



UNIVERSIDAD PERUANA  
**CAYETANO HEREDIA**

Facultad de  
**MEDICINA**

DESEMPEÑO DE LA ESPECTROMETRÍA DE MASAS MALDI BIOTYPER  
SIRIUS Y LA SECUENCIACIÓN DE GENOMA COMPLETO POR OXFORD  
NANOPORE PARA LA IDENTIFICACIÓN DE ESPECIES DE  
*Mycobacterium* spp.

PERFORMANCE OF MALDI BIOTYPER SIRIUS MASS SPECTROMETRY AND  
THE OXFORD NANOPORE WHOLE-GENOME SEQUENCING FOR THE  
IDENTIFICATION OF *Mycobacterium* spp. SPECIES

TESIS PARA OPTAR POR EL TÍTULO PROFESIONAL DE LICENCIADO  
EN TECNOLOGÍA MÉDICA EN LA ESPECIALIDAD DE LABORATORIO  
CLÍNICO Y ANATOMÍA PATOLÓGICA

AUTOR

DAVID EDUARDO HIDALGO AQUIÑO

ASESOR

PATRICIA SHEEN CORTAVARRIA

CO-ASESORES

LIDIO EDGAR NEYRA VALDEZ

MISHELL LANDA FLORES

LIMA-PERÚ

2026



## **JURADO DE TESIS**

Presidente: DR. JUAN CARLOS AGAPITO PANTA

Vocal: MG. MARIA DEL CARMEN QUISPE MANCO

Secretario: LIC. DELIA MARGOT FAUSTINO ARIAS

Fecha de aprobación: 19 de febrero del 2026

Calificación: APROBADO

**ASESORES DE TESIS**

**ASESOR**

PH.D PATRICIA SHEEN CORTAVARRIA

Laboratorio de Biología Molecular | Facultad de Ciencias e Ingeniería

ORCID:0000-0002-7118-9301

**CO-ASESOR**

M.SC. EDGAR LIDIO NEYRA VALDEZ

Unidad de Genómica | Facultad de Medicina - EPTM

ORCID: 0000-0003-2086-7245

**CO-ASESOR**

M.SC. MISHHELL LANDA FLORES

Laboratorio de Biología Molecular | Facultad de Ciencias e Ingeniería

ORCID: 0009-0005-5409-879X

## **DEDICATORIA**

En especial a mi madre, Margarita Aquíño Vega; a mi abuelita, Juana Esther Vega Támara; a mi tía, Rocío Rodríguez Vega, con profundo cariño y gratitud, dedico este trabajo, gracias por haberme dado todas las herramientas y oportunidades posibles.

A mis tíos: Zenaida Portalino, Abilio Vega Támara, Nora Alvarado y Hugo Cedrón Cheng, gracias por enseñarme a ser la mejor persona posible y por permanecer a mi lado en cada paso de este camino.

A Lizeth Mamani, por acompañarme a lo largo de todos estos años, su compañía ha sido un valioso engranaje a mi vida.

## **AGRADECIMIENTOS**

Deseo expresar mi más sincero agradecimiento a los doctores Patricia Sheen, Mirko Zimic, Louis Grandjean y Alice Omaston, por la generosa prestación de los equipos y reactivos que hicieron posible la realización de este proyecto.

Agradezco también al M.Sc. Edgar Neyra, por sus valiosos consejos durante mi etapa universitaria, y a la M.Sc. Mishell Landa, por su confianza y respaldo inquebrantables.

Un reconocimiento especial a la M.Sc. María de Grecia Cauti por haberme formado y orientado en los fundamentos de la biología, así como por su cercana amistad y por incentivar profundamente mi vocación por las ciencias biológicas.

A la Unidad de Biología Molecular/Micobacterias, por su apoyo constante durante este proyecto, en especial al Dr. Jorge Coronel, M.Sc. Candy León, al Lic. Walter Solano, a la Lic. Jimena Raez y al Lic. José Luis Pérez. Asimismo, a la Unidad de Bioinformática, por su acompañamiento y orientación, con especial reconocimiento al Lic. Paul Avellaneda.

A mis colegas, amigos y compañeros, quienes hicieron de este proceso una experiencia enriquecedora y agradable:

Kiara Aricoché, Misael Arrue, Kevin Obando, Luis Canales, Sebastián Rodríguez, Ormar Romero, Valeria Shantall, Karen López, Renato Fernández, Claudia Soto, Moe Kishima, María Del Carmen, Ada Vergara, Alessandra Vasquez.

Gracias por fomentar un ambiente de compañerismo durante mi estancia en el laboratorio.

## **FUENTES DE FINANCIAMIENTO**

Autofinanciado

## **DECLARACIÓN DE CONFLICTO DE INTERÉS**

El autor declara no tener conflictos de interés.

## DECLARACIÓN DE ORIGINALIDAD



UNIVERSIDAD PERUANA  
CAYETANO HEREDIA

### DECLARACIÓN DE ORIGINALIDAD

El egresado:

N°	APELLIDOS Y NOMBRES
1.	HIDALGO AQUIÑO DAVID EDUARDO

Pertencientes al programa de la **CARRERA PROFESIONAL DE TECNOLOGÍA MÉDICA EN LA ESPECIALIDAD DE LABORATORIO CLÍNICO Y ANATOMÍA PATOLÓGICA**, autor del trabajo titulado: **DESEMPEÑO DE LA ESPECTROMETRÍA DE MASAS MALDI BIOTYPER SIRIUS Y LA SECUENCIACIÓN DE GENOMA COMPLETO POR OXFORD NANOPORE PARA LA IDENTIFICACIÓN DE ESPECIES DE *Mycobacterium* spp.** el cual ha sido elaborado, sustentado y aprobado, según corresponda, para optar por el **TÍTULO PROFESIONAL DE LICENCIADO EN TECNOLOGÍA MÉDICA EN LA ESPECIALIDAD DE LABORATORIO CLÍNICO Y ANATOMÍA PATOLÓGICA** 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.	SHEEN CORTAVARRIA PATRICIA	MEDICINA	ASESOR
2.	NEYRA VALDEZ LIDIO EDGAR	MEDICINA	CO-ASESOR
3.	LANDA FLORES MISHELL	MEDICINA	CO-ASESOR

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 **10 %**, según el reporte emitido por el software **Turnitin®** (identificador de entrega: **trn:oid:::1:3547953121**; fecha de entrega: **23-04-2026**).
- 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, 23 de abril del 2026.**

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

Firma del Co-asesor  
N° DNI: 09608590  
ORCID: 0000-0003-2086-7245

Firma del Co-asesor  
N° DNI: 70357144  
ORCID: 0009-0005-5409-879X



## TABLA DE CONTENIDOS

	Pág.
RESUMEN	
ABSTRACT	
I. INTRODUCCIÓN	1
1.1. Justificación	5
II. OBJETIVOS	7
2.1. Objetivo general	7
2.2. Objetivos secundarios	7
III. MATERIALES Y MÉTODOS	8
3.1. Diseño del estudio	8
3.2. Población y lugar de estudio	8
3.3. Criterios de inclusión	9
3.4. Criterios de exclusión	9
3.5. Muestra	9
3.6. Procedimientos y técnicas	9
3.6.1. Siembra:	9
3.6.2. MALDI Biotyper sirius:	10
3.6.3. Secuenciación de genoma:	11
3.7. Aspectos éticos	13
3.8. Plan de análisis	13
IV. RESULTADOS	16
V. DISCUSIÓN	19
VI. LIMITACIONES	24
VII. CONCLUSIONES	25
VIII. REFERENCIAS BIBLIOGRÁFICAS	26
IX. TABLAS, GRÁFICOS Y FIGURAS	33
ANEXOS	

## RESUMEN

**Introducción:** *Mycobacterium tuberculosis* (MTB) y las micobacterias no tuberculosas (MNT) desafían la salud pública en el Perú por su carga y complejidad diagnóstica. Los métodos convencionales son limitados en rapidez, especificidad y costo. La espectrometría de masas MALDI Biotyper Sirius y la secuenciación por Oxford Nanopore podrían optimizar la identificación de *Mycobacterium* spp.

**Objetivo:** Evaluar el desempeño de MALDI Biotyper sirius y la secuenciación por Oxford Nanopore para la identificación de especies de *Mycobacterium* spp.

**Métodos:** Estudio comparativo, observacional y descriptivo. Se procesaron 186 aislados criopreservados del Laboratorio de Micobacterias (UPCH). Tras cultivo e inactivación, se procesaron por MALDI Biotyper Sirius; en paralelo se extrajo ADN para Nanopore (estándar de referencia). Se estimaron sensibilidad, especificidad, concordancia ( $\kappa$ ), curvas ROC/AUC y métricas de ensamblaje genómico (completitud y N50).

**Resultados:** MALDI Biotyper Sirius identificó 177/186 como *Mycobacterium tuberculosis* complex y 9/186 como grupo *M. chimaera/intracellulare*; solo 4/9 se subtipificaron como *M. intracellulare*. El score promedio fue 2,018. La secuenciación de genoma completo o Whole genome sequencing (WGS) asignó 177 como *M. tuberculosis* y 9 como *M. intracellulare*, con ensamblajes de alta calidad (un contig o secuencia continua de ADN generada por ensamblaje de lecturas; N50 ~4,3–4,4 Mbp; contaminación 0,02–0,91%). En la clasificación de aislamientos como MTB o MNT, MALDI Biotyper Sirius presentó una sensibilidad y especificidad de 100%, con  $\kappa=1,000$  y AUC=1,000. Para especie exacta, la sensibilidad relativa de MALDI Biotyper Sirius fue 2,15% (IC95% 0,59–5,41).

**Conclusiones:** MALDI Biotyper Sirius presentó desempeño y concordancia excelentes para diferenciar MTB vs MNT (sensibilidad/especificidades cercanas al 100%,  $\kappa=1,000$ ; AUC=1,000). Sin embargo, a nivel de especie su concordancia fue limitada por la baja/nula subtipificación, con una sensibilidad relativa frente a Nanopore de 2,15%.

**Palabras claves:** *M. tuberculosis*, MALDI-TOF, Secuenciación, Micobacterias no tuberculosas, Evaluación.

## ABSTRACT

**Introduction:** *Mycobacterium tuberculosis* (MTB) and nontuberculous mycobacteria (NTM) challenge public health in Peru due to their burden and diagnostic complexity. Conventional methods are limited in terms of speed, specificity, and cost. MALDI Biotyper Sirius mass spectrometry and Oxford Nanopore sequencing could optimize the identification of *Mycobacterium* spp.

**Objective:** To evaluate the performance of MALDI Biotyper Sirius and Oxford Nanopore sequencing for the identification of *Mycobacterium* spp. species.

**Methods:** Comparative, observational, and descriptive study. A total of 186 cryopreserved isolates from the Mycobacteria Laboratory (UPCH) were processed. After culture and inactivation, they were processed by MALDI Biotyper Sirius; in parallel, DNA was extracted for Nanopore sequencing (reference standard). Sensitivity, specificity, agreement ( $\kappa$ ), ROC/AUC curves, and genomic assembly metrics (completeness and N50) were estimated.

