	gGg	E744G	C	T	tcg	2	tTg	S450L
	Other small clades	G	A	ggg	2	gAg	G594E						
14722_7#30_H37Rv	LAM	G	T	gtg	1	Ttg	V517L	C	T	tcg	2	tTg	S450L
14722_7#34_H37Rv	T	C	A	ccg	2	cAg	P1040Q	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L
14722_7#37_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	ccg	1	Tcg	P45S
								C	T	tcg	2	tTg	S450L
14722_7#42_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	gcg	2	gTg	A286V
								C	T	tcg	2	tTg	S450L
14722_7#44_H37Rv	T	G	T	gtg	1	Ttg	V1252L	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L
14722_7#50_H37Rv	LAM	G	T	gtg	1	Ttg	V517L	C	T	tcg	2	tTg	S450L
14722_7#53_H37Rv	Other small clades	T	C	gtg	2	gCg	V1147A	C	T	tcg	2	tTg	S450L
		G	A	ggg	2	gAg	G594E						
14722_7#58_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L

Contd...

Supplementary Table 2: Contd...

Strain Number	Spoligotyping	Mutations in rpoC						Mutations in rpoB					
		Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change	Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change
14722_7#62_H37Rv	Beijing	C	G	cag	1	Gag	Q523E	C	T	tcg	2	tTg	S450L
14722_7#63_H37Rv	LAM	T	A	ctg	2	cAg	L516Q	C	T	tcg	2	tTg	S450L
14722_7#65_H37Rv	LAM	C	T	acc	2	aTc	T812I	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_7#66_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
14722_7#7_H37Rv	Other small clades	G	A	ggg	1	Agg	G571R	G	T	gtc	1	Ttc	V170F
		G	A	ggg	2	gAg	G594E						
14722_7#74_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	A	G	gag	2	gGg	E250G
								C	T	tcg	2	tTg	S450L
								C	G	ttc	3	ttG	F971L
14722_7#76_H37Rv	Other small clades	T	C	ttc	1	Ctc	F831L	C	T	tcg	2	tTg	S450L
		G	A	ggg	2	gAg	G594E						
14722_7#78_H37Rv	Beijing	T	G	ctc	2	cGc	L1245R	C	T	tcg	2	tTg	S450L
14722_7#81_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	cac	1	Tac	H445Y
14722_7#82_H37Rv	T	G	C	ggc	1	Cgc	G332R	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L
14722_7#83_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
								A	G	atg	1	Gtg	M920V
14722_7#87_H37Rv	T	G	C	ggc	1	Cgc	G332R	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L
14722_7#90_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
14722_7#95_H37Rv	Other small clades	T	G	gtg	2	gGg	V483G	C	T	tcg	2	tTg	S450L
14722_7#95_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
14722_8#11_H37Rv	LAM	C	A	aac	3	aaA	N698K	C	T	tcg	2	tTg	S450L
14722_8#14_H37Rv	LAM	A	G	atc	1	Gtc	I885V	C	T	tcg	2	tTg	S450L
14722_8#15_H37Rv	LAM	G	C	gcc	1	Ccc	A1047P	C	T	tcg	2	tTg	S450L
14722_8#16_H37Rv	Haarlem	T	C	ctg	2	cCg	L516P	C	T	tcg	2	tTg	S450L
	Haarlem	G	A	ggg	2	gAg	G594E						
14722_8#17_H37Rv	LAM	G	A	gtg	1	Atg	V1252M	C	T	tcg	2	tTg	S450L
14722_8#18_H37Rv	LAM	G	A	gtg	1	Atg	V1252M	C	T	tcg	2	tTg	S450L
14722_8#19_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	A	G	gac	2	gGc	D120G
								C	T	cac	1	Tac	H445Y
14722_8#2_H37Rv	LAM	C	A	cgc	1	Agc	R741S	C	G	cac	1	Gac	H445D

Contd...

Supplementary Table 2: Contd...

