



UNIVERSIDAD PERUANA  
CAYETANO HEREDIA

EVALUACIÓN DEL LÍMITE DE  
DETECCIÓN DE  
HETERORRESISTENCIA DE  
*Mycobacterium tuberculosis* USANDO  
SECUENCIACIÓN NANOPORE CON  
ADN GENÓMICO Y AMPLICONES A  
PARTIR DE MEZCLAS BACTERIANAS  
HETERORRESISTENTES EN  
MUESTRAS DE ESPUTO NEGATIVO

TESIS PARA OPTAR EL GRADO DE  
MAESTRO EN BIOQUÍMICA Y BIOLOGÍA  
MOLECULAR

DIEGO MANUEL RAMOS LETTE

LIMA – PERÚ

2025



**ASESOR**

DR. MIRKO JUAN ZIMIC PERALTA

**CO ASESOR**

DR. PATRICIA SHEEN CORTAVARRIA

**JURADO DE TESIS**

DR. JORGE ENRIQUE RODRIGUEZ BAILON

PRESIDENTE

MG. JUAN CARLOS AGAPITO PANTA

VOCAL

DRA. MARIA LISSETH EGUILUZ MOYA

SECRETARIO (A)

## **DEDICATORIA.**

A mi madre, por su amor incondicional y el impulso silencioso que siempre me ha sostenido en cada una de mis etapas académicas.

A quienes, que con sus palabras y ejemplo, despertaron en mí la pasión por la ciencia y el compromiso con el conocimiento.

A quienes creen en mí y en mis capacidades, por recordarme que siempre se puede seguir adelante.

Y a mí, por resistir cuando parecía más fácil rendirse; por confiar, incluso en medio de la incertidumbre científica.

## **AGRADECIMIENTO.**

Quisiera expresar mi más sincero agradecimiento a todas y cada una de las personas que me acompañaron y apoyaron a lo largo del camino de esta maestría y durante el desarrollo de esta tesis. Este proceso ha representado un esfuerzo significativo en tiempo y dedicación, así como una inversión personal que me brinda la valiosa oportunidad de seguir creciendo profesional y académicamente. Agradezco especialmente a la Dra. Paty Sheen y al Dr. Mirko Zimic, por creer en mí y por brindarme su asesoría constante, su paciencia y su apoyo académico durante todo este proceso.

Extiendo mi gratitud al equipo de Biología Molecular y Bioinformática por su colaboración y respaldo continuo en el desarrollo de mi trabajo especialmente a Carlos, Candy, Diego y Omar con quienes las tardes de trabajo eran más amenas y los debates menos tensos.

De manera muy especial, agradezco a Walter Flemil, Dr. Jorge coronel y Nataly Bailon, por su apoyo clave en los tiempos experimentales dentro del área de Bioseguridad Nivel 3 (P3) de los Laboratorios de Investigación y Desarrollo de la Universidad Peruana Cayetano Heredia.

Solo me queda decir: muchas gracias a todos ustedes.

## **FUENTE DE FINANCIAMIENTO**

Este trabajo fue financiado por el CONCYTEC-PROCIENCIA en el marco del concurso “E041-2023-02 Proyectos de Investigación Aplicada” [CONTRATO N° PE501082878-2023]. Titulado: “Desarrollo y evaluación de un protocolo basado en la secuenciación MinION (Nanopore) para determinar la heterorresistencia en pacientes con tuberculosis directamente de muestras de esputo. Agradezco profundamente este respaldo, el cual permitió la realización de las actividades académicas y experimentales requeridas para el desarrollo de este estudio.

### DECLARACIÓN DE ORIGINALIDAD

Los egresados:

| N° | APELLIDOS Y NOMBRES      |
|----|--------------------------|
| 1. | RAMOS LETTE DIEGO MANUEL |

*(Agregar filas adicionales si hay más autores)*

Pertenecientes al programa de la **MAESTRÍA EN BIOQUÍMICA Y BIOLOGÍA MOLECULAR**, autores del trabajo titulado: **EVALUACIÓN DEL LÍMITE DE DETECCIÓN DE HETERORRESISTENCIA DE *Mycobacterium tuberculosis* USANDO SECUENCIACIÓN NANOPORE CON ADN GENÓMICO Y AMPLICONES A PARTIR DE MEZCLAS BACTERIANAS HETERORRESISTENTES EN MUESTRAS DE ESPUTO NEGATIVO**, el cual ha sido elaborado, sustentado y aprobado, según corresponda, para optar por el grado de **MAESTRO EN BIOQUÍMICA Y BIOLOGÍA MOLECULAR** bajo la modalidad de **TESIS**.

En calidad de docentes asesores de la Universidad Peruana Cayetano Heredia:

| N° | APELLIDOS Y NOMBRES DEL DOCENTE | FACULTAD | NIVEL DE ASESORÍA |
|----|---------------------------------|----------|-------------------|
| 1. | ZIMIC PERALTA MIRKO JUAN        | FACI     | MAESTRÍA          |
| 2. | SHEEN CORTAVARRIA PATRICIA      | FACI     | MAESTRÍA          |


Declaramos que el contenido del presente documento es original y que las citas y referencias a otros autores cumplen con las normas académicas establecidas. En ese sentido, hacemos constar que:

- El documento presenta un porcentaje de similitud de **4%**, según el reporte emitido por el software **Turnitin®** (identificador de entrega: **2847611599**; fecha de entrega: **16-12-2025**).
- Tras una revisión detallada del reporte y del contenido del trabajo en cuestión, no se han identificado indicios de plagio.
- Se certifica que el documento respeta los principios de integridad académica y cumple con los requisitos institucionales de originalidad.

Lugar y fecha: **Lima, 16 de diciembre de 2025**



Firma del asesor  
N° DNI: 07620624  
ORCID: 0000-0002-7203-8847



Firma del Co-asesor  
N° DNI: 09541127  
ORCID: 0000-0002-7118-9301

## ÍNDICE

Resumen

Abstract

|      |   |    |
|------|---|----|
| I.   | Introducción.....   | 1  |
| II.  | Antecedentes.....   | 4  |
| III. | Planteamiento del problema. ....                              | 6  |
| IV.  | Justificación.....  | 8  |
| V.   | OBJETIVOS.....  | 10 |
|      | Objetivo general.....   | 10 |
| VI.  | Pregunta de investigación.....                                | 11 |
| VII. | Marco Teórico.....  | 12 |
|      | 7.1. <i>Mycobacterium tuberculosis</i> .....                  | 12 |
|      | 7.2. Formas clínicas y epidemiología de tuberculosis.....     | 13 |
|      | 7.3. Infección por <i>M. tuberculosis</i> .....               | 14 |
|      | 7.4. Detección de <i>M. tuberculosis</i> .....                | 15 |
|      | 7.4.1. Microscopía de esputo.....                             | 15 |
|      | 7.4.1.1 Protocolos de descontaminación.....                   | 16 |
|      | 7.4.1.1.1 Uso de NALC/NaOH al 0.5%.....                       | 16 |
|      | 7.4.1.1.2 Uso de Saponina al 5%.....                          | 16 |
|      | 7.4.2. Cultivo microbiológico.....                            | 17 |
|      | 7.4.3. Pruebas fenotípicas de susceptibilidad a fármacos..... | 17 |
|      | 7.4.3.1. Método de proporciones en Agar.....                  | 17 |
|      | 7.4.4. Pruebas moleculares rápidas.....                       | 18 |
|      | 7.4.5. Secuenciamiento por nanoporos.....                     | 20 |
|      | 7.4.5.1. Protocolos de extracción de ADN.....                 | 20 |
|      | 7.4.5.1.1. Método fenol.....                                  | 20 |
|      | 7.4.5.1.2. Método con uso de Kits comerciales.....            | 20 |
|      | 7.4.5.2. Tipos de secuenciamiento genómico.....               | 21 |
|      | 7.4.5.2.1. Secuenciamiento de genoma completo.....            | 21 |
|      | 7.4.5.2.2. Secuenciamiento de amplicones.....                 | 21 |
|      | 7.4.5.3. Tipos de librería génica.....                        | 21 |
|      | 7.4.5.3.1 <i>Native Ligation</i> .....                        | 21 |

|   |    |
|---|----|
| 7.4.5.3.2. <i>Rapid Barcoding</i> (SQK-RBK114.24).....  | 22 |
| 7.5. Muestras clínicas y cultivos puros.....  | 22 |
| 7.5.1. Muestra clínica.....   | 22 |
| 7.5.2. Cultivos Puros.....  | 23 |
| 7.6. Tratamiento antituberculosis y resistencia antibiótica .....   | 23 |
| 7.7. Principales fármacos de primera línea y la base genética de su resistencia   | 24 |
| 7.7.1. Isoniazida (H) .....   | 24 |
| 7.7.2. Rifampicina (R) .....  | 25 |
| 7.7.3. Pirazinamida (PZA) .....   | 25 |
| 7.8. Infecciones mixtas y heterorresistencia.....   | 26 |
| VIII. Metodología.....  | 29 |
| Población de estudio y muestras.. .....   | 33 |
| 8.1.1. Comparación de los métodos de descontaminación: Saponina al 5% y NAOH-NALC al 0.5% .....   | 33 |
| 8.1.2. Comparación de métodos de extracción de ADN mediante los métodos Fenol y Kit (N = 33). .....   | 34 |
| 8.1.3. Comparación de 2 métodos de preparación de librerías de ADN utilizando ADN genómico.....   | 35 |
| 8.1.4. Secuenciación y procesamiento de datos de ADN genómico .....   | 36 |
| 8.1.5. Estandarización y obtención de amplicones a partir de muestras de esputo y cultivo de <i>M. tuberculosis</i> (14 genes entre 876 y 3,662 pb.) .....                                | 40 |
| 8.1.6. Comparación de 2 métodos de preparación de librerías de ADN utilizando amplicones (14 genes). .....  | 42 |
| 8.1.7. Secuenciación de amplicones (14 genes) .....   | 42 |
| 8.1.8. Obtención de amplicones de muestras descontaminadas con saponina al 5% y NALC/NaOH (4 genes). .....  | 44 |
| 8.1.8.1 Determinación de la cantidad mínima de ADN para la obtención de amplicones usando un pool de 4 amplificaciones en muestras de cultivos primarios extraídos por el método Kit..... | 45 |
| 8.1.9. Comparación de librerías para un pool de 4 genes de resistencia utilizando los kits <i>Rapid Barcoding</i> y <i>Native Ligation</i> . .....  | 45 |
| 8.1.9.1. Generación de librerías de amplicones.....   | 45 |
| 8.1.10. Secuenciación de amplicones (4 genes). .....  | 46 |
| 8.1.11. Comparación de tipos de ADN usados para la secuenciación. ....  | 47 |

|  |    |
|--|----|
| 8.2.1 Comparación del límite de detección de heterorresistencia en mezclas bacterianas mediante secuenciación y método de proporciones. ....   | 47 |
| 8.2.2. Población de estudio y muestras .....   | 48 |
| 8.2.3. Preparación de medios de cultivo y reactivación de cepas bacterianas..  | 48 |
| 8.2.4. Preparación de mezclas bacterianas heterorresistentes .....   | 48 |
| 8.2.5. Descontaminación con el método NALC/NaOH de muestras de esputo con BK negativo.....   | 49 |
| 8.2.6. Spike de mezclas heterorresistentes en muestras de esputo negativo descontaminado y PBS.....  | 49 |
| 8.2.7. Determinación de la susceptibilidad a isoniacida, pirazinamida y rifampicina en muestras de esputo y cultivo mediante el método de proporciones en agar (cultivo).....  | 50 |
| 8.3. Extracción del ADN mediante el método Kit. ....   | 52 |
| 8.3.1. Generación de amplicones (4 genes).....   | 52 |
| 8.3.2. Preparación de librería génica. ....  | 52 |
| 8.3.3. Secuenciación.....  | 53 |
| 8.4. Análisis de datos de secuenciación.....   | 53 |
| 8.4.1. Control de calidad y validación.....  | 53 |
| 8.4.2. Identificación de lecturas “lecturas” de <i>Mycobacterium tuberculosis</i> . ..   | 54 |
| 8.4.3. Determinación de la profundidad .....   | 54 |
| 8.4.4. Análisis estadísticos .....   | 55 |
| 9. RESULTADOS .....  | 55 |
| 9.1. Estandarización de métodos de descontaminación, extracción de ADN, preparación y generación de librerías génicas y tipo de ADN usando parámetros de secuenciación a partir de muestras de esputo y cultivos primarios. .... | 55 |
| 9.1.1 Ensayo comparativo de descontaminación con saponina al 5% y NAOH-NALC al 0.5%.....   | 55 |
| 9.1.2. Comparación de métodos de extracción de ADN mediante los métodos Fenol y Kit (N = 33).....  | 56 |
| 9.1.3. Obtención de 14 amplicones a partir de muestras de esputo y cultivos puros de <i>M. tuberculosis</i> . ....   | 60 |
| 9.1.4. Obtención de 4 amplicones de un pool con 4 genes de resistencia ( <i>inhA</i> , <i>katG</i> , <i>pncA</i> y <i>rpoB</i> ).....  | 65 |
| 9.1.5. Desempeño en secuenciación – comparación de kits de biblioteca utilizando ADN genómico.....   | 68 |
| 9.1.6. Desempeño en secuenciación – comparación de kits de biblioteca utilizando amplicones (14 genes).....  | 75 |

|   |     |
|---|-----|
| 9.1.7. Desempeño en secuenciación – comparación de kits de biblioteca utilizando amplicones (4 genes).....                                      | 82  |
| 9.2.1. Determinación del límite de detección de heterorresistencia en mezclas bacterianas mediante secuenciación y método de proporciones. .... | 86  |
| 9.2.2. Cultivo de cepas y preparación de mezclas heterorresistentes.....  | 86  |
| 9.2.3. Extracción de ADN de las mezclas heterorresistentes. ....  | 87  |
| 9.2.4. Obtención de amplicones de mezclas heterorresistentes con y sin esputo.<br>88  |     |
| 9.2.5. Detección de heterorresistencia mediante el método de proporciones en Agar. 90   |     |
| 9.2.6. Detección de heterorresistencia mediante Secuenciación. ....   | 92  |
| 10. DISCUSIÓN .....   | 96  |
| 11. CONCLUSIONES .....  | 108 |
| 12. LIMITACIONES .....  | 110 |
| 13. RECOMENDACIONES .....   | 112 |
| 14. REFERENCIAS BIBLIOGRÁFICAS .....  | 114 |
| 15. ANEXOS  |     |

## LISTA DE TABLAS

|  |    |
|--|----|
| Tabla 1. Características de los 14 amplicones evaluados. ....  | 40 |
| Tabla 2. Pool de primers: inhA, katG, pncA y rpoB.....   | 44 |
| Tabla 3. Rendimiento y datos generados en la secuenciación de muestras de esputo utilizando dos tipos de kits de librerías génicas (Rapid y <i>Native Ligation</i> kit) y diferentes tipos de ADN (amplicones y ADN genómico).. ....   | 71 |
| Tabla 4. Rendimiento y datos generados en la secuenciación de muestras de cultivo utilizando dos tipos de kits de librerías génicas (Rapid y <i>Native Ligation</i> kit) y diferentes tipos de ADN (amplicones y ADN genómico).. ....  | 72 |
| Tabla 5. Lecturas obtenidas de ADN genómico de muestras extraídas por phenol y secuenciada con <i>Native Ligation Kit</i> (n = 12).....  | 73 |
| Tabla 6. Lecturas obtenidas de ADN genómico de muestras extraídas por kit y secuenciada con la librería <i>Native Ligation Kit</i> descontaminadas con saponina al 5% (n = 12). ....   | 73 |
| Tabla 7. Valores de profundidad de secuenciamiento de amplicones extraídos por Kit_FP_Columna y FC_FP_AMPure (n = 12). ....  | 79 |
| Tabla 8. Comparación de métodos de extracción de ADN.....  | 80 |
| Tabla 9. Comparación de kits de librería génica.....   | 80 |
| Tabla 10. Rendimiento y data generada en la secuenciación con muestras de esputo.  | 81 |
| Tabla 11. Cuantificación de ADN en mezclas bacterianas con y sin esputo. <b>¡Error! Marcador no definido.</b>  |    |
| Tabla 12. Resumen del secuenciamiento de amplicones. ....  | 94 |
| Tabla 13. El IC95 de los porcentajes de resistencia. Los valores fueron calculados en cada cuadrante corresponde a los valores obtenidos tras las tres repeticiones de cultivos realizados con mezclas heterorresistentes en ausencia de esputo ... <b>¡Error! Marcador no definido.</b> |    |

## LISTA DE FIGURAS

|  |    |
|--|----|
| Figura 1. Fase 1: A. Comparación de métodos de descontaminación de muestras de esputo B. Comparación de los métodos de extracción de ADN .....                                   | 31 |
| Figura 2. Fase 1. Comparación de librerías génicas en muestras de ADN a partir de esputo y cultivos. El mismo esquema se realizó con la generación de amplicones (14 genes)..... | 31 |
| Figura 3. Fase 2. Secuenciamiento y método de proporciones en agar para la detección de heterorresistencia en mezclas bacterianas con y sin esputo.....                          | 32 |
| Figura 4. Distribución de las librerías de ADN genómico en las celdas de secuenciamiento. ....   | 39 |
| Figura 5. Programación utilizada para cada reacción de PCR.....  | 41 |
| Figura 6. Distribución de las librerías de amplicones en las celdas de secuenciamiento. ....   | 43 |
| Figura 7. Distribución de las librerías de amplicones en las celdas de secuenciamiento. ....   | 47 |
| Figura 8 Distribución de muestras de esputo con spike de cepas bacterianas para detección de heterorresistencia mediante secuenciamiento .....                                   | 50 |
| Figura 9 Concentración y calidad de ADN a partir de muestras de esputo mediante descontaminación con saponina y NALC/NaOH. ....  | 56 |
| Figura 10. Comparación de la cantidad, concentración y calidad del ADN obtenido mediante dos métodos de extracción.....  | 58 |
| Figura 11. Análisis de integridad y tamaño del ADN extraído en 11 muestras mediante TapeStation; FC: Método fenol, Kit: Método Kit .....   | 59 |

|  |    |
|--|----|
| Figura 12. Análisis de 14 amplicones generados a partir de una muestra de cultivo extraída mediante el método fenol (n =12). .....   | 61 |
| Figura 13. Análisis de 14 amplicones generados a partir de una muestra de cultivo extraída mediante el método Kit (n = 12). .....  | 62 |
| Figura 14. Análisis de 14 amplicones generados a partir de una muestra de esputo extraída mediante el método fenol (n = 12). .....   | 64 |
| Figura 15. Evaluación de la inserción del nuevo gen rpoB y en un pool de cuatro amplicones mediante electroforesis en gel de agarosa.....  | 65 |
| Figura 16.. Evaluación del límite de detección de ADN para amplificación de regiones específicas.....  | 66 |
| Figura 17. Obtención de los amplicones para los genes inhA (2503pb). pncA (876pb), katG (2902pb) y rpoB (856pb) obtenidos a partir de muestras tratadas con saponina o NALC/NaOH.. .....                           | 67 |
| Figura 18. Evaluación de la calidad media, tamaño de N50 y total de bases obtenidas en el secuenciamiento de 12 muestras procesadas mediante 2 métodos de extracción y dos tipos de librería génica.....           | 74 |
| Figura 19. Profundidad media de secuenciación mediante el kit <i>Native ligation</i> para distintos genes asociados a resistencia bajo diferentes condiciones de extracción y origen de la muestra (n = 12). ..... | 77 |
| Figura 20. Comparación de profundidad entre el kit <i>Native Ligation</i> y <i>Rapid barcoding</i> (n= 12). .....  | 78 |
| Figura 21. Relación entre la profundidad de secuenciación y el tamaño del amplicon (n = 12).....   | 79 |

|  |    |
|--|----|
| Figura 22. Comparación de la profundidad media de secuenciación obtenida con los kits de librería génica <i>Rapid barcoding</i> (Rapid) y Native barcoding (Ligation kit) utilizando el nuevo pool de cuatro amplicones (n = 12). .....  | 84 |
| Figura 23. Comparativa entre el kit <i>Native ligation</i> y <i>Rapid Barcoding</i> con 4 amplicones (n = 12). Los datos están representados mediante diagramas de caja, donde el color rojo indica el <i>Native ligation</i> kit y el color azul el <i>Rapid barcoding</i> . cada punto representa a una muestra y el valor de profundidad para el gen alcanzado..... | 85 |
| Figura 24. Cultivos de cepas H37Rv y DM97 en medios de cultivo 7H10 con un tiempo de incubación de 3 semanas. ....   | 86 |
| Figura 25. Electroforesis en gel de agarosa para la detección de genes en mezclas bacterianas con y sin esputo.....  | 89 |
| Figura 26. Electroforesis en gel de agarosa de la amplificación de genes en mezclas bacterianas heterorresistentes de <i>M. tuberculosis</i> con y sin esputo en la repetición 2. Los carriles 1 y 2 corresponden a las mezclas M5 y M6 de la primera repetición. ....   | 90 |
| Figura 27. Método de proporción en agar para evaluar la resistencia a rifampicina, isoniacida y pirazinamida. Se emplearon siete mezclas de las cepas H37Rv y DM97 con el fin de determinar el porcentaje de resistencia de <i>M. tuberculosis</i> .....   | 91 |
| Figura 28. Detección de heterorresistencia en mezcla de cepas bacterianas S:R con la resistencia predicha por TB-Profiler para rifampicina, isoniazida y pirazinamida. ....  | 93 |

## LISTA DE ABREVIATURAS

**HR:** Heterorresistencia

**TB:** Tuberculosis

**NGS:** Secuenciamiento de próxima generación

**BAAR:** Bacilos ácido-alcohol resistentes

**MPA:** Método de proporciones en Agar

**WGS:** Secuenciación del genoma completo

**tNGS:** Secuenciación dirigida de próxima generación.

**MGIT:** Tubo Indicador de Crecimiento de Micobacterias

**Xpert MYCOBACTERIUM TUBERCULOSIS/RIF :** Xpert para detección de *Mycobacterium tuberculosis* y resistencia a rifampicina.

**Xpert Ultra:** Xpert ultrasensible para detección de *Mycobacterium tuberculosis* y resistencia a rifampicina.

**MYCOBACTERIUM TUBERCULOSISDRplus:** Prueba molecular para la detección de resistencia a rifampicina e isoniazida en *Mycobacterium tuberculosis*.

**MYCOBACTERIUM TUBERCULOSISDRsl:** Prueba molecular para la detección de resistencia a fluoroquinolonas y fármacos inyectables de segunda línea en *Mycobacterium tuberculosis*.

**RRDR:** Región Determinante de la Resistencia a Rifampicina

**H:** Isoniazida

**R:** Rifampicina

**PZA:** Pirazinamida

**BK:** Baciloscopia

**UFC:** Unidad formadora de colonias.

**OADC:** Suplemento de crecimiento a base de ácido oleico, albúmina, dextrosa y catalasa

**N50:** Longitud mínima de secuencia que contiene al 50% del genoma ensamblado.

**IC95:** Intervalo de Confianza del 95%

## Resumen

La heterorresistencia (HR), es definida como la coexistencia de subpoblaciones bacterianas susceptibles y resistentes a un fármaco, representa un desafío clínico importante en el tratamiento de la tuberculosis (TB), ya que puede pasar desapercibida mediante métodos convencionales de cultivo. La HR puede originarse por coinfección con múltiples cepas de *Mycobacterium tuberculosis* (una susceptible y otra resistente) o por la aparición espontánea de mutantes resistentes dentro de una población susceptible. Tradicionalmente, el método de proporciones en agar considerado el *gold standard* para la detección de resistencia permite identificar cepas resistentes incluso cuando representan apenas el 1% de la población bacteriana. Sin embargo, su aplicación rutinaria es limitada debido al tiempo prolongado de incubación (2–3 semanas) y a la infraestructura especializada que requiere. Frente a estas limitaciones, la secuenciación genética, en particular mediante la plataforma portátil Nanopore, surge como una alternativa prometedora, ofreciendo resultados más rápidos y la posibilidad de realizar análisis genómicos y detección de mutaciones asociadas a resistencia.

En este contexto, el presente estudio evaluó la capacidad de la secuenciación Nanopore para detectar HR en *M. tuberculosis* en comparación con el método de proporciones en agar. Para ello, se desarrolló un protocolo que combinó dos métodos de extracción de ADN (fenol-cloroformo purificado con perlas magnéticas y extracción mediante columnas de sílice), dos tipos de bibliotecas (*Native Ligation* y *Rapid Barcoding*), y dos tipos de material genético (ADN genómico y amplicones), utilizando muestras de esputo negativas enriquecidas con mezclas de

cepas susceptibles y resistentes en distintas proporciones. Los resultados indicaron que el método de extracción basado en fenol-cloroformo y purificación con AMPure XP beads produjo una cantidad y tamaño de fragmentos de ADN significativamente mayores que el método basado en columnas. Asimismo, la biblioteca *Native Ligation* proporcionó una mayor profundidad que la *Rapid Barcoding*, y el uso de amplicones permitió detectar mutaciones asociadas a subpoblaciones resistentes en comparación al ADN genómico donde no se pudo ni cubrir toda la secuencia de *M. tuberculosis*

En cuanto a los límites de detección, la secuenciación Nanopore permitió identificar HR cuando la cepa resistente representaba al menos ~10% de la población, mientras que el método de proporciones en agar detectó cepas resistentes desde el 1%. En conclusión, aunque la secuenciación Nanopore no sustituye completamente al método de proporciones debido a su límite de detección más alto, si permite conocer la mutación o mutaciones presentes en el gen asociado a la resistencia de manera rápida y práctica para el diagnóstico de TB HR, con el potencial de optimizar el manejo clínico de los pacientes y fortalecer la vigilancia epidemiológica.

**Palabras clave:** heterorresistencia; *Mycobacterium tuberculosis*; Nanopore, método de proporciones

## **Abstract**

Heteroresistance, defined as the coexistence of drug-susceptible and drug-resistant bacterial subpopulations, represents an important clinical challenge in the treatment of tuberculosis because it can go unnoticed by conventional culture-based methods. Heteroresistance may arise from coinfection with multiple *Mycobacterium tuberculosis* strains, one susceptible and another resistant, or from the spontaneous emergence of resistant mutants within a susceptible population. Traditionally, the agar proportion method, considered the gold standard for resistance detection, can identify resistant strains even when they represent as little as one percent of the bacterial population. However, its routine application is limited by the long incubation time required two to three weeks and by the need for specialized infrastructure. Faced with these limitations, genetic sequencing, in particular using the portable Nanopore platform, emerges as a promising alternative by offering faster results and the possibility of genomic analyses and detection of resistance-associated mutations.

In this context, the present study evaluated the ability of Nanopore sequencing to detect heteroresistance in *M. tuberculosis* compared with the agar proportion method. To that end, we developed a protocol combining two DNA extraction methods phenol-chloroform extraction followed by purification with magnetic beads and silica column extraction two library types Native Ligation and Rapid Barcoding and two types of genetic material genomic DNA and amplicons using *M. tuberculosis* negative sputum samples spiked with mixtures of susceptible and

resistant strains at different proportions. Results indicated that the phenol-chloroform extraction followed by AMPure XP bead purification yielded a significantly greater quantity and fragment size of DNA than the column-based method. Likewise, the Native Ligation library produced greater sequencing depth than the Rapid Barcoding approach, and the use of amplicons allowed detection of mutations associated with resistant subpopulations compared with genomic DNA where full coverage of the *M. tuberculosis* sequence could not be achieved.

Regarding limits of detection, Nanopore sequencing detected heteroresistance when the resistant strain represented at least approximately ten percent of the population, whereas the agar proportion method detected resistant strains from one percent. In conclusion, although Nanopore sequencing does not completely replace the agar proportion method because of its higher limit of detection, it does enable rapid and practical identification of the mutation or mutations present in the resistance-associated gene for diagnosis of heteroresistant TB, with potential to optimize clinical management and strengthen epidemiological surveillance.

**Keywords:** heteroresistance; *Mycobacterium tuberculosis*; Nanopore; agar proportion method

## **I. Introducción**

La tuberculosis (TB) es una enfermedad infecciosa crónica transmitida por vía aérea, causada por el bacilo *M. tuberculosis*, reconocido como el agente infeccioso con mayor tasa de mortalidad a nivel global hasta antes de la pandemia de COVID-19 (1). De acuerdo con la Organización Mundial de la Salud (OMS), entre 2019 y 2020 se observó una disminución de casi 20% en los casos notificados de TB a nivel mundial, y de aproximadamente 25% en Perú; sin embargo, esta aparente reducción no refleja una verdadera disminución de la incidencia, sino más bien una merma en la capacidad diagnóstica debido a la prioridad asignada a la pandemia de COVID-19 (1). La falta de detección oportuna de nuevos casos impacta negativamente en el control de la enfermedad, dado que identificar adecuadamente a los pacientes con TB, así como aquellos infectados con cepas resistentes, permite optimizar las estrategias terapéuticas y reducir la mortalidad (2).

Un paso fundamental para mejorar el control de la TB es la identificación temprana del perfil de sensibilidad de la cepa infectante, a fin de administrar un tratamiento adecuado (3). La resistencia a los fármacos de primera y segunda línea surge a partir de alteraciones genéticas, incluyendo mutaciones puntuales, inserciones, deleciones o amplificaciones en genes implicados en la activación de los medicamentos (4). Aunque los mecanismos que originan cepas extremadamente resistentes no están completamente dilucidados, la sobreexpresión de genes que codifican bombas de eflujo ha sido propuesta como un factor clave que permite a las bacterias adquirir altos niveles de resistencia con mínimo compromiso metabólico (5).

Actualmente, la prueba de baciloscopia mediante tinción de Ziehl-Neelsen, que detecta bacilos ácido-alcohol resistentes (BAAR), constituye la técnica inicial de diagnóstico para TB pulmonar (6). Por su parte, la detección de cepas resistentes a fármacos como rifampicina, isoniazida o pirazinamida se realiza tradicionalmente mediante pruebas fenotípicas de susceptibilidad en cultivo, usualmente en medios sólidos. Estas pruebas, no obstante, requieren varias semanas de incubación debido a la lenta tasa de replicación de *M. tuberculosis* (7), y su fiabilidad depende estrictamente de condiciones específicas de cultivo, como el pH, concentración exacta de los antibióticos y la muestra misma.

La presencia de subpoblaciones minoritarias resistentes dentro de una población predominantemente susceptible, fenómeno conocido como heterorresistencia (HR) (8). El método de referencia para detectar resistencia fenotípica en *M. tuberculosis* es el método de proporciones en agar (MPA), capaz de identificar subpoblaciones resistentes que constituyen alrededor del 1% de la población total. Esta elevada sensibilidad convierte al MPA en una herramienta indispensable para detectar y cuantificar la HR. Sin embargo, puede complicar el diagnóstico, ya que su detección es limitada incluso con métodos de cultivo o con algunas pruebas moleculares convencionales. Esta limitación subraya la necesidad de desarrollar metodologías más rápidas y precisas para la identificación de cepas resistentes.

En este contexto, durante los últimos años se han impulsado tecnologías de secuenciación genética, como la secuenciación del genoma completo (WGS) y la secuenciación dirigida de próxima generación (tNGS), para la detección temprana de resistencia (9). Estas técnicas han logrado reducir significativamente los tiempos

de diagnóstico y han mejorado la resolución en la detección de mutaciones; sin embargo, presentan diversos desafíos técnicos, que van desde el procesamiento adecuado de las muestras (por ejemplo, la inactivación térmica que puede degradar el ADN) hasta el análisis bioinformático de los datos obtenidos (10). En particular, la degradación del ADN durante la inactivación y la escasa proporción de ADN bacteriano en muestras clínicas (frecuentemente aproximadamente 0.01% del total de ADN) representan retos importantes para el éxito de estas tecnologías (11,12).

Frente a estas dificultades, el presente proyecto propone evaluar el límite de detección de HR de la secuenciación Nanopore a partir de un método optimizado de extracción y secuenciación de ADN utilizando mezclas de cepas resistentes y susceptibles en una matriz de esputo negativo. Para ello, se emplea la tecnología de secuenciación por nanoporos de Oxford Nanopore (MinION), basada en la medición de cambios característicos en el potencial eléctrico durante el paso de moléculas de ADN a través de nanoporos embebidos en una membrana (13). Esta tecnología ha demostrado ser prometedora en términos de rapidez, portabilidad y profundidad de cobertura en genes de resistencia (13,14).

En este estudio, se busca aprovechar estas ventajas mediante la implementación de un protocolo de secuenciación optimizado, con el objetivo de superar las limitaciones de los métodos tradicionales y mejorar la detección de HR en muestras clínicas, contribuyendo así a una intervención terapéutica más oportuna y eficaz en el control de la TB.

## II. Antecedentes

Diversos estudios preliminares han explorado aspectos relacionados con la discrepancia entre perfiles genotípicos y fenotípicos de resistencia, los métodos de extracción de ADN y el análisis de genes de resistencia aplicables a la detección de heteroresistencia (HR). Votintseva et al. (2017) no encontraron diferencias en la susceptibilidad a antibióticos de 40 cultivos de *M. tuberculosis* al comparar cultivos en medio líquido (MGIT) con secuenciación WGS mediante plataformas Illumina y Nanopore (15). Sin embargo, investigaciones posteriores con un mayor número de muestras (2099 genomas), utilizando WGS-Illumina, reportaron una concordancia del 89.2% entre perfiles genotípicos y fenotípicos, evidenciando que ciertas mutaciones no previamente reportadas explican algunas discrepancias observadas (16).

Adicionalmente, Nimmo et al. (2019) demostraron que la secuenciación directa de esputo revela mayor diversidad genética, por ejemplo, alelos heterocigotos en genes de resistencia, en comparación con la secuenciación de cultivos derivados del mismo esputo. Esto sugiere que subpoblaciones resistentes minoritarias pueden perderse durante el aislamiento en cultivo (17).

Por otra parte, revisiones sistemáticas y metaanálisis han mostrado que la HR en *M. tuberculosis* es un fenómeno global, con prevalencias que oscilan entre aproximadamente 1% y 10% según el antibiótico evaluado (18), lo que resalta la necesidad de métodos diagnósticos capaces de detectar todas las subpoblaciones resistentes que podrían pasar desapercibidas mediante cultivo convencional.

Respecto a la secuenciación dirigida, Cabibbe et al. (2020) evaluaron la plataforma Nanopore en comparación con Illumina para la secuenciación de amplicones generados con el kit Deeplex Myc-TB, encontrando que Nanopore obtuvo resultados similares a Illumina con valores de profundidad de 4151X y 4177X respectivamente y una concordancia total en la detección de mutaciones, a un costo significativamente menor y con la ventaja de ser aplicable en regiones con infraestructura limitada (10).

En cuanto a la extracción de ADN, estudios comparativos entre protocolos tradicionales y kits comerciales muestran que algunos métodos logran mayores rendimientos de ADN, aunque las diferencias en la diversidad microbiana detectada son mínimas, salvo en muestras de coinfecciones con otros patógenos (3,19). No obstante, la descontaminación de muestras de esputo, mediante agentes como NALC-NaOH (20) o la saponina, sigue siendo esencial para enriquecer la proporción de ADN micobacteriano y reducir contaminantes.

Finalmente, estudios de secuenciación directa de esputo han demostrado el potencial de detectar genes de resistencia a fármacos de primera línea (21,22); sin embargo, las cantidades de ADN obtenidas siguen siendo limitadas (aproximadamente 0.16 a 11.8 ng/ $\mu$ L) (23,24), condicionando la eficiencia de las técnicas de secuenciación. El uso de pretratamientos como la saponina ha mostrado mejorar esta limitación (25), mientras que la elección de la estrategia de preparación de bibliotecas (ADN genómico o amplicones) puede influir en la detección de variantes genéticas específicas (26,27).

### **III. Planteamiento del problema.**

La heterorresistencia (HR) se concibe como la coexistencia, en un mismo huésped, de subpoblaciones bacilares sensibles y resistentes, lo que constituye un desafío de gran relevancia para el diagnóstico y el tratamiento de la tuberculosis. La presencia de HR favorece la diseminación inadvertida de cepas resistentes, incrementa el riesgo de fracaso terapéutico y se asocia con interpretaciones erróneas de la susceptibilidad a fármacos como rifampicina, isoniazida, pirazinamida o fluoroquinolonas (5).

En la actualidad, los métodos convencionales disponibles para la detección de HR se basan predominantemente en aproximaciones fenotípicas. Aun cuando resultan recomendables, exigen periodos prolongados de cultivo y su desempeño depende de la técnica específica utilizada (8). Ensayos como la prueba de Wayne o el cultivo en medio líquido MGIT aportan ciertas ventajas, aunque presentan limitaciones para la identificación de HR, ya que solo permiten evidenciar resistencia cuando las subpoblaciones resistentes superan aproximadamente el 10% de la población total (28).

Asimismo, los sistemas de cultivo suelen favorecer el sobrecrecimiento de subpoblaciones con mayor capacidad adaptativa o “fitness” (por lo general susceptibles), mientras que las cepas resistentes, debido a su menor frecuencia y a un ritmo de crecimiento más lento, permanecen infrarepresentadas (29,30) En consecuencia, las etapas de subcultivo y de dilución introducen un sesgo que puede enmascarar la presencia de HR.

Por otro lado, si bien las pruebas moleculares rápidas han optimizado la oportunidad del diagnóstico, en general se restringen a la detección de mutaciones previamente caracterizadas y no permiten discriminar entre ADN procedente de bacilos viables y no viables (31,32). Adicionalmente, su umbral de detección frente a subpoblaciones minoritarias continúa siendo relativamente elevado.

La secuenciación genómica, si bien es una técnica que permite caracterizar mutaciones vinculadas a la resistencia, presenta importantes barreras de tipo técnico cuando se aplica directamente sobre muestras clínicas. Esto se debe a que es necesario contar con ADN de alta calidad y en cantidades suficientes (aproximadamente entre 250 y 400 ng, dependiendo del protocolo utilizado) (33). Además, la abundante presencia de ADN humano en las muestras puede dificultar o interferir en su adecuada aplicación.

