



Universidad Autónoma del Estado de México
Facultad de Odontología

Centro de Investigación y Estudios Avanzados en Odontología
“Dr. Keisaburo Miyata”

“Detección del potencial productor de moléculas de importancia biomédica de *Nocardia brasiliensis* por medio de análisis genómico comparativo”

TESIS

Que para obtener el Grado de:
Doctora en Ciencias de la Salud

Presenta

M. en C.S. Michele Guadalupe Cruz Medrano

Núm. Cuenta: 1020178

CVU: 999358

Directora

Dra. en C.B. Ninfa Ramírez Durán

Co-Directora

Dra. en C.S. Gaudy Lizeth Manzanares Leal

Tutor

Dr. en C. Ayixon Sánchez Reyes



2022-2026

Toluca, Estado de México, octubre de 2025

ÍNDICE

RESUMEN	4
INTRODUCCIÓN	5
1. ANTECEDENTES	6
1.1 Genómica microbiana	6
1.2 Clústeres de genes biosintéticos y productos naturales	7
1.2.1 Clústeres de genes biosintéticos	7
1.2.2 Productos naturales microbianos	7
1.2.3 Policétido sintasa (PKS)	8
1.2.4 Sintetasa de péptido no ribosomal (NRPS)	10
1.2.5 Métodos de detección de clústeres de genes biosintéticos en bacterias ..	13
1.2.6 Análisis genómico para la detección de clústeres de genes biosintéticos .	15
1.2.7 Aplicaciones y utilidades de los productos naturales	18
1.2.8 Principales bacterias portadoras de clústeres de genes biosintéticos	19
1.3 <i>Nocardia brasiliensis</i>	22
1.3.1 Métodos de identificación de <i>Nocardia brasiliensis</i>	23
1.3.2 Variabilidad fenotípica y genotípica intraespecie	23
1.3.3 Potencial patogénico	24
1.3.4 Potencial biotecnológico	25
2. PLANTEAMIENTO DEL PROBLEMA	29
3. JUSTIFICACIÓN	30
4. HIPÓTESIS	32
5. OBJETIVOS	33
6. MATERIALES Y MÉTODOS	34
6.1 Diseño del estudio	34
6.2 Universo	34
6.3 Muestra	34
6.4 Criterios de inclusión, exclusión y eliminación	34
6.5 Variables	35

6.6	Procedimiento	36
6.7	Consideraciones bioéticas	54
6.8	Análisis estadístico	54
7.	RESULTADOS	56
8.	DISCUSIÓN	65
9.	CONCLUSIONES	69
10.	REFERENCIAS BIBLIOGRÁFICAS	72

RESUMEN

Las actinobacterias generan la mayoría de los metabolitos secundarios o productos naturales aprovechables en el área médica para la producción de fármacos. Mientras que *Streptomyces* spp. encabezan esta producción, actinobacterias patógenas como *Nocardia* spp. han demostrado tener la capacidad para producir metabolitos con actividad biológica, pero no se han considerado como fuente de productos naturales. *Nocardia brasiliensis* es una actinobacteria patógena causante del actinomicetoma y la nocardiasis, con gran impacto en América Latina y el Caribe, pero escasamente estudiada. Es importante buscar nuevas fuentes de productos naturales para responder a problemas de salud pública mundial como la resistencia a los antimicrobianos. Este trabajo tiene como objetivo detectar el potencial productor de moléculas de importancia biomédica de *N. brasiliensis* aislada de actinomicetoma en México por medio del análisis genómico comparativo. Se incluyeron 65 cepas clínicas de *N. brasiliensis* de origen mexicano y se eligieron cuatro cepas para la extracción de DNA y secuenciación genómica. Se incluyeron 22 genomas de acceso público, con un total de 26 genomas. Se realizó la identificación genómica, se estimó la composición del pangenoma, el genoma central y la proporción de genes ortólogos para el análisis filogenómico. Se buscaron BGCs en los genomas, se identificaron los metabolitos secundarios relacionados y la actividad biológica reportada. Se calculó un pangenoma abierto, compuesto por 17,715 clústeres, el genoma central representó el 22.76% y el genoma diferencial el 48.17% de los clústeres. Se encontró que las diferencias evolutivas estaban influenciadas por el tipo de muestra y origen geográfico de las cepas. Se detectaron BGCs (PKS, NRPS, híbridos y otros) que codifican para metabolitos secundarios con actividad antimicrobiana (ϵ -Poli-L-lisina, brasiliquinonas A-B), antitumoral (rizomidas A-C, antramicina), antioxidante (isorenierateno), entre otros. Se demostró que las cepas de *N. brasiliensis* aisladas de actinomicetoma en México tienen potencial para producir moléculas con actividad biomédica.

INTRODUCCIÓN

Nocardia brasiliensis es una actinobacteria presente en el medio ambiente, tradicionalmente conocida como patógena y uno de los agentes causales más comunes del actinomicetoma y la nocardiosis. Estas patologías suelen diagnosticarse de forma tardía y en presentaciones avanzadas, periodo en el cual las secuelas son comunes. A pesar del impacto que presentan, poco se estudia acerca de estas enfermedades y de sus agentes causales.^{1,2}

En 2021, se estimó que se realizan únicamente 20 publicaciones por año sobre estas enfermedades y sus agentes causales y, únicamente, se cuentan con 22 publicaciones sobre *N. brasiliensis* entre 2017 y 2022.³ Dentro de estas publicaciones, revisiones actualizadas reportan que *N. brasiliensis* no solo es relevante clínicamente, sino que puede tener un importante papel en la biomedicina derivado de su potencial productor de moléculas de importancia biomédica. A partir de esta actinobacteria, han sido extraídas sustancias con actividad antibiótica, antineoplásica e inmunomoduladora, sin embargo, continúan siendo pocos los estudios referidos.⁴ De especial relevancia es la limitación de estudios genómicos que permitan detectar los clústeres de genes biosintéticos o BGC que codifiquen para moléculas bioactivas que sean de relevancia para la salud humana.

Derivado de lo anterior, este trabajo tiene como propósito realizar un análisis genómico comparativo de cepas de *N. brasiliensis* que permita detectar los clústeres de genes biosintéticos que presentan, así como la identificación de las moléculas de importancia biomédica relacionadas y su actividad. En este trabajo, se presentan inicialmente los antecedentes de relevancia sobre la genómica microbiana, los clústeres de genes biosintéticos y los productos naturales microbianos, así como los estudios previos realizados en *N. brasiliensis*.

1. ANTECEDENTES

1.1 Genómica microbiana

La genómica microbiana es la secuenciación de los genomas completos y su análisis.⁵ Sus primeros antecedentes datan de 1995 cuando fueron publicadas las secuencias completas de *Haemophilus influenzae*⁶ y *Mycoplasma genitalium*,⁷ permitiendo conocer de ese modo la descripción de todos los genes presentes en el DNA de ambas bacterias.

La secuenciación de los genomas provee de ventajas cuando se compara con la de un solo gen o un segmento de éste; permitiendo conocer y comprender, según los genes presentes, su tipo de organización, relación con su ambiente, patogenicidad y su evolución. Permite también realizar estudios de genómica comparativa entre distintas especies o cepas de la misma especie. De forma técnica, además, evita la manipulación de las secuencias para su estudio. Los cromosomas bacterianos son, en su gran mayoría, circulares y tienen una longitud entre 500 a 10, 000 kb, lo cual permite que la secuenciación completa sea posible y práctica.^{8,9}

Para el estudio de una especie bacteriana patógena, la secuenciación genómica puede aportar datos que permitan explicar las vías patógenas, la variabilidad relacionada con la citotoxicidad y la virulencia, permitiendo así ampliar el panorama epidemiológico subyacente.⁵

El análisis genómico permite también la construcción de árboles filogenómicos, los cuales utilizan secuencias genómicas para realizar una determinación de las relaciones evolutivas existentes entre diferentes especies de estudio. A diferencia de los estudios filogenéticos, que incluyen únicamente algunos genes específicos, la filogenómica utiliza para su estudio los genomas de las especies de interés.¹⁰

La genómica comparativa es una técnica que se ha utilizado con más frecuencia a partir de la obtención de genomas de cualquier especie, en la cual, por medio de distintas herramientas bioinformáticas, se buscan aquellas diferencias y semejanzas con otros genomas ya conocidos y depositados en bases de datos de acceso público.¹¹

Por otro lado, considerando que muchas bacterias no causan enfermedad, sino que pueden ser aprovechadas por el ser humano por medio de la biotecnología, los genomas completos de las especies de interés permiten conocer las capacidades metabólicas o de síntesis, como la presencia de clústeres de genes biosintéticos y productos naturales, así como las posibles aplicaciones que éstas puedan tener y su utilidad en la identificación y el desarrollo de moléculas de interés biomédico, de biorremediación e industrial.^{12,13}

1.2 Clústeres de genes biosintéticos y productos naturales

1.2.1 Clústeres de genes biosintéticos

También denominados grupos de genes biosintéticos (BGCs, por sus siglas en inglés), son un grupo de genes que se encuentran en proximidad y que, en conjunto, codifican para la síntesis de metabolitos secundarios por medio de rutas bioquímicas complejas.¹³ En una revisión realizada en 2021,¹⁴ se pudo hallar que la mayoría de los clústeres de genes biosintéticos encontrados en bacterias desde el 2017 provenían de actinobacterias, entre las cuales se encuentran *Streptomyces* sp., *Streptomyces artemisiae*, *Streptomyces avermitilis*, *Streptomyces seoulensis*, *Streptomyces levis*, *Streptomyces conglobatus*, *Actinoallomurus fulvus* y *Micromonospora zhangzhouensis*; mientras que entre otras especies con clústeres de genes biosintéticos están *Bacillus amyloliquefaciens*, *Pseudomonas fluorescens*, *Faecalibacterium prausnitzii*, *Lactobacillus rhamnosus* e incluso bacterias de la microbiota humana.

1.2.2 Productos naturales microbianos

Son todas aquellas moléculas que produce una bacteria, entre estas pueden encontrarse a los metabolitos primarios y los secundarios. Los metabolitos primarios son aquellos esenciales para la vida del microorganismo, mientras que los secundarios son aquellos que no se requieren estrictamente pero que interactúan con elementos intrínsecos o del medio ambiente.^{13,15}

Los policétidos (PKs) y los péptidos no ribosomales (NRPs) son dos importantes grupos de metabolitos secundarios, de utilidad clínica e industrial, sintetizados por las

enzimas policétido sintasa (PKS, por sus siglas en inglés) y sintetasa de péptido no ribosomal (NRPS, por sus siglas en inglés), respectivamente.¹⁶⁻¹⁸

1.2.3 Policétido sintasa (PKS)

Los policétidos sintasa son estructuras enzimáticas conformadas por módulos, de estructura compleja, donde participan dominios que realizan funciones específicas en la síntesis de los policétidos, como se observa en la **Figura 1**. Algunos elementos en la vía de síntesis realizada por los PKS son comparables a los realizados en la biosíntesis de ácidos grasos, con la diferencia de que los PKS pueden producir moléculas con mayor complejidad.^{19,20}

Los dominios catalizadores que componen a los PKS¹⁹⁻²² son los siguientes:

- Dominio de aciltransferasa (AT)
- Dominio de tiolación (T) o proteína transportadora de acilo (ACP)
- Dominio de cetosintasa (KS)

Los PKS presentan dominios accesorios,^{19,20,22} entre los cuales se encuentran:

- Dominio de tioesterasa (TE)
- Dominio de ceto-reductasa (KR)
- Dominio de deshidratasa (DH)
- Dominio de enoil-reductasa (ER)
- Dominio de metiltransferasa (MT)

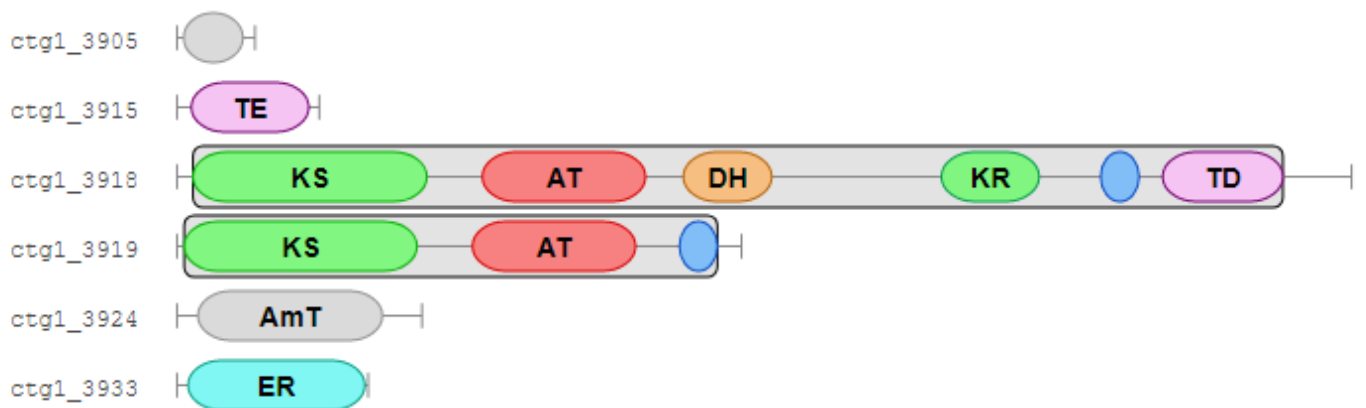


Figura 1. Estructura modular de un PKS I (Obtenido de antiSMASH²³)

Durante la producción de policétidos, el dominio AT realiza el reconocimiento y activación del estrato, posteriormente la molécula en formación será transportada por una ACP, que le llevará al dominio KS, donde ocurre la condensación descarboxilativa de Claisen y se permite la elongación de la cadena. Este proceso se repite, mientras el dominio AT añade a la molécula derivados de acetil-CoA (con mayor afinidad, malonil-CoA o metilmalonil-CoA), que son transportados al dominio KS, para finalmente ser liberada por el dominio TE, por medio de hidrólisis o macrociclación, una vez que la molécula ha completado su estructura. Los dominios accesorios realizarán modificaciones a las moléculas, derivando de ello la diversidad de productos^{19,22}. Por sus diferencias estructurales, los PKS se subdividen en tres, PKS I, PKS II y PKS III.²⁰

Algunos de los productos naturales sintetizados por los PKS son antibióticos, antiparasitarios, medicamentos de uso oncológico y moduladores inmunológicos,²⁰ como se muestra en la **Tabla 1**.