**Results:** MALDI Biotyper Sirius identified 177/186 as *Mycobacterium tuberculosis* complex and 9/186 as the *M. chimaera/intracellulare* group; only 4/9 were subtyped as *M. intracellulare*. The mean score was 2.018. Whole genome sequencing (WGS) assigned 177 as *M. tuberculosis* and 9 as *M. intracellulare*, with high-quality assemblies (one contig or continuous DNA sequence generated by read assembly; N50 ~4.3–4.4 Mbp; contamination 0.02–0.91%). In the classification of isolates as MTB or NTM, MALDI Biotyper Sirius showed a sensitivity and specificity of 100%, with  $\kappa=1.000$  and AUC=1.000. For exact species identification, the relative sensitivity of MALDI Biotyper Sirius was 2.15% (95% CI 0.59–5.41).

**Conclusions:** MALDI Biotyper Sirius showed excellent performance and agreement for differentiating MTBC vs NTM (sensitivities/specificities close to 100%,  $\kappa=1.000$ ; AUC=1.000). However, at the species level, its agreement was limited by the low/null subtyping, with a relative sensitivity of 2.15% compared with Nanopore.

**Keywords:** *M. tuberculosis*, MALDI-TOF, Sequencing, Nontuberculous mycobacteria, Evaluation.

## I. INTRODUCCIÓN

La tuberculosis (TB) en humanos es una enfermedad causada principalmente por el bacilo ácido-cohol resistente *Mycobacterium tuberculosis*, que se transmite a través de aerosoles liberados al toser por individuos infectados, esto lo hace pertenecer al grupo 3 de agentes patógenos. Esta vía de transmisión ha llevado a estimar que alrededor de una cuarta parte de la población mundial alberga infección latente y que entre el 5 % y el 10 % de las personas infectadas desarrollará enfermedad a lo largo de su vida [1]. La TB es la primera causa de muerte por un agente infeccioso en el mundo por encima del Virus de la inmunodeficiencia humana (VIH) y la malaria [2]. El Perú históricamente ha sido un país endémico de *M.tuberculosis*, actualmente se reportan entre 100 a 299 casos nuevos por 100 000 hab/año [1].

Las micobacterias no tuberculosas (MNT) son un agente patógeno emergente en personas inmunocomprometidas, su diferenciación clínica con la tuberculosis puede llegar a ser confusa, además la identificación por laboratorio, mediante el flujo de trabajo rutinario, puede llegar a ser compleja [3].

Tradicionalmente, inmediatamente después de identificar al paciente sospechoso de tuberculosis se debe proceder a la correcta recolección de la muestra de esputo, la cual permitirá la realización de la baciloscopia donde se aprecia bacilos ácido-alcohol resistentes (BAAR) de color fucsia, vistos así debido a la penetración de la fucsina fenicada en la pared celular de ácidos micólicos. *Mycobacterium tuberculosis* es un tipo de BAAR, por lo que la técnica no es discriminante [4]. Sin embargo, debido a su bajo costo esta es predilecta para tamizaje.

En el Perú, las plataformas de identificación microbiológica de rutina basadas en perfiles bioquímicos no permiten la identificación de *Mycobacterium* spp. ; por ello, la mayoría de establecimientos de salud emplea plataformas moleculares como GeneXpert para la identificación rápida de *Mycobacterium tuberculosis*, sin embargo, no pueden diferenciar si se trata de otro tipo de micobacteria u otro agente patógeno.[4] Además los consumibles como los cartuchos tienen un costo alto por paciente, sumando a eso la deficiente gestión puede incurrir en desabastecimiento de los mismos lo que conlleva a retrasos en el diagnóstico.

Los laboratorios de referencia poseen los ensayos Additional Species(AS) y Common *Mycobacterium* (CM), basadas en la hibridación por sonda, la cual son tecnologías capaces de identificar género y especie. Por un lado, AS es capaz de diferenciar algunas micobacterias no tuberculosas y el Complejo *Mycobacterium tuberculosis* (MTBC), sin embargo, no puede dilucidar algunas especies de las MNT ni las especies dentro del MTBC, para esto es necesario el ensayo Non-Tuberculous Mycobacteria Drug Resistance (NTM-DR), la cual incluye más sondas de identificación para Micobacterias no tuberculosas y el ensayo Common *Mycobacterium* (CM) respectivamente, los cuales incluyen sondas específicas para la diferenciación [3, 11].

La espectrometría de masas MALDI Biotyper Sirius no se usa ampliamente en el Perú. Permite identificar microorganismos mediante el perfil de masas de sus proteínas y, por su bajo costo por muestra (consumibles asequibles), representa una alternativa frente a otras tecnologías. [8]. La técnica consiste en mezclar la muestra directa o las proteínas extraídas con una matriz que absorbe luz ultravioleta, comúnmente ácido  $\alpha$ -ciano-4-hidroxicinámico (HCCA). Luego, esta mezcla es

irradiada con un láser, lo que provoca la desorción e ionización de las moléculas de la muestra. Los iones generados son acelerados en un tubo de vuelo y su masa es determinada en función del tiempo que tardan en alcanzar el detector. La identificación se realiza comparando el espectro de masas obtenido con una base de datos de espectros conocidos [5, 6].

La secuenciación de ADN y sobre todo en particular la del genoma completo (WGS) de un organismo es una técnica robusta y precisa para la identificación de patógenos y es aplicable a micobacterias, los estudios de epidemiología molecular han demostrado cómo la información genómica puede complementar la epidemiología tradicional. En este marco, la tecnología Nanopore ha cobrado relevancia por su capacidad de generar lecturas largas en tiempo real a partir de cambios en la corriente eléctrica producidos durante el paso del ADN por un nanoporo. En los últimos años, la calidad de los datos obtenidos mediante esta plataforma ha mejorado de forma continua y significativa. Estas lecturas largas brindan una mejor resolución de regiones complejas y disminuyen los sesgos en genomas ricos en GC, limitación aún presente en la tecnología de lecturas cortas. Se ha reportado que en un mismo estudio que con el pasar de los años, los resultados mejoraron progresivamente con cada actualización del software de basecalling y de las herramientas bioinformáticas [7]. Debido a esto la secuenciación de genoma completo mediante Oxford Nanopore es una metodología confirmatoria de alta resolución, ya que resuelve las limitaciones de otras técnicas moleculares confirmatorias, sobre todo en el contexto del diagnóstico de tuberculosis. Sin embargo, esta no es considerada para su implementación en los laboratorios de referencia del país. Además, a pesar de ser una metodología robusta, no hay

bibliografía disponible que utilice la secuenciación por Oxford Nanopore como metodología de referencia en estudios comparativos en los que se emplea espectrometría de masas para la identificación de *Mycobacterium spp* [9,10].

La problemática actual reside en la necesidad de contar con métodos de diagnóstico que sean realmente muy precisos en el contexto de la salud pública peruana. Evaluar tecnologías emergentes como el MALDI Biotyper Sirius y la secuenciación de genoma completo es fundamental para optimizar el diagnóstico de MTB y MNT, reduciendo tiempos y permitiendo una toma de decisiones a nivel epidemiológico de estas infecciones. En este sentido, estudios realizados en la región han evidenciado el potencial de tecnologías como MALDI-TOF para fortalecer la identificación de micobacterias en el laboratorio. Por ejemplo, Contreras et al. reportaron la experiencia de identificación de especies de micobacterias mediante espectrometría de masas (MALDI-TOF) en un centro de referencia en Chile. Asimismo, Vásquez et al. evaluaron la identificación de micobacterias no tuberculosas en muestras de agua potable domiciliaria, utilizando MALDI-TOF y métodos moleculares complementarios para la resolución de aislamientos [12,13]. Estos resultados refuerzan la importancia de incorporar tales herramientas en los laboratorios de rutina y de referencia en el Perú, lo que podría transformar la respuesta del sistema de salud ante estas infecciones.

En la presente tesis se evaluó si el sistema MALDI Biotyper Sirius de la empresa Bruker, el cual está basado en la tecnología MALDI-TOF, puede proporcionar una identificación de especies de micobacterias que sean tanto precisas, esto sobre 186 aislados clínicos disponibles. Se realizó una evaluación de los parámetros de calidad operacionales frente a la secuenciación por Oxford Nanopore. La

evaluación de estas tecnologías emergentes no solo aporta evidencia científica, sino que también sientan las bases para la toma de decisiones en la incorporación de innovaciones diagnósticas en los laboratorios de referencia del Perú, fortaleciendo así la capacidad de respuesta del sistema de salud ante la MTB y las MNT, al igual que otras enfermedades infecciosas.

### **1.1. Justificación**

Este estudio comparó MALDI Biotyper Sirius y la secuenciación de genoma completo por Oxford Nanopore porque son enfoques complementarios para identificar *Mycobacterium* spp. MALDI Biotyper sirius es una herramienta proteómica rápida y con menor costo por muestra, mientras que la WGS permite observar el genoma completo del aislado, lo que brinda mayor detalle para la identificación, aunque requiere más recursos y tiempo de procesamiento. Por ello, en este trabajo la secuenciación de genoma completo por Oxford Nanopore se empleó como método de referencia, ya que ofrece una base más completa para contrastar los resultados de una plataforma proteómica.

En la literatura, es común que las comparaciones se realicen usando secuenciación de genes diana y, cuando un solo gen no discrimina bien entre especies cercanas, se recurre a esquemas multigénicos. Sin embargo, estos enfoques siguen siendo parciales: dependen del marcador elegido y pueden no resolver adecuadamente especies muy relacionadas o reportar resultados en distintos niveles taxonómicos. En ese sentido, usar secuenciación de genoma completo como referencia ayuda a reducir el sesgo de comparar MALDI Biotyper Sirius contra confirmaciones incompletas.

Finalmente, en nuestro contexto los laboratorios necesitan métodos sensibles y específicos, con tiempos de respuesta cortos y costos sostenibles. Dado que la evidencia local comparativa es limitada, este estudio estimó el desempeño diagnóstico (sensibilidad/especificidad y concordancia) y describió parámetros de calidad relevantes de ambas técnicas, aportando información útil para su implementación y toma de decisiones en el flujo diagnóstico.

