



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

**FACULTAD DE ESTUDIOS SUPERIORES  
ZARAGOZA**

Identificación de variantes genéticas en  
*Mycobacterium tuberculosis* asociadas a la  
resistencia a isoniazida y rifampicina  
mediante secuenciación de nueva  
generación.

**T E S I S**

QUE PARA OBTENER EL TÍTULO DE:  
**QUÍMICA FARMACÉUTICA BIÓLOGA**

**P R E S E N T A:**  
**ARIADNA BETSABE ARROYO VARGAS**

Directora:

Dra. Mirelle Vanessa González Covarrubias

Asesora:

Dra. Raquel Retana Ugalde

2015



México, D. F.



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



## AGRADECIMIENTOS

En primer lugar a la Dra. Vanessa González Covarrubias por permitirme ser parte de su equipo de trabajo, por su confianza y apoyo brindado en todo momento.

Al Dr. Xavier Soberón Mainero por su tiempo y asesoría.

Al Instituto Nacional de Medicina Genómica, en específico a los miembros del Laboratorio de Farmacogenómica: Minerva, Omar y Araceli, muchas gracias por su amistad, asesoría y por hacer agradable mi estancia en el instituto. A la Unidad de Secuenciación: Alfredo, Salvador, Haydee y Julio; muchas gracias por su apoyo para la realización de este proyecto y por los buenos momentos compartidos.

A la Dra. Raquel Retana por sus consejos y motivación.

A los sinodales: Q.F.B Francisco Javier Parada García, M. en C. Ángel García Sánchez y Q.F.B Taide Laurita Arista Ugalde; por su asesoría, el tiempo brindado a la lectura y corrección de éste trabajo.

A la Dra. Claudia Bäcker por la oportunidad brindada, por apoyarme y creer en mí. Así mismo, a mis compañeros del Laboratorio de Micobacterias del InDRE, en especial a la M. en C. Carolina Flores Marroquín. Gracias por todo su apoyo y sus enseñanzas, ya que me fueron de gran ayuda para poder concluir este proyecto.

A la máxima casa de estudios, la Universidad Nacional Autónoma de México, que a lo largo de 8 años se convirtió en mi segundo hogar.

A la FES Zaragoza por proporcionarme todas las herramientas para mi formación académica y profesional. Es un orgullo ser egresada de esta facultad.

Gracias a todos por ser parte de este gran logro académico.

*Ariadna*

## DEDICATORIA

A mi padre Jaime Arroyo Cisneros, por ser mi modelo a seguir, mi consejero, mi apoyo, mi fuerza y mi fortaleza.

A mi madre Blanca Edith Vargas Morán†, por darme la vida, por ser mi motivación para cumplir todas mis metas y buscar ser mejor persona día a día.

A mi segunda madre Edith Marisol Morales Villegas por cuidarme, procurarme y convertirse en mi confidente, mi apoyo, mi respaldo y sobre todo mi mejor amiga.

A mis hermanas y hermanos: Wendy, Reyna, Angel y Kevin, que siempre han estado a mi lado compartiendo un sinfín de momentos inolvidables.

Al resto de mi familia que es parte fundamental de mi vida, en especial a mi abuelito León Arroyo Álvarez.

A Leonardo Airton Hernández Jiménez por todo este maravilloso tiempo que hemos compartido y los que nos faltan por vivir.

---

**CONTENIDO**

<b>ABREVIATURAS</b> .....	VII
<b>ÍNDICE DE FIGURAS</b> .....	VIII
<b>ÍNDICE DE CUADROS</b> .....	X
<b>1. RESUMEN</b> .....	1
<b>2. INTRODUCCIÓN</b> .....	3
<b>3. MARCO TEÓRICO.</b>	
3.1 Tuberculosis .....	4
3.2 Epidemiología de la Tuberculosis.....	5
3.3 Agente etiológico: <i>Mycobacterium tuberculosis</i> .	
3.3.1 Características microbiológicas de <i>Mycobacterium tuberculosis</i> .....	8
3.3.3 Características moleculares de <i>Mycobacterium tuberculosis</i> .....	10
3.4 Tratamiento y fármacorresistencia.....	11
3.4.1 Tratamiento de la Tuberculosis .....	11
3.4.2 Fármacorresistencia.....	12
3.5 Mecanismos moleculares de resistencia.....	15
3.5.1 Resistencia a isoniazida .....	16
3.5.2 Resistencia a rifampicina.....	18
3.6 Métodos diagnósticos.....	20
3.6.1 Métodos microbiológicos.....	20
3.6.1.1 Baciloscopía.....	21
3.6.1.2 Cultivo de micobacterias .....	21
3.6.1.3 Identificación.....	23
3.6.1.4 Pruebas de fármacosensibilidad .....	23
3.6.2 Métodos moleculares.....	25
3.7 Estudios previos en México .....	27
<b>4. PLANTEAMIENTO DEL PROBLEMA</b> .....	29
<b>5. HIPÓTESIS</b> .....	30
<b>6. OBJETIVOS.</b>	
6.1 General.....	31
6.2 Específicos.....	31
<b>7. MATERIAL Y MÉTODOS.</b>	
7.1 Diseño del estudio .....	32
7.1.1 Tipo de estudio .....	32
7.1.2 Universo de estudio .....	32
7.1.3 Criterios de inclusión y exclusión.	
7.1.3.1 Criterios de inclusión.....	32
7.1.3.2 Criterios de exclusión .....	32
7.2 Variables .....	33
7.3 Metodología.....	33

7.3.1 Selección de pacientes y muestras clínicas .....	34
7.3.2 Estrategia experimental.	
7.3.2.1 Análisis microbiológico.....	35
7.3.2.2 Análisis genómico.....	35
7.3.2.2.1 Extracción de ADN.....	35
7.3.2.2.2 Controles de calidad .....	40
7.3.2.2.3 Secuenciación de nueva generación.	
7.3.2.2.3.1 Selección de genes y diseño del sistema de enriquecimiento.....	42
7.3.2.2.3.2 Generación de bibliotecas y proceso de secuenciación .....	42
7.4 Análisis bioinformático.....	48
7.5 Análisis estadístico.....	55
<b>8. RESULTADOS .....</b>	<b>56</b>
<b>9. DISCUSIÓN DE RESULTADOS .....</b>	<b>82</b>
<b>10. CONCLUSIONES.....</b>	<b>90</b>
<b>11. PERSPECTIVAS.....</b>	<b>91</b>
<b>12. REFERENCIAS.....</b>	<b>92</b>
 <b>ANEXOS.</b>	
<b>Anexo 1.</b> Consentimiento informado .....	i
<b>Anexo 2.</b> Cuestionario utilizado para la recolección de información demográfica de los pacientes y resultados de las pruebas microbiológicas .....	vi
 <b>APENDICE.</b>	
<b>Apéndice 1.</b> Enriquecimiento de bibliotecas (SureSelect, Agilent Technologies) y secuenciación por síntesis (Illumina) .....	viii

---

**ABREVIATURAS**

<b>A</b>	Adenina.
<b>Abs</b>	Absorbancia.
<b>ACP</b>	Del inglés: Acyl Carrier Protein.
<b>ADN</b>	Ácido desoxirribonucleico.
<b>ARN</b>	Ácido ribonucleico.
<b>BAAR</b>	Bacilo ácido alcohol resistente.
<b>BK</b>	Baciloscopía.
<b>BLAST</b>	Basic Local Alignment Search Tool.
<b>C</b>	Citosina.
<b>CBK</b>	Cultivo del bacilo de Koch.
<b>CENAPRECE</b>	Centro Nacional de Programas Preventivos y Control de Enfermedades.
<b>CMI</b>	Concentración Mínima Inhibitoria.
<b>CMTB</b>	Complejo <i>Mycobacterium tuberculosis</i> .
<b>DM</b>	Diabetes Mellitus.
<b>E</b>	Etambutol.
<b>Eto</b>	Etionamida.
<b>Flash</b>	Fast Length Adjustment of Short reads.
<b>FR</b>	Fármacorresistencia.
<b>G</b>	Guanina.
<b>H</b>	Isoniazida.
<b>IGV</b>	Integrative Genomics Viewer.
<b>INER</b>	Instituto Nacional de Enfermedades Respiratorias.
<b>INMEGEN</b>	Instituto Nacional de Medicina Genómica.
<b>IS</b>	Inserción.
<b>IUTLD</b>	Del inglés: Internacional Union Against Tuberculosis and Lung Disease.
<b>LiPA</b>	Del inglés: Line probe assay.
<b>MFR</b>	Multifármacorresistencia.
<b>MGIT</b>	Del inglés: Mycobacteria Growth Indicator Tube.
<b>MTB</b>	<i>Mycobacterium tuberculosis</i>
<b>NAD</b>	Nicotinamida Adenina Dinucleótido.
<b>OMS</b>	Organización Mundial de la Salud.
<b>NCBI</b>	National Center for Biotechnology Information.
<b>pb</b>	Pares de bases.
<b>PCR</b>	Del inglés: Polymerase Chainr Reaction.
<b>PFS</b>	Prueba de fármacosensibilidad.
<b>PROVEAN</b>	Protein Variation Effect Analyzer.
<b>Q</b>	Del inglés: Quality.
<b>R</b>	Rifampicina.
<b>RRDR</b>	Del inglés: Rifampicin Resistance Determining Region.
<b>SNG</b>	Secuenciación de Nueva Generación.
<b>SNV</b>	Del inglés: Single Nucleotide Variant.
<b>T</b>	Timina.
<b>TAES</b>	Tratamiento Acortado Estrictamente Supervisado.
<b>TB</b>	Tuberculosis.
<b>VIH</b>	Virus de Inmunodeficiencia Humana.
<b>XFR</b>	Fármacorresistencia extendida.
<b>Z</b>	Pirazinamida.
<b>ZN</b>	Ziehl Neelsen.

---

**ÍNDICE DE FIGURAS**

	Pág.
<b>Figura 1.</b> Etapas de la infección por <i>Mycobacterium tuberculosis</i> .....	5
<b>Figura 2.</b> Tasas estimadas de incidencia de tuberculosis a nivel mundial, 2013.....	6
<b>Figura 3.</b> Tasas de incidencia de tuberculosis en México en todas sus formas de acuerdo a la Entidad Federativa.....	7
<b>Figura 4.</b> Casos de tuberculosis con fármacorresistencia, periodo de 2000-2013.....	8
<b>Figura 5.</b> Factor cordón de <i>Mycobacterium tuberculosis</i> .....	9
<b>Figura 6.</b> Estructura del genoma de <i>Mycobacterium tuberculosis</i> .....	10
<b>Figura 7.</b> Distribución de la subpoblaciones de MTB de acuerdo a su actividad metabólica, propuesta por Mitchison y colaboradores en 1979.....	14
<b>Figura 8.</b> Mecanismo de resistencia a isoniazida.....	17
<b>Figura 9.</b> Estructuras de isoniazida y etionamida.....	18
<b>Figura 10.</b> Mecanismo de acción de la rifampicina.....	19
<b>Figura 11.</b> Fases del diagnóstico microbiológico de la tuberculosis.....	21
<b>Figura 12.</b> Técnica de baciloscopía y su interpretación.....	22
<b>Figura 13.</b> Medios utilizados para el cultivo de micobacterias.....	22
<b>Figura 14.</b> Crecimiento de micobacterias en el medio Lowenstein Jensen..	24
<b>Figura 15.</b> Crecimiento de micobacterias en el medio MGIT.....	24
<b>Figura 16.</b> Identificación de micobacterias.....	24
<b>Figura 17.</b> Prueba de Fármacosensibilidad mediante el sistema BACTEC MGIT 960.....	25
<b>Figura 18.</b> Estrategia experimental.....	33
<b>Figura 19.</b> Comparación de metodologías de extracción de ADN.....	36
<b>Figura 20.</b> Parámetros de secuenciación.....	43
<b>Figura 21.</b> Preparación de bibliotecas previa a la secuenciación masiva de ADN, utilizando tecnología de enriquecimiento.....	44
<b>Figura 22.</b> Reporte de la electroforesis capilar del Bioanalyzer.....	46
<b>Figura 23.</b> Reporte de calidad y parámetros de lecturas crudas evaluadas mediante el software FastQC de la muestra 1TB (R1).....	49
<b>Figura 24.</b> Reporte de calidad y parámetros evaluados después del pre-procesamiento de lecturas, mediante el software FastQC de la muestra 1TB (R1).....	51
<b>Figura 25.</b> Cobertura y profundidad de la región <i>rpoB</i> de la muestra 1TB, mediante el visualizador del software IGV.....	52
<b>Figura 26.</b> Acercamiento a la variante identificada en el gen <i>rpoB</i> de la muestra 1TB, mediante el visualizador del software IGV.....	53
<b>Figura 27.</b> Identificación de la variante identificada en la muestra 1TB, mediante los software SnpEff y SnpSift.....	53
<b>Figura 28.</b> Reporte de la predicción funcional de la variante identificada en el gen <i>rpoB</i> de la muestra 1TB, mediante el software PROVEAN.....	54
<b>Figura 29.</b> Amplificación de genes de <i>Mycobacterium tuberculosis</i> de la muestra 3TB, por los diferentes métodos empleados de extracción.....	59

---

<b>Figura 30.</b>	Comparación de las medias de las concentraciones obtenidas, de acuerdo al método de extracción de ADN empleado.....	60
<b>Figura 31.</b>	Amplificación de genes de <i>Mycobacterium tuberculosis</i> a partir de muestras de expectoración.....	62
<b>Figura 32.</b>	Electroferograma de las diferentes condiciones utilizadas para la optimización de la hibridación con las sondas de captura de genes (SureSelect, Agilent Technologies).....	64
<b>Figura 33.</b>	Profundidad y cobertura del gen <i>katG</i> de <i>Mycobacterium tuberculosis</i> .....	69
<b>Figura 34.</b>	Profundidad y cobertura del gen <i>rpoB</i> de <i>Mycobacterium tuberculosis</i> .....	69
<b>Figura 35.</b>	Profundidad y cobertura del operón <i>mabA/inhA</i> de <i>Mycobacterium tuberculosis</i> .....	69

---

**ÍNDICE DE CUADROS**

	Pág.
<b>Cuadro 1.</b> Esquema del tratamiento primario acortado.....	12
<b>Cuadro 2.</b> Métodos diagnósticos moleculares.....	26
<b>Cuadro 3.</b> Materiales y equipos a utilizar.....	34
<b>Cuadro 4.</b> Mezcla de la reacción de PCR.....	40
<b>Cuadro 5.</b> Secuencias de “cebadores” para la amplificación de genes de <i>Mycobacterium tuberculosis</i> .....	41
<b>Cuadro 6.</b> Genes seleccionados para la secuenciación de nueva generación.....	42
<b>Cuadro 7.</b> Especificaciones y costos del equipo MiSeq-Illumina.....	48
<b>Cuadro 8.</b> Frecuencia de las variables medidas en los pacientes.....	56
<b>Cuadro 9.</b> Descripción de los pacientes que participaron en este estudio..	57
<b>Cuadro 10.</b> Comparación de los métodos de extracción empleados.....	58
<b>Cuadro 11.</b> Resultados de la prueba de ANOVA de un factor.....	59
<b>Cuadro 12.</b> Determinación de subgrupos, prueba de Tukey.....	60
<b>Cuadro 13.</b> Comparación múltiple de los métodos empleados, mediante la prueba de Tukey.....	61
<b>Cuadro 14.</b> Concentración y pureza obtenida de las extracciones de ADN a partir de muestras de expectoración. Método Proteínasa-lisozima/fenol-cloroformo.....	62
<b>Cuadro 15.</b> Modificaciones del protocolo de hibridación con las sondas de captura de genes (SureSelect, Agilent Technologies).....	63
<b>Cuadro 16.</b> Lecturas obtenidas de la secuenciación.....	66
<b>Cuadro 17.</b> Parámetros obtenidos por muestra.....	67
<b>Cuadro 18.</b> Filtrado de las lecturas obtenidas de la secuenciación.....	67
<b>Cuadro 19.</b> Profundidad y cobertura obtenida de cada una de las muestras.....	68
<b>Cuadro 20.</b> Variantes totales identificadas por muestra.....	70
<b>Cuadro 21.</b> Variantes genéticas identificadas en las regiones genómicas de interés.....	80

## 1. RESUMEN

**Introducción:** La tuberculosis es una enfermedad infecciosa, crónica y curable, es causada por *Mycobacterium tuberculosis*. Esta enfermedad se ha ido agravando por el aumento en la incidencia de cepas fármacorresistentes. Este fenómeno se da a partir del fármaco al que se presente la resistencia, siendo la isoniazida y la rifampicina los fármacos más importantes en el tratamiento de primera línea. Estudios previos ya han comenzado a buscar variantes de un sólo nucleótido que están asociadas a la resistencia a estos fármacos, pero sólo se han enfocado a buscar las variantes más frecuentes, dejando a un lado poder identificar nuevas variantes en más genes y realizar su búsqueda en muestras clínicas, para que en un futuro estas bases puedan ser utilizadas para la generación de pruebas diagnósticas rápidas.

**Objetivo:** Identificar variantes genéticas en secuencias completas de los genes *rpoB*, *katG*, *inhA*, *mabA* y la región promotora *mabA/inhA* de *Mycobacterium tuberculosis* asociadas a la resistencia a isoniazida y rifampicina, directamente en muestras de expectoración de pacientes mexicanos mediante secuenciación de nueva generación de ADN.

**Material y métodos:** En este estudio se analizaron seis muestras de expectoración de pacientes con tuberculosis activa, dos muestras con fármacorresistencia y cuatro muestras sin fármacorresistencia, de las cuales se extrajo el ADN total de cada una de ellas. Se prepararon las bibliotecas correspondientes mediante la captura de los genes seleccionados y su posterior secuenciación masiva. El análisis de las lecturas de secuenciación e identificación de variantes genéticas se llevó a cabo mediante programas bioinformáticos, utilizando el genoma de referencia de la cepa *Mycobacterium tuberculosis* H37Rv.

**Resultados:** Se identificaron un total de nueve variantes en los genes seleccionados, de las cuales tres fueron nuevas variantes, una de ellas en la posición 1,674,303 (C/T) en el gen *inhA* que produjo el cambio Ala34Ala en la proteína enoil ACP reductasa y en el gen *katG* en las posiciones 2.155. 025 (G/T) que produjo el cambio Pro29Pro y otra variante en la posición 2,155,346 (C/T) que

generó el cambio Ala256Thr, ambas en la enzima catalasa-peroxidasa, esta última variante con posible asociación a la resistencia a isoniazida. Las variantes que se identificaron en el gen *rpoB* se encontraron en la posición 761,101 (A/C) que produjo el cambio Gln510Pro y en la posición 761,155 (C/T) que dio lugar al cambio Ser531Leu, se sabe que estas variantes afectan a la subunidad  $\beta$  de la ARN polimerasa. Además, se identificó una variante en la región promotor de *mabA/inhA* en la posición 1,673,425 en el codón -15 que ha sido ampliamente reportada. Por último las variantes en las posiciones 2,154,724 (C/A) y 2,155,503 (G/A) ambas en el gen *katG* generaron los cambios Arg453Ley y Thr203Thr, así como la variante en la posición 760,115 (C/T) en el gen *rpoB* que generó el cambio Asp184Asp. Estas variantes ya han sido reportadas y asociadas a la evolución del complejo *Mycobacterium tuberculosis*.

**Conclusiones:** Se lograron identificar variantes en regiones genómicas asociadas a la resistencia a isoniazida y rifampicina, obteniendo una buena concordancia entre el fenotipo según pruebas de fármacosensibilidad y el genotipo revelados por secuenciación, a partir del análisis de la secuenciación de nueva generación.

## 2. INTRODUCCIÓN.

La tuberculosis (TB) es una enfermedad infecciosa, crónica y curable, causada por *Mycobacterium tuberculosis*, agravada en los últimos años por el aumento de la incidencia de cepas fármacorresistentes. Es la segunda causa de muerte en el mundo originada por un sólo agente infeccioso después del virus de la inmunodeficiencia humana (VIH), siendo la TB una enfermedad más común entre hombres en edad reproductiva. Aproximadamente un tercio de la población mundial está infectada con el bacilo sin sospecharlo, pero sólo del 5-10% desarrollarán la enfermedad. Sin embargo, la probabilidad de desarrollar TB es mayor entre las personas con alguna o varias comorbilidades, como es el caso del VIH y la diabetes mellitus (DM).

Actualmente la TB se considera un importante problema de salud pública. En el 2013, la Organización Mundial de la Salud (OMS), a nivel mundial estimó aproximadamente 9 millones de casos nuevos al año. El problema no sólo radica en la prevalencia de esta enfermedad sino también en la resistencia a fármacos específicos que se utilizan para su tratamiento. Dentro del tratamiento existen diversos medicamentos siendo los más importantes los utilizados en la primera línea de elección la isoniazida (H) y la rifampicina (R), la resistencia a estos medicamentos se han asociado a variantes en la secuencia de ciertos genes de *Mycobacterium tuberculosis*. Estas variantes genéticas pueden ser: variante de un sólo nucleótido (SNV, por sus siglas en inglés), inserciones o deleciones.

Estudios previos se han dado a la tarea de buscar SNV's específicos enfocándose sólo a los más frecuentes y sólo a partir de cepas obtenidas de cultivos previos, dejando a un lado el poder determinar variantes en otros genes que puedan explicar la farmacorresistencia (FR). Por lo tanto, es necesario complementar la información existente al identificar exhaustivamente variantes en secuencias de genes completos mediante nuevas tecnologías y así explorar la presencia de otros marcadores de FR. De esta manera se podrán establecer las bases para la generación de pruebas diagnósticas rápidas y específicas de cepas prevalentes en México, para su aplicación directa en muestras clínicas como es la expectoración, que al ser la muestra más representativa y de fácil obtención, agilizaría el diagnóstico y la determinación de FR.

### 3. MARCO TEÓRICO.

#### 3.1 Tuberculosis.

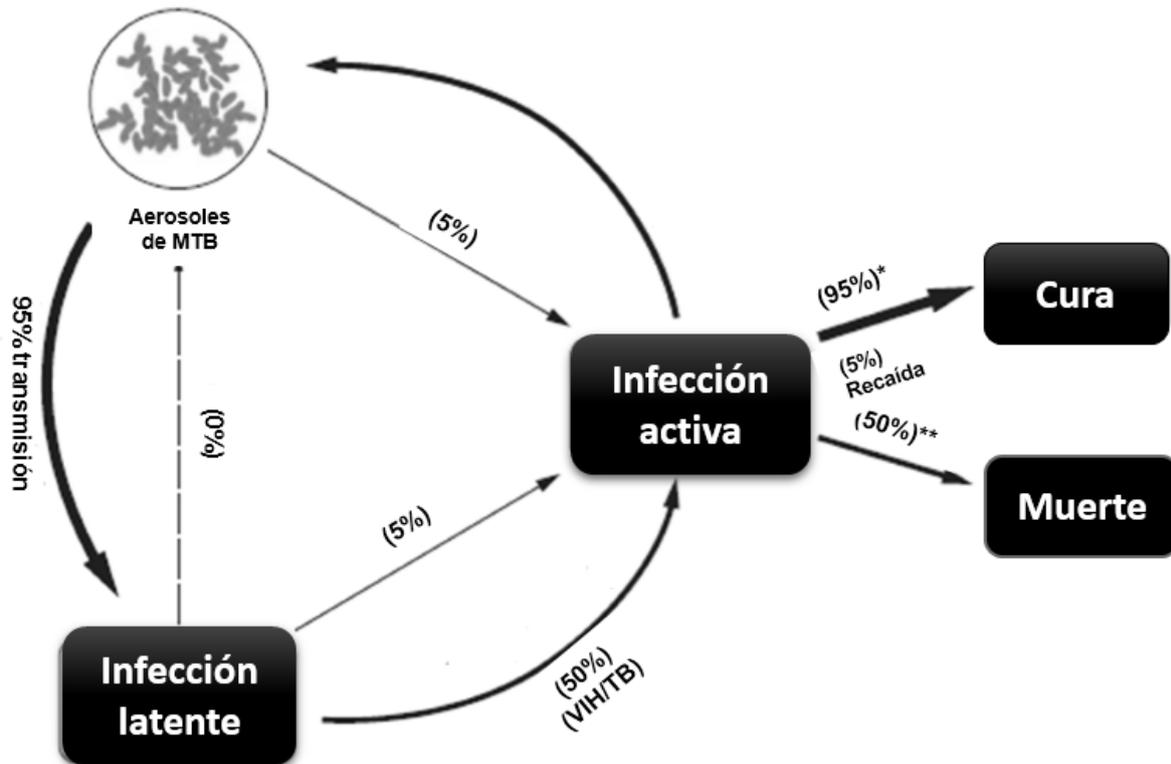
La TB es una enfermedad infecciosa originada taxonómicamente por microorganismos del orden *Actinomycetales*, de la familia *Mycobacteriaceae*, del género *Mycobacterium*, específicamente por el complejo *Mycobacterium tuberculosis* (CMTB), que comprende las especies: *M. tuberculosis* (MTB), *M. bovis*, *M. africanum*, *M. canetti*, *M. microti*, *M. pinnipedi* y *M. caprae*. La enfermedad producida por MTB es la más importante desde el punto de vista sanitario y la causa de la gran mayoría de los cuadros patológicos de esta enfermedad en humanos<sup>1,2</sup>.

El reservorio más importante de MTB y uno de los condicionantes fundamentales de la endemia es el hombre sano infectado, es decir el portador de bacilos vivos en fase latente (TB latente) sin síntomas, que puede desencadenar la enfermedad al presentar deficiencia en su inmunidad, incluso tal vez debido a la presencia de otras comorbilidades como el VIH y la DM que comprometen al sistema inmune. Estos portadores son fuente de infección al momento de enfermar (TB activa), según la localización fisiológica y avance de la patología. La TB pulmonar es la más contagiosa, debido a que los bacilos son eliminados al medio ambiente cuando el individuo tose, habla o escupe, dando lugar al mecanismo de transmisión aerógeno que es el más importante y que causa la mayor parte de los contagios<sup>2-4</sup>.

Cuando la enfermedad se presenta, los síntomas son:

- Tos con flemas durante 3 semanas o más (sintomático respiratorio), que puede acompañarse de sangrado.
- Dolores torácicos.
- Debilidad.
- Pérdida de peso.
- Fiebre.
- Escalofríos.
- Sudoración nocturna.

Existen diversas etapas de la infección por MTB (Figura 1), al inicio de la enfermedad los síntomas pueden ser leves e inespecíficos por muchos meses, por lo que personas que desarrollan la enfermedad tardan en atenderse, dando lugar al contagio de otras personas. Se sabe que por cada individuo que se encuentre enfermo, este contagiará alrededor de 10-15 personas en un año y que 5 de cada 10 enfermos morirán en 5 años si no reciben tratamiento<sup>5-7</sup>.



**Figura 1.** Etapas de la infección de *Mycobacterium tuberculosis*. Transmisión de MTB mediante aerosoles y la progresión de la enfermedad no infecciosa (latente). Una multitud considerable de personas con infección latente puede recaer en TB activa, años después de su primera exposición a la bacteria. La TB latente es comúnmente activada por la supresión inmune, como es el caso del VIH. En los casos de TB sensible a los medicamentos\*, el 95% de los pacientes se recuperan tras el tratamiento, mientras que el 5% tendrán recaídas. Si no son tratados\*\* como resultado se tendrá una alta mortalidad.

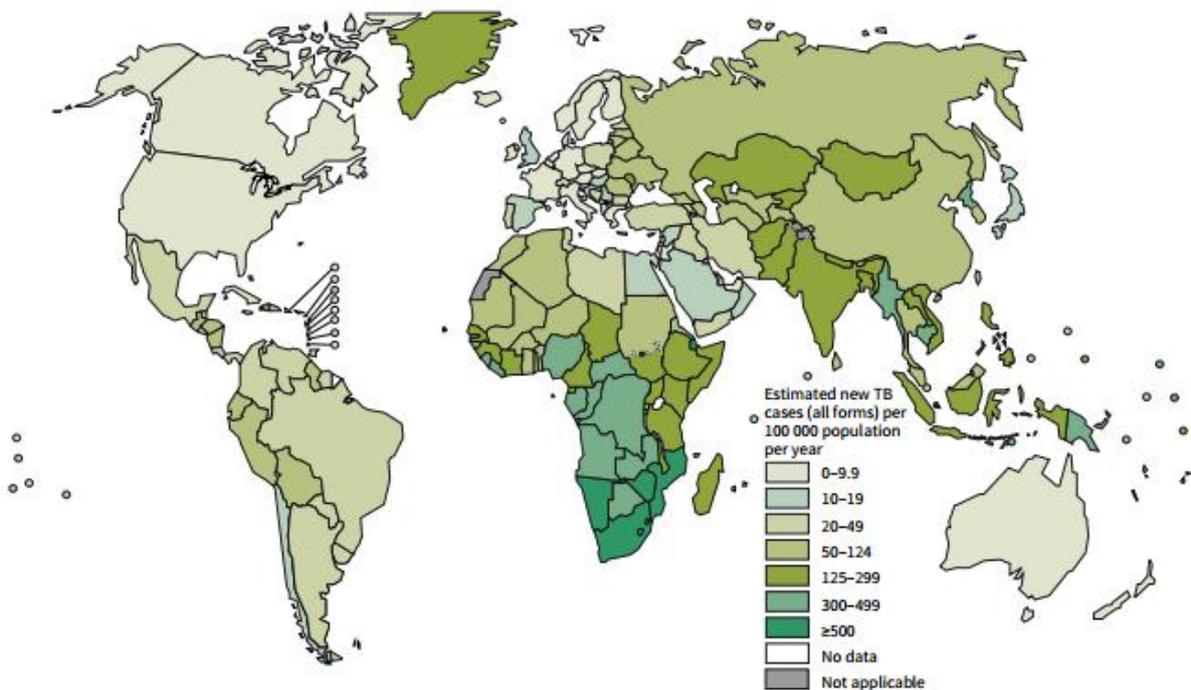
Modificado de: Koul A, Arnoult E, Lounis N, Guillemont J, Andries K. *The challenge of new drug discovery for tuberculosis. Nature.* 2011;469(7331):483-90.

### 3.2 Epidemiología de la Tuberculosis.

La TB es una de las enfermedades más antiguas de la humanidad que aún afecta grandes grupos de población, particularmente de áreas marginadas y grupos vulnerables donde predomina la pobreza, desnutrición y el hacinamiento<sup>5</sup>.

A nivel mundial de acuerdo a la OMS (2013), se estima que 9 millones de personas desarrollaron TB lo que equivale a 126 casos por cada 100,000 habitantes, de los cuales aproximadamente 550,000 eran niños, 3.3 millones eran mujeres y el resto hombres en edad reproductiva. Además, 1.5 millones de personas murieron a causa de esta enfermedad, destacando que 360,000 eran VIH-positivos, lo cual demuestra que los individuos inmunodeprimidos son altamente susceptibles. Otros factores de riesgo asociados son DM, desnutrición y terapias inmunosupresivas, también se han sugerido factores genéticos y ambientales como causa del desarrollo de la enfermedad<sup>4,6</sup>.

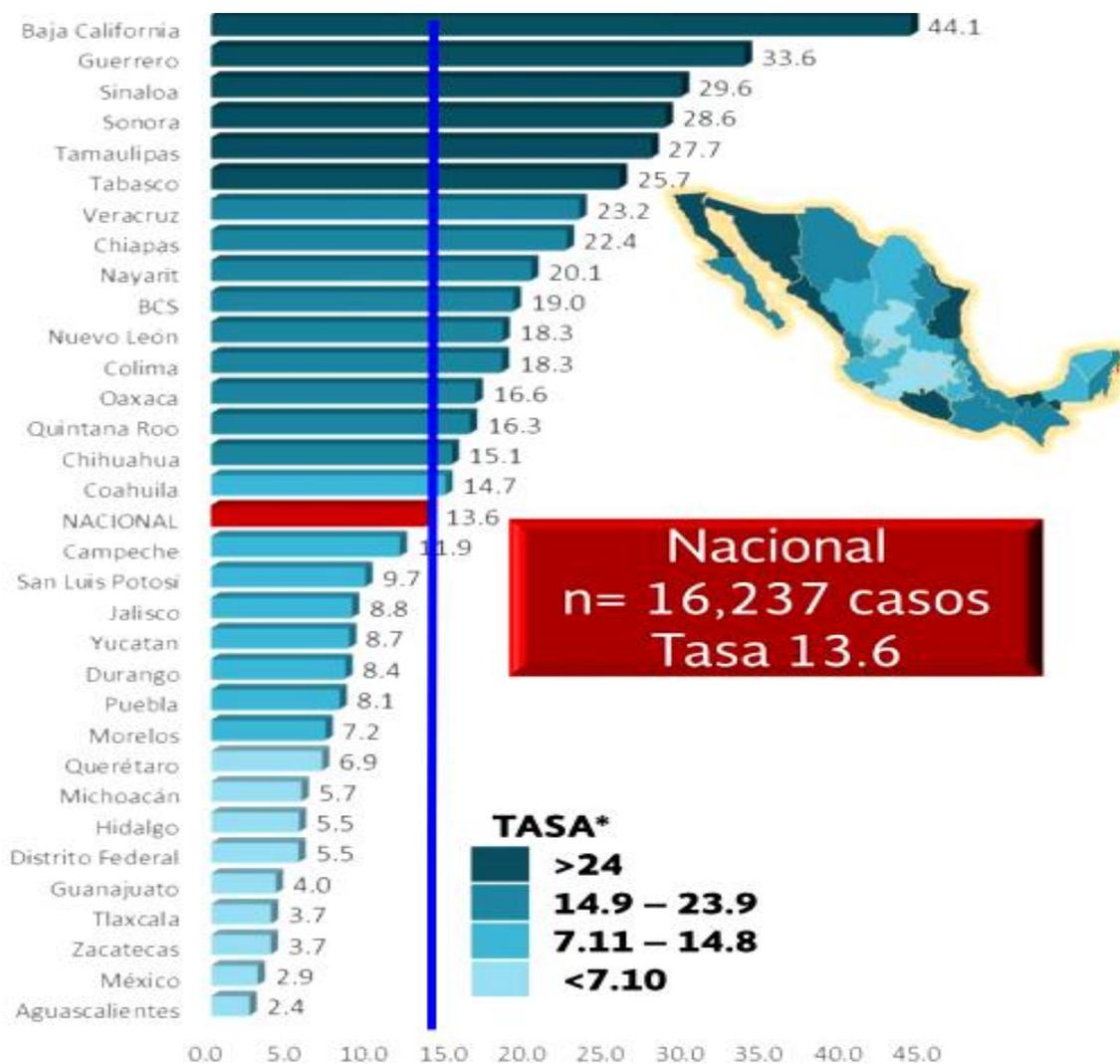
La mayor parte de los casos estimados en el año 2013 se produjeron en Asia (56%) y la Región de África (29%); en menor proporción en la Región del Mediterráneo Oriental (8%), la Región de Europa (4%) y la Región de las Américas (3%). Los países que destacan por tener el mayor número de casos incidentes fueron India, China, Nigeria, Pakistán, Indonesia y Sudáfrica (Figura 2)<sup>6</sup>.