Strain Number	Spoligotyping	Mutations in rpoC						Mutations in rpoB					
		Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change	Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change
14722_8#20_H37Rv	LAM	C	T	ccg	2	cTg	P1040L	C	T	tcg	2	tTg	S450L
14722_8#21_H37Rv	LAM	T	G	gtg	2	gGg	V483G	C	T	tcg	2	tTg	S450L
14722_8#22_H37Rv	LAM	C	A	aac	3	aaA	N698K	C	T	tcg	2	tTg	S450L
14722_8#23_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14722_8#24_H37Rv	LAM	A	G	aac	2	aGc	N698S	C	T	tcg	2	tTg	S450L
14722_8#27_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14722_8#3_H37Rv	LAM	G	T	ggc	2	gTc	G945V	C	T	tcg	2	tTg	S450L
14722_8#31_H37Rv	LAM	A	G	aac	2	aGc	N698S	C	T	tcg	2	tTg	S450L
14722_8#32_H37Rv	LAM	C	T	acc	2	aTc	T812I	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_8#36_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	A	T	cac	2	cTc	H445L
14722_8#37_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
	Haarlem	G	A	ggg	1	Agg	G571R						
14722_8#38_H37Rv	Haarlem	C	T	gcc	2	gTc	A734V	C	T	tcg	2	tTg	S450L
	Haarlem	G	A	ggg	2	gAg	G594E						
14722_8#39_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14722_8#4_H37Rv	LAM	C	T	acc	2	aTc	T812I	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_8#41_H37Rv	LAM	G	T	gtg	1	Ttg	V517L	C	T	tcg	2	tTg	S450L
14722_8#44_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	G	cac	1	Gac	H445D
14722_8#45_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
14722_8#50_H37Rv	LAM	G	A	gtg	1	Atg	V1252M	C	T	tcg	2	tTg	S450L
14722_8#53_H37Rv	LAM	G	T	gtg	1	Ttg	V517L	C	T	tcg	2	tTg	S450L
14722_8#55_H37Rv	Beijing	G	C	tgg	2	tCg	W484S	C	T	tcg	2	tTg	S450L
14722_8#56_H37Rv	LAM	A	G	atc	1	Gtc	I491V	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_8#57_H37Rv	LAM	T	G	gtg	2	gGg	V483G	C	T	tcg	2	tTg	S450L
14722_8#58_H37Rv	LAM	C	G	ccg	2	cGg	P1040R	C	T	tcg	2	tTg	S450L
14722_8#59_H37Rv	LAM	G	A	gtg	1	Atg	V1252M	C	T	tcg	2	tTg	S450L
14722_8#6_H37Rv	LAM	T	A	gtg	2	gAg	V1124E	C	T	tcg	2	tTg	S450L
14722_8#66_H37Rv	LAM	A	C	gag	2	gCg	E49A	A	T	gac	2	gTc	D435V
14722_8#69_H37Rv	LAM	A	G	aac	2	aGc	N698S	C	T	tcg	2	tTg	S450L
14722_8#70_H37Rv	LAM	G	C	gtg	1	Ctg	V1252L	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_8#75_H37Rv	LAM	G	A	gtg	1	Atg	V1252M	C	T	tcg	2	tTg	S450L
14722_8#76_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	A	T	gac	2	gTc	D435V

Contd...

Supplementary Table 2: Contd...

Strain Number	Spoligotyping	Mutations in rpoC						Mutations in rpoB					
		Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change	Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change
14722_8#77_H37Rv	LAM	C	T	acc	2	aTc	T812I	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_8#79_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	A	T	gac	2	gTc	D435V
14722_8#8_H37Rv	LAM	A	G	aac	2	aGc	N698S	C	T	tcg	2	tTg	S450L
14722_8#80_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14722_8#81_H37Rv	T	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L
14722_8#86_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
	Other small clades	T	G	gtg	2	gGg	V483G						
14722_8#87_H37Rv	LAM	C	A	cgc	1	Agc	R741S	C	T	tcg	2	tTg	S450L
14722_8#88_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450F
14722_8#9_H37Rv	LAM	C	T	acc	2	aTc	T812I	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_8#90_H37Rv	LAM	A	C	gaa	2	gCa	E757A	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14722_8#93_H37Rv	Haarlem	C	T	gcc	2	gTc	A734V	C	T	tcg	2	tTg	S450L
	Haarlem	G	A	ggg	2	gAg	G594E						
14722_8#95_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	T	G	tat	1	Gat	Y308D
								C	T	tcg	2	tTg	S450L
14892_2#1_H37Rv	LAM	G	C	gtg	1	Ctg	V517L	C	T	tcg	2	tTg	S450L
14892_2#27_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	T,C	A,T	tcg	1,2	ATG	S450M
14892_2#29_H37Rv	LAM	A	G	aac	2	aGc	N698S	C	T	tcg	2	tTg	S450L
14892_2#35_H37Rv	LAM	T	C	gtg	2	gCg	V483A	A	C	caa	2	cCa	Q432P
14892_2#39_H37Rv	LAM	C	A	cgc	1	Agc	R741S	C	G	cac	1	Gac	H445D
	LAM	G	C	ggc	2	gCc	G388A						
14892_2#41_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	A	T	gac	2	gTc	D435V
14892_2#43_H37Rv	LAM	C	A	cgc	1	Agc	R741S	C	G	cac	1	Gac	H445D
	LAM	G	C	ggc	2	gCc	G388A						
14892_2#45_H37Rv	LAM	A	G	atc	1	Gtc	I997V	C	T	tcg	2	tTg	S450L
	LAM	A	G	atc	1	Gtc	I851V						
14892_2#58_H37Rv	LAM	A	G	atc	1	Gtc	I491V	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T

Contd...

