



UNIVERSIDAD PERUANA  
**CAYETANO HEREDIA**

POTENCIALES MUTACIONES  
COMPENSATORIAS ASOCIADAS  
A LA RESISTENCIA A  
RIFAMPICINA EN *Mycobacterium  
tuberculosis*

TESIS PARA OPTAR EL GRADO DE  
MAESTRO EN INFORMÁTICA BIOMÉDICA  
EN SALUD GLOBAL CON MENCIÓN EN  
BIOINFORMÁTICA

JHONATAN RABANAL SANCHEZ

LIMA-PERÚ

2020



**Asesor de tesis**

PhD. Mirko Zimic Peralta

**JURADO DE TESIS**

DR. JESUS LORENZO CHIRINOS CACERES  
PRESIDENTE

DR. JORGE ENRIQUE RODRIGUEZ BAILON  
VOCAL

MG. JUAN CARLOS AGAPITO PANTA  
SECRETARIO

## **Agradecimientos**

Agradezco a mis padres por todo el apoyo brindado en el tiempo que estuve en la maestría, gracias a ellos pude terminar exitosamente esta tesis. Asimismo, agradezco a mi asesor de tesis y mentor, el Dr. Mirko Zimic, por permitirme realizar esta tesis en su laboratorio. Gracias a su enorme ayuda en la concepción, en el desarrollo y finalización de este estudio pude finalizar este estudio de manera exitosa. Además, agradezco a todos los integrantes del laboratorio de Bioinformática y Biología molecular de los Laboratorios de Investigación y Desarrollo (LID) por su apoyo en todo momento.

Por último, agradezco a mis amigos de la maestría y a todos los profesores y personal administrativo de la facultad de Salud Pública y Administración (FASPA) por su constante apoyo en mi desarrollo profesional.

## **FUENTES DE FINANCIAMIENTO**

Este estudio fue realizado con un aporte de FONDECYT-CONCYTEC

# Índice

<b>I.</b>	<b>Introducción</b>	1
<b>II.</b>	<b>Planteamiento de la investigación</b>	4
II.1	Planteamiento del problema	4
II.2	Marco teórico	6
II.2.1	<i>Mycobacterium tuberculosis</i>	6
II.2.2	Tuberculosis	7
II.2.3	Resistencia a rifampicina	10
II.2.4	Mutaciones compensatorias	11
II.2.5	PonA1	13
II.3	Justificación	14
II.4	Objetivos	15
II.4.1	Objetivo específicos	15
II.5	Hipótesis	15
<b>III.</b>	<b>Metodología</b>	16
III.1	Diseño de estudio	16
III.2	Población y muestra	16
III.3	Procedimiento de estudio	17
<b>IV.</b>	<b>Resultados</b>	21
<b>V.</b>	<b>Discusión</b>	47
<b>VI.</b>	<b>Conclusiones</b>	54
<b>VII.</b>	<b>Recomendaciones</b>	55
<b>VIII.</b>	<b>Referencias bibliográficas</b>	56

## Índice de tablas

Tabla 1. Antibióticos antituberculosos y sus mecanismos de acción.....	9
Tabla 2. Clasificación de los genomas teniendo en cuenta la presencia de al menos un SNP en un gen en particular.....	17
Tabla 3. Clasificación de los genomas teniendo en cuenta la presencia de al menos un SNP en el gen rpoB.....	18
Tabla 4. Clasificación de genomas teniendo en cuenta la presencia un SNP en específico.....	18
Tabla 5. Clasificación de genomas teniendo en cuenta la presencia un SNP en específico.....	19
Tabla 6. Clasificación de genomas teniendo en cuenta la presencia de al menos un SNP en el gen PonA1.....	20
Tabla 7. Clasificación de genomas teniendo en cuenta la presencia de un SNP en específico el gen PonA1.....	21
Tabla 8. Frecuencia de los genes asociados a la resistencia a rifampicina que presentaron un OR significativo.....	22
Tabla 9. Odds ratios significativos de genes asociados a la resistencia a rifampicina.....	23
Tabla 10. Frecuencia de los genes asociados a la presencia de al menos un SNP en el gen rpoB que presentaron un OR significativo.....	24
Tabla 11. OR significativos de genes asociados a la presencia de al menos un SNP en el gen rpoB.....	25
Tabla 12. Frecuencia los SNPs considerando los genomas resistentes a rifampicina como grupo expuesto.....	26
Tabla 13. OR e Índice de Youden de los SNPs de los genes asociados a la resistencia a rifampicina.....	26
Tabla 14. Frecuencia de SNPs considerando la presencia de SNPs en el gen rpoB como grupo expuesto.....	26
Tabla 15. OR e Índice de Youden de los SNPs de los genes asociados a la presencia de al menos un SNP en el gen rpoB.....	28
Tabla 16. Frecuencia de los genes asociados a la resistencia a rifampicina que presentaron un OR significativo.....	30
Tabla 17. OR e índice de Youden de genes asociados a la resistencia a rifampicina.....	30
Tabla 18. Frecuencia de los genes asociados a la presencia de al menos un SNP en el gen rpoB que presentaron un OR significativo.....	31
Tabla 19. OR e Índice de Youden de genes asociados a la presencia de SNPs en el gen rpoB.....	32
Tabla 20. Frecuencia de los SNPs considerando la resistencia a rifampicina en el grupo expuesto.....	30
Tabla 21. OR e índice de Youden de SNPs asociados a la resistencia a rifampicina.....	34
Tabla 22. Frecuencia de SNPs considerando la presencia de SNPs en el gen rpoB como grupo expuesto.....	35
Tabla 23. OR e Índice de Youden de genes asociados a la presencia de SNPs en el gen rpoB.....	37
Tabla 24. Pruebas estadísticas de los genes asociados a la resistencia a rifampicina.....	40

Tabla 25. Pruebas estadísticas de los genes asociados a la presencia de SNPs en el gen rpoB.....	41
Tabla 26. Pruebas estadísticas de los SNPs de los genes asociados a la presencia de al menos un SNP en el gen rpoB.....	42
Tabla 27. Clasificación de los genomas de las 3 bases de datos según la presencia o ausencia de al menos un SNP en el gen ponA1.....	43
Tabla 28. Clasificación de los genomas totales según la resistencia a rifampicina y la presencia de al menos un SNP en el gen ponA1.....	43
Tabla 29. OR, índice de Youden y comparación de proporciones de la asociación entre la presencia de al menos un SNP en el gen ponA1 y la resistencia a rifampicina.....	44
Tabla 30. Prueba de Chi-cuadrado.....	44
Tabla 31. Frecuencia de SNPs en el gen ponA1 según la presencia o ausencia de resistencia a rifampicina.....	45
Tabla 32. OR, índice de Youden y comparación de proporciones de la asociación entre la presencia de al menos un SNP en el gen ponA1 y la resistencia a rifampicina de los SNPs del gen ponA1.....	45
Tabla 33. Prueba de Chi-cuadrado de los SNPs.....	46

## Índice de figuras

Figura 1. Estructura química de la rifampicina.....	10
Figura 2. Puentes de hidrógeno (líneas punteadas) entre los residuos de la subunidad $\beta$ del ARN polimerasa y los átomos de oxígeno de rifampicina.....	11

## RESUMEN

Aproximadamente más del 95% de los casos de resistencia a rifampicina en *Mycobacterium tuberculosis* se deben a mutaciones en el gen rpoB. Dichas mutaciones producen efectos nocivos en las bacterias (disminuyen su fitness). Las cepas resistentes a rifampicina buscan corregir la pérdida del fitness mediante mutaciones compensatorias. El presente estudio analizó e identificó polimorfismos de un solo nucleótido (SNPs), en bases de datos de genomas de *M. tuberculosis*. En una primera etapa se analizaron todos los genes con al menos un SNP de una base de datos de 72 genomas, posteriormente se analizaron 200 genomas. Se unificaron ambas bases de datos y se calculó el odds ratio (OR), índice de Youden, razón de prevalencias y el test exacto de Fisher. Por último, se analizaron 914 genomas, para identificar la asociación entre la presencia de SNPs en el gen ponA1 y la resistencia a rifampicina. Se calculó el OR, índice de Youden, prueba de Chi-cuadrado y razón de prevalencias. Se identificaron 6 genes (dnaE, esxP, mbtB, PPE64, Rv1429 y Rv2052c) asociados estadísticamente con la presencia de SNPs en el gen rpoB (valor de  $p$  del test exacto de Fisher  $< 0.05$ ). Se encontró que los SNPs P618R (gen dnaE), T3S (gen esxP), S1000L (gen mbtB), N196Y (gen PPE64), A500T (gen PPE64), A429T (gen Rv 2052c) y R135L (gen Rv 1429) son factores asociados a la presencia de al menos un SNP en el gen rpoB. Asimismo, la presencia de al menos un SNP en el gen ponA1 es un factor asociado a la resistencia a rifampicina, al igual que los SNPs P631S y A516T de este gen.

**Palabras claves:** Tuberculosis, rifampicina, mutaciones compensatorias, resistencia antimicrobiana.

## **ABSTRACT**

Approximately more than 95% of cases of rifampicin resistance in *Mycobacterium tuberculosis* are due to mutations in the *rpoB* gene. These mutations produce harmful effects on bacteria (they decrease their fitness). Rifampicin-resistant strains seek to correct loss of fitness through compensatory mutations. The present study analyzed and identified single nucleotide polymorphisms (SNPs) in *M. tuberculosis* genome databases. In a first stage, all genes with at least one SNP from a database of 72 genomes were analyzed, later 200 genomes were analyzed. Both databases were unified and the odds ratio (OR), Youden index, prevalence ratio and Fisher's exact test were calculated. Finally, 914 genomes were analyzed to identify the association between the presence of SNPs in the *ponA1* gene and rifampicin resistance. The OR, Youden index, Chi-square test and prevalence ratio were calculated. Six genes were identified (*dnaE*, *esxP*, *mbtB*, *PPE64*, *Rv1429* and *Rv2052c*) statistically associated with the presence of SNPs in the *rpoB* gene ( $p$  value of Fisher's exact test  $<0.05$ ). SNPs P618R (*dnaE* gene), T3S (*esxP* gene), S1000L (*mbtB* gene), N196Y (*PPE64* gene), A500T (*PPE64* gene), A429T (*Rv 2052c* gene) and R135L (*Rv 1429* gene) were found to be factors. associated with the presence of at least one SNP in the *rpoB* gene. Likewise, the presence of at least one SNP in the *ponA1* gene is a factor associated with rifampicin resistance, as are the SNPs P631S and A516T of this gene.

**Keywords:** Tuberculosis, rifampicin, compensatory mutations, antimicrobial resistance.

## I. INTRODUCCIÓN

*Mycobacterium tuberculosis* es una bacteria aerobia de crecimiento lento, causante de la tuberculosis (TB) en humanos (1); a su vez esta enfermedad es la primera causa de muerte, en el mundo, por un agente bacteriano (2). En el Perú, la tuberculosis es la décimo quinta causa de muerte; y representa una carga para la sociedad, ya que es la vigésimo séptima enfermedad que genera más años de vida saludable perdidos (3).

En el 2018 la incidencia anual de tuberculosis era de 86.7 casos por 100 mil habitantes, en el Perú (4). Asimismo, en el 2014, la Organización de las Naciones Unidas (ONU), estableció como meta para el 2035 una incidencia anual máxima de 10 casos de TB por 100 mil habitantes. Por lo que se deben implementar estrategias de prevención y protocolos de tratamiento más eficientes (3).

La resistencia antimicrobiana, en *M. tuberculosis*, representa un serio problema de salud global; pues esta bacteria puede adquirir resistencia a uno o más fármacos. Esto genera un costo social alto y obliga a la sociedad a buscar nuevos fármacos que sean capaces de eliminar esta enfermedad (2)(3).

Entre los antibióticos de primera línea más usados, para el tratamiento de TB, tenemos a rifampicina, isoniacida y etambutol (2). En las últimas décadas se han reportado miles de casos de *M. tuberculosis* resistente a rifampicina e isoniacida, los cuales han sido catalogados como TB multidrogorresistente (TB-MDR). En el 2017 se reportaron, en el mundo, un total de 160 684 casos de TB-MDR y TB resistente solo a rifampicina, lo cual representa un aumento del 4.94% (153 119 casos en el 2016) (3).

La resistencia a rifampicina se debe a polimorfismos de un solo nucleótido (SNPs) en el gen *rpoB*, que inhiben la unión de la rifampicina al ARN polimerasa (5)(6)(7). Estas mutaciones reducen el fitness de estos microorganismos, lo cual genera una desventaja evolutiva (8)(9). En estas condiciones se producen mutaciones en otras regiones del ADN que compensan la reducción del fitness, se utiliza el termino mutaciones compensatorias para referirse a ellas (10) (11).

Las mutaciones en el gen *rpoB* están asociados a cambios en la composición de lípidos claves de la pared celular bacteriana (12). La mayoría de los cambios fueron reportados en lípidos asociados a factores de virulencia y que son importantes en la respuesta inmunológica en individuos con TB (12) (13) (14) (15).

Diversos estudios han encontrado principalmente mutaciones compensatorias en los genes *rpoC* y *rpoA* asociados a la resistencia a rifampicina. Estas mutaciones generan condiciones que contrarrestan los efectos nocivos de las mutaciones en el gen *rpoB*, en *M. tuberculosis* (10) (11). Sin embargo, no se ha evaluado el potencial de otras mutaciones, en otros genes, de incrementar el fitness; por lo que estas hipotéticas mutaciones serían potenciales mutaciones compensatorias.

Un potencial candidato a poseer mutaciones compensatorias es el gen *PonA1* que codifica una proteína de unión a la penicilina (*PonA1*); la cual posee actividad transglicosilasa y transpeptidasa y participa en la síntesis del peptidoglicano. En *M. tuberculosis* esta enzima posee dos dominios, el dominio transglicosilasa (residuos 37-218) y el dominio transpeptidasa (residuos 310-600) (16). Un estudio en *M. tuberculosis* encontró que el SNP Q365H le confería mayor capacidad para sobrevivir en moderadas concentraciones de rifampicina (17). Por otro lado, otro estudio reportó resultados similares en cepas con el SNP

T34D (18). Por lo tanto, algunos SNPs en PonA1 pueden alterar la susceptibilidad a rifampicina en cepas de *M. tuberculosis* (18) (19).

Nuestro estudio analizó tres bases de datos de genomas de *M. tuberculosis* e identificó potenciales mutaciones compensatorias asociadas a la resistencia a rifampicina o a la presencia de SNPs en el gen *rpoB*.