**Figura 2.** Tasas estimadas de incidencia de tuberculosis a nivel mundial, 2013.

*Tomado de: Global tuberculosis control. WHO report 2014.*

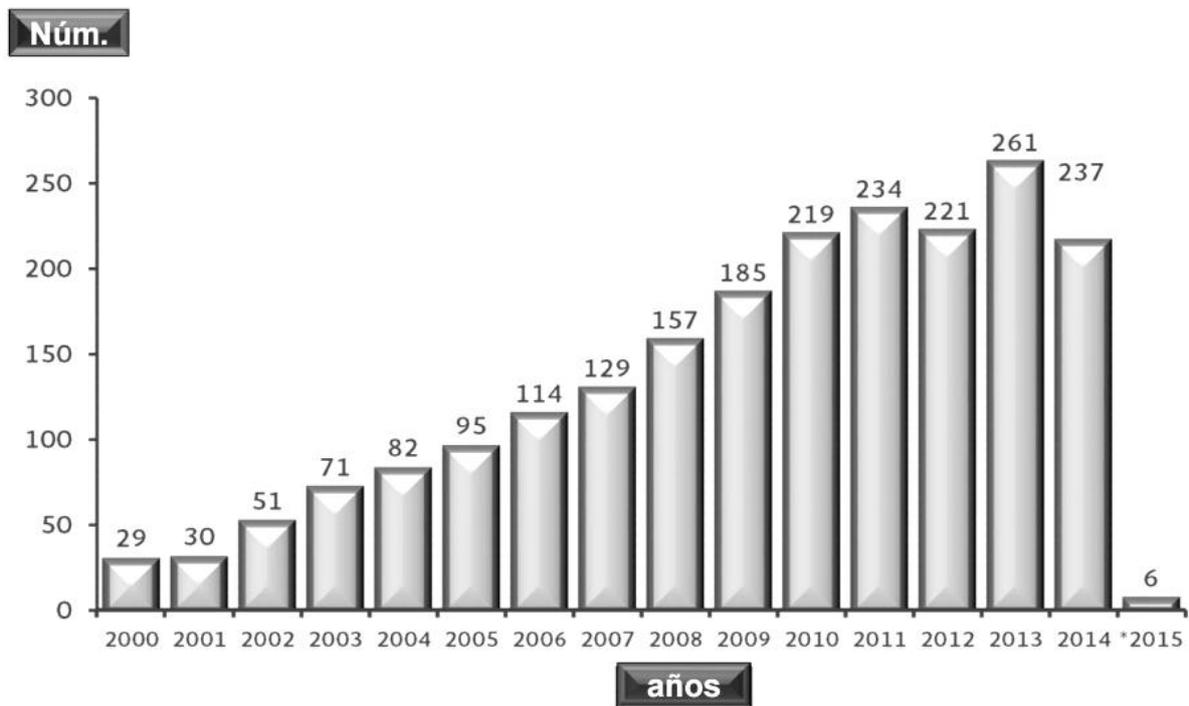
En México, en el 2014 de acuerdo al Centro Nacional de Programas Preventivos y Control de Enfermedades (CENAPRECE), se reportaron 121,881 casos en total de los cuales 20,129 fueron nuevos casos, presentando TB pulmonar (80.7%) y TB extrapulmonar (19.3%); con una tasa de incidencia de 13.6 casos por cada 100,000 habitantes. Los estados con mayor incidencia fueron: Baja California, Guerrero, Tamaulipas, Sonora y Sinaloa, siendo extremo el caso de Baja California con 54.4 casos por cada 100,000 habitantes (Figura 3). Del total de casos reportados el 63.5% tiene una enfermedad asociada, destacando DM (45%), VIH (3%), desnutrición (15%), tabaquismo (5%) y alcoholismo (8%), entre otras<sup>9,10</sup>.



**Figura 3.** Tasas de incidencia de tuberculosis en México en todas sus formas de acuerdo a la Entidad Federativa. Fuente: SUIVE/DGE/SS. 2014. Población CONAPO Proyección 2010-2050. Tasa por 100 000 habitantes.

Tomado de: *Situación actual de la Tuberculosis en México y el mundo. 2015.* Disponible en: <http://www.cenaprece.salud.gob.mx/programas/interior/micobacteriosis/descargas/pdf/1SituacionEpidemiologicaTB.pdf>

Cabe destacar que en los últimos años se ha observado un incremento en el número de casos de pacientes que presentan resistencia a alguno o varios medicamentos antituberculosis (Figura 4), lo cual es alarmante ya que la propagación de estas cepas fármacorresistentes generan un gran problema para la sociedad, desde los altos costos para cubrir tratamientos, medios de subsistencia y las vidas mismas de los pacientes. Además pone en peligro la eficacia de los programas de atención de la salud<sup>10,11</sup>.



**Figura 4.** Casos de tuberculosis con fármacorresistencia, periodo de 2000-2015.

Tomado de: *Situación actual de la Tuberculosis en México y el mundo. 2015.* Disponible en: <http://www.cenaprece.salud.gob.mx/programas/interior/micobacteriosis/descargas/pdf/1SituacionEpidemiologicaT.B.pdf>

### 3.3 Agente etiológico: *Mycobacterium tuberculosis*.

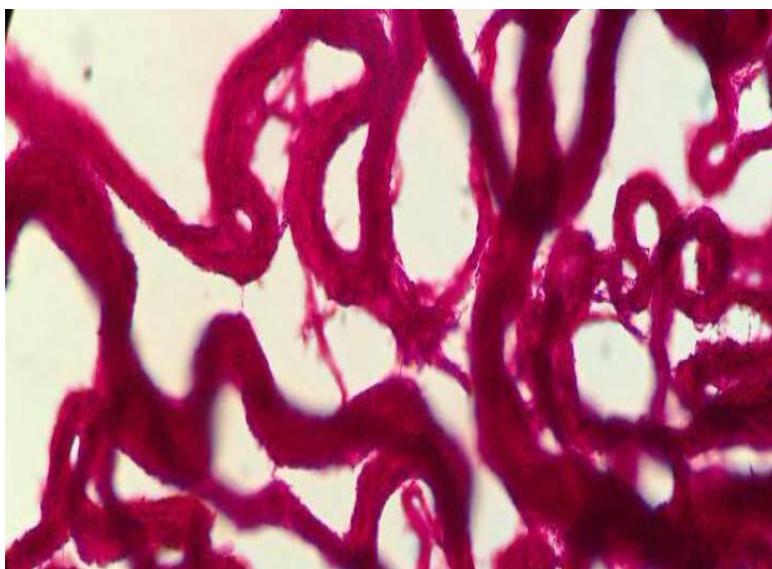
#### 3.3.1 Características microbiológicas.

El agente etiológico MTB, fue descubierto por Roberto Koch en 1882. Es un bacilo aerobio estricto, curvo o ligeramente recto, de 1-10µm de longitud, no esporulado e inmóvil. Es muy resistente al frío, congelación, desecación y desinfectantes químicos, pero sensible al calor y a la luz solar. Tiene altas capacidades de adaptación, es capaz de poner en juego diferentes vías fisiológicas para sobrevivir

en condiciones adversas. Sus condiciones óptimas de crecimiento se manifiestan a un pH de 7.0-7.4, bajo una presión de oxígeno de 100-140 mm Hg<sup>2,12</sup>.

El bacilo de la TB puede crecer utilizando glicerol como fuente de carbono, asparagina e iones de amonio como fuente de nitrógeno y micronutrientes. Otro componente clave para el desarrollo del bacilo es la albúmina, habitualmente incorporada en los medios de cultivo con el agregado de huevos de gallina o seroalbúmina bovina. Algunos medios sintéticos contienen biotina y catalasa para estimular el desarrollo de bacilos dañados<sup>12</sup>.

En su pared celular tiene un alto contenido de lípidos, de particular relevancia son los ácidos micólicos que lo hacen un bacilo ácido alcohol resistente (BAAR) al ser teñidos por la técnica de Ziehl Neelsen (ZN). En cultivo, MTB secreta un carbohidrato llamado trehalosa 6'6 dimicolato denominado "factor cordón" (Figura 5), este se asocia al alineamiento en paralelo de filas de bacilos (formación de "cordones"), que es una de las características de las cepas virulentas de MTB<sup>2,12,13</sup>.



**Figura 5.** Factor cordón de *Mycobacterium tuberculosis*.  
Observación a 100x, tinción Ziehl Neelsen.

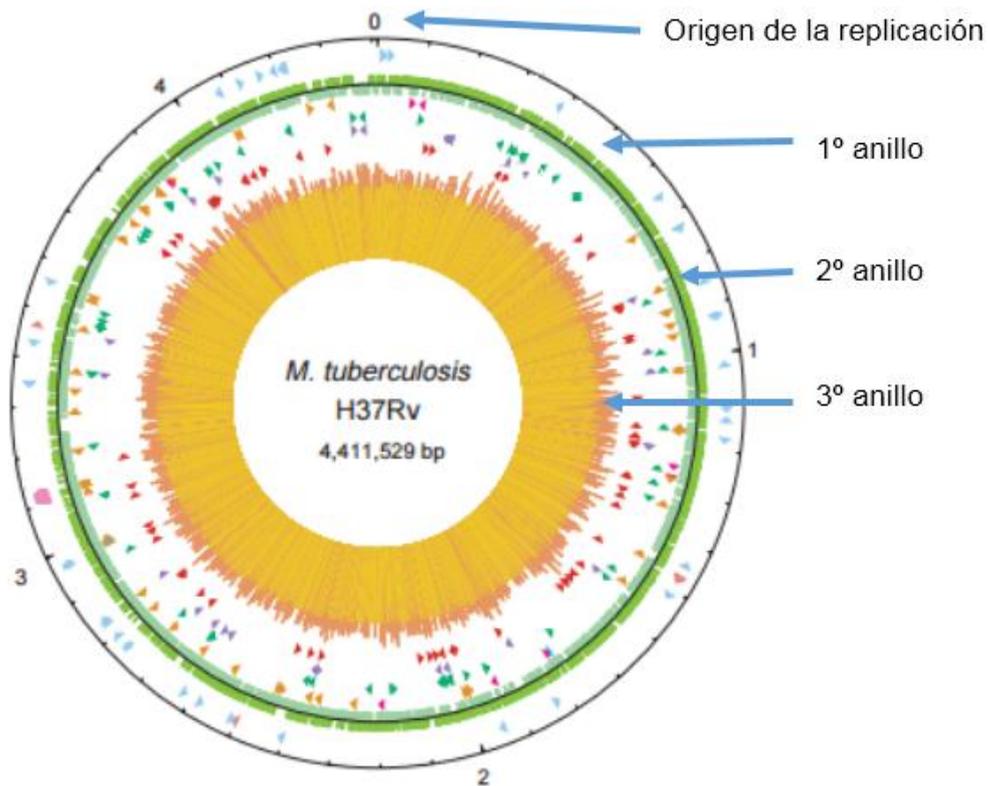
*Fuente: Cortesía del INER.*

En cuanto a su reproducción MTB se divide cada 12 a 24 horas, este ritmo es muy lento en comparación con la de la mayoría de bacterias cultivables cuya división sucede en promedio de 30 minutos a un máximo de 2 horas, esto representa una

limitante para realizar su cultivo, identificación y pruebas necesarias para el diagnóstico de la TB<sup>2,12,13</sup>.

### 3.3.2 Características moleculares de *Mycobacterium tuberculosis*.

El genoma de MTB (cepa H37Rv) fue secuenciado por Cole y colaboradores<sup>14</sup> en 1998, definiendo un cromosoma circular de 4.4 megabases, con alto contenido de guanina (G) y citosina (C) alrededor del 65%; tiene 3,924 inicios de transcripción, de los cuales el 90% codifica para proteínas y 139 genes sintetizan ácido ribonucleico (ARN). Se conocen hasta ahora el 52% de las posibles funciones supuestas en estos genes. Entre el 3 y 4% del genoma está conformado por secuencias de inserción (IS), siendo la región génica IS6110 la más empleada para la determinación del CMTB por ser altamente polimórfica<sup>14-18</sup>.



**Figura 6.** Estructura del genoma de *Mycobacterium tuberculosis*. En el exterior del círculo se muestra una escala, donde 0 representa el origen de replicación. El primer anillo denota las posiciones de los genes de ARN estables (ARN de transferencia son azules, los demás son de color rosa); el segundo anillo interior muestra la secuencia de codificación por capítulos (en sentido 5'→3', de color verde oscuro; en antisentido 3'→5', de color verde claro); el tercer anillo representa ADN repetitivo, el histograma (centro) representa el contenido de G + C, con < 65 % de G + C en amarillo, y > 65 % de G + C en rojo.

Modificado de: Cole S, et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*. 1998;393(6685):537-44.

### **3.4 Tratamiento y fármacorresistencia.**

El objetivo del tratamiento es impedir la transmisibilidad de MTB, logrando la curación del paciente, evitando complicaciones y la muerte. Para lograr el control de la TB en 1995 la OMS introdujo la estrategia de “Tratamiento Acortado Estrictamente Supervisado” (TAES) propuesto por la Unión Internacional para la Lucha contra la Tuberculosis y Enfermedades Pulmonares (IUATLD, por sus siglas en inglés)<sup>6,19,20</sup>.

El TAES se basa principalmente en la administración supervisada del tratamiento, el cual ha sido parte importante de la mejora en el pronóstico de la TB, aunque todavía falta para la erradicación de esta enfermedad, ya que de acuerdo a “The Stop TB Strategy” de la OMS plantea que para el 2050 se busca eliminar a la TB como problema de salud pública, es decir tener una incidencia de 1 caso por millón de habitantes<sup>6,20</sup>.

#### **3.4.1 Tratamiento de la Tuberculosis.**

El tratamiento de la TB inició con el descubrimiento de la estreptomicina en 1943 por Waksman y Schatz, se administró a partir de 1944, no obstante este hallazgo y el mal uso de la estreptomicina dio lugar a la resistencia de dicho medicamento. Posteriormente comenzaron con los ensayos con el ácido para-aminosalicílico en 1945. En 1952, se introdujeron a la H y a la pirazinamida (Z), posteriormente en 1966 se introdujo el etambutol (E) y en 1968 se logra el aislamiento de la R, un derivado semisintético de la rifamicina, la introducción de este potente medicamento en los esquemas antituberculosis permitió disminuir de forma notable la duración de la quimioterapia que iba de los 12 a 19 meses. De esta manera al final de la década de los 70's se establece el tratamiento primario acortado<sup>19,20</sup>.

Actualmente el tratamiento primario acortado (Cuadro 1) tiene duración de 6 meses que se administra en aproximadamente 25 semanas hasta completar 105 dosis, dividido en dos etapas: una fase intensiva (bactericida) que consiste en 60 dosis de H-R-Z-E (diariamente de lunes a sábado) y una fase de sostén o mantenimiento (esterilizante), de 45 dosis de HR administradas 3 veces a la semana<sup>1</sup>.

**Cuadro 1.** Esquema del tratamiento primario acortado.

FASES	INDICACIONES		
<b>FASE INTENSIVA:</b>	Diario, de lunes a sábado, hasta completar 60 dosis, administración en una toma.		
<u>Fármacos</u>	<u>Dosis</u>	<u>Interacciones y efectos adversos</u>	
- Rifampicina	600 mg	<ul style="list-style-type: none"> <li>• Fenitoína</li> <li>• Neuritis</li> <li>• Hepatitis</li> </ul>	<ul style="list-style-type: none"> <li>• Hipersensibilidad</li> <li>• Síndrome lupoide</li> </ul>
- Isoniazida	300 mg	<ul style="list-style-type: none"> <li>• Inhibe anticonceptivos orales</li> <li>• Hepatitis</li> <li>• Reacción febril</li> </ul>	<ul style="list-style-type: none"> <li>• Púrpura</li> <li>• Hipersensibilidad</li> <li>• Intolerancia oral</li> </ul>
- Pirazinamida	1,500 a 2,000 mg	<ul style="list-style-type: none"> <li>• Hiperuricemia</li> <li>• Hepatitis</li> <li>• Vómitos</li> </ul>	<ul style="list-style-type: none"> <li>• Artralgias</li> <li>• Hipersensibilidad cutánea</li> </ul>
- Etambutol	1,200 mg	<ul style="list-style-type: none"> <li>• Neuritis óptica</li> <li>• Discriminación rojo- verde</li> </ul>	
<b>FASE DE SOSTÉN:</b>	Intermitente, tres veces por semana, lunes, miércoles y viernes, hasta completar 45 dosis. Administración en una toma.		
<u>Fármacos</u>	<u>Dosis</u>		
- Isoniazida	800 mg		
- Rifampicina	600 mg		

Cuando el tratamiento de primera línea falla, se deben utilizar otros medicamentos denominados como segunda línea, los cuales se dividen en tres grupos farmacológicos:

- Aminoglucósidos (inyectables): kanamicina, amikacina, capreomicina.
- Fluoroquinolonas: ofloxacino, levofloxacino, moxifloxacino.
- Bacteriostáticos orales: etionamida, protionamida, cicloserina y terizidona.

### 3.4.2 Fármacorresistencia.

Entre los factores que contribuyen al resurgimiento TB a nivel mundial está el aumento en la resistencia a los medicamentos que se han utilizado por décadas con relativo éxito<sup>3,22</sup>. El fenómeno de FR se da a partir de los principales tipos de resistencia que presente la cepa infectante, ya sea la natural que da lugar a la generación de mutaciones espontáneas al azar dentro de una población bacilar; o

cuando se presenta una resistencia adquirida, es decir cuando el tratamiento ha fallado o la inicial, que depende del contagio de una cepa ya resistente<sup>3,5,22</sup>.

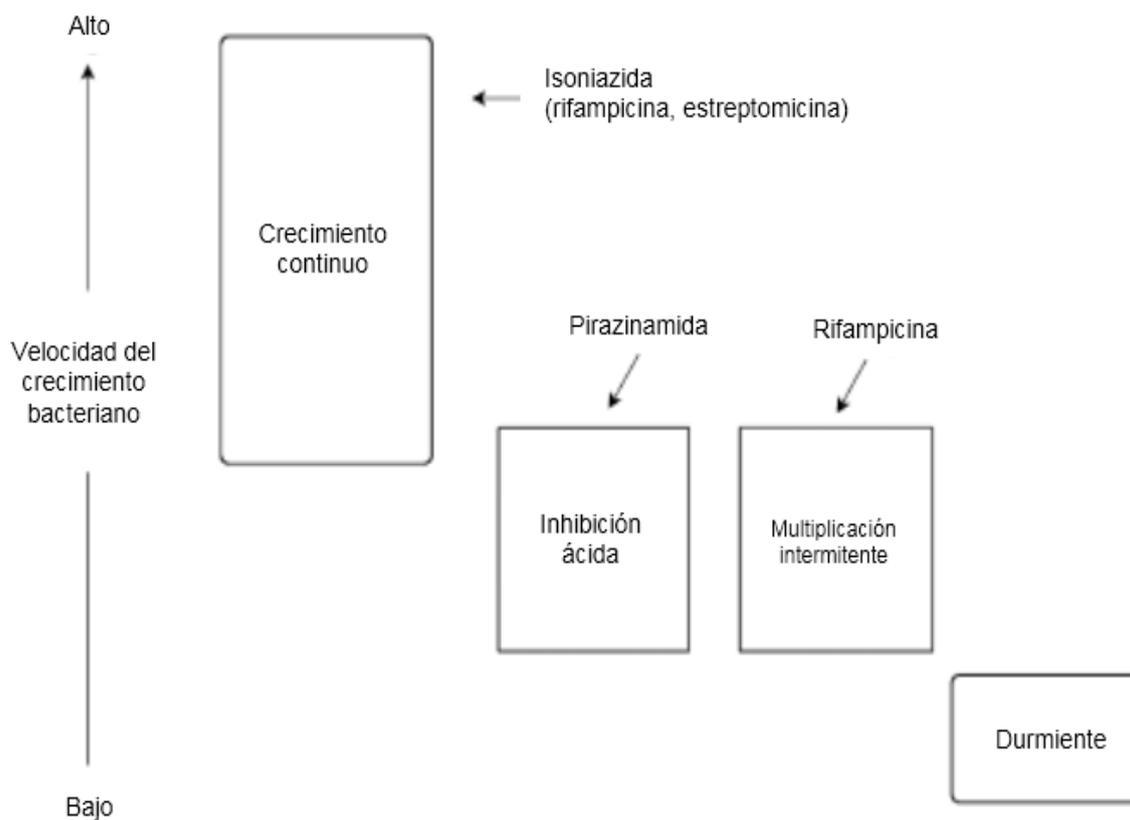
La clasificación de la FR se realiza según el número y tipo de fármaco al que se presente la resistencia:

- Monorresistencia: resistencia a sólo un fármaco de primera línea.
- Polirresistencia: resistencia a 2 o más fármacos de primera línea siempre y cuando no sean H y R simultáneamente.
- Multifármacorresistencia (MFR): resistencia simultánea a H y R.
- Extrema fármacorresistencia o fármacorresistencia extendida (XFR): es aquella que además de tener resistencia a H y R presenta resistencia a una quinolona y un aminoglucósido de segunda línea<sup>1,23</sup>.

Cabe destacar que a nivel mundial, se estima que 480,000 personas desarrollaron TB-MFR que corresponde al 5% del total de casos notificados y el 9% de personas con TB-MDR son también XFR (2014). En México del 2000 al 2015, de acuerdo al CENAPRECE se reportaron 1193 casos de TB con FR, de los cuales 163 presentaron monorresistencia, 113 polirresistencia, 898 MFR y 19 XFR<sup>6,10</sup>.

Otro problema es el hecho de que la infección no es causada por una población de MTB específica, si no por un conjunto de subpoblaciones bacilares que se denominan de acuerdo a su estado metabólico (Figura 7) y su localización fisiológica:

- a) Metabólicamente activas, en crecimiento continuo: constituyen la mayoría de bacilos, con una población estimada de  $10^7$  a  $10^9$ . Están situados extracelularmente en el interior de las paredes de la lesión cavitaria del pulmón, donde las condiciones de pH y presión de oxígeno son ideales para su reproducción. Esta población es la que determina los fracasos farmacológicos y los contagios, ya que son expulsados con la tos. Sin embargo, son rápidamente eliminados por la acción bactericida de la H.



**Figura 7.** Distribución de la subpoblaciones de MTB de acuerdo a su actividad metabólica, propuesta por Mitchison y colaboradores en 1979.

Modificado de: Caminero J. *Guidelines for clinical and operational management of drug-resistant tuberculosis*. Paris: IJTL. 2013. CDC, tuberculosis. Disponible en: <http://www.cdc.gov/tb/topic/basics/default.htm>

- b) En fase inhibida ácida: Esta subpoblación es poco numerosa, está constituida aproximadamente  $10^3$  a  $10^5$  bacilos que se localizan intracelularmente en los fagolisosomas de los macrófagos. Su crecimiento se ve inhibido por el medio ambiente ácido y por la deficiente oxigenación. Difícilmente estos bacilos pueden ser atacados por la acción de los fármacos. El fármaco más activo para esta población es la Z, ya que la condición del pH ácido (5.4 a 6.6) aumenta la actividad de este fármaco.
- c) Multiplicación intermitente: Es una población de aproximadamente  $10^3$  a  $10^5$  bacilos localizados preferentemente en el *caseum*, es decir la caseificación o necrosis de un segmento pulmonar, que se caracteriza por la formación de una materia gris amarillenta, compacta y granulosa, donde el pH es neutro.

Estos bacilos presentan largos periodos latentes, con actividad metabólicamente ocasional de pocas horas. En conjunto con los bacilos en fase inhibida ácida son los responsables de las recaídas tras el término del tratamiento. El fármaco de elección en estas circunstancias es la R por su actividad esterilizante.

- d) Persistente o durmiente: Es aquella población que permanece en latencia por periodos largos, no tienen actividad metabólica, por lo que no pueden ser destruidos por fármacos y sólo los mecanismos de defensa individuales serían capaces de controlarlos. Teóricamente esta población sería una de las responsables de la recaída en paciente con inmunodeficiencia severa<sup>2,3,18</sup>.

Esta identificación metabólica de MTB ha tenido un significado fundamental en la selección de los actuales esquemas de tratamiento y por ende se emplea una combinación estratégica de fármacos antituberculosis para combatir a la infección de forma eficiente. Se debe tomar en cuenta que existen de forma natural bacilos que son fármacorresistentes, debido a mutaciones intrínsecas, es decir que tienen una variante en un gen específico o en varios de ellos, que dan lugar a la FR. La aparición de esta resistencia también puede estar asociada a la presencia de mutaciones durante la administración de un determinado fármaco que genera una presión selectiva positiva que favorece al crecimiento de los bacilos mutantes resistentes dando lugar a la progresión de la resistencia a ese fármaco, lo que obliga a tener que buscar una alternativa para el tratamiento<sup>25-28</sup>.

### **3.5 Mecanismos moleculares de resistencia.**

La mayoría de bacterias que se convierten en resistentes, se debe a la adquisición de genes procedentes de plásmidos o de trasposones mediante transferencia horizontal de genes. En MTB no sucede este fenómeno, sólo se presenta una resistencia genética a un medicamento, a partir de mutaciones cromosómicas espontáneas que se dan a una frecuencia de  $10^{-3}$  a  $10^{-9}$  por replicación<sup>3</sup>.

### 3.5.1 Resistencia a isoniazida (H).

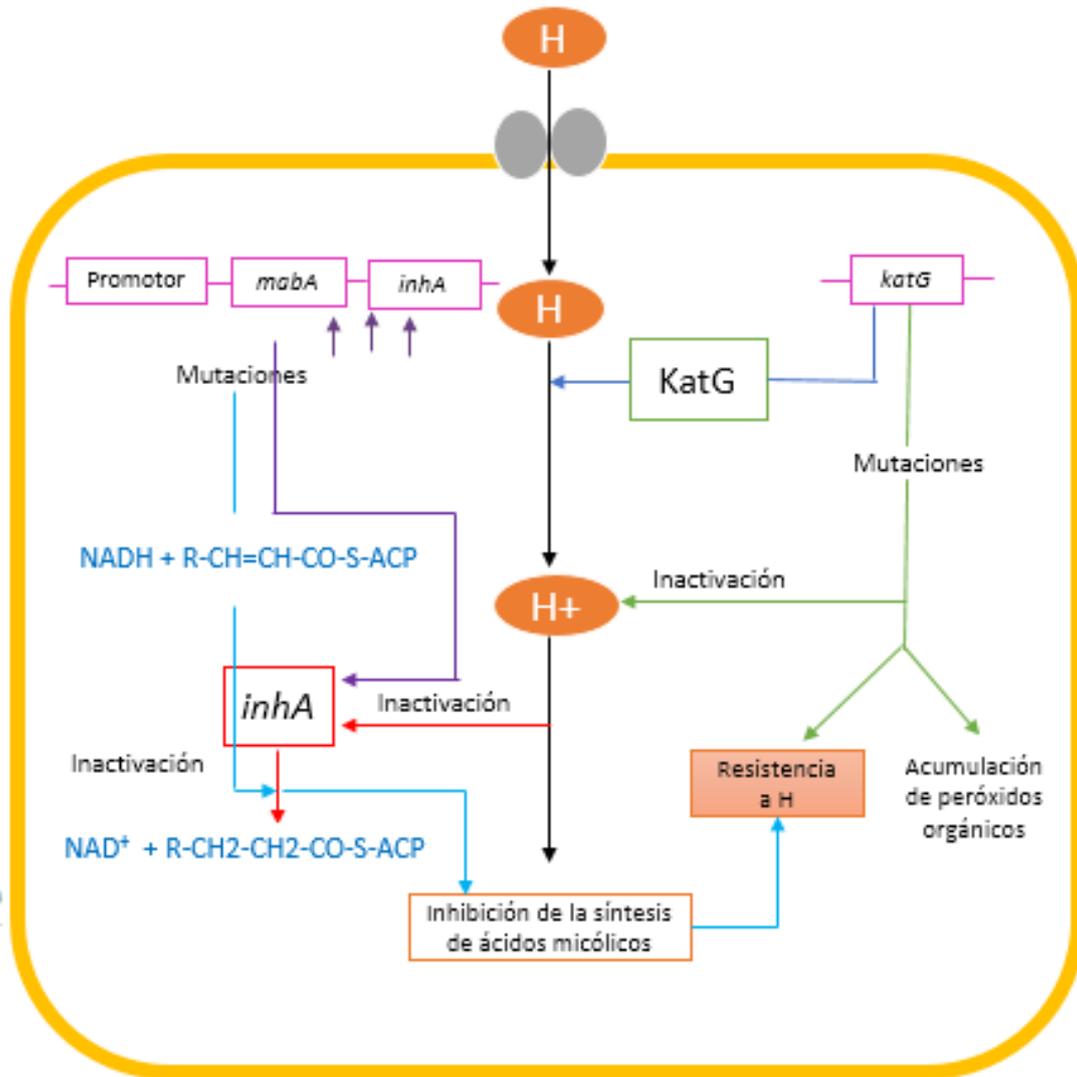
La H es un profármaco antituberculosis de primera línea ampliamente utilizado desde su descubrimiento, pues tiene un efecto bactericida. Ha sido la piedra angular de todos los regímenes efectivos para el tratamiento de la enfermedad y de la infección latente, es únicamente activa contra bacilos metabólicamente activos, tiene una concentración mínima inhibitoria (CMI)= 0.02 – 0.2 µg/mL. Su mecanismo de acción no está del todo claro. La resistencia *in vitro* a H es más frecuente que a la mayoría de los otros medicamentos, con una frecuencia de 1 resistente por cada 10<sup>5</sup> a 10<sup>6</sup> bacilos<sup>3,29</sup>.

El mecanismo de acción de la H (Figura 8), se da a partir de la activación del profármaco mediante la enzima catalasa-peroxidasa (*KatG*), codificada por el gen *katG* dando lugar al radical o anión acil isonicotínico (intermediario electrófilico), permitiendo que MTB sea susceptible a la acción de radicales de oxígeno reactivo y otros elementos tóxicos<sup>30</sup>. Se cree que el objetivo primario de la inhibición de la H es la enzima *inhA*, que codifica una enoil ACP (proteína acarreadora de acilos, por sus siglas en inglés) reductasa, involucrada en la elongación de los ácidos grasos en la síntesis de los ácidos micólicos. El anión acil isonicotínico reacciona con la NAD<sup>+</sup> (Nicotinamida Adenina Dinucleótido) para formar el complejo H-NAD que inhibe a la enzima *inhA*<sup>30,31</sup>.

En cepas resistentes a H se han observado variantes genéticas como SNV's, inserciones, deleción parcial y más raramente la deleción del gen completo. La mutación en el gen *katG* es el mecanismo principal de resistencia a H y la mutación más frecuente se encuentra en el codón de la posición 315, con los siguientes cambios de aminoácidos: Ser315Thr, Ser315Asn y Ser315Arg; que se presentan en una frecuencia del 50-90% de los casos resistentes a este fármaco<sup>29,30</sup>.

La resistencia a H puede también tener lugar a través de mutaciones en la región del promotor del operón *mabA/inhA*, siendo la más frecuente C-15T que se presenta del 5-43% en las cepas resistentes a H. Esta última mutación causa una sobreexpresión de *inhA* generando grandes cantidades de enoil ACP reductasa que conferirá resistencia a MTB al superar el poder inhibitorio de la H o a través de mutaciones en

el sitio activo de *inhA*, disminuyendo así la afinidad para formar el complejo H-NAD y por ende la síntesis de ácidos micólicos<sup>29,30,32</sup>.



**Figura 8.** Mecanismo de resistencia a isoniazida.

Modificado de: Palomino JC, Martin A. Drug Resistance Mechanisms in *Mycobacterium tuberculosis*. *Antibiotics*. 2014;3(3):317-40.

Las mutaciones en el gen *inhA* o en la región de su promotor (*mabA/inhA*) se asocian usualmente a un bajo nivel de resistencia y son menos frecuentes que las mutaciones en *katG*. La resistencia a H que alberga mutaciones en el *inhA* podría tener mutaciones adicionales en el *katG*, lo que le confiere niveles mayores de resistencia. Las mutaciones en *inhA* no sólo causan resistencia a H, sino también confieren resistencia cruzada al medicamento estructuralmente relacionado (Figura 9), la etionamida (Eto)<sup>3,30</sup>.