En conjunto, la falta de métodos sensibles y rápidos para detectar HR directamente en esputo genera una brecha importante en el diagnóstico de TB, comprometiendo la administración de tratamientos adecuados y favoreciendo la expansión de cepas resistentes (34).

#### **IV. Justificación.**

La detección precisa y oportuna de la HR es crucial para optimizar los esquemas de tratamiento y evitar solo la selección de cepas sensibles durante la terapia. En este sentido, la HR representa un obstáculo significativo para el control efectivo de la TB, dado que puede llevar a diagnósticos erróneos, fallos terapéuticos y aumento de la mortalidad.

En la actualidad, las herramientas moleculares de diagnóstico rápido como GeneXpert, presentan limitaciones importantes en cuanto a su sensibilidad. Por ejemplo, su capacidad para detectar poblaciones resistentes se sitúa en torno a aproximadamente 60%, una cifra que es considerablemente más alta que el límite del 1% que puede ser detectado por el método de proporciones en agar (35). Si bien este tipo de técnicas pueden ofrecer resultados en un menor tiempo, no aseguran una detección precisa de subpoblaciones resistentes que se encuentran en proporciones minoritarias, comprometiendo la eficacia terapéutica desde el inicio, ya que podría comprometer la elección y eficacia del tratamiento desde las etapas iniciales.

Por ello, resulta imperativo desarrollar metodologías alternativas más sensibles y accesibles que permitan la detección de HR directamente en muestras clínicas. La secuenciación mediante tecnología Nanopore ofrece una oportunidad única en este contexto: su capacidad de generar lecturas largas, su portabilidad y su facilidad para detectar cambios genéticos en poblaciones bacterianas mixtas la convierten en una herramienta prometedora.

La implementación de esta tecnología podría no solo mejorar el diagnóstico individualizado y la adecuación del tratamiento, sino también fortalecer la vigilancia epidemiológica y reducir los costos asociados a tratamientos prolongados e ineficaces. A pesar de los retos técnicos, como la baja carga bacilar o la contaminación con ADN humano, el desarrollo de protocolos optimizados permitiría superar estas barreras y contribuir de manera significativa a un mejor control de la tuberculosis en regiones endémicas.

## V. OBJETIVOS

### Objetivo general

Determinar el límite de detección de heterorresistencia de *Mycobacterium tuberculosis* mediante secuenciación Nanopore (usando ADN genómico y amplicones) a partir de mezclas bacterianas heterorresistentes en esputo negativo y comparar los resultados con los obtenidos por el método de proporciones en agar.

### Objetivos específicos

1. Comparar la concentración y calidad del ADN obtenida mediante dos métodos de extracción (fenol-cloroformo con purificación por beads magnéticas versus kit comercial de columnas) a partir de muestras de esputo baciloscopia positiva.
2. Evaluar la profundidad de cobertura en la secuenciación Nanopore utilizando dos tipos de bibliotecas (*Native Ligation* versus *Rapid Barcoding*) y dos tipos de material genético (ADN genómico versus amplicones) en muestras de *M. tuberculosis* provenientes de esputo y cultivo.
3. Determinar el límite de detección de HR en mezclas de cepas resistentes y susceptibles de *M. tuberculosis* inoculadas en una matriz de esputo negativo, comparando la sensibilidad de la secuenciación Nanopore frente al método de proporciones en agar.

## **VI. Pregunta de investigación.**

¿La secuenciación de *Mycobacterium tuberculosis* mediante la plataforma Nanopore, utilizando un protocolo optimizado de extracción de ADN y biblioteca génica, permitirá detectar heterorresistencia con una sensibilidad cercana a la obtenida por el método de proporciones en agar ( $\approx 1\%$  de la población bacteriana)?

## VII. Marco Teórico.

### 7.1. *Mycobacterium tuberculosis*

*Mycobacterium tuberculosis* es un bacilo aerobio inmóvil, cuyas dimensiones oscilan aproximadamente de 0.5 µm de ancho por 3 µm de largo. Es un tipo de bacteria que puede encontrarse de manera aislada o formando agregados, ya sea en el tejido pulmonar o en el interior de macrófagos. Su pared celular presenta una composición que es inusualmente rica en lípidos, incluyendo peptidoglucano, arabinogalactano y una alta proporción de ácidos micólicos), los cuales pueden llegar a constituir entre el 50–60% de la pared celular de la bacteria (36). Debido a esta particular composición lipídica en la estructura bacteriana, es posible realizar la detección microscópica de *M. tuberculosis* mediante la tinción de Ziehl-Neelsen, que es una de las técnicas de microscopía más utilizadas, en la cual los bacilos ácido-alcohol resistentes se observan con tonalidades rojo brillante a púrpura intenso (37).

Una de las características más peculiares de *M. tuberculosis* es su tiempo de duplicación, la cual es notablemente prologando. A diferencia de la mayoría de las bacterias, que se dividen aproximadamente cada 20 minutos, *M. tuberculosis* requiere entre 18 y 20 horas para poder completar este proceso (38). Asimismo, las cepas de *M. tuberculosis* que presentan resistencia a medicamentos a menudo muestran un ritmo de crecimiento aún más lento en comparación con las cepas susceptibles (39). *M. tuberculosis* forma parte del Complejo *M. tuberculosis* (*MTBc*) junto con otras micobacterias patógenas (*M. bovis*, *M. africanum*, etc.), compartiendo entre ellas más del 99.9% de identidad a nivel

genómico (40). Todas las especies del MTBc son capaces de causar tuberculosis en el ser humano y en otros animales, lo que evidencia su cercanía evolutiva y relevancia sanitaria.

## **7.2. Formas clínicas y epidemiología de tuberculosis**

La tuberculosis (TB) suele afectar primariamente a los pulmones (tuberculosis pulmonar), aunque puede manifestarse en casi cualquier órgano (tuberculosis extrapulmonar). Las formas pulmonares son las más comunes y contagiosas, mientras que las extrapulmonares (por ejemplo, TB pleural, ganglionar, miliar, etc.) representan alrededor del 15% de los casos, siendo más frecuentes en personas con inmunosupresión (41,42). Una forma particular es la tuberculosis miliar, diseminada por vía hematógena, que ocurre cuando hay una diseminación masiva de bacilos a múltiples órganos (43).

A nivel global, la TB continúa siendo una de las principales causas de muerte por enfermedad infecciosa. En 2021 se estimaron aproximadamente 10.6 millones de casos nuevos de TB y 1.6 millones de muertes relacionadas (44). Las regiones con mayor carga de TB son África subsahariana, el sudeste asiático y el Pacífico Occidental (45). Diversos factores contribuyen a la alta incidencia en estas zonas, incluyendo la pobreza, hacinamiento, coinfección con VIH y sistemas de salud deficientes (46). En Perú, la incidencia estimada de TB en 2020 fue de aproximadamente 116 casos por 100,000 habitantes, concentrándose los mayores números en áreas urbanas como Lima y en regiones amazónicas (44). La transmisión de *M. tuberculosis* ocurre principalmente por inhalación de aerosoles infecciosos exhalados por un paciente con TB pulmonar activa; un paciente sin

tratar puede infectar a 10–15 personas por año en promedio. La probabilidad de contagio depende de la carga bacteriana del paciente, las condiciones de ventilación del ambiente y la duración del contacto (47–49).

Luego de la inhalación del bacilo, la infección tuberculosa puede permanecer en estado latente en la mayoría de los individuos con sistemas inmunológicos competentes, ya que entre el 90 y el 95% de los individuos no llegan a desarrollar la enfermedad de forma activa. Sin embargo, se estima que alrededor del 5 al 10% de las personas infectadas desarrollarán TB activa en algún momento de su vida. Este riesgo que se incrementa significativamente en presencia de condiciones de inmunosupresión. como infección por VIH, desnutrición o presencia de diabetes, entre otras.). El control efectivo de la TB a nivel poblacional depende de múltiples estrategias complementarias. Entre ellas, se incluyen el diagnóstico y el tratamiento tempranos de los casos activos, la vacunación (BCG en neonatos para formas graves en niños), así como el abordaje de factores y/o determinantes socioeconómicos que facilitan la propagación del bacilo en la comunidad (50).

### **7.3. Infección por *M. tuberculosis***

La infección por *M. tuberculosis* se inicia cuando una persona sana susceptible inhala núcleos de gotitas que contienen bacilos viables expulsados por un paciente con TB pulmonar bacilífera. Los bacilos inhalados alcanzan los alveolos pulmonares, donde son fagocitados por macrófagos. *M. tuberculosis* ha desarrollado mecanismos para sobrevivir dentro de los macrófagos alveolares, resistiendo la acidificación y la producción de especies reactivas de oxígeno dentro del fagosoma. Esto le permite persistir y replicarse intracelularmente,

desencadenando una respuesta inmunitaria celular. En alrededor de 2 a 8 semanas post-infección, la activación de linfocitos T específicos lleva a la formación de granulomas, es decir estructuras organizadas de macrófagos infectados, células epitelioides, células gigantes de Langhans y linfocitos, que intentan contener la infección (50).

La mayoría de las infecciones primarias se controlan en esta fase, quedando los bacilos latentes dentro de los granulomas. No obstante, en algunos individuos, especialmente si tienen compromiso inmune, la infección inicial progresa a TB primaria activa o la bacteria latente se reactiva posteriormente causando TB postprimaria (51).

La coinfección con VIH es uno de los principales factores de riesgo para progresión de infección latente a TB activa, dada la depleción de linfocitos CD4 que desestabiliza los granulomas (52). Otros factores incluyen el tratamiento inapropiado o incompleto de un episodio previo de TB (que puede dejar bacilos persistentes), y condiciones como silicosis, insuficiencia renal crónica, entre otras (53).

#### **7.4. Detección de *M. tuberculosis***

La detección de TB combina métodos bacteriológicos tradicionales y técnicas moleculares detalladas a continuación.

##### **7.4.1. Microscopía de esputo**

La baciloscopia mediante tinción de Ziehl-Neelsen o auramina-rodamina es el método inicial más utilizado. Aunque rápida y económica, su sensibilidad es limitada (aproximadamente 50–60%) y requiere una alta carga bacilar (>5000 bacilos/mL) para obtener resultados positivos (54,55).

reconocer y determinar el orden nucleotídico del ADN que se esté secuenciando (56).

#### **7.4.1.1 Protocolos de descontaminación**

El proceso de descontaminación de la muestra de esputo es un proceso importante considerando que estas muestras; además del agente patógeno que se encuentre en estudio, contienen una flora oral que puede obstruir la obtención de ADN de *M. tuberculosis* por lo que es necesario realizar una etapa de descontaminación para evitar el crecimiento de bacterias y/o asegurar la viabilidad de la muestra (57).

##### **7.4.1.1.1 Uso de NALC/NaOH al 0.5%**

El método del hidróxido de sodio se usó inicialmente en 1968 y su importancia radicó en su capacidad para generar un medio alcalino que desnaturaliza las proteínas de la membrana de bacterias no micobacterias y sumado al elevado grado de iones de Na<sup>+</sup> se altera la gradiente de iones dentro de las bacterias que termina “explotando” a las bacterias (58). Es un método rápido y barato aunque se debe tener cuidado con el tiempo de exposición, ya que un tiempo prolongado de su uso puede generar incluso la lisis de micobacterias dentro de la muestra de esputo.

##### **7.4.1.1.2 Uso de Saponina al 5%**

La saponina fue reconocida en 1996 en plantas debido a su papel en la respuesta inmunitaria de plantas, (59), presenta capacidad citotóxica debido a que pueden degradar la membrana celular de bacterias por su capacidad saponificante (solubiliza los lípidos) (60)disminuyendo de esa manera material genético de células eucariotas sin alterar la diversidad biológica dependiendo del porcentaje de saponina que se utilice.

#### **7.4.2. Cultivo microbiológico**

Considerado el estándar de oro, el cultivo permite confirmar la presencia de *Mycobacterium tuberculosis* y evaluar su susceptibilidad a fármacos. Puede realizarse en medios sólidos como Löwenstein-Jensen o en sistemas líquidos automatizados como MGIT, los cuales acortan el tiempo de detección pero requieren infraestructura especializada (61).

#### **7.4.3. Pruebas fenotípicas de susceptibilidad a fármacos**

El método de proporciones en agar consiste en sembrar la muestra en medios con y sin antibióticos y comparar la proporción de crecimiento. Se considera resistente si  $\geq 1\%$  de la población crece en presencia del fármaco (62). Sistemas como BACTEC MGIT permiten realizar pruebas en medio líquido, reduciendo los tiempos de respuesta, pero con limitaciones en la detección de HR (28,55).

##### **7.4.3.1. Método de proporciones en Agar.**

El método de proporciones en agar (APM), se realiza sobre placas Middlebrook 7H10/7H11-OADC, es la prueba fenotípica de referencia para la susceptibilidad de

*M. tuberculosis* según la OMS (63). El principio de esta técnica radica en sembrar diluciones 1:10 y 1:100 de un inóculo estandarizado (0,5–1 McFarland) en medios que contienen concentraciones críticas de cada fármaco y en un control sin antibiótico; si al menos 1 % de las colonias crece en la placa con droga, la cepa se clasifica como resistente (64). Esta cuantificación directa permite evaluar simultáneamente fármacos de primera y segunda línea y mantiene costes bajos y alta reproducibilidad, aunque requiere hasta 21 días. Sus ventajas incluyen cobertura homogénea del genoma bacteriano sin sesgo molecular y la posibilidad de conservar subpoblaciones minoritarias para estudios genéticos posteriores.

#### **7.4.4. Pruebas moleculares rápidas**

En las últimas dos décadas se han introducido pruebas como Xpert MTB/RIF y Xpert Ultra del fabricante Cepheid que detectan directamente en la muestra genes específicos de *M. tuberculosis* y mutaciones de resistencia (p. ej., en *rpoB* para rifampicina) mediante PCR anidada en cartucho. Estas pruebas pueden dar un resultado en aproximadamente 2 horas (32,65). La sensibilidad de esta prueba para detectar TB pulmonar es superior a la baciloscopia (>60%). especialmente la versión Ultra (<10%) (65), y son capaces de detectar cepas resistentes al antibiótico rifampicina de manera concordante con métodos fenotípicos y atribuyendo como causante a HR cuando los resultados discrepan entre el genotipo y fenotipo (66,67).

No obstante, la capacidad de la prueba para detectar HR es limitada, ya que en muestras con mezcla de bacilos resistentes y susceptibles, la detección dependerá de qué tan abundante sea la subpoblación resistente (se estima que requieren >60–70% de bacilos resistentes para dar señal de resistencia) y solo cubren ciertos

fármacos. Otras técnicas incluyen ensayos de línea (*line probe assays*) como MYCOBACTERIUM TUBERCULOSIS DRplus y MYCOBACTERIUM TUBERCULOSIS DRsl, que identifican mutaciones en genes asociados a resistencia a rifampicina, isoniazida y fármacos de segunda línea mediante hibridación inversa en tiras, con resultados en 1–2 días (68).

La secuenciación masiva de nueva generación (NGS) ha empezado a incorporarse en algunos laboratorios de referencia para la detección simultánea de todas las mutaciones de resistencia conocidas. La secuenciación del genoma completo (WGS) de un aislamiento de cultivo puede predecir su fenotipo de resistencia con elevada concordancia para la mayoría de los fármacos de primera línea (16). Sin embargo, la aplicación directa de WGS sobre muestras clínicas es más desafiante por la escasa cantidad de ADN de *M. tuberculosis* y la presencia predominante de ADN humano y de otros microorganismos. Para lidiar con este problema se están desarrollando opciones como el uso de sondas que detecten y se acoplen al ADN micobacteriano así como la secuenciación metagenómica (9). La secuenciación mediante la plataforma Nanopore representa una de estas nuevas opciones, ya que a diferencia de Illumina (que tiene un gran desenvolvimiento con lecturas cortas), Nanopore genera lecturas largas en tiempo real, lo que facilita la detección de variantes estructurales (tales como mutaciones, inserciones y deleciones) (69,70) y potencialmente la detección de cepas en infecciones mixtas. La capacidad de Nanopore para detectar genes de resistencia de *M. tuberculosis* en muestras de esputo ya ha sido reportada y aunque su sensibilidad aún no supera a la plataforma Illumina enfoques como la reducción del ADN humano en la muestra podrían mejorar su precisión (15,17).

#### **7.4.5. Secuenciamiento por nanoporos**

El secuenciamiento nanopore, como su nombre lo menciona, basa su principio en el uso de nanoporos ubicados dentro de una membrana que divide dos áreas por las que hay un constante flujo eléctrico. Cada vez que un nucleótido atraviesa un nanoporo se genera un cambio de potencial característico de cada base nucleotídica, permitiendo así

#### **7.5.1. Protocolos de extracción de ADN**

##### **7.5.1.1. Método fenol**

El método de extracción usando fenol:cloroformo:alcohol isoamílico es bastante utilizado porque permite obtener altas concentraciones de ADN, su principio radica en la capacidad de separar el ADN del debris celular a partir de una lisis de tipo mecánica propia de esta técnica (generada por el equipo Fast prep-24 5G), debido a que el material genético no es soluble en estos compuestos orgánicos (71) permitiendo separarla en la fase acuosa. Adicionalmente, en este proceso, se hizo uso de AMPure beads para atrapar magnéticamente al ADN (72) y asegurar una muestra con calidad adecuada para un proceso de secuenciamiento.

##### **7.5.1.2. Método con uso de Kits comerciales**

El kit comercial utilizado en este estudio fue el Zymo (Quick-DNA™ Miniprep Plus Kit); al cual se le realizó cambios dentro del proceso de lisis, en donde se cambió el uso de proteinasa K por Fast prep -24 5G, Este tipo de extracción se

caracteriza por el uso de columnas de sílice que junto a sales caotrópicas (73) permiten la obtención de ADN purificado listo para ser usado en secuenciamiento.

## **7.5.2. Tipos de secuenciamiento genómico**

### **7.5.2.1 Secuenciamiento de genoma completo**

La secuenciación del genoma completo constituye una alternativa útil para identificar de forma rápida tanto enfermedades poco frecuentes como variantes genéticas codificantes y no codificantes. Este enfoque permite analizar el ADN en su totalidad, incluyendo regiones que usualmente no son exploradas con otros métodos. Además de detectar anomalías ya conocidas, ofrece la posibilidad de asociar nuevas mutaciones con diversas patologías (74).

### **7.5.2.2. Secuenciamiento de amplicones**

A diferencia del enfoque de genoma completo, la secuenciación por amplicones se enfoca en regiones específicas del ADN donde se han reportado variantes asociadas a malformaciones genéticas o, en el caso de bacterias, al desarrollo de resistencia. Al centrarse únicamente en estas zonas de interés, se puede obtener un nivel de resolución más preciso sobre los genes involucrados. Este enfoque resulta especialmente útil cuando se trabaja con muestras metagenómicas, en las que la obtención del genoma completo del patógeno no es viable (75).

## **7.5.3 Tipos de librería génica**

### **7.5.3.1 *Native Ligation* (SQK-LSK109/114)**

El kit *Native Ligation* de Oxford Nanopore consta de un paso de reparación de extremos y dA-tailing seguido de la ligación enzimática de adaptadores que portan la helicasa-motor (la cual es reconocida por el nanoporo), de modo que mantiene la longitud original de fragmentos de ADN de alta calidad. También, ofrece las lecturas más largas (frecuentemente >30 kb) y el mayor rendimiento por canal, lo que se traduce en ensamblajes más completos y uniformes, incluso en regiones con extremos citosina-guanina (CG) (76).

#### **7.5.3.2. Rapid Barcoding (SQK-RBK114.24)**

El kit *Rapid Barcoding* emplea una transposasa MuA que tagmenta (fragmenta) el ADN y añade etiquetas adaptadoras en un solo paso, así se consigue una preparación lista para cargar en  $\approx 15$  min a partir de tan solo 1–5 ng de ADN, ideal para muestras limitadas o laboratorios de campo. La rapidez y el bajo consumo se compensan con fragmentos más cortos (modal 2–3 kb) y un patrón de inserción sesgado (motivo 5'-TATGA-3') que causa cobertura irregular, especialmente en genomas con 40–70 % GC, y genera un ligero incremento de lecturas artefactuales frente al protocolo de ligación

### **7.5. Muestras clínicas y cultivos puros.**

#### **7.5.1. Muestra clínica**

La secuenciación y/o amplificación dirigida directamente sobre la muestra clínica mantiene la población total presente en paciente, por lo que permite detectar subpoblaciones resistentes que pueden representar hasta el 1 % del total y que con

frecuencia pasan inadvertidas con métodos de detección tradicional. Estudios de secuenciación de todo el genoma y de amplicones multiplex desde esputo directo permiten una detección de mutaciones de resistencia que suelen pasar por desapercibido cuando se hace uso de cultivos, reduce el tiempo de detección a solo 24–48 horas en lugar de varias semanas (77). El principal reto es técnico (baja carga bacilar y alto contenido de ADN humano), pero los protocolos de enriquecimiento y los algoritmos de filtrado han mejorado la sensibilidad sin perder especificidad.

### **7.5.2. Cultivos Puros**

El cultivo sigue siendo indispensable cuando se requiere un inóculo abundante y libre de contaminantes para pruebas fenotípicas. El uso de cultivos permite obtener una alta concentración de ADN que facilita la lectura inequívoca de mutaciones y reduce falsos negativos ligados a baja cobertura (78). Sin embargo, ya hay reportes donde se ha demostrado que el subcultivo puede generar un efecto cuello de botella: minorías resistentes detectadas en la muestra original pueden perderse, generando un falso perfil de sensibilidad y subestimando fenómenos de heterorresistencia (p. ej., rifampicina o fluoroquinolonas) (79).

## **7.6. Tratamiento antituberculosis y resistencia antibiótica**

El tratamiento estándar para la TB sensible, de acuerdo con las guías nacionales del MINSA en Perú, se organiza en dos fases diferenciadas. La primera es una fase intensiva de 2 meses y utiliza una combinación de cuatro fármacos como son la (isoniazida, la rifampicina, la pirazinamida y el etambutol. Posteriormente, se

continúa con una fase de consolidación o de continuación de 4 meses, en la cual se administran únicamente dos fármacos como son la isoniazida y la rifampicina (54). Este esquema terapéutico, es comúnmente conocido como “esquema terapéutico 2HRZE/4HR”, que ha demostrado tener altas tasas de curación superiores al 85–90% en los casos de TB pulmonar. No obstante, la duración extendida del tratamiento, junto con los efectos secundarios frecuentes, han contribuido a que muchos pacientes abandonen o interrumpan la terapia antes de tiempo, lo cual facilita la aparición y selección de cepas resistentes de esta bacteria (80).

## **7.7. Principales fármacos de primera línea y la base genética de su resistencia**

### **7.7.1. Isoniazida (H)**

Es un Antibiótico bactericida introducido en 1952, actúa sobre bacilos en replicación activa. Es un profármaco que requiere ser activado por la enzima catalasa-peroxidasa (producto del gen *katG*) para luego inhibir la síntesis de ácidos micólicos de la pared de *M. tuberculosis*, dejando al bacilo susceptible al estrés oxidativo (68). Las mutaciones en *katG* son la causa más frecuente de resistencia a isoniazida, destacando la mutación puntual S315T (Ser por Tre) que se observa en 50–70% de cepas clínicamente resistentes (81). Otra vía de resistencia involucra mutaciones en el promotor o la región estructural del gen *inhA*, que codifica una enoil-ACP reductasa en la síntesis de ácidos grasos. Las mutaciones en *inhA* (frecuentemente en su región promotora, como -15C>T) confieren resistencia de bajo nivel a isoniazida y suelen asociarse también a resistencia cruzada con etionamida (82). En conjunto, se estima que aproximadamente 70–80% de cepas

resistentes a isoniazida presentan alteraciones en *katG* y/o *inhA*, quedando un porcentaje cuyo mecanismo de resistencia no está completamente esclarecido (83).

### **7.7.2. Rifampicina (R)**

La rifampicina es clave en el tratamiento acortado de la TB, que fue introducido en 1970. Inhibe la ARN polimerasa dependiente de ADN uniéndose a la subunidad  $\beta$  de la enzima (codificada por *rpoB*), bloqueando la elongación de la transcripción (84). Más del 95% de las cepas de *M. tuberculosis* resistentes a rifampicina presentan mutaciones en una región crítica de 81 pares de bases de *rpoB* conocida como la *Región Determinante de la Resistencia a Rifampicina* (RRDR), que abarca aproximadamente los codones 507–533 (85). Las mutaciones puntuales en esta región (como S531L, H526Y/D, D516V) reducen la afinidad de la rifampicina por la ARN polimerasa, confiriendo resistencia. Debido a la alta asociación entre resistencia a R e isoniazida, la rifampicina es considerada un marcador de TB multidrogoresistente (MDR-TB) y su detección es prioritaria en el diagnóstico rápido. Cabe señalar que mutaciones fuera de la RRDR son raras pero pueden también conferir resistencia (p. ej., mutaciones en codones 534–572), lo que explica algunos casos de resistencia no detectados por pruebas genéticas enfocadas solo en la RRDR (81).

### **7.7.3. Pirazinamida (PZA)**

Es un fármaco introducido en 1952 que actúa principalmente sobre bacilos en estado latente o semilatenente dentro de los granulomas (86). La pirazinamida (PZA) es un profármaco que requiere la activación por la enzima

pirazinamidasa/nicotinamidasa, codificada por el gen *pncA*, para convertirse en ácido pirazinoico, su forma activa. En un medio ácido extracelular (como el de los macrófagos), el ácido pirazinoico protonado reingresa al bacilo y altera funciones regulatorias, conduciendo a la muerte del bacilo (87). Se estima que hasta el 80% de cepas resistentes a PZA presentan mutaciones en *pncA*, distribuidas a lo largo de todo el gen, lo que inactiva la enzima y evita la conversión a la forma activa (87). No existe una “región caliente” única de mutaciones, por lo que la detección molecular de resistencia a PZA es compleja. La pirazinamida, en combinación con rifampicina e isoniazida, permite acortar el tratamiento a 6 meses; por ello la resistencia a PZA (aislada o en cepas MDR) compromete la eficacia de los regímenes estándares.

### **7.8. Infecciones mixtas y heterorresistencia**

Se considera infección mixta o coinfección por TB cuando un mismo paciente alberga simultáneamente dos o más cepas distintas de *M. tuberculosis*. Esto puede ocurrir, por ejemplo, en entornos de alta transmisión donde es posible que una persona se infecte en momentos distintos con cepas diferentes, o en brotes en poblaciones cerradas (prisiones, albergues) donde circulan múltiples cepas. Si dentro de una infección mixta coexiste una cepa susceptible y al menos otra cepa resistente a algún fármaco, el cuadro global del paciente presentará HR para ese fármaco. En tal caso se habla de HR policlonal, ya que la resistencia proviene de cepas diferentes (17). Por otro lado, la HR monoclonal ocurre cuando una sola cepa de *M. tuberculosis*, inicialmente susceptible, adquiere una mutación de resistencia en una subpoblación de bacilos; así, genéticamente son la misma cepa, pero con

variantes susceptibles y resistentes (5). Tanto la HR policlonal como la monoclonal pueden resultar en la presencia de dos subpoblaciones en una muestra clínica, una mayoritaria sensible y una minoritaria resistente, lo que complica su detección con métodos tradicionales (88).

Las infecciones mixtas en TB no son fenómenos inusuales. Se ha observado que estudios moleculares han demostrado que, en zonas con alta incidencia de la enfermedad, entre un 9 % y un 19% de los pacientes podrían albergar simultáneamente más de una cepa de *M. tuberculosis*. En ciertos contextos particulares, las tasas reportadas son incluso más elevadas. Por ejemplo, investigaciones realizadas en países como son Sudáfrica y China, que utilizan técnicas genotípicas de alta resolución, pudieron identificar coinfecciones en aproximadamente el 20% al 30% de los casos analizados. No obstante, las estimaciones podrían variar ampliamente en función de la población estudiada y la sensibilidad de las técnicas utilizadas (desde aproximadamente 1–3% detectado con métodos convencionales hasta mayor a 50% en estudios combinando muestras pulmonares y extrapulmonares). Las infecciones mixtas suelen pasar inadvertidas en la práctica clínica habitual, ya que el aislamiento por cultivo tiende a obtener una sola colonia (o clon bacteriano), el cual puede no representar toda la diversidad de cepas presentes en el organismo del paciente. Esta limitación diagnóstica acarrea consecuencias importantes. Por ejemplo, si un paciente está coinfectado con una cepa sensible y otra resistente, podría existir el riesgo de ser diagnosticado falsamente como TB sensible, es decir que solo se detecte la cepa sensible mediante el cultivo con la consecuente administración de un tratamiento inadecuado y a la subsecuente selección y propagación de la cepa resistente (89,90).

La HR resultante de infecciones mixtas puede entonces interpretarse como HR policlonal. Clínicamente, la HR puede comprometer el éxito del tratamiento, ya que las subpoblaciones resistentes, aunque minoritarias al inicio, pueden sobrevivir a la terapia e incluso expandirse, causando recaída o fracaso terapéutico. Se ha documentado que la presencia no detectada de HR puede producir falsos negativos en las pruebas de susceptibilidad (por ejemplo, un cultivo puede salir “sensible” si la proporción de bacilos resistentes estaba por debajo del límite de detección). Además, si no se ajusta el esquema terapéutico, la subpoblación resistente continuará transmitiéndose y propagando resistencia en la comunidad. Por estas razones, detectar y cuantificar adecuadamente las infecciones mixtas y la HR asociada es crucial. Técnicas moleculares como el *spoligotyping*, MIRU-VNTR y más recientemente la secuenciación profunda, han permitido identificar patrones que evidencian más de una cepa infectando al paciente (por ejemplo, presencia de dobles bandas alélicas en varios loci MIRU-VNTR sugiere coinfección). La secuenciación de nueva generación, al proporcionar lecturas individuales de ADN, incluso permite cuantificar la proporción relativa de cada genotipo en una muestra, aportando información valiosa para distinguir HR monoclonal de policlonal.

En resumen, las infecciones mixtas representan un factor contribuyente a la HR y constituyen un desafío diagnóstico que requieren métodos avanzados para su detección. Reconocer su frecuencia, especialmente en regiones endémicas con alta transmisión y entre pacientes con antecedentes de TB, puede ayudar a interpretar mejor resultados de laboratorio inusuales (por ejemplo, cultivos con subpoblaciones) y a diseñar estrategias de tratamiento más efectivas.

## VIII. Metodología

El presente estudio experimental, de enfoque cualitativo-descriptivo, evaluó la capacidad de detección de heterorresistencia (HR) mediante secuenciación Nanopore en comparación con el método de proporciones en agar.

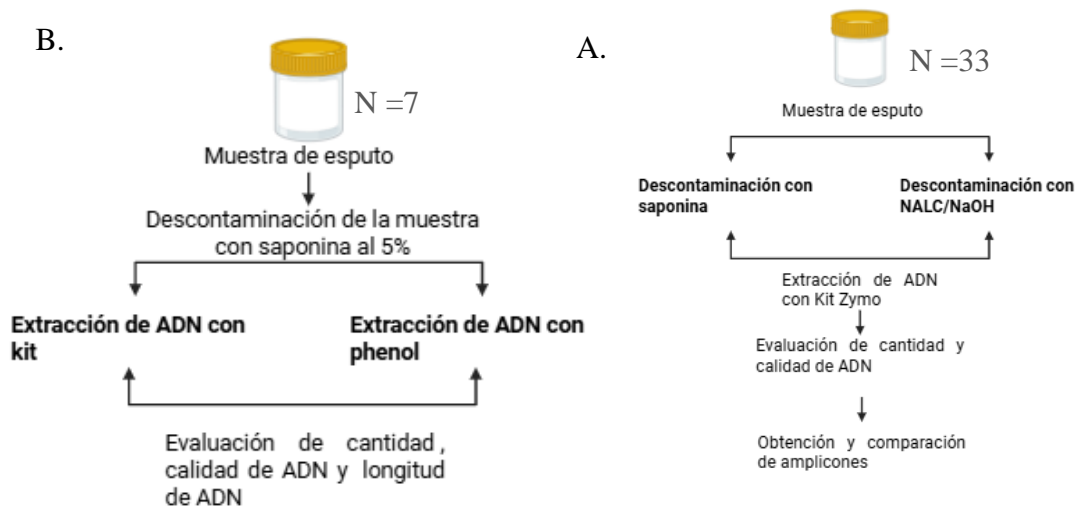
Este estudio se llevó a cabo en 2 etapas: una estandarización inicial de los métodos de descontaminación, extracción de ADN, generación de librerías génicas y elección de tipo de ADN a evaluar en la secuenciación Nanopore y una segunda etapa en donde se realiza la comparación del límite de detección de HR entre la secuenciación Nanopore y el método de proporciones en agar.

**Estandarización.** En la primera fase se comparó dos métodos de descontaminación (Figura 1); el método saponina, basado en el uso de un detergente suave capaz de lisar células humanas y NALC/NaOH, agente mucolítico y solución alcalina fuerte, capaz de lisar células bacterianas. Asimismo, se comparó dos procesos de extracción de ADN (Figura 1); el método fenol, basado en la purificación del ADN con solventes orgánicos y un kit comercial (kit), basado en la purificación del ADN con columnas de sílice. También se comparó el tipo de ADN secuenciado; ADN genómico y amplicones. Finalmente, se comparó dos formas de preparar librerías génicas para el secuenciamiento con Nanopore; método *Native Ligation*, basado en el uso de ligasas y enzimas de reparación y preparación del ADN y *Rapid Barcoding*, basado en el uso de transposasas que marcan y también fragmentan el ADN (Figura 2).

Los tipos de descontaminación de muestra y extracción de ADN fueron evaluados de acuerdo con la pureza, cantidad e integridad del ADN. El tipo de ADN, tipo de

preparación de librerías génicas fueron evaluados usando las métricas del secuenciamiento como profundidad, N50, q-score, etc.

**Comparación de límite de detección de HR.** Para simular la HR, se prepararon mezclas controladas de cepas de *M. tuberculosis* bacilos susceptibles y resistentes en diferentes proporciones (Figura 3). Adicionalmente, se utilizaron muestras de esputo humano con baciloscopia (BK) negativa obtenidas de pacientes con resultados de TB negativo por baciloscopia y GenXpert. Sobre alícuotas de esta mezcla de esputo BK negativo se realizaron los “spikes” o inoculaciones de las mezclas de cepas de *M. tuberculosis* seleccionadas, obteniendo así muestras de esputo BK-negativo (sin bacilos acidorresistentes originales) pero con proporciones conocidas de las cepas de *M. tuberculosis* de interés. Estas muestras de esputo negativo con mezcla de cepas de *M. tuberculosis* pasaron por el proceso de descontaminación, inactivación, extracción, selección de ADN genómico o amplicones y preparación de librería génica fijado en la primera fase para seguidamente evaluar tanto al secuenciamiento como al método de proporciones en



agar.

Figura 1. Estandarización: A. Comparación de métodos de descontaminación de muestras de esputo B. Comparación de los métodos de extracción de ADN.

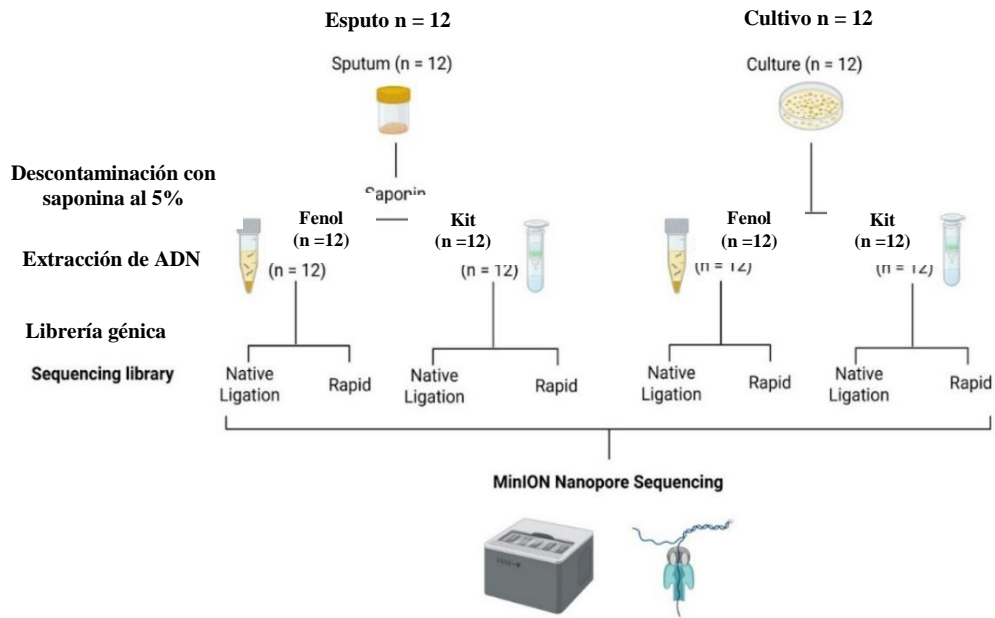
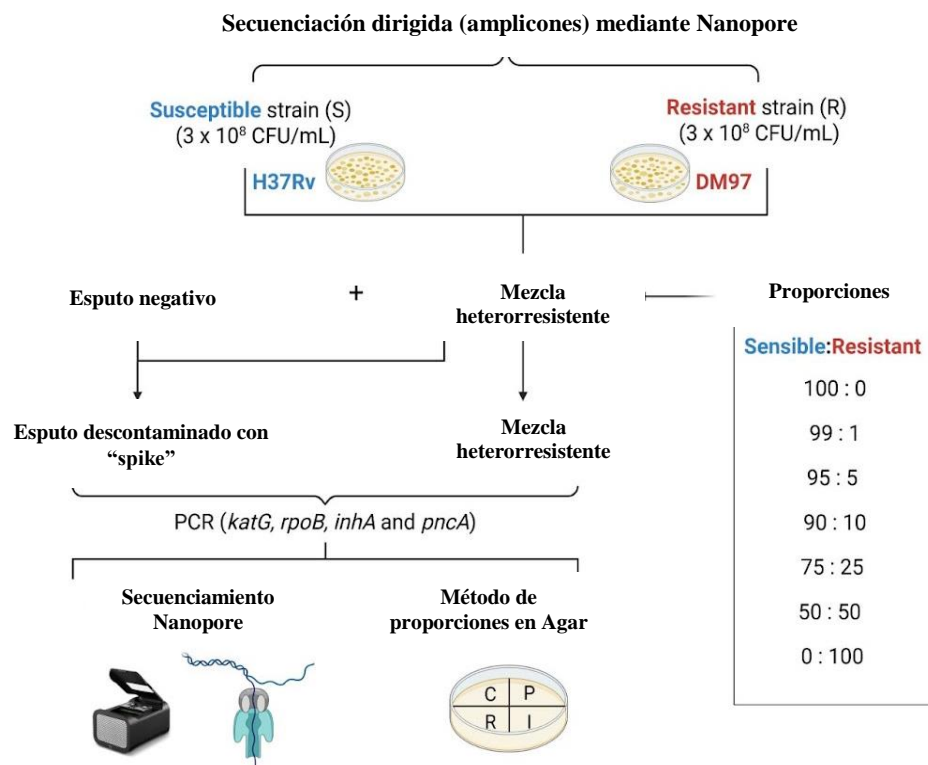


Figura 2. Estandarización. Comparación de librerías génicas en muestras de ADN a partir de esputo y cultivos. El mismo esquema se realizó con la generación de amplicones (14 genes).