Tabla 1. Características de los tipos de PKS

Tipo	PKS I	PKS II	PKS III
Estructura	Multidominios complejos	Dominios sencillos	Solo un dominio
Organismos en los cuáles se presentan	Bacterias y hongos	Bacterias	Plantas y bacterias
Especies donde se han hallado	<i>Chaetomium globosum</i> ; <i>Streptomyces seoulensis</i>	<i>Streptomyces venezuelae</i> ; <i>Streptomyces reseiscleroticus</i>	<i>Myxococcus xanthus</i> ; <i>Streptomyces coelicolor</i>
Metabolitos producidos	Quetoglobosina A, Ansaseomicina A y B	Jadomicina hidroxí-4AMBA; cromomicina	Alquilpirona-407 y393; germicidina
Actividad	Control de patógenos en agricultura; citotóxico	Citotóxico; antineoplásico	Antimicrobianos
Referencia	Gherbawy y El-Dawy, 2020; ²⁴ Liu y col., 2019 ²⁵	Robertson y col., 2018; ²⁶ Sun y col., 2018 ²⁷	Joachim y col., 2019; ²⁸ Čihák y col., 2017 ²⁹

1.2.4 Sintetasa de péptido no ribosomal (NRPS)

Las sintetetasas de péptido no ribosomal son estructuras complejas, compuestas por módulos, en los cuales se llevan a cabo distintas reacciones químicas, como la modificación de carboxilos y aminoácidos. Los módulos que componen a los NRPS pueden diferir en su número y estructura, según el metabolito que producen. Cada módulo contiene, al menos, dominios de adenilación y tiolación.^{19,21} Existen dominios catalizadores y dominios accesorios dentro de sus estructuras, como se observa en la **Figura 2.**

Los dominios catalizadores que componen los NRPS^{19,21} son los siguientes:

- Dominio de adenilación (A): Encargada de la selección, reconocimiento y activación del sustrato o aminoácido. La selección del sustrato es altamente específica y puede ser de ayuda al momento de realizar la predicción de productos.
- Dominio de tiolación (T) o proteína transportadora de péptidos (PCP): Realiza el transporte de los péptidos y de las proteínas.
- Dominio de condensación (C): Realiza los enlaces peptídicos de carbono-nitrógeno, permitiendo la elongación de la molécula. La determinación del tipo de dominio de condensación, específico para distintos aminoácidos, es de utilidad para la predicción de los productos.

Algunos de los módulos accesorios dentro de los NRPS^{19,21} se enlistan a continuación.

- Dominio de tioesterasa (TE): Realiza la liberación del producto final por medio de hidrólisis o ciclación.
- Dominio de epimerización (E)
- Dominio de metiltransferasa (MT)

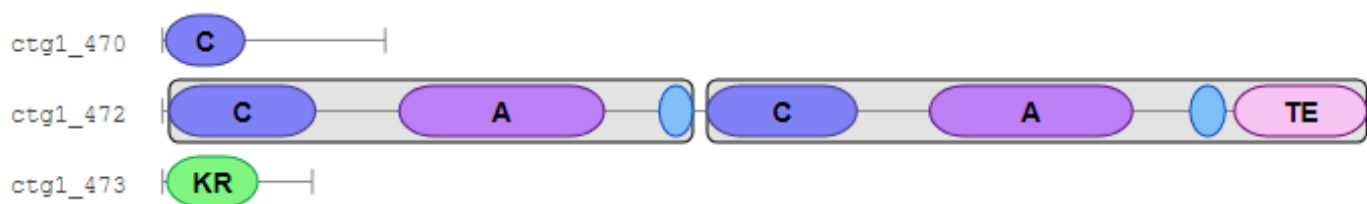


Figura 2. Estructura modular de un NRPS (Obtenido de antiSMASH (29))

La diversidad de los NRPS depende del tipo y cantidad de módulos accesorios que contenga en su estructura, principalmente.²¹ La formación de los péptidos se lleva a cabo en un proceso en el cual ocurre el reconocimiento y activación del aminoácido en el dominio A, la condensación subsecuente en el módulo C que prolonga la cadena peptídica, mientras ésta se mantiene unida a los PCP. El proceso se repetirá hasta completar la estructura del péptido final y ocurrir su liberación en el dominio TE.¹⁹

Por su estructura y ensamblaje de moléculas, los NRPS se subdividen en 2 tipos; NRPS I y NRPS II.^{30,31} Dentro de los metabolitos secundarios que pueden ser producidos por los NRPS, se encuentran los lipopéptidos (estructuras lipídicas) y los sideróforos (compuestos que se producen en medios de poca disponibilidad de hierro, ayudando al transporte, almacenamiento y competición por el elemento),¹⁹ como se muestra en la **Tabla 2**.

Tabla 2. Características de los tipos de NRPS

Tipo	NRPS I	NRPS II
Estructura involucrada en el ensamblaje de la molécula	Megaenzimas	Proteínas de transporte
Tipo de ensamblaje	Incorporación de un aminoácido en cada módulo de manera secuencial	Incorporación de los aminoácidos en dominios coordinados en cada dominio, transportados por proteínas.
Especies donde se han hallado	<i>Kroppenstedtia sanguinis</i>	<i>Streptomyces griseovariabilis</i> ; <i>Pseudomonas syringae</i>
Metabolitos producidos	Sideróforos; macrólidos	Ecinomicina/Quinomicina A; Coronatina
Actividad	No determinada; antimicrobiano	Antimicrobiano; antineoplásico
Referencia	Arthur y col., 2019; ³² Zin y col., 2020 ³³	Yang y col., 2019; ³⁴ Kim y col., 2019 ³⁵

1.2.5 Métodos de detección de clústeres de genes biosintéticos en bacterias

Las herramientas bioinformáticas son utilizadas como métodos de detección de clústeres de genes biosintéticos en secuencias de nucleótidos o aminoácidos provenientes de bacterias. Su detección se basa en la identificación de dominios proteicos específicos, algoritmos de identificación de genes y la determinación de las rutas biosintéticas.^{13,21,36} Su continuo desarrollo ha hecho posible predecir los metabolitos secundarios relacionados con los clústeres de genes biosintéticos encontrados y su potencial actividad, así como realizar la comparación con otras especies bacterianas que produzcan las mismas moléculas bioactivas. También es posible detectar genes que no se expresan, pero que por medio de la biotecnología pueden ser activados.^{13,15}

Algunas de estas herramientas bioinformáticas se enlistan a continuación.

- antiSMASH: Plataforma que realiza la identificación de clústeres de genes biosintéticos en secuencias genómicas y los compara en valor porcentual con los metabolitos secundarios que han sido publicados anteriormente.²³
- Natural Product Domain Seeker (NaPDoS2): Plataforma que realiza la identificación de PKSs y NRPSs en secuencias genéticas, genómicas o metagenómicas. Utiliza una clasificación basada en filogenia para las predicciones de estos genes.³⁷
- Minimum Information about a Biosynthetic Gene cluster (MIBiG): Plataforma que permite realizar estudios comparativos de clústeres de genes biosintéticos y sus moléculas relacionadas.³⁸
- PRISM: Software que permite la identificación de clústeres de genes biosintéticos y sus posibles estructuras resultantes a partir de secuencias genéticas.³⁹
- The Integrated Microbial Genomes Atlas of Biosynthetic gene Clusters (IMG-ABC): Base de datos que contiene la información de más de 400,000 clústeres de genes biosintéticos y sus metabolitos secundarios.⁴⁰

- BAGEL4: Plataforma que permite la identificación de bacteriocinas y RiPPs a partir de secuencias genómicas y metagenómicas.⁴¹
- The Natural Products Atlas (npatlas): Base de datos que incluye productos naturales derivados de bacterias, hongos y cianobacterias.⁴²
- Non-Ribosomal Polyketide Synthase PURified (NRPPUR): Base de datos que contiene información sobre sistemas NPRS y PKS, permitiendo su identificación en genomas por medio del programa rpsBlast.⁴³
- Natural Product Activity and Species Source database (NPASS): Base de datos que proporciona la información acerca de la estructura, actividad y especies productoras de distintos productos naturales.⁴⁴
- Antibiotic Resistant Target Seeker (ARTS): Plataforma que permite la detección de metabolitos secundarios relacionados a compuestos antibióticos a partir de genomas, así como de elementos relacionados con la resistencia antimicrobiana.⁴⁵
- Biosynthetic Gene Similarity Clustering and Prospecting Engine (BiGSCAPE): Software que realiza agrupaciones de familias de clústeres de genes biosintéticos a partir de sus similitudes.⁴⁶
- InterPro: Plataforma que realiza análisis de la estructura y los dominios de las proteínas a partir de secuencias de aminoácidos, además de su clasificación.⁴⁷
- UniProt Knowledgebase (UniProtKB): Plataforma que permite realizar el análisis de estructuras proteicas, descripción y clasificación taxonómica a partir de secuencias de aminoácidos.⁴⁸
- BLASTP: Plataforma que permite realizar la identificación y comparación de secuencias de aminoácidos con las proteínas de la base de datos de NCBI.⁴⁹

La identificación de productos naturales conlleva un largo proceso de investigación que consta de una parte inicial *in silico*, por medio de herramientas bioinformáticas y otros softwares, y una final *in vitro*, cuando se contrastan los resultados en la experimentación. En los estudios *in silico*, se obtienen las secuencias de genes o el genoma de las especies de interés, que pueden obtenerse de las bases de datos genéticas públicas o de la extracción y análisis de DNA realizados previamente.

Posteriormente se inicia la detección de los clústeres de genes biosintéticos por medio de los softwares bioinformáticos especializados, se determina la predicción los productos y su función, y se compara con los presentes en otras bacterias. Los ensayos *in vitro* engloban el aislamiento, la caracterización e identificación de las propiedades químicas de la molécula de interés, para finalmente realizar la determinación de la bioactividad.^{21,36}

Algunas limitantes en la detección de clústeres de genes biosintéticos son la disponibilidad limitada de genomas bacterianos, hallándose pocos en las bases de datos públicas, además de los costos, la complejidad y la especialización requerida para la obtención y análisis de estas secuencias.^{13,50} La búsqueda de productos naturales en especies microbianas requiere de un largo trabajo interdisciplinar de especialización en microbiología, química, genómica, filogenómica y bioinformática.⁵¹

1.2.6 Análisis genómico para la detección de clústeres de genes biosintéticos

El análisis de genomas y la genómica comparativa, permiten obtener predicciones de los clústeres de genes biosintéticos presentes en una especie bacteriana y de los metabolitos secundarios que pueden ser producidos por estos genes. Actualmente, la mayoría de los estudios de detección de clústeres de genes biosintéticos se realizan en actinobacterias, por su conocido potencial biotecnológico, pero también se han realizado en especies patógenas, en especies del medio ambiente de distintas regiones, incluyendo aquellas con condiciones extremas, y en especies nuevas.⁵²⁻⁵⁹

Mientras que las técnicas para la extracción del DNA y la secuenciación para la obtención de las secuencias del genoma pueden variar, muchos de los estudios utilizan, para la detección de clústeres de genes biosintéticos y predicción de metabolitos secundarios, herramientas bioinformáticas en común, como lo son antiSMASH, BiGSCAPE, MiBIG y BAGEL.⁵²⁻⁵⁹

Los clústeres de genes biosintéticos más comúnmente detectados son los PKS y NRPS, sin embargo, las plataformas bioinformáticas permiten hallar sistemas híbridos, sideróforos, betalactonas y terpenos conforme presentan más datos. La predicción subsecuente de los metabolitos secundarios ha permitido conocer el potencial

antimicrobiano de las especies de estudio, que puede tener acción contra patógenos humanos o de plantas.⁵²⁻⁵⁹

Mientras algunos de los estudios genómicos con búsqueda de clústeres de genes biosintéticos se realizan en muestras de suelos con la finalidad de identificar especies bacterianas ambientales conocidas o describir especies nuevas,^{52,53,56-59} también suelen reportarse comparaciones con genomas de otras especies relacionadas, que auxilien a detectar tanto las relaciones filogenéticas como las diferencias en cuanto a la biosíntesis.^{54,55} Otras comparaciones se han reportado entre genomas de la misma especie, permitiendo detectar mutaciones que tienen implicaciones en la producción de metabolitos secundarios.⁵⁷ Algunos de estos estudios genómicos realizados en 2022 incluyeron a *Streptomyces* sp.,⁵² *Streptomyces spectabilis*,⁵³ *Photobacterium* sp.,⁵⁴ *Lysobacter enzymogenes*,⁵⁵ *Corallococcus soli*,⁵⁶ *Streptomyces rapamycinicus*,⁵⁷ *Acinetobacter baumannii*⁵⁸ y *Anoxybacillus flavithermus*⁵⁹ como se resume en la **Tabla 3**.

Además, al realizar análisis en búsqueda de clústeres de genes biosintéticos en bacterias de la misma especie o con alto porcentaje de semejanza en el gen 16S rRNA, sin mutaciones conocidas, se ha demostrado que las especies de estudio presentan diferencias en cuanto a los clústeres de genes biosintéticos y sus posibles metabolitos secundarios.¹⁵

Tabla 3. Estudios genómicos comparativos realizados durante 2022

Especies bacterianas estudiadas	Clústeres de genes biosintéticos	Metabolitos secundarios	Actividad	Herramientas bioinformáticas utilizadas	Referencia
<i>Streptomyces</i> sp.	NRPS, PKS, híbridos, sideróforos, péptidos	Meridamicina	Antibacteriano y antifúngico	antiSMASH	Ashraf y col., 2022 ⁵²
<i>Streptomyces spectabilis</i> , <i>Streptomyces</i> sp.	NRPS, PKS, híbridos, terpenos	Espectinabilina, estreptorubina B	Antinematodo	antiSMASH	Yi y col., 2022 ⁵³
<i>Photobacterium</i> sp.	NRPS, híbridos, betalactonas, butirolactonas, cianobactinas, ectoína, tipo-RiPP, sideróforos	Cianobactina, Lipopolisacáridos, Ectoína, Aerobactina	Antimicrobiano	antiSMASH	Lau y col., 2022 ⁵⁴
<i>Lysobacter enzymogenes</i>	NRPS, lantipeptidos, arilpolienos, tipo-RiPP	Xanthomonadin 1, lankacidina C	Antibiótico	antiSMASH	Xu y col., 2022 ⁵⁵
<i>Coralococcus soli</i>	NRPS, PKS, híbridos, terpenos	Carotenoide, Mixochelina A y B, Geosina, Rizomida A, B, C	Antibiótico	antiSMASH, BiGSCAPE, MiBIG, Pfam	Babadi y col., 2022 ⁵⁶
<i>Streptomyces rapamycinicus</i>	Híbridos	Rapamicina	Antifúngico, antitumoral, inmunosupresor	antiSMASH	Jo y col., 2022 ⁵⁷
<i>Acinetobacter baumannii</i>	NRPS, sideróforos, betalactonas, arilpolieno	Acinetobactina, acinetoferrina, fengicina, tirosina	Antibiótico y antifúngico	antiSMASH	Poalal y col., 2022 ⁵⁸
<i>Anoxybacillus flavithermus</i>	PKS, NRPS, betalactonas, RiPPs	Amilopululanasa, pululanasa, a-amilasa	Mecanismos de adaptación en ambientes hostiles. Posible uso en astrobiología.	antiSMASH, BAGEL 4-0, PHASTER	Schultz y col., 2022 ⁵⁹