## II. PLANTEAMIENTO DE LA INVESTIGACIÓN

### II.1 Planteamiento del problema

La rifampicina es un antibiótico de primera línea en el tratamiento de la tuberculosis. Este antibiótico posee actividad bactericida, ya que inhibe la transcripción del ADN. Debido a que el mecanismo de acción de la rifampicina se basa en la unión de ésta a la subunidad  $\beta$  del ARN polimerasa (5)(6)(7) (20).

Las mutaciones en el gen *rpoB* que confieren resistencia a rifampicina, usualmente se asocian a la pérdida del fitness celular, dado que dichas mutaciones comprometen el mecanismo de transcripción y con ello el correcto balance de la expresión de proteínas y subsecuentemente el fenotipo celular (8)(9) (21). Por ejemplo, un estudio en *Neisseria meningitidis*, resistente a rifampicina, encontró alteraciones en la expresión de proteínas involucradas en el metabolismo de piruvato, plegamiento de proteínas y del ácido tricarbóxico. Estas alteraciones serían consecuencia de las mutaciones en el gen *rpoB* (22).

Un estudio en mutantes de *M. tuberculosis* demostró que un SNP en *rpoB* (H526D) genera alteraciones en la fisiología de la pared celular, aumento de la permeabilidad y reducción de la resistencia a los cambios externos (12). Estudios de metabolómica y transcriptómica, en *M. tuberculosis* resistente a rifampicina, revelaron alteraciones en la expresión de genes implicados en la síntesis de dimicocerosato de ftiocerol y otros lípidos (13). En este sentido, se han reportado cambios en la composición de los lípidos de la pared celular; entre el 2.7% a 6.2% de los lípidos detectables sufrieron cambios en su composición; por ejemplo, se reportaron alteraciones en las concentraciones de sulfoglicolípidos acilados y sideróforos de micobactina (14).

Asimismo, otro estudio en metabolómica reveló una disminución en la síntesis de ácidos grasos de cadena ramificada (15). Por lo tanto, las mutaciones en el gen *rpoB*, asociadas a la resistencia a rifampicina, producen cambios en la composición de los lípidos de la pared celular de *M. tuberculosis*. Asimismo, se reportó la sobreexpresión de los genes que codifican las proteínas *ppsA*, *ppsB* y *ppsC* (12).

La pérdida del fitness, asociada a mutaciones en el gen *rpoB*, es compensado por mutaciones en otros genes (mutaciones compensatorias). Se han reportado mutaciones compensatorias en los genes *rpoC* y *rpoA* en diversas partes del mundo. Siendo las mutaciones en *rpoC* las más comunes (23) (24).

Un estudio en *Mycobacterium leprae* encontró mutaciones con potencial acción compensatoria en los genes *gltB*, *gltD*, *esxB*, *mmpL7* y *rpoC*, en cepas resistentes a rifampicina (25).

Se han publicado una serie de estudios que demuestran el papel de *PonA1* en los procesos de crecimiento y morfología celular (16) (26). Se ha propuesto que determinados SNPs generan alteraciones en la susceptibilidad a rifampicina. Sin embargo, el mecanismo exacto se desconoce (18) (19). Sin embargo, ningún estudio ha evaluado la asociación entre la presencia de SNPs en *ponA1* y la resistencia a rifampicina. Por lo que encontrar una asociación estadística sería un primer paso para comprender los mecanismos moleculares que explicarían los resultados obtenidos en estudios previos.

Este estudio planteó la siguiente pregunta de investigación:

¿Cuál es la asociación entre la presencia de SNPs en genes y la resistencia a rifampicina o la presencia de al menos un SNP en el gen *rpoB*?

## II.2 Marco teórico

### II.2.1 *Mycobacterium tuberculosis*

*Mycobacterium tuberculosis* es un bacilo quimioorganótrofo aerobio de crecimiento lento, además no forma esporas y no presenta movilidad. Puede visualizarse mediante una coloración Ziehl-Neelsen, ya que este bacilo posee una pared celular rica en ácidos micólicos (1).

La tuberculosis es una enfermedad infecciosa causada por *Mycobacterium tuberculosis complex*, término utilizado para referirse a las especies del género *Mycobacterium* causantes de tuberculosis. Entre los miembros de este grupo tenemos: *Mycobacterium tuberculosis* (agente etiológico de la tuberculosis en humanos), *M. africanum* (causante de tuberculosis en algunas regiones de África), *M. bovis*, *M. caprae*, *M. pinnipedii* (causantes de tuberculosis en mamíferos) y *M. microtti* (causante de tuberculosis en ratones) (27).

El genoma de *M. tuberculosis* H37Rv, el cual fue el primer genoma de *Mycobacterium tuberculosis complex* en secuenciarse, tiene 4.4 Mb y un porcentaje de 65.9% de GC. Se encontró que este genoma codifica 4018 genes, los cuales en su mayoría codifican proteínas involucradas en el metabolismo de lípidos, en la estructura de la pared celular, en la supervivencia como parásito intracelular y en la evasión de la respuesta inmune (1).

La pared celular de *M. tuberculosis* está formada básicamente por una capa de peptidoglicano, la cual está unida a una capa de arabinogalactano (heteropolisacárido de arabinosa y galactosa). Asimismo, esta capa de arabinogalactano está esterificada con  $\beta$ -hidroxiácidos grasos de cadena larga (ácidos micólicos). A su vez estos ácidos micólicos se encuentran esterificados con glicerol o trehalosa; en este sentido, la trehalosa podría unirse a

dos ácidos micólicos y formar dimicolil trehalosa. Esta última molécula está involucrada en la respuesta inmune; además estimula la inflamación crónica y la formación de granulomas pulmonares. Los sulfolípidos de trehalosa constituyen un importante factor de virulencia, ya que permiten que estos bacilos evadan la acción de la fagolisosoma, cuando han sido captados por el macrófago (28) (29).

El dimicocerosato de ftiocerol es un importante factor de virulencia, presente en la pared celular de *M. tuberculosis*, permite la supervivencia de estos bacilos en los macrófagos anfitriones. Asimismo, otros lípidos importantes, en la pared celular, que funcionan como factores de virulencia son: glicolípidos fenólicos, glicopeptidolípidos y lipooligosacáridos. Además, es importante la presencia de lipoarabinomananos y fosfatidilinositolmanósidos, los cuales se anclan en la membrana celular y se extienden hasta la superficie de la pared celular (28) (29).

Como puede observarse, la mayoría de los lípidos glicosilados presentes en la pared celular de *M. tuberculosis*, desempeñan un papel como factores de virulencia y pueden estimular una respuesta inmune innata (30).

## **II.2.2 Tuberculosis**

La tuberculosis (TB) es una enfermedad infectocontagiosa causada por *Mycobacterium tuberculosis* (2). Esta enfermedad representa la primera causa de muerte, en el mundo, por un agente bacteriano. Se estima que en el 2017 existían alrededor de 10 millones de personas, en el mundo, con tuberculosis. En dicho año la TB causo, en todo el mundo, aproximadamente 1.3 millones de muertes en personas sin VIH y 300 000 muertes en personas con VIH positivo. Se calcula que el 23% de la población humana tienen una

infección latente de TB, por lo que en condiciones de inmunodepresión o inmunodeficiencia podrían desarrollar esta enfermedad (3).

La TB farmacorresistente es un problema latente, ya que cada año se reportan aislamientos de *M. tuberculosis* resistente a diversos fármacos. En el 2017 se reportaron, en todo el mundo, 558 000 casos de TB resistente a rifampicina (TB-RR) y el 85% de estos casos eran resistentes a otros fármacos (3).

*M. tuberculosis* puede causar TB pulmonar y TB extrapulmonar (pleural, ocular, esquelética, gastrointestinal y ganglionar). Aproximadamente el 90% de las personas infectadas con *M. tuberculosis* pueden controlar la TB; en estas condiciones, este bacilo, entra en un estado de latencia. *M. tuberculosis* puede permanecer por años en este estado y no causar daño al hospedero; ya que puede inhibir la respuesta inmune innata y adaptativa (31).

La producción de INF- $\gamma$ , INF- $\alpha$  y antimicrobianos de los macrófagos, son los principales mecanismos de respuesta inmune innata frente a *M. tuberculosis*. Por otro lado, la respuesta inmune adaptativa es principalmente celular, mediada en su mayoría por linfocitos T CD4, linfocitos Th17 y linfocitos T CD8. Estas dos respuestas inmunológicas mantienen al *M. tuberculosis* en un estado de latencia (31).

El pulmón es el principal órgano afectado por la TB, esto debido al modo de transmisión de dicha enfermedad. La TB pulmonar puede producir una remodelación de la arquitectura del pulmón, lo cual generalmente ocasiona diversos signos y síntomas (tos, esputo mucopurulento o purulento, hemoptisis, malestar, fiebre y pérdida de peso). Generalmente el diagnóstico se realiza mediante radiografía de tórax y análisis de esputo (32) (33).

En la tabla 1 puede observarse los antibióticos utilizados para el tratamiento de TB; asimismo, esta tabla detalla el mecanismo de acción de cada antibiótico y los genes involucrados en la resistencia a dichos antibióticos (34).

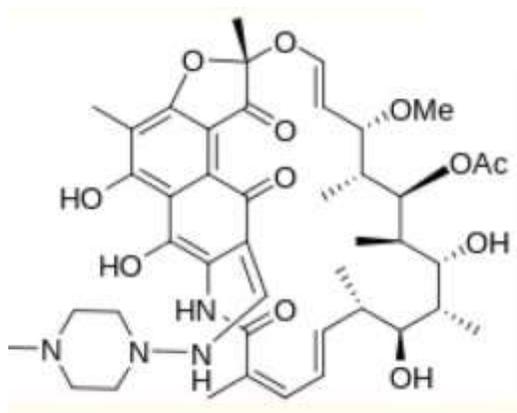
**Tabla 1.** Antibióticos antituberculosos y sus mecanismos de acción (34)

	Antibiótico	Mecanismo de acción	de	Genes involucrados en la resistencia
Primera Línea	Isoniacida	Inhíbe la síntesis de ácidos micólicos y otras vías metabólicas		KatG, ndh, inhA, ahpC
	Rifampicina	Inhíbe la transcripción del ADN	la	rpoB
	Etambutol	Inhíbe la síntesis de arabinogalactano		embCAB embR
	Pirazinamida	Acidificación del citoplasma, lo cual genera inhibición de la actividad enzimática y la síntesis de proteínas		pncA rspA
	Estreptomina	Inhíbe la síntesis de proteínas		rpsL gidB, rrs
Segunda Línea	Amikacina/Kanamicina	Inhíbe la síntesis de proteínas		rrs, eis
	Etionamida	Inhíbe la síntesis de ácidos micólicos		etA, inhA, ndh, etR, mshA
	Fluoroquinolonas	Inhíbe la ADN girasa		gyrA, gyrB

En el Perú, el Ministerio de Salud (MINSA), utiliza los siguientes términos para referirse a algunos tipos de TB resistente a antibióticos: TB resistente a rifampicina e isoniacida como TB multidrogoresistente (TB MDR) y TB resistente a rifampicina, isoniacida, a algún inyectable de segunda línea y a alguna fluoroquinolona como TB extensamente resistente (TB XDR) (2).

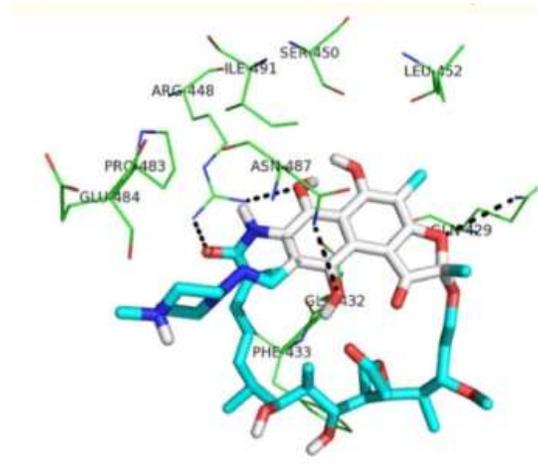
### II.2.3 Resistencia a rifampicina

La rifampicina es un antibiótico bactericida, con actividad antimicrobiana contra bacterias gram positivas y gram negativas. Este antibiótico inhibe la transcripción del ADN, ya que se une a la subunidad  $\beta$  del ARN polimerasa. En la figura 1 puede observarse la estructura química de la rifampicina; en la interacción rifampicina- subunidad  $\beta$  del ARN polimerasa participan átomos de oxígeno, los cuales forman puentes de hidrógeno con los residuos de la subunidad  $\beta$  (figura 2) (5)(6)(7).



**Figura 1.** Estructura química de la rifampicina (6)

La rifampicina puede atravesar membranas biológicas debido a su lipofilidad, por lo que podría acceder fácilmente a la subunidad  $\beta$  del ARN polimerasa y formar interacciones no covalentes con dicha polimerasa, y de esta manera impide el crecimiento de la cadena de ARN mensajero (6).



**Figura 2.** Puentes de hidrógeno (líneas punteadas) entre los residuos de la subunidad  $\beta$  del ARN polimerasa y los átomos de oxígeno de rifampicina (6)

Se han reportado SNPs en una región de 81pb del gen *rpoB*; aproximadamente más del 96% de las cepas resistentes presentan una mutación en esta región. Entre un 40% a 93% de los casos reportados de resistencia a rifampicina se debe al cambio de un aminoácido en la posición 450 (531 en el caso *Escherichia coli*), de serina a leucina (*rpoB* S450L) (6).

## II.2.4 Mutaciones compensatorias

Los mecanismos de resistencia generan carga metabólica y una disrupción de muchos procesos claves en el metabolismo de las bacterias resistentes a antibióticos, estos fenómenos disminuyen el fitness de la bacteria (8)(9) (10).

En *M. tuberculosis* se ha descrito un mecanismo que trata de explicar el proceso mediante el cual determinados SNPs, en el gen *rpoB*, disminuyen el fitness de estas bacterias. Según este modelo, las mutaciones en el gen *rpoB* alterarían el funcionamiento normal del ARN polimerasa generando una alteración en las concentraciones de nucleósidos 5-trifosfato (NTP), los cuales son importantes para el funcionamiento de enzimas como la metionina

adenosiltransferasa. Además, el GTP, un tipo de NTP, es esencial para la síntesis del dinucleótido de flavina (FAD). Asimismo, ambos, el FAD y la metionina adenosiltransferasa, participan en la biosíntesis de los ácidos micólicos (15). Este mecanismo explicaría los resultados de varios estudios que reportan alteraciones en las concentraciones de lípidos como sulfoglicolípidos acilados y sideróforos de micobactina (13) (14).