## **II. OBJETIVOS**

### **2.1. Objetivo general**

Evaluar el desempeño de la espectrometría de masas MALDI Biotyper Sirius y la secuenciación por Oxford Nanopore para la identificación de especies de *Mycobacterium spp.*

### **2.2. Objetivos secundarios**

- Determinar la sensibilidad, especificidad y rendimiento diagnóstico de la espectrometría de masas MALDI Biotyper Sirius en la identificación de especies de *Mycobacterium spp.*, utilizando la secuenciación por Oxford Nanopore como estándar de referencia.
- Analizar la concordancia entre ambos métodos en la identificación de especies de *Mycobacterium spp.*

### III. MATERIALES Y MÉTODOS

#### 3.1. Diseño del estudio

El estudio fue comparativo y observacional. Se emplearon dos métodos para la identificación de especies de *Mycobacterium*: la espectrometría de masas (MALDI Biotyper Sirius) y la secuenciación de genoma completo mediante plataforma Oxford Nanopore. Se evaluaron la precisión y la concordancia entre ambos métodos en la identificación de aislados de micobacterias criopreservados en el laboratorio. La metodología incluyó la siembra de los aislados en medios de cultivo sólidos, seguida de la inactivación y el procesamiento de estos para el análisis por MALDI Biotyper Sirius. En paralelo, se realizó la extracción de ADN y la secuenciación de genoma completo por Oxford Nanopore, utilizada como método de referencia comparativo en el estudio, además del análisis bioinformático para determinar la especie bacteriana. Los resultados se compararon estadísticamente para estimar la sensibilidad, la especificidad, la concordancia y el rendimiento de ambos métodos.

#### 3.2. Población y lugar de estudio

La población está constituida por aislados criopreservados de *Mycobacterium spp.*, procedentes del Laboratorio de Enfermedades Infecciosas: *Mycobacterium* “Luz Caviedes Rojas” (BSL-3) de la Universidad Peruana Cayetano Heredia (UPCH). El estudio se realizó en el laboratorio de Bioinformática, Biología Molecular y Desarrollos Tecnológicos (BSL-2) y BSL3 antes mencionado.

### 3.3. Criterios de inclusión

- Aislados, recuperados por subcultivo en agar 7H10 + OADC + Tween e incubados a 37 °C por 21 días, con crecimiento que permita su procesamiento, consistentes con la morfología de *Mycobacterium spp* en medio de cultivo selectivo previamente aisladas y almacenadas.

### 3.4. Criterios de exclusión

- Contaminación Evidente: Cualquier aislado que muestre signos de contaminación microbiológica cruzada durante el proceso de cultivo o en cualquier fase preparatoria del estudio se excluyó.

### 3.5. Muestra

- Se utilizó todas las aislados disponibles para este proyecto, de *Mycobacterium spp* (n=186), proporcionadas por el criobanco del Laboratorio de Micobacterias “Luz Caviedes Rojas” (LID-118) de la Universidad Peruana Cayetano Heredia.

### 3.6. Procedimientos y técnicas

#### **3.6.1. Siembra:**

En el laboratorio de Micobacterias se sembraron los aislados criopreservados a –70 °C (80 ul de cada uno) en placas descartables con agar 7H10 + OADC + Tween, y dichas placas fueron incubadas a 37 °C durante 21 días. Al vigésimo primer día, momento aproximado de la fase logarítmica y de crecimiento óptimo con mínimo riesgo de contaminación, se transfirió la biomasa micobacteriana, con un asa de

inoculación de 1  $\mu\text{L}$ , a un tubo Eppendorf Safe-Lock de 1,5 mL; tras pipetear 300  $\mu\text{L}$  de agua MiliQ y realizar la inactivación térmica a 95 °C durante 30 min en termobloque, se añadieron 900  $\mu\text{L}$  de etanol y se mezcló en vórtex.

### **3.6.2. MALDI Biotyper sirius:**

En el Laboratorio de Bioinformática, Biología Molecular y Desarrollos Tecnológicos, la suspensión micobacteriana inactivada preparada previamente (300  $\mu\text{L}$  de agua MilliQ y 900  $\mu\text{L}$  de etanol) se centrifugó a 13 000 rpm durante 2 min. Luego, se descartó el sobrenadante y se retiró el líquido residual con puntas estériles, dejando secar el pellet a temperatura ambiente durante unos minutos. Posteriormente, en el mismo tubo se colocaron perlas de circonio/sílice de 0,5 mm, se añadieron 25  $\mu\text{L}$  de acetonitrilo puro y la mezcla se agitó en vórtex a máxima velocidad durante 1 min. A continuación, se agregó un volumen igual de ácido fórmico al 70 % y se agitó nuevamente durante 5 s. Tras una segunda centrifugación a 13 000 rpm por 2 min, se depositó 1  $\mu\text{L}$  del sobrenadante sobre la placa objetivo del equipo y, una vez seco, se cubrió de inmediato con 1  $\mu\text{L}$  de matriz HCCA (Bruker n.º 8255344), previamente preparada con el disolvente estándar compuesto por 50 % de acetonitrilo, 47,5 % de agua y 2,5 % de TFA [33].

La librería MBT Mycobacteria Library 5.0 (DB-912 MSP #1850766) de Bruker contiene 912 Main Spectra Projections correspondientes a 164 especies de Mycobacterium, generadas a partir de 438 cepas de referencia y 474 cepas clínicas aportadas por 26 colaboradores de 10 países. Debido a la alta similitud de los perfiles entre algunos taxones, esta librería incluye 13 matching hints descritos en Anexos II. En estos casos, Bruker recomienda correlacionar el resultado con la

información clínica, evaluar cuidadosamente el nivel de identificación y, de ser necesario, confirmar mediante pruebas complementarias.

Respecto a la calificación de las muestras, el equipo asigna a cada espectro un score entre 0 y 3, el cual expresa la calidad de la coincidencia con la base de datos. Los valores entre 1,80 y 3,00 (verde, símbolo “+++”) indican identificación de alta confianza; los valores entre 1,60 y 1,79 (amarillo, “+”) indican concordancia moderada y requieren confirmación; y los valores entre 0,00 y 1,59 (rojo, “-”) no permiten una identificación fiable a nivel de especie. Con base en ello, cada muestra se clasificó en tres categorías de consistencia. La categoría A (alta consistencia) correspondió a muestras cuya mejor coincidencia fue de alta confianza a nivel de especie y cuya segunda coincidencia concordó al menos a nivel de género, también fue de alta confianza o no presentó identificación alternativa. La categoría B (baja consistencia) incluyó muestras en las que la primera coincidencia fue de alta o baja confianza, pero la segunda no cumplió los criterios de la categoría A. Finalmente, la categoría C (sin consistencia) agrupó los casos en los que no hubo una segunda coincidencia coherente, lo que sugiere un resultado poco fiable [32]. Esta clasificación se presenta en tablas 5 y 6 de TABLAS, GRÁFICOS Y FIGURAS.

### **3.6.3. Secuenciación de genoma:**

Se extrajo el ADN a partir de la misma placa de cultivo en el Laboratorio de Micobacterias de la UPCH (BLS3) empleando el ZymoBIOMICSTM DNA Miniprep Kit (Zymo Research, ref. D4300) según las instrucciones del fabricante. A continuación, se aplicó un paso controlado de fragmentación para reducir la longitud de los fragmentos de ADN, de más de 60 kb a aproximadamente 16 kb

mediante tubos g (Covaris, ref. 520079), utilizando 2 µg de ADN de entrada y centrifugando a 4 200 rpm durante 1 minuto hasta que la muestra descendió al fondo del tubo. El ADN fragmentado se purificó y concentró en un volumen menor mediante columnas de unión del DNA Clean & Concentrator Kit (Zymo Research, ref. D4004). [34] [35]

Posteriormente, se secuenció el ADN en cinco corridas usando tres celdas de flujo MinION R10.4.1 instaladas en un dispositivo GridION y dos celdas de flujo en un dispositivo P2 conectado al mismo GridION. Para cada corrida se prepararon bibliotecas con el Native Barcoding Kit 96 (Oxford Nanopore, SQK-NBD114.96), según indicaciones del fabricante. Tras agrupar y normalizar las muestras, se cargaron entre 18 y 85 fmol de cada grupo en las celdas de flujo y se secuenciaron hasta que, en dos corridas, el porcentaje de poros activos alcanzó el 3 % y, en las otras tres, hasta la estabilización de la producción de datos, momento en que se pausó la corrida, se lavó la celda con nucleasas y se recargó la misma biblioteca. [36]

Cada celda de flujo GridION contuvo entre 21 y 24 muestras multiplexadas (promedio 23), mientras que las R10.4.1 agrupó entre 55 y 64 muestras, incrementando progresivamente el volumen de entrada en corridas sucesivas para optimizar el rendimiento. Los archivos FASTQ se generaron con Guppy basecaller v.6.5.7 de Oxford Nanopore Technologies, aplicando un q-score mínimo de 10 y longitud mínima de 200 pb.

El control de calidad inicial se realizó con NanoPlot v.1.41.3, el mapeo de lecturas se efectuó con Minimap2 v.2.26 contra el genoma de referencia *M. tuberculosis*

H37Rv (NC\_000962.3) y el análisis de cobertura y profundidad con Samtools v.1.17. Para la clasificación taxonómica a partir de la secuenciación se utilizó la herramienta Genome Taxonomy Database-toolkit (GTDB-Tk).

### 3.7. Aspectos éticos

Este proyecto se registró en el Sistema Descentralizado de Información y Seguimiento a la Investigación (SIDISI) de la Dirección Universitaria de Investigación, Ciencia y Tecnología (DUICT) y, antes de su ejecución, fue evaluado por el Comité de Ética de la UPCH (CIE-UPCH). Además, se utilizó el formato F3 de la Oficina de Regulación y Valoración Ética de la Investigación (ORVEI), destinado a proyectos que no involucraron humanos ni animales, para que la revisión se derive a la Dirección Universitaria de Asuntos Regulatorios en Investigación (DUARI). Durante la implementación del estudio, se garantizó el respeto a los principios éticos establecidos en la Declaración de Helsinki y se acataron estrictamente las recomendaciones del CIE-UPCH.

Se garantizó la seguridad del personal mediante el uso de equipos de protección personal, cabinas de bioseguridad tanto en el ambiente de BSL-3 como BSL-2, así mismo para el manejo de químicos volátiles se realizó dentro de una campana extractora.

### 3.8. Plan de análisis

Se aplicó un análisis descriptivo en el que se calcularon las medias y desviaciones estándar de variables cuantitativas como el score del espectro, así como las frecuencias y porcentajes de las variables cualitativas. Se evaluó la concordancia

entre métodos mediante el índice Kappa, lo que permitió determinar el grado de coincidencia entre las identificaciones realizadas por cada tecnología.

Los parámetros de calidad fueron evaluados mediante QUAST, BUSCO y CheckM2 herramientas bioinformáticas empleadas para analizar métricas de ensamblaje, completitud genómica y contaminación, respectivamente. En el caso de la espectrometría de masas, se empleó los datos generados por el sistema Bruker.

Se estimó el rendimiento bajo diferentes definiciones de caso positivo, se aplicó una normalización taxonómica y recodificamos las salidas en binario (1=positivo; 0=negativo) en tres escenarios: (i) MTBC: se marcó como positivo cualquier identificación de *Mycobacterium tuberculosis complex* en MALDI Biotyper Sirius y, en el estándar de referencia, cualquier etiqueta que contuviera “s\_ *Mycobacterium tuberculosis*”; (ii) NTM: se consideró positivo toda identificación de micobacterias no tuberculosas en ambos métodos; y (iii) especie: se definió como positivo la coincidencia exacta a nivel de especie entre MALDI Biotyper Sirius y el estándar. Para cada escenario calculamos sensibilidad, especificidad, intervalos de confianza al 95%, curvas ROC con su AUC, y la concordancia mediante  $\kappa$  de Cohen, reportando los valores correspondientes en la sección de Resultados.