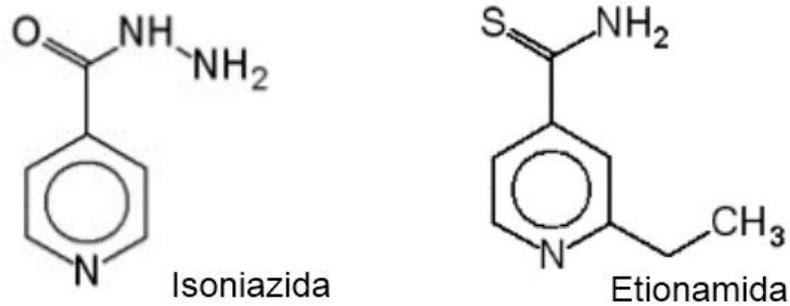


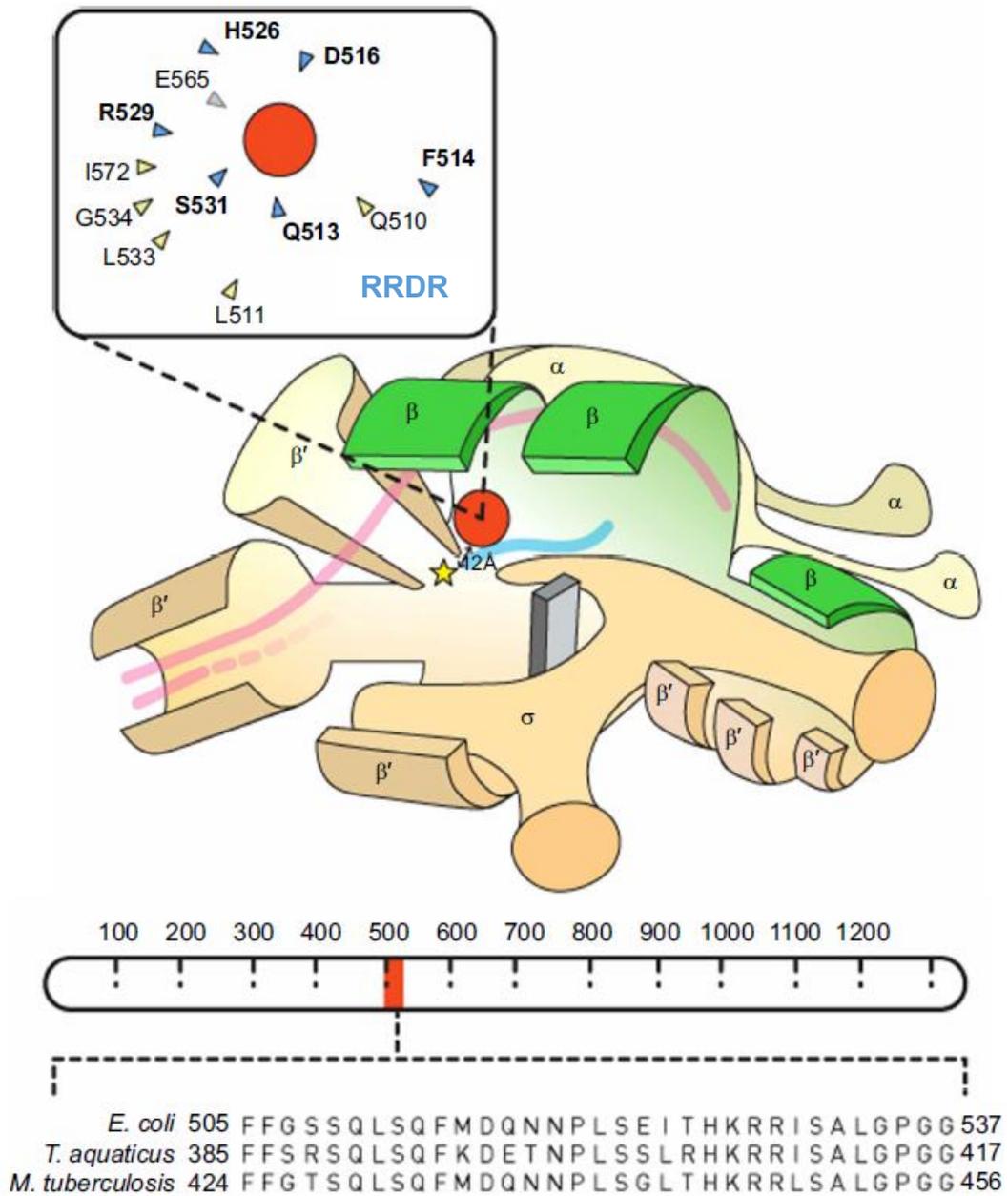
Figura 9. Estructuras de isoniazida y etionamida.

### 3.5.2 Resistencia a rifampicina (R).

La R, también un fármaco de primera línea, es bactericida y se ha empleado desafortunadamente de forma desmedida generando un incremento de cepas resistentes. Es activa contra bacilos en fase de crecimiento y en fase estacionaria con actividad metabólicamente baja, esto lo hace altamente esterilizante *in vivo* y acorta el tratamiento de 12-19 meses a solamente 9 meses<sup>3,30</sup>.

La R interfiere con la síntesis del ARN al unirse a la subunidad  $\beta$  de la ARN polimerasa. La ARN polimerasa es un oligómero que consiste de un núcleo enzimático formado por cuatro cadenas  $\alpha 2\beta\beta'$  en asociación con la subunidad  $\sigma$  para iniciar específicamente la transcripción de los promotores (Figura 10). El sitio de unión de la R está ubicado antes del centro catalítico y bloquea físicamente la elongación de la cadena de ARN, es decir impide la síntesis de ARN y por lo tanto de las proteínas. La resistencia a la R tiene lugar a una frecuencia de 1 por cada  $10^7$  a  $10^8$  bacilos<sup>8,33,34</sup>.

La resistencia a R se asocia a mutaciones en una región “hot spot” de 81 pares de bases (pb), del codón 507 al 533 (de acuerdo al sistema de numeración de *Escherichia coli*) del gen *rpoB*, también llamada región determinadora de resistencia a R (RRDR, por sus siglas en inglés), aunque en los últimos años se han observado mutaciones fuera de esta región, lo cual propone la existencia de mecanismos adicionales generadores de resistencia que sugiere el análisis de una región más amplia o analizar el gen completo y no sólo para este fármaco sino para todos los demás<sup>29,30</sup>.



**Figura 10.** Mecanismo de acción de la rifampicina. Representación esquemática se muestra la holoenzima ARN polimerasa, la subunidad  $\beta$  codificada por *rpoB* esta resaltada en verde, una estrella amarilla representa el sitio activo ARN polimerasa y un círculo rojo indica la molécula de rifampicina que se acerca a la fracción 12A del sitio activo para la inhibición la transcripción. El ADN bicatenario se representa por líneas de color rosa y, una vez desenrollada, sólo se muestra la plantilla de ADN, con color azul la creciente cadena de ARN.

Tomado de: Varma-Basil M, El-Hajj H, Colangeli R, Hazbón MH, Kumar S, Bose M, et al. Rapid detection of rifampin resistance in *Mycobacterium tuberculosis* isolates from India and Mexico by a molecular beacon assay. *J Clin Microbiol.* 2004;42(12):5512-6

Las mutaciones en el *rpoB* resultan generalmente en un alto nivel de resistencia (CMI >32µg/mL) incluyendo a todas las rifamicinas, estas mutaciones son: 531 (Ser-Leu, Thr), 526 (His- Asp, Leu, Asn, Pro, Gln, Arg), 516 (Asp-Val), se presentan del 95-98% de los casos en cepas resistentes a la R. Sin embargo, existen mutaciones específicas en los codones 516, 518, 522, 529 y 533 asociadas a un nivel menor de resistencia<sup>30</sup>.

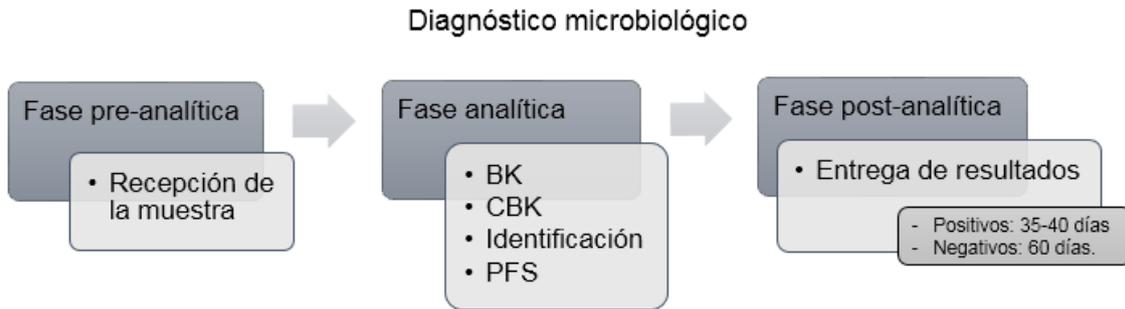
Un hecho importante es que 90% de las cepas que son resistente a R lo son también para H, sobre todo en regiones con alta prevalencia. Además la H y la R son considerados los dos fármacos más importantes en el tratamiento contra la TB, estos representan la columna vertebral de la terapia primaria y es por ello que la mayor parte de las pruebas moleculares están dirigidas a la detección de mutaciones en la región RRDR del gen *rpoB* para determinar la resistencia a R, ya que se considera un buen marcador para definir TB- MFR. De aquí la importancia de estudiar los mecanismos genómicos de resistencia a estos fármacos<sup>29,35</sup>.

### **3.6 Métodos diagnósticos.**

El diagnóstico de la TB requiere de la combinación de una serie de técnicas con elevada sensibilidad. Se puede tener sospecha de la enfermedad por el cuadro clínico, pero actualmente su confirmación se basa en la microbiología, es decir la demostración de la presencia de MTB en alguna muestra clínica, esto resulta en gran medida un problema debido a que el tiempo necesario para la confirmación del diagnóstico tarda hasta 8 semanas por el lento crecimiento de MTB, por lo que para agilizar el diagnóstico de la enfermedad y la determinación de FR es necesario implementar nuevos métodos diagnósticos<sup>2,36,37</sup>.

#### **3.6.1 Métodos microbiológicos.**

El diagnóstico microbiológico está constituido de diversas etapas que implica el uso de diversas técnicas, las cuales conllevan mucho tiempo para la obtención de un resultado (Figura 11).



**Figura 11.** Fases del diagnóstico microbiológico de la tuberculosis.

BK: Baciloscopía, CBK: cultivo de micobacterias, PFS: prueba de fármacosensibilidad.

### 3.6.1.1 Baciloscopía.

La baciloscopía (BK) es la técnica más común para diagnosticar, controlar y determinar el pronóstico de la enfermedad, se ha utilizado por más de 100 años en todo el mundo. En la BK se busca microscópicamente BAAR en un extendido de una muestra, siendo la expectoración la muestra de elección por ser la más representativa y de fácil obtención<sup>2,6</sup>.

Esta técnica se basa en la tinción clásica de ZN, permitiendo observar en el frotis bacilos de color rojo en un fondo azul (Figura 12). Es además una técnica sencilla, reproducible, rápida, de bajo costo, detecta los casos más contagiosos y tiene una alta especificidad cercana al 100%. En cuanto a su sensibilidad es relativamente baja, detecta de 5,000 a 10,000 bacilos/mL de muestra, por lo tanto a esta técnica se le pueden escapar los casos de TB pulmonar de extensión mínima o moderada, por lo que es necesario optar por otras pruebas<sup>6,12</sup>.

### 3.6.1.2 Cultivo de micobacterias.

El cultivo del bacilo de Koch (CBK) o de micobacterias, complementa a la BK ya que permite poner en evidencia bacilos viables presentes en escasa presencia en una muestra. Además a partir del aislado se puede identificar si la población bacilar es sensible o resistente a fármacos antituberculosis. Es el único método que permite un diagnóstico certero, para seguir la evolución del paciente enfermo y asegurar su curación, es por ello que es el estándar de referencia actual al tener una alta sensibilidad, ya que detecta de 10 a 100 bacilos/mL de muestra dependiendo del medio de cultivo que se utilice (Figura 13)<sup>2,3,12</sup>.

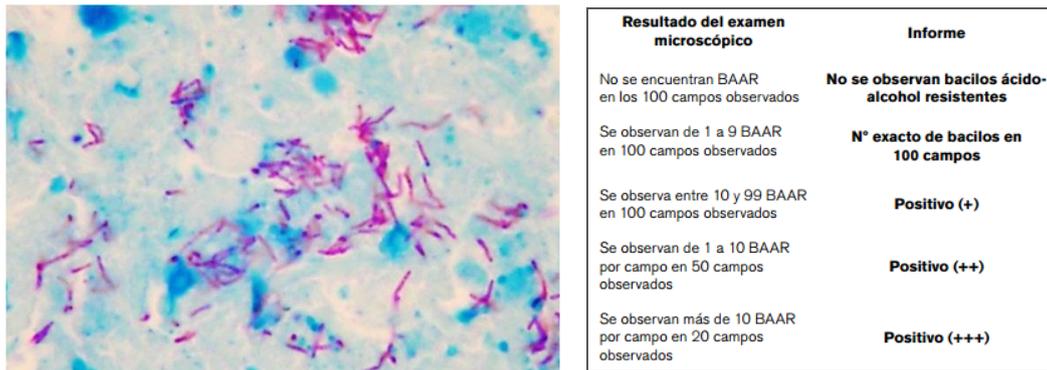


Figura 12. Técnica de baciloscopia e interpretación.

Tomado de: Manual para el diagnóstico bacteriológico de la tuberculosis. Normas y guía técnica. Parte I Baciloscopia. OPS Washington eD. CDC; 2008.

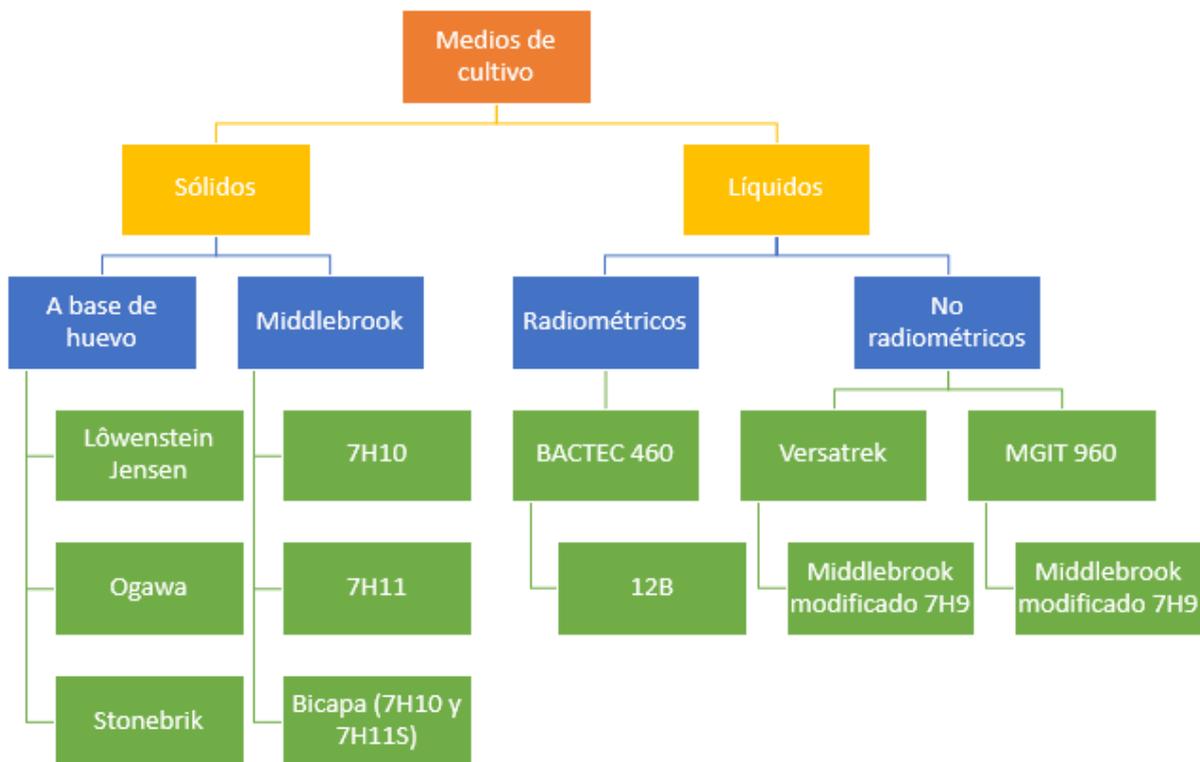


Figura 13. Medios utilizados para el cultivo de micobacterias.

Modificado de: Manual para el diagnóstico bacteriológico de la tuberculosis. Normas y guía técnica. Parte II Cultivo. OPS Washington eD. CDC; 2008.

El método de CBK se basa en la digestión y descontaminación de la muestra, para posteriormente inocularla en medios de cultivo líquidos y/o sólidos (Figuras 14 y 15). El cultivo puede dar positivo en un promedio de 15.5 días en medio líquido o de 3 a 6 semanas en medio sólido, el tiempo del cultivo es una de sus grandes desventajas para el diagnóstico, pues depende de la calidad y carga bacilar de la muestra, así como el procesamiento del cultivo<sup>2,12,37</sup>.

### **3.6.1.3 Identificación.**

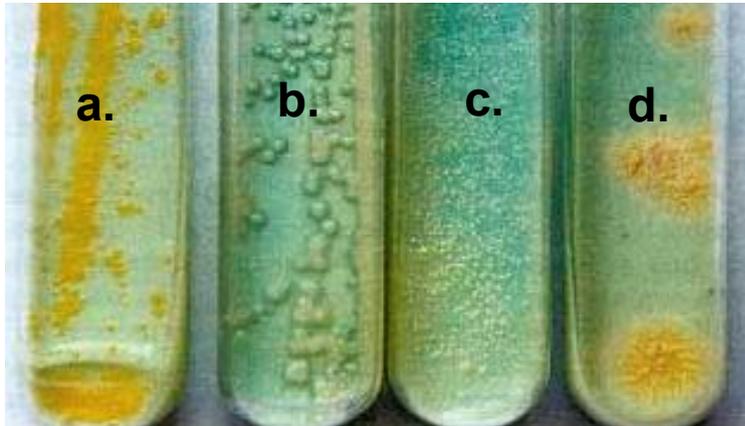
La identificación de distintas especies de micobacterias se puede realizar mediante pruebas bioquímicas sencillas. Por ejemplo, prueba de niacina, reducción de nitratos a nitritos y catalasa termolábil (Figura 16). El principal inconveniente es su lentitud y poca reproducibilidad, por lo que se deben aplicar otras técnicas de identificación molecular más sensibles como hibridación con sondas de ácidos nucleicos específicas (Accuprobe MTB, Gen-Probe) o mediante la amplificación de la región génica IS6110 mediante la reacción en cadena de la polimerasa (PCR, por sus siglas en inglés)<sup>2,12</sup>.

### **3.6.1.4. Prueba de fármacosensibilidad.**

Para la realización de la prueba de fármacosensibilidad (PFS) *in vitro*, es necesario el aislamiento e identificación de MTB. Entre las técnicas existentes se encuentra la de las proporciones de Canetti, en la cual una cepa se considera resistente cuando el 1% o más de la población bacilar se desarrolla en presencia de la concentración crítica, es decir de la CMI de un determinado fármaco que inhibe al 95% de cepas silvestres de MTB que nunca han sido expuestas al fármaco, pero que al mismo tiempo no inhibe a las cepas de MTB que han sido aisladas de pacientes que no responden a la terapia y que son consideradas resistentes.

Esta técnica fue adaptada para un sistema semiautomatizado en medio líquido denominado BACTEC MGIT 960, ya que inicialmente se empleaba para el medio sólido y se contaban las unidades formadoras de colonia en ausencia y presencia del fármaco a determinar. El método BACTEC MGIT 960 (Figura 17) es ampliamente utilizado, se basa en el uso de un medio líquido denominado tubo indicador de crecimiento de micobacterias (MGIT, por sus siglas en inglés) que contiene en el fondo un compuesto fluorescente (cloruro pentahidratado de tris-4,7difeníl-1,10-

fenantrolina rutenio) embebido en silicona, este compuesto es inhibido por el oxígeno disuelto en el medio, la micobacteria al crecer utiliza ese oxígeno y se libera la fluorescencia del compuesto, la cual es detectada por el equipo.



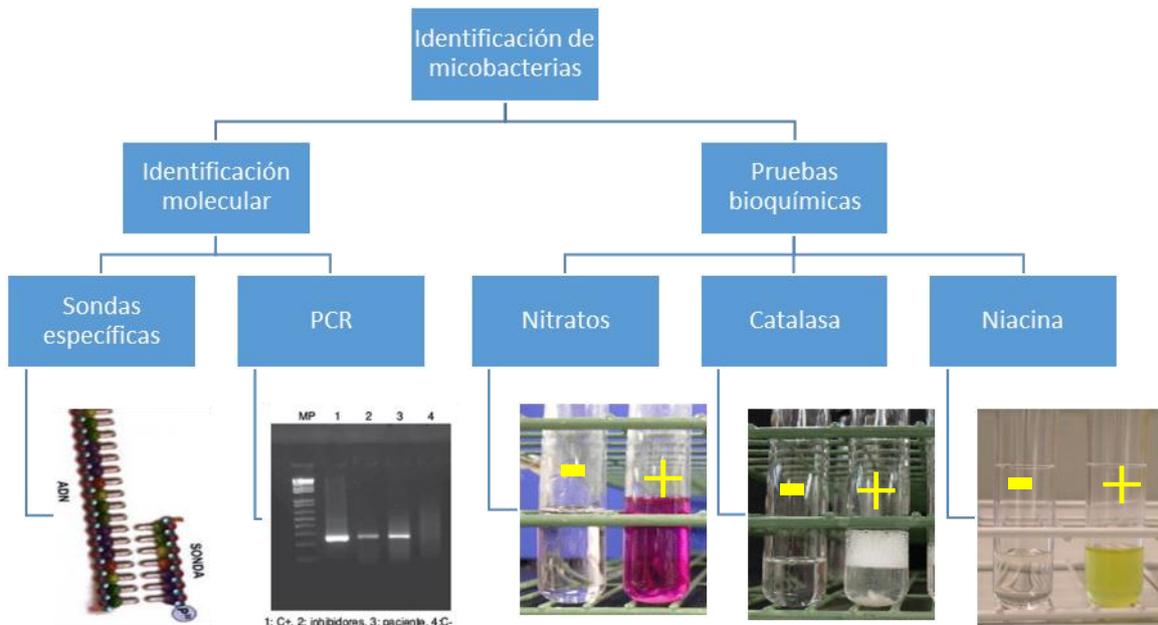
**Figura 14.** Crecimiento de micobacterias en el medio Lowenstein Jensen: a. *Mycobacterium gordonae*; b. *Mycobacterium fortuitum*; c. *Mycobacterium avium*; d. *Mycobacterium tuberculosis*.

Modificado de: Palomino JC, Leão SC, Ritacco V. Tuberculosis 2007; from basic science to patient care. 2007.



**Figura 15.** Crecimiento de micobacterias en el medio MGIT.

Fuente: Cortesía del INER.



**Figura 16.** Identificación de micobacterias.

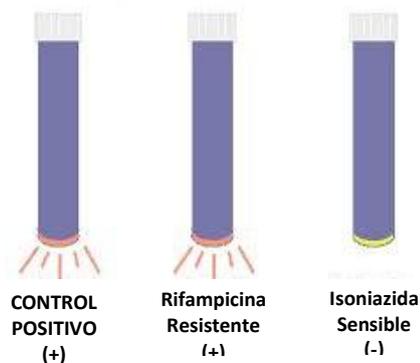
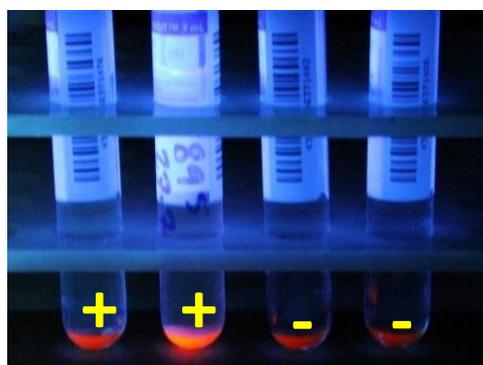
Para realizar la PFS el aislado se cultiva en presencia de una concentración crítica de un determinado fármaco y su control de crecimiento sin dicho fármaco, mientras que el equipo utiliza el análisis de la fluorescencia del tubo con antibiótico y compara con la fluorescencia del tubo de control de crecimiento para determinar los resultados de sensibilidad. La PFS de primera línea tiene una duración en promedio de 8-15 días, haciendo que esta prueba sea muy tardada para la determinación de la FR<sup>12,38-40</sup>.



a. Batería de tubos MGIT para la realización de la Prueba de Fármacosensibilidad.



b. Equipo BACTEC MGIT 960, Sistema de Detección de Micobacterias.



c. Ejemplo de la lectura de la susceptibilidad en tubos MGIT.

**Figura 17.** Prueba de Fármacosensibilidad mediante el sistema BACTEC MGIT 960.

### 3.6.2 Métodos moleculares.

A raíz de los recientes avances en el conocimiento genómico de TB, el uso de pruebas moleculares rápidas para diagnóstico e identificación de TB fármacorresistente está en aumento (Cuadro 2). Actualmente existen métodos comerciales avalados por la OMS para la identificación de MTB y la resistencia a R como es el caso de las pruebas GeneXpert MTB/RIF (Cepheid) y GenoType

MTBDR*plus* (Hain Lifescience), esta última también determina la resistencia a H<sup>6,41</sup>. Además se están utilizando otras tecnologías como microarreglos (como el chip VereMTB, Veredus laboratories), con regiones genómicas asociadas a FR ya conocidas (*rpoB*, *katG* e *inhA*) y la secuenciación de ADN que se considera como el estándar molecular para definir genotípicamente a una cepa fármacorresistente<sup>29</sup>, esta caracterización implica la identificación de mutaciones específicas en el ADN de la cepa infecciosa.

**Cuadro 2.** Métodos diagnósticos moleculares.

Método diagnóstico*	Descripción	Tiempo de resultados**	Sensibilidad	Especificidad
LiPA	Sistema colorimétrico que identifica mutaciones basado en la hibridación reversa de fragmentos de PCR de MTB biotinilados, con sondas de nucleótidos colocados en línea en una tira de membrana.	12 h	(R) 95-98%	(R) 98-100%
GenoType MTB <i>plus</i>	Técnica similar a LiPA, detecta resistencia a H y R simultáneamente.	6 h	(R) 95-99% (H) 82-83%	(H) 97-100% (R) 95-100%
Microarreglos	Fundamentado en el diagnóstico de FR en la hibridación de ADN de MTB con múltiples sondas específicas con fluorocromos adheridas a una membrana, permitiendo el análisis de miles de genes en un sólo ensayo.	6-8 h	-	-
GeneXpert***	Es una PCR semicuantitativa, integrada y en tiempo real, que se utiliza para la detección de ADN de MTB y la detección de mutaciones asociadas a la resistencia a la rifampicina del gen <i>rpoB</i> .	2 h	BAAR(+): 97-100% BAAR(-): 75-84%  (R) 91-97%	98-99%  (R) 91-97%
Secuenciación	Método que identifica y analiza la secuencia nucleotídica de un fragmento de un gen específico mediante un paquete de computo, con la finalidad de identificar mutaciones que se sabe causan FR.	8 h	100%	100%

R: rifampicina; H: isoniazida; MTB: *Mycobacterium tuberculosis*; FR: fármacorresistencia, ADN: ácido desoxirribonucleico; LiPA: ensayo de prueba en línea, por sus siglas en inglés<sup>29</sup>.

\* A partir de ADN de cepa previamente aislada de MTB.

\*\* Tiempo de resultados sin considerar el tiempo del cultivo.

\*\*\* Se realiza directamente en muestra de expectoración.

Actualmente es posible conocer la secuencia de genes completos en lugar de unas cuantas mutaciones gracias a la Secuenciación de Nueva Generación (SNG), que da la posibilidad de analizar secuencias de ADN de forma masiva (varios, decenas o

centenas de genes y muestras a la vez). Cabe destacar que la OMS en su último reporte global sobre la TB (2014) hace especial hincapié en reforzar la investigación en TB con el uso de otras tecnologías. Se ha incluido a la secuenciación de ADN como herramienta para el desarrollo de pruebas rápidas para combatir esta enfermedad y su FR. Lo anterior, con la finalidad de aumentar la tasa de curación y evitar la propagación de cepas MTB fármacorresistentes<sup>6</sup>.

En el campo de la investigación en enfermedades infecciosas la SNG de ADN es una de las herramientas de elección para la detección de agentes etiológicos de enfermedades de difícil identificación, así como para realizar inferencias acerca de la resistencia a antibióticos, virulencia, evolución y transmisibilidad<sup>42,43</sup>. Las plataformas actuales de secuenciación de ADN continuamente se han ido mejorando convirtiéndose en una tecnología más rápida, eficiente y barata, siendo la plataforma de Illumina la que ofrece mejor costo-beneficio en mercado actual<sup>44,45</sup>.

### **3.7 Estudios previos en México.**

En México se han realizado diversos estudios moleculares de FR utilizando diversas herramientas en aislados de MTB, es decir cepas aisladas de cultivos previos<sup>46</sup>. Algunos estudios sólo se han enfocado a la determinación de monorresistencia a R mediante las técnicas de PCR (Bobadilla et al.)<sup>47</sup> y otros han utilizado la secuenciación de amplicones de regiones específicas, por ejemplo la región RRDR del gen *rpoB* (Bobadilla et al., Varma et al.)<sup>48,49</sup>. También se han determinado mutaciones en los genes *rpoB* y *katG* utilizando técnicas de polimorfismos de longitud de fragmentos de restricción (RFLP, por sus siglas en inglés) e hibridación reversa, mediante la prueba INNO LiPA Rif.TB (Innogenetics) (Viader et al.)<sup>50</sup>. Otros pocos han utilizado la secuenciación de ADN para analizar regiones específicas de estos genes (López et al., Zenteno et al., Molina et al.)<sup>41,51,52</sup>.

Actualmente se han comenzado analizar otros genes asociados a la resistencia de H (*inhA* y su región promotora), así como la identificación de mutaciones en genes potencialmente asociados a la resistencia de otros fármacos de primera línea y segunda línea (Ramaswamy et al., Garza et al., Bolado et al., Cruz et al.)<sup>53-56</sup>.

Sin embargo, se ha encontrado una alta heterogeneidad entre las cepas de MTB analizadas y en algunos casos no se ha encontrado una correlación entre el genotipo obtenido con el patrón de resistencia de los aislados. Lo anterior conlleva a buscar alternativas que puedan dar lugar a la confirmación del fenotipo (correlación fenotipo-genotipo), es decir la presencia o ausencia de FR mediante el uso de herramientas moleculares como la SNG. Además, no se ha estudiado su aplicación directa en muestras clínicas, como es el caso de la expectoración. Esta muestra es la más representativa de la TB pulmonar, debido a que proviene de la zona afectada, es decir de las vías respiratorias bajas y su análisis permite conocer el estado del sistema respiratorio de forma indirecta. La expectoración es de fácil obtención a diferencia de otras muestras pulmonares, tales como biopsias o lavados bronquiolo alveolares, en los cuales se necesita realizar procedimientos invasivos en los pacientes<sup>13,20</sup>.

De esta manera el empleo de la expectoración como base de estudio, genera un reto debido a la complejidad de su naturaleza. No obstante, el empleo de la muestra y no de la cepa aislada tiene ciertas ventajas tanto para el diagnóstico, así como la investigación. La bioseguridad que se requiere es menor, debido a que la carga bacilar con la que se trabaja en la muestra es mínima en comparación de la cepa y se reduce el riesgo de un contagio<sup>57</sup>. Otra ventaja es el tiempo, es decir que la cepa aislada requiere un proceso de cultivo, que por las características de MTB tarda en crecer alrededor de 3 a 5 semanas y puede correr el riesgo de contaminarse retrasando procesos posteriores<sup>13,40</sup>. Debido a lo anterior, el empleo de la expectoración como muestra de estudio aunado al uso de nuevas tecnologías como la SNG podría establecer las bases para generar una prueba rápida que pueda disminuir el tiempo en el diagnóstico y la determinación de la FR.

#### 4. PLANTEAMIENTO DEL PROBLEMA.

La TB es una enfermedad infecciosa considerada un importante problema de salud pública. En México se han reportado un número considerable de casos MFR y la presencia de casos XFR, estos están asociados principalmente a la resistencia simultánea de H y R. Por ello es importante identificar variantes genéticas en MTB que estén asociadas a la resistencia a estos fármacos.

Actualmente no existen pruebas clínicas que determinen FR *a priori* al tratamiento y que analicen una batería comprehensiva de variantes genéticas responsables de ésta. Por lo que es necesario investigar exhaustivamente las variantes genéticas en MTB, con el fin de identificar variantes asociadas a la resistencia a H y R específicamente en cepas prevalentes en México. En este sentido, la información generada podría ser extrapolada al desarrollo de pruebas rápidas para el diagnóstico y la selección de medicamentos con información sobre FR previo a su administración.

En México ya se han realizado estudios de índole semejante, como es el caso de estudios realizados en Veracruz por Zenteno-Cuevas y colaboradores, que sólo se han enfocado a analizar regiones con variantes en genes específicos, localizando sólo las más frecuentes, limitando la exploración de otras variantes genéticas en regiones diversas con posible asociación con FR. Es decir, que no se ha realizado estudios en otros genes y en secuencias completas de genes ya relacionados a FR, menos aun directamente en muestras clínicas sino que dependen de la obtención de cultivos de muestras biológicas para su posterior aislamiento del patógeno, lo cual toma varias semanas retrasando el diagnóstico de la enfermedad.

Por lo anterior nos planteamos la siguiente pregunta de investigación:

¿Al analizar la secuencia completa de genes específicos de MTB directamente en muestras de expectoración de pacientes mexicanos, se podrán identificar variantes genéticas asociadas a la resistencia a H y R?

## 5. HIPÓTESIS.

El análisis de la secuencia completa de genes específicos de MTB permitirá identificar variantes genéticas asociadas a la resistencia a H y R a partir de muestras de expectoración de pacientes mexicanos.

## 6. OBJETIVOS.

### 6.3 General:

Identificar variantes genéticas en secuencias completas de los genes *rpoB*, *katG*, *inhA*, *mabA* y la región promotora *mabA/inhA* de MTB asociadas a la resistencia a H y R directamente en muestras de expectoración de pacientes mexicanos mediante SNG de ADN.

### 6.2 Específicos:

- Desarrollar una técnica de extracción de ADN que favorezca el enriquecimiento de ADN de MTB a partir de muestras de expectoración.
- Establecer las condiciones para la metodología de enriquecimiento de las regiones genómicas de interés de MTB para la SNG.
- Identificar variantes genéticas como SNV's, inserciones y deleciones en las secuencias de los genes de interés.
- Establecer evidencia de que el uso de SNG debe considerarse como técnica alternativa para el diagnóstico y determinación de FR de MTB.

## **7. MATERIAL Y MÉTODOS.**

### **7.3 Diseño de estudio.**

Este protocolo se desprende del proyecto de investigación “*Análisis genómico de fármaco-resistencia en Mycobacterium tuberculosis*” que se realiza en el Instituto Nacional de Medicina Genómica (INMEGEN) en el laboratorio de Farmacogenómica con la colaboración del Instituto Nacional de Enfermedades Respiratorias (INER).

#### **7.1.1 Tipo de estudio.**

Observacional, prospectivo, transversal y comparativo.

#### **7.1.2 Universo de estudio.**

Para el desarrollo de este proyecto se utilizó una muestra a conveniencia, reclutando 6 pacientes con diagnóstico de TB pulmonar activa con y sin FR, según el protocolo de la Clínica de Tuberculosis del INER.