1.2.7 Aplicaciones y utilidades de los productos naturales

Desde las primeras civilizaciones humanas, los microorganismos han sido utilizados para distintas actividades, tal es el caso de los procesos de fermentación y uso de levaduras en China, Egipto y Mesopotamia. Fue en el S. XIX cuando los microorganismos tuvieron aplicaciones directas en la salud humana, siendo precedentes de las primeras vacunas contra rabia y ántrax y, en el siglo XX, el inicio de la producción de antibióticos.⁶⁰

Contemporáneamente, se conocen diversas aplicaciones biotecnológicas de los microorganismos, refiriéndose específicamente a las bacterias, se puede afirmar que su utilidad es conocida en el área alimenticia, la agroindustria, la industria química, búsqueda de biocombustibles, la remediación ambiental y tecnología médica.^{12,18,60}

Dentro de las utilidades de las bacterias en la tecnología médica, se sabe que pueden participar en la producción de antibióticos, antifúngicos, antivirales, antiparasitarios, antineoplásicos, inmunomoduladores, medicamentos de uso en patologías neurológicas, cardiovasculares, Diabetes Mellitus, esclerosis múltiple, glaucoma, además de la producción de vacunas y el control biológico de enfermedades causadas por vector.^{60,61} Según la revisión actualizada en 2020 de Newman y Gordon⁶² sobre la fuente de nuevos fármacos desde 1981 a 2019, al menos el 45% de los fármacos aprobados durante ese periodo son productos naturales, derivados o con estructuras iguales a los de éstos, mientras que sólo el 25% fueron totalmente sintéticos. En cuanto a los agentes antibacterianos, excluyendo a las vacunas y biológicos, el 70.63% son originados a partir de productos naturales o derivados de éstos, mientras que sólo el 28.5% fueron completamente sintéticos. Lo cual confirma que la mayoría de los antibióticos disponibles en la actualidad son derivados de productos naturales.^{62,63}

Hoy en día, es bien conocido el problema mundial de resistencia a los antimicrobianos, el cual se ha ido agudizando en los últimos años. En 2015, la Organización Mundial de la Salud (OMS) publicó un plan de acción mundial ante la resistencia microbiana, catalogándola como una amenaza a la salud pública mundial.⁶⁴ Causas de este problema incluyen la prescripción indiscriminada de antibióticos por parte del personal

de salud, el uso no regulado de medicamentos, el incumplimiento de la terapéutica prescrita que termina en esquemas de antibioticoterapia incompletos, y el uso de antibióticos como suplementos agrícolas profilácticos por parte de la industria. Estos factores conllevan a la existencia de infecciones de difícil manejo, con complicaciones y tasas altas de mortalidad.^{63,65} Una revisión sistemática de casos de 204 países en 2019 estimó que se produjeron a nivel mundial 4.95 millones de muertes asociadas a resistencia bacteriana, de las cuales 1.27 millones de muertes estuvieron directamente relacionadas con ésta.⁶⁶

Una de las estrategias para contrarrestar este problema mundial es la regulación de la prescripción y distribución de antibióticos. No obstante, es innegable que, en la actualidad y a futuro, nuevos antibióticos que puedan ser eficaces contra las bacterias que presentan multiresistencia serán requeridos.^{62,63} Una de las estrategias publicadas por la OMS,⁶⁴ es mantener la prioridad en el desarrollo de nuevos antimicrobianos como una respuesta a largo plazo. Por lo tanto, la producción de nuevos antimicrobianos derivados de productos naturales es una estrategia fundamental ante el creciente aumento de microorganismos multiresistentes, razón por la cual el análisis y la detección de nuevas fuentes de compuestos con actividad antimicrobiana es requerida.^{61,65}

1.2.8 Principales bacterias portadoras de clústeres de genes biosintéticos

A partir del descubrimiento de la estreptomicina en 1947 derivada de *Streptomyces griseus*, la detección de productos naturales y derivados de utilidad en la biomedicina ha ido incrementando en las últimas décadas.^{14,15,67} Algunos de los compuestos aislados a partir de 2017, se muestran en la **Tabla 4**.

Históricamente los microorganismos que producen una mayor cantidad de metabolitos secundarios son las actinobacterias y, entre éstas, es el género *Streptomyces* el más representativo y del que derivan la mayoría de los medicamentos antimicrobianos aprobados por la Administración de Alimentos y Medicamentos de los Estados Unidos (FDA, por sus siglas en inglés).^{21,51} Además, se ha encontrado evidencia de que las actinobacterias son capaces de obtener nuevos genes codificadores para metabolitos

secundarios a través de transferencia horizontal.¹⁵ Sin embargo, se ha encontrado que otras actinobacterias patógenas, como el género *Nocardia*, presentan en su genoma clústeres de genes biosintéticos en una cantidad equiparable a *Streptomyces*, tal es el caso de *Nocardia brasiliensis*.⁶⁸

Tabla 4. Ejemplos de metabolitos secundarios de utilidad biomédica aislados de bacterias en los últimos 5 años

Metabolito secundario	Tipo de metabolito secundario	Actividad	Bacteria de origen	Referencia
Ácido esclérico	NRP	Antibiótico, antineoplásico	<i>Streptomyces sclerotialis</i>	Alberti y col., 2019 ⁶⁹
Acetato de etilo del extracto crudo de MUM256 (MUM256 EA)	Dipéptidos cíclicos, entre otros	Antineoplásico	<i>Streptomyces</i> sp.	Tan y col., 2019 ⁷⁰
Acetato de etilo del extracto MCCB 248	Poliololes	Antineoplásico	<i>Streptomyces artemisiae</i>	Dhaneesha y col., 2017 ⁷¹
Acetato de etilo del extracto OA293 (SEA4)	Complejo de sustancias	Antineoplásico	<i>Streptomyces</i> sp.	Dan y col., 2018 ⁷²
Actinoalolidos	Macrolido	Antiparasitario	<i>Actinoallomurus fulvus</i>	Inahashi y col., 2018 ⁷³
Antimicina A, E, F, G y H, N-acetyl-deformilantimicina,	Alcaloides de antimicina	Citotóxico en células transformadas por VPH	<i>Streptomyces</i> sp.	Zhang y col., 2017 ⁷⁴
Avermectina	Policétido (PK)	Antiparasitario	<i>Streptomyces avermitilis</i>	Deng y col., 2017 ⁷⁵
Cosmomicina	Policétido (PK)	Antineoplásico	<i>Streptomyces</i> sp.	Larson y col., 2017 ⁷⁶

Metabolito secundario	Tipo de metabolito secundario	Actividad	Bacteria de origen	Referencia
Estreptoseomicina	Macrolactona	Antibiótico	<i>Streptomyces seoulensis</i>	Zhang y col., 2018 ⁷⁷
Exopolisacárido BAEPS	Exopolisacárido	Antineoplásico	<i>Bacillus amyloliquefaciens</i>	Ibrahim y col., 2020 ⁷⁸
Extracto de la cepa ABRIINW111	Complejo de sustancias	Antineoplásico	<i>Streptomyces levis</i>	Faramarzian y col., 2018 ⁷⁹
Extracto de la cepa ABRIINW111	Complejo de sustancias	Antineoplásico	<i>Streptomyces levis</i>	Valipour y col., 2018 ⁸⁰
Extracto de la cepa HM134	Diterpenoide	Antineoplásico	<i>Micromonospora zhangzhouensis</i>	Fu y col., 2020 ⁸¹
Gacamida A	Lipopéptido	Antibiótico	<i>Pseudomonas fluorescens</i>	Jahanshah y col., 2019 ⁸²
Molécula microbiana antiinflamatoria (MAM)	Péptido	Antiinflamatorio	<i>Faecalibacterium prausnitzii</i>	Breyner y col., 2017 ⁸³
Neoantimicina F (NAT-F)	Depsipéptido	Antineoplásico	<i>Streptomyces conglobatus</i>	Liu y col., 2019 ⁸⁴
p8	Proteína	Antineoplásico	<i>Lactobacillus rhamnosus</i>	An y col., 2019 ⁸⁵
Propionato de sodio (SP)	Propionato	Antineoplásico	Microbiota	Kim y col., 2019 ⁸⁶
Siringolina	Complejo de policétidos	Antineoplásico	<i>Pseudomonas syringae</i>	Huang y col., 2018 ⁸⁷

1.3 *Nocardia brasiliensis*

Nocardia brasiliensis es una actinobacteria, perteneciente al género *Nocardia*, familia Nocardiaceae del orden Actinomycetales.⁸⁸ Fue reportada por primera vez en 1909⁸⁹ y descrita en 1959.⁹⁰ Es una bacteria aerobia, Gram positiva y ácido-alcohol resistente que suele encontrarse en el suelo, agua y materia orgánica en descomposición. Muestra un patrón de crecimiento común entre las actinobacterias, caracterizado por formación de micelios ramificados que se fragmentan en cocos y bacilos conforme se van alargando.^{91,92}

Las colonias son acuminadas, elevadas y rugosas, de tonalidades diversas de amarillo, naranja y café. Además, forman micelio aéreo, que al estar presente tiene una apariencia blanquecina sobre la superficie de la colonia. Metaboliza caseína y tirosina, produce ácidos a partir de inositol, manitol, glucosa y glicerol, y utiliza acetato, citrato, malato, propionato, piruvato y succinato como fuentes de carbono. Es catalasa positiva y contiene ácido micólico en la pared celular.^{91,92}

Existen pocas secuenciaciones genómicas de *N. brasiliensis*. Las secuencias más recientes fueron publicadas en 2020, reportando que se trata de un DNA circular largo, como el resto de las actinobacterias. El genoma mide aproximadamente 8 Mb con alto porcentaje de las bases nitrogenadas guanina y citocina (G y C), conformando alrededor del 68.2% del total. Presenta 7,304 genes, 7,240 regiones codificantes y 117 pseudogenes.⁹³

N. brasiliensis representa una especie relevante dentro del género *Nocardia*, principalmente a causa de su potencial patogénico, siendo causante de enfermedades como actinomicetoma y nocardiosis.⁹⁴ Sin embargo, su importancia no radica únicamente en las patologías, sino que en los últimos años se ha estudiado su potencial como especie productora de metabolitos secundarios que puedan representar un beneficio a la humanidad.⁶⁸

1.3.1 Métodos de identificación de *Nocardia brasiliensis*

El estándar de oro para la identificación molecular de *N. brasiliensis* es el gen 16S rRNA, utilizado en bacterias desde 1993 hasta la actualidad.⁹⁵ Sin embargo, se ha encontrado que el gen ribosomal contiene alta semejanza entre distintas especies del género *Nocardia*, además de múltiples copias dentro del genoma lo cual hace que el uso de este gen sea insuficiente para realizar una identificación certera.^{96,97}

Como alternativa, se ha realizado la identificación mediante Tipificación de Secuencias Multilocus (MLST), analizando genes constitutivos o housekeeping, los cuales se expresan persistentemente en el DNA del microorganismo y se hallan altamente conservados entre la especie bacteriana; a partir de estos genes se han realizado identificaciones con mayor certeza.^{98,99} Algunos de los genes utilizados con este propósito han sido la superóxido dismutasa A (*sodA*), la proteína de choque térmico 65 (*hsp65*), la fracción beta de la RNA polimerasa (*rpoB*) y la subunidad beta de la DNA girasa (*gyrB*), obteniendo en todos los estudios porcentajes de identidad mayores en comparación al gen 16S rRNA y una alta capacidad de discriminación entre las especies del género *Nocardia*.^{100–106}

1.3.2 Variabilidad fenotípica y genotípica intraespecie

Las descripciones morfológicas de la especie han mostrado variabilidad en cuanto a las características macroscópicas, como el patrón de crecimiento, el color de las colonias y la pigmentación difusible en el medio de cultivo.^{107,108} Análisis en cuanto al perfil de susceptibilidad de antibióticos en cepas de origen clínico, han mostrado variaciones en la susceptibilidad de las cepas a moxifloxacino, minociclina, ceftriaxona, doxiciclina e imipenem.^{105,109,110}

Además, se ha detectado variaciones genéticas mediante la detección de polimorfismos y haplotipos en los genes 16S rRNA,¹¹¹ *gyrB*,¹⁰⁰ *hsp65*¹⁰⁵ y *sodA*¹⁰². Lo cual sugiere que es posible encontrar diversidad dentro de la propia especie.

1.3.3 Potencial patogénico

1.3.3.1 Actinomicetoma

Es un síndrome anatómico-clínico que ha sido catalogado por la OMS como una enfermedad tropical desatendida (ETD),¹¹² ocasionada por actinobacterias, siendo *N. brasiliensis* el agente causal más común en América Latina y el Caribe.⁹⁴ La vía de acceso más común es la piel, a través de heridas expuestas.¹¹³

Se caracteriza por la presencia de lesiones nodulares y granulomatosas subcutáneas indoloras, de crecimiento progresivo, con presencia de exudado serohemático desde donde se aísla al patógeno en forma de granos. Las regiones corporales comúnmente afectadas son miembro inferior, miembro superior y tronco.¹¹² Puede presentar extensión a órganos distantes y causa deformidad e incapacidad según las zonas afectadas.¹¹⁴

El actinomicetoma se presenta en las regiones tropicales, siendo México, Brasil, Venezuela, Sudán, Nigeria y la India los países que presentan más casos. Es una enfermedad que afecta a las zonas marginadas con escaso acceso a los servicios básicos de salud y vivienda, además de que no existe ningún programa de control a nivel mundial, con excepción de Sudán.^{94,112} Todas estas características, hacen del actinomicetoma una enfermedad poco estudiada y atendida.

El tratamiento se realiza con largos esquemas de antibióticos que incluyen cuatro a ocho ciclos de trimetoprim-sulfametoxazol 8/40 mg/kg/12h, amoxicilina con ácido clavulánico 500/125 mg cada 8 horas, amikacina 15mg/kg/24h e imipenem 1500mg vía IV durante 3 a 5 semanas.^{113,115} En ocasiones, es necesaria la realización de amputaciones según la extensión de la región afectada.¹¹⁶

1.3.3.2 Nocardiosis

Se trata de una enfermedad infecciosa ocasionada por distintas especies de *Nocardia*, siendo *N. brasiliensis* la más frecuentemente aislada. La vía de entrada del patógeno, que puede ser a través de la piel o por inhalación, determina las manifestaciones clínicas, pudiendo presentarse como infección cutánea, pulmonar o sistémica.