La normalización taxonómica se ajustó al contexto peruano: la mayoría de los aislamientos humanos del complejo *M. tuberculosis* corresponde a *M. tuberculosis* sensu stricto (principalmente linaje 4 y, en menor medida, linaje 2). En Lima, *M. bovis* y *M. caprae* solo se han reportado esporádicamente, y *M. africanum* no es

endémica (restringida a África occidental). Por ello, el sesgo potencial de esta asunción sería limitado en nuestro entorno [30,31].

Los análisis se realizaron en Stata Now/SE 19.5 (StataCorp, TX, EE. UU.) y los datos se almacenaron en Excel 365 (Microsoft, Redmond, WA, EE. UU.).

#### IV. RESULTADOS

Se utilizaron las 186 aislados criopreservados disponibles de *Mycobacterium tuberculosis* y otras especies de *Mycobacterium* spp del Laboratorio de Micobacterias. Estos fueron analizados por MALDI Biotyper Sirius y, en paralelo, por secuenciación de genoma completo mediante Oxford Nanopore, el cual se utilizó como método de referencia comparativo en el estudio, para la identificación *Mycobacterium tuberculosis*.

Mediante MALDI Biotyper Sirius, 177 aislados se identificaron como *M.tuberculosis complex* y 9 aislados se identificaron como *M.chimaera\_intracellulare\_group*, de estas últimas, 4 fueron identificadas a nivel de especie como *M.intracellulare*. El score promedio de identificación fue de 2,018. En contraste mediante secuenciación por Oxford Nanopore se identificaron todas los aislados a nivel de especie, 177 como *M.tuberculosis* y 9 como *M.intracellulare*. En el caso de la secuenciación con Oxford Nanopore, se reporta mediante QUAST que los aislados presentaron ensamblajes compuestos por un único contig, con un valor de N50 equivalente al tamaño total del genoma de micobacterias (aproximadamente 4.3 a 4.4 Mbp) (Figura 1). Los niveles de contaminación oscilaron entre 0.02% y 0.91%. (Figura 2). El contenido de GC se mantuvo constante entre los aislados, con un valor aproximado del 66% (Figura 3). El porcentaje de completitud (completeness) fue prácticamente del 100% en todos los aislados (Figura 4). La evaluación mediante Benchmarking Universal Single-Copy Orthologs(BUSCO) evidenció que la mayoría de los aislados recuperaron 354 a 355 de los 356 genes BUSCO esperados, asimismo, la mayoría presentó entre 0 y 1 genes clasificados como 'Complete and Duplicated' o 'Fragmented', y un único

grupo como 'Missing' (Figura 5). Así mismo se reporta que el 90% del genoma es codificantes, con un promedio de 900 genes por megabase (Figura 6).

Los aislados clasificados como *Mycobacterium tuberculosis\_complex* (n=177) por MALDI Biotyper Sirius, un 82.5% (146) de los aislados fueron clasificadas con alta consistencia “A”, un 13.6% (24) de los aislados fueron clasificadas con baja consistencia “B”, y un 4% (7) de los aislados no cumplió con los requisitos de consistencia, siendo clasificadas como sin consistencia “C”. En el caso de *Mycobacterium chimaera\_intracellulare\_group* (n=9), un 33.33% (3) de los aislados fueron identificadas con alta consistencia “A”, un 44.44% (4) con “B” y un 22.22% (2) de los aislados no cumplieron con los requisitos de consistencia, clasificándose como sin consistencia “C” (Figura 7).

En cuanto al parámetro de confianza, un 90.45% (161) de los aislados de *Mycobacterium tuberculosis\_complex* fueron identificadas con alta confianza, mientras que 9.04% (16) de los aislados fueron clasificadas con baja confianza. En cuanto a la confianza en las identificaciones de *Mycobacterium chimaera\_intracellulare\_group*, un 55.56% (5) de los aislados fueron identificadas con alta confianza, mientras que el 44.44% (4) de los aislados mostraron baja confianza, lo que indica que la certeza en la identificación de estos aislados fue limitada (Figura 8).

En cuanto a los 4 aislados del grupo *M. chimaera/intracellulare* que alcanzaron el módulo de subtipificación a nivel de especie y fueron clasificados como *M. intracellulare*, el 100 % (4/4) presentó alta confianza; sin embargo, el 75 % (3/4) mostró consistencia “A” y el 25 % (1/4) consistencia “C”. Del resto de los aislados del grupo *M. chimaera/intracellulare* que no alcanzaron el módulo de

subtipificación, el 20 % (1/5) alcanzó alta confianza, mientras que el 80 % (4/5) alcanzó baja confianza; asimismo, el 80 % (4/5) presentó consistencia “B” y el 20 % (1/5) consistencia “C” (Figura 9).

Considerando que se buscaba identificar solo *Mycobacterium tuberculosis* y teniendo en cuenta la normalización taxonómica descrita en la sección de análisis estadístico, se alcanzó una sensibilidad del 100 % (IC 95 %: 97.9–100 %), una especificidad del 100 % (IC 95 %: 66.4%–100 %), un coeficiente de concordancia de Cohen ( $\kappa$ ) de 1.000 y un área bajo la curva ROC (AUC) de 1.000 (IC 95% 1.0–1.0).

Al considerar identificar solo Micobacterias no tuberculosas y al aplicar la normalización taxonómica descrita en plan de análisis, se alcanza una sensibilidad del 100% (IC 95%: 66.4%-100%), una especificidad del 100% (IC 95%: 97.9% - 100%) , un coeficiente de concordancia de Cohen(k) de 1 y un área bajo la curva ROC (AUC) de 1 (IC 1.0–1.0).

Considerando como criterio positivo la identificación a nivel de especie, la sensibilidad relativa del MALDI Biotyper Sirius frente al estándar de referencia fue de 2,15% (IC95% 0.59-5.41)

## V. DISCUSIÓN

La concordancia entre MALDI Biotyper Sirius y la secuenciación de genoma completo mediante Oxford Nanopore fue completa a nivel de complejo/grupo para *Mycobacterium tuberculosis complex* (MTBC; 177/177) y para el grupo *M. chimaera/intracellulare* (9/9). Sin embargo, estos resultados no deben interpretarse como un desempeño alto a nivel de especie. En este estudio, MALDI Biotyper Sirius resolvió únicamente 4/186 (2.2%) aislamientos a nivel de especie, mientras que la secuenciación asignó 186/186 aislamientos a nivel de especie. En conjunto, estos hallazgos indican que MALDI Biotyper Sirius es útil como cribado rápido para diferenciar MTBC vs MNT, mientras que la identificación taxonómica fina dependió de la secuenciación.

Estos hallazgos son consistentes con reportes previos que muestran un alto rendimiento de MALDI-TOF para distinguir MTBC de micobacterias no tuberculosas (MNT) [13]. Por ejemplo, un estudio chileno reportó 100% de identificación correcta de MTBC mediante MALDI-TOF [13]. De manera similar, una evaluación multicéntrica en Europa observó identificación correcta a nivel de complejo en el 100% de los casos analizados en 12 de 15 laboratorios [14]. En nuestro análisis, todos los aislamientos del complejo MTBC fueron reconocidos correctamente por MALDI Biotyper Sirius, lo que respalda su confiabilidad para el diagnóstico a nivel de complejo. Este punto tiene relevancia clínica, ya que la identificación rápida de MTBC facilita el inicio oportuno de medidas de control y prevención epidemiológica.

No obstante, MALDI Biotyper Sirius mostró limitaciones para la resolución de especies estrechamente relacionadas dentro de las MNT. En particular, solo 4 de 9 aislamientos del grupo *M. chimaera/intracellulare* fueron identificados a nivel de especie mediante el módulo de subtipificación. Este comportamiento concuerda con reportes que describen la dificultad de MALDI-TOF para distinguir *M. intracellulare* de *M. chimaera* debido a perfiles proteicos altamente similares. En línea con ello, la calidad interna del resultado (consistencia) también difirió entre grupos: la mayoría de los aislados de MTBC mostró consistencia A (82.49%, 146/177), mientras que el grupo *M. chimaera/intracellulare* presentó mayor incertidumbre, con consistencia A en 33.33% (3/9). En conjunto, estos datos sugieren que MTBC es más identificable de forma consistente por MALDI Biotyper Sirius que especies cercanas dentro de MAC.

Una explicación plausible para esta diferencia es la variabilidad operacional asociada al momento de procesamiento y a la edad fisiológica de la colonia. El diseño experimental contempló el procesamiento al día 21, con énfasis en micobacterias de crecimiento lento; sin embargo, las cepas del grupo *M. chimaera/intracellulare* corresponden a micobacterias de crecimiento relativamente más rápido, y para el día 21 podrían encontrarse en fase estacionaria. En esta fase se reportan cambios en la síntesis de proteínas ribosomales y en la composición de la membrana, lo que puede afectar el espectro obtenido y, por tanto, la capacidad del sistema para activar la subtipificación, la cual depende de picos biomarcadores específicos. Esta interpretación es concordante con Cuenod et al., 2021, donde se señala la edad de la colonia como un factor asociado a la calidad del espectro [15]. En este sentido, los resultados resaltan que la resolución a nivel de

especie por MALDI Biotyper Sirius puede ser sensible a condiciones preanalíticas y de procesamiento, especialmente cuando se evalúan especies estrechamente relacionadas.

Adicionalmente, se observó que el módulo de subtipificación identificó *M. intracellulare* en cuatro aislados con scores entre 1.86–1.96, todos con alta confianza; en uno de estos casos se registró consistencia C. Esto sugiere que la presencia de picos biomarcadores permitió la asignación de especie aun cuando el espectro global no alcanzó criterios de consistencia máxima. Por otro lado, se registró un aislado con score 2.02 en el que no se logró la subtipificación, lo que indica que un puntaje elevado no garantiza la resolución a nivel de especie si faltan picos discriminantes. En conjunto, estos hallazgos refuerzan que el desempeño del módulo de subtipificación puede variar según la calidad y composición del espectro, y no depende únicamente del score.

En contraste, Oxford Nanopore permitió asignación a nivel de especie en el 100% de los aislados y mostró parámetros de calidad genómica altos en todos los casos. Los genomas ensamblados consistieron en un solo contig, con tamaños entre 4.3–5.7 Mbp (compatibles con cromosomas bacterianos completos), valores de N50 equivalentes al tamaño genómico (indicando ensamblajes continuos) y contenido GC cercano a 66%, concordante con lo reportado para *Mycobacterium spp.*, particularmente *M. tuberculosis*, en estudios con Nanopore [16–19]. Asimismo, resultados previos con PacBio han reportado ensamblajes completos de *M. tuberculosis* en un solo contig (~4.41 Mbp) y GC ~65.6% [19], lo cual respalda que

los ensamblajes obtenidos en este trabajo son consistentes con referencias publicadas.