#### **7.1.3 Criterios de inclusión y exclusión.**

##### **7.1.3.1 Criterios de inclusión:**

- Hombres y mujeres mayores de 18 años.
- Pacientes con diagnóstico de TB pulmonar sensible y fármacorresistente.
- Pacientes con diagnóstico de TB pulmonar con estudios microbiológicos de BK, CBK y PFS.
- Pacientes dispuestos a firmar la carta de consentimiento informado.

##### **7.1.3.2 Criterios de exclusión.**

- Pacientes que no cuenten con los estudios de microbiología necesarios para participar en el estudio.
- Pacientes incapaces de proporcionar las muestras de expectoración necesarias o firmar el consentimiento informado.

## 7.2 Variables.

Variable a considerar:

- Variantes genéticas identificadas.
- Perfil de fármacosusceptibilidad.
- Parámetros clínico-demográficos disponibles.

## 7.3 Metodología.

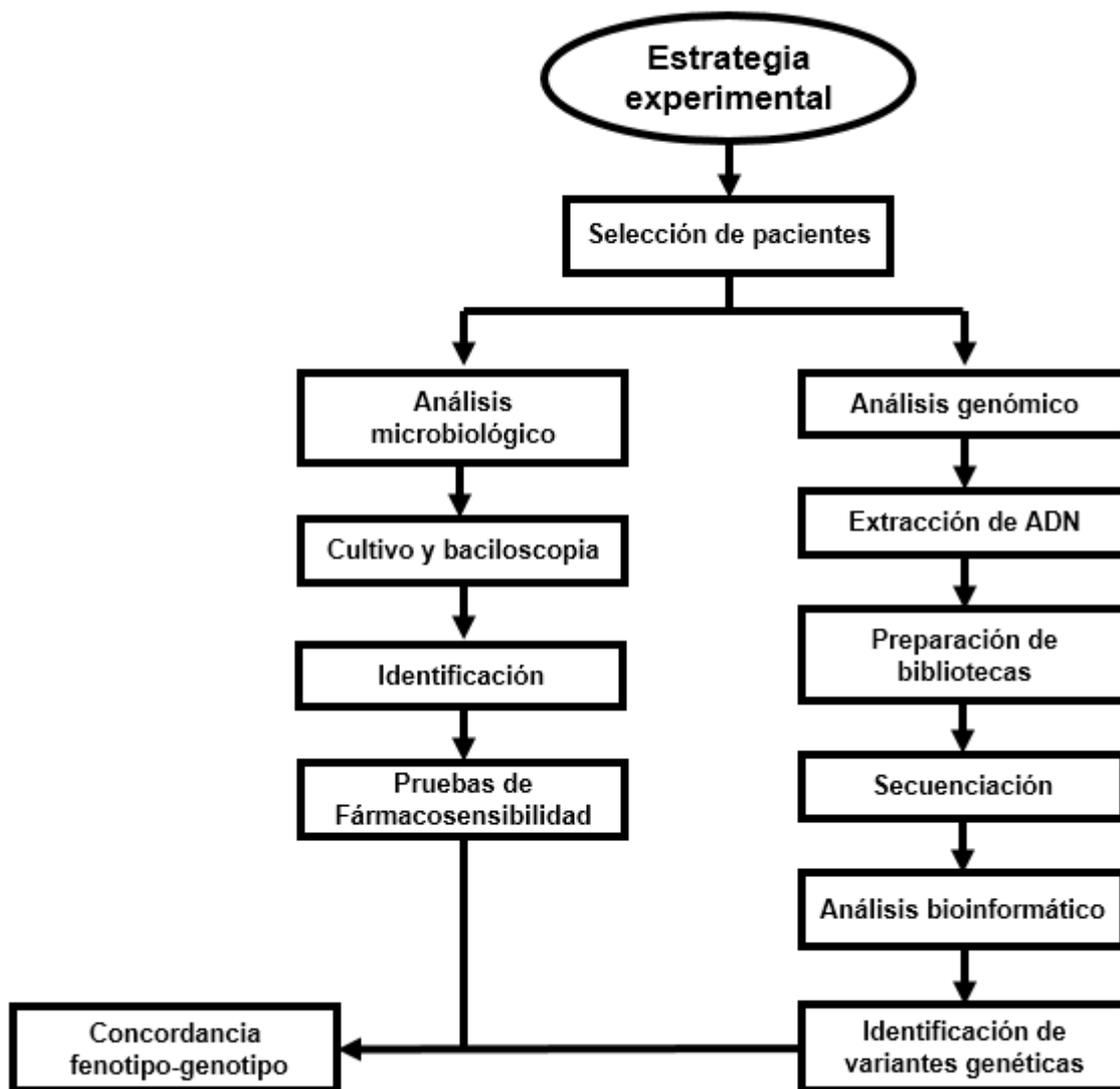


Figura 18. Estrategia experimental.

**Cuadro 3.** Materiales y equipos a utilizar.

Materiales y equipos	Utilización
<ul style="list-style-type: none"> <li>• Gabinete de bioseguridad, clase I</li> <li>• Tubos ependorf de 1.5mL</li> <li>• Micropipetas: , 0.5-2 µL, 2-10µL, 20-200µL, 100-1000µL</li> <li>• Puntas para micropipetas</li> <li>• Centrifuga refrigerada</li> <li>• Centrifuga sin refrigerar</li> <li>• Revco Ultra Low (-80°C)</li> </ul>	Extracción de ADN
<ul style="list-style-type: none"> <li>• Espectrofotometro (Nanodrop)</li> </ul>	Control de calidad: Cuantificación de ADN
<ul style="list-style-type: none"> <li>• Tubos de PCR</li> <li>• Tubos Ependorf 600µL</li> <li>• Micropipetas de 0.5-2µL, 5-20µL, 20-200µL</li> <li>• Puntas para micropipetas</li> <li>• Microcentrifuga</li> <li>• Termociclador</li> <li>• Cámara de electroforesis</li> <li>• Transiluminador UV</li> </ul>	Control de calidad: Amplificación de genes de específicos de MTB
<ul style="list-style-type: none"> <li>• Micropipetas: 0.5-2 µL, 2-10µL, 20-200µL, 100-1000µL</li> <li>• Puntas micropipetas</li> <li>• Fluorómetro (Qubit)</li> <li>• Covaris</li> <li>• Covaris microTubo con filtro AFA</li> <li>• Tubos LoBind de 1.5mL</li> <li>• Gradillas</li> <li>• Soporte magnético (Dyna DynaMag-2 magnetic stand)</li> <li>• 2100 Bioanalyzer</li> <li>• Termociclador</li> <li>• Vortex</li> <li>• Termoblock</li> <li>• Placas de 96 pozos/tubos de PCR</li> <li>• Pipetas multicanal: 2-20µL, 20-100µL</li> <li>• Concentrador a vacío</li> <li>• Microcentrifuga</li> <li>• BD Clay Adams Nutator Mixer</li> </ul>	Generación de bibliotecas y su enriquecimiento
<ul style="list-style-type: none"> <li>• Secuenciador MiSeq-Illumina</li> <li>• Cartucho de secuenciación MiSeq</li> <li>• Micropipetas: 2-10µL, 20-200µL, 100-1000µL</li> <li>• Puntas micropipetas</li> <li>• Tubos LoBind de 1.5mL</li> <li>• Gradillas</li> </ul>	Secuenciación
<ul style="list-style-type: none"> <li>• Memoria USB</li> <li>• Computadora</li> </ul>	Análisis de resultados

### 7.3.1 Selección de paciente y muestras clínicas.

Se invitó a cada uno de los pacientes a participar de manera directa y voluntaria solicitando su consentimiento por escrito, por lo que cada paciente firmó una carta de consentimiento informado y proporcionó dos muestras de expectoración que fueron entregadas en la Clínica de Tuberculosis (INER), una de las muestras fue asignada al Laboratorio de Microbiología del INER (análisis microbiológico) y otra al INMEGEN (análisis genómico), sólo se realizaron estas dos mediciones para

establecer la concordancia entre el fenotipo establecido y el genotipo obtenido. Por lo tanto no se dio seguimiento a los pacientes que participaron en este estudio. Se colectó la información clínico-demográfica mediante un cuestionario y los resultados del análisis microbiológico de cada uno de los pacientes (Anexo 1).

### **7.3.2 Estrategia experimental.**

#### **7.3.2.1 Análisis microbiológico.**

Brevemente, los protocolos del Laboratorio de Microbiología del INER incluyeron la descontaminación de las muestras de expectoración mediante la técnica de Petroff y cultivo en los medios Löwestein Jensen y MGIT. Además se realizó una BK del inóculo del cultivo para determinar la carga bacilar de la expectoración del paciente. Posterior al aislamiento de cada cepa, se realizaron las identificaciones mediante pruebas bioquímicas y sondas de hibridación de ADN (Gen-Probe, Syngene). Por último se les realizó la PFS, mediante el método fluorométrico BACTEC MGIT 960, según el protocolo del fabricante (Becton Dickinson). El laboratorio emitió un resultado que contiene el resultado de la BK, identificación de la cepa y los resultados de las PFS, esta información se recabó en la hoja de registro de datos (Anexo 2).

#### **7.3.2.2 Análisis genómico.**

Las muestras que fueron remitidas al laboratorio de Farmacogenómica del INMEGEN, se les realizó el análisis genómico por SNG de ADN. Cada una de las muestras fue separada en alícuotas de 1.5mL en tubos cónicos Eppendorf y se almacenaron a -70°C hasta su utilización.

##### **7.3.2.2.1 Extracción de ADN.**

Se probaron 4 metodologías para la extracción de ADN utilizando 6 muestras de expectoración de diferentes pacientes (alícuotas de 1.5mL), con el propósito de enriquecer la extracción del ADN de MTB y no de la biota normal o de otros patógenos.

Se utilizaron kits comerciales como Qiagen y Zymo-Research, también se experimentó con técnicas clásicas utilizando tratamientos enzimáticos con

proteasas, hidrolasas y endonucleasas; seguido de la extracción de ADN con fenol-cloroformo. Se eligió la técnica que proporcionó mayor cantidad de ADN y pureza, esta última determinada por la de relación de la absorbancia (Abs) 260/280 nm ( $Abs_{260/280} \geq 1.8$ ).

Primero, previo a cualquiera de los 4 métodos de extracción de ADN, la muestra se trató de la siguiente forma: Cada una de las muestras (1.5mL) fueron centrifugadas a 13,000rpm por 15 minutos (4°C) se eliminó el sobrenadante y a la pastilla obtenida se le agregó N-acetilcisteína y se resuspendió en agua estéril hasta obtener un volumen final de 1.5mL homogenizado completamente. La solución se dividió en 4 alícuotas, a las cuales se les asignó un protocolo de extracción de ADN, cada alícuota fue rotulada con su folio correspondiente y el método asignado (Figura 19).

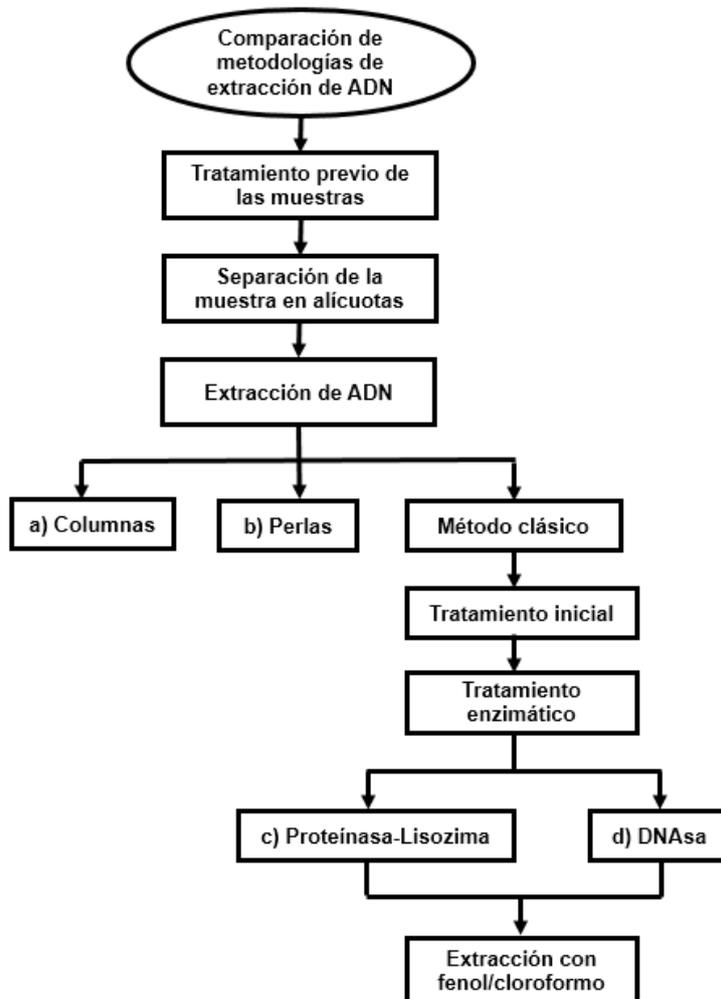


Figura 19. Comparación de metodologías de extracción de ADN.

a) Método de columnas:

El ADN fue extraído usando el kit DNasy Blood & Tissue (Qiagen), de acuerdo con las instrucciones del fabricante utilizando el protocolo de cultivo de células. Este kit se basó en el uso de columnas, las cuales se utilizaron para una rápida purificación del ADN total. La membrana DNeasy combinó las propiedades de unión de una membrana a base de sílice con tecnología microspin. El ADN fue absorbido en la membrana “DNeasy” en presencia de altas concentraciones de sal caotrópica, eliminando el agua de las moléculas hidratadas en solución. Las condiciones de amortiguación en este procedimiento permitieron la absorción específica de ADN a la membrana de sílice y la eliminación óptima de contaminantes e inhibidores enzimáticos<sup>58</sup>. Para lo cual se llevaron a cabo los siguientes pasos:

Se adicionaron 20µL de proteínasa más 200µL del buffer de lisis (AL) se incubó a 56°C por 10 minutos, seguido de la adición de 200µL de etanol (96-100%). Esta mezcla se pasó a la columna correspondiente, la cual se centrifugó a 8,000rpm por 1 minuto y se descartó el tubo de colección. Se realizaron dos lavados, para el primero se adicionaron 500µL del buffer de lavado (AW1), se centrifugó a 8,000rpm por 1 minuto y se descartó el tubo de colección; el segundo lavado se realizó adicionando 500µL del segundo buffer de lavado (AW2), se centrifugó a 14,000rpm por 3 minutos y se descartó el tubo de colección. Para la elución del ADN, la columna fue transferida a un tubo de 1.5mL y se agregó 20µL de agua libre de nucleasas, se incubó por 1 minuto a temperatura ambiente, seguida de una centrifugación por 1 minuto a 8,000rpm<sup>58</sup>. Las extracciones se almacenaron a -20°C para su posterior uso.

b) Método de perlas:

Se empleó el kit ZymoBead Genomic DNA (Zymo Research), de acuerdo con las instrucciones del fabricante utilizando el protocolo “Cell Suspensions and Proteinase K Digested Samples”. Esta metodología se empleó para el rápido aislamiento de ADN total y una máxima recuperación de ADN ultra-puro, sin contaminación de ARN. Para el procesamiento, sólo se utilizó buffer de lisis, es decir no necesitó desnaturalizantes orgánicos o digestión con proteínasa, seguido de la adición de perlas (ZymoBead) que se unieron al ADN.<sup>59</sup> Este procedimiento se realizó de acuerdo a los siguientes pasos:

Para realizar la lisis celular se adicionaron 4 volúmenes del buffer de lisis (Genomic Lysis Buffer) del volumen inicial de la muestra y se homogenizó en vórtex por 5 segundos. Posteriormente se adicionaron 10 $\mu$ L de perlas (ZymoBeads) mezclándose por inmersión e incubando por 5 minutos a temperatura ambiente, esta mezcla se centrifugó a 1,000rpm por 1 minuto y se descartó el sobrenadante. Se realizaron dos lavados, en el primero se adicionó 200 $\mu$ L del buffer de lavado (DNA Pre-Wash Buffer), se resuspendió la pastilla y se transfirió a otro tubo de 1.5mL, seguido de una centrifugación a 1,000rpm por 1 minuto, descartándose el sobrenadante. Para el segundo lavado se agregó 500 $\mu$ L del buffer de lavado (g-DNA Wash Buffer), se resuspendió la pastilla, se centrifugó a 1,000rpm por 1 minuto, se eliminó el sobrenadante y se recentrifugó para remover residuos del buffer de lavado. El ADN obtenido se eluyó en 20 $\mu$ L de agua libre de nucleasas, se resuspendió la pastilla por pipeteo, después se centrifugó a 7,000rpm por 1 minuto y se transfirió el sobrenadante a un tubo de 1.5mL. Las extracciones se almacenaron a -20°C para su posterior uso.

Método clásico:

Se experimentó con la metodología de extracción con fenol-cloroformo y se establecieron las condiciones de dos métodos por tratamiento enzimático. El primer método basado en el uso de enzimas que favorecen la lisis celular de MTB y el segundo basado en la eliminación de la mayor cantidad de ADN humano. Para llevar a cabo la extracción de ADN en ambos métodos se realizó un tratamiento inicial de digestión-descontaminación, donde a cada una de las muestras se les realizó los siguientes pasos:

Primero se centrifugó a 13,000rpm (4° C) por 10 minutos, se eliminó sobrenadante. A la pastilla obtenida se le agregó 1mL de NaOH al 4% y se incubó por 15 minutos a temperatura ambiente, seguido de una centrifugación a 13,000 rpm (4° C) por 10 minutos y la eliminación del sobrenadante. Se realizó un lavado con 1.2mL de PBS, se centrifugó a 13,000rpm (4° C) por 10 minutos y se eliminó el sobrenadante. La pastilla obtenida se lavó 3 veces con 1mL de agua estéril incubando 10 minutos, lo anterior entre cada lavado se centrifugó a 13,000rpm (4° C) por 10 minutos, esto para lisar la mayor cantidad de células humanas mediante osmosis y liberar el ADN, a la cual MTB es resistente dada la composición de su pared celular. Para la

segunda etapa del tratamiento enzimático, se separaron las alícuotas de acuerdo al método que se les asignó y se les realizó el tratamiento correspondiente:

c) Proteínasa-lisozima:

Se agregó 10 $\mu$ L de lisozima (50mg/mL) y 100 $\mu$ L de buffer de extracción (50mM Tris-HCl, 25mM EDTA, 5% glutamato monosódico, pH=7.4) se dejó incubar a 37°C por 30 minutos para romper la pared celular. Para la lisis de la membrana de MTB y degradar proteínas se adicionó 10 $\mu$ L de proteínasa K (10mg/ml) y 10 $\mu$ L de buffer 10x de proteínasa K (100mM Tris HCl, 50 mM EDTA, 5% SDS, pH=8.0), se incubó a 45°C por 30 minutos.

d) DNAsa:

Se agregó 25 $\mu$ L DNAsa I (libre de RNasas), 100 $\mu$ L de buffer (10 mM Tris-HCl, 2.5mM MgCl<sub>2</sub>, 0.5mM CaCl<sub>2</sub>, pH=7.6) y 875 $\mu$ L de agua estéril, se incubó a 37°C por 1 hora, esto para eliminar la mayor cantidad de ADN libre principalmente el humano, el ADN de MTB quedó intacto debido a que su pared lo protege de la acción de la DNAsa. Se centrifugó a 14,000rpm por 10 minutos (4°C) y se eliminó el sobrenadante.

Para la tercera etapa de ambos métodos se procedió a realizar la extracción de ADN mediante la adición de 200 $\mu$ L de agua y el doble del volumen de la muestra de fenol (fenol saturado, pH=8), se incubó por 5 minutos y se centrifugó a 10,000rpm por 5 minutos a temperatura ambiente. Se separó la fase acuosa y a ésta se le realizó una segunda extracción con fenol-cloroformo (25:25) en una relación 2:1 respecto al volumen de la muestra, se incubó por 5 minutos, se centrifugó a 10,000rpm por 5 minutos a temperatura ambiente y se separó la fase acuosa. Para la precipitación del ADN se le adicionó 1/10 del volumen de acetato de sodio (3M, pH=5.5) con respecto a la fase acuosa obtenida, más 500 $\mu$ L de isopropanol, se dejó incubar por 10 minutos a -70°C. Pasado el tiempo de incubación se prosiguió a centrifugar a 12,000rpm por 10 minutos (4°C) y se eliminó el sobrenadante. La pastilla obtenida fue lavada con 1mL de etanol al 70%, seguida de una centrifugación a 12,000rpm por 10 minutos (4°C) y se eliminó el sobrenadante. Para la elución del ADN se

adicionó 15µL de agua libre de nucleasas. Las extracciones se almacenaron a -20°C para su posterior uso.

### 7.3.2.2.2 Controles de calidad.

#### a) Concentración y pureza:

Estos parámetros fueron evaluados por espectrofotometría UV (Nanodrop, ThermoFisher), se tomaron en cuenta los valores de concentración y pureza ( $Abs_{260/280} \geq 1.8$ ).

#### b) Control de calidad del ADN extraído:

Para garantizar la presencia de MTB en la extracción de ADN de cada una de las muestras y comprobar la integridad del ADN de MTB, se amplificaron fragmentos de genes específicos de MTB como *katG* (455pb), *rpoB* (543pb) y *embB* (490pb) mediante PCR en tiempo final.

#### - Amplificación por PCR:

Se prepararon de forma independiente mezclas de reacción de 20µL, con el kit PCR Master Mix 2X (Thermo Scientific) utilizando los volúmenes del Cuadro 4 y los cebadores correspondientes a algunas de las regiones de interés específicas de MTB como *katG*, *rpoB* y *embB* (Cuadro 5).

**Cuadro 4.** Mezcla de la reacción de PCR.

Reactivo	Volumen
Buffer Master Mix 2x (buffer, MgCl <sub>2</sub> , dNTPs y Taq polimerasa)	10 µL
Cebador "reverse" de la región correspondiente	1 µL
Cebador "forward" de la región correspondiente	1 µL
ADN (100ng/ µL)	~2 µL*
Agua	~8 µL*

\* El volumen de agua y de ADN se ajustaron de acuerdo a la concentración obtenida de la extracción de ADN.

**Cuadro 5.** Secuencias de “cebadores” para la amplificación de genes de *Mycobacterium tuberculosis*.

Región	“Forward” (5'→3')	“Reverse” (5'→3')	Tamaño del amplicón (pb)
<i>katG</i>	GCGAACTCGTCGGCCAATTC	AACGACGTCGAAACAGCGGC	455
<i>rpoB</i>	CTTGACGAGGGTCAGACCA	ATCTCGTCGCTAACCACGCC	543
<i>embB</i>	CTGACCGACGCCGTGGTGATAT	TGAATGCGGCGGTAACGACG	490

Después, las mezclas de reacción se colocaron en un termociclador (GeneAmp PCR System 9700, Applied Biosystems), con el siguiente protocolo de termociclado:

- I. Desnaturalización inicial de 3 minutos.
- II. 35 ciclos de:
  - Desnaturalización a 95°C por 30 segundos.
  - Alineamiento a 55°C por 30 segundos.
  - Elongación a 72°C por 30 segundos.
- III. Elongación final a 72°C por 7 minutos.

Para la detección del producto final de cada amplificación, se realizó mediante una electroforesis en gel de agarosa, para lo cual se preparó un gel de agarosa al 1.2% con buffer TBE. Ya montado el dispositivo de electroforesis y el gel de agarosa, se agregó en el primer pozo del gel, una mezcla 6µL del marcador de peso molecular (100pb) más 3µL del buffer de carga con colorante. Seguido de la adición de 15µL de los amplicones más 3µL del buffer de carga con colorante y se colocaron en cada uno de los pozos correspondientes. La electroforesis se corrió a 100V por 1 hora y los fragmentos de PCR fueron observados en un transiluminador UV.

De acuerdo a los resultados obtenidos, se prosiguió a realizar la extracción de las 6 muestras conforme a lo establecido al protocolo seleccionado, con sus respectivos controles de calidad.

### 7.3.2.2.3 Secuenciación de nueva generación.

#### 7.3.2.2.3.1 Selección de genes y diseño del sistema de enriquecimiento.

Se diseñó el panel de genes previamente definidos (Cuadro 6) de sondas sintéticas de ARN de enriquecimiento, este diseño se llevó a cabo a través del sistema de sondas de captura de genes (SureSelect, Agilent Technologies) mediante herramientas bioinformáticas en la página del proveedor de este servicio (<http://www.genomics.agilent.com/en/NGS-Design-Tools/SureDesign-Custom-Design-Tool/?cid=AG-PT-155&tabId=AG-PR-1198&cid=AG-PT-155>).

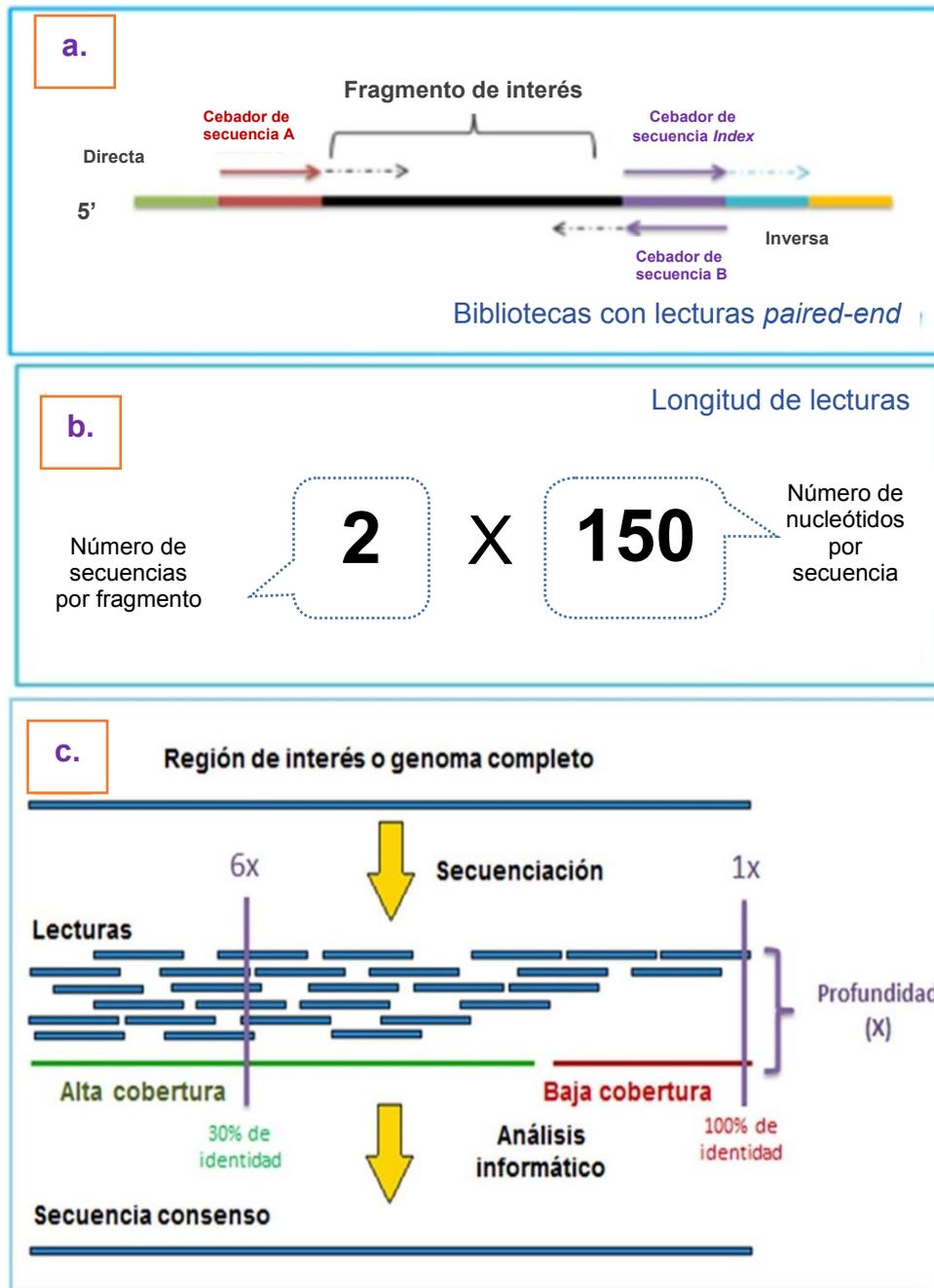
Cabe resaltar que el ADN que fue extraído de la muestra de expectoración no es exclusivo de MTB, sino que contendrá ADN de diferentes especies de microorganismos y ADN humano en gran proporción. Por lo que se seleccionó esta tecnología de enriquecimiento, previo a secuenciación para incrementar la probabilidad de secuenciar regiones de ADN de MTB y no de otras especies.

**Cuadro 6.** Genes seleccionados para la secuenciación de nueva generación.

Fármaco	Gen	Localización	Tamaño (pb)	ID	Alias
Isoniazida	<i>katG</i>	2153889-2156111	2,223	885638	Rv1908c
	<i>inhA</i>	1674192-1675011	819	886510	Rv0342
	<i>mabA</i>	1673426-1674183	757	886551	Rv1483
	<i>mabA/inhA</i>	1673200-1673425	225	-	Rv0840c- Rv0841
Rifampicina	<i>rpoB</i>	759807-763325	3,519	888164	Rv0667

#### 7.3.2.2.3.2 Generación de bibliotecas y proceso de secuenciación.

Para la SNG de ADN se utilizó la tecnología de Illumina mediante el secuenciador MiSeq, los parámetros de secuenciación (Figura 20) incluyeron una profundidad de 20x-200x y cobertura >98% (Figura 20c), en lecturas “paired-end” (Figura 20a) de 2x150 nucleótidos de longitud (Figura 20b). Estos procedimientos fueron realizados con el apoyo del personal de la Unidad de Secuenciación del INMEGEN.

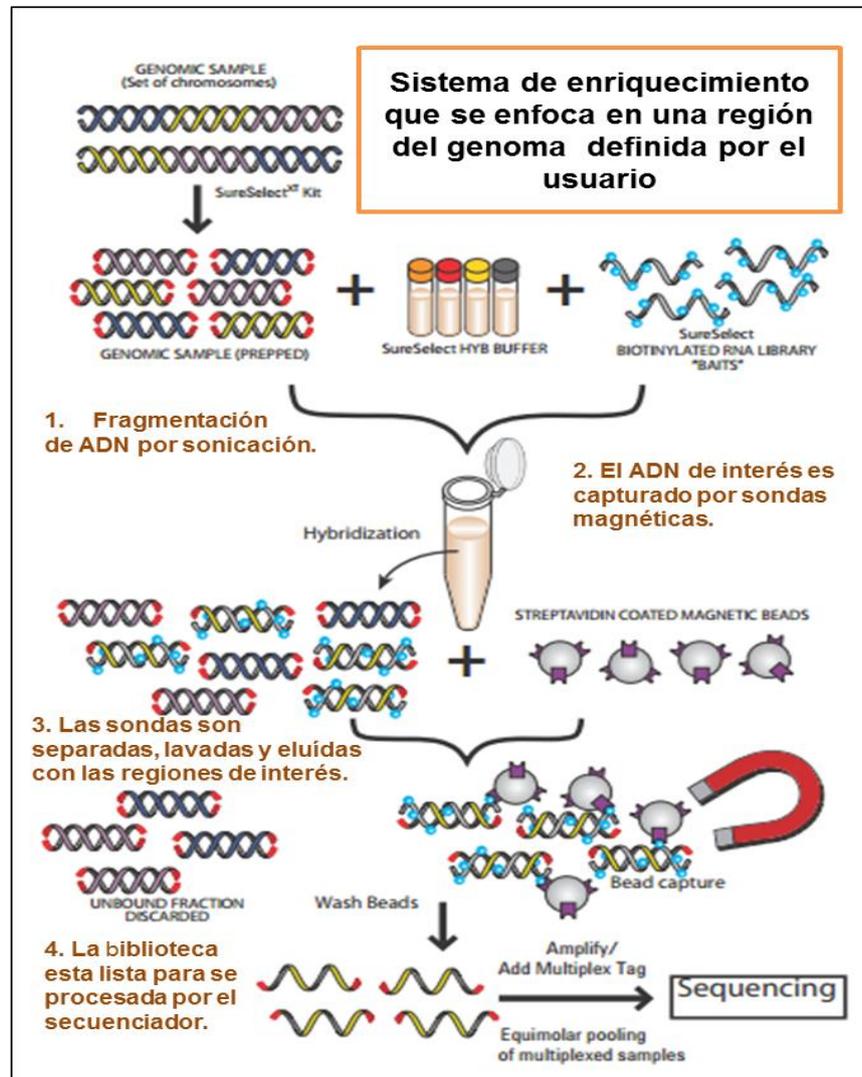


**Figura 20.** Parámetros de secuenciación. (a.) la secuenciación a partir de una biblioteca con lecturas pareadas o “paired-end”, se realiza de forma directa e inversa, de esta manera se obtienen dos lecturas por fragmento de interés, esto se establece en (b.) las longitudes de lecturas que se desean obtener por el número de nucleótidos por secuencia, para dar lugar a (c.) las lecturas de las cuales se obtendrán su profundidad (número de veces que está representada una base en una determinada región genómica) y su cobertura (porcentaje de una región genómica que ha sido cubierto con una profundidad genómica definida) y para posteriormente después de un análisis informático obtener una secuencia consenso<sup>60</sup>.

a) Preparación bibliotecas:

La generación de bibliotecas se realizó de acuerdo a las instrucciones del fabricante (SureSelect, Agilent Technologies), esta comprende la fragmentación de 200ng o 3µg de ADN mediante sonicación (Figura 21). Los protocolos se utilizaron de

acuerdo a la cantidad del ADN total obtenido de la extracción. A estos fragmentos se les realizó una ligación de adaptadores, dando lugar a las bibliotecas de secuenciación (Apéndice 1).



**Figura 21.** Preparación de bibliotecas previa a la secuenciación masiva de ADN, utilizando tecnología de enriquecimiento.

Modificado de: Agilent Technologies. SureSelectXT Target Enrichment System for Illumina Paired-End Sequencing Library. Disponible en: [http://www.chem.agilent.com/Library/usermanuals/Public/G7530-90000\\_SureSelect\\_IlluminaXTMultiplexed\\_1.8.pdf](http://www.chem.agilent.com/Library/usermanuals/Public/G7530-90000_SureSelect_IlluminaXTMultiplexed_1.8.pdf)

#### b) Enriquecimiento de bibliotecas:

Estas bibliotecas fueron hibridadas con las sondas de captura de genes (SureSelect, Agilent Technologies), estas sondas son secuencias de ARN biotinilados que se unen complementariamente a las bibliotecas que contenían las secuencias de los genes seleccionados. Posteriormente estos híbridos de ARN-ADN marcados con

biotina fueron capturados utilizando perlas magnéticas unidas a estreptavidina y aquellas que no fueron capturados se lavaron, dando lugar al enriquecimiento de las bibliotecas con las secuencias de interés.