Nota: La cepa DM97 es resistente a isoniacida, rifampicina y pirazinamida

Figura 3. Comparación de límite de detección de HR. Secuenciación y método de proporciones en agar para la detección de heterorresistencia en mezclas bacterianas con y sin esputo.

**8.1. Estandarización de métodos de descontaminación, extracción de ADN, preparación y generación de librerías génicas y tipo de ADN usando parámetros de secuenciación a partir de muestras de esputo y cultivos primarios**

**Población de estudio y muestras.** El estudio recolectó muestras de esputo de pacientes anónimos con baciloscopia positiva: 25 con baciloscopia positiva (2+) y 8 (3+), provenientes del Hospital Regional del Callao.

**8.1.1. Comparación de los métodos de descontaminación: Saponina al 5% y NAOH-NALC al 0.5%**

**Descontaminación y procesamiento de muestras de esputo.** Esta etapa del estudio permitió evaluar dos métodos de descontaminación: saponina y NALC/NaOH. El método con saponina ofrece la ventaja de eliminar el ADN del hospedero sin alterar significativamente la diversidad génica del microbiota humano, además de facilitar la obtención de altas cantidades de ADN total (Anexo 1). Por otro lado, el uso de NALC/NaOH, gracias a su alta capacidad mucolítica y descontaminante, permite recuperar un mayor porcentaje de ADN específico de *M. tuberculosis*, aunque en menor cantidad total. Ambos métodos fueron comparados considerando la cantidad y calidad del ADN obtenido, el tiempo de procesamiento de las muestras y la eficiencia en la generación de amplicones (Anexo 2).

Siete muestras de esputo BK positiva con al menos 2 mL de volumen fueron divididas en 2 alícuotas de 1 mL. El primer grupo de alícuotas fueron

descontaminadas con saponina al 5% y el segundo grupo de alícuotas fueron descontaminadas con NAOH-NALC al 0.5% tal como se muestra en la figura 1A.

### **8.1.2. Comparación de métodos de extracción de ADN mediante los métodos Fenol y Kit (N = 33).**

Se compararon dos protocolos de extracción de ADN a partir de las muestras de esputo (esputos BK positivos de pacientes, usados en una etapa inicial de estandarización): El método fenol hace uso del equipo FastPrep-24TM 5G el cual debido a la gran potencia de acción que posee y junto a las perlas de zirconia genera una lisis física sobre la muestra de esputo permitiendo la extracción del ADN, que por la carga negativa que esta presenta puede ser purificado mediante los AMPure beads. Por otro lado el método del kit también hace uso del equipo FastPrep-24TM 5G pero la obtención del ADN es mediante columnas de purificación (figura 1B).

- **Método Fenol:** Consistió en una lisis mecánica y química, al sedimento de la muestra se agregó fenol:cloroformo:alcohol isoamílico junto a perlas de zirconia y mediante el equipo FastPrep-24TM 5G se generó una lisis mecánica. El ADN crudo obtenido se purificó y concentró utilizando perlas magnéticas AMPure XP (Beckman Coulter) (Anexo 3).
- **Método Kit:** Se empleó un kit comercial de extracción de ADN genómico para micobacterias (Zymo) siguiendo el protocolo recomendado para esputo. Incluyó etapas de digestión enzimática y posterior adsorción del ADN a una membrana de sílice, con lavados y elución final en 100  $\mu$ L de buffer de elución proporcionado (Anexo 4).

A lo largo del proceso de extracción se trabajó siempre en cabina de bioseguridad y se incluyeron controles negativos (tubos sin muestra) para monitorear contaminación.

**Verificación de la calidad y tamaño del ADN genómico.** La concentración del ADN fue determinada mediante el fluorómetro Qubit 4 (Thermo Scientific); mientras que la calidad la calidad/pureza se determinó en el espectrofotómetro Nanodrop (Thermo Scientific) mediante las relaciones A260/280 y A260/230 medidas en Nanodrop. La integridad del ADN así como de la longitud de los fragmentos de ADN se determinaron mediante electroforesis en gel de agarosa al 1% y usando el Tape Station con los reactivos del kit Genomic DNA Screen Tape Analysis (Agilent) para obtener valores de DIN (*DNA Integrity Number*). En cada Screen tape se analizó 15 muestras de ADN extraído (1  $\mu$ L de muestra y 10  $\mu$ L de buffer) y el marcador de longitud de ADN.

### **8.1.3. Comparación de 2 métodos de preparación de librerías de ADN utilizando ADN genómico**

Las muestras de ADN extraídas mediante los métodos fenol y kit fueron preparadas para la secuenciación con Oxford Nanopore probando dos tipos de bibliotecas (figura 3):

**Biblioteca con *Native Ligation Barcode Kit 24*, ONT:** En este protocolo (“*Native Ligation*”), el ADN genómico extraído es unido mediante ligasas a un código de barras (barcodes) así como a los adaptadores. Se usaron 24 índices de código de barras (1 por cada muestra) para luego generar una librería múltiple con todas las

muestras con “barcodes” (multiplexar) hasta 12 muestras por proceso de secuenciación en una Flow Cell R9.4 (en 2 secuenciaciones en una misma Flow cell) (Anexo 5).

**Biblioteca con *Rapid Barcoding Kit 24, ONT***: En este protocolo (“*Rapid Barcoding*”), se hace uso de trasposones que funcionan como código de barras (barcodes) y ligasas para unir los adaptadores. Se utilizaron 24 trasposones en 2 grupos de 12, los cuales se secuenciaron en una misma celda en dos tiempos distintos (Anexo 6).

Las librerías fueron generadas del ADN obtenido de 12 esputos y 12 de cultivos primarios mediante el protocolo de Oxford Nanopore Technology para el *Rapid Barcoding Kit* y *Native Ligation Kit*.

#### **8.1.4. Secuenciación y procesamiento de datos de ADN genómico**

La corrida de secuenciación mediante tecnología Nanopore se realizó usando el software MinKNOW (versión 23.07.5), por un tiempo de hasta 48 horas o hasta consumir el 50% del total de poros iniciales. Para esta etapa, se optó por utilizar análisis en tiempo real, lo cual permitió ejecutar la asignación de bases (basecalling) de las lecturas que se hizo durante la secuenciación, esta conversión de señales fue llevada a cabo utilizando el programa Guppy (v6.1.2), empleando el modo de precisión ultra alta (súper high-accuracy), con el objeto de obtener secuencias en formato FASTQ de alta calidad.

Las lecturas resultantes fueron separadas por cada muestra individual en base a su “barcode” (“demultiplexear”) utilizando la herramienta QCAT, lo que permitió

asignar cada secuencia a la muestra correspondiente mediante códigos de barras específicas.

En una etapa posterior, se procedió a la identificación taxonómica de las lecturas obtenidas, con el propósito de distinguir aquellos provenientes o pertenecientes a *Mycobacterium tuberculosis* de las secuencias correspondientes a contaminantes, tales como ADN humano, u otras bacterias presentes en las muestras de esputo). Para lograr una clasificación robusta, se utilizó una estrategia basada en el consenso entre tres enfoques o métodos distintos: i) clasificación por *mapping* contra el genoma de referencia H37Rv (NC\_000962.3) usando Minimap2, ii) clasificación metagenómica con Kraken2 empleando una base de datos de micobacterias y genoma humano, y iii) filtrado por coincidencia con secuencias de *Mycobacterium* específicas (ej. regiones constantes del gen 16S rRNA de micobacterias). Solo se retuvieron las lecturas clasificadas como *M. tuberculosis* por al menos dos de los tres métodos. Este paso buscó eliminar lecturas (reads) de otras bacterias orales y del huésped que pudieran afectar los análisis posteriores.

Para las muestras secuenciadas como ADN genómico, de manera específica, se procedió a hacer ensamblado *de novo* mediante Flye v2.9 (cuando la cobertura lo permitía) y análisis de variantes/mutaciones respecto a H37Rv mediante los softwares *medaka* y *nanopolish*. En el caso de las muestras secuenciadas como amplicones, se realizó el alineamiento de las lecturas contra las secuencias esperadas de los genes amplificados y se determinó la frecuencia de variantes en cada posición eliminando las lecturas con un tamaño menor a 200pb para

determinar la detección de variantes nucleotídicas cuando la profundidad de esta superaba el 1% de la profundidad total de cada gen analizado (*katG*, *pncA* y *rpoB*). Adicionalmente, se calculó la profundidad de cobertura obtenida para cada método de biblioteca en cada muestra, definida como el número promedio de lecturas que cubrieron cada posición del genoma (o de los amplicones). Esto permitió comparar el rendimiento en datos entre la biblioteca *Native Ligation* y la *Rapid Barcoding*.

Para la secuenciación de ADN genómico se tomó el valor del tamaño del ADN obtenido por cada método de extracción directamente. 75 fmol de cada librería fue cargada en las “flow cells”. La determinación de fmol se realizó de la siguiente manera:

$$\text{Fmol} = 10^3 \times \text{pmol}$$

$$\text{Pmol} = \text{Concentración molar (mol/}\mu\text{L)} \times \text{volumen total (}\mu\text{L)}$$

$$\text{Concentración molar} = \frac{\text{Concentración (ng/}\mu\text{L)}}{\text{Tamaño del ADN pb} \times 660 \text{ (g/mol)}}$$

Donde, 660 g/mol: Peso promedio de un par de bases de ADN en daltons convertido a gramos por mol. Las librerías fueron distribuidas en las celdas teniendo en cuenta su procedencia (esputo o cultivo primario) y el tipo de librería (*Rapid Barcoding* o *Native Ligation*) (Figura 2).

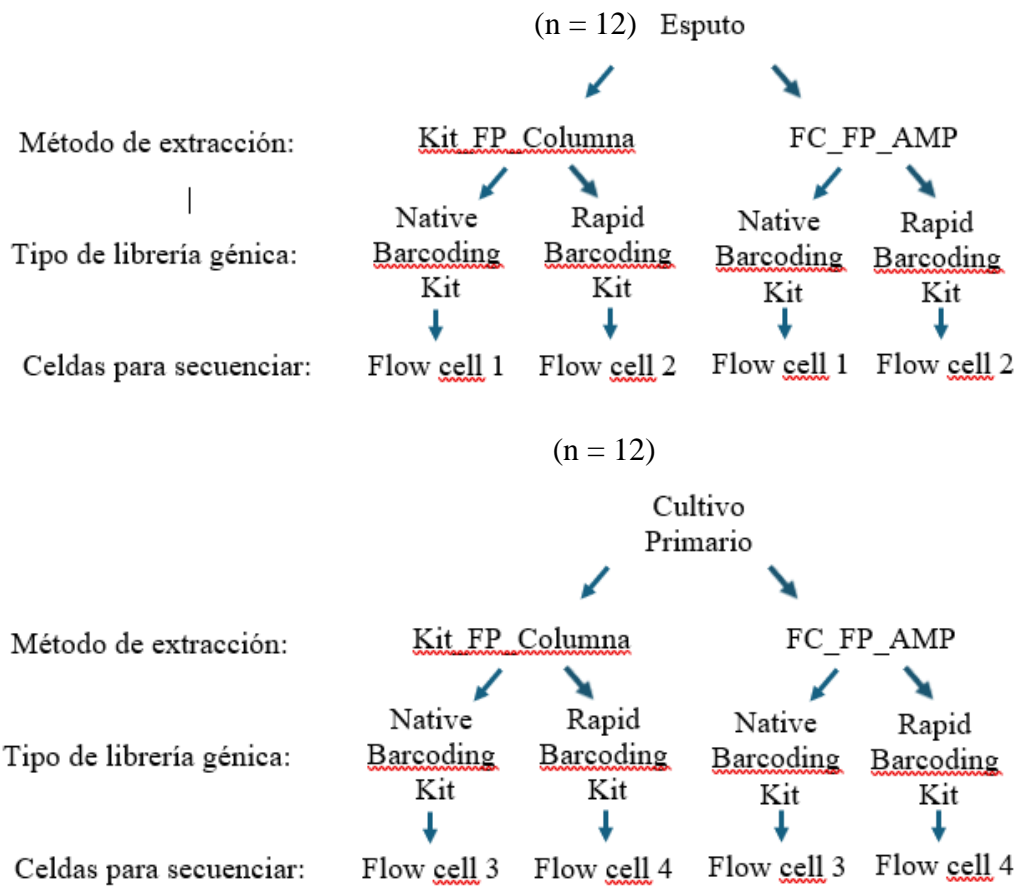


Figura 4. Distribución de las librerías de ADN genómico en las celdas de secuenciación.

El tiempo de corrida de cada celda fue de aproximadamente 20 horas o hasta que se consuma el 50% de los poros iniciales para luego cargar la segunda librería.

**8.1.5. Estandarización y obtención de amplicones a partir de muestras de esputo y cultivo de *M. tuberculosis* (14 genes entre 876 y 3,662 pb.)**

**Generación de amplicones.** Las muestras de ADN obtenidas por los dos métodos de extracción presentados previamente fueron utilizadas para obtener amplicones de 14 genes relacionados con el desarrollo de resistencia a los antibióticos de primera y segunda línea contra TB.

Los cebadores (primers) usados para amplificar 14 genes involucrados en la resistencia a los antibióticos de primera, segunda línea y fluoroquinolonas son mostrados en la tabla 1 (36). Los amplicones fueron obtenidos en 3 pools distintos en función de su tamaño (Anexo 7).

Tabla 1. Características de los 14 amplicones evaluados. Lista de los 14 amplicones utilizados en el estudio, incluyendo el gen objetivo, el tamaño del amplicón en pares de bases (pb) y las secuencias de los primers Forward y Reverse empleados para la amplificación.

| Gen                   | Tamaño gen pb | Primer Forward                | Primer Reverse                |
|-----------------------|---------------|-------------------------------|-------------------------------|
| <i>mabA-<br/>inhA</i> | 2902          | 5'-<br>AATGGAAATCGACTGGTCAGGT | 5'-<br>CTTGTAATCGCAACTCGTGGTC |
| <i>rpoB</i>           | 3662          | 5'- CTCCTCTAAGGGCTCTCGTT      | 5'- GCGGTAGCAAGACCGATG        |
| <i>ethA</i>           | 1828          | 5'- GACGGCCTCGACATTACGTT      | 5'- TTGCTGAAGTCGGTGCTCTC      |
| <i>katG</i>           | 2503          | 5'- GGACCATAACGGCTTCCTGT      | 5'- CATTTCGGCGCCCTTTCTC       |
| <i>rpsL</i>           | 939           | 5'- AATACACCGCTGCGACTAGG      | 5'- GACCAACTGCGATCCGTAGA      |

|                       |      |                               |                                |
|-----------------------|------|-------------------------------|--------------------------------|
| <i>embB</i>           | 3636 | 5'- GATCGGTGGAGCAGTACCA       | 5'-<br>GCTGAGGAGACTTTTGTGGACTG |
| <i>gyrB</i>           | 2380 | 5'- GCGCGGTTAGATGGGTAAA       | 5'- AAGCCGGAATCGAACATTG        |
| <i>Rrs</i>            | 1927 | 5'-<br>AGATCGAACGGGTATGCTGTTA | 5'- GGCCAACTTTGTTGTCATGC       |
| <i>oxyR</i>           | 1138 | 5'- GTGGTTGCTATGCGATCAGG      | 5'- GTCGAGCACTCGCAGTACCT       |
| <i>pncA</i>           | 876  | 5'- GACGGATTTGTCGCTCACTA      | 5'-<br>CCCGATGAAGGTGTCGTAGAA   |
| <i>bedA</i>           | 959  | 5'-<br>CGACCACAACCAGGATGAGCA  | 5'- CACGTGGCCGGCTCGGACAA       |
| <i>gyrA</i>           | 2978 | 5'- CGCCGCTGTACAAACTCAAG      | 5'- CCCGACTCCTAACACTCGTA       |
| <i>Eis</i>            | 1329 | 5'- CCGTACTTGCTCTGCAC         | 5'- CGAAGCAGCTGGGAATCTTT       |
| <i>embC-<br/>embA</i> | 1652 | 5'- ACAGCGGTTGACGCCTTACT      | 5'- GTACTIONGATGCGTGCCGATT     |

Las condiciones para cada PCR se muestran a continuación.

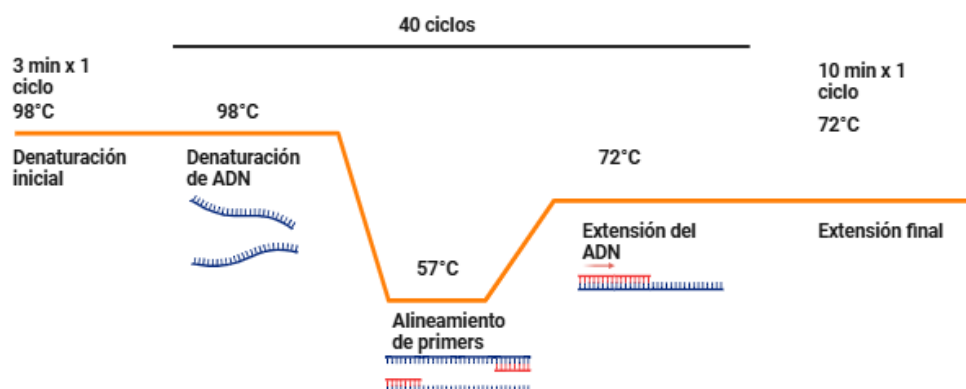


Figura 5. Programación utilizada para cada reacción de PCR

8.1.6. Comparación de 2 métodos de preparación de librerías de ADN utilizando amplicones (14 genes).

#### **Generación de librerías de amplicones.**

Las librerías fueron generadas del ADN obtenido de 12 esputos y 12 de cultivos primarios mediante el protocolo de Oxford Nanopore Technology para el *Rapid barcoding Kit* y *Native Ligation Kit* y secuenciadas separándolas por método de extracción y tipo de librería génica (figura 6).

Los “3 pools” de amplicones obtenidos por cada muestra fueron mezclados y purificados con el Kit DNA Clean & Concentrator®-5 (Anexo 8).

***Rapid barcoding Kit.*** 50 ng de amplicones, provenientes de la unión de los 3 pools con un total de 14 genes, en 9 µL de agua PCR fue usado para la generación de librerías (Anexo 9).

***Native barcoding Kit.*** 150 ng de amplicones en 11.5 µL agua PCR fue usado para la generación de la librería. La generación de librería hasta el agregado de adaptadores fue similar al realizado en la sección 8.1.3. excepto que en la etapa de reparación del ADN no se utilizó NEBNext FFPE DNA Repair Buffer y NEBNext FFPE DNA Repair Mix (Anexo 10).

#### **8.1.7. Secuenciación de amplicones (14 genes)**

Para la secuenciación de los amplicones, la longitud del ADN de cada librería fue considerado como el promedio de la longitud de todos los amplicones. 75 fmol de

cada librería fue cargada en las flow cells. La determinación de fmol se realizó de la siguiente manera:

$$\text{Fmol} = 10^3 \times \text{pmol}$$

$$\text{pmol} = \text{Concentración molar (mol/}\mu\text{L)} \times \text{volumen total (}\mu\text{L)}$$

$$\text{Concentración molar} = \frac{\text{Concentración (ng/}\mu\text{L)}}{\text{Tamaño del ADN pb} \times 660} \text{ (g/mol)}$$

Donde: 660 g/mol: Peso promedio de un par de bases de ADN en daltons convertido a gramos por mol. Las muestras fueron distribuidas en las celdas teniendo en cuenta su procedencia (esputo o cultivo primario) y el tipo de librería (*Rapid o Native barcoding*) (Figura 4).

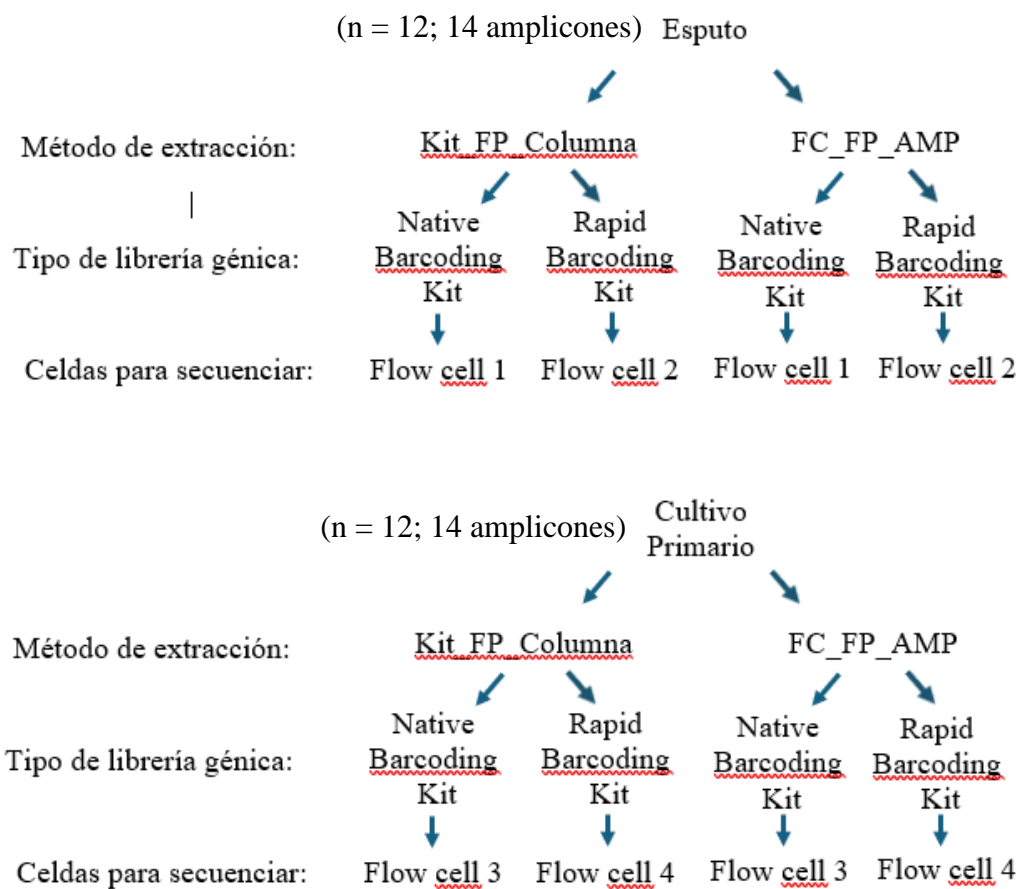


Figura 6. Distribución de las librerías de amplicones en las celdas de secuenciación.

El criterio para el cargado de la segunda librería dentro una misma Flow cell fue la misma que en el de la figura 4.

### 8.1.8. Obtención de amplicones de muestras descontaminadas con saponina al 5% y NALC/NaOH (4 genes).

Cinco  $\mu$ L de ADN de las muestras descontaminadas por NALC/NaOH y saponina al 5% fueron amplificadas usando un pool de 4 set de cebadores.

Tabla 2. Pool de primers: *inhA*, *katG*, *pncA* y *rpoB*

| Gen         | Tamaño del gen (pb) | Primer Forward                        | Primer Reverse                        |
|-------------|---------------------|---------------------------------------|---------------------------------------|
| <i>inhA</i> | 2902                | 5'-<br>AATGGAAATCGACTG<br>GTCAGGT- 3' | 5'-<br>CTTGTAATCGCAACTC<br>GTGGTC- 3' |
| <i>katG</i> | 2503                | 5'-<br>GGACCATAACGGCTTC<br>CTGT - 3'  | 5'-<br>CATTTCGGCGCCCTTT<br>CTC- 3'    |
| <i>pncA</i> | 876                 | 5'-<br>GACGGATTTGTCGCTC<br>ACTA- 3'   | 5'-<br>CCCGATGAAGGTGTC<br>GTAGAA- 3'  |
| <i>rpoB</i> | 856                 | 5'-<br>AATACACCGCTGCGAC<br>TAGG - 3'  | 5'-<br>GACCAACTGCGATCC<br>GTAGA- 3'   |

Las condiciones para el PCR fueron las mismas que en la sección 8.1.5. a excepción que en este caso el pool de primers tuvo un volumen total de 3.5  $\mu$ L en un volumen de reacción de PCR final de 25  $\mu$ L. El volumen y concentración de los primers fue

de 0.35  $\mu\text{L}$ , 0.7  $\mu\text{L}$ , 0.35  $\mu\text{L}$ , 0.35  $\mu\text{L}$  y 0.14  $\mu\text{M}$ , 0.36  $\mu\text{M}$ , 0.06  $\mu\text{M}$ , 0.08  $\mu\text{M}$  para los primers *inhA*, *katG*, *pncA*, *rpoB* forward y reverse respectivamente. Los amplicones fueron evaluados en la electroforesis en un gel de electroforesis al 0.8% a 96 V.

#### **8.1.8.1 Determinación de la cantidad mínima de ADN para la obtención de amplicones usando un pool de 4 amplificaciones en muestras de cultivos primarios extraídos por el método Kit.**

Cuatro muestras de ADN seleccionadas al azar extraídas en la sección 8.1.2. fueron llevadas a una concentración de 10 ng/ $\mu\text{L}$  en un volumen aproximado de 10  $\mu\text{L}$ , seguidamente se generaron diluciones seriadas  $10^{-1}$ ,  $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$  y  $10^{-5}$  para evaluar la mínima cantidad de ADN necesaria mediante la descontaminación con saponina para obtener los 4 amplicones del nuevo pool.

#### **8.1.9. Comparación de librerías para un pool de 4 genes de resistencia utilizando los kits *Rapid Barcoding* y *Native Ligation*.**

Los genes que confieren resistencia a 3 de los 4 antibióticos usados en la primera línea contra TB fueron amplificados (*katG*, *inhA*, *pncA* y *rpoB*).

##### **8.1.9.1. Generación de librerías de amplicones.**

Las librerías fueron generadas de amplicones obtenidos de 12 esputos obtenidos en la sección 8.1.2. mediante el protocolo de Oxford Nanopore Technology para el *Rapid barcoding Kit* y *Native Ligation Kit*.

**Rapid barcoding Kit.** Cincuenta ng de amplicones que contenían los 4 genes de interés (*inhA*, *pncA*, *katG* y *rpoB*), en 9  $\mu\text{L}$  de agua PCR fueron usados para la generación de librerías. El proceso hasta el agregado de adaptadores fue similar al descrito con ADN genómico en la sección 8.5.2

**Native Ligation Kit.** La cantidad de 150 ng de amplicones en 11.5  $\mu\text{L}$  agua PCR fue usado para la generación de la librería. La generación de librería hasta el agregado de adaptadores fue similar al realizado en la sección 8.5.2

#### **8.1.10. Secuenciación de amplicones (4 genes).**

Para la secuenciación de los amplicones, el promedio de la longitud de todos los amplicones fue considerado como la longitud del ADN de cada librería. 75 fmol de cada librería fue cargada en las flow cells. La determinación de fmol se realizó de la siguiente manera:

$$\text{Fmol} = 10^3 \times \text{pmol}$$

$$\text{Pmol} = \text{Concentración molar (mol}/\mu\text{L}) \times \text{volumen total } (\mu\text{L})$$

$$\text{Concentración molar} = \text{Concentración (ng}/\mu\text{L}) \times \text{Tamaño del ADN pb} \times 660 \text{ (g/mol)}$$

Donde:

660 g/mol: Peso promedio de un par de bases de ADN en daltons convertido a gramos por mol. Las muestras fueron distribuidas en las celdas teniendo en cuenta su procedencia (esputo o cultivo primario) y el tipo de librería (*Rapid Barcoding* o *Native Ligation*) (Figura 4).

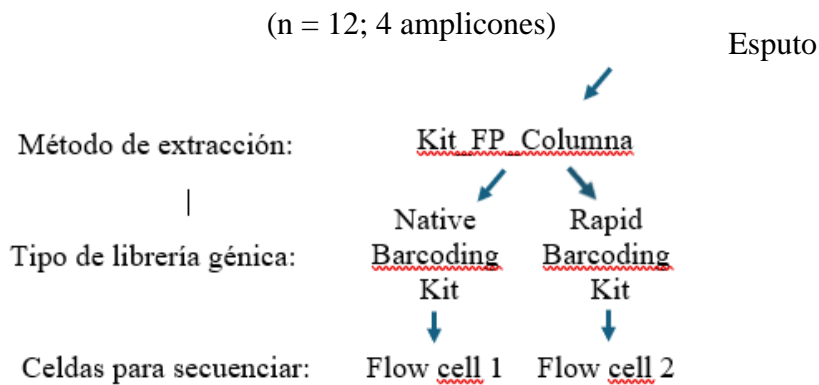


Figura 7. Distribución de las librerías de amplicones en las celdas de secuenciación.

### 8.1.11. Comparación de tipos de ADN usados para la secuenciación.

Los datos obtenidos secuenciación del ADN genómico así como de la secuenciación de amplicones permitieron la elección del tipo de ADN a trabajar en la segunda fase en función de la cobertura, profundidad y q-score que presentaron las lecturas de *M. tuberculosis* de cada secuenciación

### 8.2.1 Comparación del límite de detección de heterorresistencia en mezclas bacterianas mediante secuenciación y método de proporciones.

Las muestras de esputo (BK negativas) con “spike” de mezcla bacteriana fueron sometidas a descontaminación mediante el método NALC-NaOH. Este sedimento descontaminado se dividió para dos usos una parte para cultivo en agar y otra parte para la extracción de ADN y generación de amplicones y comparar el método de proporciones en agar y secuenciación Nanopore para la detección de heterorresistencia siguiendo el diagrama de la figura 3.

### **8.2.2. Población de estudio y muestras**

Dos cepas de referencia de *M. tuberculosis*; cepa de referencia sensible H37Rv y cepa multidrogo resistente DM97 provenientes del cepario del Laboratorio de Micobacterias de la Universidad Peruana Cayetano Heredia. Se seleccionaron muestras de esputo con baciloscopia negativa y GenXpert negativo.

### **8.2.3. Preparación de medios de cultivo y reactivación de cepas bacterianas.**

Las cepas H37Rv y DM97 se encontraron criopreservadas por lo que se necesitó reactivarlas mediante su cultivo en medio líquido y medio sólido posteriormente (Anexo 11).

### **8.2.4. Preparación de mezclas bacterianas heterorresistentes**

Las mezclas heterorresistentes se realizaron utilizando la cepa susceptible (H37Rv) y resistente (DM97). Se generaron 5 mezclas principales de bacilos susceptibles/resistentes para cada antibiótico además de una mezcla compuesta únicamente por la cepa susceptible y otra conformada exclusivamente por la cepa resistente. Las cinco mezclas principales de bacilos susceptibles/resistentes para cada antibiótico evaluado presentó las siguientes proporciones de cepa resistente: 1%, 5%, 10%, 25% y 50% en un volumen de 150  $\mu$ L con un OD 0.2, que equivale a  $1 \times 10^6$  UFC/mL (91) (Anexo 12). Cada mezcla se preparó en duplicado y se inoculó en alícuotas de 1 mL de esputo previamente homogenizado. Estas muestras simuladas de esputo HR se etiquetaron y procesaron en paralelo tanto para cultivo en agar como para secuenciación.

Para cada mezcla, se prepararon 3 repeticiones usando cultivos de las cepas de fechas diferentes.

### 8.2.5. Descontaminación con el método NALC/NaOH de muestras de esputo con BK negativo

Siete muestras de esputo con BK negativo fueron mezcladas hasta obtener un volumen de 15 mL. Las muestras fueron homogenizadas con agitación constante (180 RPM) por 10 min; luego, la muestra de esputo fue distribuida en 3 tubos con 5 mL. Cada muestra se descontaminó mediante el método NALC/NaOH y el sedimento fue resuspendido en 3 mL de PBS. 150 µL de esputo descontaminado fueron alicuotados en 14 tubos. 7 alícuotas fueron utilizadas para el método de proporciones y los 7 restantes para el método de secuenciación. Este proceso fue repetido 2 veces con nuevas muestras de esputo con BK negativo.

### 8.2.6. Spike de mezclas heterorresistentes en muestras de esputo negativo descontaminado y PBS.

**Muestras para secuenciación.** 50 µL de cada una de las 7 mezcla HR fue mezclado con 150 µL de esputo negativo descontaminado y con 150 de PBS.

**Muestras para agar en proporciones.** 50 µL de cada una de las 7 mezcla HR fue mezclado con 150 µL de esputo negativos descontaminado y con 150 µL de PBS.

#### Mezclas Heterorresistentes (n = 42; 4 amplicones)

|    | Con esputo   |    | Sin esputo   |
|----|--------------|----|--------------|
|    | M1 (100S)    |    | M1 (100S)    |
|    | M2 (99S:1R)  |    | M2 (99S:1R)  |
|    | M3 (95S:5R)  |    | M3 (95S:5R)  |
|    | M4 (90S:10R) |    | M4 (90S:10R) |
| x3 | M5 (75S:25R) | x3 | M5 (75S:25R) |
|    | M6 (50S:50R) |    | M6 (50S:50R) |
|    | M7 (100R)    |    | M7 (100R)    |



Secuenciación

Figura 8. Distribución de muestras de esputo con spike de cepas bacterianas para detección de heterorresistencia mediante secuenciación Nanopore.

### **8.2.7. Determinación de la susceptibilidad a isoniácida, pirazinamida y rifampicina en muestras de esputo y cultivo mediante el método de proporciones en agar (cultivo).**

El estudio evaluó la susceptibilidad a isoniácida, rifampicina y pirazinamida en muestras BK negativas con spike de mezclas HR de esputo y cultivo utilizando el método de proporciones en agar. Esta técnica permitió una medición cuantitativa de la HR hasta niveles tan bajos como el 1%. Para cada mezcla preparada con distintas proporciones de bacilos resistentes, se sembraron volúmenes iguales en placas con medio 7H10, distribuidos en un cuadrante sin antibiótico (control de crecimiento) y en cuadrantes con las concentraciones críticas de isoniácida, rifampicina o pirazinamida en medio 7H10. Los cultivos fueron incubados a 37 °C y monitoreados semanalmente.

Luego de 3 a 4 semanas, se contaron las colonias presentes en cada tubo y se calculó la proporción de bacilos resistentes según la fórmula:

$$(\text{UFC en medio con fármaco}/\text{UFC en medio control}) \times 100$$

Se consideró resistencia fenotípica si esta proporción era  $\geq 1\%$  (criterio estándar). De esta manera se obtuvo, para cada muestra, el porcentaje de bacilos resistentes “detectado” por el método de proporciones, que luego se comparó con el porcentaje de cepas resistentes detectados por secuenciación.

**Sembrado de las mezclas HR en placas del método de proporciones en agar.** El esputo descontaminado con las mezclas HR fue diluido 1/10 y 1/100. 30  $\mu\text{L}$  de las muestras (3 gotas equidistantes, 10  $\mu\text{L}$ /gota) fueron sembradas cada cuadrante de una placa dividida en 4 cuadrantes. Cada cuadrante contenía 7H11-OADC con lo siguiente: Control (sin antibiótico), isoniacida (0.4  $\mu\text{g}/\text{mL}$ ), rifampicina (1  $\mu\text{g}/\text{mL}$ ) y pirazinamida (100  $\mu\text{g}/\text{mL}$ ). Las placas se incubaron a 37°C durante 2-3 semanas. Este procedimiento se repitió para las diluciones 1/10 y 1/100 generando 21 placas con mezclas heterorresistentes en esputo y 21 placas con mezclas en PBS. El conteo de colonias se realizó cuando se observaron colonias separadas y se estimó el número de colonias cuando se observaron colonias confluentes.

**Conteo de colonias.** El conteo de colonias se realizó luego de 3 semanas de incubación de las placas a 37°C mediante el uso de microscopio invertido para detectar presencia o ausencia de contaminación. En cada placa se registró por separado las colonias presentes en el cuadrante control (sin antibiótico) y las detectadas en los cuadrantes que contenían las concentraciones críticas de isoniacida, rifampicina o pirazinamida. Luego, se calculó la proporción de resistencia dividiendo el número de colonias observadas en el medio con antibiótico entre el número de colonias del medio control, multiplicando el resultado por 100.

### **8.3. Extracción del ADN mediante el método Kit.**

La inactivación de las micobacterias (mezclas HR) en las muestras de esputo negativos descontaminados, extracción de ADN, amplificación, preparación de librerías y secuenciación fueron realizados según los métodos NALC/NaOH, tiocianato de guanidina, Kit y *Native Ligation* respectivamente y la secuenciación dentro de un equipo de tipo PromethION.

#### **8.3.1. Generación de amplicones (4 genes).**

Se generó un pool de 4 amplicones para la detección de los genes *inhA*, *pncA*, *katG* y *rpoB* los cuáles están relacionados a resistencia a fármacos de primera y segunda línea.