Ocasiona lesiones nodulares, inflamatorias, eritematosas, que pueden ser dolorosas y presentar exudado seroso, según el sitio de presentación. Se acompaña de crecimiento linfático doloroso. Pueden presentarse síntomas pulmonares o de sistema nervioso central, si se ubican en esas estructuras. Las regiones corporales donde es más común encontrar la enfermedad es el pulmón, piel y cerebro.^{91,92,117}

La nocardiosis se relaciona frecuentemente con pacientes inmunocomprometidos, aunque también pueden hallarse casos en individuos aparentemente sanos. No existe un grupo etario más afectado, así como tampoco se ha registrado una región geográfica de predominio. Es más común en hombres que en mujeres, en una proporción de 2:1.^{91,92}

El tratamiento para la enfermedad dependerá del agente causal, el cual puede aislarse de las lesiones, además del antibiograma individualizado. Comúnmente, se prescribe trimetoprim-sulfametoxazol 160/800 mg con doble toma cada 8 horas por 2 a 6 meses. Amoxicilina con ácido clavulánico 875/125 mg cada 12 horas y ciprofloxacino 500 mg cada 12 horas pueden acompañar el esquema antibiótico. El tratamiento quirúrgico puede ser requerido según la extensión de las lesiones.^{92,117}

1.3.4 Potencial biotecnológico

A pesar de tratarse de una especie patógena, se han aislado diversos productos naturales a partir de cepas de *N. brasiliensis*.⁴ Durante 1990 y 2005, se reportaron análisis químicos a partir de la fermentación del caldo de cultivo de las cepas, logrando aislar moléculas como la brasilcardina A, brasiliquinona A, B y C, brasilinoida B, brasilidina A y nocardiolactona. Se reportó diversidad en cuanto a la actividad biológica de estas moléculas, incluyendo antibacterianos, antifúngicos, antineoplásicos e inmunomoduladores⁴ (**Tabla 5**).

Con el advenimiento de la biología molecular y, más tarde, la secuenciación genómica, sería esperado encontrar estudios nuevos que retomaran a *N. brasiliensis* como productora de metabolitos secundarios de importancia, así como identificar y describir a los genes codificantes para las moléculas anteriormente descritas, sin embargo, son

pocos los estudios que se dedican a este patógeno, a pesar de la evidencia de su potencial biotecnológico.

Los estudios en los cuales se han realizado análisis genómicos y búsqueda selectiva de clústeres de genes biosintéticos han sido realizados tanto en cepas de origen clínico como en las secuencias genómicas disponibles en el sitio Genomes de NBCI, derivando de ello que se cuente en la actualidad con nueve secuencias genómicas accesibles.^{68,93,118}

Al realizar los análisis genómicos en búsqueda de clústeres de genes biosintéticos, las herramientas bioinformáticas más comúnmente utilizadas han sido antiSMASH, InterPro, MIBiG, UniProtKB y BLASTP, mientras que los clústeres de genes biosintéticos buscados han sido PKS I, PKS II, NRPS y betalactona. A partir de estos análisis, ha sido posible detectar en *N. brasiliensis* una cantidad de clústeres de genes biosintéticos similar a *Streptomyces* sp., aunque sin determinación específica de predicción de productos; también se han caracterizado los clústeres de genes responsables de la codificación de brasiliquinona B y nocardiolactona (**Tabla 6**). También se describió que *N. brasiliensis* presenta variabilidad en cuanto a los clústeres de genes encontrados entre las tres distintas cepas estudiadas, lo cual indicaría que estudios en muestras más grandes serían requeridas.^{68,93,118}

Realizar análisis nuevos de genómica comparativa con búsqueda de sistemas biosintéticos en *N. brasiliensis*, ampliando los parámetros a una mayor diversidad de clústeres e identificando los metabolitos secundarios relacionados, generaría información actual acerca del potencial de producción de moléculas de interés en la biomedicina y serviría como referente para futuros estudios.

Tabla 5. Moléculas aisladas del caldo de cultivo de *N. brasiliensis* durante 1990 a 2005

Molécula	Actividad biológica	Cepas de <i>N. brasiliensis</i> analizadas	Referencia
SF2457	Antibacteriano	SF2457	Itoh y col., 1992 ¹¹⁹
SO-75R1	Antibacteriano	IFM 075	Ishibashi y col., 1992 ¹²⁰
PC-766B	Antibacteriano	SC-4710, aislada de suelo	Kiyoshi y col., 1993 ¹²¹
M-3, M-4, M-13-1	Antibacteriano	IFM 0075, IFM 0075-13-1 (mutación)	Maeda y col., 1994 ¹²²
Brasilinólida A	Antibacteriano, inmunomodulador	IFM 0406	Tanaka y col., 1997 ¹²³
Brasiliquinonas A-C	Antibacteriano	IFM 0089	Nemoto y col., 1997 ¹²⁴
Brasilidina A	Antineoplásico	IFM 0089	Kobayashi y col., 1997 ¹²⁵
Notramicina	Antibacteriano	MJ896-43F17	Momose y col., 1998 ¹²⁶
Nocarinas A-C, brasiliquinona D	Antineoplásico, antibacteriano	IFM 0667	Tsuda y col., 1999 ¹²⁷
Brasilicardina A	Inmunosupresor, antineoplásico	IFM 0406	Komaki y col., 1999 ¹²⁸
Nocardiolactona	Antibacteriano	IFM 0406	Mikami y col., 1999 ¹²⁹
Eritromicina	Antibacteriano	IFM 0466	Mikami y col., 1999 ¹³⁰
Brasilinólida B	Antifúngico	IFM 0466	Mikami y col., 2000 ¹³¹
Brasilicardinas B-D	Inmunosupresor	IFM 0406	Komatsu y col., 2004 ¹³²
Brasilibactina A	Antineoplásico	IFM 0995	Tsuda y col., 2005 ¹³³

Tabla 6. Clústeres de genes biosintéticos detectados en secuencias genómicas de cepas de *Nocardia brasiliensis*

Cepas analizadas	Clústeres de genes biosintéticos detectados	País	Referencia
<i>N. brasiliensis</i> NBRC 14402T, <i>N. brasiliensis</i> IFM 10847 y <i>N. brasiliensis</i> HUJEG-1.	PKS I, NRPS e híbridos (NRPS/PKS)	Japón	Komaki y col. 2014 ⁶⁸
<i>N. brasiliensis</i> AUSMDU00024985, <i>N. brasiliensis</i> ATCC 700358, <i>N. brasiliensis</i> AUSMDU00024985, <i>N. brasiliensis</i> IFM 10847, <i>N. brasiliensis</i> NBRC 14402	PKS II (genes <i>brq</i>)	Australia y Alemania	Herisse y col. 2020 ⁹³
Cepas patógenas de <i>N. brasiliensis</i> no especificadas	β -lactona sintetasa (clúster <i>nltABCD</i>)	Estados Unidos y Australia	Robinson y col. 2020 ¹¹⁸

2. PLANTEAMIENTO DEL PROBLEMA

Nocardia brasiliensis es una actinobacteria de relevancia clínica y epidemiológica, causante de patologías como el actinomicetoma y la nocardiosis.¹¹³ El actinomicetoma se declaró una enfermedad tropical desatendida (ETD) por la OMS en 2016.¹¹² Predomina en las zonas rurales de las regiones tropicales, donde el acceso a los servicios de vivienda y salud es escaso. *N. brasiliensis* es el agente causal más común del actinomicetoma en México y en el resto de América Latina.^{94,134}

N. brasiliensis fue descubierta en 1909 y, desde entonces, se ha reconocido su importancia como patógeno, sin embargo, a pesar de que las enfermedades que produce tienen un impacto importante en las complicaciones y secuelas de los pacientes, son pocos los recursos que se destinan para su estudio.^{89,90} Actualmente, sólo se cuentan con 22 secuencias genómicas de acceso público.¹³⁵

Mediante los análisis genómicos, se ha logrado saber que *N. brasiliensis* presenta clústeres de genes biosintéticos (BGCs) en cantidad equiparable a *Streptomyces* spp., actinobacterias de referencia en la producción de metabolitos secundarios aprovechables.^{4,68,118} Sin embargo, solo se han reportado tres estudios de esta índole y no se han incluido cepas clínicas de México.

El limitado análisis de *N. brasiliensis* como potencial productor de metabolitos secundarios de importancia biomédica, restringe la aplicación de esta actinobacteria en la producción de fármacos como antibióticos, antiparasitarios, antifúngicos, antineoplásicos e inmunomoduladores, de utilidad para el ser humano.^{14,62,67} Por lo cual, el aprovechamiento en el área biomédica con impacto en la salud humana, de moléculas bioactivas derivadas de *N. brasiliensis* puede ser el resultado de la detección de clústeres de genes biosintéticos y la identificación de los metabolitos secundarios relacionados.⁴ Por lo tanto, surge la siguiente pregunta de investigación ¿Las cepas de *Nocardia brasiliensis* aisladas de actinomicetoma en México tienen potencial productor de moléculas de importancia biomédica detectado por medio de análisis genómico comparativo?

3. JUSTIFICACIÓN

Las actinobacterias son microorganismos ambientales, ocasionalmente patógenos, reconocidas por su potencial de producción de moléculas de importancia en la biomedicina, biorremediación, industria agrícola y alimentaria.¹⁴ Su relevancia en la biomedicina se conoció a partir de 1947, cuando se extrajo la estreptomicina a partir de una especie del género *Streptomyces*.¹³⁶ En la actualidad, la mayoría de los antimicrobianos aprobados por la Administración de Alimentos y Medicamentos de los Estados Unidos (FDA, por sus siglas en inglés) son derivados del mismo género.⁶² Mientras la mayor parte de las actinobacterias existen en el medio ambiente, algunas otras son patógenas en el ser humano, como diversas especies del género *Nocardia*.^{21,51}

Entre las especies patógenas del género *Nocardia*, sobresale *Nocardia brasiliensis*, el principal agente etiológico de enfermedades como el actinomictoma y la nocardiosis en las regiones de América Latina y el Caribe.⁹⁴ No obstante, aunque se conoce el papel patogénico de *N. brasiliensis*, pocos estudios se han enfocado a la búsqueda de su potencial biotecnológico.

Detectar e identificar moléculas de importancia biomédica a partir de genomas bacterianos es fundamental para el desarrollo de fármacos novedosos, siendo los microorganismos la fuente de al menos el 45% de los medicamentos aprobados entre 1981 y 2019, entre los que se incluyen antibióticos, antifúngicos, antivirales, antiparasitarios, antineoplásicos, inmunomoduladores, medicamentos de uso en patologías neurológicas y cardiovasculares.^{62,63} La producción de estos medicamentos nuevos es, además, un punto clave para la respuesta ante problemas de salud globales como lo son la resistencia antimicrobiana, la cual es causante de infecciones de difícil manejo, que suelen presentar complicaciones y llevar a la muerte.⁶⁴ En 2019, se estimó que la resistencia antimicrobiana fue causante de 4.95 millones de muertes a nivel mundial.⁶⁶

Realizar estudios enfocados en identificar el potencial de producción de metabolitos secundarios de importancia biomédica, por medio de la detección de clústeres de genes biosintéticos, en especies poco estudiadas como *N. brasiliensis*, puede aportar conocimientos nuevos sobre esta actinobacteria que sean aprovechables en el área médica e impacten en la biomedicina y la salud humana.

La trascendencia e importancia de detectar clústeres de genes biosintéticos en cepas de *N. brasiliensis* e identificar los metabolitos secundarios relacionados que tengan una potencial aplicación biomédica, por medio de un análisis genómico comparativo, permitirá que el potencial biotecnológico de esta actinobacteria poco estudiada sea aprovechable para la salud humana.

4. HIPÓTESIS

Hipótesis de investigación

Las cepas de *Nocardia brasiliensis* aisladas de actinomicetoma en México tienen potencial productor de moléculas de importancia biomédica.

Hipótesis nula

Las cepas de *Nocardia brasiliensis* aisladas de actinomicetoma en México no tienen potencial productor de moléculas de importancia biomédica.

5. OBJETIVOS

Objetivo general

Detectar el potencial productor de moléculas de importancia biomédica de *Nocardia brasiliensis* aislada de actinomicetoma en México por medio de análisis genómico comparativo.

Objetivos específicos

1. Reactivar cepas clínicas de *N. brasiliensis* para obtener DNA y los productos de la amplificación de los genes *hsp65* y 16S rRNA.
2. Analizar mediante programas y plataformas bioinformáticas las secuencias de los genes *hsp65*-16S rRNA para asignar la identidad genética de las cepas.
3. Detectar los polimorfismos genéticos e inferir la estructura poblacional de *N. brasiliensis* mediante el análisis bioinformático de los genes *hsp65* y 16S rRNA.
4. Obtener las secuencias genómicas de cuatro cepas de *N. brasiliensis* mediante la extracción y secuenciación de DNA de alto peso molecular.
5. Comparar los genomas de *N. brasiliensis* mediante los análisis bioinformáticos para la identificación genómica, la identificación de genes ortólogos, la estimación de la composición del pangenoma y el genoma central y la inferencia filogenómica.
6. Identificar los BGCs y los metabolitos secundarios con potencial actividad biomédica relacionados a partir del análisis genómico comparativo de todas las cepas de estudio.

6. MATERIALES Y MÉTODOS

6.1 Diseño del estudio

Se trata de un estudio observacional, transversal, descriptivo, comparativo y analítico.

6.2 Universo

65 cepas de *Nocardia brasiliensis* obtenidas de casos clínicos de actinomicetoma en humanos.

6.3 Muestra

4 cepas de *Nocardia brasiliensis* obtenidas de casos clínicos de actinomicetoma en humanos.

6.4 Criterios de inclusión, exclusión y eliminación

Criterios de inclusión

- Cepas que pertenezcan a la colección de actinobacterias aisladas de casos clínicos de actinomicetoma.
- Cepas de actinobacterias que sean identificadas como *Nocardia brasiliensis* por medio de los genes *hsp65* y 16S rRNA.
- Cepas que presenten variabilidad en las características morfológicas macroscópicas.

Criterios de exclusión

- Cepas de las cuales no se cuente con material biológico para su reactivación.

Criterios de eliminación

- Cepas que se contaminen durante el procedimiento o que no puedan reactivarse adecuadamente.
- Cepas de las cuales no sea posible obtener el material genético o realizar la secuenciación genómica.

6.5 Variables

Cuadro de operacionalización de variables

Variable		Definición		Tipo de variable	Escala de medición	Análisis estadístico
		Conceptual	Operativa			
Independiente	Genoma de <i>Nocardia brasiliensis</i>	ADN total presente en la especie bacteriana <i>Nocardia brasiliensis</i>	Secuencia total de ADN que poseen las cepas aisladas de casos clínicos de actinomietoma en humanos e identificadas por medio de los genes <i>hsp65</i> y 16S rRNA como <i>Nocardia brasiliensis</i>	Cualitativa	Nominal	Distribución de frecuencias
Dependiente	Potencial productor de moléculas de importancia biomédica	Virtud de poseer la capacidad de producir moléculas con aplicación en biomedicina	Capacidad de producción de antimicrobianos, antineoplásicos e inmunomoduladores a partir de metabolitos secundarios codificados por clústeres de genes biosintéticos NRPS I, NRPS II, PKS I, PKS II, PKS III, betalactonas, sideróforos, terpenos, péptidos e híbridos.	Cualitativa	Nominal dicotómica	Distribución de frecuencias

6.6 Procedimiento

6.6.1 Cepas de estudio

Se trata de una colección de actinobacterias aisladas de casos clínicos de actinomicetoma en humanos durante los años 1944 hasta 1999, recolectadas por el Instituto de Diagnóstico y Referencia Epidemiológica (InDRE) desde distintos institutos de salud en América.