De manera similar, los genomas del grupo *M. intracellulare* mostraron tamaños mayores que los de *M. tuberculosis*, lo cual es esperado para especies de MAC [Imagen 6]. Estudios de secuenciación completa han reportado, por ejemplo, el genoma de *M. intracellulare* ATCC 13950 (5,402,402 pb; GC ~68.1%) y la cepa clínica MOTT-02 (5,409,696 pb), dentro del rango de ~5.3–5.4 Mb y ~68–69% GC descrito en distintas cepas [20–22]. Estas cifras son concordantes con lo observado en este estudio y respaldan la adecuada calidad del procesamiento y ensamblaje de estos aislados.

Desde el punto de vista clínico y epidemiológico, la imposibilidad de distinguir entre *M. chimaera* y *M. intracellulare* puede ser relevante, especialmente en el contexto de vigilancia de infecciones asociadas a atención de salud. *M. chimaera* se ha asociado a infecciones nosocomiales vinculadas a cirugías cardíacas con circulación extracorpórea [23,24]. No diferenciar entre ambas especies dificulta establecer si un caso corresponde a un evento asociado a brote hospitalario o a una fuente ambiental. Este problema se agrava considerando la latencia prolongada reportada para *M. chimaera*, con intervalos de hasta más de 6 años desde la cirugía hasta la presentación clínica [23]. Asimismo, se han reportado tasas de mortalidad elevadas en series clínicas, por ejemplo 50% en un estudio europeo (n=10) y 60% en un estudio del Reino Unido (n=30) [25,26]. Por ello, la identificación a nivel de especie puede ser clave para decisiones de control epidemiológico y seguimiento de casos.

Finalmente, la diferenciación a nivel de especie también puede aportar información para la optimización terapéutica. Se ha descrito que algunas cepas de *M. chimaera* presentan mayor resistencia a macrólidos en comparación con *M. intracellulare*, y los macrólidos forman parte del esquema estándar para infecciones por MAC (macrólido + rifampicina + etambutol). En ese contexto, la identificación específica podría ser clínicamente relevante, especialmente ante falla terapéutica [27,28]. En cuanto a perspectivas de mejora, la identificación microbiana por espectrometría de masas aún enfrenta retos por la necesidad de cultivo aislado para resultados confiables; sin embargo, se han propuesto estrategias para avanzar hacia identificación más rápida. Por ejemplo, se ha evaluado MALDI-TOF MS en esputo directo junto con examen microscópico, con sensibilidad de 97% y especificidad de 77% para diferenciar tuberculosis pulmonar de controles no TB [29], lo que sugiere que futuras optimizaciones podrían ampliar la aplicabilidad clínica de MALDI-TOF.

## **VI. LIMITACIONES**

Una limitación del estudio fue el uso de aislados criopreservados disponibles (186) para la presente tesis, lo que limita la generalización de los hallazgos a otros laboratorios o periodos. Además, el desempeño de MALDI Biotyper Sirius depende de la calidad del espectro y de la cobertura de la base de datos, por lo que especies poco representadas o taxonómicamente cercanas pueden presentar identificaciones con menor confianza o discrepancias. En cuanto a la secuenciación por Oxford Nanopore, la asignación de especie se realizó mediante un pipeline bioinformático específico; por ello, este aspecto debe considerarse al extrapolar la concordancia.

## VII. CONCLUSIONES

Objetivo 1: Bajo la normalización taxonómica, MALDI Biotyper Sirius mostró un desempeño muy alto para clasificar entre el complejo *M. tuberculosis* (MTBC) y las micobacterias no tuberculosas (MNT), con estimaciones compatibles con sensibilidad y especificidad cercanas al 100%. Asimismo, mostró un alto poder discriminante para diferenciar MTBC frente a MNT (AUC=1,000). En cambio, al exigir la identificación de la especie exacta, la sensibilidad relativa de MALDI Biotyper Sirius frente a Nanopore fue de 2,15%.

Objetivo 2: La concordancia a nivel de complejo/grupo entre ambas técnicas fue plena ( $\kappa=1,000$ ), lo que indica que MALDI Biotyper Sirius y Oxford Nanopore son intercambiables para diferenciar entre MTBC y MNT. A nivel de especie, la concordancia se reduce por la nula tasa de subtipificación de MALDI Biotyper Sirius en MTBC y baja en MNT.

## VIII. REFERENCIAS BIBLIOGRÁFICAS

1. World Health Organization. Global tuberculosis report 2022. Geneva, Switzerland: World Health Organization; 2022.
2. Alarcón V, Alarcón E, Figueroa C, Mendoza-Ticona A. Tuberculosis en el Perú: situación epidemiológica, avances y desafíos para su control. Rev Peru Med Exp Salud Publica. 2017;34(2):299. Available from: [http://www.scielo.org.pe/scielo.php?script=sci\\_arttext&pid=S1726-46342017000200021](http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1726-46342017000200021)
3. Barbeito Castiñeiras G, Coira Nieto MA, Pérez del Molino Bernal ML. Micobacterias no tuberculosas, más allá de la montaña mágica. Arch Bronconeumol .2021;57(3):156–7. Disponible en: <http://dx.doi.org/10.1016/j.arbres.2020.02.012>
4. Ministerio de Salud del Perú. Norma Técnica de Salud para el cuidado integral de la persona afectada por tuberculosis, familia y comunidad (NTS N° 200-MINSA/DGIESP-2023). Lima: Ministerio de Salud; 2023.
5. Cercenado Mansilla E, Cantón Moreno R, Oviaño García M, Rodríguez Sánchez B, Caballero Pérez Jd, Muñoz Bellido JL. Aplicaciones de la espectrometría de masas MALDI-TOF en Microbiología Clínica. Procedimientos en Microbiología Clínica. 2019;65., <https://seimc.org/contenidos/documentoscientificos/procedimientosmicrobiologia/seimc-procedimientomicrobiologia65.pdf>
6. Bacanelli, G., Araujo, F. R., & Verbisck, N. V. (2023). Improved MALDI-TOF MS identification of Mycobacterium tuberculosis by use of an enhanced cell

- disruption protocol. *Microorganisms*, 11(7), 1692.  
<https://doi.org/10.3390/microorganisms11071692>
7. Hall MB, Rabodoarivelo MS, Koch A, Dippenaar A, George S, Grobbelaar M, et al. Evaluation of Nanopore sequencing for *Mycobacterium tuberculosis* drug susceptibility testing and outbreak investigation: a genomic analysis. *Lancet Microbe* [Internet]. 2022;4(2):e84–92. Disponible en: [http://dx.doi.org/10.1016/S2666-5247\(22\)00301-9](http://dx.doi.org/10.1016/S2666-5247(22)00301-9)
  8. Legarraga P, Moraga M, Lam M, Geoffroy E, Zumarán C, García P. Impacto de la espectrometría de masas por MALDI-TOF MS en la identificación rápida de bacterias aeróbicas y anaeróbicas de importancia clínica. *Rev Chilena Infectol* [Internet]. 2013;30(2):140–6. Disponible en: <http://dx.doi.org/10.4067/s0716-10182013000200004>
  9. Sun X, Song J, Leng X, Li F, Wang H, He J, et al. A preliminary evaluation of targeted nanopore sequencing technology for the detection of *Mycobacterium tuberculosis* in bronchoalveolar lavage fluid specimens. *Front Cell Infect Microbiol* [Internet]. 2023;13. Disponible en: <http://dx.doi.org/10.3389/fcimb.2023.1107990>
  10. Wang Y, Yang Q, Wang Z. The evolution of nanopore sequencing. *Front Genet* [Internet]. 2015;5. Disponible en: <http://dx.doi.org/10.3389/fgene.2014.00449>
  11. Ministerio de Salud. Manual de pruebas moleculares para el diagnóstico bacteriológico y de sensibilidad de la tuberculosis [Internet]. Lima: MINSA; 2022 [citado el 1 de noviembre de 2024]. Disponible en : <https://bvs.minsa.gob.pe/local/fi-admin/RM-906-2022-MINSA.pdf>

12. Vásquez M, Balcells ME, García P, Moreno-Switt AI, Fredes-García D, Castillo C, Pizarro E. Identificación de micobacterias no tuberculosas en red de agua potable domiciliaria en la Región Metropolitana de Chile. *Rev Chilena Infectol.* 2025;42(3):205-215. doi:10.4067/S0716-10182025000300129.
13. Contreras S, Rodríguez D, Vera F, Balcells ME, Celis L, Legarraga P, Román JC, García P. Identificación de especies de micobacterias mediante espectrometría de masas (MALDI-TOF). *Rev Chil Infectol.* 2020;37(3). doi:10.4067/s0716-10182020000300252.
14. Rodriguez-Temporal D, Alcaide F, Mareković I, O'Connor JA, Gorton R, van Ingen J, et al. Multicentre study on the reproducibility of MALDI-TOF MS for nontuberculous mycobacteria identification. *Sci Rep.* 2022 Jan 24;12(1):1237. doi:10.1038/s41598-022-05315-7.
15. Cuénod A, Foucault F, Pflüger V, Egli A. Factors associated with MALDI-TOF mass spectral quality of species identification in clinical routine diagnostics. *Front Cell Infect Microbiol* [Internet]. 2021;11:646648. Disponible en: <http://dx.doi.org/10.3389/fcimb.2021.646648>
16. Ciuffreda L, Rodríguez-Pérez H, Flores C. Nanopore sequencing and its application to the study of microbial communities. *Comput Struct Biotechnol J* [Internet]. 2021;19:1497–511. Disponible en: <http://dx.doi.org/10.1016/j.csbj.2021.02.020>
17. Bainomugisa A, Duarte T, Lavu E, Pandey S, Coulter C, Marais BJ, et al. A complete high-quality MinION nanopore assembly of an extensively drug-resistant *Mycobacterium tuberculosis* Beijing lineage strain identifies novel

- variation in repetitive PE/PPE gene regions. *Microb Genom.* 2018 Jul;4(7):e000188. doi:10.1099/mgen.0.000188.
18. Andrievskaia O, Garceac A, Lloyd D, Savic M, Duceppe MO. Complete genome sequence of a *Mycobacterium bovis* strain associated with a bovine tuberculosis outbreak on a cattle farm in Saskatchewan, Canada. *Microbiol Resour Announc.* 2025 Feb 11;14(2):e0109024. doi:10.1128/mra.01090-24.
  19. Rodríguez JG, Pino C, Tauch A, Murcia MI. Complete genome sequence of the clinical Beijing-like strain *Mycobacterium tuberculosis* 323 using the PacBio real-time sequencing platform. *Genome Announc.* 2015 Apr 30;3(2):e00371-15. doi:10.1128/genomeA.00371-15.
  20. Kim BJ, Choi BS, Lim JS, Choi IY, Lee JH, Chun J, Kook YH, Kim BJ. Complete genome sequence of *Mycobacterium intracellulare* strain ATCC 13950(T). *J Bacteriol.* 2012 May;194(10):2750. doi:10.1128/JB.00295-12
  21. Kim BJ, Choi BS, Lim JS, Choi IY, Lee JH, Chun J, Kook YH, Kim BJ. Complete genome sequence of *Mycobacterium intracellulare* clinical strain MOTT-02. *J Bacteriol.* 2012 May;194(10):2771. doi:10.1128/JB.00365-12.
  22. Bouso JM, Planet PJ. Complete nontuberculous mycobacteria whole genomes using an optimized DNA extraction protocol for long-read sequencing. *BMC Genomics.* 2019 Oct 30;20(1):793. doi:10.1186/s12864-019-6134-y.
  23. Lecorche E, Pean de Ponfilly G, Mougari F, Benmansour H, Poisnel E, Janvier F, et al. Disseminated *Mycobacterium chimaera* following open-heart surgery, the heater-cooler unit worldwide outbreak: case report and minireview. *Front Med (Lausanne).* 2020 Jun 16;7:243. doi:10.3389/fmed.2020.00243.