Previo a la realización de todas bibliotecas se realizó el establecimiento de las condiciones de hibridación para garantizar una óptima captura de las regiones de interés de MTB y no de otros microorganismos. Se realizó la hibridación de acuerdo a las indicaciones del fabricante (Agilent Technologies) y algunas modificaciones con las que establecieron las mejores condiciones para llevar a cabo la hibridación con las sondas de captura de genes (SureSelect, Agilent Technologies)<sup>61</sup>.

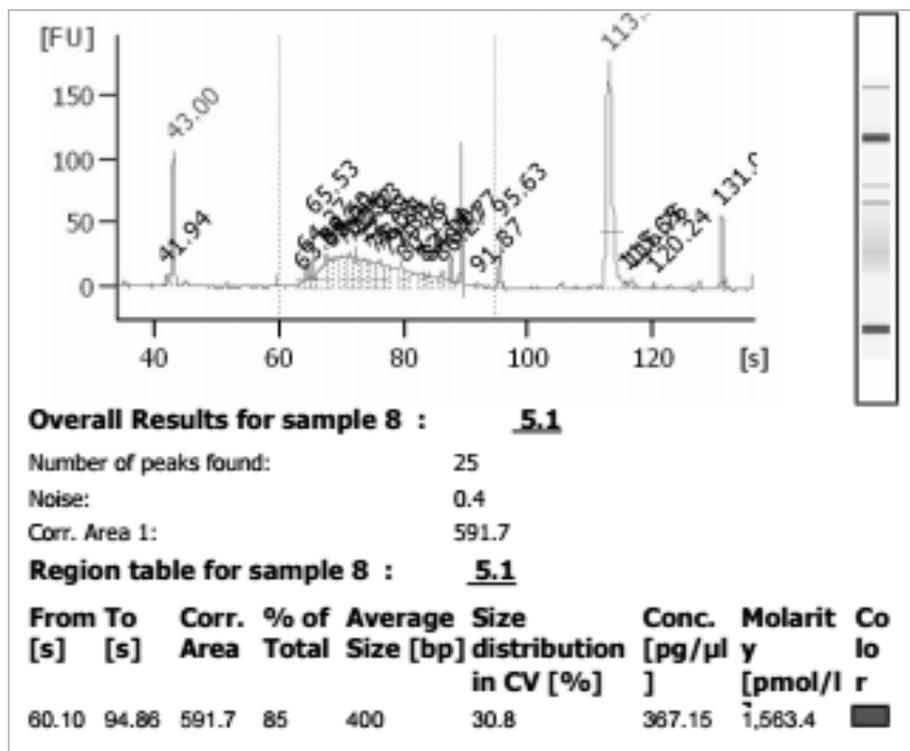
- Condiciones de la hibridación:

Se determinaron las condiciones de hibridación con el panel de captura de genes (SureSelect, Agilent Technologies), para ello se prepararon dos bibliotecas, una a partir de una muestra positiva a MTB (muestra 4TB) y otra a partir de una muestra de ADN humano utilizado como control negativo. Las hibridaciones se montaron en termocicladores (GeneAmp PCR System 9700, Applied Biosystems) de acuerdo a cada una de las condiciones a evaluar. Se realizaron los siguientes experimentos:

- I. La hibridación se realizó mediante las condiciones establecidas en el protocolo de hibridación mediante la utilización de 750ng de la biblioteca de la muestra 4TB, a una temperatura de 65°C y 24 horas de hibridación.
- II. Se evaluó que las sondas no hibridaran con ADN humano, mediante la utilización de 750ng de la biblioteca de ADN humano, a una temperatura de 65°C y 24 horas de hibridación.
- III. Se evaluó el aumento de la cantidad utilizada de biblioteca, mediante la utilización de 1500ng de la biblioteca de la muestra 4TB, a una temperatura de 65°C y 24 horas de hibridación.
- IV. Se evaluó el aumento de la temperatura de hibridación, mediante la utilización de 750ng de la biblioteca de la muestra 4TB, a una temperatura de 68°C y 24 horas de hibridación.
- V. Se evaluó el aumento del tiempo y la temperatura de hibridación, mediante la utilización de 750ng de la biblioteca de la muestra 4TB, a una temperatura de 68°C y 40 horas de hibridación.

VI. Se evaluó que las sondas no hibridaran con ADN humano cuando se aumentó la temperatura de hibridación, mediante la utilización de 750ng de la biblioteca de ADN humano, a una temperatura de 68°C y 40 horas de hibridación.

Posteriormente los productos de cada una de las hibridaciones, se les realizó la captura de las regiones mediante el panel de capturas de genes (SureSelect, Agilent Technologies), utilizando perlas magnéticas unidas a estreptavidina, con procesos de lavado y elución de la biblioteca enriquecida en agua libre de nucleasas. Para evaluar que experimento dio mejores resultados se utilizó el equipo 2100 Bioanalyzer (Agilent Technologies) mediante el chip "High sensitivity DNA". Se realizó la electroforesis capilar para la separación de ácidos nucleicos, determinando su nivel de integridad y la concentración mediante un electroferograma (Figura 22). El electroferograma es un gráfico que contrasta el tamaño de las pares de bases (pb) del producto analizado contra unidades de fluorescencia (FU) obtenidas de la electroforesis capilar. Esta fluorescencia es proporcional a la concentración del ácido nucleico medido<sup>63</sup>.



Para ello se colocó en el chip la matriz para electroforesis con el colorante fluorescente del proveedor y se distribuyó por presurización en todo el chip con ayuda de los capilares. Se agregó 1µL de cada una de las bibliotecas enriquecidas en el pozo correspondiente. El chip se colocó en el equipo para su corrimiento mediante el programa 2100 Expert Software (<http://www.genomics.agilent.com/en/Bioanalyzer-System/2100-Expert-Software/?cid=AG-PT-106&tabId=AG-PR-1002>), obteniéndose de esta manera un electroferograma, así como la concentración final del enriquecimiento de las bibliotecas<sup>63</sup>.

De acuerdo a los resultados obtenidos, se eligieron aquellas condiciones con las cuales se obtuvo una mayor cantidad de biblioteca enriquecida y se procedió a realizar la preparación de bibliotecas de las demás muestras bajo las mismas condiciones de hibridación. Por último, a las bibliotecas enriquecidas se realizó una amplificación por PCR para incrementar la presencia de los fragmentos de interés, incluyendo una ligación de secuencias identificadoras para cada muestra, que fueron específicas para llevar a cabo la secuenciación en plataformas de Illumina<sup>62</sup> (Apéndice 1).

c) Secuenciación:

Posteriormente a la generación de bibliotecas y su enriquecimiento, se tomó el volumen correspondiente a 4 nanomolar de cada una de las bibliotecas enriquecidas, se diluyeron con NaOH 0.2N a una concentración de 20 picomolar, después de mezclaron y 600µL de esta solución se añadió a la celda de flujo en el cartucho de secuenciación, originando la hibridación de los adaptadores seguido de la secuenciación.

La tecnología Illumina se basó en la secuenciación por síntesis, en la cual la biblioteca fue hibridada a la celda de flujo, se enriqueció por medio de una amplificación en puente que generó grupos o “clusters” de cadenas del mismo fragmento, estos fueron utilizados como entidades a secuenciar. La secuenciación se realizó entonces alternando ciclos de terminación reversible cíclica y de la toma de imágenes. En la reacción de terminación reversible cíclica se agregaron dNTP's (cada uno posee un fluoróforo de color particular), los cuales se agregaron de

acuerdo a la complementariedad de la secuencia de la biblioteca, después se realizó la excitación con un láser y se capturó la imagen. La imagen capturada en una posición específica corresponde al primer nucleótido y los ciclos se repiten hasta determinar la secuencia de la biblioteca, una a la vez (Apendice 1) <sup>64,65</sup>.

Un equipo MiSeq-Illumina tiene una capacidad de 4GB, en el Cuadro 7 se muestran las especificaciones del equipo y costos.

**Cuadro 7.** Especificaciones y costos del equipo MiSeq-Illumina.

<b>Especificaciones de MiSeq</b>	
Numero de muestras por corrida	50
Tiempo de corrida	36 horas
Longitud promedio de lecturas	2x150
Capacidad	4 Gb
Secuenciación por corrida	\$25,300.00*
Secuenciación por muestra	\$506.00
Costo de enriquecimiento del panel de captura de genes para biblioteca por muestra	\$7,079.43
Preparación de bibliotecas por muestra	\$2,500.00
<b>Costo por muestra</b>	<b>\$10,085.43</b>

\*Costo según la Unidad de Secuenciación del INMEGEN.

#### **7.4 Análisis bioinformático.**

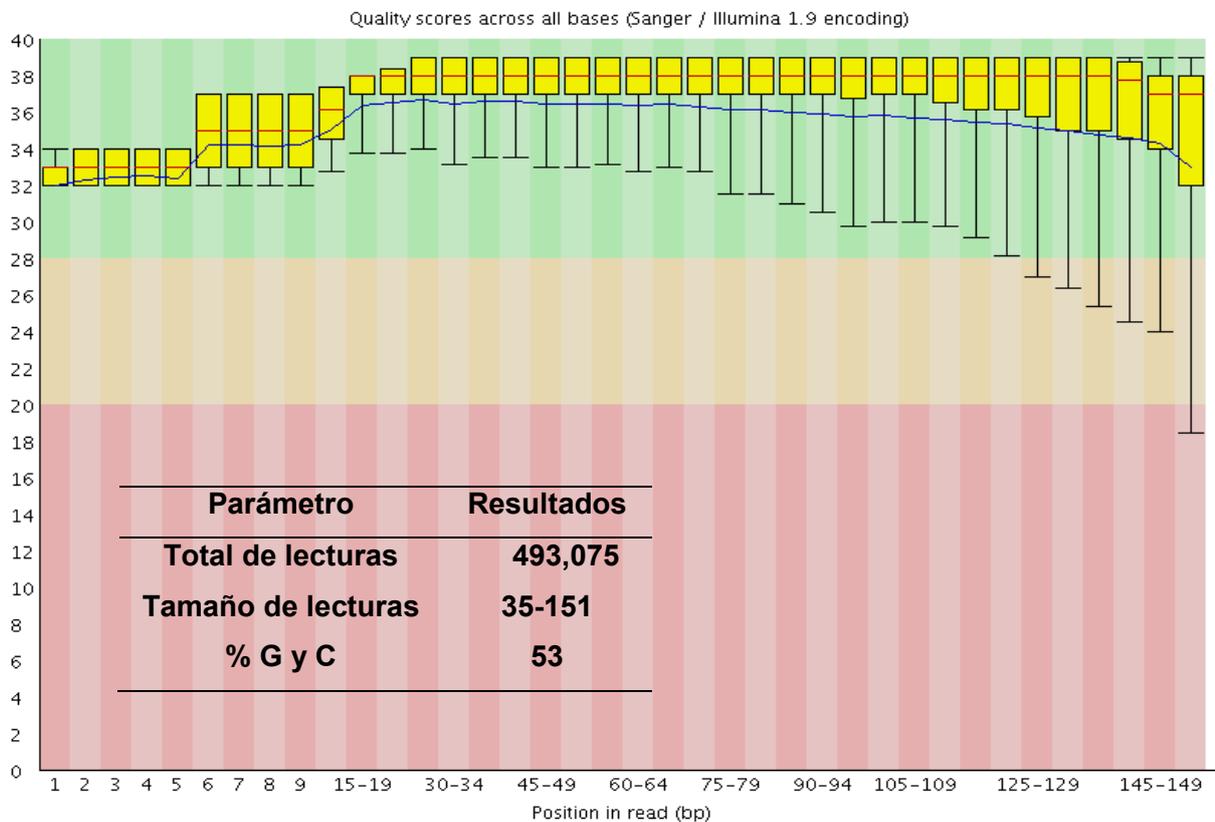
El análisis bioinformático fue realizado con la asesoría de la compañía Winter Genomics, quienes desarrollaron un “Pipeline” o un archivo de órdenes basado en comandos para su ejecución en el sistema operativo Linux. El análisis se llevó a cabo de forma automatizada para cada una de las muestras.

Después de la secuenciación, los resultados fueron extraídos del equipo MiSeq (Illumina) en archivos .fastq.gz y fueron almacenados en una memoria USB, para ser transportados a la computadora donde se realizó el análisis bioinformático. Esta computadora, tuvo los requerimientos específicos para realizar el análisis, como un sistema operativo basado en Linux (Biolinux, <http://environmentalomics.org/biolinux/>), con la paquetería necesaria con la que se realizó el análisis.

El “Pipeline” comprendió de las siguientes etapas:

I. Calidad de las secuencias obtenidas:

La calidad se determinó mediante el software FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), se ingresaron los archivos de las lecturas obteniéndose un informe que contenía las estadísticas básicas que incluían el total y el tamaño de lecturas, así como el contenido de G y C (Figura23). Así como la visualización de la calidad (Q, por su sigla en inglés) de las lecturas de secuenciación, siendo aquellas de buena calidad las que tuvieron una  $Q \geq 30$ , es decir 1 error cada 1,000 bases, localizadas en la zona de color verde.



**Figura23.** Reporte de calidad y parámetros de lecturas crudas evaluadas por el software FastQC de la muestra 1TB (R1).

II. Alargamiento de las lecturas:

El alargamiento de las lecturas se realizó mediante el software Fast Length Adjustment of Short reads (Flash) (<http://ccb.jhu.edu/software/FLASH/>). Debido a que las lecturas de secuenciación pueden o no haberse obtenido del acuerdo al tamaño establecido (150 nucleótidos), sobre todo cuando la fragmentación del ADN se

realizó al azar por sonicación, fue necesario incluir este paso para evitar problemas en procesos posteriores en la realización del análisis. El software calculó en donde las lecturas sobrelapan y las unió formando lecturas más largas.

### III. Pre-procesamiento de datos crudos:

Estas lecturas fueron sometidas a un proceso de control de calidad, para obtener sólo aquellas que cumplieron con ciertos requerimientos. Para lo cual se realizaron los siguientes pasos:

#### a) “Trimmer” o corte de lecturas de mala calidad y remoción de adaptadores:

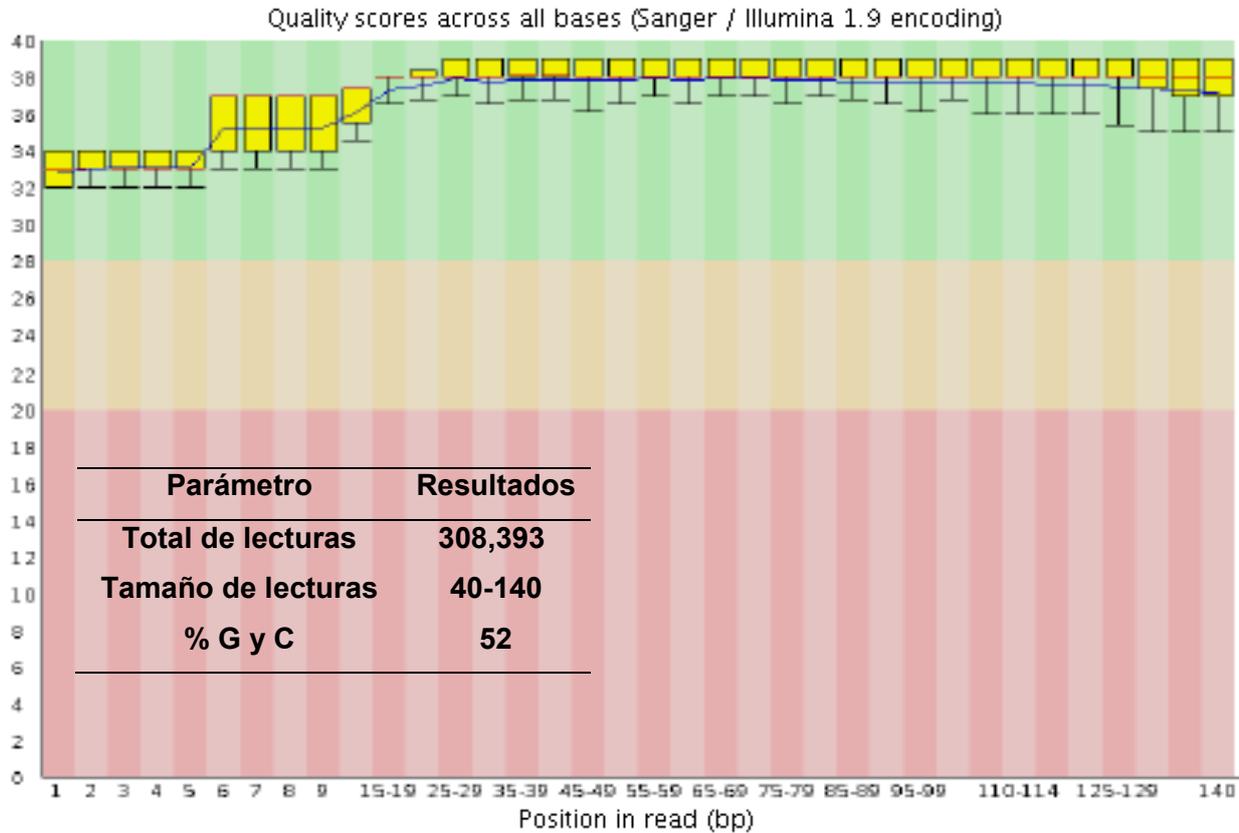
Al haber obtenido lecturas de baja calidad ( $Q < 30$ ) de acuerdo a lo observado en el reporte de FastQC, se realizó el corte en la parte final de las lecturas a partir del nucleótido 140 y se eliminaron aquellas que no tuvieron al menos el 90% de las lecturas con una  $Q \geq 30$ . También se eliminaron las probables lecturas que correspondieron a los adaptadores que fueron utilizados en la preparación de bibliotecas. Para ello, se empleó el software Fastx toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)), utilizando la herramienta Trimmomatic y el software Cutadapt (<https://pypi.python.org/pypi/cutadapt/1.4.2>).

#### b) Filtrado de ADN humano:

Debido a que el ADN utilizado en secuenciación contenía una gran cantidad de ADN humano, se realizó el filtrado de éste, para garantizar que todas las lecturas obtenidas fueran de origen bacteriano, para lo cual se realizó el alineamiento de las lecturas obtenidas con el genoma de referencia humano (Hg19). El alineamiento es el proceso por el cual se descubre cómo y dónde las secuencias de lectura son similares a la secuencia de referencia<sup>60</sup>. Este proceso de alineamiento se realizó con el programa Bowtie2 (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>), generando los índices (reordenamiento de cada uno de los nucleótidos de las secuencias) del genoma de referencia y de las lecturas, después se alineo y se desearon estas lecturas, este último mediante un comando específico.

#### c) Calidad de las secuencias obtenidas:

Posteriormente se volvió a realizar el análisis en FastQC (Figura 24), para comprobar que la calidad mejoró al realizar dichos cortes y filtrado de ADN humano.



**Figura 24.** Reporte de calidad y parámetros evaluados después del pre-procesamiento de lecturas, mediante el software FastQC de la muestra 1TB (R1).

a) Alineamiento de lecturas con las sondas de captura de genes:

Para este paso, se procedió a seleccionar sólo aquellas lecturas que correspondieron a las regiones enriquecidas de acuerdo al diseño de las sondas de captura de genes (SureSelect, Agilent Technologies), debido a la probabilidad de haber capturado genes de otros microorganismos que pudieron estar presentes en la muestra y por lo tanto de haber sido secuenciados. Esto se realizó mediante el software Bowtie2<sup>66</sup>.

Posteriormente, se construyeron los índices del diseño de las sondas de captura de genes y de las lecturas ya procesadas. La finalidad de hacer lo anterior, fue para capturar las secuencias específicas de las lecturas correspondientes al diseño de las sondas y no de artefactos de secuenciación, esto es para disminuir sesgos al momento de realizar el llamado de variantes.

## IV. Llamado de variantes:

## a) Alineamiento con el genoma de MTB (referencia).

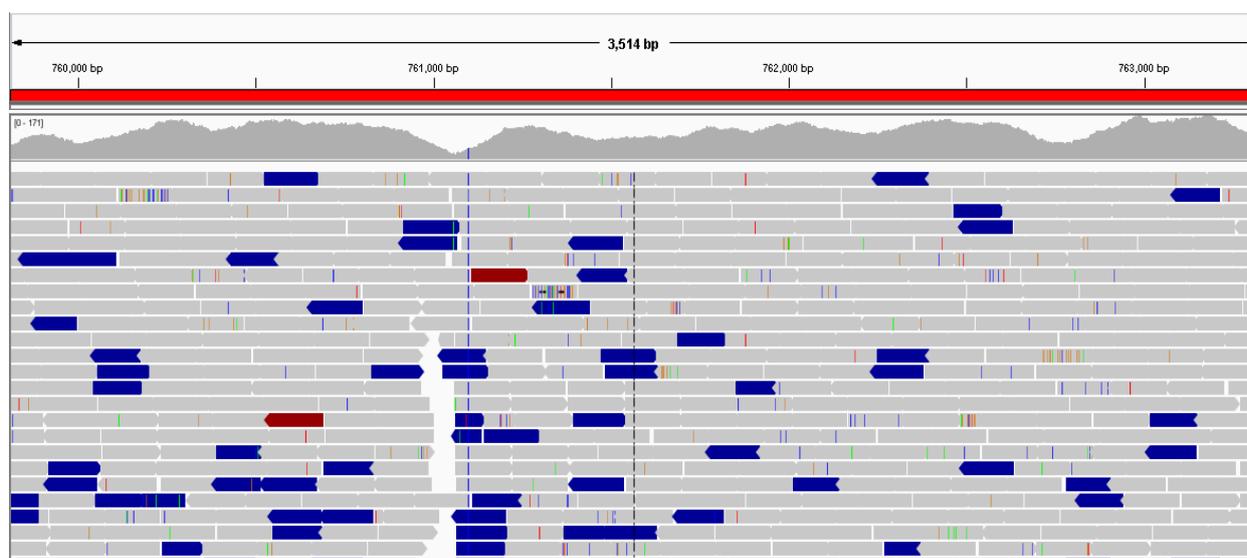
El alineamiento con el genoma de referencia MTB (H37Rv), se realizó mediante el software BWA<sup>67</sup> (<http://bio-bwa.sourceforge.net/>) mediante el algoritmo MEM. Se prosiguió a la generación de los índices tanto del genoma de referencia como de las lecturas y se realizó el alineamiento.

## b) Remoción de duplicados:

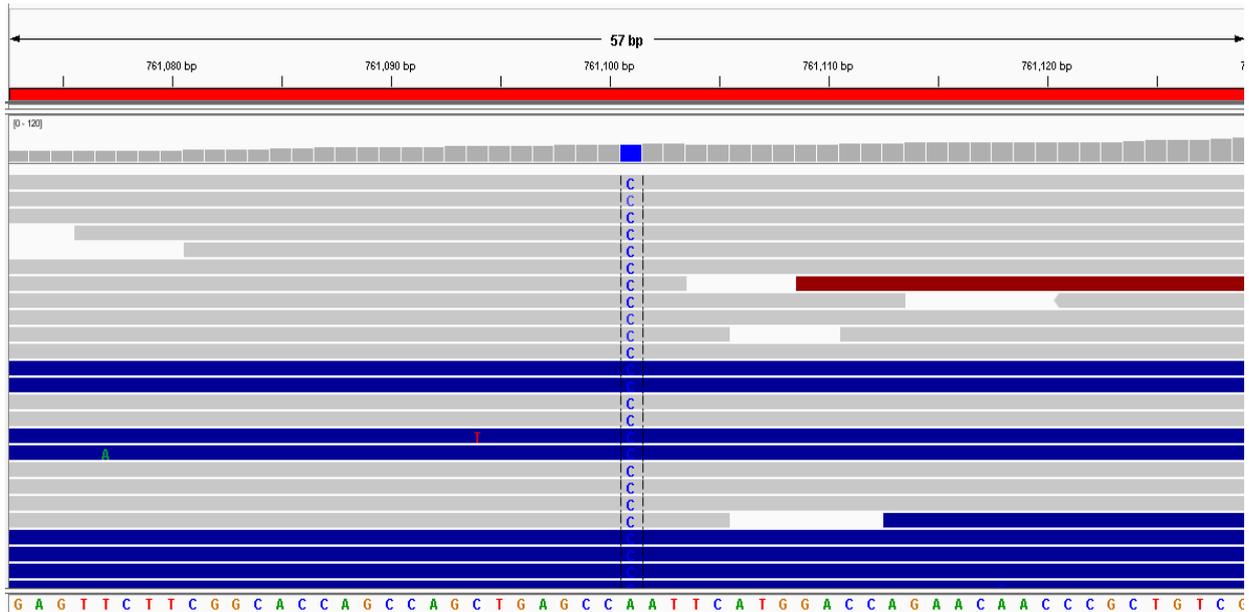
Debido a que en la preparación de bibliotecas se realizó la amplificación de las regiones de interés capturadas con las sondas de captura de genes (SureSelect, Agilent Technologies), obteniéndose más entidades a secuenciar, generando duplicados de lecturas, estas lecturas fueron eliminadas para poder identificar las verdaderas variantes. Lo anterior se realizó mediante las herramientas del software Picard (<http://broadinstitute.github.io/picard/>).

## c) Visualización de la profundidad y de cobertura:

Después del alineamiento con el genoma de referencia y la remoción de duplicados, se utilizó el visualizador del software Integrative Genomics Viewer (IGV) (<https://www.broadinstitute.org/igv/>), para observar la profundidad y la cobertura obtenidas de la secuenciación de las regiones de interés (Figura 25), además permitió observar las variantes encontradas (Figura 26).



**Figura 25.** Cobertura y profundidad de la región *rpoB* de la muestra 1TB, mediante el visualizador del software IGV.



**Figura 26.** Acercamiento a la variante identificada en la posición 761,101 del gen *rpoB* de la muestra 1TB, mediante el visualizador del software IGV.

d) Identificación de variantes:

Para la identificación de variantes se utilizaron los software SnpEff y SnpSift (<http://snpeff.sourceforge.net/>). El primero, se utilizó como una herramienta de predicción de la variante de anotación y efecto, es decir, se usó un archivo que contiene la información establecida de cada nucleótido de todo el genoma en cuestión y predice el efecto que se puede presentar si se tiene una variante en una posición determinada. El software SnpSift se usó como una herramienta para exportar a hojas de cálculo o archivos de texto los resultados obtenidos del software SnpEff (Figura 27).

#CHROM	POS	ID	REF	ALT	FILTER	ANN[*].ALLEL	ANN[*].EFFEC	ANN[*].IMPA	ANN[*].GENE	ANN[*].GENE
NC:000962	761101	.	A	C	.	C	nosynonymous	MODIFIER	rpoB	Rv0667
NC:000962	2154724	.	C	A	.	A	nosynonymous	MODIFIER	katG	RV1908c
NC:000962	2155503	.	G	A	.	A	synonymous	LOW	katG	RV1908c
NC:000962	2156025	.	G	T	.	T	synonymous	LOW	katG	RV1908c

**Figura 27.** Variantes identificada en la muestra 1TB, mediante los software SnpEff y SnpSift.

V. Predicción del impacto funcional de las variantes genéticas obtenidas como no sinónimas:

Esta predicción se realizó mediante el software Protein Variation Effect Analyzer (PROVEAN, <http://provean.jcvi.org/index.php>). El cual se utilizó para determinar si una variante no sinónima afecta o no a la función de la proteína. Esta predicción se

estableció mediante un umbral basado en las versiones de referencia y de la variante de una secuencia, con respecto a los homólogos de secuencia obtenidos de la base de datos de proteínas del National Center for Biotechnology Information (NCBI) a través de Basic Local Alignment Search Tool (BLAST). Con este software se calculó la probabilidad de que la variante altere ( $score < -2.5$ ) o no afecte ( $score > -2.5$ ) la función de la proteína en cuestión (Figura 28).

## PROVEAN Result (Download)

**PROVEAN Prediction - Job ID: 289851152093340**

- **Query sequence** (fasta)
- **Supporting sequence set used for prediction**  
 Number of sequences: **625** (fasta, E-values)  
 Number of clusters: **30**
- **Score thresholds for prediction**  
 (1) Default threshold is -2.5, that is:  
 -Variants with a score equal to or below -2.5 are considered "deleterious,"  
 -Variants with a score above -2.5 are considered "neutral."  
 (2) How to use a more stringent threshold.

Variant	PROVEAN score	Prediction (cutoff= -2.5)
Q510P	-5.594	Deleterious

- Submitted at 17:52:18 EDT, Sunday, May 17, 2015
- Started at 17:57:35 EDT, Sunday, May 17, 2015
- Finished at 17:58:20 EDT, Sunday, May 17, 2015

\* The results are kept for 48 hours.

**Figura 28.** Reporte de la predicción funcional de la variante identificada en el gen *rpoB* de la muestra 1TB, mediante el software PROVEAN.

### **7.5 Análisis estadístico.**

El análisis estadístico se realizó mediante el programa IBM SPSS Statistics 21.

- Descripción de la población de estudio:

Se calcularon las frecuencias de las variables clínico-demográficas.

- Establecimiento de la metodología de extracción:

Se determinaron las diferencias de las varianzas de las concentraciones obtenidas de los diferentes métodos extracción a partir de muestras independientes, para ello se realizó la prueba ANOVA de un factor, con un valor de significancia de  $p=0.05$ . Para determinar que método fue el que generó la diferencia estadística se realizó una prueba Post hoc, la prueba de Tukey.

## 8. RESULTADOS.

### a) Características clínico-demográficas de los pacientes reclutados:

La selección de pacientes se realizó en la Clínica de Tuberculosis del INER teniendo un total de 6 pacientes con TB con y sin FR, de los cuales se recabó su información y se determinó la frecuencia de las variables medidas (Cuadro 8). El resumen de las variables clínico-demográfica de los pacientes se establece en el Cuadro 9.

**Cuadro 8.** Frecuencia de las variables medidas en los pacientes.

<b>Variables</b>	<b>Pacientes N=6</b>
Edad	42.3 ± 6.9
Sexo	
- Hombre	5 (83.3%)
- Mujer	1 (17.7%)
Lugar de residencia	
- Oaxaca	1 (17.7%)
- Veracruz	2 (33.3%)
- Distrito Federal	2 (33.3%)
- México	1 (17.7%)
Factores de riesgo	
- Diabetes	3 (50.0%)
- Contacto	1 (17.7%)
- Alcoholismo	2 (33.3%)
- Humo de leña	2 (33.3%)
- Tabaquismo	2 (33.3%)
Tipo de caso	
- Previamente tratado	2 (33.3%)
- Sin tratamiento	4 (66.7%)
Baciloscopia	
- Negativa	1 (17.7%)
- Positiva 1+	2 (33.3%)
- Positiva 3+	3 (50.0%)
Perfil de susceptibilidad	
- Sensible	4 (66.7%)
- Fármacorresistente	2 (33.3%)

**Cuadro 9.** Descripción de los pacientes que participaron en este estudio.

Muestra	Edad	Sexo	Lugar de residencia	Factores de riesgo	Tipo de caso	Baciloscopia	Bacteria aislada	Perfil de susceptibilidad	Resistencia fenotípica
1TB	37	Mujer	Oaxaca	Diabetes	PT	Positivo 3+	CMTB	Polirresistencia	R,Z
2TB	38	Hombre	Veracruz	Contacto Alcoholismo Humo de leña	PT	Positivo 1+	CMTB	MFR	H,R,Eto
3TB	51	Hombre	Veracruz	Diabetes Tabaquismo	ST	Positivo 3+	CMTB	Sensible	-
4TB	48	Hombre	Distrito Federal	Ninguno	ST	Positivo 3+	CMTB	Sensible	-
5TB	34	Hombre	Estado de México	Diabetes Humo de leña	ST	Negativa	CMTB	Sensible	-
6TB	46	Hombre	Distrito Federal	Tabaquismo Alcoholismo	ST	Positivo 3+	CMTB	Sensible	-

PT: previamente tratado; ST: sin tratamiento; CMTB: complejo *Mycobacterium tuberculosis*; MFR: multifármacorresistente; R: rifampicina, Z: pirazinamida; H: isoniazida, Eto: etionamida.

- b) Establecimiento de la metodología de extracción de ADN a partir de muestras de expectoración:

Se emplearon 4 metodologías buscando el enriquecimiento de la extracción de ADN de MTB, para establecer la eficiencia de estos métodos se determinó la concentración obtenida de la extracción, el rendimiento ( $\%R = \frac{\text{concentración de ADN extraído}}{\text{Volumen inicial de la muestra}} \times 100$ ), su pureza (Cuadro 10) e integridad.

**Cuadro 10.** Comparación de los métodos de extracción utilizados\*.

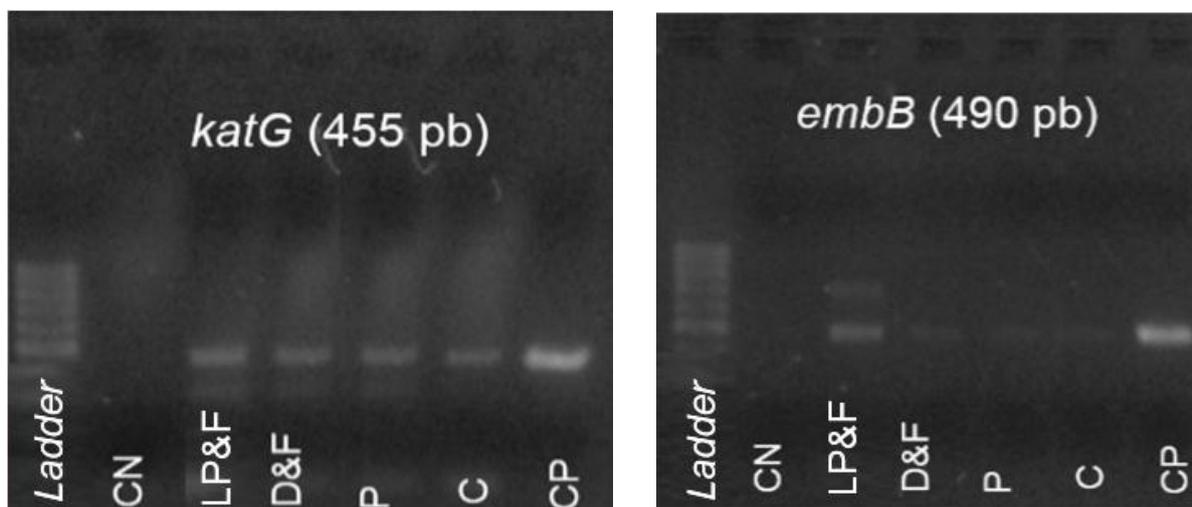
Métodos de extracción N=4	Concentración (ng/μL)	Pureza (260/280nm)	Rendimiento (%)
a) Columnas, Qiagen DNeasy Blood & Tissue Kit.	14.2 ± 11.1	1.4 ± 0.4	3.8 ± 2.9
b) Perlas, ZymoBead Genomic DNA Kit.	28.4 ± 25.0	1.0 ± 0.2	6.7 ± 6.6
c) Proteínasa-lisozima/fenol-cloroformo	104.9 ± 70.9	1.7 ± 0.1	28.0 ± 18.8
d) DNAsa y fenol-cloroformo	34.8 ± 27.8	1.6 ± 0.1	12.2 ± 7.4

\* Realizado a partir de 6 muestras de expectoración de diferentes pacientes.