Las muestras de actinomicetoma fueron tomadas de lesiones en distintas regiones corporales, incluyendo miembro superior e inferior, pulmón, cabeza y cuello y zona dorsal. Posteriormente fueron cultivadas y conservadas en criotubos a -20°C hasta su reactivación. Se trabajó con las cepas provenientes de esta colección, identificadas previamente por métodos fenotípicos como *Nocardia brasiliensis*.

Adicionalmente, se utilizaron 22 secuencias genómicas de cepas de acceso público de *N. brasiliensis*, obtenidas del Centro Nacional para la Información Biotecnológica (NCBI),¹³⁵ con los números de acceso GCF_002209125.2, GCF_011801125.1, GCF_000308475.2, GCF_000710915.1, GCF_001689765.1, GCF_900454265.1, GCF_015476985.1, GCF_015478325.1, GCF_029867465.1, GCF_029867765.1, GCF_029868505.1, GCF_029868525.1, GCF_029868545.1, GCF_029868565.1, GCF_029868585.1, GCF_029869025., GCF_029869565.1, GCF_029869625.1, GCF_029869645.1, GCF_963869805.1 y GCF_040629275.1.

6.6.2 Reactivación y mantenimiento de las cepas

Para la reactivación de las cepas, se realizó la inoculación en cajas Petri de 90 mm con medio de cultivo agar dextrosa Sabouraud (BD Bioxon™ número de catálogo 210700) adicionada con papa deshidratada al 1% y en agar Bennett. Posteriormente se incubaron a 37°C durante tres semanas o hasta que mostraran crecimiento.

Composición del medio de cultivo agar dextrosa Sabouraud
modificado con papa deshidratada al 1%

Agar dextrosa Sabouraud	65 g
Papa deshidratada	10 g
Agua destilada	1000 ml

Composición del medio de cultivo agar Bennett

Extracto de levadura	1 g
Dextrosa	10 g
N-Z amina	2 g
Agar bacteriológico	15 g
Extracto de carne	1 g
Agua destilada	1000 ml

Se realizó la inspección macroscópica, visualizando la calidad, cantidad y morfología de las colonias obtenidas, y la microscópica, por medio de tinción de Gram y revisión en microscopía óptica, descartando posibles contaminaciones. Se realizó la resiembra periódica de la colección.

6.6.3 Conservación de las cepas

Posterior a la obtención de cepas puras y con un adecuado crecimiento, se siguió el protocolo de conservación en criotubos que fueron almacenados a -20°C . Las especificaciones de la conservación se describen a continuación.

1. Preparar una solución de glicerol al 40% con 30 ml de agua destilada y 20 ml de glicerol.
2. Preparar 50 ml de medio de cultivo caldo dextrosa Sabouraud.
3. Esterilizar material.
4. Colocar en criotubos estériles, 500 μl de glicerol y 500 μl de caldo dextrosa Sabouraud. Mezclar.
5. Con ayuda de un asa estéril, obtener 2 a 3 colonias bacterianas a partir de las cajas Petri y depositarlas en los criotubos. Rotular cada tubo correctamente, indicando clave de cepa, número de copia y fecha de conservación.
6. Almacenar a -20°C .

6.6.4 Determinación de la curva de crecimiento de *N. brasiliensis*

Una vez que se obtuvieron colonias con buen crecimiento y sin contaminación a partir del medio de cultivo sólido, se inocularon las cepas 4037 y 4113, elegidas aleatoriamente a partir de la colección de estudio, en medio de cultivo líquido dextrosa

Sabouraud (BD Bioxon™, número de catálogo 222400) en matraz Erlenmeyer de 50 ml, permaneciendo en incubación a 37°C durante 15 a 20 días. Posteriormente, se verificó su crecimiento y pureza, y se utilizó como inóculo inicial para la determinación de la curva de crecimiento bacteriana.

La medición de la curva de crecimiento se realizó en placas de cultivo celular de 12 pocillos, con 3 ml de caldo Dextrosa Sabouraud sin inóculo como blanco por duplicado y 200 µl de cultivo bacteriano de cada cepa por quintuplicado, como se muestra en la **Figura 3**.

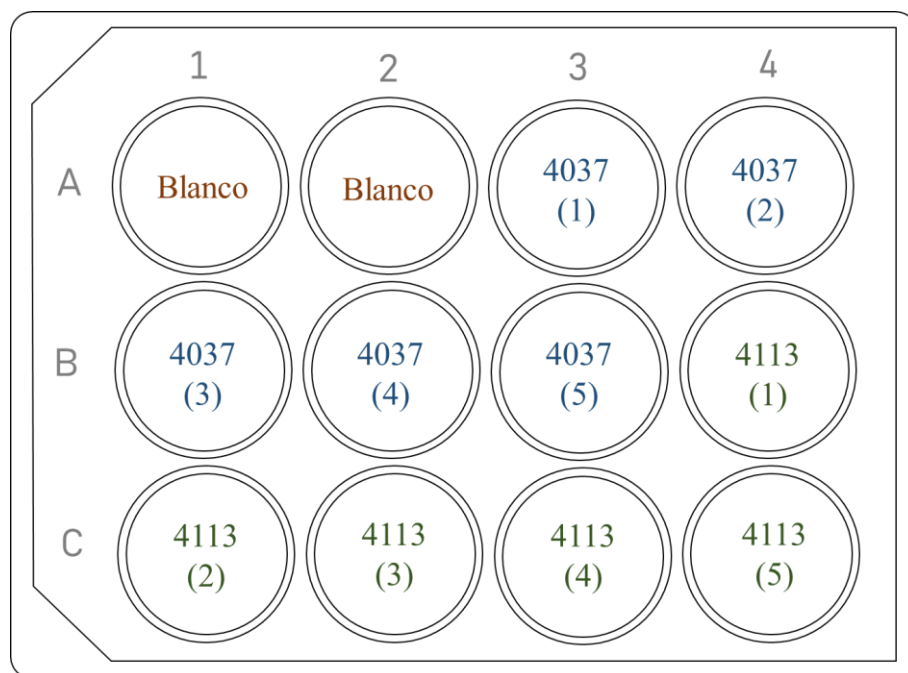


Figura 3. Medición de curva de crecimiento de *N. brasiliensis* en placas de 12 pocillos

Se realizó la medición del crecimiento por medio de espectrofotometría UV (EPOCH, Biotech), previa creación de protocolo de lectura de absorbancia para *N. brasiliensis* con los siguientes parámetros: placa de 12 pocillos, longitud de onda de 600 nm y escaneo en área con tamaño de matriz de 5 x 5. Las lecturas se realizaron cada 24 horas durante 23 días.

6.6.5 Antibiograma de cepas de *N. brasiliensis* por difusión en disco Kirby-Bauer

Se realizó la estandarización del protocolo de antibiograma por difusión en disco Kirby-Bauer para cepas de lento crecimiento de *N. brasiliensis*. El procedimiento, modificado de Kosova-Maali, et.al,¹⁰⁵ se describe a continuación.

1. Realizar la preparación de un inóculo bacteriano con patrón de turbidez 0.5 McFarland. Para esto, se recolecta una colonia fresca a partir de medio de cultivo sólido y se transfiere con un palillo estéril a un tubo de cultivo de vidrio con 6 ml de solución salina 0.85% y 7 perlas de vidrio estériles. Posteriormente, se separan las células con agitación en vórtex de 3 a 5 minutos.
2. En una caja Petri de 90 mm con agar Mueller-Hinton de 4 mm de espesor a temperatura ambiente, se realiza la siembra del inóculo bacteriano por confluencia en tres direcciones con ayuda de un hisopo estéril.
3. Se coloca un disco de antibiótico en el centro de la caja, sin dejar pasar más de 10 minutos a partir de la siembra.
4. Se incuban las cajas a 37°C durante 72 horas.
5. Se realiza la lectura de la presencia o ausencia del halo de inhibición, cuantificando en milímetros el diámetro del halo cuando está presente.

6.6.6 Extracción de DNA para amplificación de genes housekeeping *hsp65* y *sodA*

A partir de las cepas activas y purificadas, en fase de crecimiento exponencial, se obtuvo DNA por medio del método de choque térmico (modificado de Ribeiro Junior J. C, et.al¹³⁷). Para realizarlo, se siguió el protocolo que se describe a continuación.

1. Colocar 50 µl de buffer TE en un tubo de PCR de 0.2 ml.
2. Tomar una a dos colonias bacteriana fresca con la ayuda de un palillo de madera estéril. Resuspender en el buffer TE, rompiendo la colonia hasta homogeneizar el contenido.
3. Incubar a 96°C durante 10 minutos dentro del termociclador Axygen® MaxyGene II Thermal Cyclers.

4. Incubar las muestras a -20°C durante 10 minutos.
5. Centrifugar a 10,000 rpm durante 10 minutos, hasta precipitar el sedimento.
6. Recolectar el DNA contenido en 25 µl del sobrenadante y transferir a un tubo Eppendorf estéril de 0.6 ml.
7. Utilizar inmediatamente el DNA o almacenar a -20°C.

El DNA obtenido se utilizó para la amplificación de genes por medio de reacción en cadena de la polimerasa (PCR).

6.6.7 Amplificación del gen 16S rRNA por medio de PCR

La amplificación del gen *sodA* se realizó en un termociclador Axygen® MaxyGene II Thermal Cycler, con el siguiente protocolo.

Primer F 27F 5'-AGAGTTTGATCMTGGCTCAG-3'

Primer R 1492R 5'-TACGGYTACCTTGTTACGACTT-3'

Cálculo para las reacciones de PCR para amplificación del gen 16S rRNA.

- | | |
|---------|---|
| 16.3 µl | Agua libre de nucleasas |
| 2.5 µl | Buffer PCR Bioline, Meridian (5X) |
| 2.5 µl | Primer 27F (10 µM) |
| 2.5 µl | Primer 1492R (10 µM) |
| 0.20 µl | MyTaq™ DNA Polymerase, Meridian, Bioline, BIO 21105, Meridian (5U/µl) |
| 1 µl | DNA |

Preparación de mezcla maestra

1. Añadir 1 µl del DNA extraído en un tubo para PCR de 0.2 ml y reservar.
2. En un tubo Eppendorf de 1.5 ml, añadir el agua libre de nucleasas.
3. Descongelar y homogeneizar el buffer de PCR previamente a ser añadido.
4. Descongelar y homogeneizar cada uno de los primers previamente a ser añadidos.
5. Añadir Taq™ DNA polimerasa.
6. Homogeneizar la mezcla en vórtex por 10 segundos.
7. Adicionar 24 µl de la mezcla maestra a cada tubo con el DNA.

8. Colocar en el vórtex los tubos con DNA y mezcla de PCR durante 10 segundos.
9. Centrifugar los tubos y precipitar el contenido previo a ser colocados dentro del termociclador.

Protocolo de amplificación de PCR

1. Ingresar los tubos con DNA y mezcla de PCR al termociclador Axygen® MaxyGene II Thermal Cycller
2. Seleccionar el programa correspondiente a 16S rRNA, integrado por un ciclo de desnaturalización inicial a 94°C durante 5 minutos; 30 ciclos desnaturalización a 94°C durante 1 minuto, alineación a 59°C durante 30 segundos y elongación a 72°C durante 1 minuto; con una elongación final a 72°C durante 10 minutos. El programa se ilustra a continuación en la **Figura 4**.

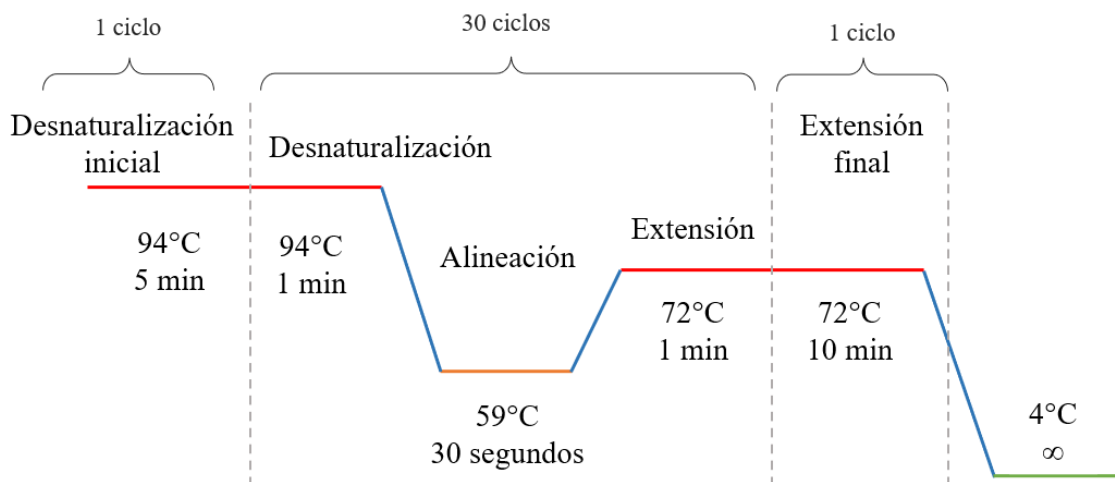


Figura 4. Condiciones del termociclador para PCR del gen 16S rRNA

Los productos de amplificación fueron observados en una electroforesis en gel de agarosa al 1% (BioBasic, número de catálogo D0012) con 3µl de bromuro de etidio (1mg/ml) y marcador de peso molecular de 1 kb (Thermo Scientific™, número de catálogo SM1163). Para la corrida se utilizó buffer TAE 1X (Invitrogen, número de catálogo 24710-030) a 120V durante 30 minutos en una caja de gel horizontal (Axygen™, número de catálogo HGB-15). El gel se transfirió al fotodocumentador y la imagen fue capturada y digitalizada. El amplicón se identificó como una banda de 1500

pb en la imagen resultante. Posteriormente, los fragmentos amplificados fueron enviados al servicio de secuenciación de Macrogen en Maryland, Estados Unidos.

6.6.8 Amplificación del gen *sodA* por medio de PCR

La amplificación del gen *sodA* se realizó en un termociclador Axygen® MaxyGene II Thermal Cycler, con el siguiente protocolo.

Primer F	SODV1	5'-CACCA ^Y W ^S CAAGCACCA-3' (Donde Y = C or T, W = A or T, and S = C or G)
Primer R	SODV2	5'-CCTTAGCGTTCTGGTACTG-3' (Donde Y = C or T, W = A or T, and S = C or G)

Cálculo para las reacciones de PCR para amplificación del gen *sodA*

8.5 µl	Agua libre de nucleasas
3.75 µl	Buffer PCR Bioline, Meridian (5X)
5.0 µl	Primer SODV1 (10 µM)
5.0 µl	Primer SODV2 (10 µM)
1.25 µl	Dimetilsulfóxido (DMSO) al 5%
0.50 µl	MyTaq™ DNA Polymerase, Meridian, Bioline, BIO 21105, Meridian (5U/µl)
1 µl	DNA

Preparación de mezcla maestra

1. Añadir 1 µl del DNA extraído en un tubo para PCR de 0.2 ml y reservar.
2. En un tubo Eppendorf de 1.5 ml, añadir el agua libre de nucleasas.
3. Descongelar y homogeneizar el buffer de PCR previamente a ser añadido.
4. Descongelar y homogeneizar cada uno de los primers previamente a ser añadidos.
5. Añadir Taq™ DNA polimerasa.
6. Homogeneizar la mezcla en vórtex por 10 segundos.
7. Adicionar 24 µl de la mezcla maestra a cada tubo con el DNA.
8. Colocar en el vórtex los tubos con DNA y mezcla de PCR durante 10 segundos.
9. Centrifugar los tubos y precipitar el contenido previo a ser colocados dentro del termociclador.