Diversos estudios han demostrado que las mutaciones compensatorias (MC) son uno de los mecanismos que las bacterias utilizan para aumentar el fitness (8)(9) (10). Estudios *in vivo* e *in vitro* en *E. coli* resistente a estreptomicina, encontraron que el fitness aumentó debido a la presencia de MC (35). En este sentido, estudios en *Salmonella typhimurium* encontraron mutaciones compensatorias, y dichas MC permitirían, a estas bacterias, aumentar el fitness (36). Otros estudios en *M. tuberculosis* (37), *Streptococcus pneumoniae* (38) y *Staphylococcus aureus* (39); sugieren que las MC desempeñan un papel en la evolución de la resistencia antimicrobiana y permiten aumentar el fitness (37) (38) (39).

En este sentido, diversas investigaciones, en *M. tuberculosis* resistente a rifampicina, han encontrado mutaciones compensatorias en los genes *rpoA* y *rpoC*, los cuales codifican las subunidades  $\alpha$  y  $\beta'$  del ARN polimerasa respectivamente (10) (11) (23) (24). Estudios realizados en varios países encontraron mutaciones compensatorias en los genes *rpoA* y *rpoC*; asimismo, más del 90% de estas fueron localizadas en el gen *rpoC*. Las mutaciones en *rpoC* son muy diversas, pero principalmente se da en los codones 483 y 491 (10) (23) (24).

Se ha sugerido que las mutaciones en *rpoC* permiten la interacción de las subunidades  $\beta$ ,  $\beta'$  y  $\alpha$  del ARN polimerasa. Ya que las mutaciones en *rpoB* producen una alteración en estas interacciones (21). Asimismo, estas mutaciones aumentan el fitness en la fosforilación

oxidativa y en el metabolismo de aminoácidos (40). Se ha propuesto que mutaciones en el gen *rpoC* están asociados a una mayor propagación de cepas TB-MDR (11).

### **II.2.5 PonA1**

La proteína de unión a la penicilina de *M. tuberculosis* (PonA1) participa en la biosíntesis del peptidoglicano. Es una proteína transmembrana con dos dominios bien definidos (transglicosilasa y transpeptidasa) (16), el dominio transglicosilasa cataliza la formación de los enlaces  $\beta$  (1 $\rightarrow$ 4) que unen el ácido N-acetilmurámico con la N-acetil-glucosamina y el dominio transpeptidasa cataliza la formación de enlaces entre una D-alanina y un ácido meso-diaminopimérico (enlaces 3 $\rightarrow$ 4), los cuales provienen de una cadena peptídica unida al ácido N-acetilmurámico. También se han descrito enlaces, menos frecuentes, entre dos ácidos meso-diaminopimérico (enlaces 3 $\rightarrow$ 3). Por tanto, de esta manera se ensambla la estructura básica del peptidoglicano (41).

PonA1 desempeña una función clave en el crecimiento y la morfología bacteriana, estudios recientes señalan que PonA1 regula la actividad de RipA mediante su unión al dominio endopeptidasa, en dicha interacción participa la región C-terminal de PonA1, la cual es una región de alrededor de 150 residuos rica en prolina (16) (26).

Las familias de genes PE y PPE, las cuales suelen ser muy polimórficas, codifican proteínas con regiones ricas en prolina-glutamato y prolina-prolina-glutamato respectivamente, generalmente son proteínas excretadas o están localizadas en la membrana celular. Se ha sugerido que estas proteínas participan en la respuesta inmune y pueden actuar como factores de virulencia (42). En este sentido, estudios recientes señalan que el gen *ponA1*, el cual pertenece a la familia PPE, es un importante regulador de la morfología y crecimiento

bacteriano (16) (26). Asimismo, se lo ha descrito como un posible biomarcador de tuberculosis latente (43).

Se ha descrito que un SNP en el dominio transglicosilasa (T34D) puede alterar la concentración inhibitoria mínima (MIC) de rifampicina en *M. tuberculosis* (17), se ha observado resultados similares en el dominio transpeptidasa (Q365H) (18). El knock-out del dominio transglicosilasa incrementa el porcentaje de supervivencia de *M. tuberculosis* en concentraciones moderadas de rifampicina (0.00195 ug/ml); sin embargo, el knock-out del dominio transpeptidasa disminuye ligeramente el porcentaje de supervivencia comparado con las cepas wild type. Por tanto, los cambios en la actividad enzimática de PonA1 están asociados a una alteración de la susceptibilidad de *M. tuberculosis* a rifampicina (17).

Recientemente se ha identificado la sobreexpresión del gen ponA1 en pacientes con tuberculosis latente, por lo que se podrían implementar nuevos métodos de laboratorio para la identificación temprana de tuberculosis mediante la medición de esta proteína en líquidos biológicos (26).

### **II.3 Justificación**

La resistencia a la rifampicina representa un problema de salud global, ya que el número de casos de TB-MDR y TB resistente a rifampicina viene en aumento. Por lo que entender los mecanismos moleculares que llevan a la adquisición de resistencia antimicrobiana en *Mycobacterium tuberculosis*, representa un paso previo para la búsqueda de nuevos enfoques terapéuticos y métodos diagnósticos. Asimismo, la presencia de mutaciones compensatorias podría generar una alteración en la expresión de proteínas, por lo que las concentraciones alteradas de determinadas proteínas o metabolitos podrían ser usados como un método para la detección temprana de tuberculosis resistente a rifampicina.

## **II.4 Objetivos**

-Identificar potenciales mutaciones compensatorias asociadas a la resistencia a rifampicina en *Mycobacterium tuberculosis*.

### **II.4.1 Objetivos específicos**

-Identificar SNPs asociados a la resistencia a rifampicina o a la presencia de SNPs en el gen *rpoB* de *Mycobacterium tuberculosis*.

-Describir la posible asociación entre la presencia de al menos un SNP en genes, distintos a *rpoA* y *rpoC*, y la resistencia a rifampicina.

-Describir la posible asociación entre la presencia de al menos un SNP en genes, distintos a *rpoA* y *rpoC*, y la presencia de al menos un SNP en el gen *rpoB*.

-Describir la posible asociación entre la presencia de SNPs en *ponA1* y la resistencia a rifampicina.

## **II.5 Hipótesis de trabajo**

Existen SNPs en diversos genes asociados a la resistencia a rifampicina, además de los reportados en la literatura, o a la presencia de al menos un SNP en el gen *rpoB*, que podrían ser potenciales mutaciones compensatorias en *Mycobacterium tuberculosis*.

### **III. METODOLOGÍA**

#### **III.1 Diseño de estudio**

Estudio observacional descriptivo de tipo transversal analítico.

#### **III.2 Población y muestra**

Nuestro estudio se basó en el análisis de 3 bases de datos secundarias, las cuales provienen de 3 estudios previos realizados en el Laboratorio de Bioinformática y Biología molecular de los Laboratorios de Investigación y Desarrollo de la Universidad Peruana Cayetano Heredia (LID-UPCH). Estas 3 bases de datos contienen en total 914 genomas de *M. tuberculosis*, los cuales provienen de cepas aisladas en el periodo 2001-2016 de pacientes con tuberculosis de Lima y Callao.

Para la selección de las muestras, estos estudios, utilizaron los siguientes criterios: las muestras fueron aisladas de pacientes con tuberculosis activa y no relacionados, además se utilizó la prueba de sensibilidad a medicamentos observada microscópicamente (MODS) para identificar a las cepas sensibles y resistentes a rifampicina.

La mayoría de los genomas analizados en este estudio presentan polirresistencia microbiana, por lo que son resistentes a muchas drogas de primera línea (incluida la rifampicina). Siendo la mayor parte genomas de cepas MDR.

### III.3 Procedimiento de estudio

#### III.3.1 Análisis de 72 genomas para determinar la asociación entre la presencia de al menos un SNP en genes y la resistencia a rifampicina

Se analizó una base de datos de 72 genomas de *M. tuberculosis* los cuales proceden de muestras tomadas en el periodo 2001-2007. Para lo cual se procedió a realizar el siguiente análisis:

**Tabla 2.** Clasificación de los genomas teniendo en cuenta la presencia de al menos un SNP en un gen en particular

	Con al menos un SNP en un gen "X"	Sin ningún SNP en un gen "X"	Total
Resistente a rifampicina	A	B	A+B
Sensibles a rifampicina	C	D	C+D
Total	A+C	B+D	A+B+C+D

Este análisis se realizó para todos los genes que presenten al menos un SNP de los 72 genomas estudiados. Luego se realizaron las siguientes pruebas estadísticas:

-Odds ratio (OR):  $A(D)/B(C)$

La cual se realizó para hallar la asociación entre resistencia a rifampicina y la presencia de al menos un SNP en un gen en específico (44). Cuando el valor del denominador del OR fue cero, se le sumó 0.5 al A, B, C y D (45).

-Índice de Youden:  $\{[A/(A+C)] + [D/(B+D)] - 1\}$

La cual se realizó para determinar la capacidad que tiene la presencia de al menos un SNP, en un gen en específico, de identificar las cepas resistentes a rifampicina (44).

Además, se procedió a realizar un análisis similar teniendo en cuenta la presencia o ausencia de al menos un SNP en el gen rpoB. Por lo que A+B fueron los genomas con al menos un SNP en rpoB y C+D los genomas sin ningún SNP en rpoB (tabla 3). De igual forma, como en el apartado anterior, se calculó el OR y el índice de Youden. De igual manera, cuando el valor del denominador del OR fue cero, se le sumó 0.5 al A, B, C y D (45).

**Tabla 3.** Clasificación de los genomas teniendo en cuenta la presencia de al menos un SNP en el gen rpoB

	Con al menos un SNP en un gen "X"	Sin ningún SNP en un gen "X"	Total
Con al menos un SNP en el gen rpoB	A	B	A+B
Sin ningún SNP en el gen rpoB	C	D	C+D
Total	A+C	B+D	A+B+C+D

Las pruebas estadísticas se realizaron teniendo en cuenta un nivel de confianza del 95%. Asimismo, para la clasificación y el análisis estadístico se usó el lenguaje de programación R.

### III.3.2 Análisis de 72 genomas para determinar la asociación entre la presencia de SNPs y la resistencia a rifampicina

Se realizó el análisis de los SNPs de los genes cuyos odds ratios presentaron un intervalo de confianza que no incluyó la unidad. Para lo cual se realizó el siguiente análisis:

**Tabla 4.** Clasificación de genomas teniendo en cuenta la presencia un SNP en específico

	Con al menos un SNP "Y" en un gen "X"	Sin ningún SNP "Y" en un gen "X"	Total
Resistente a rifampicina	A	B	A+B
Sensibles a rifampicina	C	D	C+D
Total	A+C	B+D	A+B+C+D

Luego se calculó el OR y el índice de Youden de cada SNP. Además, se procedió a realizar un análisis similar teniendo en cuenta la presencia o ausencia de al menos un SNP en el gen rpoB. Por lo que A+B fueron los genomas con al menos un SNP en el gen rpoB y C+D los genomas sin ningún SNP en el gen rpoB (tabla 5). De igual forma, como en el apartado anterior, se calculó el OR y el índice de Youden. Además, cuando el valor del denominador del OR fue cero, se le sumó 0.5 al A, B, C y D (45).

**Tabla 5.** Clasificación de genomas teniendo en cuenta la presencia un SNP en específico

	Con al menos un SNP “Y” en un gen “X”	Sin ningún SNP “Y” en un gen “X”	Total
Con al menos un SNP en el gen rpoB	A	B	A+B
Sin ningún SNP en el gen rpoB	C	D	C+D
Total	A+C	B+D	A+B+C+D

Las pruebas estadísticas se realizaron teniendo en cuenta un nivel de confianza del 95%. Asimismo, para la clasificación y el análisis estadístico se usó el lenguaje de programación R.

### **III.3.3 Análisis de 200 genomas para determinar la asociación entre la presencia de al menos un SNP en genes y la resistencia a rifampicina**

Se realizó el análisis descrito en la sección III.3.1 para los genes de la base de 72 genomas que tuvieron una asociación estadísticamente significativa entre la presencia de al menos un SNP un gen y la resistencia a rifampicina. Asimismo, para los genes que presentaron asociación con la presencia de al menos un SNP en el gen rpoB, se procedió de manera similar. En esta sección se analizaron un total de 200 genomas.

### III.3.4 Análisis de 200 genomas para determinar la asociación entre la presencia de SNPs y la resistencia a rifampicina

Se realizó el análisis descrito en la sección III.3.2 para los SNPs de los genes que tuvieron una asociación estadísticamente significativa. Se analizaron 200 genomas.

Luego se unificaron los datos de ambas bases de datos obteniéndose 272 genoma, se analizaron los genes que presentaron un OR estadísticamente significativo en el análisis descrito en la sección III.3.1. Se calculó el OR, índice de Youden, razón de prevalencias y el test exacto de Fisher.

### III.3.5 Análisis de 914 genomas para determinar la asociación entre la presencia de SNPs en el gen ponA1 y la resistencia a rifampicina

Se analizaron 3 bases de datos de 914 genomas de *M. tuberculosis* los cuales proceden de muestras tomadas en el periodo 2001-2016. Para lo cual se procedió a realizar el siguiente análisis:

**Tabla 6.** Clasificación de genomas teniendo en cuenta la presencia de al menos un SNP en el gen PonA1

	Con al menos un SNP en el gen ponA1	Sin ningún SNP en el gen ponA1	Total
Resistente a rifampicina	A	B	A+B
Sensibles a rifampicina	C	D	C+D
Total	A+C	B+D	A+B+C+D

Luego se procedió a calcular el OR, índice de Youden, prueba de Chi-cuadrado y la razón de prevalencias. Más adelante se realizaron estas pruebas estadísticas para cada uno de los SNPs de ponA1 utilizando la metodología descrita en la sección III.3.2 (tabla 7).

**Tabla 7.** Clasificación de genomas teniendo en cuenta la presencia de un SNP en específico el gen PonA1

	Con al menos un SNP “Y” en el gen ponA1	Sin ningún SNP “Y” en el gen ponA1	Total
Resistente a rifampicina	A	B	A+B
Sensibles a rifampicina	C	D	C+D
Total	A+C	B+D	A+B+C+D

Las pruebas estadísticas se realizaron teniendo en cuenta un nivel de confianza del 95%.

Asimismo, para la clasificación y el análisis estadístico se usó el lenguaje de programación

R.

## IV. RESULTADOS

### IV.1 Análisis de 72 genomas para determinar la asociación entre la presencia de SNPs en un gen y la resistencia a rifampicina

Se encontró que 2044 genes en los 72 genomas de *M. tuberculosis* presentaron al menos un SNP. Se encontraron dos genes (plcA y Rv1999c) cuyo odds ratios fueron estadísticamente significativos (tabla 9). Asimismo, la frecuencia de genomas con SNPs en estos genes fue mayor en los genomas resistentes a rifampicina (tabla 8). Se excluyeron genes asociados a la resistencia a otras drogas, ya que la mayoría de los genomas estudiados proviene de cepas multidrogorresistente.