Supplementary Table 2: Contd...

Strain Number	Spoligotyping	Mutations in rpoC						Mutations in rpoB					
		Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change	Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change
14892_2#59_H37Rv	T	T	G	gtg	2	gGg	V483G	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L
14892_2#63_H37Rv	Beijing	G	C	tgg	2	tCg	W484S	C	T	tcg	2	tTg	S450L
14892_2#65_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450F
14892_2#66_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14892_2#67_H37Rv	LAM	G	T	gtg	1	Ttg	V517L	C	T	tcg	2	tTg	S450L
14892_2#69_H37Rv	LAM	C	T	ccg	2	cTg	P615L	C	T	tcg	2	tTg	S450L
14892_2#70_H37Rv	LAM	G	T	gtg	1	Ttg	V517L	C	T	tcg	2	tTg	S450L
14892_2#71_H37Rv	LAM	G	A	gtg	1	Atg	V1252M	C	T	tcg	2	tTg	S450L
14892_2#74_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	ccg	1	Tcg	P45S
								C	T	tcg	2	tTg	S450L
14892_2#76_H37Rv	LAM	G	A	gag	1	Aag	E1033K	C	G	ttc	3	ttG	F424L
								A	T	atc	1	Ttc	I491F
14892_2#80_H37Rv	LAM	G	A	gtg	1	Atg	V1252M	C	T	tcg	2	tTg	S450L
14892_2#82_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
14892_2#85_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
	Haarlem	T	C	ctg	2	cCg	L516P						
14892_2#88_H37Rv	T	T	G	gtg	2	gGg	V483G	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L
14892_2#89_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	G	T	gtc	1	Ttc	V170F
	Other small clades	T	A	tcg	1	Acg	S428T	A	C	atc	1	Ctc	I488L
14892_2#9_H37Rv	LAM	C	A	cgc	1	Agc	R741S	C	T	tcg	2	tTg	S450F
14892_2#90_H37Rv	Haarlem	C	T	gcc	2	gTc	A734V	C	T	tcg	2	tTg	S450L
	Haarlem	G	A	ggg	2	gAg	G594E						
14892_2#91_H37Rv	LAM	T	G	gtg	2	gGg	V483G	C	T	tcg	2	tTg	S450L
14892_2#94_H37Rv	Haarlem	C	T	gcc	2	gTc	A734V	C	T	tcg	2	tTg	S450L
	Haarlem	G	A	ggg	2	gAg	G594E						
14893_2#12_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
14893_2#22_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	A	T	atc	1	Ttc	I491F
14893_2#23_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14893_2#30_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	G	T	gtc	1	Ttc	V170F
	Other small clades	G	A	ggg	1	Agg	G571R						

Contd...

Supplementary Table 2: Contd...

Strain Number	Spoligotyping	Mutations in rpoC						Mutations in rpoB					
		Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change	Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change
14893_2#36_H37Rv	LAM	C	T	acc	2	aTc	T812I	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14893_2#37_H37Rv	LAM	A	G	aac	2	aGc	N698S	C	T	tcg	2	tTg	S450L
	LAM	C	A	cgc	1	Agc	R741S						
14893_2#38_H37Rv	LAM	G	T	gtg	1	Ttg	V517L	C	T	tcg	2	tTg	S450L
14893_2#40_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14893_2#43_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
								C	T	cgc	1	Tgc	R827C
14893_2#45_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14893_2#46_H37Rv	LAM	A	G	atc	1	Gtc	I491V	C	T	tcg	2	tTg	S450L
								G	A	gcc	1	Acc	A692T
14893_2#47_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
								G	A	ggc	2	gAc	G973D
14893_2#49_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	G	cac	1	Gac	H445D
14893_2#52_H37Rv	T	T	G	gtg	2	gGg	V483G	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L
14893_2#54_H37Rv	Beijing	C	T	gcg	2	gTg	A788V	G,A	T,T	gac	1,2	TTC	D435F
14893_2#6_H37Rv	Haarlem	G	T	gtg	1	Ttg	V1252L	C	T	tcg	2	tTg	S450L
	Haarlem	G	A	ggg	2	gAg	G594E						
14893_2#63_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	A	T	gac	2	gTc	D435V
14893_2#64_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	A	G	gag	2	gGg	E250G
14893_2#71_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
14893_2#72_H37Rv	T	T	G	gtg	2	gGg	V483G	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L
14893_2#74_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	G	tcg	2	tGg	S450W
	Haarlem	T	G	gtg	2	gGg	V1206G						
	Haarlem	G	A	ggg	2	gAg	D1218A						
14893_2#75_H37Rv	LAM	G	A	gtg	1	Atg	V1252M	C	T	tcg	2	tTg	S450L
14893_2#79_H37Rv	LAM	G	T	gtg	1	Ttg	V517L	C	T	tcg	2	tTg	S450L
14893_2#8_H37Rv	LAM	C	G	ccg	1	Gcg	P1040A	C	T	tcg	2	tTg	S450L
14893_2#80_H37Rv	LAM	G	C	ggc	1	Cgc	G332R	C	T	tcg	2	tTg	S450L
14893_2#84_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	A	T	gac	2	gTc	D435V
14893_2#86_H37Rv	Beijing	G	C	tgg	2	tCg	W484S	C	T	tcg	2	tTg	S450L
14893_2#87_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
	Other small clades	T	C	gtg	2	gCg	V483A						

Contd...