Protocolo de amplificación de PCR

1. Ingresar los tubos con DNA y mezcla de PCR al termociclador Axygen® MaxyGene II Thermal Cycler
2. Seleccionar el programa correspondiente a *sodA*, integrado por un ciclo de desnaturalización inicial a 94°C durante 5 minutos; 35 ciclos de desnaturalización a 94°C durante 1 minuto, alineación a 55°C durante 1 minuto y elongación a 72°C durante 1 minuto; con una elongación final a 72°C durante 10 minutos. El programa se ilustra en la **Figura 5**.

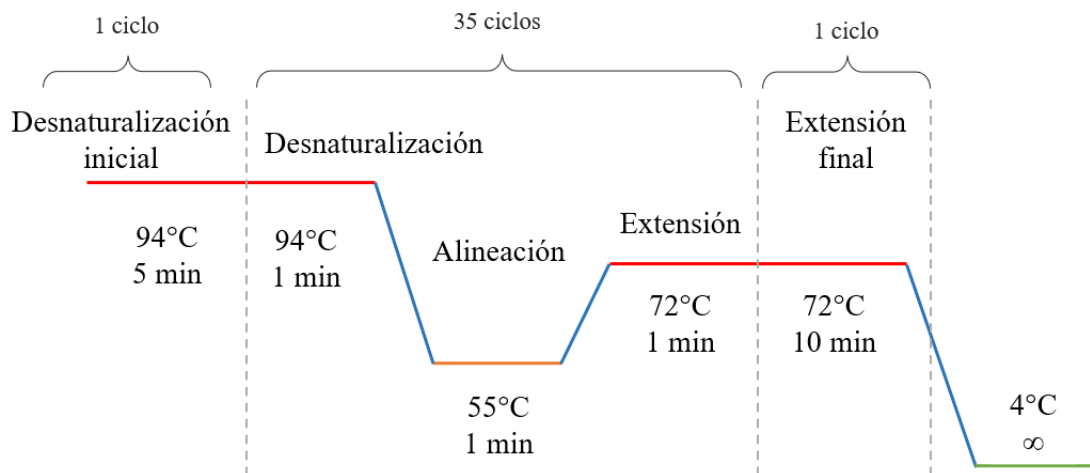


Figura 5. Condiciones del termociclador para PCR del gen *sodA*

Los productos de amplificación fueron observados en una electroforesis en gel de agarosa al 1% (BioBasic, número de catálogo D0012) con 3µl de bromuro de etidio (1mg/ml) y marcador de peso molecular de 100 pb (Thermo Scientific™, número de catálogo SM1143). Para la corrida se utilizó buffer TAE 1X (Invitrogen, número de catálogo 24710-030) a 90V durante 30 minutos en una caja de gel horizontal (Axygen™, número de catálogo HGB-15). El gel se transfirió al fotodocumentador y la imagen fue capturada y digitalizada. El amplicón se identificó como una banda de 440 pb en la imagen resultante. Posteriormente, los fragmentos amplificados fueron enviados al servicio de secuenciación de Macrogen en Seúl, Corea del Sur.

6.6.9 Amplificación del gen *hsp65* por medio de PCR

La amplificación del gen *hsp65* se realizó en un termociclador Axygen® MaxyGene II Thermal Cyclers, con el siguiente protocolo.

Primer F	TB11	5'-ACCAACGATGGTGTGTCCAT-3'
Primer R	TB12	5'-CTTGTCGAACCGCATACCCT-3'

Cálculo para las reacciones de PCR para amplificación del gen *hsp65*

16.3 µl	Agua libre de nucleasas
2.5 µl	Buffer PCR Bioline, Meridian (5X)
2.5 µl	Primer TB11 (10 µM)
2.5 µl	Primer TB12 (10 µM)
0.20 µl	MyTaq™ DNA Polymerase, Meridian, Bioline, BIO 21105, Meridian (5U/µl)
1 µl	DNA

Preparación de mezcla maestra

1. Añadir 1 µl del DNA extraído en un tubo para PCR de 0.2 ml y reservar.
2. En un tubo Eppendorf de 1.5 ml, añadir el agua libre de nucleasas.
3. Descongelar y homogeneizar el buffer de PCR previamente a ser añadido.
4. Descongelar y homogeneizar cada uno de los primers previamente a ser añadidos.
5. Añadir Taq™ DNA polimerasa.
6. Homogeneizar la mezcla en vórtex por 10 segundos.
7. Adicionar 24 µl de la mezcla maestra a cada tubo con el DNA.
8. Colocar en el vórtex los tubos con DNA y mezcla de PCR durante 10 segundos.
9. Centrifugar los tubos y precipitar el contenido previo a ser colocados dentro del termociclador.

Protocolo de amplificación de PCR

1. Ingresar los tubos con DNA y mezcla de PCR al termociclador Axygen® MaxyGene II Thermal Cyclers.

2. Seleccionar el programa correspondiente a *hsp65*, integrado por un ciclo de desnaturalización inicial a 94°C durante 5 minutos; 35 ciclos desnaturalización a 94°C durante 1 minuto, alineación a 55°C durante 1 minuto y elongación a 72°C durante 1 minuto; con una elongación final a 72°C durante 5 minutos. El programa se ilustra en la **Figura 6**.

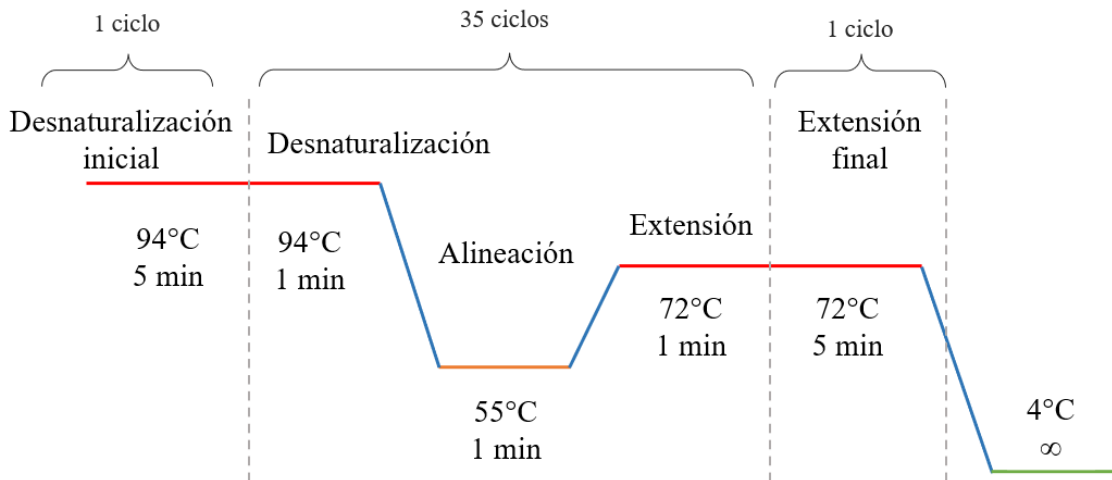


Figura 6. Condiciones del termociclador para PCR del gen *hsp65*

Los productos de amplificación fueron observados en una electroforesis en gel de agarosa al 1% (BioBasic, número de catálogo D0012) con 3 μ l de bromuro de etidio (1mg/ml) y marcador de peso molecular de 100 pb (Thermo Scientific™, número de catálogo SM1143). Para la corrida se utilizó buffer TAE 1X (Invitrogen, número de catálogo 24710-030) a 90V durante 35 minutos en una caja de gel horizontal (Axygen™, número de catálogo HGB-15). El gel se transfirió al fotodocumentador y la imagen fue capturada y digitalizada. El amplicón se identificó como una banda de 401 pb en la imagen resultante. Posteriormente, los fragmentos amplificados fueron secuenciados por medio de Sanger en el servicio de secuenciación de Macrogen en Seúl, Corea del Sur.

6.6.10 Limpieza y ensamblaje de secuencias de los genes *sodA*, *hsp65* y 16S rRNA

Las secuencias recibidas fueron analizadas por medio de programas bioinformáticos, donde se realizó la limpieza en ChromasPro versión 1.5 (Technelysium, 2003-2009; RRID:SCR_000229) y el ensamblaje en BioEdit versión 7.0.9.0,¹³⁸ obteniendo así la secuencia consenso en formato FASTA para cada uno de los genes amplificados. Posteriormente, se realizó la concatenación y almacenamiento en formato FASTA de las secuencias en orden *sodA-hsp65*-16S rRNA y *hsp65*-16S rRNA.

6.6.11 Identificación molecular de las cepas de *N. brasiliensis* por medio de los genes *sodA*, *hsp65* y 16S rRNA

Las secuencias genéticas fueron procesadas en la plataforma BLAST⁴⁹ de NCBI, donde se compararon con secuencias de referencia para la especie, arrojando datos sobre los porcentajes de similitud y de cobertura. Se realizó una identificación preliminar con el gen 16S rRNA y dos identificaciones con las secuencias concatenadas *sodA-hsp65*-16S rRNA y *hsp65*-16S rRNA.

6.6.12 Detección de variaciones genéticas y haplotipos en los genes *hsp65* y 16S rRNA

Se realizó el alineamiento múltiple de las secuencias del gen *hsp65* con la plataforma Clustal Omega,¹³⁹ se recortaron las secuencias alineadas por medio de la plataforma Gblocks 0.91.1¹⁴⁰ y se recopilaron en un archivo en formato fasta (.fas). Se definieron los ORF (marco abierto de lectura) mediante la herramienta ORF finder de NCBI,¹⁴¹ utilizando las secuencias de referencia *N. brasiliensis* NBRC 14402 y *N. brasiliensis* 4004. Se repitió el mismo procesamiento para las secuencias del gen 16S rRNA. Con el software DnaSP versión 6¹⁴² se realizaron las pruebas de aproximación estadística a la adaptación molecular con métodos basados en la distribución de polimorfismos en las secuencias de DNA, utilizando los parámetros por defecto del software. Se hizo la detección de sitios polimórficos o segregativos (S), diversidad nucleotídica (π) con corrección de Jukes y Cantor, parámetro de diversidad genética (θ), estimación de haplotipos, polimorfismos de inserción-delección no superpuestos (InDel) mediante un

modelo dialélico, divergencia de DNA entre poblaciones, regiones conservadas de DNA, eventos de recombinación y la prueba de Tajima (a partir del número de sitios segregativos) para evaluar la adaptación molecular. Se definieron dos poblaciones: Núm. 1 con las secuencias de referencia y Núm. 2 las secuencias clínicas de México.

Finalmente, se realizó un análisis de significancia estadística por coalescencia, a partir del número de sitios segregativos y valor theta por gen y por sitio, donde la hipótesis nula fue el modelo neutral estándar (MNE) y se estableció un nivel de significancia α de 0.005. Las pruebas fueron realizadas para los valores obtenidos de diversidad nucleotídica (π), parámetro de diversidad genética (θ), número de haplotipos (h), diversidad de haplotipos (H_d), eventos de recombinación (R_m) y prueba de Tajima (TD).

6.6.13 Inferencia poblacional basada en los genes *hsp65*-16S rRNA

A partir de los haplotipos para los genes *hsp65* y 16S rRNA inferidos en DnaSP, se realizó un análisis de inferencia poblacional en STRUCTURE.¹⁴³ Se incluyó información del origen de las cepas (población 1: cepas de referencia, población 2: cepas clínicas de México) y del fenotipo según su caracterización morfológica. Se utilizó el modelo Admixture, con un burnin length de 100,000 y el número de repeticiones MCMC en 1,000,000. Se predefinieron 12 clústeres o K, basándose en la información obtenida en el estudio de RFLP realizada en el gen 16S rRNA en la misma muestra y 10 repeticiones para cada una. Los resultados generados se visualizaron en StructureSelector,¹⁴⁴ donde se implementó el método de Evanno para determinar el mejor número de poblaciones obtenidas.

6.6.14 Análisis filogenético y pruebas de especiación con los genes *hsp65*-16S rRNA

Las secuencias alineadas y recortadas de los genes *hsp65*-16S rRNA se concatenaron y fueron analizadas para realizar la inferencia filogenética en el programa IQ-TREE versión 2.2.0.3.^{145,146} Se utilizó la función ModelFinder¹⁴⁷ para encontrar el mejor modelo para la estimación filogenética. El árbol filogenético de Maximum-likelihood

(ML) fue inferido a partir de 1,000 réplicas de Bootstrap. Finalmente, se utilizó FigTree versión 1.4.4¹⁴⁸ para visualizar y exportar el árbol filogenético.

Las pruebas de especiación se realizaron mediante los métodos de la implementación Bayesiana del modelo de Poisson (bPTP) y K/θ (KoT) para obtener bases estadísticas confiables y sólidas en cuanto a la diferenciación a nivel especie. La delimitación de especies se realizó con bPTP a través de la plataforma <https://github.com/zhangjjajie/PTP>,¹⁴⁹ aplicando el modelo Bayesiano de procesos de árbol de Poisson utilizando un muestreo de cadenas de Markov de Monte Carlo (MCMC) con 500, 000 generaciones, verificando posteriormente la convergencia en las cadenas. La limitación de especies se evaluó con el método de KoT,¹⁵⁰ el cual se calcula a partir de la relación K/θ, donde K es la distancia genética media entre dos clados putativos a nivel especie y θ es la diversidad genética dentro de los clados.

6.6.15 Elección de cepas representativas de la colección para secuenciación genómica

Cuatro cepas representantes de la colección fueron elegidas de acuerdo con los siguientes criterios: disponibilidad, características morfológicas, perfil de susceptibilidad a antibióticos, identificación por medio de los genes *hsp65-16S rRNA* y datos obtenidos de estudios previos en la colección de cepas donde se determinó el perfil de restricción enzimática en el gen 16S rRNA¹¹¹ y la presencia de sistemas biosintéticos PKS y NRPS.¹⁵¹

6.6.16 Obtención de biomasa de las cepas de *N. brasiliensis* para extracción de DNA genómico

Una vez elegidas las cepas representativas y que se obtuvieron colonias con buen crecimiento y sin contaminación a partir del medio de cultivo sólido, se transfirió una colonia, con ayuda de un palillo de madera estéril, a un tubo de cultivo de vidrio de 10 ml, conteniendo 5 ml de solución salina 0.85% y siete perlas de vidrio estériles. La homogeneización de las células se realizó por medio de agitación en vórtex durante 3 minutos, hasta que el inóculo obtuvo un patrón de turbidez 0.5 McFarland.