#### **8.3.2. Preparación de librería génica.**

Se utilizó el kit *Native Ligation* debido a la mejor calidad de lecturas obtenidos, profundidad y cantidad de data producida en comparación al kit *Rapid Barcoding*.

### **8.3.3. Secuenciación.**

La calidad y filtrado de datos se realizó utilizando los mismos criterios que en la primera fase agregando la detección de HR detallada a continuación:

Se definió como detección de HR por secuenciación el hallazgo de mutaciones puntuales de resistencia en un gen (p. ej. mutación S315T en *katG* para isoniazida, S531L en *rpoB* para rifampicina) con una frecuencia de lecturas portadoras mayor o igual al porcentaje de cepas resistentes iniciales que había en la mezcla HR. Por ejemplo, si en la muestra M3 10/90 (10% resistente rifampicina) el alineamiento de lecturas mostraba que el 12% de las lecturas de *rpoB* contenían la mutación S531L y el resto presentaba la secuencia silvestre, se consideró que la secuenciación detectó HR (con la subpoblación resistente representada por aproximadamente 12% de las lecturas).

## **8.4. Análisis de datos de secuenciación.**

### **8.4.1. Control de calidad y validación.**

La data cruda de lecturas fue analizada con el software Dorado v0.7.3 usando el modelo de precisión ultra alta (High Super accuracy) (dna\_r10.4.1\_e8.2\_400bps\_sup). La eliminación de las secuencias de los adaptadores, secuencias cortas (< 200 pb) y lecturas de baja calidad se realizó mediante el software de Nanopore-MinKNOW v23.07.5. La calidad de los resultados fue verificada a través de la inclusión de los siguientes controles: 1) Filtración de las lecturas (reads) con un q-score menor a 20 para asegurar una precisión del 99% dentro del ADN secuenciado; 2) Ensamblaje y anotación se hizo

mediante un consenso de 3 softwares para asegurar una alta fidelidad de la clasificación taxonómica y detección de mutaciones.

Los datos brutos de secuenciación (FASTQ) fueron depositados en el repositorio público NCBI SRA, y el análisis bioinformático reproducible se documentó mediante scripts en Python y R disponibles en GitHub (GitHub - epi2me-labs/wf-tb-amr).

#### **8.4.2. Identificación de lecturas “lecturas” de *Mycobacterium tuberculosis*.**

Las lecturas de *M. tuberculosis* fueron identificados usando un consenso de 3 enfoques. En el primer enfoque se clasificó taxonómicamente a nivel de lecturas, utilizando los softwares Kraken2 combinado con BRAKEN. Las abundancias fueron estimadas a partir de la base de datos estándar (secuencias génicas ya reportadas). El segundo método se basó en el alineamiento de las lecturas de Nanopore con respecto al genoma de referencia de *M. tuberculosis* H37Rv, empleando minimap2 para la alineación. En el tercer enfoque se usó el software KrakenUnique, basada en el uso de una base de datos de patógenos especializada.

#### **8.4.3. Determinación de la profundidad**

La profundidad de cobertura de los genes de resistencia en las muestras de amplicones se determinó mediante el mapeo de las secuencias de ADN usando como referencia el genoma de *M. tuberculosis* H37Rv y el programa Mosdepth. En el análisis de las secuencias (archivos “. bed”) se excluyó los alineamientos no primarios, así como las lecturas complementarias y suplementarias. Las lecturas obtenidas de los amplicones a partir de los cultivos primarios fueron comparados a los resultados obtenidos de los amplicones de las muestras de esputo.

#### **8.4.4. Análisis estadísticos**

Para la determinación del tamaño muestral que se utilizó para comparar los métodos de extracción se utilizó la fórmula de tamaño de muestra para poblaciones finitas.

Para la comparación de los métodos de extracción de ADN se utilizó la prueba no paramétrica Wilcoxon en donde se comparó la cantidad y calidad de ADN obtenida mediante cada método.

### **9. RESULTADOS**

#### **9.1. Estandarización de métodos de descontaminación, extracción de ADN, preparación y generación de librerías génicas y tipo de ADN usando parámetros de secuenciación a partir de muestras de esputo y cultivos primarios.**

##### **9.1.1 Ensayo comparativo de descontaminación con saponina al 5% y NAOH-NALC al 0.5%.**

Los métodos de descontaminación se compararon utilizando 7 muestras de esputo con baciloscopia positiva utilizando el método de extracción Kit. La figura 9 muestra la diferencia entre las cantidades de ADN que se obtienen tras procesar las muestras mediante saponina al 5% o NALC/NaOH al 0.5%, además de los “ratios” de calidad mostrando a la saponina como la alternativa adecuada para ensayos en donde se requiera altas cantidades de ADN. Sin embargo, se debe considerar que esta diferencia en la cantidad de ADN obtenida mediante cada método se debe a su capacidad para eliminar el ADN de bacterias así como de células humanas durante su proceso por lo que una elevada cantidad de ADN no indica necesariamente mayor cantidad de *M. tuberculosis* dentro de la muestra.

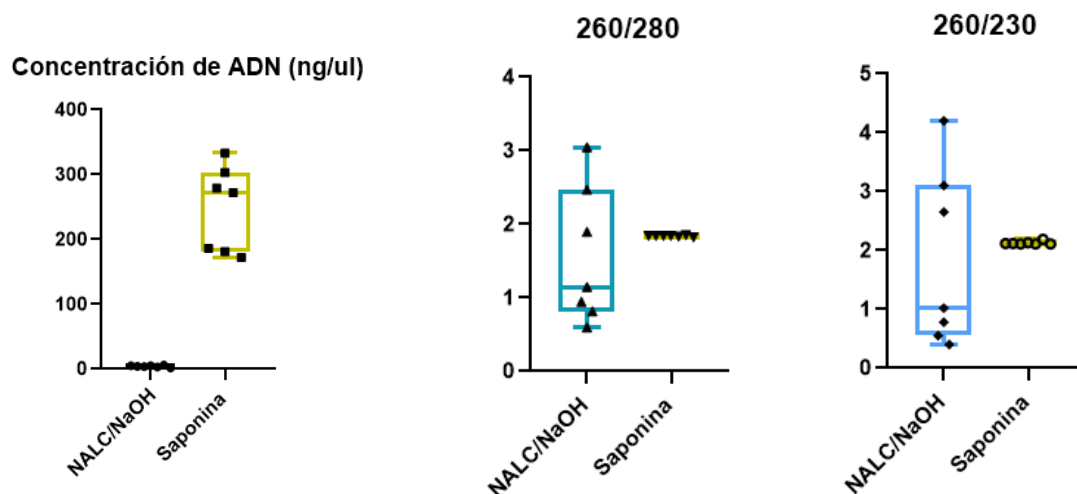


Figura 9. Concentración y calidad de ADN a partir de muestras de esputo mediante descontaminación con saponina y NALC/NaOH.

### 9.1.2. Comparación de métodos de extracción de ADN mediante los métodos Fenol y Kit (N = 33).

Los resultados de la extracción de ADN de 33 muestras de esputo BK positivas usando el método Fenol y Kit en términos de rendimiento (cantidad total de ADN en ng), concentración (ng/ $\mu$ L) y ratios de calidad espectrofotométricos (260/280 para pureza respecto a proteínas y 260/230 para presencia de contaminantes orgánicos y sales) son mostrados en la figura 10.

El método Fenol produjo en promedio  $18.4 \pm 3.5$   $\mu$ g de ADN por muestra, superior al rendimiento del kit comercial ( $7.6 \pm 2.1$   $\mu$ g;  $p < 0.01$ ) (Figura 10). La pureza del ADN fue adecuada en ambos métodos (ratios A260/280  $\sim 1.8$ – $1.9$ ), aunque el método de Kit mostró ligeramente menor relación A260/230 en algunas muestras (indicativo de trazas de sales). Cabe destacar que el método con beads AMPure

permitió concentrar eficientemente el ADN en volúmenes menores (50  $\mu\text{L}$  vs 100  $\mu\text{L}$ ), lo cual resultó beneficioso para cumplir los requerimientos de entrada en la preparación de bibliotecas.

La Figura 10 también, ilustra los rendimientos individuales de ADN por cada muestra y método, evidenciando la superioridad consistente del método Fenol. Este método utiliza solventes orgánicos y beads magnéticos que capturan el ADN de forma más efectiva incluso en matrices complejas. En cambio, el método de Kit está limitado por la capacidad de adsorción de la membrana de sílice ( $\sim 25 \mu\text{g}$  máximo por columna según el fabricante) y mostró saturación en muestras con alto contenido de ADN. Asimismo, se observó que muestras con esputo muy viscoso tendieron a obstruir las columnas, reduciendo su eficiencia, mientras que los beads, aunque también pueden verse afectadas por exceso de mucina, permitieron una recuperación más uniforme tras el pretratamiento con NALC y/o saponina. En síntesis, tanto el método Fenol como el kit proporcionaron una cantidad y calidad de ADN óptimas para los posteriores análisis de secuenciación.

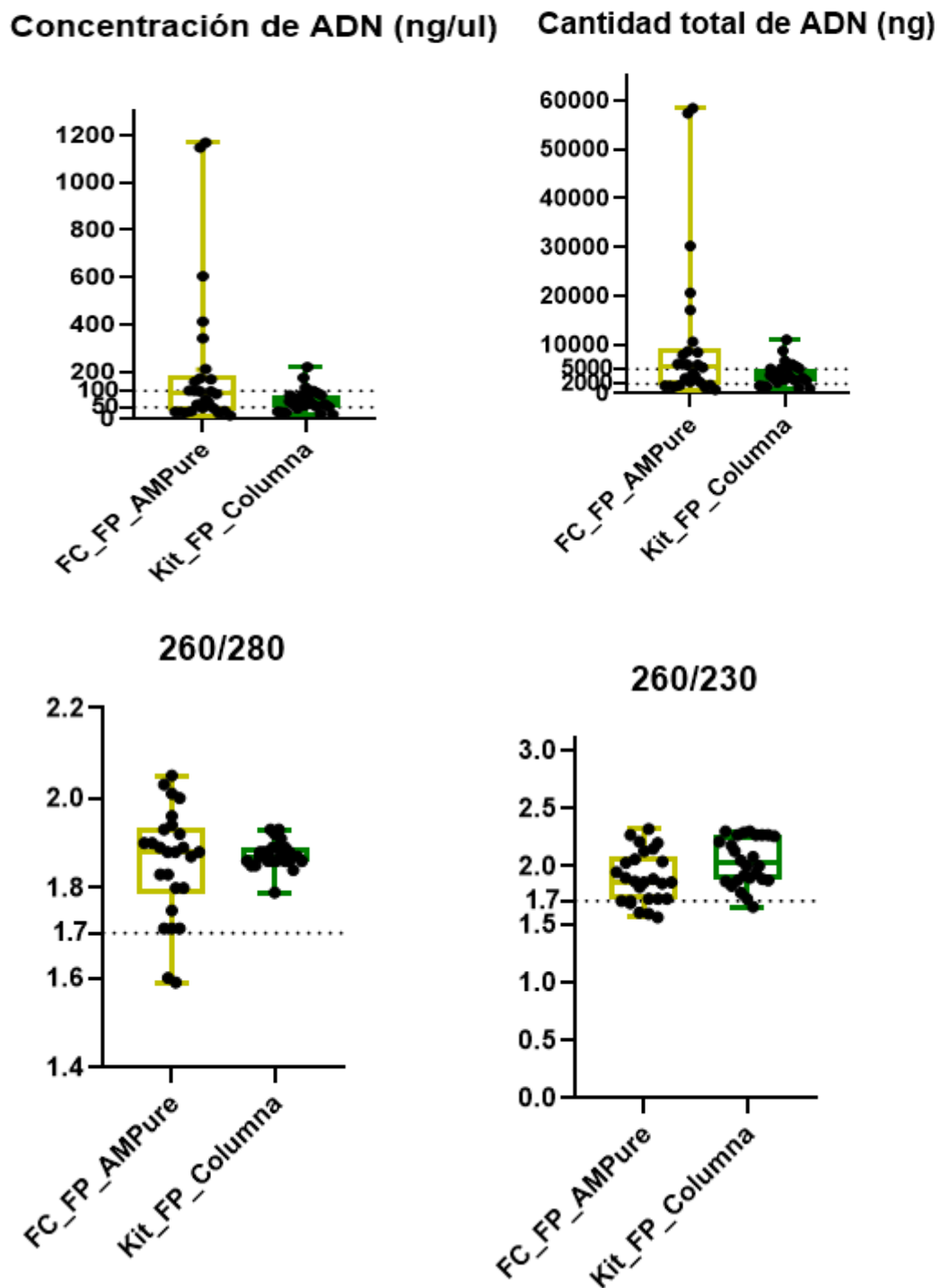


Figura 10. Comparación de la cantidad, concentración y calidad del ADN obtenido mediante dos métodos de extracción

**Evaluación de integridad de muestras extraídas por Kit y FC.** Del total de las 33 muestras extraídas, se hizo una selección al azar de 12 de estas a las cuales se evaluó la integridad y tamaño del ADN tal como se presenta en la figura 11. Se muestra la comparación de dos métodos de extracción de 11 muestras, evaluando la integridad y distribución del tamaño del ADN mediante TapeStation. Se determinaron los perfiles electroforéticos y los valores de DNA Integrity Number (DIN) para estimar la calidad del ADN extraído.

En términos de integridad, el ADN extraído con el método Fenol presentó un tamaño de fragmento predominante de ~9–10 kb, mientras que el extraído por columnas presento mayor longitud (~15 kb). Si bien ambos métodos de extracción muestran una longitud de ADN en mayor proporción, se debe considerar que dependiendo de la muestra estos valores pueden variar. Muestras de esputo con muchos días de guardado pueden llegar a generar ADN degradado y esto impacta

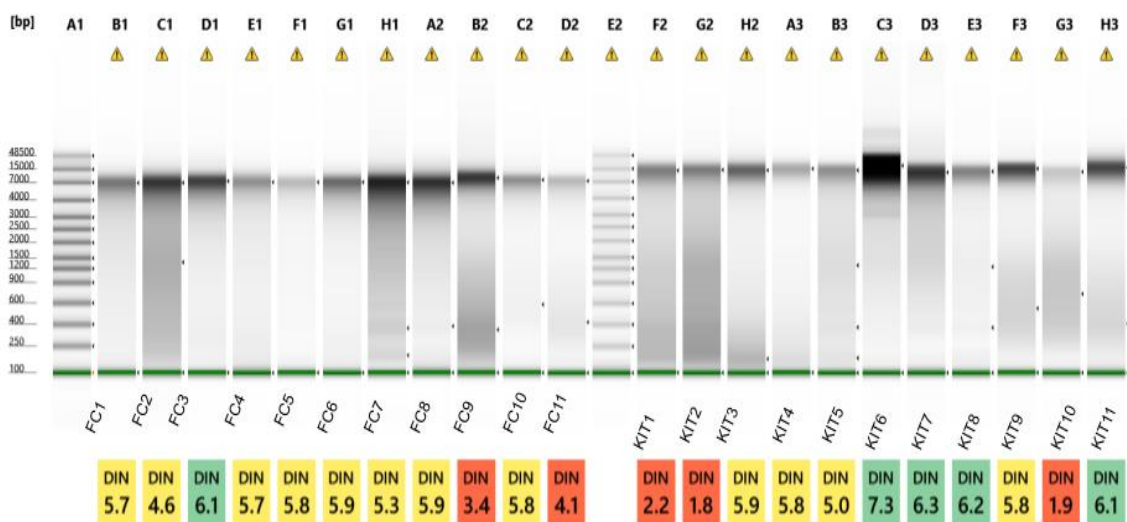


Figura 11. Análisis de integridad y tamaño del ADN extraído en 11 muestras mediante TapeStation; FC: Método fenol, Kit: Método Kit

en el valor DIN y en los cálculos requeridos para el input de fmoles en secuenciación.

### **9.1.3. Obtención de 14 amplicones a partir de muestras de esputo y cultivos puros de *M. tuberculosis*.**

Las 12 muestras de ADN pertenecientes a pacientes con baciloscopia positiva y extraídas con ambos métodos de extracción fueron alicuotados para la obtención de amplicones (figura 12). En contraste, se extrajo ADN de 12 muestras de cultivo de *M. tuberculosis H37Rv* mediante los métodos de extracción fenol y kit, a las cuales también se les generó amplicones como se observa en las figuras 13 y 14.

La amplificación por PCR de los 14 genes seleccionados fue exitosa en las muestras provenientes de cultivos puros, mostrando bandas definidas para cada gen. Sin embargo, en las muestras directas de esputo, la PCR multiplex presentó algunas inespecificidades y menor intensidad en los productos amplificados. En los geles de electroforesis, las muestras de esputo mostraron bandas más tenues y en algunos casos bandas adicionales débiles, indicativas de amplificación inespecífica posiblemente debida a la presencia de ADN de otras bacterias o del propio humano en la reacción. Aun así, se lograron visualizar las 14 bandas esperadas por muestra, confirmando la amplificación de todos los genes objetivo, aunque con variaciones en intensidad.

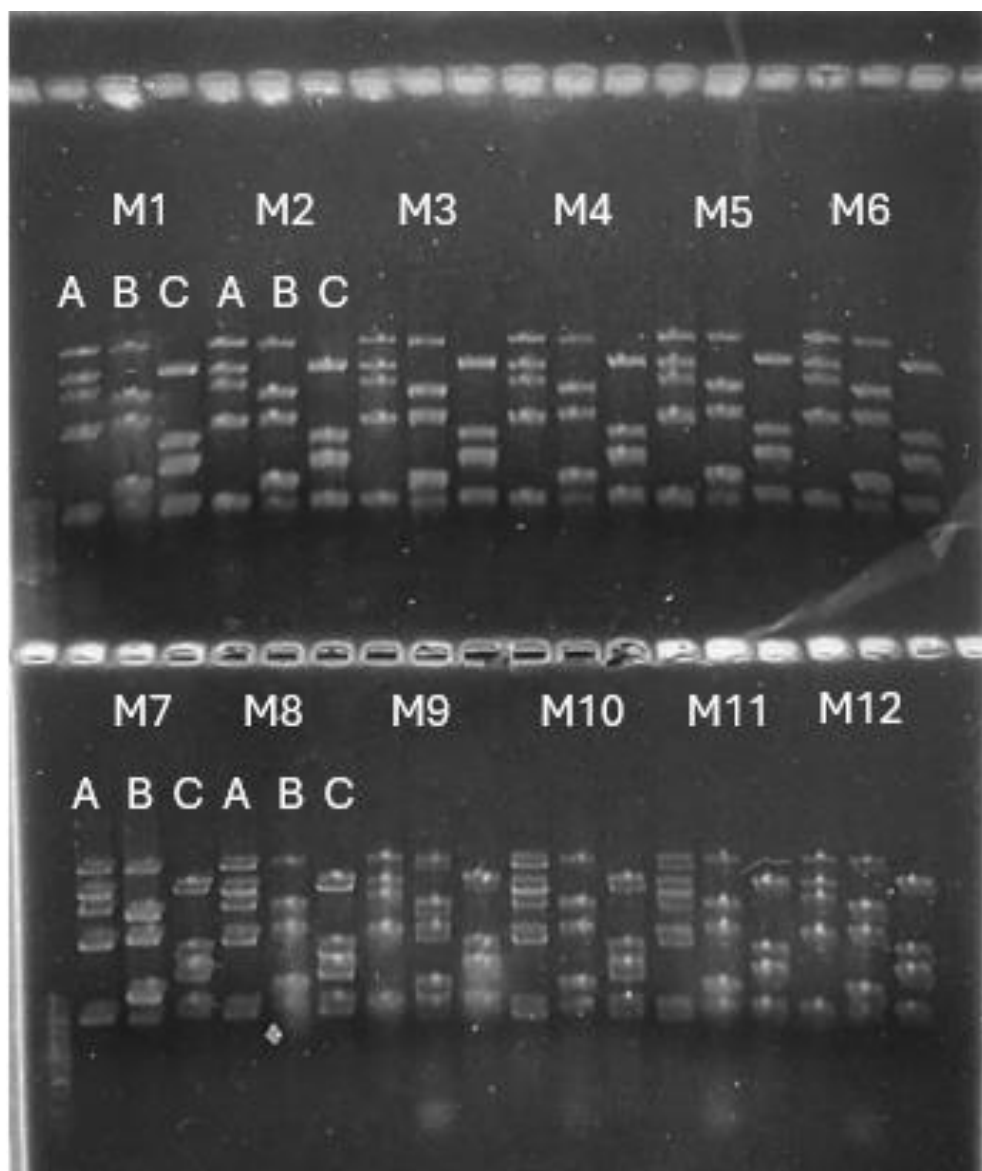


Figura 12. Análisis de 14 amplicones generados a partir de una muestra de cultivo extraída mediante el método fenol (n =12).

Se evaluó la amplificación de 14 fragmentos específicos, organizados en tres pools de amplicones, mediante electroforesis en gel de agarosa. La separación de los fragmentos confirma la presencia de los productos esperados para cada pool. Los 3 primeros carriles corresponden al Pool A, B y C respectivamente para cada una de las 12 muestras. Pool A: *rpoB* (3662 pb), *inhA* (2902 pb), *katG* (2503 pb), *ethA*

(1828 pb), *rpsL* (939 pb); Pool B: *embB* (3636 pb), *gyrB* (2380 pb), *rrs* (1927 pb), *oxyR* (1138 pb), *pncA* (876 pb); Pool C: *gyrA* (2978 pb), *embC* (1652 pb), *eis* (1329 pb), *BedA* (959 pb).

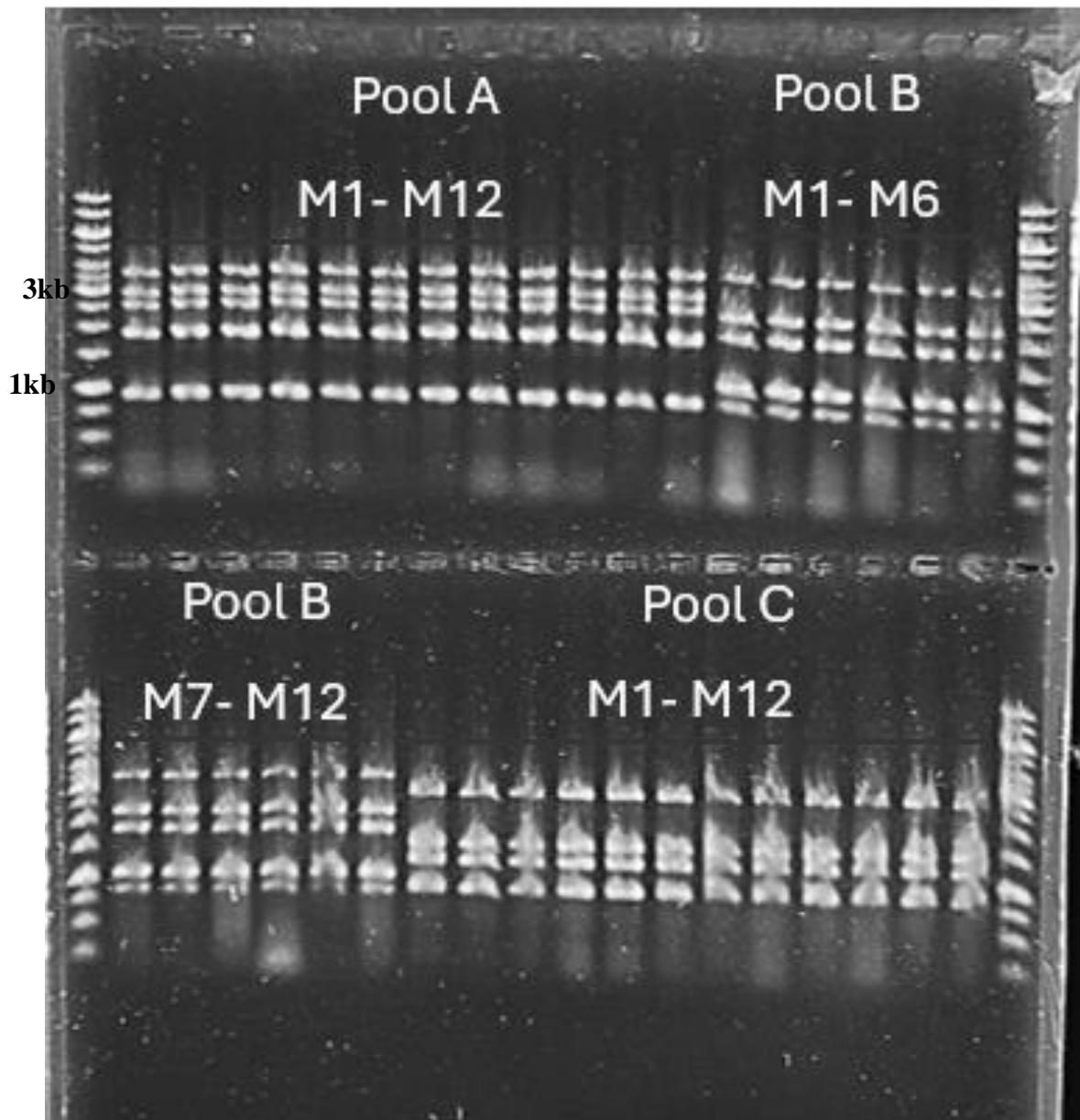


Figura 13. Análisis de 14 amplicones generados a partir de una muestra de cultivo extraída mediante el método Kit (N = 12).

Se evaluó la amplificación de 14 fragmentos específicos, organizados en tres pools de amplicones, mediante electroforesis en gel de agarosa. La separación de los fragmentos confirma la presencia de los productos esperados para cada pool. Los 12 primeros carriles corresponden al Pool A para cada una de las 12 muestras, así como los 12 siguientes para el pool B y pool C respectivamente. Pool A, B y C respectivamente para cada una de las 12 muestras. Pool A: *rpoB* (3662 pb), *inhA* (2902 pb), *katG* (2503 pb), *ethA* (1828 pb), *rpsL* (939 pb); Pool B: *embB* (3636 pb), *gyrB* (2380 pb), *rrs* (1927 pb), *oxyR* (1138 pb), *pncA* (876 pb); Pool C: *gyrA* (2978 pb), *embC* (1652 pb), *eis* (1329 pb), *BedA* (959 pb).

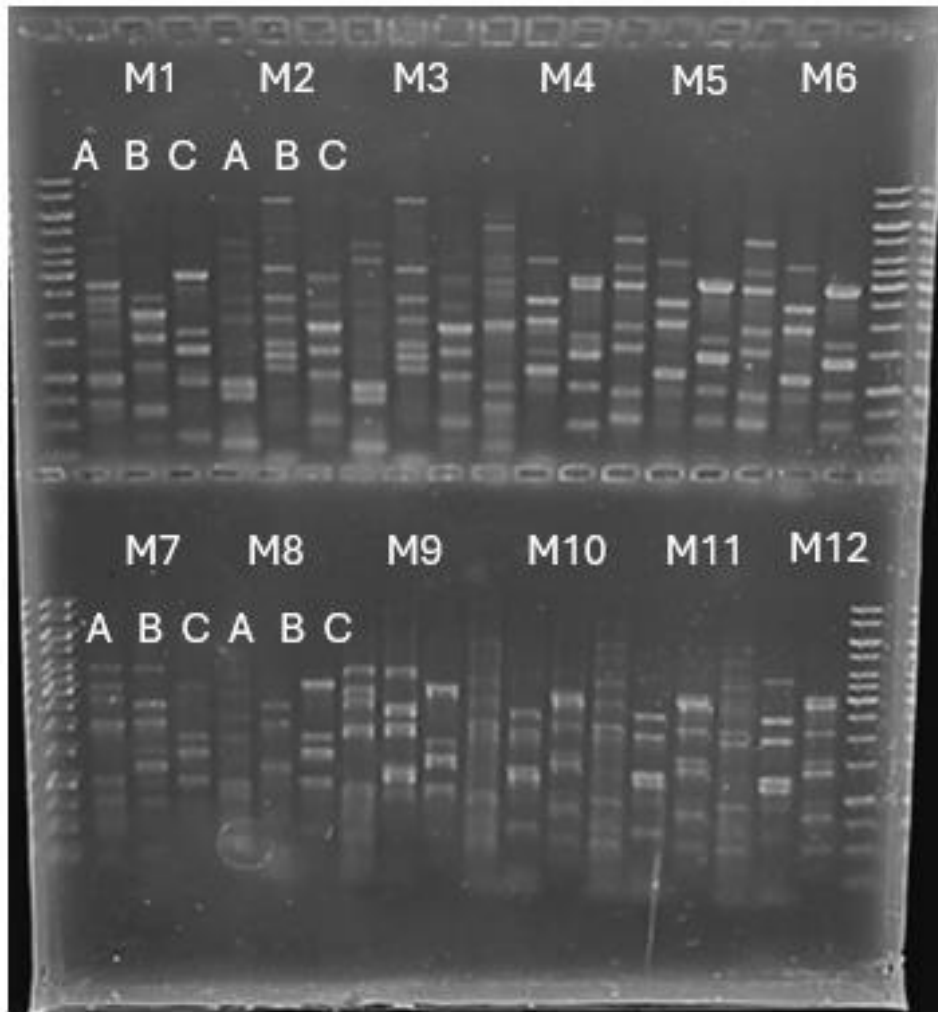


Figura 14. Análisis de 14 amplicones generados a partir de una muestra de esputo extraída mediante el método fenol (N = 12).

Se evaluó la amplificación de 14 fragmentos específicos, organizados en tres pools de amplicones, mediante electroforesis en gel de agarosa. La separación de los fragmentos confirma la presencia de los productos esperados para cada pool. Los 3 primeros carriles corresponden al Pool A, B y C respectivamente para cada una de las 12 muestras. Se observa la presencia de fragmentos inespecífico-amplificados propios de una muestra de esputo.

#### 9.1.4. Obtención de 4 amplicones de un pool con 4 genes de resistencia (*inhA*, *katG*, *pncA* y *rpoB*).

Una vez que se obtuvieron 14 amplicones de una muestra se decidió evaluar un pool conformado por los 4 principales genes involucrados en el desarrollo de resistencia a los medicamentos de primera y segunda línea.

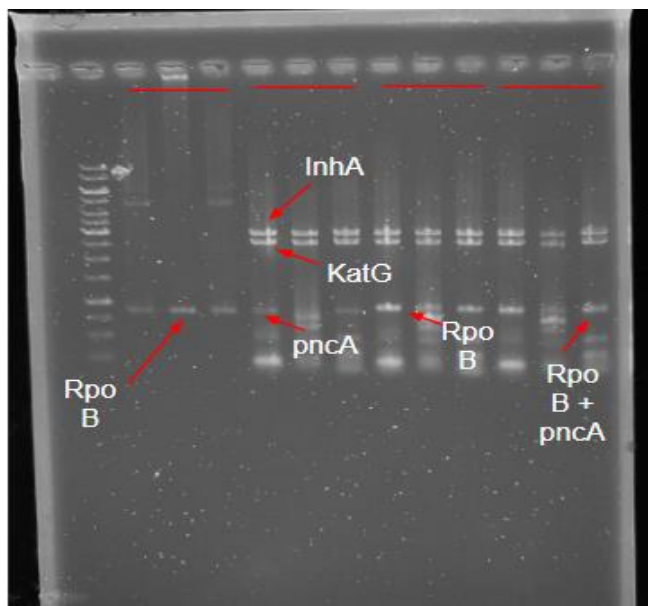


Figura 15. Evaluación de la inserción del nuevo gen *rpoB* y en un pool de cuatro amplicones mediante electroforesis en gel de agarosa.

Se observan bandas correspondientes a los genes *inhA*, *katG*, *pncA* y *rpoB*, indicando la amplificación individual de cada uno. Los 3 primeros carriles indican la amplificación de solo el gen *rpoB*. Los carriles 4 al 6 muestran el pool de 3 genes *inhA*, *katG* y *pncA* sin el *rpoB*. Los carriles 7 al 9 muestran la inserción del gen *rpoB* en lugar del gen *pncA* dentro del pool. En los últimos carriles se muestra la combinación de amplicones (*rpoB*, *pncA*, *inhA* y *katG*), confirmando la correcta generación del pool. Las bandas bien definidas sugieren una amplificación

específica, mientras que la presencia de bandas adicionales podría corresponder a productos inespecíficos o fragmentos parciales de amplificación.

**Evaluación de cantidad mínima de ADN para la obtención de amplicones en muestras de cultivos primarios. extraídos por el método Kit.**

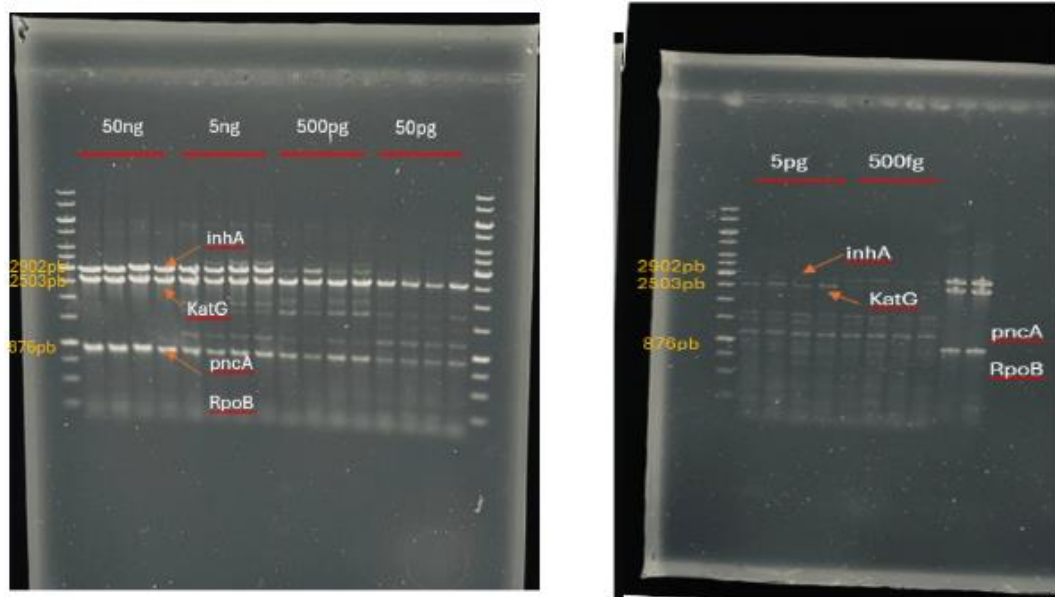


Figura 16. Evaluación del límite de detección de ADN para amplificación de regiones específicas.

Electroforesis en gel de agarosa que muestra la amplificación de los genes *inhA*, *katG*, *pncA* y *rpoB* a diferentes concentraciones de ADN. Se evaluaron cantidades de ADN desde 50 ng hasta 50 pg (izquierda) y desde 5 pg hasta 500 fg (derecha). Los fragmentos amplificados fueron separados en un gel de agarosa y comparados con un marcador de peso molecular (escalas indicadas en pb). Se observa que la amplificación es robusta hasta los 5ng de ADN mientras que en las concentraciones más bajas, la intensidad de las bandas disminuye.

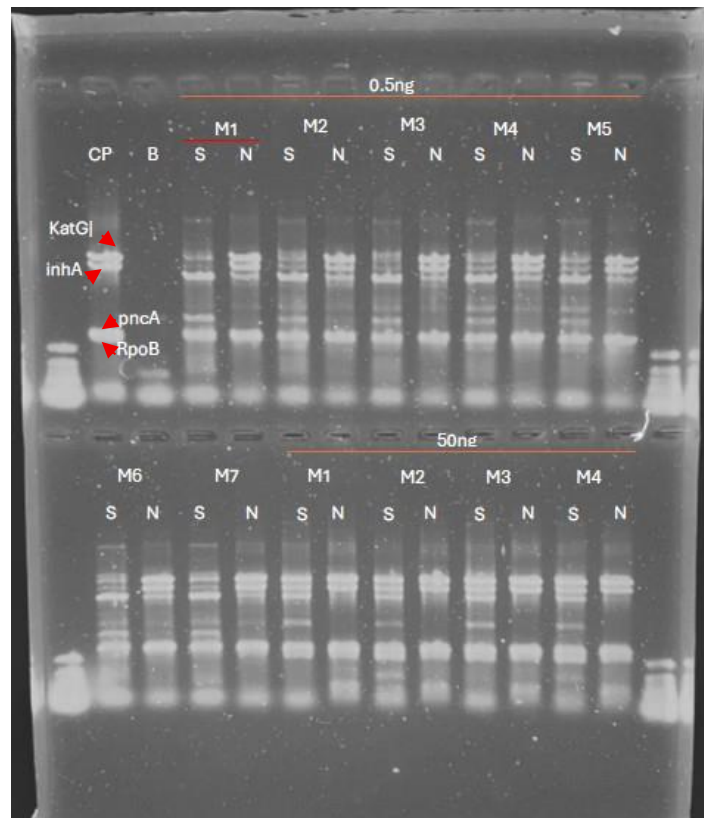


Figura 17. Obtención de los amplicones para los genes *inhA* (2503pb). *pncA* (876pb), *katG* (2902pb) y *rpoB* (856pb) obtenidos a partir de muestras tratadas con saponina o NALC/NaOH. M1, M2, M3, M4, M5, M6 y M7 corresponden al número de muestra, S: descontaminada con saponina; N: descontaminada con NALC/NaOH. Se evaluó la obtención de amplicones en muestras descontaminadas con saponina y NALC/NaOH.

En la figura 17 se puede apreciar los 4 genes de resistencia obtenidos en muestras tratadas con saponina al y NALC/NaOH, se observa además la presencia de bandas inespecíficas, que pueden corresponder a microorganismos de la flora oral, en las muestras tratadas con saponina a diferencia de las muestras tratadas con NALC/NaOH. Esta ausencia de inespecificidad en muestras tratadas con

NALC/NaOH muestran a este método como uno con mayor capacidad para eliminar material genético no correspondiente a *M. tuberculosis*, mostrándose como la adecuada en caso se busque trabajar solo con amplicones.