En la clasificación de estos genes A representa los genomas resistentes a rifampicina y con algún SNP en el gen estudiado, B son los genomas resistentes a rifampicina y sin ningún SNP en dicho gen, C son los genomas sensibles a rifampicina y con algún SNP en el gen estudiado y D son los genomas sensibles a rifampicina y sin ningún SNP en dicho gen.

**Tabla 8.** Frecuencia de los genes asociados a la resistencia a rifampicina que presentaron un OR significativo

Gen	A	B	C	D
plcA	7	28	1	36
Rv1999c	8	27	2	35

A (genomas resistentes a rifampicina y con algún SNP en el gen estudiado), B (genomas resistentes a rifampicina y sin ningún SNP en dicho gen), C (genomas sensibles a rifampicina y con algún SNP en el gen estudiado) y D (genomas sensibles a rifampicina y sin ningún SNP en dicho gen)

**Tabla 9.** Odds ratios significativos de genes asociados a la resistencia a rifampicina

Gen	OR	Intervalo de confianza bajo del OR	Intervalo de confianza alto del OR	Valor de $\rho$ del OR	Índice de Youden
plcA	9	1.045	77.476	0.045	0.39
Rv1999c	5.19	1.017	26.433	0.047	0.17

La presencia de al menos un SNP en el gen plcA presentan mayor capacidad para clasificar correctamente en resistente o sensible a rifampicina que la presencia de al menos un SNP en el gen Rv1999c (tabla 9).

Se encontraron 21 genes que tuvieron odds ratios estadísticamente significativos, con un nivel de confianza del 95% (tabla 10). Se consideró a los genomas con algún SNP en rpoB como el grupo expuesto. Donde A representa los genomas con algún SNP en rpoB y con algún SNP en el gen estudiado, B son los genomas con algún SNP en rpoB y sin ningún SNP en dicho gen, C son los genomas sin ningún SNP en rpoB y con algún SNP en el gen estudiado y D son los genomas sin ningún SNP en rpoB y sin ningún SNP en dicho gen. Los odds ratios son estadísticamente significativos, por lo que serían factores asociados a la presencia de al menos un SNP en rpoB (tabla 11). En este análisis el gen Rv1999c presentó un mayor índice de Youden, lo cual es indicativo que posee mayor capacidad para clasificar cepas respecto a si poseen o no algún SNP en rpoB.

**Tabla 10.** Frecuencia de los genes asociados a la presencia de al menos un SNP en el gen rpoB que presentaron un OR significativo

Gen	A	B	C	D
dnaE	11	29	0	32
dppA	10	30	1	31
esxI	12	28	2	30
esxP	12	28	2	30
esxV	11	29	3	29
guaB2	14	26	2	30
lipY	20	20	8	24
lldD2	17	23	3	28
mbtB	13	27	3	29
narJ	8	32	1	31
pdv	12	28	1	31
pimE	12	28	1	31
PPE64	10	30	0	32
Rv0380c	19	21	6	26
Rv0987	10	30	1	31
Rv1429	9	31	0	32
Rv2052c	10	30	1	31
Rv2308	13	27	2	30
Rv3729	19	21	7	25
Rv3798	17	23	4	28
Rv1999c	9	31	1	31

A (genomas con algún SNP en rpoB y con algún SNP en el gen estudiado), B (genomas con algún SNP en rpoB y sin ningún SNP en dicho gen), C (genomas sin ningún SNP en rpoB y con algún SNP en el gen estudiado) y D (genomas sin ningún SNP en rpoB y sin ningún SNP en dicho gen)

**Tabla 11.** Odds ratios significativos de genes asociados a la presencia de al menos un SNP en el gen rpoB

Gen	OR	Intervalo de confianza bajo del OR	Intervalo de confianza alto del OR	Valor de $\rho$ del OR	Índice de Youden
dnaE	28.51	1.61	503.39	0.02	0.52
dppA	10.33	1.245	85.751	0.02	0.42
esxI	6.43	1.320	31.307	0.02	0.37
esxP	6.67	1.37	32.52	0.02	0.38
esxV	5.69	1.159	27.922	0.03	0.35
guaB2	8.08	1.677	38.898	0.009	0.41
lipY	3	1.090	8.254	0.03	0.26
lldD2	6.90	1.796	26.495	0.004	0.4
mbtB	4.65	1.194	18.142	0.02	0.33
narJ	9	1.074	75.368	0.04	0.4
pdc	13.29	1.621	108.826	0.01	0.45
pimE	13.29	1.621	108.826	0.01	0.45
PPE64	22.38	1.256	398.549	0.03	0.47
Rv0380c	3.231	1.139	9.168	0.03	0.27
Rv0987	10.333	1.245	85.750	0.03	0.42
Rv1429	19.603	1.094	351.241	0.04	0.46
Rv1999c	9	1.074	75.368	0.04	0.4
Rv2052c	10.333	1.245	85.750	0.03	0.42
Rv2308	8.078	1.677	38.898	0.009	0.41
Rv3729	3.231	1.388	9.168	0.02	0.27
Rv3798	5.174	1.526	17.541	0.008	0.36

#### IV.2 Análisis de 72 genomas para determinar la asociación entre la presencia de SNPs y la resistencia a rifampicina

Los SNPs de los genes que presentaron odds ratios estadísticamente significativos, fueron más frecuentes en los genomas resistentes a rifampicina (tabla12). Los SNPs de los genes plcA y Rv1999c presentan un OR mayor a 1, pero no son estadísticamente significativos con un nivel de confianza del 95%. Asimismo, el SNP G337A del gen plcA y el SNP N271D del gen Rv1999c presentaron una mayor capacidad para clasificar correctamente en resistente o sensible a rifampicina que los otros SNPs presentados (tabla 13).

**Tabla 12.** Frecuencia los SNPs considerando los genomas resistentes a rifampicina como grupo expuesto

Gen	SNP	A	B	C	D
plcA	G337A	3	32	0	37
plcA	T446A	4	31	1	36
Rv1999c	A47T	1	34	0	37
Rv1999c	W94C	2	33	1	36
Rv1999c	L240M	4	31	2	35
Rv1999c	N271D	3	32	0	37

A (genomas resistentes a rifampicina y con el SNP estudiado), B (genomas resistentes a rifampicina y sin el SNP estudiado), C (genomas sensibles a rifampicina y con el SNP estudiado) y D (genomas sensibles a rifampicina y sin el SNP estudiado)

**Tabla 13.** OR e Índice de Youden de los SNPs de los genes asociados a la resistencia a rifampicina

Gen	SNP	OR	Intervalo de confianza bajo del OR	Intervalo de confianza alto del OR	Valor de p del OR	Índice de Youden
plcA	G337A	8.08	0.4	162.27	0.17	0.54
plcA	T446A	4.65	0.49	43.78	0.18	0.34
Rv1999c	A47T	3.26	0.13	82.76	0.47	0.52
Rv1999c	W94C	2.18	0.19	25.2	0.53	0.19
Rv1999c	L240M	2.26	0.39	13.19	0.37	0.20
Rv1999c	N271D	8.08	0.4	162.27	0.17	0.54

Los SNPs de los genes asociados a la presencia de al menos un SNP en rpoB, son más frecuentes en genomas con al menos un SNP en el gen rpoB (tabla 14).

**Tabla 14.** Frecuencia de SNPs considerando la presencia de SNPs en el gen rpoB como grupo expuesto

Gen	SNP	A	B	C	D
Rv1999c	A47T	1	39	0	32
Rv1999c	W94C	3	37	0	32
Rv1999c	L240M	5	35	1	31
Rv1999c	N271D	3	37	0	32
dnaE	P618R	7	33	0	32
dnaE	D961N	2	38	0	32
dnaE	A1031S	3	37	0	32
dppA	Q4R	5	35	1	31
dppA	A287V	1	39	0	32

dppA	P484S	4	36	0	32
esxI	Q20L	10	30	2	30
esxI	S23L	5	35	1	31
esxP	T3S	12	28	2	30
esxV	Q20L	10	30	2	30
esxV	S23L	6	34	1	31
guaB2	H38D	3	37	0	32
guaB2	A391T	5	35	2	30
guaB2	R529H	7	33	0	32
lipY	G58A	5	35	1	31
lipY	G197A	8	32	4	28
lipY	P427L	12	28	7	25
lipY	S49A	3	37	0	32
lldD2	V3I	3	37	2	30
lldD2	G59A	5	35	1	31
lldD2	V253M	13	27	3	29
mbtB	S412R	1	39	2	30
mbtB	V674L	5	35	1	31
mbtB	S1000L	7	33	0	32
narJ	P179R	5	35	1	31
narJ	P195L	4	36	0	32
pdC	P247T	3	37	0	32
pdC	D297N	2	38	1	31
pdC	S494R	7	33	0	32
pimE	T4P	5	35	1	31
pimE	G53S	7	33	0	32
PPE64	N196Y	7	33	0	32
PPE64	A500T	3	37	0	32
Rv0380c	M76V	5	35	1	31
Rv0380c	A171T	4	36	0	32
Rv0380c	M76V	5	35	1	31
Rv0380c	G180D	11	29	5	27
Rv0987	V83F	2	38	0	32
Rv0987	T142A	1	39	0	32
Rv0987	A372V	5	35	1	31
Rv0987	F717V	5	35	1	31
Rv1429	P91L	2	38	0	32
Rv1429	R135L	7	33	0	32
Rv2052c	P83L	1	39	0	32
Rv2052c	V422I	3	37	1	31
Rv2052c	A429T	3	37	0	32
Rv2052c	R512W	3	37	0	32
Rv2308	T31A	16	33	0	32

Rv2308	R179P	2	38	1	31
Rv2308	D212N	4	36	1	31
Rv3729	R24C	11	29	5	27
Rv3729	V142A	11	29	5	27
Rv3729	P269S	5	35	1	31
Rv3729	Q516H	5	35	1	31
Rv3729	R549K	3	37	0	32
Rv3729	G555R	4	36	0	32
Rv3798	L433R	17	23	4	28

A (genomas con algún SNP en rpoB y con el SNP estudiado), B (genomas con algún SNP en rpoB y sin el SNP estudiado), C (genomas sin ningún SNP en rpoB y con el SNP estudiado) y D (genomas sin ningún SNP en rpoB y sin el SNP estudiado)

Asimismo, los SNPs Q20L, T3S, V253M y L433R de los genes *esxI*, *esxP*, *lldD2* y *Rv3798* respectivamente, presentaron un OR que no incluye la unidad, por lo que son estadísticamente significativos con un nivel de confianza del 95%. Los SNPs P618R, R529H, S494R, G53S, N196Y, R135L y T31A de los genes *dnaE*, *guaB2*, *pdv*, *pimE*, *PPE64*, *Rv1429* y *Rv2308* respectivamente, presentaron una mayor capacidad para clasificar correctamente en 2 grupos, teniendo en cuenta la presencia o no algún SNP en el gen *rpoB* (tabla 15).

**Tabla 15.** OR e Índice de Youden de los SNPs de los genes asociados a la presencia de al menos un SNP en el gen *rpoB*

Gen	SNP	OR	Intervalo de confianza bajo del OR	Intervalo de confianza alto del OR	Valor de p del OR	Índice de Youden
Rv1999c	A47T	2.47	0.1	62.66	0.58	0.45
Rv1999c	W94C	6.07	0.302	121.88	0.24	0.46
Rv1999c	L240M	4.29	0.49	40	0.19	0.30
Rv1999c	N271D	6.07	0.302	121.88	0.24	0.46
<i>dnaE</i>	P618R	14.55	0.8	265.31	0.07	0.49
<i>dnaE</i>	D961N	2.47	0.1	62.66	0.58	0.45
<i>dnaE</i>	A1031S	6.07	0.302	121.88	0.24	0.46
<i>dppA</i>	Q4R	4.29	0.49	40	0.19	0.3
<i>dppA</i>	A287V	2.47	0.1	62.66	0.58	0.45
<i>dppA</i>	P484S	8.01	0.415	154.62	0.17	0.47
<i>esxI</i>	Q20L	5	1.01	24.77	0.049	0.33
<i>esxI</i>	S23L	4.29	0.49	40	0.19	0.3

esxP	T3S	6.43	1.32	31.31	0.02	0.37
esxV	Q20L	3.22	0.81	12.9	0.1	0.26
esxV	S23L	5.47	0.62	48.02	0.13	0.33
guaB2	H38D	6.07	0.302	121.88	0.24	0.46
guaB2	A391T	2.14	0.39	11.87	0.38	0.18
guaB2	R529H	14.55	0.8	265.31	0.07	0.49
lipY	G58A	4.29	0.49	40	0.19	0.3
lipY	G197A	1.75	0.48	6.44	0.4	0.13
lipY	P427L	1.53	0.52	4.49	0.44	0.1
lipY	S49A	6.07	0.302	121.88	0.24	0.46
lldD2	V3I	1.27	0.19	7.76	0.84	0.05
lldD2	G59A	4.29	0.49	40	0.19	0.3
lldD2	V253M	4.65	1.19	18.14	0.03	0.33
mbtB	S412R	0.38	0.03	4.44	0.44	-0.23
mbtB	V674L	4.29	0.49	40	0.19	0.3
mbtB	S1000L	14.13	0.76	257.47	0.07	0.48
narJ	P179R	3.44	0.37	32.46	0.28	0.26
narJ	P195L	8.01	0.415	154.62	0.17	0.47
pdC	P247T	6.07	0.302	121.88	0.24	0.46
pdC	D297N	1.63	0.14	18.85	0.7	0.12
pdC	S494R	14.55	0.8	265.31	0.07	0.49
pimE	T4P	4.29	0.49	40	0.19	0.3
pimE	G53S	14.55	0.8	265.31	0.07	0.49
PPE64	N196Y	14.55	0.8	265.31	0.07	0.49
PPE64	A500T	6.07	0.302	121.88	0.24	0.46
Rv0380c	M76V	4.29	0.49	40	0.19	0.3
Rv0380c	A171T	8.01	0.415	154.62	0.17	0.47
Rv0380c	M76V	4.29	0.49	40	0.19	0.3
Rv0380c	G180D	2.66	0.76	9.33	0.13	0.22
Rv0987	V83F	4.22	0.2	91.12	0.36	0.46
Rv0987	T142A	2.47	0.1	62.66	0.58	0.45
Rv0987	A372V	4.29	0.49	40	0.19	0.3
Rv0987	F717V	4.29	0.49	40	0.19	0.3
Rv1429	P91L	4.22	0.2	91.12	0.36	0.46
Rv1429	R135L	14.55	0.8	265.31	0.07	0.49
Rv2052c	P83L	2.47	0.1	62.66	0.58	0.45
Rv2052c	V422I	2.51	0.25	25.4	0.43	0.21
Rv2052c	A429T	6.07	0.302	121.88	0.24	0.46
Rv2052c	R512W	6.07	0.302	121.88	0.24	0.46
Rv2308	T31A	14.55	0.8	265.31	0.07	0.49
Rv2308	R179P	1.63	0.14	18.85	0.7	0.12
Rv2308	D212N	3.44	0.37	32.46	0.28	0.26
Rv3729	R24C	2.05	0.63	6.67	0.23	0.17

Rv3729	V142A	2.05	0.63	6.67	0.23	0.17
Rv3729	P269S	4.29	0.49	40	0.19	0.3
Rv3729	Q516H	4.29	0.49	40	0.19	0.3
Rv3729	R549K	6.07	0.302	121.88	0.24	0.46
Rv3729	G555R	8.01	0.415	154.62	0.17	0.47
Rv3798	L433R	5.17	1.53	17.54	0.008	0.36

### IV.3 Análisis de 200 genomas para determinar la asociación entre la presencia de SNPs en un gen y la resistencia a rifampicina

La presencia de al menos un SNP en los genes analizados es más frecuente en los genomas resistentes a rifampicina (tabla 16) y en los genomas con al menos un SNP en rpoB (tabla 18). La mayoría de los genes presentan odds ratios menor a 1 (tablas 17 y 19), lo cual es diferente a lo observado en el análisis anterior. Sin embargo, ninguno OR presentó un intervalo de confianza que no incluya a la unidad. Por lo que serían estadísticamente no significativos, con un nivel de confianza del 95%.