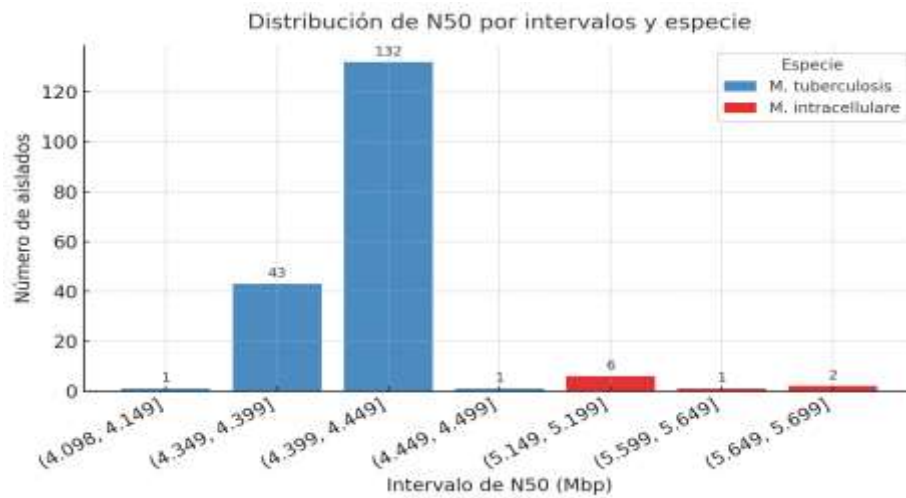
24. Diekema DJ. Mycobacterium Chimaera infections after cardiovascular surgery: Lessons from a global outbreak. *Trans Am Clin Climatol Assoc.* 2019;130:136–44.
25. Kohler P, Kuster SP, Bloemberg G, Schulthess B, Frank M, Tanner FC, et al. Healthcare-associated prosthetic heart valve, aortic vascular graft, and disseminated Mycobacterium chimaera infections subsequent to open heart surgery. *Eur Heart J* [Internet]. 2015;36(40):2745–53. Disponible en: <http://dx.doi.org/10.1093/eurheartj/ehv342>
26. Scriven JE, Scobie A, Verlander NQ, Houston A, Collyns T, Cajic V, et al. Mycobacterium chimaera infection following cardiac surgery in the United Kingdom: clinical features and outcome of the first 30 cases. *Clin Microbiol Infect* [Internet]. 2018;24(11):1164–70. Disponible en: <http://dx.doi.org/10.1016/j.cmi.2018.04.027> 30pacientes
27. Fernandez-Pittol M, Batista-Arnau S, Román A, San Nicolás L, Oliver L, González-Moreno O, et al. Differences in drug-susceptibility patterns between Mycobacterium avium, Mycobacterium intracellulare, and Mycobacterium chimaera clinical isolates: prospective 8.5-year analysis by three laboratories. *Antibiotics (Basel)*. 2023 Jan;12(1):64. Epub 2022 Dec 29. doi:10.3390/antibiotics12010064.
28. Kasperbauer SH, Daley CL. Mycobacterium chimaera infections related to the heater-cooler unit outbreak: A guide to diagnosis and management. *Clin Infect Dis* [Internet]. 2019;68(7):1244–50. Disponible en: <http://dx.doi.org/10.1093/cid/ciy789>

29. Dao TL, Hoang VT, Ly TDA, Lagier JC, Baron SA, Raoult D, et al. Sputum proteomic analysis for distinguishing between pulmonary tuberculosis and non-tuberculosis using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS): preliminary results. *Clin Microbiol Infect* [Internet]. 2021;27(11):1694.e1-1694.e6. Disponible en: <http://dx.doi.org/10.1016/j.cmi.2021.02.031>
30. Shrestha A, Picoy J, Torres A, Moore DA, Gilman RH, Coronel J, et al. A case report of transmission and disease caused by *Mycobacterium caprae* and *Mycobacterium bovis* in Lima, Peru. *BMC Infect Dis* [Internet]. 2021;21(1):1265. Disponible en: <http://dx.doi.org/10.1186/s12879-021-06944-5>
31. Gehre F, Kumar S, Kendall L, Ejo M, Secka O, Ofori-Anyinam B, et al. A Mycobacterial perspective on tuberculosis in West Africa: Significant geographical variation of *M. africanum* and other *M. tuberculosis* complex lineages. *PLoS Negl Trop Dis* [Internet]. 2016;10(3):e0004408. Disponible en: <http://dx.doi.org/10.1371/journal.pntd.0004408>
32. Bruker Daltonik GmbH. Release Notes – MBT Mycobacteria Library 5.0 (DB-912 MSP, #1850766). Revision B. Bruker; 2017 Apr. 9 p
33. Bruker. MALDI Biotyper: Standard Operating Procedure. Mycobacteria Extraction (MycoEX) Method. Revision 3. January 2014.
34. Zymo Research Corporation. Quick-DNA™ Miniprep Plus Kit (Catalog numbers: D4068, D4068T, D4069). Instruction manual. Ver. 2.0.1
35. Elton L, Aydin A, Stoker N, Rofael S, Wildner LM, Abdul JBPAA, et al. A pragmatic pipeline for drug resistance and lineage identification in

Mycobacterium tuberculosis using whole genome sequencing. Muñiz-Salazar R, editor. PLOS Global Public Health. 2025 Feb 10;5(2):e0004099.

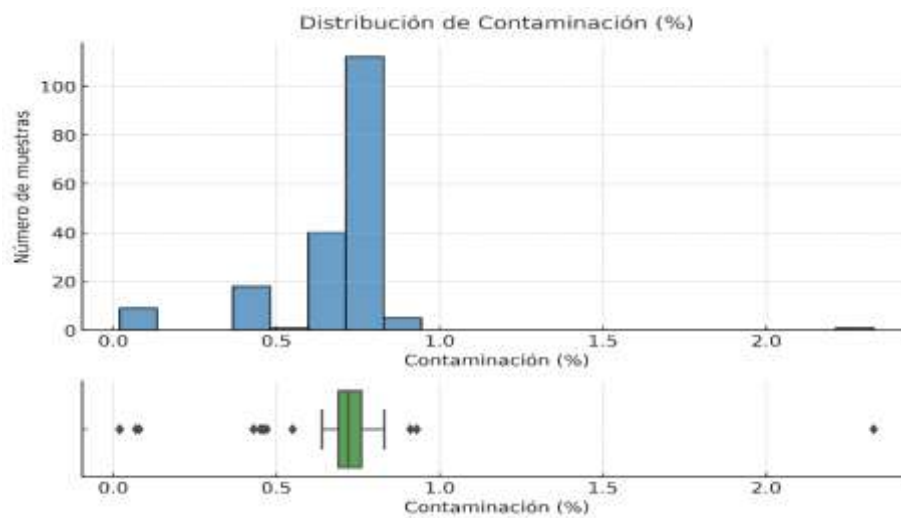
36. Oxford Nanopore Technologies. MinION: Protocol. Ligation sequencing gDNA – Native Barcoding Kit 96 V14 (SQK-NBD114.96). Version NBE\_9171\_v114\_revT\_03Oct2025. 03 Oct 2025

## IX. TABLAS, GRÁFICOS Y FIGURAS



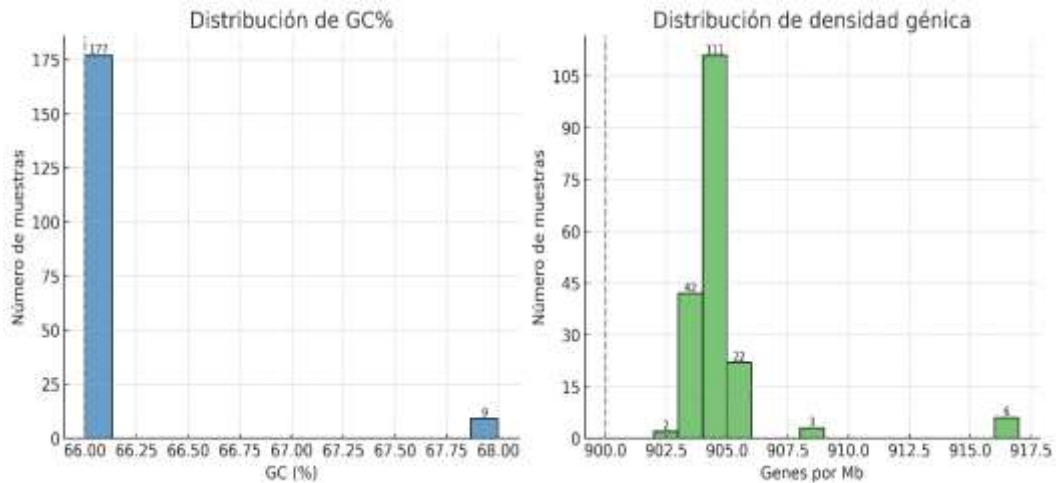
**Figura 1. Distribución del N50 por intervalos y especie.**

El gráfico presenta el número de aislados en cada intervalo de N50 (Mbp), agrupados según la especie identificada: *M. tuberculosis* (azul) y *M. intracellulare* (rojo). Los valores sobre las barras indican la frecuencia de aislados por intervalo.