#### **9.1.5. Desempeño en secuenciación – comparación de kits de biblioteca utilizando ADN genómico.**

Se llevaron a cabo dos corridas de secuenciación Nanopore, una con bibliotecas *Native Ligation* y otra con bibliotecas *Rapid Barcoding*; ambos con ADN genómico y amplicones, llegando a secuenciar al mismo tiempo hasta 12 muestras.

La producción de datos (rendimiento) fue notablemente superior con la biblioteca de *Native Ligation*. En las muestras de esputo, el kit *Native Ligation* generó entre ~8.0 y 10.4 gigabases (Gb) por corrida (lecturas clasificadas como *Mycobacterium tuberculosis*), mientras que el kit *Rapid Barcoding* produjo entre ~1.2 y 6 Gb. Esto representa hasta ~9 veces más datos a favor del método de ligación. En las muestras de cultivo, la diferencia fue aún más marcada en el caso de *Native Ligation* alcanzó hasta ~27 Gb, versus ~10 Gb máximo con *Rapid* (Tabla 3 y 4).

En cuanto a la longitud de las lecturas, medida a través del valor N50 (longitud por debajo de la cual se encuentra el 50% de las bases secuenciadas), también fue mayor con *Native Ligation*. Para muestras de esputo, el N50 con *Native* tuvo un valor entre 1.4 kb y ~3.9 kb, mientras que con *Rapid* fueron valores menores a 2.6 kb; en cultivo, *Native* alcanzó un N50 ~8.2 kb vs ~4.9 kb con *Rapid*. Esto era esperable, ya que la biblioteca *Rapid* hace uso de transposasas, las cuales fragmentan el ADN lo que ocasiona un menor tamaño de estos. (Tabla 3 y 4)

La data clasificada también influyó en los datos útiles obtenidos. El porcentaje de lecturas clasificadas efectivamente como *M. tuberculosis* fue mayor con *Native Ligation*, en esputo se clasificaron hasta ~8.2 Gb con *Native* vs solo ~2.5 Gb con *Rapid*. En cultivo, *Native* clasificó ~22.7 Gb vs 0.046 Gb (46.47 Mb) con *Rapid*. Esto indica que el kit *Rapid* generó una proporción mucho mayor de lecturas no clasificadas (posiblemente por ser muy cortas o de menor calidad). En cambio, *Native* produce lecturas largas de mayor calidad que se alinean más claramente con el genoma de referencia. (Tabla 3 y 4)

Independientemente del tipo de ADN se observa que el kit *Native* genera fragmentos de mayor tamaño en comparación con el kit *Rapid* lo cual no generaría un gran impacto en el ensamblaje cuando se usa ADN genómico pero si en caso se esté secuenciando amplicones con un tamaño menor a 300 o 500 pb.

En cuanto al número total de lecturas generadas, el kit *Native* también superó a *Rapid*. En esputo, *Native* produjo hasta ~16.8 millones de lecturas, comparado con ~5.6 millones en *Rapid*; en cultivo, ~8.7 millones vs ~4.1 millones, respectivamente. Esto sugiere que, además de generar lecturas más largas, el protocolo *Native* aprovechó mejor los poros disponibles en la celda de flujo para secuenciar más moléculas. (Tabla 3 y 4).

Adicionalmente, al evaluar el rendimiento de las librerías génica *Native Ligation* usando ADN genómico, las tablas 5 y 6 muestran el bajo porcentaje de ADN perteneciente a *M. tuberculosis* que se obtuvo en las 12 muestras secuenciadas extraídas mediante el método fenol y kit. El número de lecturas de *M. tuberculosis*

apenas llegó, en el mejor de los casos, a un total de 9510 lo que porcentualmente no representaba ni el 1% del total de ADN dentro de la muestra.

Un aspecto importante de la secuenciación también es la calidad de las lecturas que se obtengan y sobre estos el filtro de q-score que se utilice para realizar los análisis bioinformáticos. En la figura 18 se observa que las lecturas obtenidas mediante la librería *Native Ligation* presentan un q-score cercano y/o superior a 20 lo que indica una precisión de 99% en las bases nucleotídicas mientras que *Rapid* presenta valores cercanos a 10 lo que indica una precisión del 90%. Esto muestra claramente que además de obtener suficiente data, esta debe presentar un valor de q-score que asegure la mayor tasa de fiabilidad respecto a las bases nucleotídicas para un adecuado análisis.

Al comparar el desempeño según el tipo de muestra y de ADN, se notó que el kit *Native* fue especialmente eficiente con ADN genómico de cultivos (el escenario ideal, con lecturas muy largas y muchas secuencias). Con ADN de esputo, su rendimiento disminuyó (menor rendimiento y tamaño de lecturas) pero aun así superó ampliamente a *Rapid*. En cambio, el kit *Rapid* mostró rendimientos relativamente similares entre esputo y cultivo en términos de datos, sugiriendo que el cuello de botella principal era el propio método de biblioteca más que la fuente de ADN.

En general, el protocolo *Native Ligation* demostró ventajas en cuanto a mayor cantidad de datos, lecturas más largas, mayor tasa de lecturas utilizables y menor error. Estos factores hacen que sea la opción preferida cuando se busca obtener la máxima información posible (por ejemplo, para ensamblar genomas completos o

detectar variaciones con confianza). No obstante, el kit *Rapid* tiene como punto a favor la rapidez y menor consumo de ADN de partida. En aplicaciones donde la cantidad de ADN es extremadamente limitada o se requiere un resultado en pocas horas, el método *Rapid* podría ser útil, aunque sacrificando información. En nuestro caso, para la detección de HR, finalmente optamos por priorizar la estrategia de amplicones con secuenciación *Native Ligation*, combinando así la sensibilidad de la PCR para subpoblaciones minoritarias con la calidad de la biblioteca de ligación

Tabla 3. Rendimiento de datos generados en la secuenciación de muestras de esputo utilizando dos tipos de kits de librerías génicas (*Rapid Barcoding* y *Native Ligation kit*) y diferentes tipos de ADN (amplicones y ADN genómico) (N= 12).

| Tipo de ADN         | Amplicones |        |        |        | ADN Genómico       |                  |                |                   |
|---------------------|------------|--------|--------|--------|--------------------|------------------|----------------|-------------------|
|                     | FENOL      |        | KIT    |        | FENOL              |                  | KIT            |                   |
| Kit de librería     | Rapid      | Native | Rapid  | Native | Rapid              | Native           | Rapid          | Native            |
| Rendimiento         | 1.18Gb     | 8.38Gb | 1.42Gb | 8.18Gb | 3.49Gb             | 7.97Gb           | 6.01Gb         | 10.4Gb            |
| N50 aprox           | 995b       | 1.4kb  | 1.16kb | 1.38kb | 2.66kb,<br>2.22kb* | 922b,<br>1.18kb* | 462b,<br>792b* | 3.9kb,<br>3.51kb* |
| Data clasificada    | 30Mb       | 7.58Gb | 20Mb   | 6.46Gb | 2.5Gb              | 5.67Gb           | 3.36Gb         | 8.2Gb             |
| Data no clasificada | 1.15Gb     | 800Mb  | 1.4Gb  | 1.72Gb | 996Mb              | 2.3Gb            | 2.65Gb         | 2.2Gb             |

|                    |       |       |       |       |        |       |       |        |
|--------------------|-------|-------|-------|-------|--------|-------|-------|--------|
| Total de lecturas  | 1.51M | 8.23M | 1.58M | 8.4M  | 2.62Mb | 8.31M | 5.65M | 11.68M |
| Lecturas aprobadas | 1.29M | 6.91M | 1.39M | 6.79M | 2.22Mb | 6.81M | 5M    | 10.95M |

Tabla 4. Rendimiento y datos generados en la secuenciación de muestras de cultivo utilizando dos tipos de kits de librerías génicas (*Rapid Barcoding* y *Native Ligation kit*) y diferentes tipos de ADN (amplicones y ADN genómico). (N = 12).

| Tipo de ADN         | Amplicones |         |         |         | ADN Genómico |          |         |         |
|---------------------|------------|---------|---------|---------|--------------|----------|---------|---------|
|                     | Fenol      |         | KIT     |         | Fenol        |          | KIT     |         |
| Kit de librería     | Rapid      | Native  | Rapid   | Native  | Rapid        | Native   | Rapid   | Native  |
| Rendimiento         | 2.05Gb     | 4.52Gb  | 29.18Mb | 8.16Gb  | 167.97Mb     | 26.97 Gb | 10.41Gb | 18.95Gb |
| N50 Aprox           | 1.27kb     | 1.47Kb  | 1.09kb  | 1.65Kb  | 4.9kb        | 8.25kb   | 5.23kb  | 4.91kb  |
| Data clasificada    | 10Mb       | 4.07Gb  | 6kb     | 7.27Gb  | 46.47Mb      | 22.77Gb  | 8.88Gb  | 17.09Gb |
| Data no clasificada | 2.05Gb     | 512.2Mb | 29.12Mb | 888.9Mb | 121.5Mb      | 4.2Gb    | 1.53Gb  | 1.86Gb  |
| Total de lecturas   | 2.11M      | 3.73M   | 33.67k  | 5.39M   | 54.34k       | 5.55M    | 4.15M   | 8.73M   |
| Lecturas aprobadas  | 1.85M      | 3.39M   | 29.81k  | 4.91M   | 31.06k       | 5.05M    | 3.72M   | 8.07M   |

Tabla 5. Lecturas obtenidas de ADN genómico de muestras descontaminadas con saponina al 5%, extraídas por fenol y secuenciada con *Native Ligation Kit* (N = 12).

| Barcode | BAAR | Cantidad de ADN (ng) | %r lecturas mapeados de TB | Número de lecturas |
|---------|------|----------------------|----------------------------|--------------------|
| 1       | ++   | 17100                | 0.34                       | 3889               |
| 2       | ++   | 5900                 | 0.34                       | 5375               |
| 3       | ++   | 5900                 | 0.32                       | 4999               |
| 4       | +++  | 1720                 | 0.36                       | 3375               |
| 5       | ++   | 4140                 | 0.05                       | 1661               |
| 6       | ++   | 1760                 | 0.47                       | 6807               |
| 7       | ++   | 5860                 | 0.21                       | 6778               |
| 8       | +    | 1374                 | 0.48                       | 4913               |
| 9       | +++  | 3600                 | 0.97                       | 9510               |
| 10      | +++  | 1872                 | 0.8                        | 5462               |
| 11      | +++  | 912                  | 0.61                       | 7161               |
| 12      | +++  | 966                  | 0.58                       | 6356               |

Tabla 6. Lecturas obtenidas de ADN genómico de muestras extraídas por Kit y secuenciada con la librería *Native Ligation Kit* descontaminadas con saponina al 5% (N = 12).

| Barcode | BAAR | Cantidad de ADN (ng) | % lecturas mapeadas de TB | Número de lecturas total |
|---------|------|----------------------|---------------------------|--------------------------|
| 1       | ++   | 5050                 | 0.58                      | 1614                     |

|   |     |      |      |      |
|---|-----|------|------|------|
| 2 | ++  | 6700 | 0.3  | 1585 |
| 3 | ++  | 6700 | 0.28 | 1056 |
| 4 | +++ | 2260 | 0.11 | 905  |
| 5 | ++  | 6000 | 0.06 | 787  |
| 6 | ++  | 1440 | 0.07 | 1351 |
| 7 | ++  | 2418 | 0.15 | 1812 |
| 8 | +   | 2274 | 0.08 | 1779 |
| 9 | +++ | 3060 | 0.2  | 5048 |

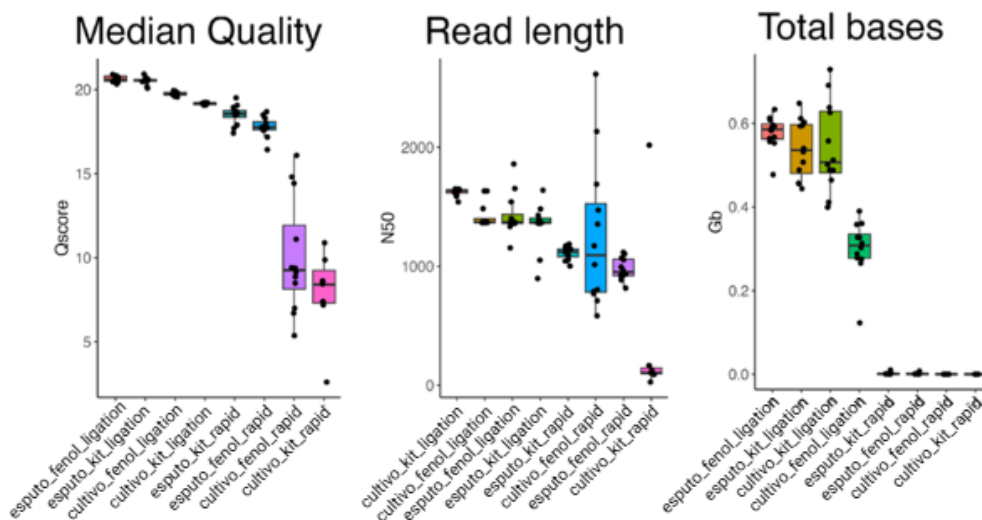


Figura 18. Evaluación de la calidad media, tamaño de N50 y total de bases obtenidas en la secuenciación de 12 muestras procesadas mediante 2 métodos de extracción y dos tipos de librería génica

Se compararon los métodos de extracción Fenol y Kit, además del tipo de muestra inicial, esputo o cultivo, analizando métricas clave de la secuenciación, como calidad media (Phred score), tamaño de N50 y total de bases generadas. Se presentan los resultados para las librerías génicas *Rapid Barcoding* y *Native Ligation*).

#### **9.1.6. Desempeño en secuenciación – comparación de kits de biblioteca utilizando amplicones (14 genes).**

La secuenciación de amplicones se hizo posteriormente al de ADN genómico, cuando se utilizó el kit *Native Ligation* se produjo un volumen de datos notablemente mayor que la preparada con el kit *Rapid*. Específicamente, el protocolo *Native Ligation* rindió en promedio ~8,3 gigabases de secuencia por corrida, en comparación con solo ~1,2–1,4 Gb obtenidos con *Rapid Barcoding* bajo las mismas condiciones. Esto correspondió a aproximadamente 8 millones de lecturas generadas con la librería de ligación, frente a apenas ~1,5 millones de lecturas con la librería *Rapid* (Tabla 10).

Consecuentemente, la cobertura promedio de los amplicones fue muy superior con la biblioteca de ligación. La profundidad de secuenciación alcanzó valores cercanos >1000× con *Native Ligation*, mientras que con *Rapid Barcoding* la cobertura media se mantuvo por debajo de 30× (Tabla 7) (figura 19 y 20). Esta diferencia clara indica que el kit *Rapid* no logró secuenciar los amplicones en suficiente profundidad, potencialmente limitando la sensibilidad para detectar variantes en muestras con baja proporción de ADN blanco. Además, se observa que gran parte de los datos generados con *Rapid Barcoding* correspondieron a lecturas no clasificadas,

sugiriendo una fracción considerable de secuencias espurias o fragmentos truncados. En las corridas con 14 amplicones, más del 90% de las lecturas secuenciadas con la librería de ligación pudieron asignarse a *M. tuberculosis*, mientras que en la librería rápida la proporción fue mínima (p. ej., solo ~30 Mb clasificaron como *M. tuberculosis* frente a >1 Gb de datos no clasificables).

En cuanto a la calidad de las lecturas, las bibliotecas por ligación también mostraron una ventaja notable. El Q-score promedio de las lecturas con *Native Ligation* se encontró en un rango de ~18–21, cercano o superior al umbral Q20 (indicativo de 99% de precisión por lectura), mientras que las lecturas del kit *Rapid* presentaron calidades inferiores, con Q-scores típicamente entre 9 y 16 (Tabla 9). Esto implica que una fracción sustancial de las lecturas de la librería de ligación alcanzó alta fidelidad (muchas superando Q20), mientras que casi ninguna lectura de la librería rápida logró exceder Q20. En la práctica, se estimó que menos del 5% de las lecturas generadas con *Rapid Barcoding* superaron un Q-score de 20, frente a una proporción considerablemente mayor en la condición de ligación nativa (más del 50%, dependiendo de la muestra, fueron Q>20). Esta disparidad refleja las diferencias en química de secuenciación: el protocolo de ligación aprovecha plenamente la última química Q20+ de ONT, mientras que el método *Rapid*, si bien utiliza la misma química v14, podría estar penalizado por la longitud más corta de los fragmentos y la menor eficiencia de incorporación de adaptadores, resultando en promedios de calidad más bajos.

## Profundidad Media( Ligation)

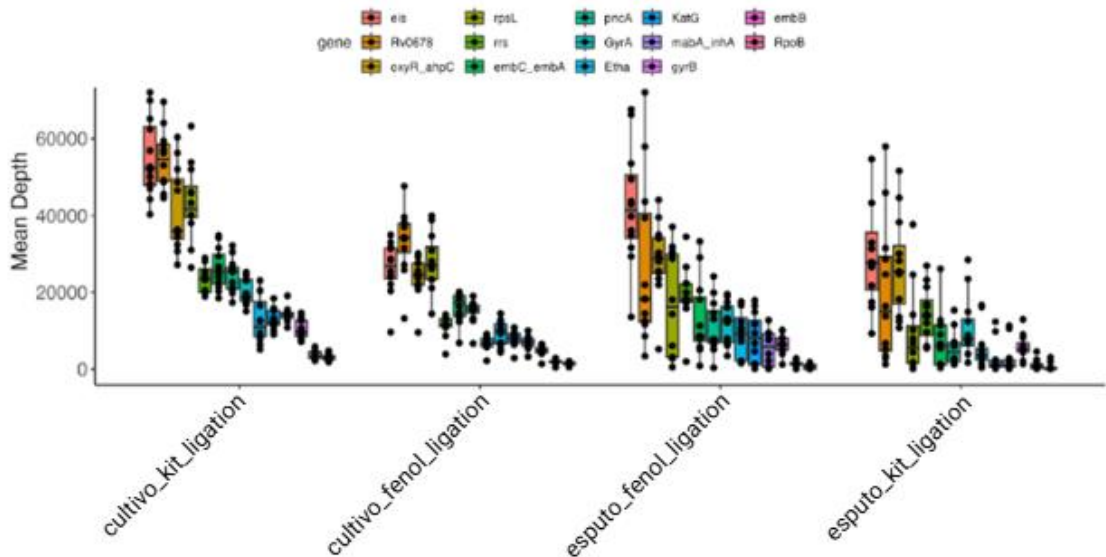


Figura 19. Profundidad media de secuenciación mediante el kit *Native Ligation* para distintos genes asociados a resistencia bajo diferentes condiciones de extracción y origen de la muestra (N = 12).

En la figura 19, el eje Y se representa la profundidad media de cobertura, mientras que en el eje X se agrupan las condiciones experimentales, muestras provenientes de cultivo o esputo y extraídas mediante los métodos Kit y Fenol. Cada color representa un gen específico según la leyenda superior. Los puntos individuales reflejan la variabilidad entre muestras, y las barras de dispersión indican la distribución de la profundidad de secuenciación en 12 muestras.

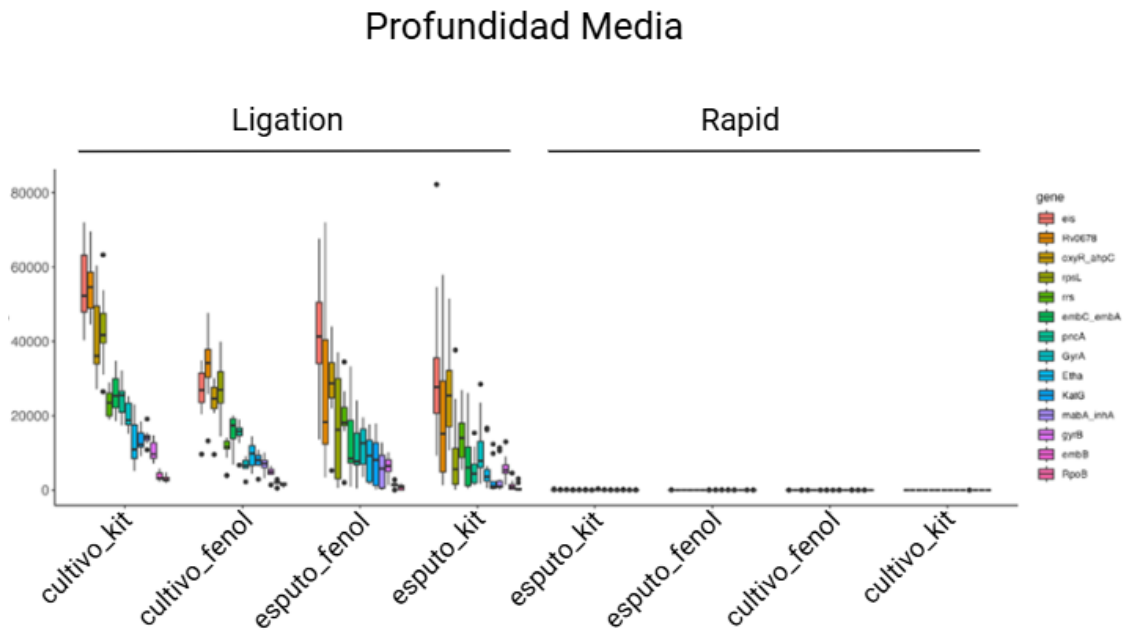


Figura 20. Comparación de profundidad entre el kit *Native Ligation* y *Rapid barcoding* (N= 12).

En esta figura se incluyen datos generados mediante el kit de librería *Rapid barcoding*, permitiendo comparar su desempeño frente a la librería *Native Ligation* tanto para muestras de esputo y cultivo extraídos por los métodos de extracción Kit y Fenol. Cada color representa un gen específico según la leyenda superior y los puntos individuales reflejan la variabilidad en la cobertura entre muestras.

Tabla 7. Valores de profundidad de secuenciación de amplicones extraídos por kit y fenol (N = 12).

| Characteristic | culture_miniprep_ligation, N = | culture_miniprep_rapid, N = | culture_fenol_ligation, N = | culture_fenol_rapid, N = | sputum_fenol_ligation, N = | sputum_fenol_rapid, N = | sputum_miniprep_ligation, N = | sputum_rapid_miniprep, N = |
|----------------|--------------------------------|-----------------------------|-----------------------------|--------------------------|----------------------------|-------------------------|-------------------------------|----------------------------|
|                | 12'                            | 7'                          | 12'                         | 12'                      | 12'                        | 12'                     | 12'                           | 12'                        |
| Rv0678         | 54,616 (48,901, 58,612)        | 0 (0, 0)                    | 34,176 (30,391, 37,879)     | 0 (0, 1)                 | 18,276 (12,356, 40,428)    | 6 (3, 12)               | 15,104 (4,825, 29,333)        | 7 (1, 24)                  |
| eis            | 52,278 (47,916, 63,119)        | 0 (0, 0)                    | 26,911 (23,570, 31,494)     | 0 (0, 1)                 | 41,336 (34,039, 50,501)    | 15 (6, 24)              | 27,719 (20,683, 35,633)       | 24 (11, 42)                |
| rpsL           | 41,706 (39,494, 47,564)        | 0 (0, 0)                    | 26,986 (23,402, 31,823)     | 0 (0, 1)                 | 16,222 (3,072, 30,010)     | 7 (4, 10)               | 5,608 (1,553, 11,207)         | 4 (0, 13)                  |
| oxyR_ahpC      | 36,079 (33,922, 49,487)        | 0 (0, 0)                    | 24,601 (21,901, 27,600)     | 0 (0, 0)                 | 28,728 (24,813, 34,273)    | 18 (7, 34)              | 25,414 (17,038, 32,104)       | 17 (11, 37)                |
| embC_embA      | 25,321 (22,170, 29,955)        | 0 (0, 0)                    | 17,410 (13,799, 19,158)     | 0 (0, 1)                 | 8,459 (7,313, 18,730)      | 9 (2, 13)               | 6,065 (1,136, 11,513)         | 6 (1, 13)                  |
| rrs            | 23,521 (19,964, 26,044)        | 0 (0, 0)                    | 11,639 (11,048, 13,176)     | 1 (0, 1)                 | 18,075 (17,397, 22,230)    | 15 (5, 23)              | 14,005 (8,740, 18,027)        | 14 (9, 34)                 |
| GyrA           | 18,747 (17,624, 23,371)        | 0 (0, 0)                    | 6,611 (5,843, 7,961)        | 0 (0, 0)                 | 12,635 (6,936, 16,392)     | 11 (6, 21)              | 7,848 (6,106, 13,099)         | 20 (10, 38)                |
| pncA           | 25,489 (20,913, 26,577)        | 0 (0, 0)                    | 15,729 (14,667, 16,676)     | 0 (0, 0)                 | 7,736 (6,848, 15,254)      | 7 (3, 10)               | 4,383 (1,885, 6,949)          | 2 (1, 7)                   |
| mabA_inhA      | 14,114 (13,376, 14,701)        | 0 (0, 0)                    | 7,116 (6,053, 7,970)        | 0 (0, 0)                 | 5,788 (543, 9,322)         | 3 (2, 8)                | 911 (627, 2,407)              | 2 (1, 7)                   |
| KatG           | 12,265 (11,536, 15,346)        | 0 (0, 0)                    | 8,056 (6,776, 9,333)        | 0 (0, 0)                 | 8,090 (1,196, 12,636)      | 7 (3, 16)               | 972 (528, 2,262)              | 3 (1, 10)                  |
| gyrB           | 9,674 (8,524, 12,689)          | 0 (0, 0)                    | 4,832 (4,276, 5,547)        | 0 (0, 0)                 | 6,445 (4,982, 8,100)       | 11 (8, 23)              | 5,299 (4,525, 6,813)          | 18 (10, 36)                |
| EthA           | 10,923 (8,434, 17,533)         | 0 (0, 0)                    | 9,845 (6,610, 11,873)       | 0 (0, 0)                 | 9,235 (2,137, 13,414)      | 10 (6, 22)              | 3,611 (2,410, 5,428)          | 8 (3, 19)                  |
| RpoB           | 2,758 (2,437, 3,428)           | 0 (0, 0)                    | 1,434 (1,198, 1,852)        | 0 (0, 1)                 | 797 (58, 1,246)            | 3 (1, 8)                | 156 (121, 445)                | 1 (0, 6)                   |
| embB           | 3,162 (2,910, 4,582)           | 0 (0, 0)                    | 1,952 (1,819, 2,173)        | 0 (0, 1)                 | 1,470 (1,154, 1,850)       | 6 (3, 11)               | 784 (256, 1,376)              | 6 (2, 12)                  |

<sup>1</sup>Median (IQR)

### Relación entre profundidad obtenida en base al tamaño del amplicon

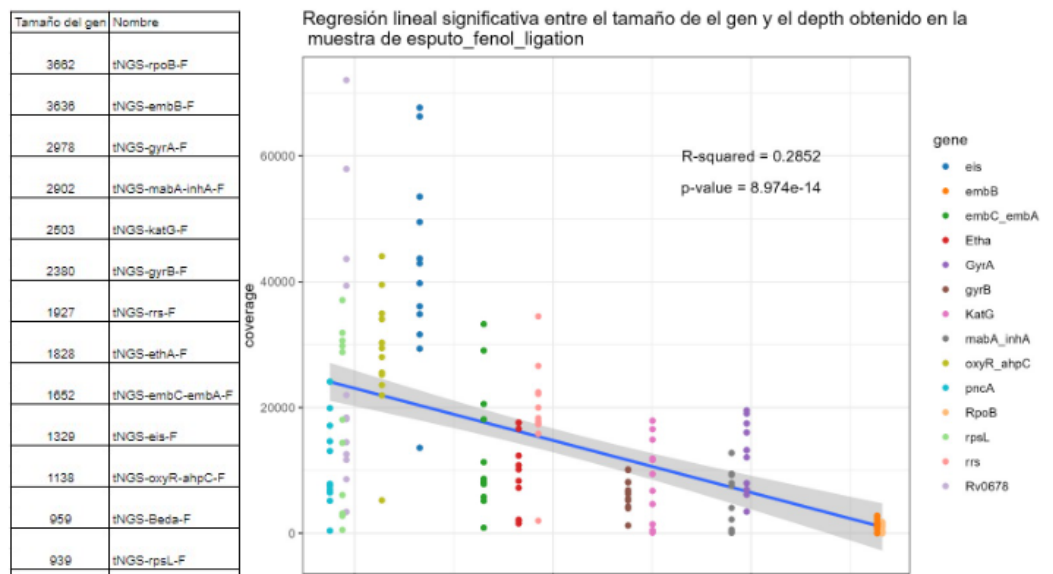


Figura 21. Relación entre la profundidad de secuenciación y el tamaño del amplicon (N = 12).

Se presenta un análisis de regresión lineal, donde el eje Y muestra la cobertura obtenida para cada gen y el eje X representa el tamaño del amplicón en pares de bases. La tendencia negativa observada indica una correlación significativa entre el tamaño del amplicón y la profundidad de secuenciación ( $R^2 = 0.2852$ ,  $p < 0.001$ ), sugiriendo que los amplicones de menor tamaño tienden a obtener mayor cobertura

en esta condición experimental. La tabla a la izquierda muestra los genes analizados junto con sus respectivos tamaños en pares de bases.

Tabla 8. Comparación de métodos de extracción de ADN.

|                                     | Fenol             | Kit             |
|-------------------------------------|-------------------|-----------------|
| Tiempo de procesamiento por muestra | 20 min/muestra    | 15 min/muestra  |
| Costo por muestra                   | S/. 21.6          | S/. 13.5        |
| Nivel de complejidad                | Elevado           | Bajo            |
| Cantidad de ADN total               | <1.41ug - 58.4ug> | <1.4ug - 8.7ug> |
| Tamaño de fragmentos                | <9kb - 10kb>      | <14kb - 17kb>   |
| DIN                                 | <3.4 - 6.1>       | <1.9 - 7.3>     |

Tabla 9. Comparación de kits de librería génica

|                       | <i>Native Ligation</i> |            | <i>Rapid Barcoding</i> |            |
|-----------------------|------------------------|------------|------------------------|------------|
| Tiempo de proceso     | 150 min/muestra        |            | 90 min/muestra         |            |
| Costo por 12 muestras | S/. 5197               |            | S/.3981                |            |
| Complejidad           | Elevado                |            | Bajo                   |            |
| Q-score               | <18-21>                |            | <9 -16>                |            |
|                       | ADN<br>genómico        | Amplicones | ADN<br>genómico        | Amplicones |

|             |     |        |     |      |
|-------------|-----|--------|-----|------|
| Profundidad | <1x | >1000x | <1x | <30x |
|-------------|-----|--------|-----|------|

Tabla 10. Rendimiento y data generada en la secuenciación con muestras de esputo.

| Tipo de ADN          | Amplicones  |              |             |              | ADN Genómico   |                |           |                |
|----------------------|-------------|--------------|-------------|--------------|----------------|----------------|-----------|----------------|
| Método de extracción | FENOL       |              | KIT         |              | FENOL          |                | KIT       |                |
| Kit de librería      | Rapid       | Ligatio<br>n | Rapid       | Ligatio<br>n | Rapid          | Ligation       | Rapid     | Ligation       |
| Rendimiento (Gb)     | 1.18        | 8.38         | 1.42        | 8.18         | 3.49           | 7.97           | 10.4      | 6.01           |
| N50 Aprox (kb)       | 0.99        | 1.4          | 1.16        | 1.38         | 2.66,<br>2.22* | 922b,<br>1.18* | 3.6, 3.5* | 0.46,<br>0.79* |
| Data clasificada     | 30Mb        | 7.58Gb       | 20Mb        | 6.46Gb       | 2.5Gb          | 5.67Gb         | 8.2Gb     | 3.36Gb         |
| Data no clasificada  | 1.15Gb      | 800Mb        | 1.4Gb       | 1.72Gb       | 996Mb          | 2.3Gb          | 2.2Gb     | 2.65Gb         |
| Total lecturas       | 1.51M       | 8.23M        | 1.58M       | 8.4M         | 2.62M<br>b     | 21.68M         | 5.65M     | 16.81M         |
| Lecturas aprobadas   | 1.29M       | 6.91M        | 1.39M       | 6.79M        | 2.22M<br>b     | 20.95M         | 5M        | 18.31M         |
| Lecturas rechazadas  | 226.5<br>9k | 683k         | 188.3<br>4k | 840.5k       | 42k            | 1.7M           | 584.2k    | 2.07M          |

\* Se presentan 2 valores debido a que la secuenciación se realizó en 2 etapas por la falta de espacio para almacenar los datos en el equipo.

### **9.1.7. Desempeño en secuenciación – comparación de kits de biblioteca utilizando amplicones (4 genes).**

Cuando se evaluó los kits de librería génica utilizando solo 4 amplicones se mantuvo la tendencia general observada con 14 amplicones, aunque con diferencias menos extremas. En este ensayo, ambos kits se compararon bajo condiciones idénticas (1200 nanoporos activos y 6 horas de secuenciación). Nuevamente, el protocolo de *Native Ligation* proporcionó una mayor profundidad de lectura en todos los genes objetivo en comparación con el kit *Rapid*. La cobertura mediana alcanzada por amplicón fue consistentemente superior con *Native Ligation*, destacando especialmente en los genes *katG* e *inhA* donde la brecha de profundidad entre métodos fue más pronunciada (figura 23). En valores absolutos, el kit de *Native Ligation* logró medianas de cobertura del orden de varios cientos a miles de lecturas por base en cada uno de los 4 amplicones, mientras que el kit *Rapid barcoding* se quedó en coberturas medianas menores.

Estos resultados indican que el kit *Rapid barcoding* presentó rendimiento menos consistente entre muestras: algunas muestras obtuvieron coberturas moderadas, mientras otras quedaron cercanas al límite inferior aceptable (1000x), produciendo resultados heterogéneos. Por el contrario, el kit *Native Ligation* mostró resultados más uniformes, con coberturas altas en prácticamente todas las muestras para cada uno de los cuatro genes (figura 23).

En términos de rendimiento total, con solo 4 amplicones en el pool, ambos métodos lograron elevar la profundidad respecto a cuando se usa un pool de 14 amplicones. Sin embargo, el método *Native Ligation* mantuvo una ventaja clara en cantidad de

lecturas útiles generadas. Por ejemplo, en la comparación directa del gen *katG*, la librería de ligación alcanzó coberturas que superaban en varias veces a las obtenidas por la librería *Rapid*, pese a haber partido de la misma cantidad de ADN de entrada y haberse secuenciado por igual tiempo. Este resultado confirma que la química de ligación aprovecha más eficientemente la capacidad del flujo de nanoporos produciendo mayor rendimiento por unidad de tiempo que la química rápida.

En resumen, incluso en un panel reducido, el kit de *Native Ligation* aseguró coberturas ampliamente suficientes en todos los genes (factor crítico para una posterior confiabilidad en análisis de variantes), mientras que el kit *Rapid* quedó más cerca del umbral mínimo de secuenciación (1000X) en uno o más blancos, lo cual podría ser un limitante en aplicaciones críticas.

No obstante, cabe señalar que la brecha de rendimiento entre kits se redujo ligeramente al secuenciar menos amplicones. En este escenario de 4 genes, el número total de lecturas generadas por *Rapid barcoding* representó una fracción mayor del obtenido con *Ligation* que en el panel de 14. Esto sugiere que, cuando la demanda de secuenciación por muestra es baja (pocos genes en la muestra), el protocolo *Rapid* puede proveer datos suficientes en ciertos casos. Aun así, para garantizar coberturas profundas en todos los blancos y minimizar la variabilidad, la estrategia de ligación demostró ser más confiable y con mayor margen de seguridad técnica.

| Rapid                                |                                 | Ligación Kit                         |                                   |
|--------------------------------------|---------------------------------|--------------------------------------|-----------------------------------|
| N° de poros: 1200<br>Tiempo: 6 horas |                                 | N° de poros: 1200<br>Tiempo: 6 horas |                                   |
| Characteristic                       | amplicones, N = 12 <sup>1</sup> | Characteristic                       | amplicones<br>N = 12 <sup>1</sup> |
| inhA                                 | 3,095 (1,931, 5,063)            | katG                                 | 6,242 (5,260, 7,843)              |
| fabG1                                | 3,078 (1,919, 4,960)            | fabG1                                | 6,648 (5,330, 7,926)              |
| katG                                 | 2,815 (1,713, 4,454)            | inhA                                 | 6,495 (5,223, 7,711)              |
| rpoB                                 | 2,127 (1,249, 3,317)            | pncA                                 | 4,940 (4,077, 5,689)              |
| pncA                                 | 1,187 (637, 1,541)              | rpoB                                 | 1,420 (1,144, 1,706)              |
| <sup>1</sup> Median (IQR)            |                                 | <sup>1</sup> Median (Q1, Q3)         |                                   |

Figura 22. Comparación de la profundidad media de secuenciación obtenida con los kits de librería génica *Rapid barcoding* (Rapid) y Native barcoding (Ligation kit) utilizando el nuevo pool de cuatro amplicones (N = 12).

Se presentan los valores medianos de profundidad (IQR) para cada amplicón en ambos métodos. En ambas condiciones se utilizó el mismo número de poros (1200) y un tiempo de secuenciación de 6 horas. Se observa que el Kit *Native Ligation* obtuvo una mayor profundidad media en todos los amplicones, especialmente en *katG*, e *inhA*, mientras que en el Kit *Rapid barcoding* la profundidad fue menor y más variable.