6.6.17 Extracción de DNA de alta pureza apto para secuenciación genómica

A partir de la biomasa obtenida, se realizó la extracción del DNA, siguiendo las indicaciones y especificaciones del kit de extracción MasterPure™ Complete DNA & RNA Purification Kit (LGC Biosearch Technologies, MV89010). El protocolo de extracción se describe a continuación.

1. Diluir 25 µl de proteínasa K (kit RNA and DNA Purification Kit Spin Column DaAn Gene) en 300 µl de solución de lisis de tejidos y células. Preparar una dilución para cada muestra de extracción.
2. Obtener un pellet por centrifugación. Centrifugar 500 µl de inóculo bacteriano con patrón 1 McFarland y desechar el sobrenadante, dejando aproximadamente 25 µl de líquido.
3. Resuspender el sedimento celular o pellet.
4. Añadir a cada muestra los 300 µl de solución de lisis de tejidos y células con 25 µl de proteínasa K. Mezclar por inversión.
5. Incubar a 65°C durante 15 minutos en agitación a 120 rpm.
6. Enfriar las muestras en la incubadora durante 10 minutos para que alcancen los 37°C. Posteriormente, adicionar 3 µl de RNAsa A (5 mg/ml) a cada muestra y mezclar por inversión.
7. Incubar a 37°C durante 30 minutos.
8. Colocar las muestras en una hielera con bolsas de gel refrigerante durante 5 minutos.
9. Añadir 150 µl de reactivo de precipitación de proteína MPC a la muestra (328 µl) que se encuentra lisada en ese momento. Mezclar bien por inversión durante 10 segundos.
10. Centrifugar a 4°C durante 10 minutos a 11,000 rpm. Si el sedimento resultante es transparente, pequeño o suelto, agregar 25 µl adicionales de reactivo de precipitación de proteínas MPC, mezclar y centrifugar nuevamente.
11. Transferir el sobrenadante a tubo Eppendorf de microcentrífuga de 1.5 ml limpio y estéril. Desechar sedimento.

12. Agregar 500 μ l de isopropanol frío al sobrenadante recuperado en el tubo nuevo. Mezclar por inversión de 30 a 40 veces.
13. Centrifugar a 4°C durante 10 minutos a 11,000 rpm.
14. Desechar cuidadosamente el isopropanol, teniendo cuidado de no desprender el pellet de DNA.
15. Enjuagar dos veces con 100 μ l de etanol al 70% por inversión. Centrifugar 2 minutos a 14,000 rpm para su precipitación en el tubo. Retirar todo el etanol con una pipeta y dejar secar brevemente.
16. Resuspender el DNA en 35 μ l de Buffer TE. Conservar a temperatura ambiente.

En caso de requerir eliminar completamente el RNA de la muestra, llevar a cabo el protocolo adicional que se describe a continuación.

1. Agregar 1 μ l de RNasa A a la muestra. Mezclar bien.
2. Incubar a 37°C durante 30 minutos.
3. Añadir 14 μ l de Buffer TE y 50 μ l de solución de lisis 2X T y C a cada muestra.
4. Colocar las muestras en hielera con bolsas de gel refrigerante durante 3 a 5 minutos. Agregar 100 μ l de reactivo de precipitación de proteína MPC y mezclar vigorosamente en vórtex durante 10 segundos.
5. Centrifugar los desechos a 4°C durante 10 minutos a 11,000 rpm.
6. Transferir el sobrenadante a un tubo Eppendorf de microcentrífuga limpio y estéril. Desechar el sedimento.
7. Agregar 200 μ l de isopropanol al sobrenadante recuperado. Mezclar por inversión de 30 a 40 veces.
8. Centrifugar el DNA a 4°C durante 10 minutos a 11,000 rpm.
9. Desechar cuidadosamente el isopropanol sin desalojar el sedimento de DNA.
10. Enjuagar dos veces con 100 μ l de etanol al 70%, teniendo cuidado de no desprender el pellet. Volver a centrifugar si hay desprendimiento de pellet. Retirar todo el etanol con una pipeta y dejar secar brevemente.
11. Resuspender el DNA en 35 μ l de Buffer TE.

6.6.18 Secuenciación de DNA genómico

Para llevar a cabo la secuenciación de DNA, se evaluó la concentración y pureza del DNA extraído mediante espectrofotometría UV (EPOCH, Biotech). Los parámetros óptimos que se buscaron fueron: concentración superior a 80 ng/μl, pureza medida con las relaciones de absorbancia A260-280 en un rango de 1.8-2.0 y relaciones de absorbancia A260-230, en un rango de 2.0-2.2. La integridad del DNA se analizó mediante electroforesis en gel de agarosa al 1% bajo las siguientes condiciones: 120V, 300mA y 30 min de corrida. El DNA se sometió a secuenciación híbrida, utilizando las plataformas Oxford Nanopore Technologies (ONT) de lecturas largas con la librería v14 library prep chemistry, e Illumina sequencing and array technologies en el servicio de secuenciación Plasmidsaurus (Eugene, Oregon, US).

6.6.19 Análisis de calidad, ensamblaje y anotación del genoma

Inicialmente, las lecturas obtenidas con Nanopore fueron sometidas a un análisis de calidad por medio del programa bioinformático FASTQC versión 0.12.0.¹⁵² La limpieza de las lecturas se realizó con Filtrlong versión 0.2.1, removiendo aquellas de peor calidad, utilizando los parámetros de default. Un primer ensamblaje se generó con Flye versión 2.9.1 con parámetros para lecturas de alta calidad. Se realizó un mapeo de las lecturas con Miniasm versión 0.3, utilizado para mejorar el ensamblaje con Medaka versión 1.8.0. Posteriormente, se utilizaron las lecturas de Illumina para mejorar las secuencias por medio de Polypolish versión 0.6.0, generando un nuevo ensamblaje híbrido. El contig generado fue analizado en la paquetería de QUAST versión 5.2¹⁵³ para evaluar la calidad de la secuencia obtenida. A continuación, se realizó la anotación de los genomas por medio de Bakta versión 1.6.1¹⁵⁴ y Prokka versión 1.14.6.¹⁵⁵

6.6.20 Obtención de genomas de acceso público de *Nocardia brasiliensis*

Se realizó la búsqueda y obtención de todos los genomas de referencia y de acceso público disponibles de *N. brasiliensis* en el sitio Genomes del NCBI,¹³⁵ utilizando la búsqueda guiada por organismo y por procariotas. Los genomas y sus archivos de

anotación estructural se obtuvieron en formato FASTA para su posterior análisis bioinformático.

6.6.21 Identificación genómica

Se calcularon los valores de identidad nucleótida media (ANI), identidad aminoácida media (AAI), hibridación ADN-ADN (DDH) y distancia Mash (estimación de la distancia entre dos secuencias) de los genomas con los programas RaPDTool versión 2.1.0,¹⁵⁶ Mash versión 2.3¹⁵⁷ y EzAAI versión 1.2.3.¹⁵⁸ A partir de estos valores, se realizó la identificación genómica de las cepas clínicas y las secuencias de acceso público. Se generó la matriz de ANI all-vs-all para todos los genomas y se graficó en un mapa de calor con pyani versión 0.2.13.1¹⁵⁹ y fastANI versión 1.34.¹⁶⁰

6.6.22 Análisis de calidad de los genomas

Se obtuvieron los valores de completitud y redundancia de las secuencias genómicas por medio de RaPDTool versión 2.1.0¹⁵⁶ que utiliza las dependencias de miComplete. A partir de esta puntuación se determinó el porcentaje de genes secuenciados y de secuencias repetidas.

6.6.23 Mapeo de los genomas

Se graficaron los mapas circulares de cada genoma, incluyendo la anotación genómica y la detección de los genes RAM con las herramientas Bakta y CARD, respectivamente. Se graficó un mapa de similitud genómica con la herramienta de comparación de secuencias de BLAST. Todos los mapas circulares se realizaron en la plataforma Proksee.¹⁶¹

6.6.24 Análisis del genoma central (coregenoma) y pangenoma

A partir de las secuencias de aminoácidos de cada genoma, se determinó la composición del genoma central y pangenoma por medio del programa GET_HOMOLOGUES.¹⁶² Para calcular el genoma central se utilizaron los algoritmos de BDBH, COGtriangles y OrthoMCL, para generar un consenso de cada uno de los resultados obtenidos con la herramienta compare_clusters.pl. En el cálculo del pangenoma se utilizaron los algoritmos de COGtriangles y OrthoMCL, generando

posteriormente un consenso de los resultados obtenidos con la herramienta `compare_clusters.pl`. Se definió el genoma central estricto como el presente en el 100% de los genomas, el genoma central relajado con >95% de presencia, el genoma periférico como el compartido por el 10%-90% de los genomas, y el genoma diferencial como el presente en <10% de los genomas.

La composición del genoma central y pangenoma se graficó con la herramienta `parse_pangenome_matrix.pl`. La matriz del pangenoma se graficó en un mapa de calor con la función `plot_matrix_heatmap.sh`. La curva de acumulación de genes del genoma central y pangenoma con los ajustes de Tettelin, Willenbrock y el ajuste exponencial de Tettelin se graficó con la función `plot_pancore_matrix.pl`.

6.6.25 Análisis de genes ortólogos

Las secuencias de aminoácidos de los genomas fueron analizadas con el programa OrthoFinder versión 2.5.5,¹⁶³ utilizando los algoritmos de DIAMOND, Markov Clustering (MCL) y FastME para reportar los genes ortólogos, grupos de genes ortólogos (ortogrupos) y duplicaciones genéticas.

6.6.26 Análisis filogenómico

Se infirió un árbol filogenómico de especies en OrthoFinder a partir de todos los ortogrupos en los genomas analizados, utilizando el método STAG.¹⁶⁴ Se calculó la matriz de distancia por pares para cada ortogrupo y se infirió cada árbol con el algoritmo de FastME.¹⁶⁵ Posteriormente, se calculó un consenso para generar un árbol de especies final. El árbol filogenómico generado fue visualizado en la plataforma Interactive Tree Of Life (iTOL) versión 6.¹⁶⁶

6.6.27 Detección de clústeres de genes biosintéticos

Las secuencias genómicas de *N. brasiliensis* fueron analizadas para la detección de clústeres de genes biosintéticos por medio de la herramienta bioinformática antiSMASH.²³ Los criterios de búsqueda de clústeres de genes biosintéticos fueron “strict” para la detección de clusters bien definidos y, como características adicionales, la elección de KnownClusterBlast, ClusterBlast, SubClusterBlast, MIBiG cluster

comparison, ActiveSiteFinder, RREFinder, Cluster Pfam analysis, Pfam-based GO term annotation y TIGRFam analysis. Se realizó la búsqueda de todos los clústeres de genes biosintéticos disponibles, entre los que se encontraron NRPS I, NRPS II, PKS I, PKS II, PKS III, betalactonas, sideróforos, terpenos, péptidos e híbridos.

6.6.28 Determinación de metabolitos secundarios y posible actividad biológica

Se realizó una búsqueda de la información mínima acerca de los clústeres de genes biosintéticos en la base de datos de MIBiG³⁸ para encontrar la relación entre la producción de metabolitos secundarios conocidos y nuevos, además de realizar la comparación con otras bacterias que hayan sido reportadas anteriormente como productoras de los mismos compuestos. Posteriormente, se realizó una búsqueda en PubMed para obtener la información acerca de la actividad biológica reportada previamente para cada uno de los metabolitos secundarios relacionados a los BGC detectados en los genomas de *N. brasiliensis*.

6.7 Consideraciones bioéticas

Todos los procedimientos se realizaron conforme los protocolos de manejo de microorganismos patógenos que establece la Ley General de Salud en Materia de Investigación para la Salud de 2014, siguiendo en todo momento las medidas de seguridad adecuadas. Se siguieron los lineamientos establecidos para el manejo de residuos peligrosos biológico-infecciosos, como indica la NOM-087-ECOL-SSA1-2002. El estudio siguió las metas de investigación en Ciencias de la Salud presentes en el Reglamento de la Investigación Universitaria de la Universidad Autónoma del Estado de México, priorizando en todo momento la búsqueda de soluciones a los problemas de salud por medio de métodos científicos y herramientas tecnológicas.

6.8 Análisis estadístico

Recolección de datos: Cada una de las secuencias genéticas y genómicas fueron recopiladas en los formatos especificados en los parámetros de los programas bioinformáticos utilizados, incluyendo a los formatos Fasta, GenBank y EMBL. Los

datos obtenidos a partir de los análisis en los programas bioinformáticos se recopilaron en forma de bases de datos y matrices.

Análisis de datos: Las bases de datos y matrices obtenidas se curaron manualmente en software de hojas de cálculo y editores de texto y código fuente. Los archivos fueron utilizados para la generación de las tablas y los gráficos necesarios. Finalmente, se realizaron análisis de estadística descriptiva para la presentación de los datos.


7. RESULTADOS

7.1. Artículos científicos producidos


7.1.1. Primer artículo científico (publicado)

Aceptación y publicación del artículo original de investigación titulado “**Genetic variability of the 16S rRNA gene of *Nocardia brasiliensis*, the most common causative agent of actinomycetoma in Latin America and the Caribbean**” en la Revista do Instituto de Medicina Tropical de São Paulo, indexada en JCR (InCites Journal Citation Reports), Scopus y Web of Science, con eISSN 1678-9946. DOI de la publicación: <https://doi.org/10.1590/s1678-9946202365031>
Se incluye, a continuación, la carta de envío y aceptación del artículo.

ScholarOne Manuscripts

 Revista do Instituto de Medicina Tropical de São Paulo

Submission Confirmation

 Print

Thank you for your submission

Submitted to
Revista do Instituto de Medicina Tropical de São Paulo

Manuscript ID
RIMTSP-2022-0204

Title
Genetic variability of the 16S rRNA gene of *Nocardia brasiliensis*, the most common causative agent of actinomycetoma in Latin America and the Caribbean

Authors
Cruz-Medrano, Michele Guadalupe
Manzanares-Leal , Gaudy Lizeth
González-Nava, Janette
Moreno-Pérez, Pablo Antonio
Sandoval-Trujillo , Horacio
Ramírez-Durán, Ninfa

Decision Letter (RIMTSP-2022-0204.R1)

From: revimtsp@usp.br

To: ninfard@hotmail.com

CC:

Subject: Revista do Instituto de Medicina Tropical de São Paulo - Decision on Manuscript ID RIMTSP-2022-0204.R1

Body: 24-Feb-2023

Dear Dr. Ramírez-Durán:

It is a pleasure to accept your manuscript entitled "Genetic variability of the 16S rRNA gene of *Nocardia brasiliensis*, the most common causative agent of actinomycetoma in Latin America and the Caribbean" in its current form for publication in the Revista do Instituto de Medicina Tropical de São Paulo. The comments of the reviewers who reviewed your manuscript are included at the foot of this letter.

Thank you for your fine contribution. We look forward to your continued contributions to the Journal.