El método de columnas ofreció un rendimiento del 3.8% y una pureza de 1.4, mientras que el método de perlas dio un 6.7% de rendimiento y una pureza de 1.0, estos valores obtenidos se encontraron por debajo de los requerimientos para llevar a cabo la SNG. A diferencia de los métodos basados en la extracción con fenol-cloroformo, siendo el método que se realizó con el tratamiento de proteínasa-lisozima el que dio un mayor rendimiento del 28%; en comparación con el tratamiento con DNAsa con un 12.2%, ambos métodos ofrecieron una mayor pureza de 1.7 y 1.6 respectivamente (Cuadro 10).

Para confirmar la presencia de ADN de MTB en la extracción y su integridad, se realizó una PCR en tiempo final, mediante la amplificación de regiones de algunos genes como *katG* y *embB* (Figura 29). En el caso del gen *katG* (450 pb) todas las extracciones amplificaron, pero cuando se trató de amplificar un fragmento de mayor tamaño como el del gen *embB* (490 pb), las bandas de amplificación de los métodos comerciales se observaron más tenues o incluso no se observó amplificación, en

comparación con las metodologías que se basaron en la extracción con fenol-cloroformo.



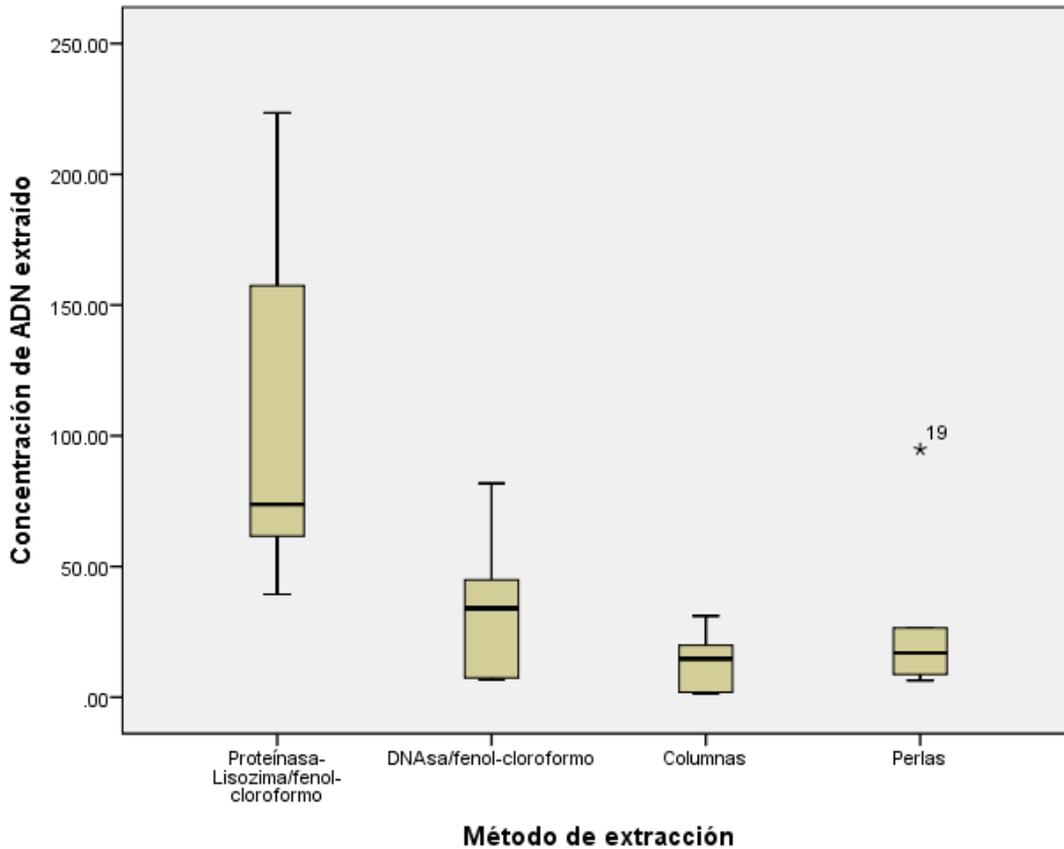
**Figura 29.** Amplificación genes de *Mycobacterium tuberculosis* de la muestra 3TB, por los diferentes métodos empleados de extracción.

CN: Control negativo; PL&F: Proteínasa-lisozima/fenol-cloroformo; D&F: DNAsa/feno-cloroformol; P: Perlas; C: Columnas; CP: Control positivo.

De acuerdo al análisis estadístico realizado mediante la prueba de ANOVA de un factor (Cuadro 11), se estableció que de los métodos comparados al menos uno es diferente estadísticamente debido a que se obtuvo una  $p=0.004$ , que está por debajo del nivel de significancia del 5%.

**Cuadro 11.** Resultados de la prueba de ANOVA de un factor.

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	29577.660	3	10086.04	6.16	0.004
Intra-grupos	35175.373	20	1637.11		
Total	64753.033	23			



**Figura 30.** Comparación de las medias de las concentraciones obtenidas, de acuerdo al método de extracción de ADN empleado.

Por lo cual se procedió a hacer la prueba de Tukey para determinar que método o métodos fueron diferentes entre sí, dando como resultado que los métodos se agruparon en dos subconjuntos, donde se observó que el método proteínasa-lisozima/fenol-cloroformo fue el que difirió a los otros tres métodos, que a su vez estos tres métodos fueron iguales teniendo una  $p=0.314$  (Cuadro 12). Lo anterior se confirmó al analizar el gráfico de “boxplots” (Figura 30), donde se observaron similitudes en el comportamiento de las medias entre los métodos de DNAsa/fenol-cloroformo, perlas y columnas, a diferencia del método de proteínasa-lisozima/fenol-cloroformo.

**Cuadro 12.** Determinación de subgrupos, prueba de Tukey.

GRUPO	N	Subconjuntos para alfa=0.05		
		1	2	3
Proteínasa-lisozima/fenol-cloroformo	6	104.9	-	-
DNAsa/fenol-cloroformo	6	-	34.81	-
Columnas	6	-	14.21	-
Perlas	6	-	24.95	-
p-valor	-	-	1.00	0.314

Al hacer las comparaciones múltiples (Cuadro 13), se observó que el método de proteínasa-lisozima/fenol-cloroformo fue casi igual al método de DNAsa-lisozima/fenol-cloroformo ( $p=0.048$ ). Los métodos DNAsa-lisozima/fenol-cloroformo, columnas y perlas fueron iguales entre sí ( $p>0.05$ ) y su vez estos difirieron al método proteínasa-lisozima/fenol-cloroformo.

**Cuadro 13.** Comparación múltiple de los métodos empleados, mediante la prueba de Tukey,

Grupos	Grupos	Diferencias de medias	p-valor
Proteínasa-lisozima/ fenol-cloroformo	DNAsa/fenol-cloroformo	-70.1	0.048
	Columnas	-90.7	0.011
	Perlas	-80.0	0.026
DNAsa/ fenol-cloroformo	Proteínasa-lisozima/fenol-cloroformo	-70.1	0.048
	Columnas	-20.6	0.123
	Perlas	-9.9	0.533
Columnas	Proteínasa-lisozima/fenol-cloroformo	-90.7	0.011
	DNAsa/fenol-cloroformo	-20.6	0.123
	Perlas	10.7	0.360
Perlas	Proteínasa-lisozima/fenol-cloroformo	-80.0	0.026
	DNAsa/fenol-cloroformo	-9.9	0.533
	Columnas	10.7	0.360

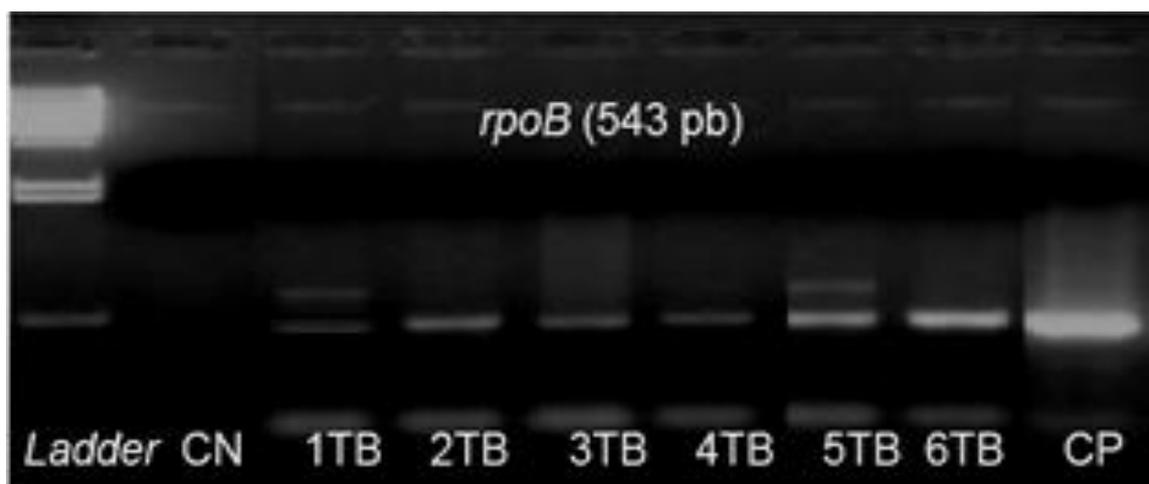
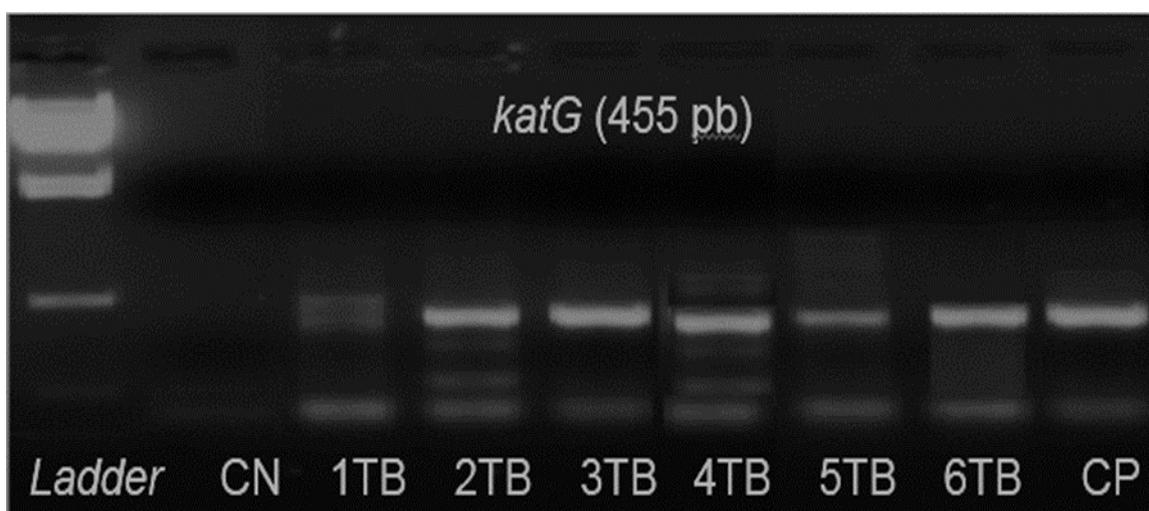
c) Extracción de ADN y su control de la calidad:

De acuerdo a los resultados anteriores se estableció que el mejor método de extracción fue el de proteínasa-lisozima/fenol-cloroformo. Por lo tanto, se realizó la extracción de las todas las muestras con dicho protocolo.

De cada muestra se obtuvo una concentración de ADN con una media de 87.7ng/μL y una pureza con una media de 1.73 (Cuadro 14). Se realizó la amplificación de todas las muestras del fragmento de la región de *katG* (455pb) y de una región de mayor tamaño como la del gen *rpoB* (543pb), lo que garantizó una mayor integridad del ADN extraído por el método seleccionado. El único inconveniente fue la aparición de bandas inespecíficas en algunos producto de PCR, como fue el caso de las muestras 1TB, 2TB, 4TB y 5TB (Figura 31).

**Cuadro 14.** Concentración y pureza obtenida de las extracciones de ADN a partir de muestras de expectoración. Método proteínasa- lizozima/fenol-cloroformo.

Muestra	Concentración (ng/ $\mu$ L)	Pureza (260/280nm)
1TB	117.4	1.67
2TB	71.4	1.73
ss3TB	74.4	1.66
4TB	52.5	1.86
5TB	120.6	1.71
6TB	90.1	1.79
$\bar{x} \pm DE$	$87.7 \pm 27.0$	$1.73 \pm 0.1$

**Figura 31.** Amplificación de genes de *Mycobacterium tuberculosis* a partir de muestras de expectoración.

d) Establecimiento de las condiciones de hibridación con las sondas de captura: El kit SureSelect (Agilent Technologies) fue diseñado primordialmente para un sólo microorganismo, es decir exclusivo de MTB. Se establecieron algunas modificaciones para garantizar la captura óptima de las regiones génicas de interés de MTB. Para ello se prepararon dos bibliotecas una a partir de ADN humano (2876.7pg/μL) y otra con una muestra positiva a MTB la muestra 4TB (4,645.4pg/μL). De estas bibliotecas se tomó el volumen necesario de acuerdo a las especificaciones para llevar a cabo cada uno de los experimentos (Cuadro 15).

**Cuadro 15.** Modificaciones del protocolo de hibridación con sondas de captura (SureSelect, Agilent Technologies).

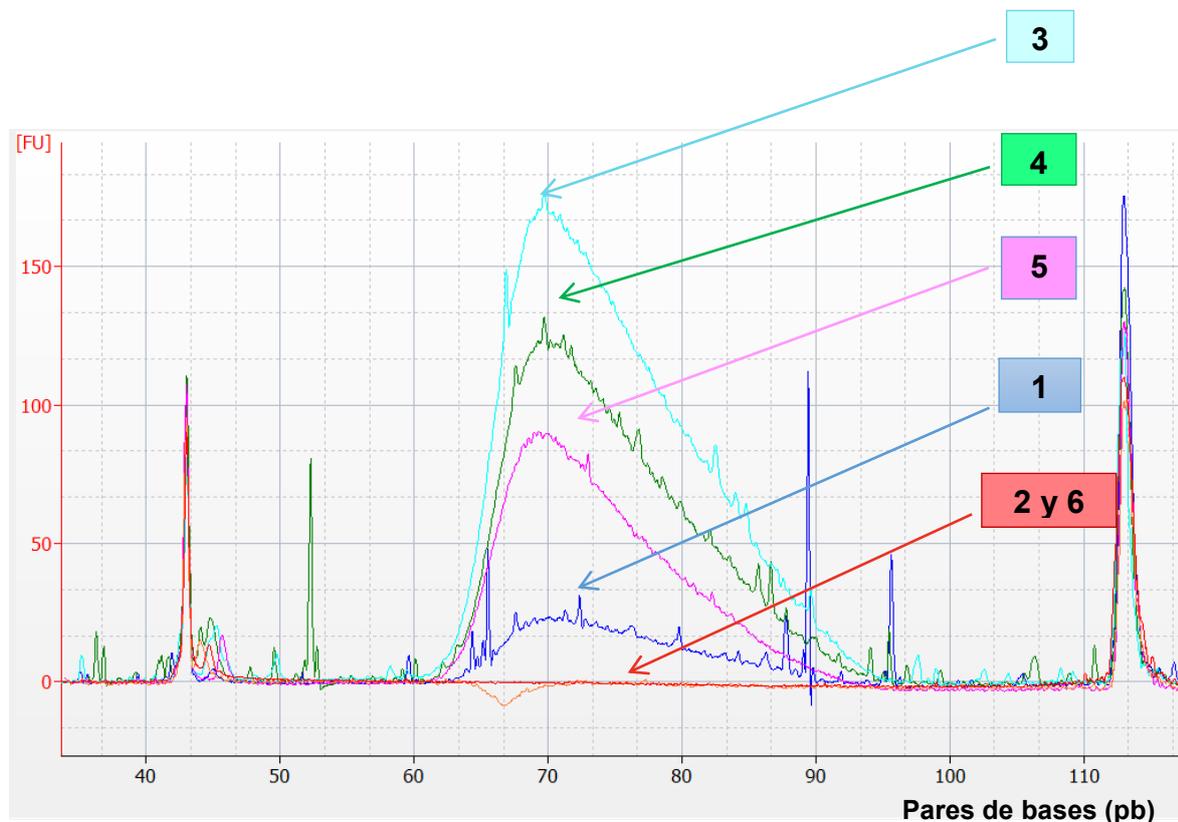
Experimento	ADN utilizado para la hibridación	Cantidad de ADN utilizado (ng)	Temperatura de hibridación ( °C)	Tiempo de hibridación (horas)	Concentración final (pg/μL)*	Rendimiento de la hibridación (%)
1	4TB	750	65	24	367.15	7.9
2	Humano (CN)	750	65	24	83.55	2.9
3	4TB	1,500	65	24	3,376.97	72.7
4	4TB	750	68	24	2,082.88	44.8
5	4TB	750	68	40	1,469.08	31.6
6	Humano (CN)	750	68	40	39.46	1.4

\*Concentración final de la biblioteca enriquecida obtenida del electroferograma.  
CN: control negativo.

Los productos de hibridación fueron sometidos a una electroforesis capilar, obteniéndose un electroferograma (Figura 30) y la concentración final de la hibridación. Se calculó el rendimiento de la hibridación ( $\%R = \frac{\text{Concentración final de la hibridación}}{\text{Concentración de la biblioteca}} \times 100$ ), para determinar cuáles fueron las mejores condiciones para llevar a cabo la hibridación con las sondas de captura de genes (SureSelect, Agilent Technologies).

Para los controles negativos se utilizaron bibliotecas a partir de ADN humano, esto para garantizar que las sondas hibridaron específicamente con ADN de MTB. De acuerdo a las condiciones de los experimento 2 y 6, se obtuvieron un rendimiento del 2.9% y 1.4%, respectivamente. Lo anterior indicó que si hubo una captura de ADN humano, pero esta fue menor cuando se incrementó la temperatura de 65°C

(experimento 2) a 68°C (experimento 6), lo que sugiere que la hibridación se hizo más específica.



**Figura 32.** Electroferograma de las diferentes condiciones utilizadas para la optimización de la hibridación con las sondas de captura de genes (SureSelect, Agilent Technologies), mediante el software 2100 Expert.

En el protocolo inicial se indicó la utilización de 750ng de biblioteca, 65°C y 24 horas de hibridación (experimento 1) ofreciendo un rendimiento del 7.9%, que fue menor de lo esperado. Cuando se aumentó la cantidad de biblioteca utilizada (experimento 3) de 750ng a 1,500ng de biblioteca, con 65°C y 24 horas de hibridación se incrementó el rendimiento a 72.7%. Cuando se incrementó la temperatura de hibridación (experimento 4) al utilizar 750ng de biblioteca, 68°C y 24 horas de hibridación, se obtuvo un rendimiento del 44.8%. Al aumentar la temperatura y tiempo de hibridación (experimento 5) al utilizar 750ng de biblioteca, 68°C y 40 horas de hibridación, se obtuvo un rendimiento del 31.6%.

En el electroferograma (Figura 30) se observaron diversas campanas obtenidas a partir de la electroforesis capilar, a partir de cada uno de los productos de hibridación de cada experimento y aquella campana que dio el punto más alto, fue el

experimento con el que se obtuvo una mayor concentración de biblioteca enriquecida con las regiones génicas de interés, siendo este el experimento 3 el que dio un mayor rendimiento. Sin embargo, debido a lo observado en los controles negativos, el aumento de la temperatura de hibridación disminuyó la captura de ADN humano, por lo tanto se eligieron las condiciones del experimento 4 (750ng de biblioteca, 68°C y 24 horas de hibridación) debido a que el aumento de la temperatura hizo más específica la captura con las sondas de captura de genes (SureSelect, Agilent Technologies).

Una vez seleccionadas las mejores condiciones de hibridación se prepararon las bibliotecas, enriquecidas en cuatro genes y una región promotora de MTB asociadas a la resistencia a H y R, a partir del ADN total de 6 muestras de expectoración.

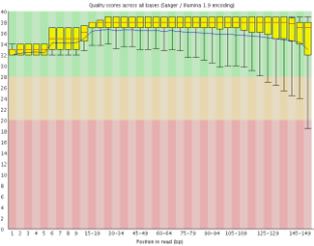
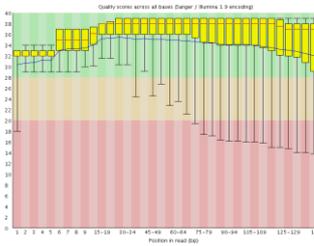
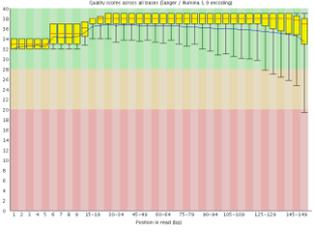
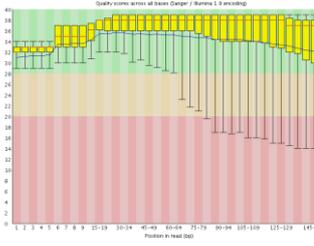
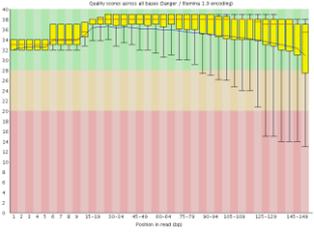
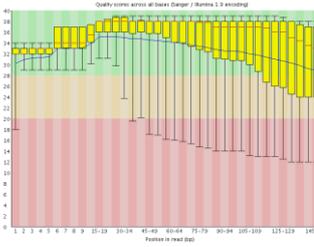
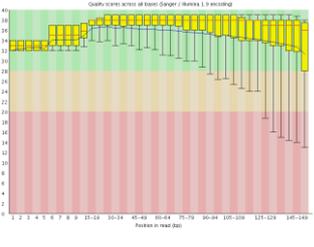
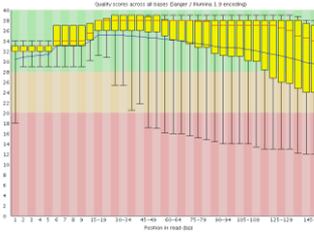
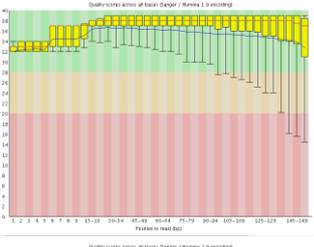
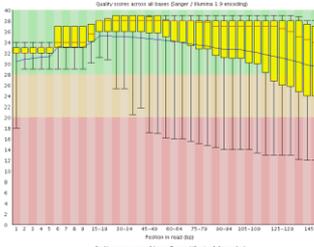
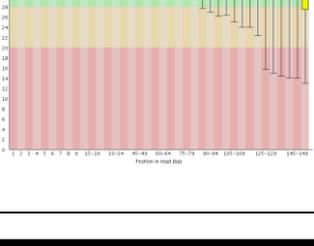
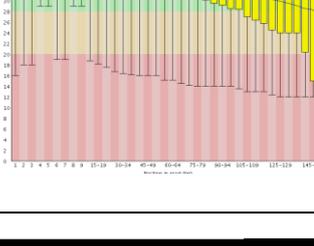
e) Análisis bioinformático:

- Lecturas obtenidas de la secuenciación.

Los parámetros establecidos para la realización de la secuenciación en el equipo MiSeq (Illumina) fueron: profundidad de 20x-200x, cobertura >98%, lecturas “paired-end” con un tamaño de lectura de 150 nucleótidos. Debido a que la secuenciación fue “paired-end”, se obtuvieron dos lecturas por cada muestra, una lectura sentido 5'→3' (R1) y la otra lectura antisentido 3'→5' (R2).

Se realizó el análisis de calidad de las lecturas obtenidas en el software FastQC, se obtuvo información como: calidad, total de lecturas, el tamaño de lecturas y el contenido de G y C. En la Cuadro 16 se indica los parámetros obtenidos de cada una de las 6 muestras analizadas. En cuanto a la calidad, estas se definieron en aquellas lecturas que tuvieron una  $Q \geq 30$  (1 error cada 1,000 bases), que se observó en cada una de las figuras en la zona de color verde. Sin embargo, en todas las lecturas al final de estas se observó un decaimiento de la calidad, lo cual fue un indicativo de lecturas de mala calidad que fue inherente al proceso de secuenciación.

**Cuadro 16.** Lecturas obtenidas de la secuenciación.

Muestra	Parámetro		R1		R2
1TB	Lecturas totales	493,075		493,075	
	Tamaño de las lecturas	35-151		35-151	
	%GC	57		57	
2TB	Lecturas totales	257,676		257,676	
	Tamaño de las lecturas	35-151		35-151	
	%GC	56		56	
3TB	Lecturas totales	212,833		212,833	
	Tamaño de las lecturas	35-151		35-151	
	%GC	63		63	
4TB	Lecturas totales	286,910		286,910	
	Tamaño de las lecturas	35-151		35-151	
	%GC	62		62	
5TB	Lecturas totales	344,145		344,145	
	Tamaño de las lecturas	35-151		35-151	
	%GC	63		63	
6TB	Lecturas totales	206,321		206,321	
	Tamaño de las lecturas	35-151		35-151	
	%GC	63		63	

Por cada muestra se obtuvo un promedio de 600,320 lecturas, no obstante, se esperaba que todas las lecturas tuvieran un tamaño de 150 nucleótidos, de los cuales se obtuvieron lecturas que van desde los 35 a los 151 nucleótidos. El contenido de G-C en el genoma de MTB se encontró alrededor de 65%, obteniéndose un promedio de 59% de esta relación (Cuadro 17).

**Cuadro 17.** Promedios de los parámetros obtenidos por muestra.

Total de lecturas (R1+R2)	Tamaño de lecturas	Contenido de G-C (%)
600,320	35-151 nucleótidos	59

- Pre-procesamiento de las lecturas de secuenciación.

De acuerdo al reporte obtenido del software FastQC (Cuadro 17) se observaron lecturas de baja calidad ( $Q < 30$ ), por lo cual se realizó el filtrado de las lecturas crudas, cortando el extremo de menor calidad dejando que al menos el 90% de las lecturas que tuvieran una  $Q \geq 30$ . Además se incluyó la remoción de lecturas provenientes de adaptadores. De este proceso se perdieron alrededor de 33.1% de lecturas totales, dejando un promedio de 414,967 de lecturas (67.9%) para los posteriores pasos del análisis bioinformático (Cuadro 18).

**Cuadro 18.** Filtrado de las lecturas obtenidas de la secuenciación.

Muestra	Total de lecturas crudas	Pre-procesamiento		ADN humano filtrado		Lecturas filtradas		Alineamiento con genoma de MTB	
		Lecturas	%	Lecturas	%	Lecturas	%	Lecturas	%
1 TB	986,150	744,466	75.5	87,402	8.9	657,064	66.6	331,854	33.7
2 TB	515,352	376,910	73.1	5,296	1.0	371,614	72.1	216,194	42.0
3 TB	425,666	273,364	64.2	19,494	4.6	253,870	59.6	216,204	50.8
4 TB	573,820	371,646	64.8	48,989	8.5	322,657	56.2	258,883	45.1
5 TB	688,290	472,202	68.6	184,424	26.8	287,778	41.8	171,570	24.9
6 TB	412,642	251,214	60.9	8,015	1.9	243,199	58.9	198,543	48.1
<b>Promedio</b>	<b>600,320</b>	<b>414,967</b>	<b>67.9</b>	<b>58,937</b>	<b>8.6</b>	<b>356,030</b>	<b>59.2</b>	<b>232,208</b>	<b>40.8</b>

\*Porcentajes establecidos a partir del total de lecturas crudas.

El siguiente paso fue la filtración del ADN humano y de acuerdo al establecimiento de la hibridación con las sondas de captura de genes (SureSelect, Agilent Technologies), se observó que estas sondas capturaron ADN humano con un rendimiento del 1.4% (Cuadro 15), lo que pudo impactar en el proceso de secuenciación, ya que al filtrar las lecturas en promedio se obtuvieron que el 8.6% correspondieron al ADN humano (Cuadro 18).

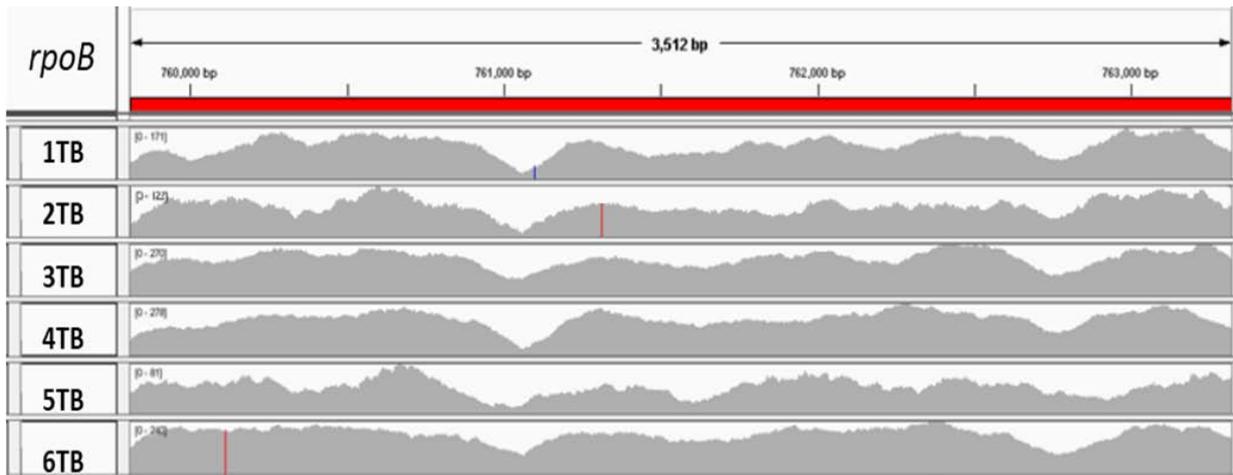
A partir de las secuencias filtradas, se realizó el alineamiento con el diseño de las sondas de captura de genes y el alineamiento con el genoma de referencia de MTB (H37Rv), con lo que se aseguró que todas las lecturas sobrantes correspondieron a MTB, de las cuales se obtuvieron un promedio del 40.8%, es decir menos de la mitad de las lecturas totales obtenidas de la secuenciación.

Después de realizar el alineamiento con el genoma de referencia de MTB (H37Rv), se obtuvieron las profundidades y coberturas de cada una de las muestra que fueron analizadas (Cuadro 19), mostrando que en promedio se obtuvo una profundidad de 125.7X, es decir que las regiones de interés fueron secuenciadas 125.7 veces con una cobertura del 100%.

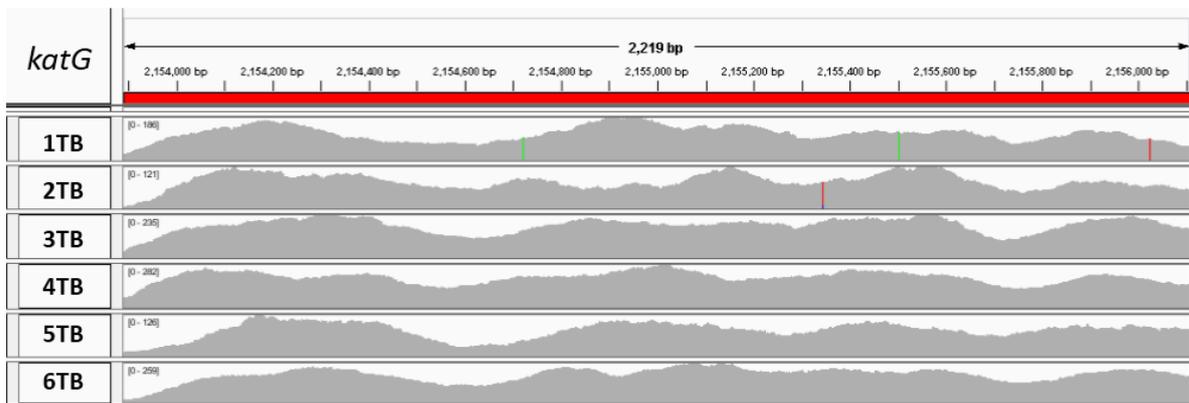
**Cuadro 19.** Profundidad y cobertura obtenida de cada una de las muestras.