### Comparación del ligación kit vs Rapid

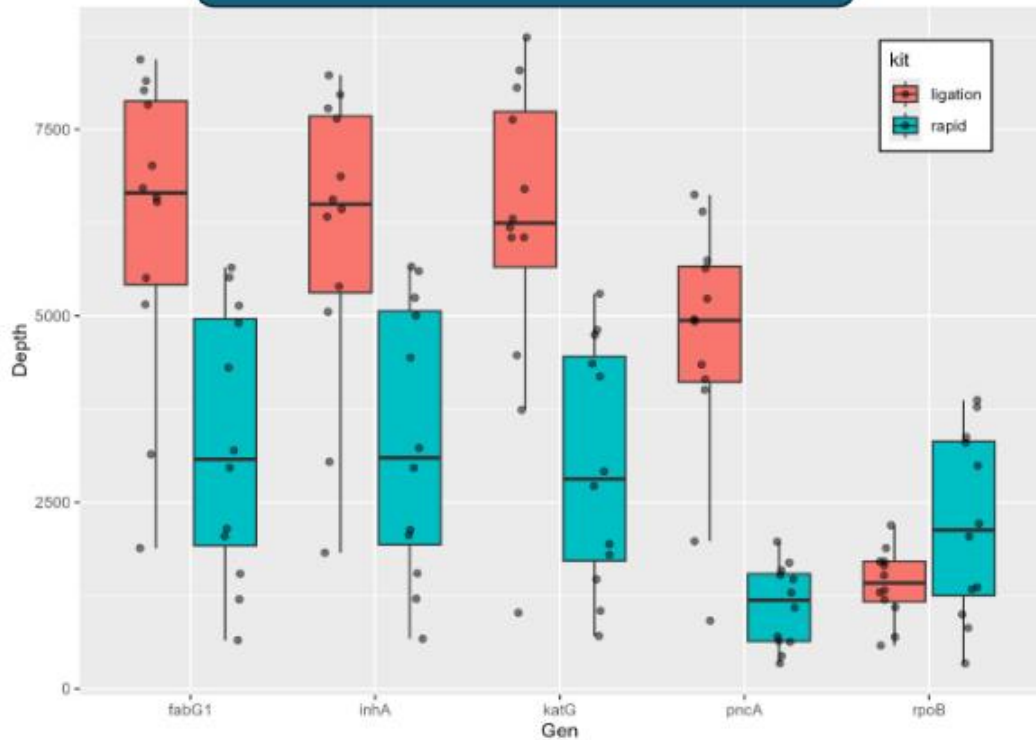


Figura 23. Comparativa entre el kit *Native Ligation* y *Rapid Barcoding* con 4 amplicones (N= 12). Los datos están representados mediante diagramas de caja, donde el color rojo indica el *Native Ligation* kit y el color azul el *Rapid barcoding*. cada punto representa a una muestra y el valor de profundidad para el gen alcanzado.

### 9.2.1. Determinación del límite de detección de heterorresistencia en mezclas bacterianas mediante secuenciación y método de proporciones.

El principal objetivo del estudio fue comparar el límite de detección de HR mediante secuenciación Nanopore frente al método de proporciones, por lo que en la Tabla 13 y figura 28 se resume, para cada mezcla preparada, la proporción real de cepa resistente, la proporción detectada por secuenciación y la detectada por cultivo en agar.

### 9.2.2. Cultivo de cepas y preparación de mezclas heterorresistentes.

En la figura 24 se muestra los cultivos de cepas de *M. tuberculosis* sensible (H37Rv) y resistente (DM97) que se utilizaron para generar una mezcla HR y que se obtuvieron luego de un periodo de incubación de 21 días.

### CULTIVOS DE CEPAS

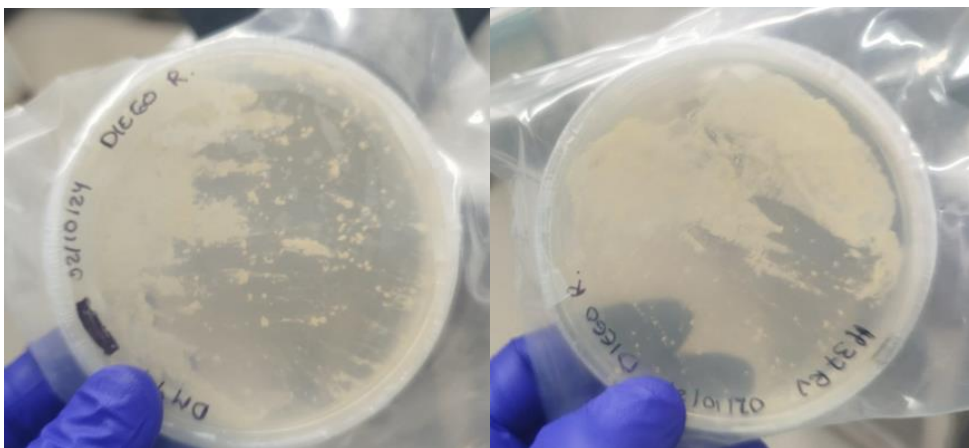


Figura 24. Cultivos de cepas H37Rv y DM97 en medios de cultivo 7H10 con un tiempo de incubación de 3 semanas.

Con los cultivos crecidos se preparó las mezclas heterorresistentes y las alícuotas destinadas al proceso de secuenciación fueron sometidas a la extracción de ADN mostrando cantidades que no pudieron ser cuantificadas mediante el Qubit en algunos casos debido a su baja concentración de ADN (Tabla 11).

### 9.2.3. Extracción de ADN de las mezclas heterorresistentes.

La extracción de ADN de las mezclas heterorresistentes se realizó mediante el método Kit, para lo cual se generó un grupo de mezcla bacteriana con y sin esputo.

Tabla 11. Cuantificación de ADN en mezclas bacterianas con y sin esputo. Las mezclas están representadas por su proporción de cepa sensible (S) a resistente (R), desde 100:0 (M1) hasta 0:100 (M7). Se observa variabilidad en la cuantificación del ADN entre las distintas condiciones, con diferencias en la detección dependiendo de la presencia de esputo en la muestra.

| Proporción<br>mezclas (S:R) | Mezclas bacterianas sin esputo |                 |                 | Mezclas bacterianas con esputo |                 |                 |
|-----------------------------|--------------------------------|-----------------|-----------------|--------------------------------|-----------------|-----------------|
|                             | Rep1<br>(ng/μl)                | Rep2<br>(ng/μl) | Rep3<br>(ng/μl) | Rep1<br>(ng/μl)                | Rep2<br>(ng/μl) | Rep3<br>(ng/μl) |
| 100:0 (M1)                  | 1.54                           | too low         | 1.56            | 3.28                           | 5.88            | 9.96            |
| 99:1 (M2)                   | 2.34                           | 0.11            | 3.25            | 3.88                           | 4.60            | 8.80            |
| 95:5 (M3)                   | too low                        | 0.15            | 7.644           | 3.88                           | 6               | 10.7            |
| 90:10 (M4)                  | 8.21                           | 0.15            | too low         | 3.98                           | 6.92            | 10.2            |

|            |      |         |         |      |      |      |
|------------|------|---------|---------|------|------|------|
| 75:25 (M5) | 5.17 | too low | too low | 4.94 | 4.5  | 9    |
| 50:50 (M6) | 3.98 | too low | 4.27    | 4.38 | 5.84 | 10.6 |
| 0:100 (M7) | 5.3  | too low | 4.65    | 8.18 | 5.12 | 15.3 |

Rep: Repetición

#### **9.2.4. Obtención de amplicones de mezclas heterorresistentes con y sin esputo.**

Posteriormente, y aunque se contaba con una cantidad de ADN no detectable en algunos casos se logró obtener los amplicones para los 4 genes de resistencia (*inhA*, *katG*, *pncA* y *rpoB*) en todas las muestras de mezclas heterorresistentes preparadas (figura 25 y 26).

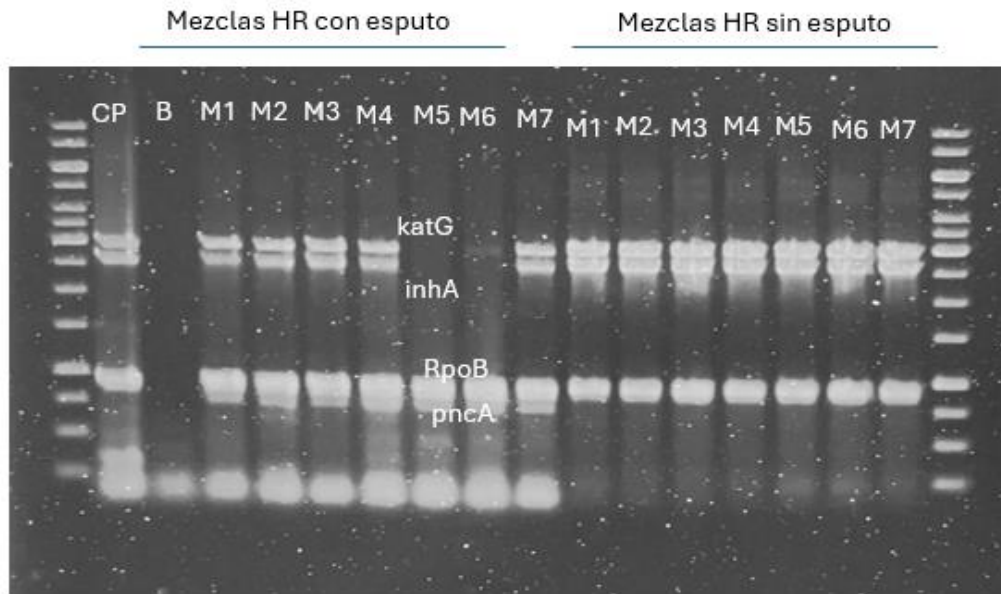


Figura 25. Electroforesis en gel de agarosa para la detección de genes en mezclas bacterianas con y sin esputo.

La figura 25 es una imagen representativa de la amplificación de los genes *katG*, *inhA*, *rpoB* y *pncA* en distintas proporciones de mezclas bacterianas con esputo (izquierda) y sin esputo (derecha). Las muestras corresponden a diferentes proporciones de cepas resistentes y sensibles (HR) de *M. tuberculosis*, etiquetadas de M1 a M7. Se observa variabilidad en la intensidad de las bandas, lo que sugiere diferencias en la detección y amplificación en presencia o ausencia de esputo. Se incluyen controles positivos (CP) y negativos (B).

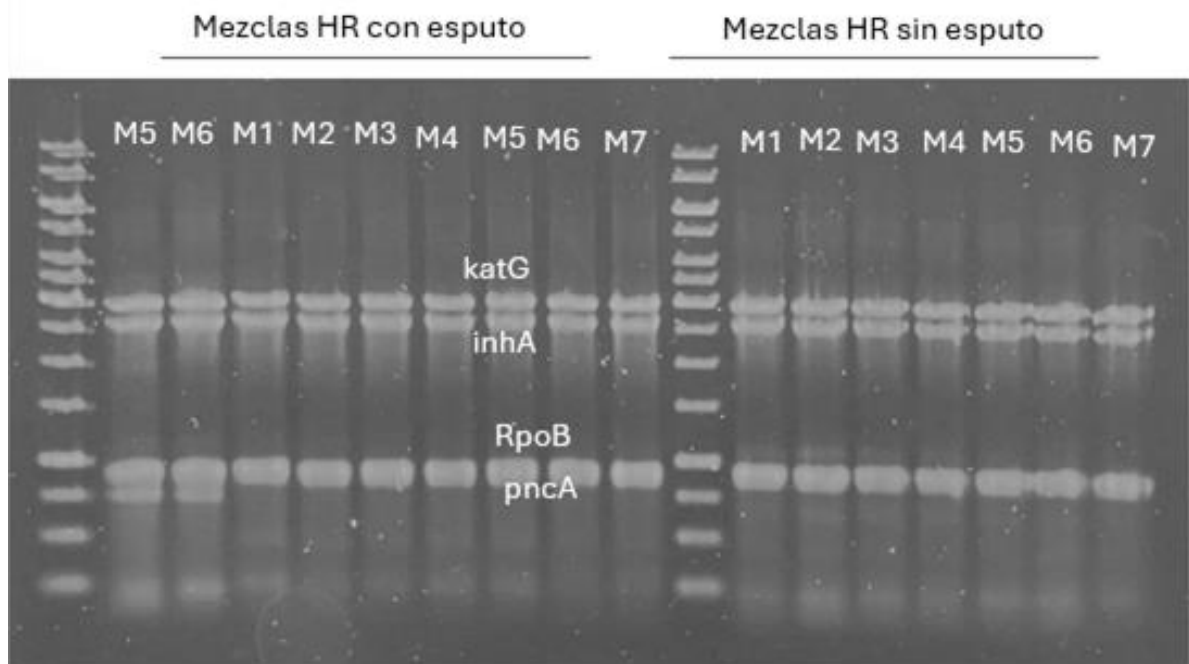


Figura 26. Electroforesis en gel de agarosa de la amplificación de genes en mezclas bacterianas heterorresistentes de *M. tuberculosis* con y sin esputo en la repetición 2. Los carriles 1 y 2 corresponden a las mezclas M5 y M6 de la primera repetición.

#### 9.2.5. Detección de heterorresistencia mediante el método de proporciones en Agar.

Paralelamente, las mezclas preparadas para la detección de HR mediante el método de proporciones fueron sembradas en placas 7H11-OADC y se contabilizó el número de colonias a partir de los 21 días de cultivo como se muestra en la figura 27. En el caso de rifampicina, por ejemplo, el método de proporciones en agar, como era de esperarse, logró detectar crecimiento resistente (colonias en medio con rifampicina) en todas las mezclas donde la cepa resistente representaba  $\geq 1\%$  (mezclas M2 en adelante). En la mezcla M2 (1% resistente), el MPA también evidenció unas colonias en el medio con rifampicina, dando una proporción

calculada de ~14% (lo cual técnicamente ya se interpreta como resistencia presente). Para isoniazida y pirazinamida, se observaron resultados similares, donde el agar identificó resistencia incluso en las mezclas 1% (aunque en PZA no fue posible dado que requiere de valores de pH específicos que pudieron sufrir variación) (Tabla 13).

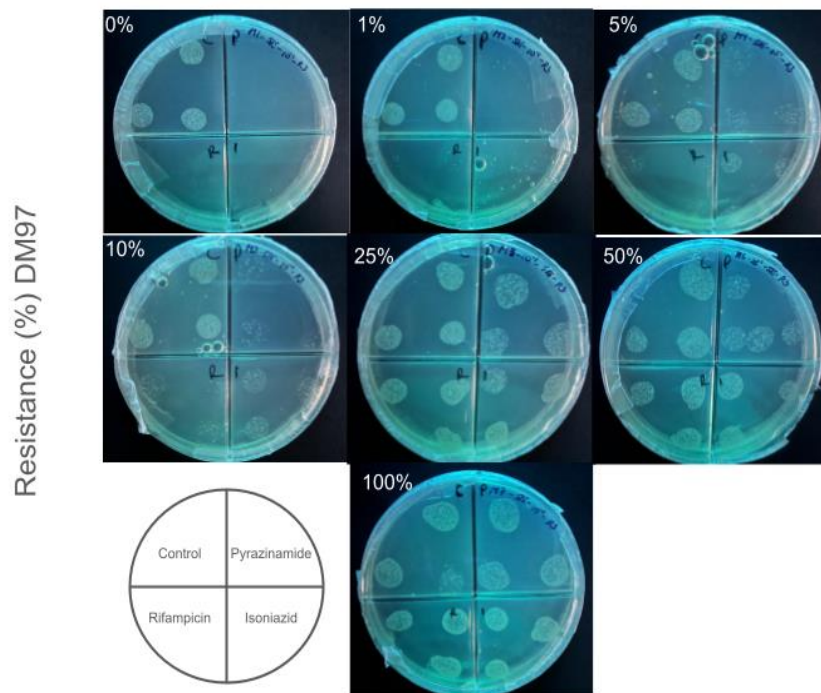


Figura 27. Método de proporción en agar para evaluar la resistencia a rifampicina, isoniacida y pirazinamida. Se emplearon siete mezclas de las cepas H37Rv y DM97 con el fin de determinar el porcentaje de resistencia de *M. tuberculosis*.

Tabla 13. El IC95 de los porcentajes de resistencia. Los valores fueron calculados en cada cuadrante corresponde a los valores obtenidos tras las tres repeticiones de cultivos realizados con mezclas heterorresistentes en ausencia de esputo

| (S/R) | Isoniacida (%) | Rifampicina (%) | Pirazinamida (%) |
|-------|----------------|-----------------|------------------|
|-------|----------------|-----------------|------------------|

|          |              |                |              |
|----------|--------------|----------------|--------------|
| M1 100/0 | 0.00         | 0.00           | 0.00         |
| M2 99/1  | 3.41 ± 2.49  | 14.15 ± 14.74  | 0.00         |
| M3 95/5  | 7.17 ± 5.17  | 38.17 ± 12.16  | 0.00         |
| M4 90/10 | 19.93 ± 11.6 | 38.27 ± 32.78  | 0.00         |
| M5 75/25 | 23.57 ± 6.16 | 162.34 ± 18.77 | 9.40 ± 5.72  |
| M6 50/50 | 20.42 ± 8.40 | 190.57 ± 61.26 | 5.10 ± 2.14  |
| M7 0/100 | 29.83 ± 4.02 | 185.12 ± 14.33 | 13.28 ± 2.11 |

### 9.2.6. Detección de heterorresistencia mediante Secuenciación.

Por su parte, la secuenciación Nanopore mostró una capacidad de detección dependiente de la proporción de cepa resistente. En las mezclas con  $\geq 10\%$  de bacilos resistentes, la secuenciación identificó claramente las mutaciones de resistencia correspondientes en las lecturas. Por ejemplo, en la mezcla 10/90 de isoniazida se detectó la mutación *katG S315T* en  $\sim 12\%$  de las lecturas, concordando con la proporción esperada. Sin embargo, en las mezclas de 5% y 1%, la presencia de la mutación no pudo diferenciarse del ruido de fondo o error de secuenciación. En la mezcla 5/95 de rifampicina, por ejemplo, la frecuencia de la mutación *rpoB S531L* en las lecturas fue  $\sim 2\%$ , un valor cercano al límite de detección del análisis, por lo que no se pudo afirmar con confianza la presencia de la subpoblación resistente (se requeriría quizás mayor cobertura o una estrategia de enriquecimiento para confirmarla). En cambio, para pirazinamida, la mutación *pncA* de la cepa resistente (un cambio puntual) se detectó en  $\sim 7\%$  de lecturas en la mezcla 10% pero no en la 5%.

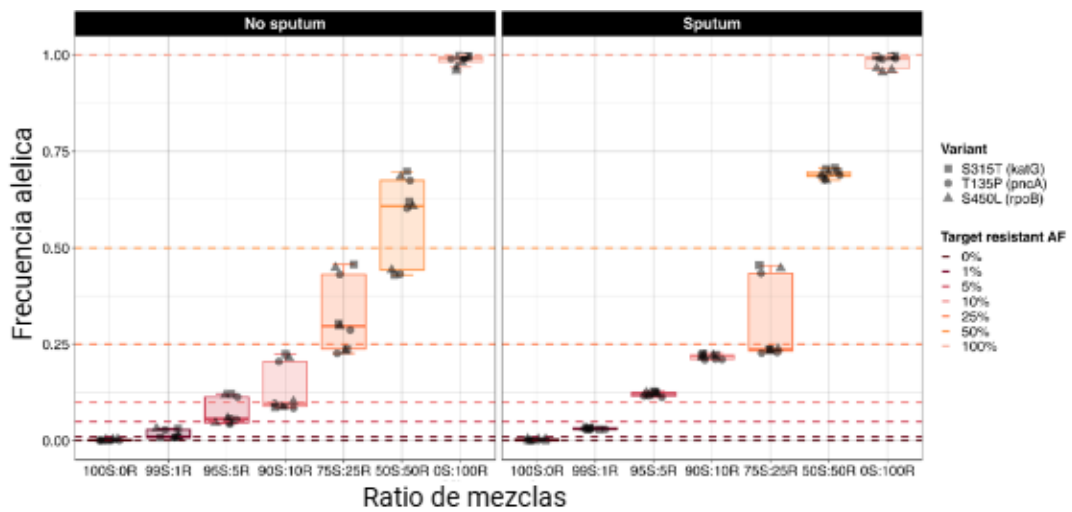


Figura 28. Detección de heterorresistencia en mezcla de cepas bacterianas S:R con la resistencia predicha por TB-Profiler para rifampicina, isoniazida y pirazinamida.

Gráfico de cajas que muestra la relación entre las proporciones de cepas sensibles (S) y resistentes (R) en muestras analizadas y la resistencia predicha a rifampicina utilizando TB-Profiler. Se presentan dos categorías: CS (con esputo) y SS (sin esputo), comparando diferentes proporciones de S:R, desde 100:0 hasta 0:100. Los colores representan diferentes niveles de resistencia esperados según el porcentaje objetivo, desde 1% hasta 100%.

En resumen, el límite de detección de HR por secuenciación Nanopore en este protocolo fue cercano al 10% de población resistente. Por debajo de ese umbral, la fiabilidad disminuyó. Esto implica que, si la subpoblación resistente es muy pequeña (<5–10%), es probable que la secuenciación pase por alto su presencia o no la distinga del error inherente. Si bien la secuenciación mejoró notablemente ese límite respecto a Xpert, aún quedó por encima del método de proporciones (1%).

Un hallazgo interesante fue que la rifampicina mostró cierta detección por secuenciación incluso a niveles muy bajos, en la mezcla 1% de rifampicina, aunque en general no se detectó la mutación de forma consistente, en una de las réplicas se obtuvo un porcentaje inusualmente alto de lecturas mutantes (~14%). En cualquier caso, no fue reproducible, y para fines prácticos consideramos que <5% no es confiable.

Finalmente, comparando ambos métodos, la secuenciación Nanopore identificó HR en todas las muestras que tenían  $\geq 10\%$  de cepas resistentes (mostrando las mutaciones correspondientes) (figura 28), mientras que el método de proporciones identificó HR hasta en las muestras con ~1%. Por lo tanto, no se logró equiparar completamente la sensibilidad del “gold standard”. Sin embargo, la secuenciación aportó información adicional valiosa, por ejemplo, confirmó la presencia de mutaciones específicas, permitió distinguir qué mecanismo de resistencia estaba involucrado (algo que el cultivo no brinda) e incluso en mezclas  $\geq 25\%$  pudo inferirse la proporción aproximada de la subpoblación resistente a partir de la frecuencia de lecturas mutantes.

Tabla 12. Resumen de los parámetros de secuenciación de amplicones: Rendimiento total de datos, la calidad promedio de las lecturas (Q-score) y el tamaño de N50, profundidad de secuenciación promedio por amplicón.

|             | Secuenciación de amplicones |
|-------------|-----------------------------|
| Rendimiento | 80.36Gb                     |
| Q-score     | 20                          |

|                                   |         |
|-----------------------------------|---------|
| N50                               | 2.49kb  |
| Data clasificada                  | 76.76Gb |
| Data no clasificada               | 3.6Gb   |
| Profundidad por amplicon promedio | 20000X  |

## 10. DISCUSIÓN

La detección temprana y precisa de la HR en *M. tuberculosis* es crítica para el manejo adecuado de la TB resistente, especialmente en regiones de alta prevalencia como el Perú (Callao, Loreto) (92), donde la TB multidrogoresistente representa un gran desafío de salud pública. En estos entornos, caracterizados por alta transmisión, recursos diagnósticos limitados y tratamientos complejos, es imprescindible implementar técnicas avanzadas que permitan identificar HR oportunamente. En este contexto, este estudio evaluó un método molecular innovador, el cual fue la secuenciación Nanopore, frente a una técnica tradicional de referencia, que fue el método de proporciones en agar, con miras a proponer un método que podría optimizar el diagnóstico y la gestión clínica de pacientes con TB HR.

El proceso de secuenciación de muestras de ADN requiere determinar críticamente cada paso en la preparación del material genético, por lo que en este trabajo se comparó dos protocolos de descontaminación del esputo: el uso de NALC/NaOH frente a la alternativa basada en el uso de saponina. Los resultados de la figura 9 muestran al NALC/NaOH como el método con menor rendimiento en cantidad y calidad de ADN (9.2ng/μL y ratio 260/280 de 1.56 en promedio) (Tabla S1) que comparados con los resultados obtenidos por Dippenaar, Anzaan et al (93) al utilizar NALC/NaOH como método de descontaminación y Cetyl Trimethyl Ammonium Bromide (CTAB) como método de extracción obtuvo valores similares de concentración de ADN excepto cuando agregó los reagentes MycoPrep y Myco-TB (los cuáles no se utilizaron en este estudio) lo que le

permitió alcanzar concentraciones de hasta 79.1 ng/  $\mu$ L con ratios de calidad 260/280 mayores a 1.7; Por otro lado, la cantidad de ADN obtenida usando saponina como método de descontaminación (246 ng/ $\mu$ L en promedio) fue mayor a lo reportado por Kot et al. (>60 ng/ $\mu$ L) (94) quien uso tres kits comerciales para el proceso de extracción: DNeasy UltraClean Microbial Kit (cat N° 12224), QIAamp DNA Mini Kit (cat N°51304) y AllPrep PowerFecal DNA/RNA Kit (cat N° 80244); sin embargo, aunque en este estudio se utilizó un kit comercial (Zymo) para la extracción, este tuvo modificaciones como el uso del Fastprep-24 5G lo que explicaría la alta cantidad de ADN obtenido. Adicionalmente, la diferencia entre el rendimiento de ADN obtenido por cada método de descontaminación se explica por la naturaleza de los agentes: NALC/NaOH genera lisis en las bacterias presentes en la muestra de esputo (95) liberando su ADN y descartándolo durante el proceso mientras que la saponina se enfoca en eliminar el ADN de células humanas (96). Debe considerarse también, que al ser ambos métodos de descontaminación, la presencia de inhibidores obtenidos mediante cada método puede impactar en la cantidad de ADN detectado en procesos posteriores de amplificación como ya se ha reportado en ensayos en donde se usa métodos de descontaminación de tipo alcalinizante (NALC/NOH) y es que la presencia de este tipo de bases alcalinas inhiben y/o retardan la acción de la enzima (97) a así como interfieren en la detección de algunas regiones génicas sobre otras pero es un método bastante usado cuando se trabaja con cepas debido a la alta especificidad que genera (98). La saponina por su parte ha mostrado no afectar la acción de la enzima en la detección del gen diana y aunque tiene una gran capacidad para “depletar” (eliminar) el ADN del hospedero (humano) en las muestras de esputo no elimina el material genético de las demás

bacterias que no corresponden a *Mycobacterium tuberculosis* (99) por lo que la alta cantidad de ADN que permite obtener no asegura la presencia de *M. tuberculosis* en la muestra y usar un exceso de ADN en una reacción de PCR puede inhibir la acción de la enzima (100). Consecuentemente, se muestra a la saponina como el método a utilizar cuando se requiere una alta cantidad de ADN directo y al NALC/NaOH como el método a utilizar si se trabaja con cultivos en donde se requiere generar amplicones.

El proceso de extracción también es una etapa de gran importancia dentro de la preparación de la muestra, la figura 10 muestra a ambos métodos como opciones viables debido a que el método Fenol permitió obtener una elevada cantidad de ADN respecto a lo obtenido mediante el método kit. Al enfocarnos en la cantidad de ADN obtenido por el método fenol se observa un valor máximo de 6µg de ADN obtenido que supera a los valores reportados anteriormente por de Almeida et al. (101) de 2 µg como máximo. Por otro lado, al comparar los resultados obtenidos del método kit con los de Oriano M. et al. (102). quien evaluó hasta 3 kits de extracción basados en columnas de purificación: High Pure PCR Template Preparation Kit (Roche - cat N° 11796828001), Quick-DNA Universal Kit (Zymo – cat N° D4068) y DNeasy PowerLyzer PowerSoil DNA Isolation Kit (Qiagen – cat N° 12855-50) con 2 métodos de descontaminación: ditioneitol (DTT) y lisozima, se puede notar la diferencia de cantidad de ADN obtenido. La obtención de ADN utilizando el kit Miniprep (Zymo) permitió obtener cantidades totales de ADN de 10µg en promedio por muestra partiendo desde 1mL de esputo mientras que Oriano obtuvo hasta 20µg de ADN total haciendo uso del Quick-DNA Universal Kit además de lisozima y ditioneitol (DTT) pero con 5 mL de esputo

inicial, por lo que se corrobora que la inclusión del Fastprep-24 5G dentro del método de extracción como lisis mecánica propició la obtención de mayor cantidad de ADN justificando la modificación del protocolo de extracción con el kit Miniprep (Zymo).

Aunque ambos métodos proporcionaron suficiente cantidad de ADN para la secuenciación usando ADN genómico. La figura 11 permite observar la integridad del ADN obtenido mediante cada método siendo el kit el que brindo los fragmentos de mayor tamaño, lo cual lo hace una mejor opción en caso se busque secuenciar directamente DNA genómico o amplicones. Estos valores de DIN (6 aproximadamente) son similares a los reportados por Nazlı Arslan et al. (103) quien comparó distintos métodos de extracción de ADN enzimáticos y mecánicos en muestras de cultivo de *M. tuberculosis*. obteniendo valores de DIN alrededor de 6. Por otro lado, el uso del método fenol genera fragmentos de menor tamaño lo que indica una mayor fragmentación del ADN y aunque se muestra como una opción viable a utilizar en caso se requiera amplicones hay bibliografía en donde se muestra que el uso de fenol puede interferir en la inespecificidad de amplicones a obtener además de inhibir ensayos posteriores de PCR (104,105).

Respecto a la generación de amplicones, proceso bastante conocido, debe considerarse que es un proceso dependiente de la enzima de amplificación. Trabajos presentados como el de Filges S et al. (106) mencionan que una de las principales fuentes que pueden generar error en el proceso de secuenciación provienen del tipo de enzima con el que se esté generando los amplicones siendo la Platinum Taq la más susceptible a generar copias complementarias con nucleótidos que no

representan al ADN inicial (9X de fidelidad de amplificación frente a 25X obtenida por la enzima Phusion). Esto es reforzado por Brandariz-Fontes C et al. (107) quien también evaluó distintas enzimas de amplificación y su impacto en procesos de secuenciación por pirosecuenciación encontrado que enzimas como Onetaq y Biotaq generaron un 53% de lecturas correctas (que correspondían a la secuencia amplificada) mientras que la enzima Phusion alcanzó valores de hasta 92% de lecturas correctas lo cual fundamenta el uso de la enzima Phusion para obtener amplicones dentro de este estudio esperando disminuir en lo máximo posible el error que se pueda generar en esta etapa de amplificación lo que también es respaldado por Guido N et al. (108) al mencionar que secuencias ricas en GC requieren enzimas de alta fidelidad como es la Phusion junto a dimetilsulfóxido (DMSO).

Adicionalmente, se ha registrado que la presencia de inhibidores en el esputo (mucina, sangre, detritos celulares) pudo afectar la eficiencia de la PCR. Cauwelaert et al. reportaron que la amplificación directa desde esputo suele ser menos consistente, observándose bandas más débiles incluso en muestras con alta carga bacilar, recomendando estrategias como PCR anidada para mejorar la sensibilidad (109). En concordancia, nuestros resultados sugieren que la preparación de amplicones desde muestras clínicas requiere optimizaciones adicionales para la eliminación de inespecificidades (Figura 14) (por ejemplo, un paso de purificación del ADN de *M. tuberculosis* para remover ADN humano, o el uso de aditivos PCR que contrarresten inhibidores). Dou et al. (2025) describen que la inclusión de buffers termoestables y pretratamientos para reducir el contenido de ADN humano aumenta considerablemente la intensidad y reproducibilidad de los amplicones

obtenidos directamente de esputo. Por otro lado, las muestras provenientes de cultivo (ADN de colonias) proporcionaron un ADN más puro y concentrado, lo que facilitó una amplificación PCR más específica y robusta (Figura 12 y 13). En estas, se obtuvieron amplicones de intensidad homogénea para todos los genes, sin amplificaciones inespecíficas que además concordaban con lo mencionado por Yang et al. (2018), quienes reportaron especificidad del 100% y sensibilidad >93% al amplificar directamente ADN de cultivo para detectar mutaciones en *rpoB* y *katG*, obteniendo bandas claras y reproducibles (110), lo cual coincide con nuestros hallazgos en muestras de cultivo.

La generación de librerías génicas y el rendimiento de estos se observan en las tablas 3,4, 5 y 6 en donde se observa una gran diferencia en la cantidad de data generada por cada tipo de librería. La importancia de la comparación de kits de librería génica se sustenta en trabajos como el de Sauvage T et al. (111) en donde se comparó 3 tipos de librerías génicas (*Rapid*, *Native Ligation* y *PCR Barcoding*) y sustenta que si se cuenta con suficiente ADN para el uso del kit *Native Ligation* es la opción más recomendable y si lo que se busca es rapidez a costa de mayor ruido y menor rendimiento el *kit Rapid Barcoding* sería la opción a elegir. Dentro de sus resultados se observa que al utilizar el kit *Native Ligation* y trabajando con un q-score > 7 obtuvo un rendimiento de hasta 33.6Gb con un porcentaje de lecturas aprobadas de 92.9% que fueron similares a los resultados obtenidos en este estudio, aunque con un score > 10 (solo en la etapa de estandarización) (26.97 Gb y un 90.1% de lecturas aprobadas) cuando se utilizó muestras provenientes de cultivo puro; sin embargo, al evaluarlo frente a los resultados en muestras de esputo (10.4Gb y 93.7%) el rendimiento decae hasta un tercio, esto es explicado por trabajos como

los de Brandt C. et al. (112) y Maghini D. G. et al. (113) quienes mencionan que las muestras metagenómicas presentan fragmentos de ADN de menor tamaño que las muestras provenientes de cultivo que terminan disminuyendo el rendimiento de los poros dentro de una celda de flujo Nanopore. De hecho, se observó que la tasa de lecturas descartadas por calidad fue consistentemente menor con *Native* (promedio 5%) que con *Rapid* (promedio 15%), lo cual concuerda con estudios previos donde el protocolo de ligación mostró menor error de secuenciación (114,115) Adicionalmente, Chen, Ziming et al. (116) menciona que el kit *Rapid Barcoding* presenta problemas al secuenciar regiones con una elevada cantidad de GC en su secuencia y que debido a esto puede causar una subrepresentación de la variedad poblacional dentro de la muestra lo que es muy importante al evaluar mutaciones en una sola especie y organismo como *M. tuberculosis* (117).

Aunque la bibliografía presentada respalda el uso del kit *Native Ligation* para muestras con alto contenido GC en su secuencia hay evidencia como la presentada por Wick RR et al. (118) en donde presenta al *Rapid Barcoding* como una opción viable si lo que se busca es secuenciar plásmidos debido a la rapidez de este kit. Por otro lado, usando como fuente de ADN los amplicones también se presenta al Kit *Native Ligation* como la mejor opción; sin embargo, Player R et al. (119) presenta además cambios en el proceso de preparación de librería (incluir una PCR y un software ONT-DART) para mejorar la detección de la totalidad de variantes dentro de la muestra y disminuir la cantidad de falsos positivos que puedan generarse.

Aunque ambos kits de librería génica no generaron resultados que permitan trabajar directamente el ADN genómico los resultados obtenidos por el kit *Rapid Barcoding*

fueron ampliamente superados *por Native Ligation* que sumado a su principio de funcionamiento; en donde, no fragmenta el ADN se muestra como la opción correcta para un proceso de secuenciación de alto rendimiento.

Respecto al tipo de ADN a utilizar, nuestros resultados mostraron que utilizando ADN genómico como fuente de información no se logró cubrir el genoma completo de *M. tuberculosis* lo que es concordante con trabajos como el de Andrzej Pastusiak et al. (120) en donde no logró cubrir el genoma completo de *Plasmodium falciparum* en una muestra de tipo metagenómica proveniente de mosquitos y Kok NA et al. (121) que tampoco pudo cubrir el genoma completo de *M abscessus* en una muestra de esputo, explicando que la naturaleza propia de una muestra metagenómica (diversidad de especies en una muestra) no permite cubrir el genoma del patógeno de interés a menos que se realice un proceso de eliminación de material genético indeseable (122).

la tabla 5 y 6 muestran que no se logró obtener una profundidad aceptable (>1X) del genoma de *Mycobacterium tuberculosis* para considerarla como una opción viable por lo que la segunda fase se realizó utilizando amplicones.

En este estudio se encontró que, utilizando un enfoque basado en amplicones, fue posible identificar mutaciones de resistencia a isoniazida, rifampicina y pirazinamida en mezclas bacterianas heterogéneas, siempre que la proporción de bacilos resistentes superara dicho umbral (~10%). Por debajo de este porcentaje, la detección no fue confiable, estableciendo así el límite práctico de la técnica en nuestras condiciones experimentales. En cambio, el método fenotípico de proporciones en agar confirmó la presencia de HR incluso con subpoblaciones

resistentes del 1%, reafirmando su alta sensibilidad. No obstante, la secuenciación Nanopore ofreció importantes ventajas, entre ellas redujo el tiempo de obtención de resultados de semanas a tan solo 2–3 días (incluyendo extracción, preparación de bibliotecas y análisis), identificó directamente las mutaciones genéticas responsables de la resistencia, y permitió analizar múltiples muestras simultáneamente mediante la asignación de código de barras a cada muestra, evidenciando su potencial como herramienta complementaria en el diagnóstico de TB resistente.

Los resultados obtenidos se alinean parcialmente con la literatura previa. Por un lado, confirman que la detección de HR mediante métodos moleculares tiene un límite superior al del cultivo en agar y no todos los métodos microbiológicos pueden detectar HR. Estudios como el de Werngren et al. ya habían señalado que pruebas que hacen uso de cultivos como la prueba de Wayne no detectan HR a pirazinamida hasta que la cepa resistente es claramente predominante (>10%) (28). En este caso, la secuenciación, aunque más sensible que Xpert u otras pruebas moleculares, aún requirió una proporción significativa de bacilos resistentes. Por otro lado, se coincidió con Nimmo et al. (17,123) en que el análisis directo de muestras clínicas puede revelar poblaciones que se pierden en el subcultivo por un efecto “cuello de botella”, se observa que la secuenciación de amplicones desde esputo logró detectar la coexistencia de genotipos que un análisis post-cultivo podría subestimar como en el caso de la rifampicina en el método de proporciones en Agar (probablemente debido a una variación en el pH del medio de cultivo) (124). Esto refuerza la idea de que, a pesar de las dificultades técnicas, los métodos moleculares directos brindan una “fotografía” más fiel de la diversidad de cepas in vivo. Esta observación

está alineada con lo reportado para métodos moleculares, donde estudios han señalado que técnicas como MGIT (líquido) detectan  $\approx 10\%$  como límite y Xpert requiere  $>60\%$  (23,61) Adicionalmente, estos hallazgos concuerdan con Cabibbe et al. (2020), quienes destacaron la viabilidad de Nanopore para detección de resistencia a menor costo y con infraestructura mínima (10). En su estudio, la concordancia de Nanopore con Illumina fue alta, aunque no enfrentaron escenarios de HR deliberada.