**Tabla 16.** Frecuencia de los genes asociados a la resistencia a rifampicina que presentaron un OR significativo

Gen	A	B	C	D
plcA	7	70	1	122
Rv1999c	10	67	16	107

A (genomas resistentes a rifampicina y con algún SNP en el gen estudiado), B (genomas resistentes a rifampicina y sin ningún SNP en dicho gen, C (genomas sensibles a rifampicina y con algún SNP en el gen estudiado) y D (genomas sensibles a rifampicina y sin ningún SNP en dicho gen)

**Tabla 17.** OR e índice de Youden de genes asociados a la resistencia a rifampicina

Gen	OR	Intervalo de confianza bajo del OR	Intervalo de confianza alto del OR	Valor de p del OR	Índice de Youden
plcA	1.23	0.55	2.77	0.61	0.05
Rv1999c	0.998	0.43	2.33	0.996	0

**Tabla 18.** Frecuencia de los genes asociados a la presencia de al menos un SNP en el gen rpoB que presentaron un OR significativo

Gen	A	B	C	D
dnaE	0	80	0	120
dppA	6	74	14	106
esxI	56	24	76	44
esxP	0	80	0	120
esxV	72	8	92	28
guaB2	6	74	12	108
lipY	22	58	58	62
lldD2	20	60	28	92
mbtB	14	66	14	106
narJ	6	74	10	110
pdc	8	72	16	104
pimE	8	72	14	106
PPE64	6	74	6	114
Rv0380c	20	60	40	80
Rv0987	10	70	20	100
Rv1429	4	76	6	114
Rv2052c	6	74	6	114
Rv2308	12	68	26	94
Rv3729	18	62	46	74
Rv3798	12	68	28	92
Rv1999c	10	70	16	104

A (genomas con algún SNP en rpoB y con algún SNP en el gen estudiado), B (genomas con algún SNP en rpoB y sin ningún SNP en dicho gen), C (genomas sin ningún SNP en rpoB y con algún SNP en el gen estudiado) y D (genomas sin ningún SNP en rpoB y sin ningún SNP en dicho gen)

**Tabla 19.** OR e Índice de Youden de genes asociados a la presencia de SNPs en el gen *rpoB*

Gen	OR	Intervalo de confianza bajo del OR	Intervalo de confianza alto del OR	Valor de p del OR	Índice de Youden
<i>dnaE</i>	1.5	0.029	76.21	0.84	0.10
<i>dppA</i>	0.61	0.23	1.67	0.34	-0.11
<i>esxI</i>	1.35	0.74	2.47	0.33	0.07
<i>esxP</i>	1.5	0.029	76.21	0.84	0.10
<i>esxV</i>	2.74	1.18	6.37	0.02	0.22
<i>guaB2</i>	0.73	0.26	2.03	0.55	-0.07
<i>lipY</i>	0.41	0.22	0.74	0.004	-0.21
<i>lldD2</i>	1.1	0.57	2.12	0.79	0.02
<i>mbtB</i>	1.61	0.72	3.58	0.25	0.12
<i>narJ</i>	0.89	0.31	2.56	0.83	-0.03
<i>pdC</i>	0.72	0.29	1.78	0.48	-0.08
<i>pimE</i>	0.84	0.34	2.11	0.71	-0.04
<i>PPE64</i>	1.54	0.48	4.96	0.47	0.11
<i>Rv0380c</i>	0.67	0.35	1.26	0.21	-0.10
<i>Rv0987</i>	0.71	0.32	1.62	0.42	-0.08
<i>Rv1429</i>	1	0.27	3.66	1	0.00
<i>Rv2052c</i>	1.54	0.48	4.96	0.47	0.11
<i>Rv2308</i>	0.64	0.3	1.35	0.24	-0.10
<i>Rv3729</i>	0.47	0.25	0.89	0.02	-0.17
<i>Rv3798</i>	0.53	0.25	1.11	0.09	-0.14
<i>Rv1999c</i>	0.93	0.4	2.16	0.86	-0.02

Sin embargo, los genes *plcA*, *dnaE*, *esxI*, *esxP*, *esxV*, *lldD2*, *mbtB*, *PPE64* y *Rv2052c* presentaron un OR mayor a 1, similar a lo observado en el anterior análisis, aunque ninguno presentó un OR estadísticamente significativo (tabla 19).

#### IV 4. Análisis de 200 genomas para determinar la asociación entre la presencia de SNPs y la resistencia a rifampicina

La frecuencia de SNPs en los genomas resistentes a rifampicina es baja (tabla 20), al igual que en los genomas con al menos un SNP en el gen *rpoB* (tabla 22). Asimismo, los SNPs G337A, P90S y K234N del gen *plcA* presentan un OR mayor a 1, pero no son estadísticamente significativos. De la misma manera, el SNP N270D del gen *Rv1999c* presenta un OR mayor a 1 (tabla 21).

**Tabla 20.** Frecuencia de los SNPs considerando la resistencia a rifampicina en el grupo expuesto

Gen	SNP	A	B	C	D
<i>plcA</i>	G337A	3	74	0	123
<i>plcA</i>	T446A	4	73	1	122
<i>plcA</i>	P90S	0	77	0	123
<i>plcA</i>	K234N	0	77	0	123
<i>Rv1999c</i>	A47T	0	77	0	123
<i>Rv1999c</i>	W94C	0	77	6	117
<i>Rv1999c</i>	L240M	6	71	10	113
<i>Rv1999c</i>	N271D	4	73	0	123
<i>Rv1999c</i>	N271H	0	77	2	121

A (genomas resistentes a rifampicina y con el SNP estudiado), B (genomas resistentes a rifampicina y sin el SNP estudiado), C (genomas sensibles a rifampicina y con el SNP estudiado) y D (genomas sensibles a rifampicina y sin el SNP estudiado)

**Tabla 21.** OR e índice de Youden de SNPs asociados a la resistencia a rifampicina

Gen	SNP	OR	Intervalo de confianza bajo del OR	Intervalo de confianza alto del OR	Valor de p del OR	Índice de Youden
plcA	G337A	15.12	0.802	284.92	0.07	0.63
plcA	T446A	0.62	0.19	2.05	0.43	-0.11
plcA	P90S	3.32	0.59	18.55	0.17	0.29
plcA	K234N	8.18	0.39	172.68	0.18	0.62
Rv1999c	A47T	1.59	0.03	81.15	0.82	0.11
Rv1999c	W94C	0.12	0.007	2.1	0.15	-0.40
Rv1999c	L240M	0.95	0.33	2.74	0.93	-0.01
Rv1999c	N271D	15.12	0.802	284.92	0.07	0.63
Rv1999c	N271H	0.31	0.015	6.62	0.46	-0.39

El SNP S412R del gen mbtB presenta un OR de 21.3 y un valor de p de 0.04, por lo que sería estadísticamente significativo (tabla 23). Sin embargo, este mismo SNP, en la tabla 15, presenta un OR menor a 1 y no es estadísticamente. Adicionalmente, los SNPs P247T, A500T, A429T de los genes pdc, PPE64, Rv2052c respectivamente; presentan un OR mayor a 10 (tabla 23) y estos mismos SNPs en la tabla 15 presentan un OR mayor a 6. Sin embargo, en ambos casos estos resultados no son estadísticamente significativos.

**Tabla 22.** Frecuencia de SNPs considerando la presencia de SNPs en el gen *rpoB* como grupo expuesto

GEN	SNP	A	B	C	D
Rv1999c	A47T	0	80	0	120
Rv1999c	W94C	0	80	6	114
Rv1999c	L240M	6	74	10	110
Rv1999c	N271D	4	76	0	120
Rv1999c	N271H	0	80	2	118
dnaE	P618R	0	80	0	120
dnaE	D961N	0	80	0	120
dnaE	A1031S	0	80	0	120
dppA	Q4R	6	74	10	110
dppA	A287V	0	80	0	120
dppA	P484S	0	80	2	118
dppA	P273S	0	80	2	118
esxI	Q20L	36	44	52	68
esxI	S23L	38	42	56	64
esxI	S33R	6	74	2	118
esxP	T3S	0	80	0	120
esxV	Q20L	18	62	36	84
esxV	S23L	60	20	86	34
esxV	S33R	26	54	50	70
esxV	S39G	2	78	0	120
esxV	T37A	2	78	0	120
guaB2	H38D	0	80	0	120
guaB2	A391T	4	76	10	110
guaB2	R529H	0	80	2	118
guaB2	A499V	2	78	0	120
lipY	G58A	6	74	10	110
lipY	G197A	10	70	18	102
lipY	P427L	14	66	44	76
lipY	S49A	2	78	0	120
lipY	G278S	0	80	2	118
lldD2	V3I	8	72	14	106
lldD2	G59A	6	74	10	110
lldD2	V253M	10	70	14	106
mbtB	S412R	6	74	0	120
mbtB	V674L	6	74	8	112
mbtB	S1000L	2	78	2	118

mbtB	S178T	0	80	2	118
narJ	P179R	6	74	10	110
narJ	P195L	0	80	0	120
narJ	E122V	2	78	0	120
pdC	P247T	4	76	0	120
pdC	D297N	2	78	14	106
pdC	S494R	2	78	2	118
pdC	T396A	0	80	2	118
pimE	T4P	6	74	10	110
pimE	G53S	2	78	2	118
pimE	R363P	0	80	2	118
PPE64	N196Y	2	78	2	118
PPE64	A500T	4	76	0	120
PPE64	G75S	0	80	2	118
PPE64	T50P	0	80	2	118
Rv0380c	M76V	6	74	4	116
Rv0380c	A171T	0	80	0	120
Rv0380c	M76V	0	80	6	114
Rv0380c	G180D	12	68	30	90
Rv0380c	H86D	2	78	0	120
Rv0987	V83F	2	78	4	116
Rv0987	T142A	0	80	0	120
Rv0987	A372V	6	74	10	110
Rv0987	F717V	6	74	10	110
Rv0987	W797X	2	78	0	120
Rv0987	L415F	0	80	4	116
Rv0987	E854A	0	80	4	116
Rv0987	L264I	0	80	2	118
Rv1429	P91L	2	78	4	116
Rv1429	R135L	2	78	2	118
Rv2052c	P83L	0	80	0	120
Rv2052c	V422I	2	78	0	120
Rv2052c	A429T	4	76	0	120
Rv2052c	R512W	0	80	6	114
Rv2308	T31A	2	78	0	120
Rv2308	R179P	2	78	14	106
Rv2308	D212N	6	74	10	110
Rv2308	L145P	2	78	0	120
Rv2308	R106P	0	80	2	118

Rv3729	R24C	10	70	16	104
Rv3729	V142A	10	70	16	104
Rv3729	P269S	6	74	10	110
Rv3729	Q516H	6	74	10	110
Rv3729	R549K	0	80	6	114
Rv3729	G555R	0	80	2	118
Rv3729	D110N	0	80	2	118
Rv3729	R694Q	0	80	2	118
Rv3798	L433R	12	68	28	92

A (genomas con algún SNP en rpoB y con el SNP estudiado), B (genomas con algún SNP en rpoB y sin el SNP estudiado), C (genomas sin ningún SNP en rpoB y con el SNP estudiado) y D (genomas sin ningún SNP en rpoB y sin el SNP estudiado)