<b>Muestra</b>	<b>Lecturas alineadas con el genoma de referencia</b>	<b>Profundidad (X)</b>	<b>Cobertura (%)</b>
1TB	331,854	179.7	100
2TB	216,194	117.0	100
3TB	216,204	117.1	100
4TB	258,883	140.2	100
5TB	171,570	92.9	100
6TB	198,543	107.5	100
<b>Promedio</b>	<b>232,208</b>	<b>125.7</b>	<b>100</b>

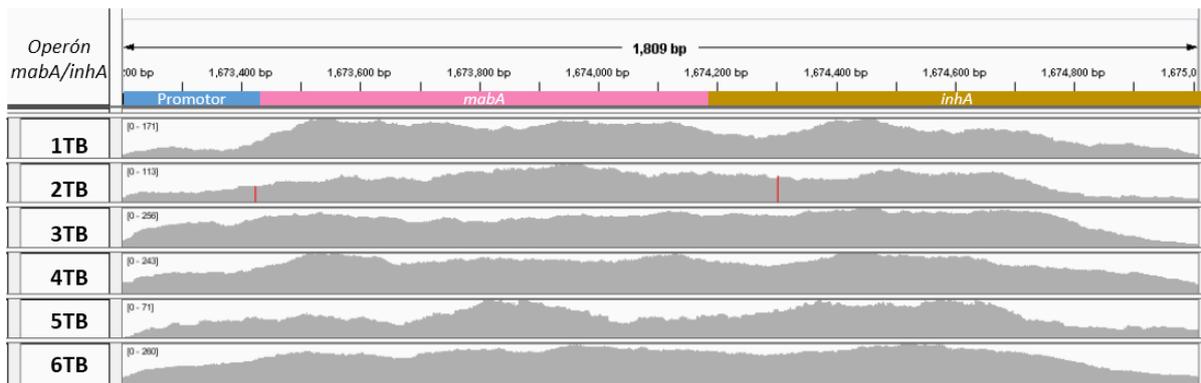
En las figuras 33, 34 y 35, se muestran las coberturas y profundidades de las regiones analizadas de cada una de las muestras. Donde la altura de las barras grises representó el número de lecturas que cubren cada posición y se marcaron con color las variantes encontradas (rojo: T, azul: C, verde: A, café: G), definidas como aquellas posiciones donde el nucleótido difiere de la referencia en más de 10 de lecturas de buena calidad.



**Figura 33.** Profundidad y cobertura del gen *rpoB* de *Mycobacterium tuberculosis*. Utilizando el visualizados del programa IGV.



**Figura 34.** Profundidad y cobertura del gen *katG* de *Mycobacterium tuberculosis*. Utilizando el visualizador del programa IGV.



**Figura 35.** Profundidad y cobertura del operón *mabA/inhA* de *Mycobacterium tuberculosis*. Utilizando el visualizador del programa IGV.

- Variantes genéticas identificadas.

De acuerdo al visualizador de IGV, se observaron el total de variantes encontradas por cada muestra, que se resumen en la Cuadro 20. Teniendo que en la muestra 1TB y 2TB se identificaron cuatro variantes, a diferencias de la muestra 6TB que sólo se identificó una variante. En las muestras 3TB, 4TB y 5TB no se identificaron ningún tipo de variante.

**Cuadro 20.** Variantes totales identificadas por muestra.

Muestra	Gen	Variante genética	Tipo de variante	Posición
1TB	<i>rpoB</i>	A/C	SNV	761,101
	<i>katG</i>	C/A	SNV	2,154,724
		G/A	SNV	2,155,503
		G/T	SNV	2,156,025
2TB	<i>rpoB</i>	C/T	SNV	761,155
	<i>katG</i>	C/T	SNV	2,155,346
	<i>inhA</i>	C/T	SNV	1,674,303
	<i>mabA/inhA</i>	C/T	SNV	1,673,425
3TB	-	-	-	-
4TB	-	-	-	-
5TB	-	-	-	-
6TB	<i>rpoB</i>	C/T	SNV	760,115

Posterior a la identificación de variantes, de acuerdo a la platilla de anotación de la base del software SnpEff utilizando la referencia de MTB (H37Rv), se determinaron los cambios obtenidos al tener una variante en cierta posición de acuerdo al nucleótido cambiado. Esta determinó si la variante identificada cambia o no la codificación del aminoácido (variante sinónima o variante no sinónima) y por ende el impacto en la proteína. Se realizó una predicción del efecto funcional de las variantes no sinónimas, mediante el software PROVEAN donde se determinó, si el cambio de aminoácido afecta o no a la función de la proteína en cuestión (Cuadro 21).

**Cuadro 21.** Variantes genéticas identificadas en las regiones genómicas de interés.

Muestra	Gen	Variante genética	Posición	Cambio de aminoácido	Efecto	Predicción ***	PFS	Concordancia con PFS
<b>1TB</b>	<i>rpoB</i>	A/C	761,101	Gln432Pro-Gln510Pro*	No sinónima	Afecta la función	Resistente a R	Si
	<i>katG</i>	C/A	2,154,724	Arg463Leu	No sinónima	No afecta la función	Sensible a H	Si
		G/A	2,155,503	Thr203Thr	Sinónima	-		
		G/T	2,156,025	Pro29Pro**	Sinónima	-		
	<i>inhA</i>	-	-	-	-	-		
	<i>mabA</i>	-	-	-	-	-		
<i>mabA/inhA</i>	-	-	-	-	-			
<b>2TB</b>	<i>rpoB</i>	C/T	761,155	Ser450Leu-Ser531Leu*	No sinónima	Afecta la función	Resistente a R	Si
	<i>katG</i>	C/T	2,155,346	Ala256Thr**	No sinónima	Afecta la función	Resistente a H	Si
	<i>inhA</i>	C/T	1,674,303	Ala34Ala**	Sinónima	-		
	<i>mabA</i>	-	-	-	-	-		
	<i>mabA/inhA</i>	C/T	1,673,425	-15	Intergénica	-		
<b>3TB</b>	<i>rpoB</i>	-	-	-	-	-	Sensible a R	Si
	<i>katG</i>	-	-	-	-	-	Sensible a H	Si
	<i>inhA</i>	-	-	-	-	-		
	<i>mabA</i>	-	-	-	-	-		
	<i>mabA/inhA</i>	-	-	-	-	-		

\* Variante determinada de acuerdo a la nomenclatura de *Escherichia coli*,

\*\*Nueva variante identificada.

\*\*\* Predicción en función del producto de codificación de cada gen, establecido por el programa PROVEAN.

PFS: Prueba de fármacosensibilidad; H: isoniazida; R: rifampicina.

**Cuadro 21.** Variantes genéticas identificadas en las regiones genómicas de interés. *Continuación.*

Muestra	Gen	Variante genética	Posición	Cambio de aminoácido	Efecto	Predicción ***	PFS	Concordancia con PFS
<b>4TB</b>	<i>rpoB</i>	-	-	-	-	-	Sensible a R	Si
	<i>katG</i>	-	-	-	-	-	Sensible a H	Si
	<i>inhA</i>	-	-	-	-	-		
	<i>mabA</i>	-	-	-	-	-		
	<i>mabA/inhA</i>	-	-	-	-	-		
<b>5TB</b>	<i>rpoB</i>	-	-	-	-	-	Sensible a R	Si
	<i>katG</i>	-	-	-	-	-	Sensible a H	Si
	<i>inhA</i>	-	-	-	-	-		
	<i>mabA</i>	-	-	-	-	-		
	<i>mabA/inhA</i>	-	-	-	-	-		
<b>6TB</b>	<i>rpoB</i>	C/T	760,115	Asp103Asp- Asp184Asp*	Sinónima	-	Sensible a R	Si
	<i>katG</i>	-	-	-	-	-	Sensible a H	Si
	<i>inhA</i>	-	-	-	-	-		
	<i>mabA</i>	-	-	-	-	-		
	<i>mabA/inhA</i>	-	-	-	-	-		

\* Variante determinada de acuerdo a la nomenclatura de *Escherichia coli*,

\*\*Nueva variante identificada.

\*\*\* Predicción en función del producto de codificación de cada gen, establecido por el programa PROVEAN.

PFS: Prueba de fármacosensibilidad; H: isoniazida; R: rifampicina.

## 9. DISCUSIÓN DE RESULTADOS.

### a) Características clínico-demográficas de los pacientes reclutados:

Las muestras de expectoración se obtuvieron a partir de cinco pacientes del sexo masculino (83.3%) y una de un paciente del sexo femenino (16.7%). El intervalo de edad encontrado fue de 34 a 51 años, con una media de 42.3 años. Los pacientes de las muestras 1TB (Oaxaca, 17.7%), 2TB y 3TB (Veracruz, 33.3%) residen en estados de alta incidencia de TB, así como el resto viven en lugares de menor incidencia, tales como los pacientes de la muestras 5TB (Estado de México, 17.7%), 4TB y 6TB (Distrito Federal, 33.3%).

El factor de riesgo que se observó como el más frecuente entre los pacientes fue la diabetes (50%), seguido del alcoholismo (33%), humo de leña (33%) y tabaquismo (33%); dejando en último lugar al contacto con personas con TB activa (17.7%). De acuerdo a la BK realizada en el Laboratorio de Microbiología del INER, tres de los pacientes dieron una BK positiva de 3+ (50%), lo cual indicó una alta carga de BAAR en su expectoración, se tuvieron dos pacientes (33.3%) con una BK de 2+ y sólo uno (17.7%) presentó una BK negativa, esta última no fue indicativo de que el paciente no expectoró bacilos, sino que la carga bacilar podría ser menor como para ser detectada por esta técnica debido a su baja sensibilidad. Los pacientes de las muestras 1TB y 2TB presentaron resistencia a diversos medicamentos (33.3%), estos pacientes fueron previamente tratados, lo que sugiere que esta resistencia se pudo haber adquirido, debido a un mal diagnóstico o tratamiento<sup>2,3</sup>.

La cepa que se aisló de cada uno de los pacientes correspondió al CMTB, debido a que la identificación en un laboratorio de diagnóstico sólo puede discernir entre el CMTB y las micobacterias no tuberculosas. La única especie del CMTB que se puede diferenciar en estas condiciones es *Mycobacterium bovis*, mediante la realización de la prueba bioquímica niacina, pero con el inconveniente de que para realizar dicha prueba es necesario un cultivo maduro de al menos 4 semanas de crecimiento<sup>12,13</sup>.

a) Establecimiento del método de extracción de ADN.

El método de extracción con fenol-cloroformo es el método más empleado, debido al alto rendimiento de ADN puro que se obtiene, es altamente recomendado para la extracción de ADN de MTB<sup>41-51</sup>. Tiene el inconveniente de ser un método sumamente laborioso y puede ser un poco peligroso para el analista, debido a la exposición a quemaduras o por la inhalación de los gases tóxicos del fenol. Por estas últimas razones, se decidió probar métodos comerciales.

Los métodos comerciales de columnas y perlas ofrecieron un bajo rendimiento del 6.8% y del 3.7% respectivamente, así como una pureza que no cumple con los requerimientos para llevar a cabo la SNG ( $A_{260/280} \geq 1.8$ ). A diferencia de los métodos basados en la extracción con el fenol-cloroformo que además de extraer mayor cantidad de ADN, se pudo enriquecer la presencia de ADN de MTB. Las dos variantes que se hicieron a este método, en los cuales se emplearon tratamientos previos a la extracción con fenol-cloroformo, por un lado el tratamiento con proteínasa-lisozima y por otro el tratamiento con DNAsa. El primero dio un mayor rendimiento del 28%, debido a que estas enzimas liberaron la mayor cantidad de ADN de MTB, así como de otros microorganismos que pudieran estar presentes en la muestra sirviendo como acarreador del ADN de MTB. Mientras que el segundo tratamiento, que al haber lisado el ADN humano presente en la muestra por acción de la DNAsa, el rendimiento obtenido fue del 12.2% y aunque fue menor pudiera estar más enriquecido con ADN de MTB. Sin embargo, el problema de este método fue que la cantidad de ADN extraído fue insuficiente para su utilización en etapas posteriores. Por lo tanto, el método que cumplió con las especificaciones para ser utilizado en la generación de bibliotecas y por ende en la secuenciación, fue el método de extracción proteínasa-lisozima/fenol-cloroformo.

Al establecer las diferencias entre los métodos de extracción, mediante la prueba estadística de ANOVA de un factor se determinó que al menos un método fue diferente estadísticamente ( $p=0.004$ ) y de acuerdo a la prueba de Tukey se determinó que el método de proteínasa-lisozima/fenol-cloroformo fue el único diferente estadísticamente a los métodos comerciales empleados ( $p>0.05$ ) y fue casi igual al método de DNAsa/fenol-cloroformo ( $p=0.048$ ), esto era de esperarse debido

a que ambos métodos se basaron en la extracción con fenol-cloroformo con la variante en el tratamiento enzimático realizado previo a la extracción.

b) Extracción de ADN y su control de calidad:

Ya establecido el método de extracción de ADN a partir de muestras de expectoración, se procedió a realizar la extracción de las seis muestras obteniendo la cantidad suficiente de ADN, entre 200ng y 3µg totales para la realización del protocolo de preparación de bibliotecas, con una pureza de alrededor del 1.8. Al determinar la presencia e integridad del ADN de MTB, mediante la amplificación de las regiones de los genes *katG* y *rpoB*, se observaron bandas inespecíficas en algunas muestras (Figura 29), lo que podría tratarse del ADN de otros microorganismos, ya que estos genes constituyen a otros genomas bacterianos que pudieron estar presentes en el ADN total que fue extraído de las muestras de la expectoración<sup>69,70</sup>.

c) Establecimiento de la hibridación con sondas de captura de genes:

Las bibliotecas enriquecidas se analizaron mediante un electroferograma (Figura 32), a partir de diferentes experimentos realizados, desde lo sugerido en el protocolo del fabricante (SureSelect, Agilent Technologies) en el experimento 1 (750ng de biblioteca, con 65°C y 24 horas de hibridación) dando el rendimiento más bajo del 7.9%. Al analizar las curvas obtenidas de los demás experimentos se observó que la mejor curva fue aquella del experimento 3 (1,500ng de biblioteca, con 65°C y 24 horas de hibridación), proporcionando la mayor cantidad de biblioteca enriquecida del 72.7%, la cual tuvo la variante de aumentar la cantidad de biblioteca utilizada de 750ng a 1500ng. No obstante, cuando se observaron los controles negativos que difieren en la temperatura de hibridación empleada se observó que en condiciones establecidas a 65°C (experimento 2) hubo una mayor captura de ADN humano con un rendimiento del 2.9%, en comparación del aumento de la temperatura de hibridación a 68°C (experimento 6) con el que se obtuvo un rendimiento del 1.4%. Lo anterior pudo haber impactado en el rendimiento del experimento 3, ya que al aumentar la cantidad de la biblioteca utilizada, se pudo aumentar la probabilidad de capturar gran cantidad de bibliotecas que contenían ADN humano disfrazando el rendimiento obtenido.

Por lo tanto, la mejor opción fue aumentar la temperatura de hibridación para garantizar la captura de MTB, debido a que en su ADN esta micobacteria posee alrededor del 65% de G y C<sup>12</sup>. Esto nos sugiere que necesita una mayor temperatura alineación (energía calórica) de las bases complementarias, para la formación de 3 puentes de hidrógeno (triple enlace), contribuyendo a hacer una captura más específica<sup>71,72</sup>. Siendo entonces las condiciones del experimento 4, es decir la utilización de 750ng de biblioteca, 68°C y 24 horas de hibridación como las mejores condiciones que fueron empleadas, ofreciendo un rendimiento del 44.8% y la probabilidad de reducir la captura de bibliotecas con ADN humano. Por lo cual fueron elegidas estas condiciones para realizar el procedimiento de la hibridación con las sondas de captura de genes (SureSelect, Agilent Technologies).

Ya establecidas las condiciones de hibridación se prosiguió con la preparación de bibliotecas y la secuenciación de los genes *rpoB*, *katG*, *mabA*, *inhA* y la región promotora de *mabA/inhA*, a partir del ADN total de las muestras de expectoración.

a) Análisis bioinformático:

Después del pre-procesamiento de las lecturas de secuenciación, la filtración del ADN humano y el alineamiento con el genoma de referencia de MTB (H37Rv), se obtuvieron un promedio de 600,320 lecturas crudas (Cuadro 18), en promedio el 40.8% de estas lecturas correspondieron al genoma de MTB. Lo anterior puede estar asociado a lo observado en el PCR realizado para la comprobación de la presencia e integridad del ADN de MTB (Figura 24), probablemente es debido al tipo muestra con la que se trabajó, ya que la expectoración es una muestra muy compleja, pues contiene una gran diversidad de microorganismos que muchos de ellos comparten genes en común como *katG* y *rpoB*, que pudieron ser capturados mediante las sondas y por lo tanto secuenciados, presentándose un fenómeno similar cuando se amplificaron por PCR regiones de estos genes dando lugar a la presencia de bandas inespecíficas.

b) Variantes genéticas encontradas:

Las variantes genéticas encontradas se trataron de SNV's, se identificaron variantes sinónimas, es decir sin cambios en la codificación del aminoácido y variantes no sinónimas, las cuales al sustituir un nucleótido cambió la codificación a otro

aminoácido, que depende de la localización de la variante y si esta se encuentra en una región muy conservada la probabilidad de que se afecte la función de la proteína es muy alta, este parámetro se evaluó mediante el software PROVEAN<sup>73,74</sup>.

- Muestra 1TB:

En esta muestra fueron identificadas en total cuatro variantes en los genes *rpoB* y *katG*. En el gen *rpoB* se determinó una variante en la posición 761,101 (A/C) que generó un cambio de aminoácidos en la posición 510 (Gln→Pro) en la subunidad β de la ARN polimerasa, afectando su función de acuerdo a la predicción realizada con el software PROVEAN. Esta variante sólo ha sido reportada en México en un estudio realizado por Zenteno-Cuevas y colaboradores en el estado de Veracruz<sup>51,52</sup>, tomando en cuenta este antecedente y el hecho de que esta muestra provino de un paciente del estado de Oaxaca, lo que podría sugerir la presencia local de esta variante en México.

En cuanto al gen *katG*, fueron identificadas tres variantes, de las cuales dos se localizaron en las posiciones 2,155,503 (G/A) y 2,155,025 (G/T) que no generaron cambios de aminoácidos, estos se localizaron en las posiciones 29 (Pro→Pro) y 203 (Thr→Thr) respectivamente, por lo que se sugiere que no afectan la función de la enzima catalasa-peroxidasa. La primera no ha sido reportada en otros estudios y no se conoce si altere la función de la enzima catalasa-peroxidasa, aunque es posible predecir que no se afecte la función por tratarse de una variante sinónima. La variante en la posición 2,155,503 (G/A) ya ha sido reportada en asociación a la evolución del CMTB, sin asociación a resistencia a H<sup>76</sup>.

La tercera variante que se identificó en el gen *katG*, se encontró en la posición 2,154,724 (C/A) que generó un cambio de aminoácidos en la posición 463 (Arg→Leu) en la enzima catalasa-peroxidasa y de acuerdo al software PROVEAN se predijo que no se afecta la función de esta enzima. Esta variante ya ha sido reportada en otros estudios sin asociación a la resistencia a H, por lo que la predicción está en acuerdo con lo reportado anteriormente; además de que está asociada al proceso evolutivo de las micobacterias del CMTB y es característica de *Mycobacterium bovis* que sirve como marcador filogenético<sup>76</sup>, lo cual se puede sustentar con el hecho de que la cepa aislada presentó fenotípicamente resistencia

a Z (Cuadro 9), esta resistencia es intrínseca de *Mycobacterium bovis*<sup>53</sup>, pero para asegurar de que se trate específicamente de esta micobacterias es necesario la implementación de otras técnicas moleculares como el “spoligotyping”<sup>1</sup>.

Esta muestra presentó una resistencia fenotípica a R, que concuerda con lo encontrado en la identificación de variantes, ya que la única variante que tiene un efecto funcional fue la encontrada en la posición 761,101 (A/C) en el gen *rpoB*, dando un genotipo resistente a R y en el caso de los genes asociados a la resistencia a H, no se encontró ninguna variante que afecte a ninguna de las funciones de regiones génicas del operón *mabA/inhA*, prediciendo un genotipo sensible a H.

- Muestra 2TB:

En esta muestra fueron encontradas en total cuatro variantes. En el gen *rpoB* se identificó una variante en la posición 761,155 (C/T) que generó un cambio de aminoácidos en la posición 531 (Ser→Leu) que de acuerdo al software PROVEAN, esta variante afecta la función de la subunidad  $\beta$  de la ARN polimerasa. Esta variante ya ha sido reportada y es una de las variantes más frecuentes asociadas a la resistencia a R, debido a que se localiza en la región RRDR que ha sido ampliamente estudiada<sup>30,35</sup>.

En el gen *katG* se identificó una variante en la posición 2,155,346 (C/T) que generó un cambio de aminoácido en la posición 256 (Ala→Thr) en la enzima catalasa-peroxidasa, esta variante no ha sido reportada en ningún estudio previo y se predijo que afecta la función de dicha enzima de acuerdo al software PROVEAN, lo cual establecería una nueva variante asociada a la resistencia a H en MTB, debido a que variantes en el gen *katG* generan la principal vía de resistencia a dicho fármaco. Al analizar la presencia de las variantes genéticas en el operón *mabA/inhA* se

---

<sup>1</sup> **Spoligotyping:** Método basado en la presencia o ausencia de 43 secuencias (llamados espaciadores), dentro de una región específica del genoma de MTB denominada región de “repetidos directos”, esta región está conformada por secuencias repetidas intercaladas por espaciadores no repetidos, el conjunto de las secuencias repetidas y un espaciador no repetido dan lugar a la unidad Variable de Repeticiones Directas, las cuales son altamente variables entre las diversas cepas de MTB tanto por deleciones como por inserciones. Para su identificación es necesaria la amplificación de la zona “repetidos directos” por medio de PCR utilizando el primer reverso marcado con biotina, después de la obtención del material amplificado, este es hibridado con una membrana de nylon con oligonucleótidos inmovilizados que se unen covalentemente a los oligonucleótidos de la muestra<sup>16,17</sup>.

encontraron dos variantes, una variante en la posición 1,674,303 (C/T) en el gen *inhA* que no generó cambio de aminoácidos en la posición 34 (Ala→Ala) en la enoil ACP reductasa, esta variante no ha sido reportada anteriormente. La otra variante que se identificó, se encontró en la región promotora en la posición 1,673,425 en el codón -15 (C/T), esta es una de las variantes más comunes asociadas a la resistencia a H y se sabe que aumenta la expresión de la enoil ACP reductasa sobrepasando el poder inhibitorio de la H, dando lugar a la resistencia a este fármaco<sup>31</sup>.

De acuerdo a los resultados obtenidos de la PFS, la cepa que se aisló de este paciente fue resistente a R, H y Eto. Al haber analizado el gen *rpoB*, se identificó una variante en la posición 761,155 (C/T) que se predijo como un genotipo resistente a R, que concuerda con el fenotipo obtenido. Además se identificaron dos variantes, una en la posición 2,155,346 (C/T) en el gen *katG* y el cambio de una C→T en la posición 1,673,425 de la región promotora del operón *mabA/inhA*, lo cual cumplió con uno de los mecanismos propuestos de la resistencia a H, cuando se tienen ambas variantes se presenta un mayor nivel de resistencia a H<sup>29,30,77</sup>. Otro hecho importante es que *inhA* es también blanco de la Eto, la presencia de una variante en el codón C-15T en la región promotora, esta aumenta la expresión de la enoil ACP reductasa que podría inhibir también a la Eto por una resistencia cruzada<sup>30</sup>, correlacionando los resultados de la PFS con el genotipo propuesto.

- Muestras 3TB, 4TB y 5TB:

Al realizar el análisis de estas muestras no se encontraron variantes en las regiones genómicas analizadas, por lo que los resultados obtenidos del genotipo concordaron con los resultados de las PFS's, indicando que en los tres casos se trataron de cepas sensibles.

- Muestra 6TB:

En esta muestra sólo se encontró una variante en gen *rpoB* en la posición 760,115 (C/T) que no generó un cambio de aminoácidos en la posición 184 (Asp→Asp) en la subunidad  $\beta$  de la ARN polimerasa, esta variante ya ha sido reportada sin asociación a la resistencia a R<sup>78</sup> y se predijo que no afecta la función de la subunidad  $\beta$  de la ARN polimerasa por el hecho de que se trató de una variante no sinónima, lo que

sugiere un genotipo de sensible a R. Este resultado concuerda con lo obtenido a la PFS, la cual indicó que esta cepa fue sensible a los fármacos de primera línea.

Debido a que el tamaño de muestra empleado fue muy pequeño (N=6), es necesario aumentarlo para que los resultados puedan ser sometidos a un análisis estadístico y de esta manera poder establecer tal vez en un futuro una prueba diagnóstica. Sin embargo, con este estudio se demostró que la tecnología de secuenciación masiva o SNG puede ser utilizada como una herramienta molecular que ofrece información sobre el total de variantes genéticas potencialmente asociadas a FR. Además de que sirvió como prueba de concepto, pues la preparación de la muestra sólo incluye un paso de extracción de ácidos nucleicos directamente en expectoración, previo al análisis genómico, ahorrando así semanas de cultivo microbiológico y PFS. Lo anterior podría ser aprovechado para en un futuro apoyar el diagnóstico y selección de medicamento antituberculosis previo a su administración.

## 10. CONCLUSIONES.

- La técnica proteínasa-lisozima/fenol-cloroformo favoreció la de extracción de ADN de MTB a partir de muestras de expectoración para su posterior utilización en la SNG.
- Se establecieron las condiciones para la preparación de librerías de secuenciación mediante captura de genes con sondas sintéticas (SureSelect (Agilent Technologies)), que consistieron en la utilización de 750ng de biblioteca y 24 horas de hibridación a 68°C.
- El análisis de la secuencia completa de genes de MTB permitió identificar variantes genéticas asociadas a FR y otras variantes específicas del CMTB, obteniéndose tanto variantes ya reportadas así como nuevas variantes.
- Las variantes genéticas asociadas a resistencia a H y R que fueron observadas en este estudio concordaron con los resultados de las PFS's.
- Este estudio demostró la utilidad de la SNG de ADN como herramienta molecular, como apoyo de la determinación de FR, que podría agilizar el diagnóstico y la terapéutica contra la TB.

## 11.PERSPECTIVAS.

El estudio realizado permitió la identificación de variantes nuevas y ya conocidas, asociadas a resistencia y otras variantes genéticas que se han pasado de generación en generación asociadas a la evolución del CMTB.

En cuanto a la posible nueva variante identificada en la posición 2,155,346 en el gen *katG* en el codón 256, que se predijo que afecta a la función de la enzima catalasa-peroxidasa, es necesario realizar la demostración bioquímica para confirmar su asociación a resistencia a H. Sin embargo, se necesita aumentar el número de muestras para poder evaluar estadísticamente la presencia de esta variante y poder establecer si es propia de pacientes mexicanos.

El análisis de la FR mediante la secuenciación de regiones genómicas completas, sugiere una alternativa prometedora para apoyar en un futuro el diagnóstico mediante la generación de pruebas diagnósticas rápidas, que puedan ofrecer mayor información acerca de la resistencia a medicamentos y de esta manera contribuir a la elección del mejor esquema terapéutico en cada paciente.

Actualmente, existen investigaciones en el laboratorio de Farmacogenómica en el INMEGEN que buscan la identificación de variantes genómicas en MTB, en más de 100 genes potencialmente asociados a FR.

## 12. REFERENCIAS.

1. Norma Oficial Mexicana NOM-006-SSA2-1993, Para la prevención y control de la tuberculosis en la atención primaria a la salud. [Consultado 18 Noviembre 2014]. Disponible en: <http://www.salud.gob.mx/unidades/cdi/nom/m006ssa23.html>
2. Farga V., Caminero J. Tuberculosis. 3ra Ed. Buenos Aires: Editorial Mediterraneo; 2011.
3. Caminero J. Guidelines for clinical and operational management of drug-resistant tuberculosis. Paris: IJTLD. 2013.CDC, tuberculosis. Disponible en: <http://www.cdc.gov/tb/topic/basics/default.htm>
4. Secretaria de Salud. Perfil epidemiológico de la Tuberculosis en México. Disponible en: [http://www.epidemiologia.salud.gob.mx/doctos/infoepid/publicaciones/2012/Monografias5\\_Tuberculosis\\_Mex\\_junio12.pdf](http://www.epidemiologia.salud.gob.mx/doctos/infoepid/publicaciones/2012/Monografias5_Tuberculosis_Mex_junio12.pdf)
5. Global tuberculosis control: WHO report 2014: World Health Organization; 2014. [Consultado 18 Noviembre 2014]. Disponible en: [http://www.who.int/tb/publications/global\\_report/en/](http://www.who.int/tb/publications/global_report/en/)
6. Parrish NM, Dick JD, Bishai WR. Mechanisms of latency in Mycobacterium tuberculosis. Trends Microbiol. 1998;6(3):107-12.
7. Koul A, Arnoult E, Lounis N, Guillemont J, Andries K. The challenge of new drug discovery for tuberculosis. Nature. 2011;469(7331):483-90
8. Reunión Regional de Jefes de Programas Nacionales de Control de la Tuberculosis Organización Panamericana de la Salud Programa Regional de Tuberculosis Ciudad de México, México - 26 al 28 de mayo del 2014. Disponible en: [http://www.paho.org/hq/index.php?option=com\\_topics&view=article&id=59&Itemid=40776](http://www.paho.org/hq/index.php?option=com_topics&view=article&id=59&Itemid=40776)
9. CENAPRECE. Hospital General de México: Situación actual de la Tuberculosis en México y el mundo. 2015. Disponible en: <http://www.cenaprece.salud.gob.mx/programas/interior/micobacteriosis/descargas/pdf/1SituacionEpidemiologicaTB.pdf>.
10. PA, Emdelaom. Estrategia mundial de la OMS para contener la resistencia a los antimicrobianos. Disponible en: <http://www.who.int/drugresistance/SpGlobal2.pdf>
11. Palomino JC, Leão SC, Ritacco V. Tuberculosis 2007; from basic science to patient care. 2007.
12. Manual para el diagnóstico bacteriológico de la tuberculosis. Normas y guía técnica. Parte II cultivo. OPS Washington eD. C DC; 2008.
13. Gorocica P, Jiménez-Martínez MdC, Garfias Y, Sada I, Lascurain R. Componentes glicosilados de la envoltura de Mycobacterium tuberculosis que intervienen en la patogénesis de la tuberculosis. Rev Inst Nal Enf Resp. 2005;18(2):142-53.
14. Cole S, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature. 1998;393(6685):537-44.
15. Cole ST. Comparative and functional genomics of the Mycobacterium tuberculosis complex. Microbiology. 2002;148(10):2919-28.
16. Kamerbeek J, Schouls L, Kolk A, Van Agterveld M, Van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of

- Mycobacterium tuberculosis for diagnosis and epidemiology. *J Clin Microbiol.* 1997;35(4):907-14.
17. Barnes PF, Cave MD. Molecular epidemiology of tuberculosis. *N Engl J Med.* 2003;349(12):1149-56.
  18. Benedek TG. The history of gold therapy for tuberculosis. *J Hist Med Allied Sci.* 2004;59(1):50-89.
  19. Caminero Luna JA. Guía de la tuberculosis para médicos especialistas. Paris: Unión Internacional Contra la Tuberculosis y Enfermedades Respiratorias. 2003:312-23.
  20. Santos-Preciado JI, Franco-Paredes C. Tratamiento Acortado Estrictamente Supervisado (TAES o DOTS) para tuberculosis en poblaciones con niveles moderados de farmacorresistencia: perspectiva del impacto internacional. *Revista de Investigación Clínica.* 2005;57(3):488-90.
  21. Maxmen, A. Drug development: A combined effort. *Nature.* 2013;502(7470):S4-S6.
  22. Campbell PJ, Morlock GP, Sikes RD, Dalton TL, Metchock B, Starks AM, et al. Molecular detection of mutations associated with first-and second-line drug resistance compared with conventional drug susceptibility testing of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother.* 2011;55(5):2032-41.
  23. Guzman JD, Montes-Rincón X, Ribón W. Research and Development of New Drugs Against Tuberculosis. INTECH Open Access Publisher, 2013.
  24. Eldholm V, Norheim G, von der Lippe B, Kinander W, Dahle UR, Caugant DA, et al. Evolution of extensively drug-resistant *Mycobacterium tuberculosis* from a susceptible ancestor in a single patient. 2014;15(11):490.
  25. Canetti G. Present aspects of bacterial resistance in tuberculosis. *Am Rev Respir Dis.* 1965;92(5):687.
  26. Mitchinson D. Mecanismos de acción de los medicamentos en la quimioterapia de corta duración. *UICTER.* 1985;60:3640.
  27. Boldú J, Cebollero P, Abu J, De Prado A. Tratamiento de la tuberculosis pulmonar Treatment of pulmonary tuberculosis. *An Sist Sanit Navar.* 2007;30(Supl 2):99-115.
  28. Cuevas-Córdoba B, Zenteno-Cuevas R. Tuberculosis drogorresistente: mecanismos moleculares y métodos diagnósticos. *Enferm Infecc Microbiol Clin.* 2010;28(9):621-8.
  29. Zhang Y, Yew W. Mechanisms of drug resistance in *Mycobacterium tuberculosis* [State of the art series. Drug-resistant tuberculosis. Edited by CY. Chiang. Number 1 in the series]. *IJTLD.* 2009;13(11):1320-30.
  30. Slayden RA, Barry CE. The genetics and biochemistry of isoniazid resistance in *Mycobacterium tuberculosis*. *Microbes Infect.* 2000;2(6):659-69.
  31. Palomino JC, Martin A. Drug Resistance Mechanisms in *Mycobacterium tuberculosis*. *Antibiotics.* 2014;3(3):317-40.
  32. Goldstein BP. Resistance to rifampicin: a review. *J Antibiot.* 2014;67(9):625-30.
  33. Koch A, Mizrahi V, Warner DF. The impact of drug resistance on *Mycobacterium tuberculosis* physiology: what can we learn from rifampicin? *Emerging Microbes and Infections.* 2014;3(3):e17.
  34. Varma-Basil M, El-Hajj H, Colangeli R, Hazbón MH, Kumar S, Bose M, et al. Rapid detection of rifampin resistance in *Mycobacterium tuberculosis* isolates from India and Mexico by a molecular beacon assay. *J Clin Microbiol.* 2004;42(12):5512-6