La identificación de HR en muestras de esputo con *M. tuberculosis* ya ha sido reportada anteriormente. 4 de 499 aislados clínicos resultaron heterorresistentes a rifampicina según Zheng C. et al. (125), con un porcentaje de lecturas mutantes cercanos al 10%, lo que se asemeja a lo mostrado en este estudio; sin embargo, la mutación reportada fue la de Ser531Leu mientras que en este estudio, al trabajar con cepas H37Rv y DM97, la mutación causante de resistencia a RIF fue la S450L mostrando la necesidad de ampliar el espectro de mutaciones por cada gen de resistencia a evaluar. Paralelamente, en India (126) se reportó la presencia de casos de HR a más de un fármaco mostrando las mutaciones S531L para el gen *rpoB*, S315T para el gen *katG* y mutaciones en la región *eis*. Nuestros resultados muestran similitud en la mutación S315T reportada para el gen *katG*; sin embargo, no cubre las demás mutaciones reportadas debido a que la evaluación de HR fue con la cepa DM97 (perteneciente al Laboratorio de Micobacterias de la Universidad Peruana Cayetano Heredia) y si bien es MDR-RR no presenta toda la amplia gama de mutaciones causantes de resistencia a los antibióticos. Tal es el caso del gen *inhA*, en este estudio inicialmente se incluyó este gen para la evaluación de HR pero se desconocía (hasta el momento de la secuenciación) que la cepa DM97 no presentaba

la mutación en ese gen por lo que no se pudo evaluar HR utilizando esa región génica.

Adicionalmente, se ha visto que la detección de HR también puede depender mucho del análisis bioinformático que se esté llevando a cabo. Trabajos como el de Werngren et al. (127) y Whitfield et al. (128) detectaron HR al 10% y al 5% cuando el proceso de filtrado de datos fue más flexible, lo que a su vez aumenta la probabilidad de estar reportando falsos positivos.

La presente investigación demostró que la secuenciación mediante la plataforma Nanopore es capaz de detectar subpoblaciones resistentes de *M. tuberculosis* cuando estas constituyen al menos alrededor de 10% de la población total, lo cual ya se había propuesto en organismos como *E. coli* pero solo haciendo uso de Machine learning luego de una secuenciación de genoma completo (129) de cultivo un aspecto que aún necesita de mayor investigación en muestras clínicas.

Nuestros resultados respaldan y son respaldados por bibliografía presentada, la utilización de la secuenciación Nanopore como una herramienta complementaria de diagnóstico de TB resistente, capaz de detectar HR con mayor rapidez que los métodos tradicionales. Si bien no alcanza la sensibilidad absoluta del método de proporciones, aporta información genotípica valiosa y su desempeño podría mejorarse aún más con optimizaciones técnicas. Estudios futuros deberán validar este enfoque en cohortes de pacientes y evaluar su impacto en los resultados del tratamiento.

Finalmente, la importancia de estos hallazgos radica en su contribución a comprender las fortalezas y debilidades de los métodos diagnósticos actuales para TB resistente. La capacidad de la secuenciación Nanopore de detectar HR en niveles alrededor del 10%, especialmente frente a fármacos críticos como rifampicina, podría ser clave para implementar estrategias diagnósticas avanzadas en programas de control de TB-MDR. Por lo tanto, futuras investigaciones deben enfocarse en validar estos hallazgos en muestras clínicas reales y en optimizar metodologías que permitan reducir los sesgos y aumentar la precisión en la cuantificación de resistencias, asegurando así un tratamiento más efectivo y disminuyendo la transmisión de cepas resistentes.

## 11. CONCLUSIONES

**Estandarización de métodos de descontaminación, extracción de ADN, preparación y generación de librerías génicas y tipo de ADN usando parámetros de secuenciación a partir de muestras de esputo y cultivos primarios.**

- El método de descontaminación de muestras de esputo basado en saponina permite recuperar aproximadamente 4 µg de ADN, y es adecuado para la secuenciación de genomas completos. En cambio, el método NALC/NaOH, con un rendimiento cercano a 20 ng, resulta suficiente para la secuenciación de amplicones.
- Los protocolos de extracción de ADN con Fenol y con Kit obtuvieron ADN de alta pureza e integridad, adecuado para ensayos de secuenciación y con suficiente material de *M. tuberculosis*. Sin embargo, se seleccionó el protocolo basado en Kit por ser un método más sencillo y rápido.
- La secuenciación que usó ADN genómico extraído de muestras de esputo solo alcanzó una profundidad cercana a 1×, lo cual es insuficiente para identificar variantes génicas; en cambio, el uso amplicones logró profundidades de 1 000× a 60 000×.
- El kit de librería génica *Native Ligation* generó mayor rendimiento de data (4 -26 Gb), mejor valor de q-score (18-21), profundidad (1000X – 60000x) y mayor tamaño de N50 en comparación al kit *Rapid barcoding* usando ADN genómico así como amplicones, aunque eso implique un tiempo de preparación mayor.

**Determinación del límite de detección de heterorresistencia en mezclas bacterianas mediante secuenciación y método de proporciones.**

- La secuenciación usando amplicones permitió la detección de HR en mezclas de cepas susceptibles/resistentes con y sin esputo.
- La secuenciación Nanopore detectó heterorresistencia solo cuando las subpoblaciones resistentes superaron ~10% del total; por debajo de ese umbral (<5–10%), la detección fue inconsistente.
- Aunque la secuenciación Nanopore permitió detectar HR en muestras simuladas aún no se puede generalizar su comparación frente al método de proporciones en agar por la falta de robustez en la cantidad de muestras evaluadas.

## 12. LIMITACIONES

- Si bien la calidad de secuenciación obtenida con la plataforma Nanopore en nuestro estudio fue alta y comparable a la reportada con Illumina en otros contextos, sería recomendable realizar secuenciaciones paralelas con Illumina (u otra plataforma de alta fidelidad) para validar de manera independiente la precisión en la detección de mutaciones y subpoblaciones minoritarias. Esta validación cruzada no se llevó a cabo por restricciones de tiempo y presupuesto.
- Durante la estandarización del protocolo nos enfocamos en seleccionar métodos que maximicen la cantidad total de ADN extraído, atendiendo a los requerimientos mínimos de las bibliotecas de secuenciación. Sin embargo, esto no garantiza que la fracción de ADN micobacteriano sea suficiente en todos los casos. Una limitación fue no cuantificar directamente el porcentaje de ADN de *M. tuberculosis*. humano en cada muestra; en el futuro, técnicas como qPCR específica podrían implementarse para evaluar la eficiencia de enriquecimiento de *Mycobacterium tuberculosis* en cada extracción.
- La tecnología de Nanopore está en continua evolución. En este estudio se usaron flow cells R9.4.1 con química de basecalling supuesta en 2023. Los resultados pueden variar con futuras actualizaciones (por ejemplo, las flow cells R10 con diferentes dinámicas de error). Es posible que alguna variabilidad observada se deba a la versión de la tecnología empleada, lo cual escapa a nuestro control. Se intentó mitigar manteniendo actualizados los algoritmos de análisis (Guppy, MinKNOW) a sus últimas versiones estables.

- Las mezclas bacterianas heterorresistentes simuladas en este trabajo se generaron a partir de cepas distintas (HR policlonal). Esto reflejó un escenario de coinfección con cepas resistentes y susceptibles (muy relevante clínicamente); no obstante, no exploramos la HR monoclonal derivada de mutantes emergentes de una misma cepa. Sería conveniente realizar estudios complementarios donde se induzca resistencia en laboratorio sobre una cepa inicial para evaluar si la secuenciación detecta esas subpoblaciones derivadas con la misma sensibilidad.
- A pesar de haber considerado trabajar la generación de amplicones utilizando una enzima de alta fidelidad, lecturas con un q-score > 20 (99% de precisión) y valores de profundidad mayores a 1000x para cada gen aún se mantiene la posibilidad de generación de errores por la tasa intrínseca de error en secuenciación Nanopore (mayor que Illumina), lo que podría llevar a falsos positivos.
- En este estudio se trabajó con un número limitado de cepas y mezclas simuladas; aunque representativas, no capturan toda la diversidad genética de *M. tuberculosis*. Por ejemplo, solo se evaluó una mutación por fármaco (*katG* S315T, *rpoB* S531L, *pncA* C14T), cuando en la clínica existen múltiples mutaciones posibles cuya detectabilidad podría variar (especialmente *pncA*, con decenas de alelos).
- El uso de un número limitado de muestras analizadas (N =21, 7 mezclas por triplicado) en la etapa de comparación de límite de detección no permite la generalización de los resultados; sin embargo, se puede tomar como ensayos

pilotos que muestran a la secuenciación como una opción a considerar en la detección de HR.

### **13. RECOMENDACIONES**

Este trabajo de investigación no solo brinda resultados al conocimiento científico, sino que también permite hacer una serie de recomendaciones. Entre ellas, implementar estudios piloto futuros con muestras de esputo de pacientes TB reales, aplicando el protocolo de secuenciación Nanopore optimizado en paralelo al método de proporciones tradicional. Esto permitirá corroborar el desempeño en condiciones clínicas, evaluar la concordancia de resultados y determinar en qué proporción de pacientes se detecta HR oculta que pudiera pasar desapercibida con los métodos actuales. Los hallazgos de tal piloto facilitarán la adopción informada de esta tecnología en laboratorios de referencia.

De otro lado con respecto a la optimización de la sensibilidad, se recomienda explorar modificaciones al protocolo para mejorar la detección de subpoblaciones minoritarias. Por ejemplo, incorporar una etapa de *deep sequencing* focalizada en genes de resistencia clave (p. ej., mediante enriquecimiento por captura de híbridos de las regiones *rpoB*, *katG*, etc., antes de la secuenciación) podría aumentar la cobertura sobre esos loci y reducir el límite de detección por debajo de 10%. Asimismo, ensayar la utilización de un número mayor de ciclos de PCR o PCR anidada para amplicones en muestras paucibacilares podría amplificar señales débiles de resistencia.

Dado que encontramos que las infecciones mixtas pueden detectarse y caracterizarse con secuenciación, se recomienda implementar la tipificación molecular (por ejemplo, por *spoligotyping* o análisis de variantes) de las lecturas secuenciadas para identificar la presencia de más de un genotipo de *Mycobacterium tuberculosis* en la misma muestra. Esto aportaría información epidemiológica valiosa y ayudaría a distinguir HR policlonal de monoclonal. En contextos de alta incidencia, se sugiere vigilar la frecuencia de coinfecciones, ya que estas pueden requerir consideraciones especiales en el tratamiento y el control de brotes.

Es necesario continuar investigando tecnologías complementarias para mejorar la detección de HR. Por ejemplo, la combinación de Nanopore con técnicas de depleción de ADN del huésped (como el uso de enzimas exonucleasas que degraden ADN eucariota) podría aumentar la fracción de lecturas *Mycobacterium tuberculosis*, incrementando la sensibilidad global. Además, implementar algoritmos de *machine learning* que analicen las señales de corriente de Nanopore de manera directa podría, en un futuro, identificar variaciones sutiles sin necesidad de una cobertura extremadamente alta, lo que revolucionaría la detección de HR en tiempo real.

Finalmente, estas recomendaciones buscan avanzar hacia un diagnóstico más oportuno y preciso de la TB resistente. Integrar nuevas herramientas como Nanopore en los programas de control de TB podría mejorar significativamente el manejo de la enfermedad, beneficiando tanto al paciente individual como a la salud pública al frenar la propagación de cepas resistentes.

#### **14. REFERENCIAS BIBLIOGRÁFICAS**

1. Bagcchi S. WHO's Global Tuberculosis Report 2022. *Lancet Microbe*. 2023;
2. Rendon A, Centis R, D'Ambrosio L, Migliori GB. Estrategias de la OMS para el tratamiento de la tuberculosis resistente. *Arch Bronconeumol*. 2017 Mar;53 (3):95–7.
3. Amirkhani A, Humayun M, Ye W, Worku Y, Yang Z. Patient characteristics associated with different types of prison TB: an epidemiological analysis of 921 TB cases diagnosed at an Ethiopian prison. *BMC Pulm Med*. 2021 Dec 1;21 (1).
4. Eldholm V, Monteserin J, Rieux A, Lopez B, Sobkowiak B, Ritacco V, et al. Four decades of transmission of a multidrug-resistant *Mycobacterium tuberculosis* outbreak strain. *Nat Commun*. 2015 May 11;6.
5. Schmalstieg AM, Srivastava S, Belkaya S, Deshpande D, Meek C, Leff R, et al. The antibiotic resistance arrow of time: Efflux pump induction is a general first step in the evolution of mycobacterial drug resistance. *Antimicrob Agents Chemother*. 2012 Sep;56 (9):4806–15.
6. Belknap RW. Current Medical Management of Pulmonary Tuberculosis. Vol. 29, Thoracic surgery clinics. NLM (Medline); 2019. p. 27–35.
7. Lam E, Nateniyom S, Whitehead S, Anuwatnonthakate A, Monkongdee P, Kanphukiew A, et al. Use of drug-susceptibility testing for management of drug-resistant tuberculosis, Thailand, 2004-2008. *Emerg Infect Dis*. 2014 Mar;20 (3):400–8.

8. Andersson DI, Nicoloff H, Hjort K. Mechanisms and clinical relevance of bacterial heteroresistance. Vol. 17, *Nature Reviews Microbiology*. Nature Publishing Group; 2019. p. 479–96.
9. Shi CL, Han P, Tang PJ, Chen MM, Ye ZJ, Wu MY, et al. Clinical metagenomic sequencing for diagnosis of pulmonary tuberculosis. *Journal of Infection*. 2020 Oct 1;81 (4):567–74.
10. Cabibbe AM, Spitaleri A, Battaglia S, Colman RE, Suresh A, Uplekar S, et al. Application of targeted next-generation sequencing assay on a portable sequencing platform for culture-free detection of drug-resistant tuberculosis from clinical samples. *J Clin Microbiol*. 2020 Oct 1;58 (10).
11. Lindahl T. Instability and decay of the primary structure of DNA. 1993;362 (6422):709-15.
12. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol*. 2016 Dec;17 (1).
13. Golparian D, Donà V, Sánchez-Busó L, Foerster S, Harris S, Endimiani A, et al. Antimicrobial resistance prediction and phylogenetic analysis of *Neisseria gonorrhoeae* isolates using the Oxford Nanopore MinION sequencer. *Sci Rep*. 2018 Dec 1;8 (1).
14. Woods GL. Susceptibility testing of mycobacteria, nocardiae, and other aerobic actinomycetes: approved standard. *Clinical and Laboratory Standards Institute*; 2011. 61 p.
15. Votintseva AA, Bradley P, Pankhurst L, Del Ojo Elias C, Loose M, Nilgiriwala K, et al. Same-day diagnostic and surveillance data for

- tuberculosis via whole-genome sequencing of direct respiratory samples. *J Clin Microbiol.* 2017 May 1;55 (5):1285–98.
16. Walker TM, Kohl TA, Omar S V., Hedge J, Del Ojo Elias C, Bradley P, et al. Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: A retrospective cohort study. *Lancet Infect Dis.* 2015 Oct 1;15 (10):1193–202.
  17. Nimmo C, Shaw LP, Doyle R, Williams R, Brien K, Burgess C, et al. Whole genome sequencing *Mycobacterium tuberculosis* directly from sputum identifies more genetic diversity than sequencing from culture. *BMC Genomics.* 2019 May 20;20 (1).
  18. Ye M, Yuan W, Molaeipour L, Azizian K, Ahmadi A, Kouhsari E. Antibiotic heteroresistance in *Mycobacterium tuberculosis* isolates: a systematic review and meta-analysis. Vol. 20, *Annals of Clinical Microbiology and Antimicrobials.* BioMed Central Ltd; 2021.
  19. Oriano M, Terranova L, Teri A, Sottotetti S, Ruggiero L, Tafuro C, et al. Comparison of different conditions for DNA extraction in sputum - A pilot study. *Multidiscip Respir Med.* 2019 Jan 31;14 (1).
  20. Satapathy P, Das D, Murmu BN, Kar SK. Decontamination of sputum for longer time in sodium hydroxide for isolation of *Mycobacterium tuberculosis*. *Int J Mycobacteriol.* 2014 Dec 1;3 (4):290–2.
  21. Kambli P, Ajbani K, Kazi M, Sadani M, Naik S, Shetty A, et al. Targeted next generation sequencing directly from sputum for comprehensive genetic information on drug resistant *Mycobacterium tuberculosis*. *Tuberculosis.* 2021 Mar 1;127.

22. Murphy SG, Smith C, Lapierre P, Shea J, Patel K, Halse TA, et al. Direct detection of drug-resistant *Mycobacterium tuberculosis* using targeted next generation sequencing. *Front Public Health*. 2023;11.
23. Dippenaar A, Ismail N, Grobbelaar M, Oostvogels S, de Vos M, Streicher EM, et al. Optimizing liquefaction and decontamination of sputum for DNA extraction from *Mycobacterium tuberculosis*. *Tuberculosis*. 2022 Jan 1;132.
24. Kok NA, Peker N, Schuele L, de Beer JL, Rossen JWA, Sinha B, et al. Host DNA depletion can increase the sensitivity of *Mycobacterium* spp. detection through shotgun metagenomics in sputum. *Front Microbiol*. 2022 Oct 25;13.
25. Yang L, Haidar G, Zia H, Nettles R, Qin S, Wang X, et al. Metagenomic identification of severe pneumonia pathogens in mechanically-ventilated patients: A feasibility and clinical validity study. *Respir Res*. 2019 Nov 27;20 (1).
26. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples [Internet]. 2017. Available from: <http://biorxiv.org/lookup/doi/10.1101/201178>
27. Head SR, Kiyomi Komori H, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR, et al. Library construction for next-generation sequencing: Overviews and challenges. *Biotechniques*. 2014 Feb;56 (2):61–77.
28. Werngren J, Mansjö M, Glader M, Hoffner S, Forsman LD. Detection of Pyrazinamide heteroresistance in *mycobacterium tuberculosis*. *Antimicrob Agents Chemother*. 2021 Sep 1;65 (9).

29. O'Sullivan DM, McHugh TD, Gillespie SH. Analysis of *rpoB* and *pncA* mutations in the published literature: An insight into the role of oxidative stress in *Mycobacterium tuberculosis* evolution? *Journal of Antimicrobial Chemotherapy*. 2005 May;55 (5):674–9.
30. Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, et al. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet*. 2012 Jan;44 (1):106–10.
31. Tagliani E, Anthony R, Kohl TA, De Neeling A, Nikolayevskyy V, Ködmön C, et al. Use of a whole genome sequencing based approach for *Mycobacterium tuberculosis* surveillance in Europe in 2017-2019: An ECDC pilot study. *European Respiratory Journal*. 2021 Jan 1;57 (1).
32. Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, et al. Rapid Molecular Detection of Tuberculosis and Rifampin Resistance. *New England Journal of Medicine*. 2010 Sep 9;363 (11):1005–15.
33. Ligation sequencing gDNA-Native Barcoding Kit 24 V14 (SQK-NBD114.24) VNBE\_9169\_v114\_revR\_15Sep2022.
34. Danchuk SN, Solomon OE, Kohl TA, Dreyer V, Barilar I, Utpatel C, et al. Challenging the gold standard: The limitations of molecular assays for detection of *Mycobacterium tuberculosis* heteroresistance. *Thorax*. 2024;
35. Danchuk SN, Solomon OE, Kohl TA, Dreyer V, Barilar I, Utpatel C, et al. Challenging the gold standard: The limitations of molecular assays for detection of *Mycobacterium tuberculosis* heteroresistance. *Thorax*. 2024 Jun 14;

36. Poole RK. *Advances in microbial physiology*. London: Academic Press; 1998. 421 p.
37. Fodor T. BULK STAINING OF SPUTUM BY THE ZIEHL-NEELSEN METHOD. Vol. 65, *Tubercle*. 1984.
38. Cole, S. T., et al. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*, 393 (6685), 537–544.
39. Meacci F, Orrù G, Iona E, Giannoni F, Piersimoni C, Pozzi G, et al. Drug resistance evolution of a *Mycobacterium tuberculosis* strain from a noncompliant patient. *J Clin Microbiol*. 2005 Jul;43 (7):3114–20.
40. Kanabalan RD, Lee LJ, Lee TY, Chong PP, Hassan L, Ismail R, et al. Human tuberculosis and *Mycobacterium tuberculosis* complex: A review on genetic diversity, pathogenesis and omics approaches in host biomarkers discovery. Vol. 246, *Microbiological Research*. Elsevier GmbH; 2021.
41. Pascual-Pareja JF, Carrillo-Gómez R, Hontañón-Antoñana V, Martínez-Prieto M. Treatment of pulmonary and extrapulmonary tuberculosis. *Enferm Infecc Microbiol Clin*. 2018 Oct 1;36 (8):507–16.
42. Chaves W, Buitrago JF, Dueñas A, Bejarano JC. Acerca de la tuberculosis extrapulmonar. *Repertorio de Medicina y Cirugía*. 2017 Apr;26 (2):90–7.
43. Osorio Aira S, Casadiego Matarranz L, Andrés García N, López Pedreira MR. Miliary tuberculosis induced by intravesical instillation of bacillus Calmette-Guérin. *Radiología (English Edition)*. 2019 Jul;61 (4):337–40.
44. *Global Tuberculosis Report 2021*. World Health Organization; 2021.

45. Seung KJ, Keshavjee S, Rich ML. Multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis. *Cold Spring Harb Perspect Med.* 2015 Sep 1;5 (9).
46. Parsons LM, Somoskövi Á, Gutierrez C, Lee E, Paramasivan CN, Abimiku A, et al. Laboratory diagnosis of tuberculosis in resource-poor Countries: Challenges and opportunities. Vol. 24, *Clinical Microbiology Reviews.* 2011. p. 314–50.
47. Kyu HH, Maddison ER, Henry NJ, Ledesma JR, Wiens KE, Reiner R, et al. Global, regional, and national burden of tuberculosis, 1990–2016: Results from the Global Burden of Diseases, Injuries, and Risk Factors 2016 Study. *Lancet Infect Dis.* 2018 Dec 1;18 (12):1329–49.
48. Riley RL, Mills CC, O’grady F, Sultan LU, Witt Stadt F, Shivpuri DN. INFECTIONOUSNESS OF AIR FROM A TUBERCULOSIS WARD. Ultraviolet Irradiation of Infected Air: COmparative Infectiousness of Different Patients.
49. Fennelly KP, Martyny JW, Fulton KE, Orme IM, Cave DM, Heifets LB. Cough-generated Aerosols of Mycobacterium tuberculosis. *Am J Respir Crit Care Med.* 2004 Mar 1;169 (5):604–9.
50. González-Martín J, García-García JM, Aníbarro L, Vidal R, Esteban J, Blanquer R, et al. Consensus document on the diagnosis, treatment and prevention of tuberculosis. *Enferm Infecc Microbiol Clin.* 2010;28 (5).
51. Machuca I, Vidal E, de la Torre-Cisneros J, Rivero-Román A. Tuberculosis in immunosuppressed patients. *Enferm Infecc Microbiol Clin.* 2018 Jun 1;36 (6):366–74.

52. Alcaide F, Coll P. Advances in rapid diagnosis of tuberculosis disease and anti-tuberculous drug resistance. *Enferm Infecc Microbiol Clin*. 2011 Mar;29 (SUPPL. 1):34–40.
53. Varela L, Ortiz J, García PK, Luna L, Fuertes-Bucheli JF, Pacheco R. Factors associated with unsuccessful treatment outcome for tuberculosis in previously treated patients in Cali, Colombia, during the period 2015-2019. *Biomedica*. 2023;43 (3):360–73.
54. Guía Técnica NY. MANUAL PARA EL DIAGNÓSTICO BACTERIOLÓGICO DE LA TUBERCULOSIS PARTE I BACIIOSCOPIA 2008.
55. Nagamiah S, Selvakumar N, Gomathi Sekar M, Kumar V, Vijaya D, Rao B, et al. Sensitivity of Ziehl-Neelsen method for centrifuged deposit smears of sputum samples transported in cetyl-pyridinium chloride [Internet]. Article in *The Indian Journal of Medical Research*. 2006. Available from: <https://www.researchgate.net/publication/6640957>
56. MacKenzie M, Argyropoulos C. An Introduction to Nanopore Sequencing: Past, Present, and Future Considerations. Vol. 14, *Micromachines*. MDPI; 2023.
57. Peres RL, Palaci M, Loureiro RB, Dietze R, Johnson JL, Golub JE, et al. Evaluation of oral antiseptic rinsing before sputum collection to reduce contamination of mycobacterial cultures. *J Clin Microbiol*. 2011;49 (8):3058-60.

58. Cadmus SIB, Falodun OI, Fagade OE. Methods of sputum decontamination with emphasis on local tuberculosis laboratories. *Afr J Med Med Sci.* 2011;40 (1):5-14.
59. Osbourn A. Saponins and plant defence—a soap story. *Trends Plant Sci.* 1996;1 (1):4–9.
60. Honey-Escandón M, Arreguín-Espinosa R, Solís-Marín FA, Samyn Y. Biological and taxonomic perspective of triterpenoid glycosides of sea cucumbers of the family Holothuriidae (Echinodermata, Holothuroidea). Vol. 180, *Comparative Biochemistry and Physiology Part - B: Biochemistry and Molecular Biology.* Elsevier Inc.; 2015. p. 16–39.
61. Ma Y, Fan J, Li S, Dong L, Li Y, Wang F, et al. Comparison of Lowenstein-Jensen medium and MGIT culture system for recovery of *Mycobacterium tuberculosis* from abscess samples. *Diagn Microbiol Infect Dis.* 2020 Apr 1;96 (4).
62. Xie YL, Chakravorty S, Armstrong DT, Hall SL, Via LE, Song T, et al. Evaluation of a Rapid Molecular Drug-Susceptibility Test for Tuberculosis. *New England Journal of Medicine.* 2017 Sep 14;377 (11):1043–54.
63. WHO operational handbook on tuberculosis.
64. Satana D, Coban AY, Uzun M. Testing susceptibility of multidrug-resistant *Mycobacterium tuberculosis* to second-line drugs by use of blood agar. *J Clin Microbiol.* 2010 Nov;48 (11):4291–3.
65. Zifodya JS, Kreniske JS, Schiller I, Kohli M, Dendukuri N, Schumacher SG, et al. Xpert Ultra versus Xpert MTB/RIF for pulmonary tuberculosis and rifampicin resistance in adults with presumptive pulmonary tuberculosis.

Vol. 2021, Cochrane Database of Systematic Reviews. John Wiley and Sons Ltd; 2021.

66. Bokop C, Faye LM, Apalata T. Analysis of Discordance between Genotypic and Phenotypic Assays for Rifampicin-Resistant *Mycobacterium tuberculosis* Isolated from Healthcare Facilities in Mthatha. *Pathogens*. 2023 Jul 1;12 (7).
67. Mboowa G, Namaganda C, Ssenooba W. Rifampicin resistance mutations in the 81 bp RRDR of *rpoB* gene in *Mycobacterium tuberculosis* clinical isolates using Xpert® MTB/RIF in Kampala, Uganda: a retrospective study [Internet]. 2014. Available from: <http://www.biomedcentral.com/1471-2334/14/481>
68. Vilchèze C, Jacobs WR. The mechanism of isoniazid killing: Clarity through the scope of genetics. Vol. 61, *Annual Review of Microbiology*. 2007. p. 35–50.
69. Gomes LC, Campino S, Marinho CRF, Clark TG, Phelan JE. Whole genome sequencing reveals large deletions and other loss of function mutations in *Mycobacterium tuberculosis* drug resistance genes. *Microb Genom*. 2021;7 (12).
70. Johnsen CH, Clausen PTLC, Aarestrup FM, Lund O. Improved resistance prediction in *Mycobacterium tuberculosis* by better handling of insertions and deletions, premature stop codons, and filtering of non-informative sites. *Front Microbiol*. 2019;10 (OCT).
71. Barnett R, Larson G. A phenol-chloroform protocol for extracting DNA from ancient samples. *Methods in Molecular Biology*. 2012;840:13–9.

72. McGaughey KD, Yilmaz-Swenson T, Elsayed NM, Cruz DA, Rodriguez RR, Kritzer MD, et al. Comparative evaluation of a new magnetic bead-based DNA extraction method from fecal samples for downstream next-generation 16S rRNA gene sequencing. *PLoS One*. 2018 Aug 1;13 (8).
73. Research Corp Z. A Comprehensive Solution for DNA and RNA Clean-Up Sequenced DNA Samples ZR DNA Sequencing Clean-up Kits <sup>TM</sup> Agarose Gel Excisions Zymoclean <sup>TM</sup> Gel DNA Recovery Kits Zymoclean <sup>TM</sup> Gel RNA Recovery Kits DNA Clean-up Oligonucleotide Clean-up RNA Clean-up Large DNA Clean-up DNA Size Selection Zymo-Spin <sup>TM</sup> Technology Ensures High-quality DNA or RNA Low Recovery Following Clean-up with the DCC<sup>TM</sup>, Multiple Bands Appear in an Agarose Gel [Internet]. Available from:  
[www.zymoresearch.com/Overview/Learnmoreatwww.zymoresearch.com/zymo-spin](http://www.zymoresearch.com/Overview/Learnmoreatwww.zymoresearch.com/zymo-spin)Phone:
74. Brlek P, Bulić L, Bračić M, Projić P, Škaro V, Shah N, et al. Implementing Whole Genome Sequencing (WGS) in Clinical Practice: Advantages, Challenges, and Future Perspectives. Vol. 13, *Cells*. Multidisciplinary Digital Publishing Institute (MDPI); 2024.
75. Pei XM, Yeung MHY, Wong ANN, Tsang HF, Yu ACS, Yim AKY, et al. Targeted Sequencing Approach and Its Clinical Applications for the Molecular Diagnosis of Human Diseases. Vol. 12, *Cells*. MDPI; 2023.
76. Ligation sequencing gDNA (Q-SQK-LSK109).
77. Dou HY, Huang TS, Wu HC, Hsu CH, Chen FJ, Liao YC. Targeted sputum sequencing for rapid and broad drug resistance of *Mycobacterium*

- tuberculosis. *Infection* [Internet]. 2025 Jan 16; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/39821740>
78. Mariner-Llicer C, Goig GA, Torres-Puente M, Vashakidze S, Villamayor LM, Saavedra-Cervera B, et al. Genetic diversity within diagnostic sputum samples is mirrored in the culture of *Mycobacterium tuberculosis* across different settings. *Nature Communications*. 2024 Dec 1;15 (1).
  79. Metcalfe JZ, Streicher E, Theron G, Colman RE, Penaloza R, Allender C, et al. *Mycobacterium tuberculosis* Subculture Results in Loss of Potentially Clinically Relevant Heteroresistance [Internet]. 2017. Available from: <https://journals.asm.org/journal/aac>
  80. Molina Rueda MJ, Martín-Vivaldi Jiménez A, Molina Rueda F. Tratamiento de la tuberculosis: ¿qué pacientes y por qué lo abandonan? Vol. 44, *Atencion Primaria*. 2012.
  81. Zhang Y, Yew WW. Mechanisms of drug resistance in *Mycobacterium tuberculosis*: Update 2015. Vol. 19, *International Journal of Tuberculosis and Lung Disease*. International Union Against Tuberculosis and Lung Disease; 2015. p. 1276–89.
  82. Seifert M, Catanzaro D, Catanzaro A, Rodwell TC. Genetic mutations associated with isoniazid resistance in *Mycobacterium tuberculosis*: A systematic review. *PLoS One*. 2015 Mar 23;10 (3).
  83. Bakhtiyariniya P, Khosravi AD, Hashemzadeh M, Savari M. Detection and characterization of mutations in genes related to isoniazid resistance in *Mycobacterium tuberculosis* clinical isolates from Iran. *Mol Biol Rep*. 2022 Jul 1;49 (7):6135–43.

84. Jagielski T, Bakuła Z, Roeske K, Kamiński M, Napiórkowska A, Augustynowicz-Kopec E, et al. Detection of mutations associated with isoniazid resistance in multidrug-resistant *Mycobacterium tuberculosis* clinical isolates. *Journal of Antimicrobial Chemotherapy*. 2014;69 (9):2369–75.
85. Hofmann-Thiel S, Van Ingen J, Feldmann K, Turaev L, Uzakova GT, Murmusaeva G, et al. Mechanisms of heteroresistance to isoniazid and rifampin of *Mycobacterium tuberculosis* in Tashkent, Uzbekistan. *European Respiratory Journal*. 2009 Feb;33 (2):368–74.
86. Mitchison DA, Davies GR. Assessment of the Efficacy of New Anti-Tuberculosis Drugs. *Open Infect Dis J*. 2009 Jan 5;2 (1):59–76.
87. Zhang Y, Shi W, Zhang W, Mitchison D. Mechanisms of Pyrazinamide Action and Resistance. *Microbiol Spectr*. 2014 Aug 15;2 (4).
88. Aung YW, Faksri K, Sangka A, Tomanakan K, Namwat W. Heteroresistance of *Mycobacterium tuberculosis* in the Sputum Detected by Droplet Digital PCR. *Biology (Basel)*. 2023 Apr 1;12 (4).
89. Mei Z, Sun Z, Bai D, Xu Y, Li Z, Huang H, et al. Discrepancies in Drug Susceptibility Test for Tuberculosis Patients Resulted from the Mixed Infection and the Testing System. *Biomed Res Int*. 2015;2015.
90. Zheng C, Li S, Luo Z, Pi R, Sun H, He Q, et al. Mixed infections and rifampin heteroresistance among *mycobacterium tuberculosis* clinical isolates. *J Clin Microbiol*. 2015 Jul 1;53 (7):2138–47.
91. Peñuelas-Urquides K, Villarreal-Treviño L, Silva-Ramírez B, Rivadeneyra-Espinoza L, Said-Fernández S, Bermúdez De León M. Measuring of

- Mycobacterium tuberculosis growth. A correlation of the optical measurements with colony forming units [Internet]. 2013. Available from: [www.sbmicrobiologia.org.br](http://www.sbmicrobiologia.org.br)
92. Julia D, María R, Vidal R, Ejecutiva D. Situación y desafíos en el control de la TB en el Perú.
  93. Dippenaar A, Ismail N, Grobbelaar M, Oostvogels S, de Vos M, Streicher EM, et al. Optimizing liquefaction and decontamination of sputum for DNA extraction from Mycobacterium tuberculosis. Tuberculosis. 2022 Jan 1;132.
  94. Kok NA, Peker N, Schuele L, de Beer JL, Rossen JWA, Sinha B, et al. Host DNA depletion can increase the sensitivity of Mycobacterium spp. detection through shotgun metagenomics in sputum. Front Microbiol. 2022 Oct 25;13.
  95. Prajwal P, Neary T, Rohrbach K, Bittel P, Göller PC, Buch T, et al. Optimizing mycobacteria molecular diagnostics: No decontamination! Human DNA depletion? Greener storage at 4 °C! Front Microbiol. 2023;14.
  96. Longhi G, Argentini C, Fontana F, Tarracchini C, Mancabelli L, Lugli GA, et al. Saponin treatment for eukaryotic DNA depletion alters the microbial DNA profiles by reducing the abundance of Gram-negative bacteria in metagenomics analyses. Microbiome Research Reports. 2023;2 (1).
  97. Desjardin LE, Perkins MD, Teixeira L, Cave MD, Eisenach KD. Alkaline Decontamination of Sputum Specimens Adversely Affects Stability of Mycobacterial mRNA. Vol. 34, JOURNAL OF CLINICAL MICROBIOLOGY. 1996.

98. Prajwal P, Neary T, Rohrbach K, Bittel P, Göller PC, Buch T, et al. Optimizing mycobacteria molecular diagnostics: No decontamination! Human DNA depletion? Greener storage at 4 °C! *Front Microbiol.* 2023;14.
99. Kok NA, Peker N, Schuele L, de Beer JL, Rossen JWA, Sinha B, et al. Host DNA depletion can increase the sensitivity of *Mycobacterium* spp. detection through shotgun metagenomics in sputum. *Front Microbiol.* 2022 Oct 25;13.
100. Latham S, Hughes E, Budgen B, Morley A. Inhibition of the PCR by genomic DNA. *PLoS One.* 2023 Apr 1;18 (4 April).
101. Neves De Almeida I, Da W, Carvalho S, Rossetti ML, Dalla Costa R, Spindola De Miranda S. Evaluation of six different DNA extraction methods for detection of *Mycobacterium tuberculosis* by means of PCR-IS6110: preliminary study [Internet]. 2013. Available from: <http://www.biomedcentral.com/1756-0500/6/561>
102. Oriano M, Terranova L, Teri A, Sottotetti S, Ruggiero L, Tafuro C, et al. Comparison of different conditions for DNA extraction in sputum - A pilot study. *Multidiscip Respir Med.* 2019 Jan 31;14 (1).
103. Nazlı Arslan. Comparison of Four Different DNA Isolation Methods from MGIT Culture for Long-Read Whole Genome Sequencing of *Mycobacterium tuberculosis*.
104. Claycamp HG. Phenol sensitization of DNA to subsequent oxidative damage in 8-hydroxyguanine assays [Internet]. Vol. 13. 1992. Available from: <http://carcin.oxfordjournals.org/>
105. Srirungruang S, Mahajindawong B, Nimitpanya P, Bunkasem U, Ayuyoe P, Nuchprayoon S, et al. Comparative Study of DNA Extraction Methods for

- the PCR Detection of Intestinal Parasites in Human Stool Samples. *Diagnostics*. 2022 Nov 1;12 (11).
106. Filges S, Yamada E, Ståhlberg A, Godfrey TE. Impact of Polymerase Fidelity on Background Error Rates in Next-Generation Sequencing with Unique Molecular Identifiers/Barcodes. *Sci Rep*. 2019 Dec 1;9 (1).
  107. Brandariz-Fontes C, Camacho-Sanchez M, Vilà C, Vega-Pla JL, Rico C, Leonard JA. Effect of the enzyme and PCR conditions on the quality of high-throughput DNA sequencing results. *Sci Rep*. 2015;5:8056.
  108. Guido N, Starostina E, Leake D, Saaem I. Improved PCR amplification of broad spectrum GC DNA templates. *PLoS One*. 2016 Jun 1;11 (6).
  109. Cauwelaert ND, Ramarokoto H, Ravololonandriana P, Richard V, Rasolofo V. DNA extracted from stained sputum smears can be used in the MTBDRplus assay. *J Clin Microbiol*. 2011 Oct;49 (10):3600–3.
  110. Dou HY, Huang TS, Wu HC, Hsu CH, Chen FJ, Liao YC. Targeted sputum sequencing for rapid and broad drug resistance of *Mycobacterium tuberculosis*. *Infection* [Internet]. 2025 Jan 16; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/39821740>
  111. Sauvage T, Cormier A, Delphine P. A comparison of Oxford nanopore library strategies for bacterial genomics. *BMC Genomics*. 2023 Dec 1;24 (1).
  112. Brandt C, Bongcam-Rudloff E, Müller B. Abundance tracking by long-read nanopore sequencing of complex microbial communities in samples from 20 different biogas/wastewater plants. *Applied Sciences (Switzerland)*. 2020 Nov 1;10 (21).