**Tabla 23.** OR e Índice de Youden de genes asociados a la presencia de SNPs en el gen rpoB

GEN	SNP	OR	Intervalo de confianza bajo del OR	Intervalo de confianza alto del OR	Valor de p del OR	Índice de Youden
Rv1999c	A47T	1.5	0.029	76.21	0.84	0.10
Rv1999c	W94C	0.11	0.006	1.97	0.13	-0.34
Rv1999c	L240M	0.89	0.31	2.56	0.83	-0.02
Rv1999c	N271D	14.18	0.753	267.04	0.08	0.51
Rv1999c	N271H	0.29	0.014	6.21	0.43	-0.24
dnaE	P618R	1.5	0.029	76.21	0.84	0.10
dnaE	D961N	1.5	0.029	76.21	0.84	0.10
dnaE	A1031S	1.5	0.029	76.21	0.84	0.10
dppA	Q4R	0.89	0.31	2.56	0.83	-0.02
dppA	A287V	1.5	0.029	76.21	0.84	0.10
dppA	P484S	0.29	0.014	6.21	0.43	-0.24
dppA	P273S	0.29	0.014	6.21	0.43	-0.24
esxI	Q20L	1.07	0.61	1.89	0.82	0.02
esxI	S23L	1.03	0.59	1.82	0.91	0.01
esxI	S33R	4.78	0.94	24.33	0.06	0.34
esxP	T3S	1.5	0.029	76.21	0.84	0.10

esxV	Q20L	0.68	0.35	1.3	0.24	-0.09
esxV	S23L	1.19	0.62	2.26	0.6	0.04
esxV	S33R	0.67	0.37	1.22	0.19	-0.09
esxV	S39G	7.68	0.361	162.01	0.19	0.44
esxV	T37A	7.68	0.361	162.01	0.19	0.44
guaB2	H38D	1.5	0.029	76.21	0.84	0.10
guaB2	A391T	0.58	0.175	1.91	0.37	-0.11
guaB2	R529H	0.29	0.014	6.21	0.43	-0.24
guaB2	A499V	7.68	0.361	162.01	0.19	0.44
lipY	G58A	0.89	0.31	2.56	0.83	-0.02
lipY	G197A	0.81	0.35	1.86	0.62	-0.05
lipY	P427L	0.37	0.18	0.73	0.004	-0.22
lipY	S49A	7.68	0.361	162.01	0.19	0.44
lipY	G278S	0.29	0.014	6.21	0.43	-0.24
lldD2	V3I	0.84	0.34	2.11	0.71	-0.04
lldD2	G59A	0.89	0.31	2.56	0.83	-0.02
lldD2	V253M	1.08	0.46	2.57	0.86	0.02
mbtB	S412R	21.03	1.17	378.71	0.04	0.55
mbtB	V674L	1.14	0.38	3.41	0.82	0.03
mbtB	S1000L	1.51	0.21	10.97	0.68	0.10
mbtB	S178T	0.29	0.014	6.21	0.43	-0.24
narJ	P179R	0.89	0.31	2.56	0.83	-0.02
narJ	P195L	1.5	0.029	76.21	0.84	0.10
narJ	E122V	7.68	0.361	162.01	0.19	0.44
pdC	P247T	14.18	0.753	267.04	0.08	0.51
pdC	D297N	0.19	0.04	0.88	0.03	-0.28
pdC	S494R	1.51	0.21	10.97	0.68	0.10
pdC	T396A	0.29	0.014	6.21	0.43	-0.24
pimE	T4P	0.89	0.31	2.56	0.83	-0.02
pimE	G53S	1.51	0.21	10.97	0.68	0.10

pimE	R363P	0.29	0.014	6.21	0.43	-0.24
PPE64	N196Y	1.51	0.21	10.97	0.68	0.10
PPE64	A500T	14.18	0.753	267.04	0.08	0.51
PPE64	G75S	0.29	0.014	6.21	0.43	-0.24
PPE64	T50P	0.29	0.014	6.21	0.43	-0.24
Rv0380c	M76V	2.35	0.64	8.61	0.2	0.20
Rv0380c	A171T	1.5	0.029	76.21	0.84	0.10
Rv0380c	M76V	0.11	0.006	1.97	0.13	-0.34
Rv0380c	G180D	0.53	0.252	1.11	0.09	-0.14
Rv0380c	H86D	7.68	0.361	162.01	0.19	0.44
Rv0987	V83F	0.74	0.13	4.16	0.74	-0.05
Rv0987	T142A	1.5	0.029	76.21	0.84	0.10
Rv0987	A372V	0.89	0.31	2.56	0.83	-0.02
Rv0987	F717V	0.89	0.31	2.56	0.83	-0.02
Rv0987	W797X	7.68	0.361	162.01	0.19	0.44
Rv0987	L415F	0.16	0.009	3.03	0.22	-0.31
Rv0987	E854A	0.16	0.009	3.03	0.22	-0.31
Rv0987	L264I	0.29	0.014	6.21	0.43	-0.24
Rv1429	P91L	0.74	0.13	4.16	0.74	-0.05
Rv1429	R135L	1.51	0.21	10.97	0.68	0.10
Rv2052c	P83L	1.5	0.029	76.21	0.84	0.10
Rv2052c	V422I	7.68	0.361	162.01	0.19	0.44
Rv2052c	A429T	14.18	0.753	267.04	0.08	0.51
Rv2052c	R512W	0.11	0.006	1.97	0.13	-0.34
Rv2308	T31A	7.68	0.361	162.01	0.19	0.44
Rv2308	R179P	0.19	0.04	0.88	0.03	-0.28
Rv2308	D212N	0.89	0.31	2.56	0.83	-0.02
Rv2308	L145P	7.68	0.361	162.01	0.19	0.44
Rv2308	R106P	0.29	0.014	6.21	0.43	-0.24
Rv3729	R24C	0.93	0.4	2.16	0.86	-0.01

Rv3729	V142A	0.93	0.4	2.16	0.86	-0.01
Rv3729	P269S	0.89	0.31	2.56	0.83	-0.02
Rv3729	Q516H	0.89	0.31	2.56	0.83	-0.02
Rv3729	R549K	0.11	0.006	1.97	0.13	-0.34
Rv3729	G555R	0.29	0.014	6.21	0.43	-0.24
Rv3729	D110N	0.29	0.014	6.21	0.43	-0.24
Rv3729	R694Q	0.29	0.014	6.21	0.43	-0.24
Rv3798	L433R	0.58	0.28	1.22	0.15	-0.12

En la tabla 24 y 25 puede apreciarse los OR y el índice de Youden de los datos unificados de las bases de 72 y 200 genomas. Se puede observar que la presencia de al menos un SNP en los genes dnaE, esxP, mbtB, PPE64, Rv1429 y Rv2025c está asociado significativamente a la presencia de al menos un SNP en el gen rpoB.

**Tabla 24.** Pruebas estadísticas de los genes asociados a la resistencia a rifampicina

Gen	OR	Valor de p del OR	Intervalo de confianza del OR	Índice de Youden	Test exacto de Fisher (valor de p)	Razón de prevalencias	Intervalos de confianza del RP
plcA	0.94	0.83	0.51-1.73	-0.01	0.88	0.96	0.62-1.46
Rv1999c	1.5	0.25	0.75-3.05	0.10	0.28	1.26	0.87-1.8

**Tabla 25.** Pruebas estadísticas de los genes asociados a la presencia de SNPs en el gen rpoB

Gen	OR	Valor de p del OR	Intervalo de confianza del OR	Índice de Youden	Test exacto de Fisher (valor de p)	Razón de prevalencias	Intervalos de confianza del RP
dnaE*	35.14	0.01	2.06-599.89	0.58	0.00003	ID	ID
dppA	1.41	0.37	0.66-2.97	0.08	0.44	1.35	0.7-2.6
esxI	1.24	0.38	0.77-2.01	0.05	0.39	1.1	0.89-1.38
esxP*	8.41	0.004	1.9-38.36	0.44	0.002	7.66	1.7-33.59
esxV	1.38	0.21	0.83-2.3	0.08	0.25	1.12	0.94-1.33
guaB2	1.97	0.07	0.95-4.09	0.17	0.1	1.8	0.95-3.43
lipY	0.70	0.16	0.43-1.15	-0.09	0.17	0.81	0.59-1.09
lldD2	1.73	0.05	0.99-3.001	0.14	0.07	1.5	0.99- 2.27
mbtB*	2.31	0.01	1.19-4.47	0.21	0.01	2.01	1.15-3.51
narJ	1.83	0.15	0.81-4.15	0.15	0.15	1.73	0.82-3.62
pdc	1.59	0.19	0.79-3.19	0.12	0.21	1.49	0.82-2.72
pimE	1.83	0.1	0.89-3.74	0.15	0.1	1.69	0.9-3.16
PPE64*	3.74	0.008	1.42-9.89	0.31	0.006	3.38	1.36-8.37
Rv0380c	1.08	0.78	0.64-1.8	0.02	0.79	1.05	0.74-1.49
Rv0987	1.25	0.51	0.64-2.43	0.05	0.61	1.21	0.69-2.12
Rv1429*	2.96	0.03	1.09-8.03	0.26	0.03	2.74	1.08-7.01
Rv2052c*	3.19	0.01	1.27-8.02	0.28	0.01	2.9	1.23-6.81
Rv2308	1.22	0.51	0.67-2.23	0.05	0.54	1.18	0.73-1.9
Rv3729	0.83	0.48	0.5-1.39	-0.04	0.52	0.88	0.63-1.25
Rv3798	1.11	0.73	0.63-1.95	0.02	0.77	1.08	0.7-1.67
Rv1999c	1.49	0.26	0.74-3.02	0.05	0.28	1.42	0.77-2.6

\*Genes cuyos resultados son estadísticamente significativos. ID: indeterminado.

**Tabla 26.** Pruebas estadísticas de los SNPs de los genes asociados a la presencia de al menos un SNP en el gen rpoB

Gen	SNP	OR	Valor de $\rho$ del OR	Razón de prevalencias	Intervalos de confianza de RP	Valor de $\rho$ del test exacto de Fisher	Índice de Youden
dnaE	P618R*	20.15	0.04	ID	ID	0.003	0.57
dnaE	D961N	3.83	0.41	ID	ID	0.44	0.56
dnaE	A1031S	9.09	0.15	ID	ID	0.08	0.57
esxP	T3S*	8.33	0.006	7.6	1.73-33.31	0.002	0.44
mbtB	S412R	4.65	0.06	4.43	0.94-20.95	0.047	0.35
mbtB	V674L	1.6	0.31	1.55	0.66-3.61	0.35	0.12
mbtB	S1000L*	6.08	0.02	5.7	1.25-25.89	0.01	0.39
mbtB	S178T	0.25	0.37	ID	ID	0.51	-0.44
PPE64	N196Y*	6.08	0.02	5.7	1.25-25.89	0.01	0.39
PPE64	A500T*	20.15	0.04	ID	ID	0.003	0.57
PPE64	G75S	0.25	0.37	ID	ID	0.51	-0.44
PPE64	T50P	0.25	0.37	ID	ID	0.51	-0.44
Rv2052c	V422I	6.57	0.09	6.33	0.75-53.49	0.09	0.40
Rv2052c	A429T*	20.15	0.04	ID	ID	0.003	0.57
Rv2052c	R512W	0.62	0.51	0.63	0.16-2.48	0.74	-0.11
Rv2052c	P83L	3.83	0.41	ID	ID	0.44	0.56
Rv1429	P91L	1.28	0.73	1.27	0.32-4.96	0.73	0.06
Rv1429	R135L*	6.08	0.02	5.7	1.25-25.89	0.01	0.39

\* SNPs cuyos resultados son estadísticamente significativos. ID: indeterminado.

Se encontró que los SNPs P618R (gen dnaE), T3S (gen esxP), S1000L (gen mbtB), N196Y (gen PPE64), A500T (gen PPE64), A429T (gen Rv 2052c) y R135L (gen Rv 1429) son factores asociados significativamente a la presencia de al menos un SNP en rpoB. Además, son más prevalentes en los genomas con algún SNP en rpoB.

#### IV.5 Análisis de 914 genomas para determinar la asociación entre la presencia de SNPs en el gen ponA1 y la resistencia a rifampicina

Este estudio evaluó 3 bases de datos de 469, 72 y 373 genomas las cuales fueron obtenidas en los períodos 2007-2012, 2001-2007 y 2016, respectivamente. Los genomas fueron clasificados según el criterio descrito en la sección III.3.5 (tabla 27).

**Tabla 27.** Clasificación de los genomas de las 3 bases de datos según la presencia o ausencia de al menos un SNP en el gen ponA1

Bases de datos	A	B	C	D
Base de 469 genomas	218	128	89	34
Base de 72 genomas	20	15	21	16
Base de 373 genomas	37	21	223	92
<b>Total</b>	<b>275</b>	<b>164</b>	<b>333</b>	<b>142</b>

A (genomas resistentes a rifampicina y con algún SNP en ponA1), B (genomas resistentes a rifampicina y sin ningún SNP en ponA1), C (genomas sensibles a rifampicina y con algún SNP en ponA1) y D (genomas sensibles a rifampicina y sin ningún SNP en ponA1)

El número de genomas con al menos un SNP en ponA1 duplica el número de genomas sin ningún SNP en ponA1. Por lo que se puede afirmar que los SNPs en este gen son bastante frecuentes y es ligeramente mayor el número de cepas, sensibles a rifampicina, con SNPs en ponA1 que cepas resistentes con SNPs en ponA1 (tabla 28).

**Tabla 28.** Clasificación de los genomas totales según la resistencia a rifampicina y la presencia de al menos un SNP en el gen ponA1

	Con al menos un SNP en el gen ponA1	Sin ningún SNP en el gen ponA1	Total
Resistente a rifampicina	275	164	439
Sensibles a rifampicina	333	142	475
<b>Total</b>	<b>608</b>	<b>306</b>	<b>914</b>

La presencia de al menos un SNP en el gen ponA1 es un factor asociado a la resistencia a rifampicina, ya que tiene un OR, estadísticamente significativo, menor a 1. Es más probable encontrar SNPs del gen ponA1 en cepas sensibles a rifampicina que en cepas resistentes, ya que la razón de prevalencias es menor a 1. Además, determinar la presencia de SNPs en ponA1 no es una prueba muy útil si queremos determinar si una cepa es resistente a rifampicina, ya que el índice de Youden fue muy bajo (tabla 29).

**Tabla 29.** OR, índice de Youden y comparación de proporciones de la asociación entre la presencia de al menos un SNP en el gen ponA1 y la resistencia a rifampicina

Gen	OR	Intervalo de confianza del OR	Valor de $\rho$ del OR	Índice de Youden	Razón de prevalencias	Intervalo de confianza del RP
ponA1	0.72	0.543-0.942	0.017	-0.08	0.89	0.81-0.98

**Tabla 30.** Prueba de Chi-cuadrado

Gen	$X^2$ calc	Valor de $\rho$ del $X^2$ calc
ponA1	5.71	0.017

Existe una asociación estadísticamente significativa entre la presencia de al menos un SNP y la resistencia a rifampicina, ya que el valor de  $\rho$  de la prueba de Chi-cuadrado es menor a 0.05 (tabla 30).

El cambio de prolina a serina en la posición 631 es más frecuente que todos los otros SNPs juntos. Además, este SNP es ligeramente más frecuente en cepas sensibles a rifampicina. Por el contrario, el SNP A516T es ligeramente más frecuente en cepas resistentes a rifampicina (tabla 31).

**Tabla 31.** Frecuencia de SNPs en el gen ponA1 según la presencia o ausencia de resistencia a rifampicina

SNP	A	B	C	D
P631S	252	187	303	172
A516T	70	369	51	424
T658A	5	434	4	471
I41M	19	420	20	455
P630S	0	439	2	473
A292T	0	439	1	474
Q415H	0	439	1	474
P6L	0	439	1	474
T9I	0	439	3	472
A185V	1	438	0	475

El cambio de prolina a serina en la posición 631 de PonA1 es un factor asociado estadísticamente a la resistencia a rifampicina. Además, es ligeramente más prevalente en cepas sensibles a rifampicina. El SNP A516T es un factor asociado estadísticamente a la resistencia a rifampicina (tabla 32), este SNP es más frecuente en cepas resistentes a rifampicina.