35. Chegou NN, Hoek KG, Kriel M, Warren RM, Victor TC, Walzl G. Tuberculosis assays: past, present and future. *Expert Rev Anti Infect Ther.* 2011;9(4):457-69.
36. Canetti G, Froman S, Grosset J, Hauduroy P, Langerová M, Mahler HT, et al. Mycobacteria: laboratory methods for testing drug sensitivity and resistance. *Bull World Health Organ.* 1963;29(5):565-78.
37. BD. BBL MGIT. Mycobacteria Growth Indicator Tube, OADC Enrichment, PANTA Antibiotic Mixture. Inserto de medio de cultivo. 2012.
38. BD. BACTEC MGIT 960 SIRE kit. Inserto de medio de cultivo. 2012.
39. Secretaria de Salud, Dirección General de Epidemiología, Instituto de Diagnostico y Referencia Epidemiologicos. Lineamientos para la Vigilancia Epidemiológica de la Tuberculosis por laboratorio, 2015. [Consultado 18 de Enero 2015]. Disponible en: [http://www.indre.salud.gob.mx/opencms/opencms/sites/indre/descargas/pdf/Lineamientos/lineamientos\\_para\\_la\\_vigilancia\\_de\\_tuberculosis.pdf](http://www.indre.salud.gob.mx/opencms/opencms/sites/indre/descargas/pdf/Lineamientos/lineamientos_para_la_vigilancia_de_tuberculosis.pdf)
40. Lopez-Alvarez R, Badillo-Lopez C, Cerna-Cortes JF, Castillo-Ramirez I, Rivera-Gutierrez S, Helguera-Repetto AC, et al. First insights into the genetic diversity of Mycobacterium tuberculosis isolates from HIV-infected Mexican patients and mutations causing multidrug resistance. *BMC Microbiol.* 2010;10:82.
41. Frey KG, Herrera-Galeano JE, Redden CL, Luu TV, Servetas SL, Mateczun AJ, et al. Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood. *BMC genomics.* 2014;15(1):96.
42. Miller RR, Montoya V, Gardy JL, Patrick DM, Tang P. Metagenomics for pathogen detection in public health. *Genome Med.* 2013;5(9):81.
43. Barba M, Czosnek H, Hadidi A. Historical perspective, development and applications of next-generation sequencing in plant virology. *Viruses.* 2014;6(1):106-36.
44. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, et al. Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol.* 2012;30(5):434-9.
45. Flores-Trevino S, Mendoza-Olazarán S, Garza-Gonzalez E. Drug resistance and molecular epidemiology of Mycobacterium tuberculosis in Mexico: a systematic review. *Salud Publica Mex.* 2014;56(1):63-77.
46. Alvarado-Esquivel C, Rossau R, Martínez-García S, Cisneros-Martínez J, Mijs W, Nevarez-Najera A, et al. Characterization of rpoB gene mutations in rifampicin resistant Mycobacterium tuberculosis strains isolated from pulmonary tuberculosis patients at 5 Mexican public hospitals. *Rev Invest Clin.* 2000;53(6):526-30.
47. Varma-Basil M, El-Hajj H, Colangeli R, Hazbón MH, Kumar S, Bose M, et al. Rapid Detection of Rifampin Resistance in Mycobacterium tuberculosis Isolates from India and Mexico by a Molecular Beacon Assay. *J Clin Microbiol.* 2004;42(12):5512-6.
48. Bobadilla-del-Valle M, Ponce-de-León A, Arenas-Huertero C, Vargas-Alarcon G, Kato-Maeda M, Small PM, et al. rpoB Gene mutations in rifampin-resistant Mycobacterium tuberculosis identified by polymerase chain reaction single-stranded conformational polymorphism. *Emerg Infect Dis.* 2001;7(6):1010-3.
49. Viader-Salvadó JM, Luna-Aguirre CM, Reyes-Ruiz JM, Valdez-Leal R, de los Angeles del Bosque-Moncayo

- M, Tijerina-Menchaca R, et al. Frequency of mutations in *rpoB* and codons 315 and 463 of *katG* in rifampin-and/or isoniazid-resistant *Mycobacterium tuberculosis* isolates from northeast Mexico. *Microb Drug Resist.* 2003;9(1):33-8.
50. Zenteno-Cuevas R, Zenteno J, Cuellar A, Cuevas B, Sampieri C, Riviera J, et al. Mutations in *rpoB* and *katG* genes in *Mycobacterium tuberculosis* isolates from the Southeast of Mexico. *Mem Inst Oswaldo Cruz.* 2009;104(3):468-72.
51. Molina-Torres CA, Moreno-Torres E, Ocampo-Candiani J, Rendon A, Blackwood K, Kremer K, et al. *Mycobacterium tuberculosis* Spoligotypes in Monterrey, Mexico. *J Clin Microbiol.* 2010;48(2):448-55.
52. Ramaswamy SV, Dou S-J, Rendon A, Yang Z, Cave MD, Graviss EA. Genotypic analysis of multidrug-resistant *Mycobacterium tuberculosis* isolates from Monterrey, Mexico. *J Med Microbiol.* 2004;53(2):107-13.
53. Garza-González E, González GM, Rentería A, Cruz-Pulido W, Rivera G, Bocanegra-García V. A pyrosequencing method for molecular monitoring of regions in the *inhA*, *ahpC* and *rpoB* genes of *Mycobacterium tuberculosis*. *Clin Microbiol Infect.* 2010;16(6):607-12.
54. Bolado-Martínez E, Pérez-Mendoza A, Alegría-Morquecho FM, Candia-Plata MdC, Aguayo-Verdugo MdR, Álvarez-Hernández G. DNA mutations associated to rifampicin or isoniazid resistance in *M. tuberculosis* clinical isolates from Sonora, Mexico. *Salud Publica Mex.* 2012;54(2):167-70.
55. Cruz-Correa, O.F. Pruebas genómicas prioritarias para la salud pública de México: Aspectos generales y el caso de farmacorresistencia en tuberculosis. España: Académica Española; 2013.
56. Organización Mundial de la Salud. Manual de bioseguridad en el laboratorio de Tuberculosis, 2013. [Consultado 25 de Septiembre 2014]. Disponible en: [http://apps.who.int/iris/bitstream/10665/92661/1/9789243504636\\_spa.pdf](http://apps.who.int/iris/bitstream/10665/92661/1/9789243504636_spa.pdf)
57. Qiagen. DNeasy Blood & Tissue Kit. [Consultado 10 de Noviembre 2014]. Disponible en: <http://www.qiagen.com/products/catalog/sample-technologies/dna-sample-technologies/genomic-dna/dneasy-blood-and-tissue-kit#productdetails>
58. Zymo Reserch. Zymobead Genomic DNA kit. [Consultado 10 de Noviembre 2014]. Disponible en: <http://www.zymoresearch.com/downloads/dl/file/id/17/d3004i.pdf>
59. Ekblom R, Wolf JB. A field guide to whole-genome sequencing, assembly and annotation. *Evol Appl.* 2014;7(9):1026-42.
60. Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, Brockman W, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol.* 2009;27(2):182-9.
61. Agilent Technologies. SureSelectXT Target Enrichment System for Illumina Paired-End Sequencing Library. Protocolo Versión 1.8. 2014. [Consultado 25 de Noviembre 2014]. Disponible en: [http://www.chem.agilent.com/Library/usermanuals/Public/G7530-90000\\_SureSelect\\_IlluminaXTMultiplexed\\_1.8.pdf](http://www.chem.agilent.com/Library/usermanuals/Public/G7530-90000_SureSelect_IlluminaXTMultiplexed_1.8.pdf)
62. Agilent Technologies. Agilent High Sensivity DNA kit Guide, 2013. [Consultado 15 de Marzo 2014]. Disponible en: [http://www.chem.agilent.com/library/usermanuals/Public/G2938-90321\\_SensitivityDNA\\_KG\\_EN.pdf](http://www.chem.agilent.com/library/usermanuals/Public/G2938-90321_SensitivityDNA_KG_EN.pdf)

63. Mardis ER. Next-generation sequencing platforms. *Annu Rev Anal Chem*. 2013;6:287-303.
64. Illumina. MiSeq System User Guide. 2013. [Consultado 25 de Noviembre 2014]. Disponible en: [http://support.illumina.com/download/miseq\\_system\\_user\\_guide\\_15027617.html](http://support.illumina.com/download/miseq_system_user_guide_15027617.html)
65. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Meth*. 2012;9(4):357-9.
66. Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2010;26(5):589-95.
67. Ullibarri Galparsoro L, Pita Fernández S. Medidas de concordancia: el índice de Kappa. *Cad Aten Primaria*. 1999;6:169-71.
68. Abisheganaden J, Lim A, Pua S. Sputum Microbiome And Clinical Features During Acute Exacerbation Of Bronchiectasis. *Am J Respir Crit Care Med*. 2015;191:A2441.
69. Lim YW, Evangelista JS, Schmieder R, Bailey B, Haynes M, Furlan M, et al. Clinical insights from metagenomic analysis of sputum samples from patients with cystic fibrosis. *J Clin Microbiol*. 2014;52(2):425-37.
70. Angerer LM, Stoler MH, Angerer RC. In situ hybridization with RNA probes: an annotated recipe. In *situ hybridization: applications to neurobiology*: Oxford University Press New York; 1987. p. 42-70.
71. Herráez Á, Sánchez ÁH. *Biología Molecular e Ingeniería Genética*: Elsevier Health Sciences; 2012.
72. Rocha EP, Smith JM, Hurst LD, Holden MT, Cooper JE, Smith NH, et al. Comparisons of dN/dS are time dependent for closely related bacterial genomes. *J Theor Biol*. 2006;239(2):226-35.
73. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS one*. 2012;7(10):e46688.
74. Hillemann D, Rüsç-Gerdes S, Boehme C, Richter E. Rapid molecular detection of extrapulmonary tuberculosis by the automated GeneXpert MTB/RIF system. *J Clin Microbiol*. 2011;49(4):1202-5.
75. Frothingham R, Strickland PL, Bretzel G, Ramaswamy S, Musser JM, Williams DL. Phenotypic and Genotypic Characterization of *Mycobacterium africanum* Isolates from West Africa. *J Clin Microbiol*. 1999;37(6):1921-6.
76. Tseng S-T, Tai C-H, Li C-R, Lin C-F, Shi Z-Y. The mutations of *katG* and *inhA* genes of isoniazid-resistant *Mycobacterium tuberculosis* isolates in Taiwan. *J Microbiol Immunol Infect*. 2015;48(3):249-55.

## ANEXOS

## Anexo 1. Consentimiento informado.



**INER**

INSTITUTO NACIONAL DE ENFERMEDADES RESPIRATORIAS  
ISMAEL COSÍO VILLEGAS  
DIRECCIÓN DE INVESTIGACIÓN



Instituto Nacional de  
Medicina Genómica  
INMEGENOMICA

**ESTUDIO: Análisis genómico de fármaco-resistencia en  
*Mycobacterium tuberculosis***

***CARTA DE CONSENTIMIENTO PARA PARTICIPAR EN ESTUDIO DE  
INVESTIGACION***

***Invitación a participar y descripción del proyecto***

Estimado Señor o Señora, se le hace una cordial invitación para participar en este estudio. Este formulario forma parte de un proceso de consentimiento informado para un estudio de investigación, que le proporcionará información útil para decidir si desea o no participar voluntariamente en este estudio. Por favor, tome todo el tiempo que sea necesario para leer este formato, preguntar a la persona que le brinda este documento o al responsable del proyecto de investigación (cuyos datos aparecen más adelante) o incluso usted puede consultarlo con alguien de su confianza.

Una vez que usted haya leído el formulario, y se hayan respondido todas sus preguntas, si desea participar en este estudio de investigación se le pedirá que firme este documento y se le entregará una copia del mismo.

Es importante que sepa que usted que el participar en este estudio no afectará su atención médica en este instituto, ni tampoco está renunciando a ninguno de sus derechos legales por firmar este formulario de consentimiento.

***A continuación se le presentan una serie de preguntas y respuestas sobre el proyecto, que lo guiarán para saber si desea o no participar en el estudio.***

**1. ¿Cuáles son los objetivos de este estudio?**

A usted le han informado que tiene una enfermedad llamada tuberculosis pulmonar (TBP). La TBP es una enfermedad infecciosa ocasionada por una bacteria que se llama *Mycobacterium tuberculosis* (M. tb). La TBP se transmite a través de gotitas de saliva cargadas de la bacteria que son expulsadas de los pacientes enfermos de tuberculosis por lo que las personas cercanas pueden respirar estas gotitas y contagiarse de tuberculosis. La tuberculosis es una enfermedad curable, una vez que se establece el diagnóstico y el tratamiento de manera oportuna. Lamentablemente, el tratamiento de la tuberculosis es largo, requiere de 6 meses cuando se trata de una forma sensible a fármacos, pero en presencia de farmacoresistencia. En México el tratamiento de la tuberculosis es gratuito aun en sus formas resistentes.

El propósito de este estudio es investigar e identificar las variantes genéticas en 200 genes de la bacteria que causa la tuberculosis (*Mycobacterium tuberculosis*) y que potencialmente podrían estar asociados a fármaco-resistencia.

En este proyecto, se plantea estudiar 100 pacientes con diagnóstico de TBP: 50 pacientes con TB farmacoresistente y 50 con TB farmacosensible

**2. ¿Por qué me han pedido que participe?**

Versión: 28 de abril de 2014

RECIBIDA  
16 JUN. 2014  
COMITÉ DE ÉTICA  
INMEGENOMICA

1/5



INSTITUTO NACIONAL DE ENFERMEDADES RESPIRATORIAS  
ISMAEL COSÍO VILLEGAS  
DIRECCIÓN DE INVESTIGACIÓN



**ESTUDIO: Análisis genómico de fármaco-resistencia en  
*Mycobacterium tuberculosis***

**2. ¿Por qué me han pedido que participe?**

Usted ha sido diagnosticado con tuberculosis, por lo que queremos analizar una muestra de su saliva para poder estudiar la información de sus genes, y una muestra de su expectoración (flema) para obtener información genética de la bacteria que causa la tuberculosis, y de esta manera ver la interacción entre ambas.

También se le invita a participar porque al residir en México, queremos identificar si hay una relación entre las distintas cepas de la bacteria y la región geográfica de donde usted proviene.

**3. ¿En qué consiste mi participación?**

Para participar en este estudio, se le solicitarán 2 muestra de expectoración (flema) antes de iniciar el tratamiento de la tuberculosis, además se le solicitará una muestra de saliva la cual se obtendrá mediante un hisopo (o bastoncillo de algodón) que se frotará en el interior de la boca. Este procedimiento no será doloroso, no representa ningún riesgo y no tendrá ninguna molestia como consecuencia del mismo.

También se coleccionarán sus datos personales y la información de su expediente clínico que incluirá su edad, diagnóstico, historial de enfermedades, tratamientos médicos y respuesta a tratamientos.

**4. ¿Cuáles son los riesgos posibles a los que me enfrento por participar en el estudio?**

La toma de muestra de expectoración (flema) no tiene ningún riesgo. La toma de muestra de saliva y/o mucosa no tiene riesgos, sin embargo usted presentará una pequeña molestia al momento de frotar el hisopo en la boca. Es importante que sepa que esta muestra se tomará con equipo totalmente nuevo y esterilizado y será realizado por personas capacitadas para ello.

**5. ¿Cuáles son los beneficios que obtengo por participar en el estudio?**

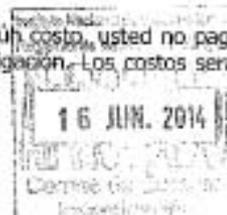
Usted no recibirá beneficios personales por su participación en el estudio. La principal razón por la cual quiera participar en el estudio es la de contribuir a que se tenga un conocimiento más profundo sobre la enfermedad y la interacción genética entre los humanos y la bacteria que causa la tuberculosis. Los resultados de este estudio podrían ser de utilidad para encontrar mejores formas de diagnosticar la enfermedad y la presencia de tuberculosis farmaco resistente, para que de esta manera los pacientes afectados puedan recibir tratamiento oportuno y adecuado. Su participación ayudará a otros pacientes con la misma enfermedad que usted en México y en el resto del mundo.

Usted no recibirá reportes personales acerca de los resultados genéticos o hallazgos que se encuentren en su muestra.

**6. ¿Tiene algún costo participar en el estudio? ¿Se me pagará por participar?**

Participar en este estudio de investigación no tiene ningún costo, usted no pagará ninguno de los estudios que se realizarán en el laboratorio de investigación. Los costos serán pagados por

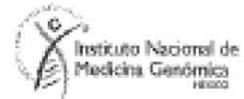
Versión: 28 de abril de 2014



2/5



INSTITUTO NACIONAL DE ENFERMEDADES RESPIRATORIAS  
ISMAEL COSÍO VILLEGAS  
DIRECCION DE INVESTIGACION



**ESTUDIO: Análisis genómico de fármaco-resistencia en  
*Mycobacterium tuberculosis***

muestras de expectoración (flema) será utilizada para este fin, el costo del estudio estará cubierto por el proyecto de investigación y se le entregará una copia de los resultados.

Usted, únicamente tendrá que cubrir los costos de su atención médica y de los estudios de laboratorio que le solicite su médico durante su tratamiento y el seguimiento de acuerdo a la clasificación socioeconómica que se le otorgue.

Usted no recibirá ningún pago, ni incentivo por la participación en el estudio, su participación es voluntaria y altruista.

**7. ¿Participar en este estudio afecta el tratamiento que voy a recibir?**

Es importante que este informado que participar en este estudio no afecta para nada el tratamiento antituberculosis que usted va a recibir.

El tratamiento de la tuberculosis en México es gratuito, aun en sus formas resistentes.

**8. ¿Cómo se manejarán mis muestras y registros médicos?**

Las muestras e información personal obtenida para participar en este estudio es totalmente confidencial. Las muestras de esputo y de mucosa así como los registros médicos serán marcados con un código (codificados) para su seguimiento y con el fin de mantener la confidencialidad. Todos los datos iniciales serán depositados en una base de datos segura, donde sólo un grupo de investigadores y médicos autorizados, que se han comprometido a proteger los datos de los participantes en el estudio, tendrán acceso a la base de datos. Nadie que no esté autorizado podrá tener acceso a su información.

**9. ¿A dónde irán mis muestras e información?**

Las muestras de saliva y expectoración (flema) codificadas, es decir, no etiquetadas con su nombre sino identificadas por código, serán procesadas en el **Instituto Nacional de Medicina Genómica** y el que cuenta con el equipo y el personal capacitado para mantener el resguardo de sus muestras y de su información personal.

Haremos todos los esfuerzos posibles para mantener sus datos confidenciales. Solo un grupo pequeño de investigadores tendrán acceso a la información y estos se mantendrán bajo resguardo.

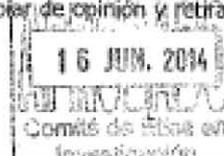
**10. ¿Seré contactado (a) de nuevo en el futuro?**

Si en el futuro se requiere obtener una muestra adicional, para analizar la presencia de la bacteria de tuberculosis en su cuerpo o se requiere información adicional acerca de su estado de salud, una persona del INMEGEN o del INER lo contactará si usted acepta.

**11. ¿Y si en el futuro ya no deseo continuar en el estudio?**

Usted es libre de decidir si desea participar o no en este estudio de investigación. **Si decide no participar, esto no afectará su relación con los médicos de su hospital tratante ni con el INMEGEN.** Si decide participar, puede cambiar de opinión y retirarse del estudio en el

Versión: 28 de abril de 2014



3/5



INSTITUTO NACIONAL DE ENFERMEDADES RESPIRATORIAS  
ISMAEL COSÍO VILLEGAS  
DIRECCIÓN DE INVESTIGACIÓN



**ESTUDIO: Análisis genómico de fármaco-resistencia en  
*Mycobacterium tuberculosis***

momento que lo decida. En este caso, sus muestras de saliva y mucosa, y todos los datos clínicos ligados a las muestras serán destruidos. Por razones de bioseguridad, su muestra no le puede ser devuelta bajo ninguna circunstancia.

**12. ¿Dónde puedo acudir si requiero mayor información acerca del estudio?**

Fue necesario utilizar lenguaje técnico en este formato de consentimiento informado, si tiene alguna pregunta sobre la participación en este estudio o si cree que ha sufrido una lesión relacionada con la investigación, puede contactar a los investigadores a cargo:

Dra. Marcela Muñoz Torrico  
Investigador en Ciencias Médicas A  
Instituto Nacional de Enfermedades Respiratorias (INER)  
Calzada de Tlalpan 4502, Col. Sección XVI C.P. 14080  
México D.F.  
Tel.: 54871712

o

Dra. Minerva Mata Rocha  
Investigador en Ciencias Médicas B  
Instituto Nacional de Medicina Genómica (INMEGEN)  
Periférico Sur 4809, Col. Arenal Tepepan C.P. 14610  
México D.F.  
Tel: 5350 1900 Ext. 1212

Si tiene alguna pregunta sobre sus derechos como sujeto de una investigación, puede llamar a:

Dra. Rocío Chapela  
Presidenta del Consejo de Revisión Institucional  
Instituto Nacional de Enfermedades Respiratorias (INER)  
Tel.: 55-54871702

**ACUERDO PARA PARTICIPAR**

Si Usted ha leído este formato, está de acuerdo en participar, sabiendo que es libre para hacerlo o no hacerlo y que esto no afectaría su atención médica, por favor escriba su nombre completo, firme en el espacio correspondiente y escriba la fecha. Esto se hace por duplicado, para que usted reciba una copia del documento.

Nombre del sujeto: \_\_\_\_\_

Firma del sujeto: \_\_\_\_\_ Fecha: \_\_\_\_\_

Versión: 28 de abril de 2014



4/5



INSTITUTO NACIONAL DE ENFERMEDADES RESPIRATORIAS  
ISMAEL COSÍO VILLEGAS  
DIRECCIÓN DE INVESTIGACIÓN



**ESTUDIO: Análisis genómico de fármaco-resistencia en  
*Mycobacterium tuberculosis***

Investigador/Persona que obtiene el consentimiento: \_\_\_\_\_

Firma: \_\_\_\_\_ Fecha: \_\_\_\_\_

**Testigo 1**

Nombre: \_\_\_\_\_

Firma: \_\_\_\_\_ Fecha: \_\_\_\_\_

**Firma del testigo 2**

Nombre: \_\_\_\_\_

Firma: \_\_\_\_\_ Fecha: \_\_\_\_\_

**USO FUTURO DE MUESTRAS DE SANGRE O TEJIDO SOBRANTES  
O INTENCIÓN DE CONSERVAR LAS MUESTRAS:**

Escriba sus iniciales en cada una de las siguientes afirmaciones que aplique en su caso:

Los investigadores del estudio pueden comunicarse con usted en el futuro para que participe en estudios de investigación adicionales.

Sí \_\_\_\_\_ No \_\_\_\_\_

Sus muestras de sangre/tejido se pueden conservar para estudios futuros.

Sí \_\_\_\_\_ No \_\_\_\_\_

Comprende que si no quiere que sus muestras se conserven y se usen en estudios futuros, no podrá participar en este estudio.







INSTITUTO NACIONAL DE ENFERMEDADES RESPIRATORIAS  
 ISMAEL COSIO VILLEGAS  
 DIRECCION DE INVESTIGACION



**ESTUDIO: Análisis genómico de fármaco-resistencia en  
*Mycobacterium tuberculosis***

**FÁRMACOS DE PRIMERA LÍNEA:** H Isoniacida, R Rifampicina, Z Pirazinamida, E Etambutol, S Estreptomina  
**FÁRMACOS DE SEGUNDA LÍNEA:** 1) Inyectables: Am Amikacina, Km Kanamicina; Cm Capreomicina; 2) Fluoroquinolonas: Cfx Ciprofloxacina, Lfx Levofloxacina, Ofx Ofloxacina; Mfx Moxifloxacina  
 3) Carbodonamidas: Pto Profenamida, Eto Etonamida; 4) Imidafenazina: Cfz Clofazimina; 5) Otras: Tl Terecizóna, Amx/Ac Clav Amoxicilina y Ácido Clavulánico

Otros: \_\_\_\_\_

**ESTATUS BACTERIOLÓGICO AL INICIO DE TRATAMIENTO**

PFS	1
Fecha de toma	
Fecha de reporte	
Laboratorio que realizó	
Muestra	
Folio	
Baciloscopia*	
Cultivo*	
Identificación de la cepa	
Isoniacida (0.1 mcg/ml)	
Isoniacida (0.4 mcg/ml)	
Rifampicina	
Estreptomina (2mcg/ml)	
Estreptomina (8mcg/ml)	
Etambutol (5 mcg/ml)	
Etambutol (8mcg/ml)	
Pirazinamida	
Amikacina	
Ofloxacina	
Levofloxacina	
Profenamida	
Otro1:	
Otro2:	
Tipo de resistencia***	

\* Negativo: 0; Muy escasos: +; Escasos: ++; Numerosos: +++/++++

\*\* S: Sensible, R: Resistente, NR: No realizado

\*\*\* 1. Mono-resistencia, 2. Poli-resistencia, 3. MDR, 4. XDR, 5. XX-DR

Elaboro: \_\_\_\_\_

Fecha: \_\_\_\_\_ Correo  
 electrónico: \_\_\_\_\_

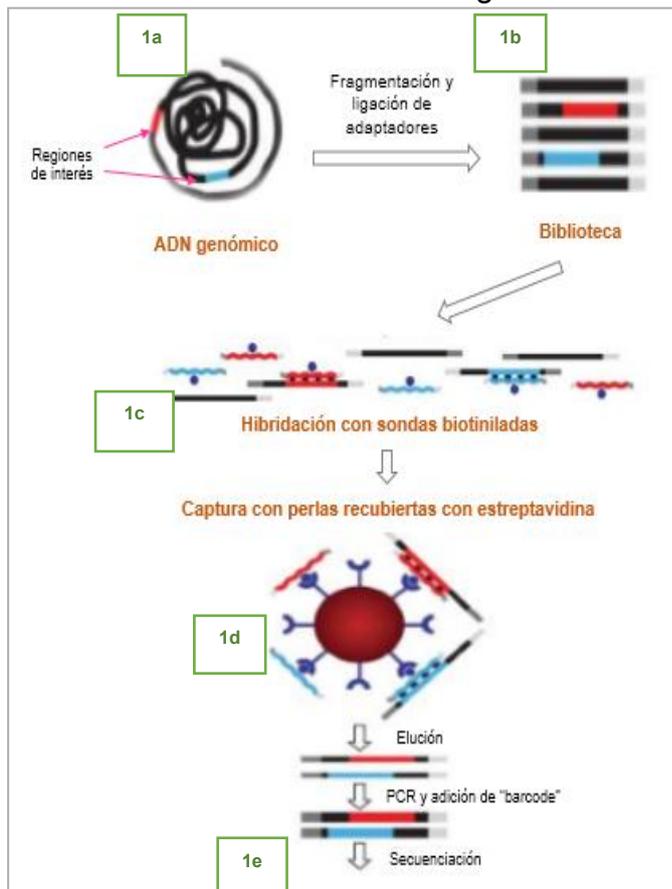


**Apéndice 1.** Enriquecimiento de bibliotecas (SureSelect) y secuenciación por síntesis (Illumina).

b) Enriquecimiento de bibliotecas (SureSelect, Agilent Technologies):

El enriquecimiento de SureSelect, es un conjunto de sondas de ARN que son complementarias a las regiones de interés (Figura 1a) que se desean analizar mediante la secuenciación de ADN. Previamente se debe realizar el proceso de generación de bibliotecas (Figura 1b) mediante la fragmentación del ADN, reparación de extremos, adenilación y la ligación de adaptadores, para dar paso a la hibridación con las sondas de SureSelect, estas sondas u oligonucleótidos de ARN están biotiniladas.

La hibridación consiste en adicionar al templado estas sondas más una solución bloqueadora de ARNasas y diferentes buffers de hibridación, para ser incubados a 95°C durante 5 minutos dando lugar a la desnaturalización del templado, seguido de



una incubación por 24 horas a una temperatura de alineamiento de 65°C. La hibridación de las sondas de ARN con la hebra monocatenaria del templado genera una unión muy estable, aunado a la solución bloqueadora de ARNasas que no permite su degradación durante la larga incubación (Figura 1c).

Posteriormente se realiza la captura de los fragmentos de las bibliotecas hibridadas con las sondas biotiniladas mediante el uso de perlas magnéticas marcadas con estreptavidina y que son capturadas con un potente imán (figura 1d), esto

**Figura 1.** Tecnología de SureSelect para el enriquecimiento de bibliotecas.

Modificado de: Gnirke A, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature biotechnology*. 2009;27(2):182-9.

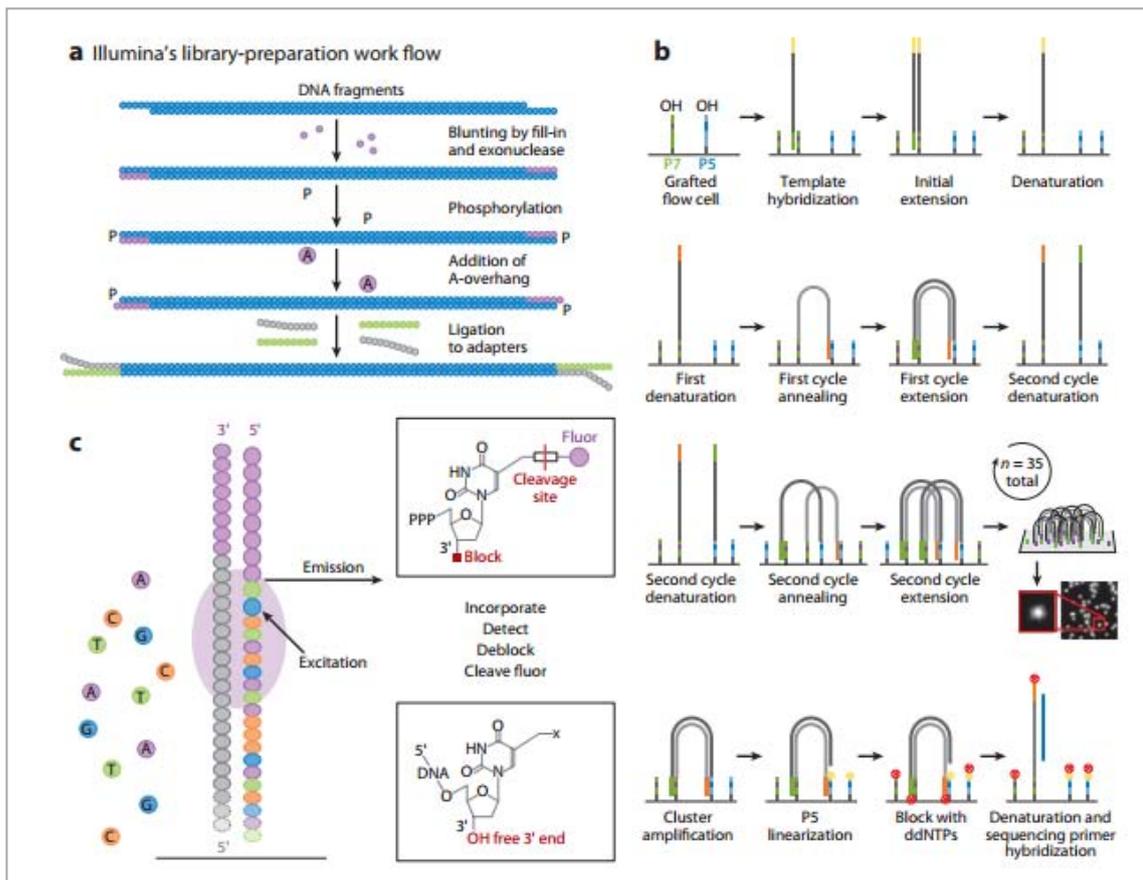
para eliminar las bibliotecas que no contengan los fragmentos de interés mediante el uso de buffers de lavado. Por último se agregan unos ciclos de amplificación para aumentar la concentración de las regiones de interés y la adición de un “barcode” para su identificación (Figura 1e), dando como resultado la biblioteca enriquecida lista para ser cargada en el instrumento de secuenciación.

c) Secuenciación por síntesis (Illumina).

En esta tecnología, se utilizan nucleótidos terminadores al igual que en la secuenciación de Sanger pero estos son marcados con moléculas fluorescentes y basada en la secuenciación masiva en paralelo (es decir la capacidad de realizar millones de secuencias en cada corrida), la diferencia con el método convencional es la posibilidad de eliminar la fluorescencia una vez obtenida la imagen, al desbloquear el carbono 3' de modo que pueda aceptar una nueva base para continuar la reacción de secuenciación, haciendo que la incorporación de un nucleótido terminador sea reversible.

Los primeros pasos consisten en la generación de bibliotecas (Figura 2a), después un paso de amplificación en una superficie sólida denominada celda de flujo o “flow cell”, dónde también se dará la reacción de secuenciación. La celda de flujo contiene un conducto que se compone de un vidrio plano con ocho canales de microflúidos, cada uno decorado con uniones covalentes de adaptadores de las secuencias complementarias a los adaptadores de la biblioteca. Se debe tomar en cuenta la cuantificación de la concentración de biblioteca, ya que se debe agregar una solución diluida de la biblioteca que se amplifica *in situ* en la superficies de la celda de flujo mediante el uso de una etapa de amplificación en puente para producir focos para la secuenciación (clusters) (Figura 1b). Un paso posterior afecta químicamente la liberación de extremos del fragmento que lleva el adaptador, para alinearse con un ADN sintético complementario (cebador), esto para proporcionar grupos 3'-OH que se pueden extender en las reacciones de secuenciación para pasos posteriores. Después se ofrecen los cuatro nucleótidos que son proporcionados en cada ciclo, en el cual cada nucleótido lleva un marcador fluorescente de identificación. La secuenciación se produce como reacciones de adición de un sólo nucleótido debido a que existe un grupo de bloqueo en la posición de la ribosa 3'-OH, evitando reacciones adicionales de incorporación de otras bases por la polimerasa. Como tal,

la serie de eventos en cada paso incluye lo siguiente, en orden de aparición: (a) El nucleótido se añade por la polimerasa, (b) los nucleótidos no incorporados se eliminan por lavado, (c) la celda de flujo forma la imagen en ambas superficies interiores para identificar cada clúster que está reportando una señal fluorescente, (d) los grupos fluorescentes son químicamente eliminados, y (e) el 3'-OH se desbloquea químicamente (Figura 3c). Esta serie de pasos se repite con la capacidad para 150 reacciones de adición de nucleótidos, con lo cual se comienza con la lectura del fragmento pareado. Para hacer la lectura desde el extremo opuesto de cada grupo de fragmentos, el instrumento elimina primero las hebras sintetizadas por desnaturalización y regenera los grupos mediante la realización de un puente limitando la amplificación para mejorar la relación señal-ruido en la segunda lectura. Después de la etapa de amplificación, los extremos opuestos de los fragmentos son liberados de las superficies de la celda de flujo por una sustancia química diferente reactivo de escisión (correspondiente a un grupo lábil en el adaptador inverso), y el cebador inverso se alinea con los fragmentos. La secuenciación procede como se ha descrito anteriormente.



**Figura 2.** Tecnología Illumina: (a) bibliotecas-preparación de la celda de trabajo, (b) generación de clúster mediante la amplificación en puente y (c) secuenciación por síntesis con terminadores marcados reversibles.

Tomado de: Mardis ER. Next-generation sequencing platforms. *Annual review of analytical chemistry*. 2013;6:287-303.