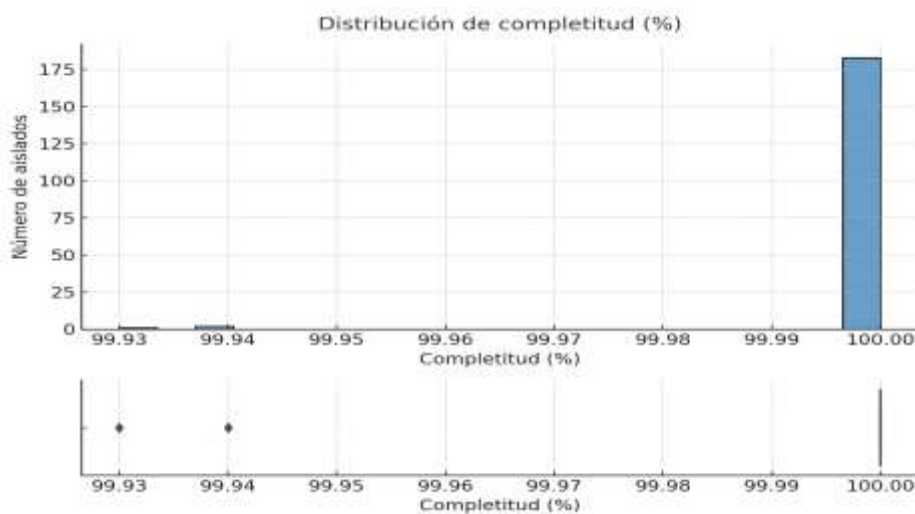
**Figura 2. Distribución de los niveles de contaminación en los genomas ensamblados.**

Se muestra la frecuencia de los aislados en función del porcentaje de contaminación estimado, mientras que el boxplot inferior resume su distribución estadística. La gran mayoría de los genomas presentan valores de contaminación entre 0.02% y 0.91%, con una mediana cercana a 0.6%, sin presencia de outliers significativos.



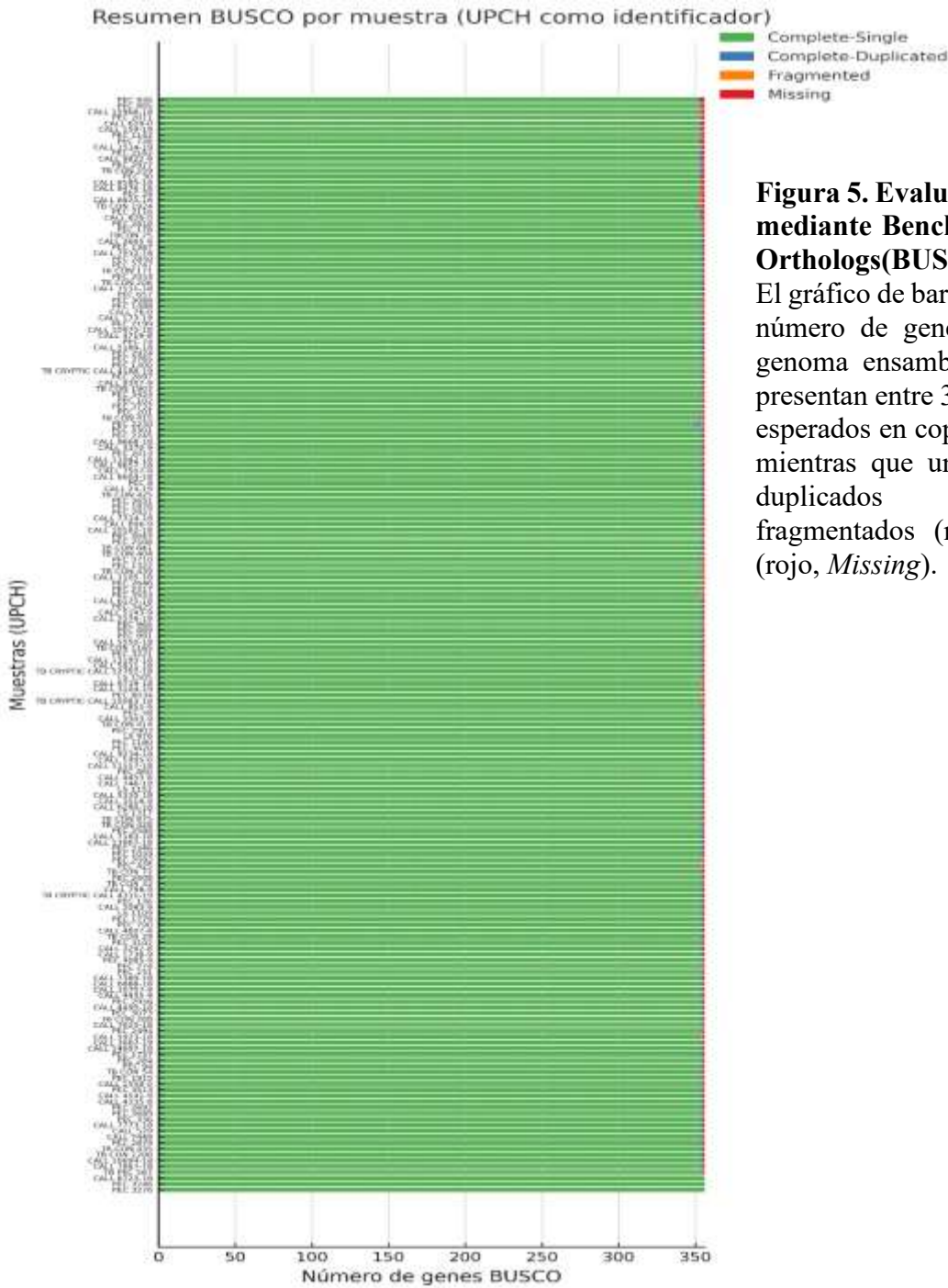
**Figura 3. Distribución del contenido de Guanina y Citocina (GC) y densidad génica en los genomas ensamblados.**

Los histogramas muestran la frecuencia de los aislados según el porcentaje de GC (izquierda) y la densidad génica expresada como número de genes por megabase (derecha). En el caso del contenido de GC, prácticamente todos los aislados se concentran estrechamente alrededor del 66%, sin variación significativa entre genomas. De manera similar, la densidad génica se distribuye en torno a 900 genes/Mb.



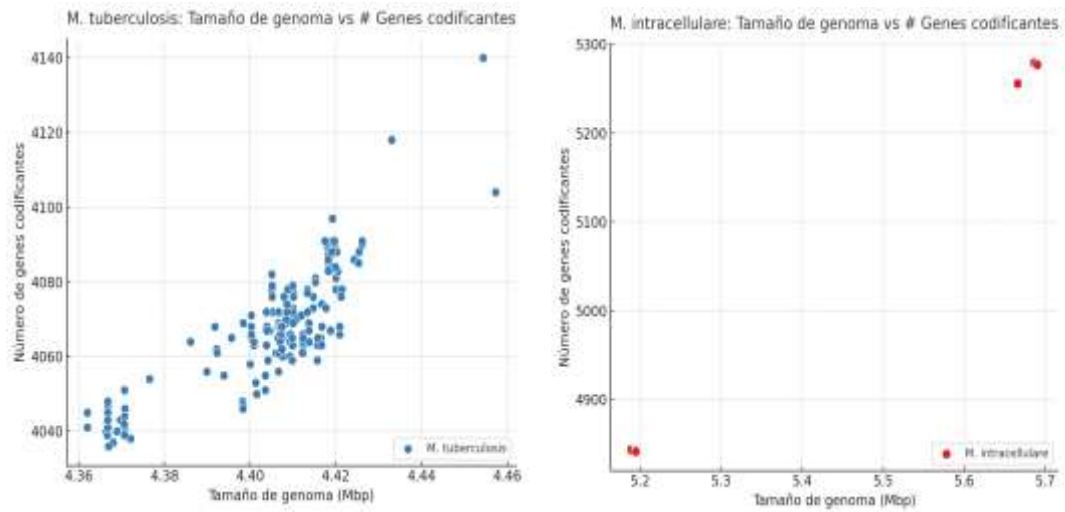
**Figura 4. Distribución del porcentaje de completitud en los genomas ensamblados.**

Se muestra la frecuencia de los aislados en función de su porcentaje de completitud, mientras que el boxplot inferior resume su distribución. La gran mayoría de los genomas presentan valores de completitud cercanos al 100%, sin presencia de valores atípicos ni dispersión significativa.



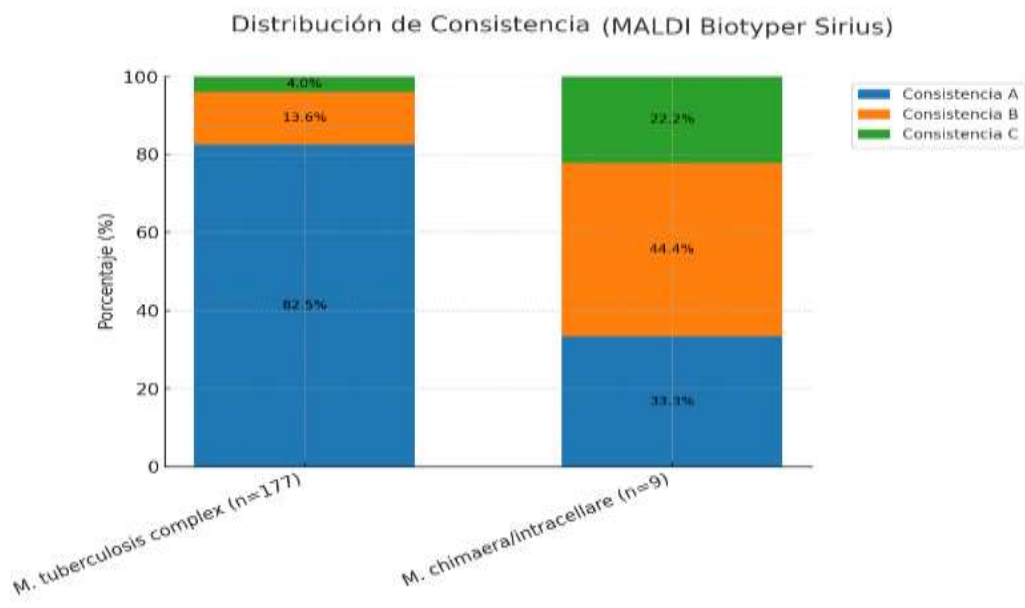
**Figura 5. Evaluación de la completitud genómica mediante Benchmarking Universal Single-Copy Orthologs(BUSCO).**

El gráfico de barras horizontales apiladas muestra el número de genes BUSCO identificados en cada genoma ensamblado. La mayoría de los genomas presentan entre 354 y 355 de los 356 genes BUSCO esperados en copia única (verde, *Complete-Single*), mientras que un número reducido aparece como duplicados (azul, *Complete-Duplicated*), fragmentados (naranja, *Fragmented*) o ausentes (rojo, *Missing*).