**Tabla 32.** OR, índice de Youden y comparación de proporciones de la asociación entre la presencia de al menos un SNP en el gen ponA1 y la resistencia a rifampicina de los SNPs del gen ponA1

SNP	OR	Intervalos de confianza del OR	Valor de p del OR	Índice de Youden	Razón de prevalencias	Intervalos de confianza del RP
P631S*	0.77	0.586-0.998	0.048	-0.07	0.9	0.81-0.9998
A516T*	1.58	1.07-2.32	0.02	0.11	1.49	1.06-2.08
T658A	1.36	0.36-5.08	0.65	0.08	1.35	0.37-5.01
I41M	1.03	0.54-1.96	0.93	0.01	1.03	0.56-1.9
P630S	0.22	0.01-4.5	0.32	-0.48	ID	ID
A292T	0.36	0.01-8.86	0.53	-0.48	ID	ID
Q415H	0.36	0.01-8.86	0.53	-0.48	ID	ID

P6L	0.36	0.01-8.86	0.53	-0.48	ID	ID
T9I	0.15	0.007-2.96	0.22	-0.48	ID	ID
A185V	3.25	0.13-80.07	0.47	0.52	ID	ID

\*SNPs estadísticamente significativos. ID: indeterminado

**Tabla 33.** Prueba de Chi-cuadrado para los SNPs

SNP	$X^2$ <i>calc</i>	Valor de $p$ del $X^2$ <i>calc</i>
P631S	3.9	0.048
A516T	5.39	0.02

Se encontró que los SNPs P631S y A516T están asociados estadísticamente a la resistencia a rifampicina, cuando se los evaluó con la prueba de Chi-cuadrado (valor de  $p > 0.05$ ).

## V. DISCUSIÓN

Las mutaciones en los genes *rpoC* y *rpoA* incrementan el fitness de las cepas resistentes a rifampicina, las cuales presentan una disminución del fitness debido a SNPs en el gen *rpoB*. Se ha propuesto el término mutaciones compensatorias para referirse a ellas (10). Sin embargo, las mutaciones en los genes *rpoC* y *rpoA* no explicarían del todo las características fenotípicas observadas en las cepas resistentes a rifampicina (12) (13). Por lo que, existirían mutaciones en otros genes que estarían asociados a la resistencia a rifampicina, como las que fueron encontradas en otros microorganismos (25).

Este estudio analizó 3 bases de datos de genomas de *M. tuberculosis* con el objetivo de identificar potenciales mutaciones compensatorias asociadas a la resistencia a rifampicina. En una primera etapa se evaluaron todos los genes que presentaron al menos un SNP de una base de datos de 72 genomas. Se encontró que la presencia de al menos un SNP en dos genes (*plcA* y *Rv1999c*), cuando se consideró a los genomas resistentes a rifampicina como el grupo expuesto, son factores asociados a la resistencia a rifampicina. Sin embargo, cuando se consideró a los genomas con algún SNP en el gen *rpoB* como el grupo expuesto, se encontró que 21 genes serían factores asociados a la presencia de al menos un SNP en el gen *rpoB*. Esto se debe a que no todos los SNPs en el gen *rpoB* generan resistencia a rifampicina; sin embargo, este estudio incluyó estas mutaciones ya que podrían contribuir con la disminución del fitness.

En una segunda etapa se evaluó solo los genes que presentaron un OR estadísticamente significativo, en la primera etapa. Se analizaron 200 genomas y se encontró que solo el gen *esxV* presentó un OR con significancia estadística con un nivel de confianza del 95%, además su OR fue mayor que 1; por lo que sería un factor asociado para la presencia de al menos un

SNP en el gen *rpoB*. Sin embargo, cuando se unificó estos datos con la base de 72 genomas, se encontró que este gen presentaba un odds ratio estadísticamente no significativo. Asimismo, se encontró otros genes que también presentaron odds ratios mayor que 1 (*plcA*, *dnaE*, *esxI*, *esxP*, *lldD2*, *mbtB*, *PPE64* y *Rv2052c*).

Posteriormente se unificaron ambas bases de datos obteniéndose un total de 272 genomas, donde la presencia de al menos un SNP en 6 genes (*dnaE*, *esxP*, *mbtB*, *PPE64*, *Rv1429* y *Rv2052c*) serían factores asociados significativamente a la presencia de al menos un SNP en el gen *rpoB*. En este análisis no se encontró que la presencia de al menos un SNP en algún gen sea un factor asociado para la resistencia a rifampicina. Ya que los genes *plcA* y *Rv1999c* presentaron odds ratios estadísticamente no significativos con un nivel de confianza del 95%. Asimismo, los genes *dnaE*, *esxP* y *PPE64* presentaron un índice de Youden igual a 0.58, 0.44 y 0.31 respectivamente, los cuales fueron los valores más altos encontrados en este grupo de 6 genes. Por lo que tendrían mayor capacidad para clasificar genomas o cepas en según si presentan algún SNP en el gen *rpoB* o no. Los genes *esxP*, *mbtB*, *PPE64*, *Rv1429* y *Rv2052c*, en esta unificación de datos, presentan razones de prevalencia mayores a 1; asimismo, estos resultados son estadísticamente significativos con un nivel de confianza del 95%. Por tanto, cuando se dan SNPs en estos genes es más probable que estas cepas también presenten SNPs en el gen *rpoB*. Ya que los SNPs en estos genes son más prevalentes en el grupo con SNPs en el gen *rpoB*.

El gen *dnaE* codifica la subunidad  $\alpha$  del ADN polimerasa III, la cual participa en la replicación del ADN y en la reparación de errores (presenta actividad 3'-exonucleasa y 5'-exonucleasa) (46). Nuestro estudio encontró SNPs en este gen cuando se analizó la base de datos de 72 genomas, pero en la base de datos de 200 genomas no se encontró ningún SNPs.

Además, el cambio de prolina a arginina en la posición 618 (P618R) es el SNP más frecuente. Los SNPs en este gen podrían generar un cambio en la reparación de errores, por lo que las cepas de *M. tuberculosis* con SNPs en este gen serían más propensos a mutaciones (47). Nuestro estudio propone que el SNP P618R podría generar un cambio conformacional en esta subunidad del ADN polimerasa III, ya que la prolina es un aminoácido apolar y la arginina presenta polaridad (48); por lo que esta mutación generaría una región con polaridad que alteraría las posiciones de aminoácidos vecinos.

El gen *esxP* codifica una proteína similar a la ESAT-6, cuya función aún no se conoce. Sin embargo, se ha encontrado niveles altos en pacientes con TB activa y se ha sugerido que podría ser importante para la propagación de *M. tuberculosis* en el entorno alveolar (49). Nuestro estudio encontró solo el SNP T3S en el gen *esxP* en la base de datos de 72 genomas; sin embargo, en la base de datos de 200 genomas no se encontró ningún SNP en este gen. El cambio de serina a treonina no produciría cambios importantes en la proteína codificada por el gen *esxP*, ya que ambos aminoácidos presentan un grupo hidroxilo en su cadena lateral (48); por lo que no se darían cambios estructurales importantes.

Por otro lado, el gen *mbtB* codifica la feniloxazolina sintasa, la cual está involucrada en la síntesis de micobactina (50). Recientemente se ha propuesto que esta proteína interactúa con la proteína codificada por el gen *Rv0148*, la cual sería una deshidrogenasa/ reductasa de cadena corta y su sobreexpresión estaría relacionada con la resistencia a aminoglucósidos (51). Nuestro estudio encontró que el SNP S1000L era más común en cepas con algún SNP en el gen *rpoB* cuando se analizó la base de 72 genomas; sin embargo, esto fue refutado cuando se analizó la base de 200 genomas. Además, se encontró el SNP S412R en solamente los genomas con algún SNP en el gen *rpoB* en la base de 200 genomas y en la base de 72

genomas su frecuencia fue baja y estaba presente en los grupos con SNP y sin SNP en el gen *rpoB*. Sin embargo, en forma global la presencia de algún SNP en este gen es un factor asociado a la presencia de algún SNP en el gen *rpoB*. Además, es más frecuente en cepas con alguna mutación en el gen *rpoB*. Nuestro estudio plantea que el SNP S1000L podría generar cambios estructurales en la enzima feniloxazolina sintasa, ya que el cambio de serina (un aminoácido polar) a leucina (un aminoácido apolar) (48) podría alterar las posiciones de aminoácidos vecinos. Sin embargo, el cambio de serina a arginina (S412R) no produciría cambios importantes, ya que ambos aminoácidos son polares.

La familia de genes PPE codifica proteínas con regiones ricas en repeticiones de prolina-prolina-glutamato, las cuales suelen desencadenar respuestas inmunes innatas y son blanco del sistema inmune adaptativo. Además, las regiones ricas en estos aminoácidos suelen ser muy polimórficas (42). Nuestro estudio encontró el SNP N196Y en el gen PPE64, un gen de la familia PPE, estaba presente en solamente genomas con algún SNP en el gen *rpoB* en la base de 72 genomas; sin embargo, en la base de 200 genomas este SNP estaba presente en igual cantidad en ambos grupos, aunque en menor frecuencia que en la primera base de datos. Como en el caso anterior, cuando se unificó ambas bases de datos, se encontró que la presencia de algún SNP en este gen es un factor asociado a la presencia de algún SNP en el gen *rpoB*. Nuestro estudio propone que el SNP N196Y podría generar cambios estructurales en la proteína codificada por el gen PPE64, debido a que ambos aminoácidos presentan diferencias estructurales (la asparagina presenta un grupo carboxamida y la tirosina un grupo aromático en sus cadenas laterales) (48).

El gen Rv1429 y el gen Rv2052 codifican proteínas hipotéticas cuya función no se conoce (52). Nuestro estudio encontró los SNPs P91L y R135L en el gen RV1429, este último en

mayor frecuencia. Este último SNP podría generar cambios estructurales en la proteína hipotética codificada por este gen, ya que la arginina presenta un grupo guanidino lo que le confiere polaridad y la leucina presenta un grupo isobutilo que le confiere apolaridad (48). Por otro lado, el SNP A429T del gen Rv2052 se encontró solo en los genomas que tenían algún SNP en el gen rpoB, cuando se analizaron los 270 genomas; sin embargo, la frecuencia de este SNP fue baja. Este SNP también generaría cambios estructurales, ya que la alanina es apolar y la treonina polar (48).

El gen ponA1 es un gen que pertenece a la familia PPE y codifica una proteína bifuncional de unión a la penicilina con actividad transglicosilasa y transpeptidasa. Participa en la biosíntesis del peptidoglicano y está involucrado en procesos de crecimiento y morfología celular (16) (26). Se ha reportado que el SNP Q365H altera la susceptibilidad de *M. tuberculosis* a rifampicina, mediante el cambio de la concentración inhibitoria mínima (MIC) (18). Se observó resultados similares con el SNP T34D (17). Por lo tanto, determinados SNPs en PonA1 podrían alterar la susceptibilidad a rifampicina en cepas de *M. tuberculosis* (18). Nuestro estudio encontró que la presencia de al menos un SNP en ponA1 es un factor asociado significativamente a la resistencia a rifampicina, ya que presentó un odds ratio en un intervalo de confianza que no incluyó a la unidad. Asimismo, es más probable encontrar estos SNP de ponA1 en cepas sensibles a rifampicina que en resistentes; ya que se encontró una razón de prevalencias de 0.89. Estos resultados se obtuvieron luego de evaluar 914 genomas de *M. tuberculosis*.

El cambio de prolina a serina en la posición 631 es el SNP más frecuente de ponA1, encontrado en este estudio. Se observó que la presencia de este SNP es un factor asociado significativamente a la resistencia a rifampicina, ya que posee un OR de 0.77 y un intervalo

de confianza que no incluye la unidad. Asimismo, presenta una razón de prevalencia de 0.9 por lo que es ligeramente más probable encontrar este SNP en cepas sensibles a rifampicina que en resistentes. Un análisis proteómico identificó a la mutación P631S de PonA1 en cepas de *M. tuberculosis* H37Ra, las cuales son avirulentas, por lo que se sugirió que esta proteína está asociada a la virulencia y mutaciones puntuales en la región terminal rica en prolina podrían ser claves para el desarrollo de avirulencia (53). Otro estudio que evaluó las familias de genes PE y PPE sugirió que la mutación P631S en PonA1 sería un marcador de resistencia a rifampicina (42).

Estudios recientes señalan que PonA1 regula la actividad de RipA mediante su unión al dominio endopeptidasa, en dicha interacción participa la región C-terminal de PonA1, la cual es una región de alrededor de 150 residuos rica en prolina (16). Recientemente se encontró que mutaciones en esta región pueden alterar capacidad de unión de PonA1 con RipA y estas alteraciones generan cambios fenotípicos como crecimiento lento o alteraciones en la morfología celular (26). Por lo que los resultados presentados en este estudio se deberían a que el cambio de prolina a serina en la posición 631 alteraría la capacidad de unión de PonA1 con RipA.

Nuestro estudio encontró que el cambio de alanina a treonina en la posición 516 de PonA1 es un factor asociado significativamente a la resistencia a rifampicina, ya que posee un OR de 1.58 y un intervalo de confianza que no incluye a la unidad. Asimismo, se encontró una razón de prevalencia igual a 1.49; por tanto, es más probable encontrar este SNP en cepas resistentes a rifampicina. El dominio transpeptidasa de PonA1 abarca desde el residuo 310 al 600 (16); por lo que este SNP podría alterar la actividad transpeptidasa de PonA1. Esta alteración generaría alteraciones en la estructura y/o composición del peptidoglicano, lo cual

generaría cambios fenotípicos en cepas con esta mutación. Alguno de estos cambios sería una menor permeabilidad de la rifampicina.

Nuestro estudio analizó bases de datos de genomas obtenidos de cepas con polirresistencia microbiana; por lo que los resultados presentados podrían deberse a factores distintos a la resistencia a rifampicina o a la presencia de SNPs en el gen *rpoB*. Debido a esto nuestro estudio analizó distintas bases de datos de periodos de tiempo distintos para disminuir al máximo este tipo de sesgos. Asimismo, debido a que no es posible determinar la temporalidad de los SNPs estudiados, no es posible determinar si estos SNPs ocurrieron antes o después de la aparición de SNPs en el gen *rpoB*. Razón por la cual no es posible determinar si los SNPs encontrados son factores de riesgo o factores de protección asociados a la resistencia a rifampicina.

## VI CONCLUSIONES

-Se encontraron 6 genes (dnaE, esxP, mbtB, PPE64, Rv1429 y Rv2052c) asociados significativamente a la presencia de al menos un SNP en el gen rpoB, cuyos SNPs serían potenciales mutaciones compensatorias.

- La presencia de al menos un SNP en estos genes (dnaE, esxP, mbtB, PPE64, Rv1429 y Rv2052c) es más probable en cepas con al menos un SNP en rpoB.

-Los SNPs P618R (gen dnaE), T3S (gen esxP), S1000L (gen mbtB), N196Y (gen PPE64), A500T (gen PPE64), A429T (gen Rv 2052c) y R135L (gen Rv 1429) son factores asociados significativamente a la presencia de al menos un SNP en el gen rpoB.