113. Maghini DG, Moss EL, Vance SE, Bhatt AS. Improved high-molecular-weight DNA extraction, nanopore sequencing and metagenomic assembly from the human gut microbiome. *Nat Protoc.* 2021 Jan 1;16 (1):458–71.
114. Gamaarachchi H, Samarakoon H, Jenner SP, Ferguson JM, Amos TG, Hammond JM, et al. Fast nanopore sequencing data analysis with SLOW5. *Nat Biotechnol.* 2022 Jul 1;40 (7):1026–9.
115. Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. Nanopore sequencing technology, bioinformatics and applications. Vol. 39, *Nature Biotechnology.* Nature Research; 2021. p. 1348–65.
116. Chen Z, Ong CT, Nguyen LT, Lamb HJ, González-Recio O, Gutiérrez-Rivas M, et al. Biases from Oxford Nanopore library preparation kits and their effects on microbiome and genome analysis. *BMC Genomics.* 2025 Dec 1;26 (1).
117. Tyler AD, Christianson S, Knox NC, Mabon P, Wolfe J, Van Domselaar G, et al. Comparison of sample preparation methods used for the next-generation sequencing of mycobacterium tuberculosis. *PLoS One.* 2016 Feb 1;11 (2).
118. Wick RR, Judd LM, Wyres KL, Holt KE. Recovery of small plasmid sequences via oxford nanopore sequencing. *Microb Genom.* 2021;7 (8).
119. Player R, Verratti K, Staab A, Forsyth E, Ernlund A, Joshi MS, et al. Optimization of Oxford Nanopore Technology Sequencing Workflow for Detection of Amplicons in Real Time Using ONT-DART Tool. *Genes (Basel).* 2022 Oct 1;13 (10).

120. Pastusiak A, Reddy MR, Chen X, Hoyer I, Dorman J, Gebhardt ME, et al. A metagenomic analysis of the phase 2 *Anopheles gambiae* 1000 genomes dataset reveals a wide diversity of cobionts associated with field collected mosquitoes. *Commun Biol.* 2024 Dec 1;7 (1).
121. Kok NA, Peker N, Schuele L, de Beer JL, Rossen JWA, Sinha B, et al. Host DNA depletion can increase the sensitivity of *Mycobacterium* spp. detection through shotgun metagenomics in sputum. *Front Microbiol.* 2022 Oct 25;13.
122. Anthony WE, Allison SD, Broderick CM, Chavez Rodriguez L, Clum A, Cross H, et al. From soil to sequence: filling the critical gap in genome-resolved metagenomics is essential to the future of soil microbial ecology. Vol. 19, *Environmental Microbiome*. BioMed Central Ltd; 2024.
123. Nimmo C, Brien K, Millard J, Grant AD, Padayatchi N, Pym AS, et al. Dynamics of within-host *Mycobacterium tuberculosis* diversity and heteroresistance during treatment. *EBioMedicine.* 2020 May 1;55.
124. Santucci P, Greenwood DJ, Fearn A, Chen K, Jiang H, Gutierrez MG. Intracellular localisation of *Mycobacterium tuberculosis* affects efficacy of the antibiotic pyrazinamide. *Nat Commun.* 2021 Dec 1;12 (1).
125. Zheng C, Li S, Luo Z, Pi R, Sun H, He Q, et al. Mixed infections and rifampin heteroresistance among *mycobacterium tuberculosis* clinical isolates. *J Clin Microbiol.* 2015 Jul 1;53 (7):2138–47.
126. Asim Khan M, Meravi K. ORIGINAL RESEARCH Heteroresistance to rifampicin & isoniazid in clinical samples of patients with presumptive drug-resistant tuberculosis.

127. Werngren J, Mansjö M, Glader M, Hoffner S, Forsman LD. Detection of Pyrazinamide heteroresistance in mycobacterium tuberculosis. *Antimicrob Agents Chemother*. 2021 Sep 1;65 (9).
128. Whitfield MG, Engelthaler DM, Allender C, Folkerts M, Heupink TH, Limberis J, et al. Comparative Performance of Genomic Methods for the Detection of Pyrazinamide Resistance and Heteroresistance in Mycobacterium tuberculosis [Internet]. 2022. Available from: <https://www.R-project.org>;
129. Guliaev A, Hjort K, Rossi M, Jonsson S, Nicoloff H, Guy L, et al. Machine learning detection of heteroresistance in Escherichia coli. *EBioMedicine*. 2025 Mar 1;113.
130. Zymoresearch. Zymo Quick DNA Plus. 2018.
131. Mtafya B, Qwaray P, John J, Sichone E, Shoo A, Gillespie SH, et al. A practical approach to render tuberculosis samples safe for application of tuberculosis molecular bacterial load assay in clinical settings without a biosafety level 3 laboratory. *Tuberculosis*. 2023 Jan 1;138.

## 15. ANEXOS.

### **Anexo 1.**

**Descontaminación de muestras de esputo con saponina.** Aproximadamente 1 mL de las muestras de esputo fueron homogenizado con un agitador VORTEX-GENIE2 a 2500 RPM durante 1 min y mantenido en reposo por 2 min para reducir la formación de aerosoles. La muestra fue dividida en 2 alícuotas de 500  $\mu$ L, cada alícuota fue centrifugada a 13 000 RPM por 10 min, el sobrenadante fue descartado y el sedimento fue resuspendido y homogenizado mediante pipeteo con 625  $\mu$ L de Buffer fosfato salino (PBS). Luego, 500  $\mu$ L de saponina (5% concentración final) fue agregada (24). La mezcla fue incubada durante 10 min a temperatura ambiente y homogenizada cada 3 min. Finalmente 200  $\mu$ L de agua y 30  $\mu$ L de NaCl fue agregados para desnaturalizar proteínas y neutralizar la carga negativa del ADN humano en el sobrenadante, el cual fue descartado.

El sedimento formado que contenía a las micobacterias en los 2 tubos fue mezclado y lavado 3 veces con 1000, 800 y 400  $\mu$ L de PBS, respectivamente y centrifugado a 6000g por 3 min. El sedimento fue resuspendido en 400  $\mu$ L de PBS y divididos en 2 alícuotas de 200  $\mu$ L para ser usados en los 2 métodos de extracción de ADN (23).

### **Anexo 2.**

**Descontaminación de muestras de esputo mediante NaOH – NALC al 0.5%.** La solución stock de descontaminación contiene NaOH y citrato de sodio al 4% y 2.9%, respectivamente. Para preparar la solución NaOH-NALC al 0.5% se

tomó 12 mL de la solución stock y se añadió 0.6 g de N-acetil-L-cisteína (NALC) agitando levemente para su uso inmediato.

1 mL de esputo fue mezclado con 1 mL de NaOH-NALC, homogenizado mediante vortex e incubado por 15 min con agitaciones leves cada 5 min. Luego, PBS fue agregado hasta 15 mL y homogenizado suavemente y centrifugado a 3000 RPM por 15 min. Luego, el sobrenadante fue descartado y el sedimento fue resuspendido en 1.6 mL de PBS. Este volumen fue distribuido en 7 alícuotas de 200  $\mu$ L.

### **Anexo 3.**

**Método de extracción de ADN: Fenol.** A 200  $\mu$ L de esputo descontaminado se le agregó 200  $\mu$ L de fenol:cloroformo:alcohol isoamílico, la mezcla fue homogenizada por pipeteo e incubada por 1 hora para inactivar a las micobacterias. Toda la mezcla fue transferida a un tubo con 100 mg de beads de sílice 0.1 mm sellado con parafilm. La muestra fue homogenizada usando el FastPrep-24TM 5G por 2 veces por 40 sec a 5.5 metros/segundo. con intervalos de 5 minutos en hielo entre agitaciones. La muestra fue centrifugada a 14000 rpm por 10 min; el sobrenadante conteniendo el ADN fue recuperado. Un volumen de cloroformo: alcohol isoamílico (24:1) igual al volumen del sobrenadante fue agregado. La mezcla fue homogenizada mediante vórtex y centrifugado a 14 000 RPM por 10 min. El sobrenadante fue transferido a otro tubo y un volumen de AMPure XP beads fue agregado, previamente descongelado por 30 min, e incubado por 15 min. La mezcla conteniendo al ADN unido a los AMPure XP beads fue lavada 3 veces con 1mL de etanol al 80% usando un rack magnético. El AMPure XP beads unido al ADN fue eluido con 15-30  $\mu$ L de Buffer TE (10 mM Tris, pH 8.5, 0.1 mM EDTA) a 37°C.

#### **Anexo 4.**

**Método de extracción de ADN: Kit.** A los 200  $\mu\text{L}$  de muestra de esputo descontaminado se agregó 200  $\mu\text{L}$  de BioFluid & Cell Buffer (130). El tiocianato contenido en el BioFluid & Cell Buffer inactiva a los micobacterios (131). La mezcla fue homogenizada mediante pipeteo e incubada por 1 hora a temperatura ambiente. Toda la mezcla fue transferida a un tubo con 0.1 g de beads de sílice de 0.1 mm y sellada con parafilm. Las micobacterias fueron lisadas en el FastPrep-24<sup>TM</sup>5G con programación de 2 veces por 40 sec a 5.5 m/seg, con intervalos de 5 min en hielo, para luego ser centrifugada a 12 000 g por 10 min. El sobrenadante fue transferido a una columna Zymo-Spin<sup>TM</sup> IIC-XLR dentro de un tubo de recolección (Quick-DNA<sup>TM</sup> Miniprep Plus Kit); luego 400  $\mu\text{L}$  de Genomic Binding Buffer fue agregado. La mezcla fue centrifugada a 12 000 g por 1 min. El ADN fue lavado 2 veces con 400  $\mu\text{L}$  y 700  $\mu\text{L}$  de DNA- Pre Wash Buffer y g-DNA Wash Buffer respectivamente. con centrifugaciones de 12 000 g x 1 min y un último lavado con 200  $\mu\text{L}$  de g-DNA Wash Buffer. El ADN fue incubado por 5 min a temperatura ambiente y eluido con 30  $\mu\text{L}$  de DNA elution buffer.

#### **Anexo 5.**

**Biblioteca de ADN genómico con *Native Ligation kit*.** La reparación y preparación de los extremos de ADN fue realizado de la siguiente manera: A 400 ng de ADN en 11  $\mu\text{L}$  agua PCR fue mezclada con 1  $\mu\text{L}$  de Dilued DNA Control Sample. 0.875  $\mu\text{L}$  de NEBNext FFPE DNA Repair Buffer, 0.875  $\mu\text{L}$  de Ultra II End-prep Reaction Buffer, 0.75  $\mu\text{L}$  de Ultra II End-prep Enzyme Mix y 0.5  $\mu\text{L}$  de NEBNext FFPE DNA Repair Mix. La mezcla fue incubada a 20°C por 5 min seguido de 65°C por 5 min. Luego, 15  $\mu\text{L}$  de AMPure XP Beads fue agregado y

homogenizada mediante vórtex e incubado en un Hula mixer por 5 min a temperatura ambiente. Luego, las muestras fueron lavadas 3 veces con etanol al 80% usando un soporte magnético. El ADN fue eluido con 10  $\mu$ L de agua y cuantificado. Luego, los “barcodes” fueron agregados a una cantidad de ADN de cada muestra de esputo que previamente fue normalizada con la muestra con menor cantidad de ADN en 7.5  $\mu$ L de agua PCR. 10  $\mu$ L de la enzima Blunt/TA Ligase Master Mix y 2.5  $\mu$ L de “barcode” fue agregado e incubado por 20 min a temperatura ambiente. Luego, 2  $\mu$ L de EDTA fue agregado.

Las 12 muestras de ADN con sus “barcodes” fueron mezcladas (librería); 0.4X de AMPure XP Beads fue agregado e incubado por 10 min en un Hula Mixer a temperatura ambiente. Lo mismo se realizó para las 12 muestras de ADN de los cultivos primarios. Las librerías fueron colocadas en un rack magnético donde se realizaron 3 lavados con etanol al 80%. El ADN fue eluido con 35  $\mu$ L de agua PCR y cuantificado.

Finalmente, se adicionaron los adaptadores. A 30  $\mu$ L de librería se le agregó 5  $\mu$ L de Native Adapter, 10  $\mu$ L de NEBNext Quick Ligation Reaction Buffer (5X) y 5  $\mu$ L de Quick T4 DNA Ligase e incubado a temperatura ambiente por 20 min. 20  $\mu$ L de AMPure XP Beads fue agregado e incubado en el Hula Mixer por 10 min. Luego, las muestras fueron lavadas 2 veces usando el rack magnético con 125  $\mu$ L de Short Fragment Buffer (SFB). Las librerías fueron eluidas con 15  $\mu$ L de buffer de elución. La cantidad de ADN de cada librería fue cuantificada mediante Qubit y el tamaño, longitud de los fragmentos de ADN fue determinado mediante el Tape Station. Las librerías fueron guardadas a 4°C por 2 horas hasta la secuenciación.

## **Anexo 6.**

***Biblioteca de ADN genómico con Rapid barcoding Kit.*** A 250 ng de ADN se le agregó “barcodes”. El ADN, ajustado a 9  $\mu\text{L}$  de agua libre de nucleasas, fue mezclado con 1  $\mu\text{L}$  del mix de fragmentación (barcodes) del *Rapid barcoding kit 24 V14* (SQK-RBK114.24) y centrifugado con un pulso e incubado en un termociclador a 30°C por 2 min, 80°C por 2 min y luego, enfriada en hielo. Las 12 muestras de ADN de esputo con su “barcode” fueron mezcladas (librería). Una librería adicional fue generada para las 12 muestras provenientes de cultivos primarios. Las librerías fueron suspendidas con AMPure XP Beads y homogenizados mediante suave pipeteo. Las librerías con los AMPure XP Beads fueron lavados con etanol al 80% y eluidos con 15  $\mu\text{L}$ . 800 ng de ADN como máximo fue completado con buffer de elución hasta 11  $\mu\text{L}$ . Luego, 1  $\mu\text{L}$  de adaptador preparado (1.5  $\mu\text{L}$  de rapid adapter y 3.5  $\mu\text{L}$  de buffer adapter) fue agregado a la librería y mezclado suavemente. La cantidad de ADN de cada librería fue cuantificada mediante Qubit y el tamaño, longitud de los fragmentos de ADN fue determinado mediante Tape Station. Las librerías fueron guardadas a 4°C por 2 horas hasta la secuenciación.

## **Anexo 7.**

**Estandarización y obtención de amplicones a partir de muestras de esputo y cultivo de *M. tuberculosis* (14 genes entre 876 y 3,662 pb.)** Cada reacción de PCR por pool contenía: 12.5  $\mu\text{L}$  de Phusion® High-Fidelity PCR Master Mix with HF Buffer (2X), 0.75  $\mu\text{L}$  de dimetilsulfóxido (DMSO), 3.25  $\mu\text{L}$  de agua, 50 ng de ADN en 5  $\mu\text{L}$  y 3.5  $\mu\text{L}$  del pool de primers correspondiente.

Pool 1. 0.35  $\mu\text{L}$ , 1.75  $\mu\text{L}$ , 0.225  $\mu\text{L}$ , 0.9  $\mu\text{L}$  y 0.2  $\mu\text{L}$  de los primers *mabA-inhA*, *rpoB*, *ethA*, *katG* y *rpsL* forward y reverse con concentraciones finales de 0.14  $\mu\text{M}$ , 0.7  $\mu\text{M}$ , 0.09  $\mu\text{M}$ , 0.36  $\mu\text{M}$  y 0.08  $\mu\text{M}$  respectivamente.

Pool 2. 2  $\mu\text{L}$ , 0.7  $\mu\text{L}$ , 0.325  $\mu\text{L}$ , 0.2  $\mu\text{L}$  y 0.15  $\mu\text{L}$  de los primers *embB*, *gyrB*, *rrs*, *oxyR* y *pncA* forward y reverse con concentraciones finales de 0.8  $\mu\text{M}$ , 0.28  $\mu\text{M}$ , 0.13  $\mu\text{M}$ , 0.08  $\mu\text{M}$  y 0.06  $\mu\text{M}$  respectivamente.

Pool 3. 0.2  $\mu\text{L}$ , 0.825  $\mu\text{L}$ , 0.3  $\mu\text{L}$  y 0.225  $\mu\text{L}$  de los primers *bedA*, *gyrA*, *eis* y *embC-embA* forward y reverse con concentraciones finales de 0.08  $\mu\text{M}$ , 0.33  $\mu\text{M}$ , 0.12  $\mu\text{M}$  y 0.09  $\mu\text{M}$  respectivamente.

#### **Anexo 8.**

Cinco volúmenes de DNA binding buffer fueron agregados a un volumen de amplicones. La mezcla fue homogenizada brevemente mediante vortex y transferida a una columna Zymo-Spin™. La columna fue colocada dentro de un tubo de colección y centrifugada a 13 000 RPM por 30 seg, el eluido fue descartado. El ADN fue lavado 2 veces con 200  $\mu\text{l}$  de DNA wash buffer. Finalmente, los amplicones fueron eluidos con 15  $\mu\text{l}$  de DNA elution buffer.

#### **Anexo 9.**

**Biblioteca de amplicones con *Rapid barcoding Kit*.** A 50 ng de ADN se le agregó “barcodes”. El ADN, ajustado a 9  $\mu\text{L}$  de agua libre de nucleasas, fue mezclado con 1  $\mu\text{L}$  del mix de fragmentación (barcodes) del *Rapid barcoding kit 24 V14* (SQK-RBK114.24) y centrifugado con un pulso e incubado en un termociclador a 30°C por 2 min, 80°C por 2 min y luego, enfriada en hielo. Las 12 muestras de ADN de esputo con su “barcode” fueron mezcladas (librería). Una librería adicional fue generada para las 12 muestras provenientes de cultivos primarios. Las librerías

fueron suspendidas con AMPure XP Beads y homogenizados mediante suave pipeteo. Las librerías con los AMPure XP Beads fueron lavados con etanol al 80% y eluidos con 15  $\mu$ L. 800 ng de ADN como máximo fue completado con buffer de elución hasta 11  $\mu$ L. Luego, 1  $\mu$ L de adaptador preparado (1.5  $\mu$ L de rapid adapter y 3.5  $\mu$ L de buffer adapter) fue agregado a la librería y mezclado suavemente. La cantidad de ADN de cada librería fue cuantificada mediante Qubit y el tamaño, longitud de los fragmentos de ADN fue determinado mediante Tape Station. Las librerías fueron guardadas a 4°C por 2 horas la secuenciación.

#### **Anexo 10.**

**Biblioteca de amplicones con *Native Ligation kit*.** La reparación y preparación de los extremos de ADN fue realizado de la siguiente manera: A 150 ng de ADN en 11  $\mu\text{L}$  agua PCR fue mezclada con 1  $\mu\text{L}$  de Dilued DNA Control Sample. 0.875  $\mu\text{L}$  de NEBNext FFPE DNA Repair Buffer, 0.875  $\mu\text{L}$  de Ultra II End-prep Reaction Buffer, 0.75  $\mu\text{L}$  de Ultra II End-prep Enzyme Mix y 0.5  $\mu\text{L}$  de NEBNext FFPE DNA Repair Mix. La mezcla fue incubada a 20°C por 5 min seguido de 65°C por 5 min. Luego, 15  $\mu\text{L}$  de AMPure XP Beads fue agregado y homogenizada mediante vórtex e incubado en un Hula mixer por 5 min a temperatura ambiente. Luego, las muestras fueron lavadas 3 veces con etanol al 80% usando un soporte magnético. El ADN fue eluido con 10  $\mu\text{L}$  de agua y cuantificado. Luego, los “barcodes” fueron agregados a una cantidad de ADN de cada muestra de esputo que previamente fue normalizada con la muestra con menor cantidad de ADN en 7.5  $\mu\text{L}$  de agua PCR. 10  $\mu\text{L}$  de la enzima Blunt/TA Ligase Master Mix y 2.5  $\mu\text{L}$  de “barcode” fue agregado e incubado por 20 min a temperatura ambiente. Luego, 2  $\mu\text{L}$  de EDTA fue agregado.

Las 12 muestras de ADN con sus “barcodes” fueron mezcladas (librería); 0.4X de AMPure XP Beads fue agregado e incubado por 10 min en un Hula Mixer a temperatura ambiente. Lo mismo se realizó para las 12 muestras de ADN de los cultivos primarios. Las librerías fueron colocadas en un rack magnético donde se realizaron 3 lavados con etanol al 80%. El ADN fue eluído con 35  $\mu\text{L}$  de agua PCR y cuantificado.

Finalmente, se adicionaron los adaptadores. A 30  $\mu\text{L}$  de librería se le agregó 5  $\mu\text{L}$  de Native Adapter, 10  $\mu\text{L}$  de NEBNext Quick Ligation Reaction Buffer (5X) y 5  $\mu\text{L}$  de Quick T4 DNA Ligase e incubado a temperatura ambiente por 20 min. 20

$\mu\text{L}$  de AMPure XP Beads fue agregado e incubado en el Hula Mixer por 10 min. Luego, las muestras fueron lavadas 2 veces usando el rack magnético con 125  $\mu\text{L}$  de Short Fragment Buffer (SFB). Las librerías fueron eluidas con 15  $\mu\text{L}$  de buffer de elución. La cantidad de ADN de cada librería fue cuantificada mediante Qubit y el tamaño, longitud de los fragmentos de ADN fue determinado mediante el Tape Station. Las librerías fueron guardadas a 4°C por 2 horas hasta la secuenciación.

### **Anexo 11.**

**Preparación de medios de cultivo y reactivación de cepas bacterianas.** Las cepas de *M. tuberculosis* H37Rv y DM97 fueron descongeladas y reactivadas en 5 mL de medio líquido 7H9 enriquecido con 10% de OADC (ácido oleico, dextrosa, catalasa y albúmina), 0.2% de glicerol y 0.05% de Tween-OADC a 37°C con agitación constante (180 RPM) por 2 semanas. Luego, las cepas fueron subcultivadas en 3 placas 7H10 enriquecido con 5% de OADC, 0.2% de glicerol y 0.05% de Tween-OADC durante 10 – 15 días. Las placas que contenían medio 7H11-OADC también fueron suplementadas con OADC al 5%, junto con la adición de Tween al 0.05%. Para evitar el crecimiento de microorganismos contaminantes, se incorporó un coctel de antibióticos compuesto por polimixina B (200 U/mL), anfotericina B (10  $\mu\text{g}/\text{mL}$ ), carbenicilina (50  $\mu\text{g}/\text{mL}$ ) y trimetoprima (20  $\mu\text{g}/\text{mL}$ ). Además, se adicionó un antibiótico específico en cada uno de los cuadrantes de la placa para evaluar la susceptibilidad. En uno de los cuadrantes se incluyó isoniacida a una concentración de 0.4  $\mu\text{g}/\text{mL}$ , en otro, se agregó rifampicina 1  $\mu\text{g}/\text{mL}$ ; y en un tercer cuadrante se agregó pirazinamida a 100  $\mu\text{g}/\text{mL}$ . El cuarto cuadrante se mantuvo sin antibiótico, sirviendo como control para comparar el crecimiento bacteriano en ausencia de fármacos.

## Anexo 12.

**Preparación de mezclas bacterianas heterorresistentes.** Una asada de cultivo bacteriano fue mezclado con 100 µL de Tween 0.04% con agua estéril y 3-4 perlas de zirconia de 2.3 mm. La suspensión fue homogenizada con el vortex durante 5 min; luego se dejó reposar 10 min. A la suspensión, 5 mL de Tween 0.04% fue agregada, homogenizada con el vortex; y luego se dejó reposar por 10 min. En un tubo adicional con 8 mL de Tween 0.04% se fue agregando de 100 -300 µL de la suspensión con cepas hasta llegar a un OD 0.2, que equivale a  $1 \times 10^6$  ufc/mL (91)

Tabla S1. Concentración y calidad de ADN extraída mediante los métodos de descontaminación NALC/NaOH y saonina

| Muestra | NALC/NaOH              |         |         | Saponina 5%            |         |         |
|---------|------------------------|---------|---------|------------------------|---------|---------|
|         | Concentración (ng/ uL) | 260/280 | 260/230 | Concentración (ng/ uL) | 260/280 | 260/230 |
| 1       | 10.2                   | 0.82    | 3.1     | 279                    | 1.84    | 2.19    |
| 2       | 8.4                    | 2.47    | 0.55    | 186                    | 1.85    | 2.12    |
| 3       | 10.6                   | 1.15    | 1.02    | 272                    | 1.82    | 2.11    |
| 4       | 7.6                    | 0.6     | 4.2     | 172                    | 1.84    | 2.11    |
| 5       | 10                     | 1.9     | 0.4     | 303                    | 1.83    | 2.11    |
| 6       | 9.1                    | 3.05    | 0.78    | 333                    | 1.84    | 2.13    |
| 7       | 8.6                    | 0.95    | 2.65    | 181                    | 1.84    | 2.12    |



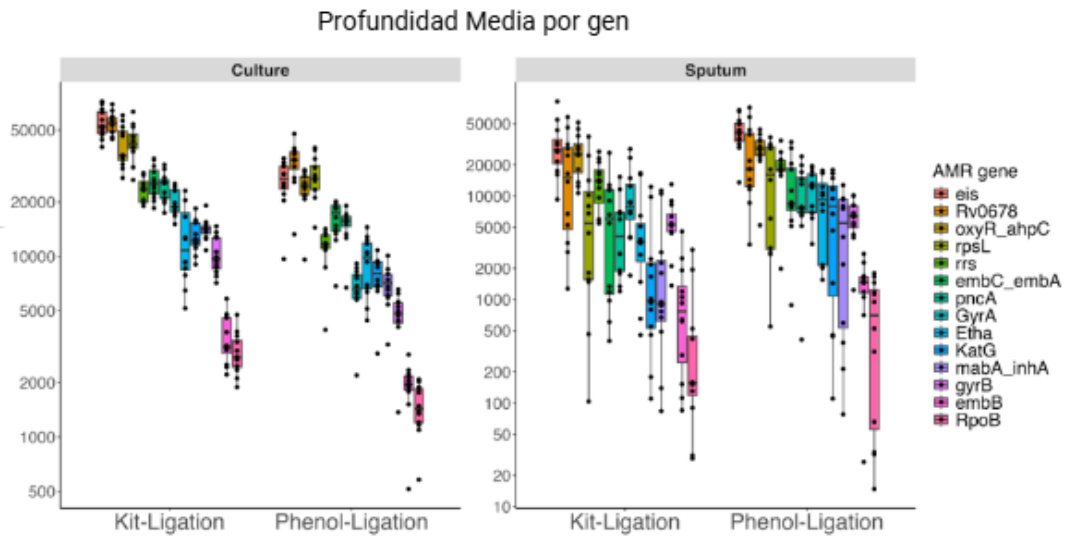


Figura S3. Profundidad media del panel de 14 genes para muestras de cultivo y esputo de acuerdo con el método de extracción. Los genes de mayor tamaño presentaron los menores valores aunque mayores a 1000x.

### Frecuencia alélica en la mezcla de cepas

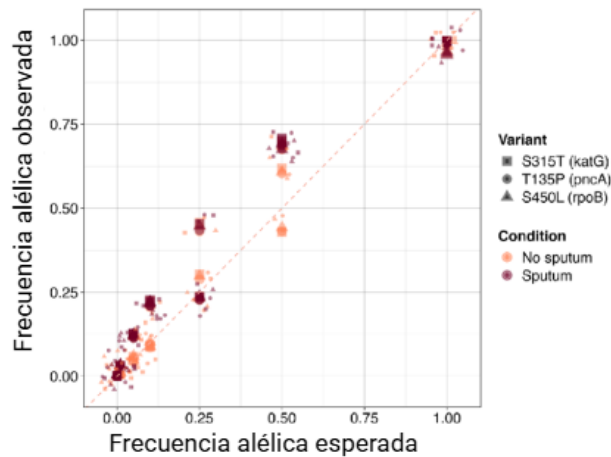


Figura S4. Gráfico de dispersión de las frecuencias alélicas observadas versus las esperadas para variantes seleccionadas asociadas a resistencia (*katG* S315T, *pncA* T135P, *rpoB* S450L) en diferentes mezclas, destacando la concordancia bajo condiciones con esputo (púrpura) y sin esputo (naranja).

Tabla S2. Mutaciones asociadas a la cepa resistente de *Mycobacterium tuberculosis* DM97

Mutaciones asociadas a la cepa DM97

| Genome position | Locus tag | Gene name | Change      | Drug         |
|-----------------|-----------|-----------|-------------|--------------|
| 761155          | Rv0667    | rpoB      | p.Ser450Leu | rifampicin   |
| 2155168         | Rv1908c   | katG      | p.Ser315Thr | isoniazid    |
| 2288839         | Rv2043c   | pncA      | p.Thr135Pro | pyrazinamide |
| 4247574         | Rv3795    | embB      | p.Asp354Ala | ethambutol   |

Tabla S3. Concentración, cantidad de ADN y ratios de calidad de 33 muestras de esputo mediante el método fenol y kit.

| N° de muestra | Concentración de ADN (ng/μL) |      | Cantidad de ADN total (ng) |      | 260/280 |      | 260/230 |      |
|---------------|------------------------------|------|----------------------------|------|---------|------|---------|------|
|               | Fenol                        | Kit  | Fenol                      | Kit  | Fenol   | Kit  | Fenol   | Kit  |
| 1             | 604                          | 59.2 | 30200                      | 2960 | 1.89    | 1.88 | 2.32    | 2.13 |
| 2             | 342                          | 101  | 17100                      | 5050 | 2.03    | 1.89 | 2.27    | 2    |
| 3             | 212                          | 111  | 10600                      | 5550 | 2.05    | 1.87 | 2.21    | 1.87 |
| 4             | 121                          | 52.4 | 6050                       | 2620 | 1.88    | 1.84 | 2.13    | 2    |
| 5             | 158                          | 63.2 | 7900                       | 3160 | 1.83    | 1.86 | 1.9     | 2.3  |

|    |       |       |       |      |      |      |      |      |
|----|-------|-------|-------|------|------|------|------|------|
| 6  | 47    | 31.6  | 2350  | 1580 | 1.93 | 1.93 | 2.04 | 1.72 |
| 7  | 168   | 86.4  | 8400  | 4320 | 1.87 | 1.86 | 1.95 | 2.26 |
| 8  | 118   | 134   | 5900  | 6700 | 2.01 | 1.88 | 2.2  | 2.27 |
| 9  | 4.64  | 14.3  | 232   | 715  | 1.95 | 1.89 | 1.07 | 1.73 |
| 10 | 73.6  | 58.6  | 3680  | 2930 | 1.96 | 1.91 | 2.03 | 1.77 |
| 11 | 16.1  | 22    | 805   | 1100 | 1.94 | 1.93 | 1.7  | 1.65 |
| 12 | 0.707 | 0.362 | 35.35 | 18.1 | 2.2  | 1.91 | 3    | 0.31 |
| 13 | 34.4  | 45.2  | 1720  | 2260 | 2    | 1.92 | 2.15 | 2.05 |
| 14 | 24.8  | 20.2  | 1240  | 1010 | 1.71 | 1.79 | 1.72 | 1.94 |
| 15 | 173   | 102   | 8650  | 5100 | 1.8  | 1.87 | 1.7  | 2.3  |
| 16 | 82.8  | 120   | 4140  | 6000 | 1.9  | 1.89 | 1.72 | 2.18 |
| 17 | 412   | 108   | 20600 | 5400 | 1.75 | 1.86 | 1.68 | 2.29 |
| 18 | 1148  | 175   | 57400 | 8750 | 1.8  | 1.86 | 1.82 | 2.27 |
| 19 | 99.6  | 116   | 4980  | 5800 | 1.59 | 1.86 | 1.6  | 0.36 |
| 20 | 107.2 | 62.6  | 5360  | 3130 | 1.71 | 1.88 | 1.72 | 2.27 |
| 21 | 35.2  | 28.8  | 1760  | 1440 | 1.6  | 1.85 | 1.59 | 2.08 |

|    |       |      |       |       |      |      |      |      |
|----|-------|------|-------|-------|------|------|------|------|
| 22 | 117.2 | 80.6 | 5860  | 4030  | 1.71 | 1.86 | 1.6  | 2.27 |
| 23 | 1168  | 220  | 58400 | 11000 | 1.9  | 1.88 | 2.06 | 2.21 |
| 24 | 28.2  | 30.2 | 1410  | 1510  | 1.59 | 1.87 | 1.56 | 1.9  |
| 25 | 45.8  | 75.8 | 2290  | 3790  | 1.88 | 1.85 | 1.85 | 1.88 |
| 26 | 120   | 102  | 6000  | 5100  | 1.92 | 1.87 | 1.87 | 1.92 |
| 27 | 62.4  | 62.4 | 3120  | 3120  | 1.89 | 1.86 | 1.86 | 1.89 |
| 28 | 30.4  | 74.8 | 1520  | 3740  | 1.88 | 1.86 | 1.86 | 1.88 |
| 29 | 32.2  | 52.2 | 1610  | 2610  | 1.83 | 1.89 | 1.89 | 1.83 |
| 30 | 86.7  | 65.2 | 4335  | 3260  | 1.88 | 1.86 | 1.86 | 1.88 |
| 31 | 124   | 68.9 | 6200  | 3445  | 1.83 | 1.85 | 1.79 | 2.1  |
| 32 | 264   | 84.5 | 13200 | 4225  | 1.89 | 1.86 | 1.84 | 2.0  |
| 33 | 79    | 64.3 | 3950  | 3215  | 1.85 | 1.87 | 1.86 | 1.98 |

Tabla S4. Cantidades de ADN de amplicones purificados

| Tipo de muestra      | Espuito        |                 |         |              |                 |         | Cultivo Primario |                 |         |                |                 |         |
|----------------------|----------------|-----------------|---------|--------------|-----------------|---------|------------------|-----------------|---------|----------------|-----------------|---------|
| Método de extracción | Kit_FP_Columna |                 |         | FC_FP_AMPure |                 |         | FC_FP_AMPure     |                 |         | Kit_FP_Columna |                 |         |
| N° de muestra        | AD N           | 260/280 (ng/ul) | 260/230 | AD N         | 260/280 (ng/ul) | 260/230 | AD N             | 260/280 (ng/ul) | 260/230 | AD N           | 260/280 (ng/ul) | 260/230 |
| 1                    | 279            | 1.91            | 2.2     | 243          | 1.9             | 2.2     | 308              | 1.9             | 2.2     | 183            | 1.9             | 2.2     |
| 2                    | 186            | 1.9             | 2.2     | 321          | 1.9             | 2.2     | 339              | 1.9             | 2.2     | 206            | 1.9             | 2.2     |
| 3                    | 272            | 1.9             | 2.2     | 303          | 1.9             | 2.2     | 319              | 1.9             | 2.2     | 202            | 1.9             | 2.2     |

|    |     |     |     |     |     |   |     |     |   |     |     |   |
|----|-----|-----|-----|-----|-----|---|-----|-----|---|-----|-----|---|
| 4  | 172 | 1.9 | 2.2 | 448 | 1.9 | 2 | 313 | 1.9 | 2 | 224 | 1.9 | 2 |
| 5  | 303 | 1.9 | 2.2 | 307 | 1.9 | 2 | 334 | 1.9 | 2 | 214 | 1.9 | 2 |
| 6  | 333 | 1.9 | 2.2 | 245 | 1.9 | 2 | 292 | 1.9 | 2 | 216 | 1.9 | 2 |
| 7  | 181 | 1.9 | 2.2 | 259 | 1.9 | 2 | 319 | 1.9 | 2 | 212 | 1.9 | 2 |
| 8  | 230 | 1.9 | 2.2 | 11  | 1.9 | 2 | 400 | 1.9 | 2 | 206 | 1.9 | 2 |
| 9  | 238 | 1.9 | 2.2 | 271 | 1.9 | 2 | 391 | 1.9 | 2 | 202 | 1.9 | 2 |
| 10 | 146 | 1.9 | 2.2 | 263 | 1.9 | 2 | 332 | 1.9 | 2 | 212 | 1.9 | 2 |
| 11 | 148 | 1.9 | 2.2 | 349 | 1.9 | 2 | 387 | 1.9 | 2 | 206 | 1.9 | 2 |

|    |     |     |     |     |     |    |     |     |    |     |    |    |
|----|-----|-----|-----|-----|-----|----|-----|-----|----|-----|----|----|
| 12 | 165 | 1.9 | 2.2 | 285 | 1.9 | 2. | 480 | 1.9 | 2. | 159 | 1. | 2. |
|    |     |     |     |     |     | 2  |     |     | 2  |     | 9  | 2  |