Supplementary Table 2: Contd...

Strain Number	Spoligotyping	Mutations in rpoC						Mutations in rpoB					
		Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change	Original base	Mutated base	Reference codon	Codon position	Mutations codon	aa change
14893_2#89_H37Rv	Beijing	G	C	tgg	2	tCg	W484S	C	T	tcg	2	tTg	S450L
14893_2#9_H37Rv	T	T	G	ttg	1	Gtg	L507V	C	T	tcg	2	tTg	S450L
14893_2#90_H37Rv	Beijing	T	C	gtg	2	gCg	V483A	C	T	tcg	2	tTg	S450L
14893_2#92_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	G	agc	3	agG	S428R
								A	C	cac	2	cCc	H445P
14893_2#93_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
	Other small clades	T	C	gtg	2	gCg	V483A						
14893_2#94_H37Rv	Haarlem	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
	Haarlem	G	A	ggg	1	Agg	G571R						
14893_2#95_H37Rv	LAM	T	G	gtg	2	gGg	V483G	C	T	tcg	2	tTg	S450L
14893_2#85_H37Rv	Other small clades	G	A	ggg	2	gAg	G594E	C	T	tcg	2	tTg	S450L
14893_2#88_H37Rv	T	T	G	gtg	2	gGg	V483G	C	T	tcg	2	tTg	S450L
								G	C	gtg	1	Ctg	V695L

Supplementary Table 3: Putative rpoC compensatory mutations

Mutation	aa change	Spoligotyping	Reported
aac - aaA	N698K	LAM (2)	Yes
atc - Gtc	I491V	Haarlem (1), LAM (6), other small clades (1)	Yes
cac - caG	H525Q	Beijing	Yes
ccg - cGg	P1040R	Beijing, LAM (2)	Yes
ctg - cCg	L516P	Haarlem (2)	Yes
gtg - Atg	V1252M	LAM (8)	Yes
gtg - gCg	V483A	Beijing (1), LAM (2), T (1), other small clades (2)	Yes
gtg - gGg	V483G	T (5), other small clades (2), LAM (4)	Yes
gtg - Ttg	V1252L	T, LAM, Haarlem	Yes
gtg - Ttg	V517L	LAM (10), NA (1)	No
tcg - Acg	S428T	LAM (11), NA (1), other small clades (1)	No
tgg - tCg	W484S	Beijing (6)	No
ttc - Ctc	F831L	Other small clades (1)	No
ttg - Gtg	L507V	T (1)	No
	L527V	T (1), other small clades (1)	No
aac - aGc	N698S	LAM (6)	No
acc - aTc	T812I	LAM (7)	No
atc - Gtc	I851V	LAM (1)	No
	I997V	LAM (1)	No
	I885V	LAM (1)	No
cag - Gag	Q523E	Beijing (1)	No
ccg - cAg	P1040Q	T (1)	No
ccg - cTg	P1040L	LAM (1)	No
	P615L	LAM (1)	No
ccg - Gcg	P1040A	LAM (14), T (1)	No
cgc - Agc	R741S	LAM (8), NA (1), T (1)	No
ctc - cGc	L1245R	Beijing (1)	No
ctg - cAg	L516Q	LAM (1)	No
gaa - gCa	E757A	LAM (1)	No
gaa - gGa	E488G	Haarlem (1)	No
gag - Aag	E1033K	LAM (1)	No
gag - gCg	E49A	LAM (1)	No
gag - gGg	E744G	Other small clades (1)	No
gcc - Ccc	A1047P	LAM (1)	No
gcc - gTc	A734V	Haarlem (4)	No
gcg - gTg	A788V	Beijing (1)	No
ggc - Cgc	G332R	T (3), LAM (1)	No
ggc - gCc	G388A	LAM (2)	No
ggc - gTc	G945V	LAM (1)	No
ggg - Agg	G571R	Haarlem (3), other small clades (1)	No
ggg - gGg	V1206G	Haarlem (1)	No
gtg - gAg	V1124E	LAM (1)	No
gtg - gCg	V1147A	Other small clades (1)	No
gtg - gGg	D1218A	Haarlem (1)	No

Numbers in brackets indicate the number of isolates in each strain