-La presencia de al menos un SNP en el gen ponA1 es un factor asociado significativamente a la resistencia a rifampicina. De igual forma el cambio de prolina a serina en la posición 631 (P631S) y el SNP A516T son factores asociados significativamente a la resistencia a rifampicina.

## VII RECOMENDACIONES

- Analizar más bases de datos de genomas de *M. tuberculosis* para confirmar o refutar los resultados presentados. Ya que las frecuencias de muchos de los SNPs encontrados fueron bajas y se requiere una evidencia estadística robusta para proponer hipótesis que expliquen la relación de estos SNPs con la presencia de SNPs en el gen *rpoB*.
- Evaluar experimentalmente el efecto de los SNPs en PonA1 sobre la actividad transpeptidasa y transglicosilasa.
- Evaluar experimentalmente el efecto de los SNPs en PonA1 sobre la resistencia a rifampicina en *M. tuberculosis*.

## VIII REFERENCIAS BIBLIOGRÁFICAS

1. Gordon SV, Parish T. Microbe Profile: Mycobacterium tuberculosis: Humanity's deadly microbial foe. *Microbiology*. 2018; 164(4): 437-39. Doi: 10.1099/mic.0.000601.
2. Alarcón V, Alarcón E, Figueroa C, Mendoza-Ticona A. Tuberculosis en el Perú: Situación epidemiológica, avances y desafíos para su control. *Rev Peru Med Exp Salud Publica*. 2017;34(2): 299-310. Doi: 10.17843/rpmesp.2017.342.2384.
3. World Health Organization. Global Tuberculosis Report 2018 [Internet]. France: WHO; 2018. [Citado el 25 de enero del 2019]. Disponible en: <http://apps.who.int/iris/bitstream/handle/10665/274453/9789241565646-eng.pdf>.
4. Organización Panamericana de la Salud. Sala de Información y análisis en salud 2019 [Internet]. Lima: OPS; 2019. [Citado el 3 de enero del 2020]. Disponible en: <http://bvspers.paho.org/SIA/index.php/sala-de-situacion#/categorie/1/enfermedades-transmisibles>.
5. Goldstein BP. Resistance to rifampicin: a review. *J Antibiot (Tokyo)*. 2014; 67(9): 625-30. Doi: 10.1038/ja.2014.107
6. Kumar S, Jena L. Understanding Rifampicin Resistance in Tuberculosis through a Computational Approach. *Genomics Inform*. 2014; 12(4): 276-82. Doi: 10.5808/GI.2014.12.4.276
7. Molodtsov V, Scharf NT, Stefan MA, Garcia GA, Murakami KS. Structural basis for rifamycin resistance of bacterial RNA polymerase by the three most clinically important RpoB mutations found in *Mycobacterium tuberculosis*. 2017; 103(6): 1034-45. Doi: 10.1111/mmi.13606

8. Dookie N, Rambaran S, Padayatchi N, Mahomed S, Naidoo K. Evolution of drug resistance in *Mycobacterium tuberculosis*: a review on the molecular determinants of resistance and implications for personalized care. *J Antimicrob Chemother.* 2018; 73(5): 1138-51. Doi: 10.1093/jac/dkx506
9. Handel A, Regoes RR, Antia R. The role of compensatory mutations in the emergence of drug resistance. *PLoS Comput Biol.* 2006;2(10): e137.
10. 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; 44:106–11. Doi:10.1038/ng.1038
11. Li QJ, Jiao WW, Yin QQ, Xu F, Li QJ, 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(25): 2807- 12. Doi:10.1128/AAC.02358-15
12. Campodónico VL, Rifat D, Chuang YM, Ioerger TR, Karakousis PC. Altered *Mycobacterium tuberculosis* Cell Wall Metabolism and Physiology Associated With RpoB Mutation H526D. *Front Microbiol.* 2018; 9:494. Doi:10.3389/fmicb.2018.00494
13. Bisson GP, Mehaffy C, Broeckling C, Prenni J, Rifat D Lun DS, et al. Upregulation of the phthiocerol dimycocerosate biosynthetic pathway by rifampin-resistant, rpoB mutant *Mycobacterium tuberculosis*. *J Bacteriol.* 2012;194(23):6441-52. Doi: 10.1128/JB.01013-12

14. Lahiri N, Shah RR, Layre E, Young D, Ford C, Murray M B, et al. Rifampin resistance mutations are associated with broad chemical remodeling of *Mycobacterium tuberculosis*. *J Biol Chem*. 2016; 291: 14248–256.
15. Du Preez I, Loots du T. Altered fatty acid metabolism due to rifampicin-resistance conferring mutations in the rpoB Gene of *Mycobacterium tuberculosis*: mapping the potential of pharmaco-metabolomics for global health and personalized medicine. *OMICS*. 2012; 16: 596–603.
16. Filippova EV, Kieser KJ, Luan CH, Wawrzak Z, Kiryukhina O, Rubin EJ, et al. Crystal structures of the transpeptidase domain of the *Mycobacterium tuberculosis* penicillin-binding protein PonA1 reveal potential mechanisms of antibiotic resistance. *FEBS J*. 2016; 283(12):2206-18. Doi: 10.1111/febs.13738.
17. 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. Doi: 10.1038/ng.2735.
18. Kieser KJ, Boutte C, Kester JC, Baer CE, Barczak AK, Meniche X, et al. Phosphorylation of the Peptidoglycan Synthase PonA1 Governs the Rate of Polar Elongation in *Mycobacteria*. *PLOS Pathog*. 2015; 11(6): e1005010. Doi: 10.1371/journal.ppat.1005010.
19. Farhat MR, Shapiro BJ, Kieser KJ, Sultana R, Jacobson KR, Victor TC, et al. Genomic analysis identifies targets of convergent positive selection in drugresistant *Mycobacterium tuberculosis*. *Nat Genet*. 2013; 45: 1183–9. Doi: 10.1038/ng.2747.
20. Nguyen L. Antibiotic resistance mechanisms in *M. tuberculosis*: an update. *Arch Toxicol*. 2016; 90(7): 1585-604. Doi: 10.1007/s00204-016-1727-6

21. Brandis G, Wrande M, Liljas L, Hughes D. Fitness-compensatory mutations in rifampicin-resistant RNA polymerase. *Mol Microbiol.* 2012; 85:142–151. Doi:10.1111/j.1365-2958.2012.08099
22. Neri A, Mignogna G, Fazio C, Giorgi A, Schinina ME, Stefanelli P. *Neisseria meningitidis* rifampicin resistant strains: analysis of protein differentially expressed. *BCM Microbiology.* 2010; 10: 246.
23. Casali N, Nikolayevskyy V, Balabanova Y, Harris SR, Ignatyeva O, Kontsevaya I, et al. Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nat Genet.* 2014; 46:279–286. Doi:10.1038/ng.2878.
24. 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; 57:827–832. Doi:10.1128/AAC.01541-12.
25. Lavania M, Singh I, Turankar RP, Gupta AK, Ahuja M, Pathak V, et al. Enriched whole genome sequencing identified compensatory mutations in the ARN polymerase gene of rifampicin-resistant *Mycobacterium leprae* strains. 2018; 11: 169- 75.
26. Gao B, Wang J, Huang X, Sha W & Quin L. The dynamic region of the peptidoglycan synthase gene, Rv0050, induces the growth rate and morphologic heterogeneity in *Mycobacteria*. *Infect Genet Evol.* 2019; 72: 86-92. Doi: 10.1016/j.meegid.2018.12.012.
27. Forrellad MA, Klepp LI, Gioffre A, Sabio y Garcia J, Morbidoni HR, Paz-Santangelo M, et al. Virulence factors of the *Mycobacterium tuberculosis* complex. *Virulence.* 2013; 4(1): 3-66. Doi:10.4161/viru.22329

28. Jackson M. The mycobacterial cell envelope-lipids. *Cold Spring Harb Perspect Med.* 2014;4(10): a021105.
29. Queiroz A, Riley LW. Bacterial immunostat: *Mycobacterium tuberculosis* lipids and their role in the host immune response. *Revista da Sociedade Brasileira de Medicina Tropical.* 2017; 50(1): 9- 18. Doi: 10.1590/0037-8682-0230-2016
30. Delogu G, Sali M, Fadda G. The biology of mycobacterium tuberculosis infection. *Mediterr J Hematol Infect Dis.* 2013;5(1): e2013070. Doi:10.4084/MJHID.2013.070
31. Chai Q, Zhang Y, Liu CH. *Mycobacterium tuberculosis*: An Adaptable Pathogen Associated With Multiple Human Diseases. *Front Cell Infect Microbiol.* 2018; 8:158. Doi:10.3389/fcimb.2018.00158
32. Campbell IA, Bah-Show O. Pulmonary tuberculosis: diagnosis and treatment. *BJM.* 2006; 332(7551): 1194-1197. Doi: 10.1136/bmj.332.7551.1194
33. Romero-Marin MP, Romero-Rondon SK, Sanchez-Robayo KJ, Santamaria-Alza Y, Mendoza-Herrera T, Bolivar-Grimaldos F. Secuelas estructurales y funcionales de tuberculosis pulmonar: una revisión del tema. *Revista Americana de Medicina Respiratoria.* 2016; 6(2): 163-169.
34. Smith T, Wolff KA, Nguyen L. Molecular Biology of Drug Resistance in *Mycobacterium tuberculosis*. *Curr Top Microbiol Immunol.* 2013; 374: 53- 80. Doi: 10.1007/822012279
35. Angst DC, Hall AR. The cost of antibiotic resistance depends on evolutionary history in *Escherichia coli*. *BMC Evol Biol.* 2013; 13:163.

36. Maisnier-Patin S, Berg OG, Lijas L, Andersson DI. Compensatory adaptation of the deleterious effect of antibiotic resistance in *Salmonella typhimurium*. *Mol Microbiol*. 2002; 46(2): 355-66.
37. Sherman DR, Mdluli K, Hickey MJ, Arain TM, Morris SL, Barry CE 3rd, et al. Compensatory *ahpC* gene expression in isoniazid-resistant *Mycobacterium tuberculosis*. *Science*. 1996; 272:1641–1643.
38. Maskell JP, Sefton AM, Hall LM. Multiple mutations modulate the function of dihydrofolate reductase in trimethoprim-resistant *Streptococcus pneumoniae*. *Antimicrob Agents Chemother*. 2001; 45:1104–1108. DOI: 10.1128/AAC.45.4.1104-1108.2001
39. Nagaev I, Björkman J, Andersson DI, Hughes D. Biological cost and compensatory evolution in fusidic acid-resistant *Staphylococcus aureus*. *Mol Microbiol*. 2001; 40:433–439.
40. Xu Z, Zhou A, Wu J, Zhou A, Li J, Zhang S, et al. Transcriptional Approach for Decoding the Mechanism of *rpoC* Compensatory Mutations for the Fitness Cost in Rifampicin-Resistant *Mycobacterium tuberculosis*. *Front Microbiol*. 2018; 9: 2895. Doi: 10.3389/fmicb.2018.02895.
41. Raghavendra T, Patil S & Mukherjee R. Peptidoglycan in Mycobacteria: chemistry, biology and intervention. *Glycoconj J*. 2018; 35(5): 421-32. Doi: 10.1007/s10719-018-9842-7.
42. Phelan JE, Coll F, Bergval I, Anthony RM, Warren R, Sampson SL, et al. Recombination in *pe/ppa* genes contributes to genetic variation in *Mycobacterium tuberculosis* lineages. *BMC Genomics*. 2016; 17:151.

43. Chiliza TE, Pillay M, Naidoo K & Pillay B. Immunoscreening of the *M. tuberculosis* F15/LAM4/KZN secretome library against TB patients' sera identifies unique active –and latent-TB specific biomarkers. *Tuberculosis (Edinb)*. 2019; 115: 161-170. Doi: 10.1016/j.tube.2019.03.005.
44. Sheen P, Requena D, Gushiken E, Gilman RH, Antiparra R, Lucero B, et al. A multiple genome analysis of *Mycobacterium tuberculosis* reveals specific novel genes and mutations associated with pyrazinamide resistance. *BMC Genomics*. 2017; 18: 769. Doi: 10.1186/s12864-017-4146-z.
45. Pagano M & Gauvreau K. *Principles of biostatistics*, 2nd. ed. Belmont CA: Brooks/Cole; 2000.
46. Chhabra G, Dixit A & Garg L. DNA polymerase III  $\alpha$  subunit from *Mycobacterium tuberculosis* H37Rv: Homology modeling and molecular docking of its inhibitor. *Bioinformatics*. 2011; 6(2): 69-73. Doi: 10.6026/97320630006069.
47. Baños-Mateos S, Van Roon AM, Lang UF, Maslen SL, Skehel JM & Lamers MH. High-fidelity DNA replication in *Mycobacterium tuberculosis* relies on a trinuclear zinc center. *Nat Commun*. 2017; 8(11): 855-65. Doi: 10.1038/s41467-017-00886-w.
48. Dwyer DS. Electronic properties of the amino acid side chains contribute to the structural preferences in protein folding. *J Biomol Struct Dyn*. 2001;18(6):881-892. Doi:10.1080/07391102.2001.10506715.
49. Ryndak MB, Singh KK, Peng Z & Laal S. Transcriptional profile of *Mycobacterium tuberculosis* replicating in type II alveolar epithelial cells. *PloS One*. 2015; 10(4): e0123745. Doi: 10.1371/journal.pone.0123745.
50. De Voss JJ, Rutter K, Shroeder BG, Su H, Zhu Y & Barry CE 3rd. The salicylate-derived mycobactin siderophores of *Mycobacterium tuberculosis* are essential for

- grown in macrophages. *Proc Natl Acad Sci USA*. 2000; 97(3): 1252-7. Doi: 10.1073/pnas.97.3.1252.
51. Sharma D, Singh R, Deo N & Bisht D. Interactome analysis of Rv0148 to predict potential targets and their pathways linked to aminoglycosides drug resistance: An insilico approach. *Microb Pathog*. 2018; 121: 179-183. Doi: 10.1016/j.micpath.2018.05.034.
52. Lew JM, Kapopoulou A, Jones LM & Cole ST. TubercuList—10 years after. *Tuberculosis (Edinb)*. 2011; 91(1): 1-7. Doi: 10.1016/j.tube.2010.09.008.
53. Jena L, Kashikar S, Kumar S & Harinath BC. Comparative proteomic analysis of *Mycobacterium tuberculosis* strain H37Rv versus H37Ra. *Int J Mycobacteriol*. 2013; 2(4): 220-6. Doi: 10.1016/j.ijmyco.2013.10.004.