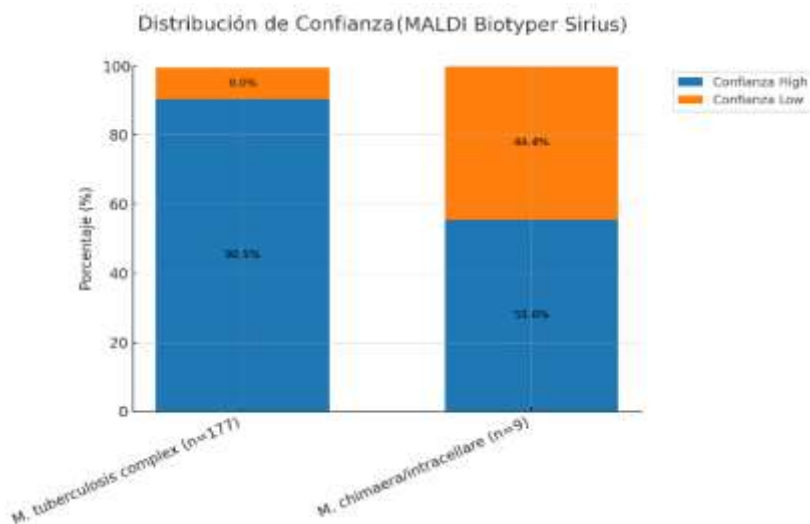
**Figura 6.** Comparación del tamaño genómico y número de genes codificantes entre *M. tuberculosis* y *M. intracellulare*.

El gráfico de dispersión muestra la relación entre el tamaño del genoma (Mbp) y el número total de genes codificantes predichos para cada muestra. Las cepas de *M. tuberculosis* (azul) se agrupan de manera compacta en torno a ~4.3–4.4 Mbp y ~4000 genes, reflejando genomas más reducidos y altamente conservados. En contraste, los aislados clasificados como *M. intracellulare* (rojo) presentan genomas considerablemente más grandes (5.1–5.7 Mbp) y un mayor número de genes (4800–5300).



**Figura 7. Distribución porcentual de la consistencia por genero en el sistema MALDI Biotyper sirius**

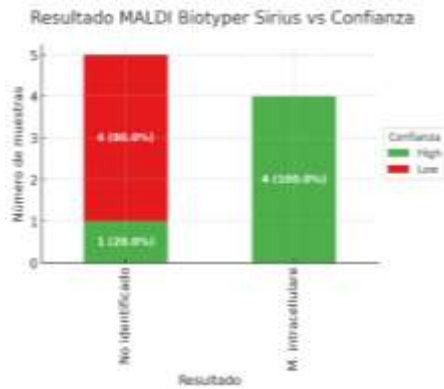
La figura muestra un gráfico de barras apiladas que compara dos grupos taxonómicos: *Mycobacterium tuberculosis complex* y *M. chimaera/intracellare group*. Cada barra está segmentada según el porcentaje de aislamientos clasificados en las categorías de consistencia A (azul), B (anaranjado) y C (verde). Para *M. tuberculosis complex* se obtuvieron 146 aislamientos en consistencia A, 24 en consistencia B y 7 en consistencia C. En el caso de *M. chimaera/intracellare group*, 3 aislamientos correspondieron a consistencia A, 4 a consistencia B y 2 a consistencia C.



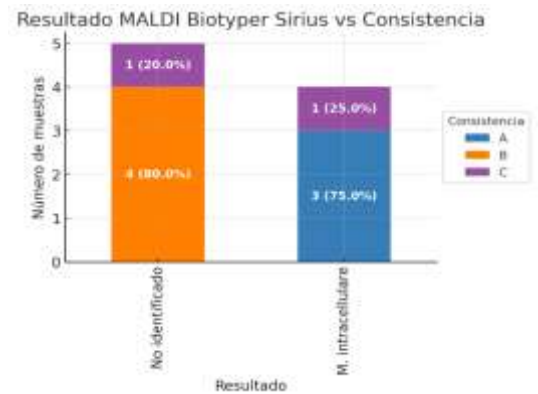
**Figura 8. Distribución porcentual del parámetro de confianza en el sistema MALDI Biotyper sirius**

La figura muestra un gráfico de barras apiladas que compara dos grupos taxonómicos: *Mycobacterium tuberculosis complex* y *M. chimaera/intracellare group*. Cada barra está segmentada según el porcentaje de aislamientos clasificados en las categorías de confianza alta (azul) y confianza baja (anaranjado).

A



B



**Figura 9. A) Distribución porcentual del parámetro de confianza en el sistema MALDI Biotyper sirius respecto a *M.chiamaera/intracellulare* group** La figura muestra un gráfico de barras apiladas que compara dos grupos: *M. intracellulare*, identificada a nivel de especie, y las que no lograron ser identificadas a nivel de especie. Cada barra está segmentada según el porcentaje de aislamientos clasificados en las categorías de confianza alta (verde) y confianza baja (roja). **B) Distribución porcentual del parámetro de consistencia en el sistema MALDI Biotyper sirius respecto a *M.chiamaera/intracellulare* group.** La figura muestra un gráfico de barras apiladas que compara dos grupos: *M. intracellulare*, identificada a nivel de especie, y las que no lograron ser identificadas a nivel de especie. Cada barra está segmentada según el porcentaje de aislamientos clasificados en las categorías de consistencia A (azul), B (anaranjado) y C (morado)

Tabla 1.

### Distribución de identificaciones por MALDI Biotyper Sirius en 186 aislados

Resultado MALDI Biotyper Sirius	Frecuencia	Porcentaje	Porcentaje acumulado
<i>Mycobacterium chimaera/intracellulare</i> group	9	4.84	4.84
<i>Mycobacterium tuberculosis</i> complex	177	95.16	100
Total	186	100	

Tabla 2.

**Consistencia de la identificación por MALDI Biotyper Sirius según especie**

Especie identificada por MALDI Biotyper sirius	Consistencia A	Consistencia B	Consistencia C	Total
<i>Mycobacterium chimaera/intracellulare_group</i>	3 (33.33%)	4 (44.44%)	2 (22.22%)	9 (100%)
<i>Mycobacterium tuberculosis_complex</i>	146 (82.49%)	24 (13.56%)	7 (3.95%)	177 (100%)
Total	149 (80.11%)	28 (15.05%)	9 (4.84%)	186 (100%)

Tabla 3.

**Nivel de confianza del resultado MALDI Biotyper Sirius por grupo taxonómico.**

Especie identificada	Confianza High	Confianza Low	Total
<i>Mycobacterium chimaera_intracellulare_group</i>	5(55.56%)	4 (44.44%)	9(100%)
<i>Mycobacterium tuberculosis_complex</i>	161 (90.45%)	16 (9.04%)	177(100%)
Total	166(88.77%)	20(10.75%)	186(100%)

Tabla 4.

**Métricas Diagnosticas IC 95%**

Métrica	Estimación	IC Lower	IC Upper
Sensibilidad	1	0.97	0.99
Especificidad	1	0.7	0.99
Kappa de Cohen	1	1	1
ROC AUC	1	1	1

Tabla 5.

**Interpretación de *log(score)* en MALDI Biotyper sirius para micobacterias**

<b>Rango de <i>log(score)</i></b>	<b>Interpretación</b>	<b>Señal en reporte (referencial)</b>
1,80 – 3,00	Identificación de <b>alta confianza</b>	Verde / “+++”
1,60 – 1,79	Identificación de <b>baja confianza</b> (requiere confirmación)	Amarillo / “+”
0,00 – 1,59	<b>No</b> es posible una identificación fiable	Rojo / “-”

Tabla 6.

**Definición de categorías de consistencia (A–C) en MALDI Biotyper**

<b>Categoría</b>	<b>Criterio operativo (cómo se asigna)</b>	<b>Lectura práctica</b>
A (alta consistencia)	La mejor coincidencia es una identificación de alta confianza. La segunda mejor es: (1) alta confianza con misma especie, o (2) baja confianza con misma especie o mismo género, o (3) sin identificación.	El ranking respalda fuertemente al primer hit (resultado robusto).
B (baja consistencia)	No cumple criterios de A. La mejor coincidencia puede ser de alta o baja confianza. La segunda mejor es: (1) alta o baja confianza con mismo género, o (2) sin identificación.	Hay coherencia parcial (al menos a nivel de género), pero menor solidez.
C (sin consistencia)	No cumple criterios de A ni de B.	Coherencia insuficiente entre los hits (resultado poco fiable).

## ANEXOS

### I. Agrupaciones de micobacterias con espectros MALDI Biotyper Sirius similares:

(1) *M. austroafricanum* con *M. vanbaallenii*; (2) *M. canariasense* con *M. cosmeticum*; (3) *M. chubuense*, *M. chlorophenicum* y *M. psychrotolerans*; (4) *M. crocinum* con *M. pallens*; (5) miembros del complejo *M. fortuitum* (especialmente *M. farcinogenes*, *M. fortuitum* subsp. *fortuitum*, *M. porcinum* y *M. senegalense*); (6) *M. gadium* con *M. tusciae*; (7) *M. intracellulare* con *M. chimaera*; (8) *M. mucogenicum* con *M. phocaicum*; (9) *M. murale* con *M. tokaiense*; (10) *M. sherrisii* con *M. simiae*; (11) miembros del complejo *M. tuberculosis*; (12) *M. chelonae* con *M. salmoniphilum* (posibles coincidencias de alta confianza en casos muy raros, distinguibles por pruebas fisiológicas o secuenciación); y (13) *M. aichiense* con *M. aromaticivorans*.

### II. Glosario de Términos

- MALDI Biotyper sirius: Plataforma de espectrometría de masas MALDI-TOF (desorción/ionización láser asistida por matriz-tiempo de vuelo) utilizada para la identificación de microorganismos.
- OXFORD NANOPORE: Tecnología de secuenciación de ADN de “lectura larga” que determina la secuencia al hacer pasar moléculas individuales por un nanoporo y medir cambios en la corriente eléctrica; permite obtener lecturas extensas y realizar secuenciación de genoma completo para identificación taxonómica
- WGS (Whole Genome Sequencing / secuenciación de genoma completo): Método que obtiene la secuencia del genoma completo de un microorganismo, permitiendo identificarlo con alta resolución.
- BAAR (bacilo ácido-alcohol resistente): Bacteria cuya pared rica en ácidos micólicos resiste la decoloración con ácido-alcohol (por ejemplo *M. tuberculosis*).
- Barcoding (etiquetado): Adición de secuencias cortas (barcodes) a bibliotecas para mezclar varias muestras y luego separar lecturas por muestra (demultiplexación).
- Basecalling: Conversión de señales crudas (eléctricas en ONT) a bases A/C/G/T con un modelo (por ejemplo Dorado).
- BUSCO: Herramienta para estimar “completitud” de un ensamblaje/genoma a partir de genes ortólogos conservados.
- CheckM2: Herramientas para estimar completitud y contaminación de genomas ensamblados a partir de marcadores.
- GTDB-Tk: Herramienta para clasificar taxonómicamente genomas bacterianos/arqueanos según la base GTDB.
- N50 (de ensamblaje): Longitud L tal que el 50 % de la suma total del ensamblaje está en contigs con longitud  $\geq L$  (contigs ordenados de mayor a menor).

- Contig: Secuencia de ADN contigua reconstruida por el ensamblador a partir de lecturas solapadas. y representa un fragmento continuo del genoma.
- Mbp (mega pares de bases): Unidad que equivale a un millón de pares de bases ( $1 \text{ Mbp} = 10^6 \text{ bp}$ ); se usa para expresar el tamaño de genomas o de secuencias/ensamblajes