Kindly,

Thelma Suely Okay, MD
Editor-in-Chief

Revista do Instituto de Medicina Tropical de São Paulo

Entire Scoresheet:
Reviewer: 1

Recommendation: Accept

Comments:
Revision completed

Additional Questions:

Does the manuscript contain new and significant information to justify publication?: Yes

Does the Abstract (Summary) clearly and accurately describe the content of the article?: Yes

Is the problem significant and concisely stated?: Yes

Are the methods described comprehensively?: Yes

Are the interpretations and conclusions justified by the results?: Yes

Is adequate reference made to other work in the field?: Yes

Is the language acceptable?: Yes

Please rate the priority for publishing this article (1 is the highest priority, 10 is the lowest priority): 8

Please state any conflict(s) of interest that you have in relation to the review of this paper (state "none" if this is not applicable).: None

Rating:

Interest: 2. Good

Quality: 3. Average

Originality: 3. Average

Overall: 3. Average

Date Sent: 24-Feb-2023

Resumen:

El micetoma fue declarado una enfermedad tropical desatendida (ETD) declarada por la Organización Mundial de la Salud (OMS) en 2016. Se caracteriza por el crecimiento progresivo de nódulos y lesiones granulomatosas en las piernas, los brazos y el tronco. Es potencialmente desfigurante y causa discapacidad o amputaciones en personas en edad productiva de zonas marginadas. Los agentes causantes pueden ser hongos (eumicetoma) o actinobacterias (actinomicetoma), siendo este último el más común en América y Asia. *Nocardia brasiliensis* es el agente causal más importante del actinomicetoma en América. Se han reportado dificultades en la asignación taxonómica al momento de identificar esta especie, por lo que el objetivo de este estudio fue detectar las variaciones del gen 16S rRNA de cepas de *N. brasiliensis* utilizando la técnica de restricción enzimática *in silico*.

El estudio incluyó cepas procedentes de casos clínicos de actinomicetoma en México, aisladas de humanos y previamente identificadas como *N. brasiliensis* mediante métodos fenotípicos tradicionales. Las cepas se caracterizaron macroscópica y microscópicamente y luego se sometieron a extracción de ADN para la amplificación del gen 16S rRNA mediante reacción en cadena de la polimerasa (PCR). Los productos de amplificación se secuenciaron y se ensamblaron las secuencias consenso que fueron utilizadas para la identificación genética y el análisis de restricción enzimática *in silico* con el programa bioinformático New England BioLabs® NEBcutter. Todas las cepas del estudio se identificaron molecularmente como *N. brasiliensis*; sin embargo, el análisis de restricción *in silico* detectó diversidad en los patrones de restricción que finalmente se agruparon y subclasificaron en 7 ribotipos. Este hallazgo confirma la existencia de subgrupos dentro de *N. brasiliensis*. Los resultados respaldan la necesidad de considerar a *N. brasiliensis* como un complejo de especies.

7.1.2. Segundo artículo científico (enviado)

Envío del artículo original de investigación titulado “**Unmasking the Genetic Landscape of Clinically Relevant *Nocardia brasiliensis* Strains in Mexico: Insights from *hsp65* and 16S rRNA Markers**” para su revisión y consideración de publicación en la revista PLOS Neglected Tropical Diseases, indexada en DOAJ, Google Scholar, PubMed, Scopus y Web of Science, con eISSN: 1935-2735. Se anexa, a continuación, la carta de envío.

PLOS Neglected Tropical Diseases: Notification of co-authorship on manuscript - [EMID:5505c6812be5eb99]

1 mensaje

PLOS Neglected Tropical Diseases <em@editorialmanager.com>
Responder a: PLOS Neglected Tropical Diseases <plosntds@plos.org>
Para: Michele Guadalupe Cruz-Medrano <cruzmed.mi@gmail.com>

4 de septiembre de 2025, 6:31 p.m.

PNTD-D-25-01588

Dear Dr. Michele Cruz-Medrano,

You are receiving this email because you have been listed as an author on a manuscript recently submitted to PLOS Neglected Tropical Diseases, which is entitled "Unmasking the Genetic Landscape of Clinically Relevant *Nocardia brasiliensis* Strains in Mexico: Insights from *hsp65* and 16S rRNA Markers".

The corresponding author for the submission process is: Dr Ayixon Sánchez Reyes

The full author list for the submission is: Michele Guadalupe Cruz-Medrano; Gaudy Lizeth Manzanares-Leal; Ayixon Sánchez Reyes; Ninfa Ramírez-Durán

The Corresponding Author has entered your name into our online submission system as indicated below. Please ensure your name is entered correctly, as this affects indexing upon publication. If amendments are required, please reply to this email or contact us at plosntds@plos.org.

First Name: Michele
Middle Name/Initial: Guadalupe
Last Name: Cruz-Medrano

If you would like to add an ORCID iD you are able to do so via the following link.
<https://www.editorialmanager.com/pntd/l.asp?i=2230052&l=XK0LPW8Y>

Please note that if you would like to link your ORCID iD to the submission, you will need to log in to your Editorial Manager account to do so. If you do not have an Editorial Manager account, you can register here: <https://www.editorialmanager.com/pntd/>.

If you are not aware of this submission, or if you should not be listed as a co-author, then please contact the journal office at plosntds@plos.org. For more information on PLOS Neglected Tropical Diseases's authorship requirements, please visit: <https://journals.plos.org/plosntds/s/authorship>.

Kind regards,
PLOS Neglected Tropical Diseases
plosntds@plos.org

In compliance with data protection regulations, you may request that we remove your personal registration details at any time. (Use the following URL: <https://www.editorialmanager.com/pntd/login.asp?a=r>). Please contact the publication office if you have any questions.

PLOS Neglected Tropical Diseases

Unmasking the Genetic Landscape of Clinically Relevant *Nocardia brasiliensis* Strains in Mexico: Insights from *hsp65* and 16S rRNA Markers

--Manuscript Draft--

Manuscript Number:	
Full Title:	Unmasking the Genetic Landscape of Clinically Relevant <i>Nocardia brasiliensis</i> Strains in Mexico: Insights from <i>hsp65</i> and 16S rRNA Markers
Short Title:	Population Genetics in Mexican Strains of <i>Nocardia brasiliensis</i>
Article Type:	Research Article
Keywords:	<i>Nocardia brasiliensis</i> ; population genetics; <i>hsp65</i> -16S rRNA molecular markers; genetic variability; population divergence
Corresponding Author:	Ayixon Sánchez Reyes, PhD UNAM: Universidad Nacional Autónoma de México Cuernavaca, México MEXICO
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	UNAM: Universidad Nacional Autónoma de México
Corresponding Author's Secondary Institution:	
First Author:	Michele Guadalupe Cruz-Medrano, M.H.Sc.
First Author Secondary Information:	
Order of Authors:	Michele Guadalupe Cruz-Medrano, M.H.Sc. Gaudy Lizeth Manzanares-Leal, PhD Ayixon Sánchez Reyes, PhD Ninfa Ramírez-Durán, PhD
Order of Authors Secondary Information:	

Powered by Editorial Manager® and Prodxion Manager® from Aries Systems Corporation

Resumen:

Nocardia brasiliensis es el agente causal más importante del actinomicetoma en las regiones de América Latina y El Caribe. Se ha reportado dificultad en la identificación a nivel especie y detección de polimorfismos en genes constitutivos, pero no se cuenta con estudios de genética de poblaciones en América Latina. El objetivo de este estudio fue conocer la diversidad genética intraespecie y la estructura poblacional de las cepas clínicas mexicanas de *N. brasiliensis* mediante el estudio de genética de poblaciones bajo el perfil de los marcadores moleculares *hsp65*-16S rRNA. Se utilizaron 60 cepas provenientes de actinomicetoma en México y se amplificaron los genes *hsp65* y 16S rRNA, utilizados como marcadores moleculares para la identificación genética. Se analizó la distribución de polimorfismos en los genes *hsp65* y 16S rRNA para la detección de variabilidad genética. Se realizó una inferencia poblacional y filogenética bajo el perfil de los marcadores *hsp65*-16S rRNA. En los resultados, se encontraron diferencias en la caracterización morfológica de las cepas analizadas. 59 cepas fueron identificadas como *N. brasiliensis* y una como *N. vulneris*. Para el gen *hsp65*, se encontró una diversidad nucleotídica intraespecie moderadamente alta ($\pi=0.04083$), se identificaron 12 haplotipos con una diversidad de haplotipos (H_d) de 0.4479 y se calculó una selección negativa por medio de la D de Tajima ($TD = -2.56064$). Entre poblaciones, se detectaron ocho eventos de polimorfismo de inserción-delección (InDel) y un índice de fijación alto ($F_{ST}=0.70302$). Para los marcadores *hsp65*-16S rRNA se infirió una estructura poblacional óptima en 4 clústeres ($K=4$), lo cual fue consistente con el análisis filogenético, donde las cepas clínicas se distribuyeron en cuatro principales clados. Este estudio demuestra la utilidad de los marcadores *hsp65*-16S rRNA para la identificación de *N. brasiliensis* y sustenta la propuesta de su consideración para la creación de un perfil MLST para la especie. Se evidenció la existencia de variabilidad genética intraespecífica en *N. brasiliensis* y la distribución de las cepas en cuatro clústeres, sugiriendo una mayor complejidad taxonómica en la especie. Se reveló diferenciación genética de las cepas mexicanas en comparación a las cepas de referencia, suponiendo la existencia de linajes diferenciados que no han sido estudiados en la región.

7.1.3. Tercer artículo científico (enviado)

Envío del artículo original de investigación titulado “**Genomic diversity and distribution of biosynthetic clusters in the pangenome of *Nocardia brasiliensis***” para su revisión y consideración de publicación en la revista Microbiological Research, indexada en Scopus, Medline, Science Citation Index Expanded (SCIE), SCImago Journal Rank (SJR) y SNIP, con elISSN: 1618-0623. Se anexa, a continuación, la carta de envío.

MICRES-D-25-04107 - Confirming your submission to Microbiological Research

1 mensaje

De: em.micres.0.971daa.63019cae@editorialmanager.com <em.micres.0.971daa.63019cae@editorialmanager.com> en nombre de Microbiological Research <em@editorialmanager.com>

Enviado: viernes, 31 de octubre de 2025 14:21

Para: Ninfa Ramírez Durán <nramirez@uaemex.mx>

Asunto: MICRES-D-25-04107 - Confirming your submission to Microbiological Research

This is an automated message.

Genomic diversity and distribution of biosynthetic clusters in the pangenome of *Nocardia brasiliensis*

Dear Dr. Ramírez-Durán,

We have received the above referenced manuscript you submitted to Microbiological Research. It has been assigned the following manuscript number: **MICRES-D-25-04107**.

To track the status of your manuscript, please log in as an author at <https://www.editorialmanager.com/micres/>, and navigate to the "Submissions Being Processed" folder.

Thank you for submitting your work to this journal.

Kind regards,
Microbiological Research

FAQ: How can I reset a forgotten password?

https://service.elsevier.com/app/answers/detail/a_id/28452/supporthub/publishing/

For further assistance, please visit our customer service site: <https://service.elsevier.com/app/home/supporthub/publishing/>

Here you can search for solutions on a range of topics, find answers to frequently asked questions, and learn more about Editorial Manager via interactive tutorials. You can also talk 24/7 to our customer support team by phone and 24/7 by live chat and email

This journal uses the Elsevier Article Transfer Service. This means that if an editor feels your manuscript is more suitable for an alternative journal, then you might be asked to consider transferring the manuscript to such a journal. The recommendation might be provided by a Journal Editor, a dedicated Scientific Managing Editor, a tool assisted recommendation, or a combination. For more details see the journal guide for authors.

At Elsevier, we want to help all our authors to stay safe when publishing. Please be aware of fraudulent messages requesting money in return for the publication of your paper. If you are publishing open access with Elsevier, bear in mind that we will never request payment before the paper has been accepted. We have prepared some guidelines (<https://www.elsevier.com/connect/authors-update/seven-top-tips-on-stopping-apc-scams>) that you may find helpful, including a short video on Identifying fake acceptance letters (<https://www.youtube.com/watch?v=o518thD9XIE>). Please remember that you can contact Elsevier's Researcher Support team (<https://service.elsevier.com/app/home/supporthub/publishing/>) at any time if you have questions about your manuscript, and you can log into Editorial Manager to check the status of your manuscript (https://service.elsevier.com/app/answers/detail/a_id/29155/c/10530/supporthub/publishing/kw/status/).

#AU_MICRES#

To ensure this email reaches the intended recipient, please do not delete the above code

In compliance with data protection regulations, you may request that we remove your personal registration details at any time. ([Remove my information/details](#)). Please contact the publication office if you have any questions.

Microbiological Research
Genomic diversity and distribution of biosynthetic clusters in the pangenome of
Nocardia brasiliensis
--Manuscript Draft--

Manuscript Number:	MICRES-D-25-04107
Article Type:	Research Paper
Section/Category:	Microbial Ecology
Keywords:	<i>Nocardia brasiliensis</i> , comparative genomics, biosynthetic gene clusters, natural products, pangenome
Corresponding Author:	Ninfa Ramírez-Durán Universidad Autonoma del Estado de Mexico Facultad de Medicina MEXICO
First Author:	Michele Guadalupe Cruz-Medrano, MsD
Order of Authors:	Michele Guadalupe Cruz-Medrano, MsD Gaudy Lizeth Manzanares-Leal, PhD Ayixon Sánchez-Reyes, PhD Ninfa Ramírez-Durán

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Resumen:

Las actinobacterias generan la mayoría de los productos naturales aprovechables en la medicina. Aquellas de origen ambiental son ampliamente estudiadas, pero poco se conoce sobre el potencial de producción de estos metabolitos secundarios por actinobacterias patógenas, como *Nocardia* spp. En este estudio se analizó el potencial de *Nocardia brasiliensis* como fuente de productos naturales mediante un análisis genómico comparativo. Se incluyeron cuatro cepas clínicas de México y 22 genomas de acceso público. Se realizó la identificación genómica, se estimó la composición del pangenoma, el genoma central y la proporción de genes ortólogos para la inferencia filogenómica. Se buscaron los BGCs conocidos en todos los genomas, se identificaron los metabolitos secundarios relacionados y los datos de la actividad biológica reportada. Se obtuvo el cálculo del pangenoma compuesto por 17,715 clústeres, donde el genoma central representó el 22.76% y el genoma diferencial el 48.17%. La tendencia de la curva de acumulación de genes indicó un pangenoma abierto para la especie. Se identificó la agrupación filogenómica de las cepas en clados de acuerdo con su origen geográfico y tipo de muestra. Se detectaron BGCs (PKS, NRPS, NAPAA, terpenos, ácido aminopolicarboxílico, híbridos y otros) que codifican para metabolitos secundarios con actividad antimicrobiana (ϵ -Poli-L-lisina, brasiliquinonas A-B), antitumoral (rizomidas A-C, antramicina), antioxidante (isorenierateno) y un fertilizante para suelos calcáreos ([S,S]-EDDS). Estos resultados resaltan la amplia diversidad genética presente en *N. brasiliensis*, reflejada en la variedad de BGCs presentes en su genoma, convirtiendo a la especie en potencial candidata para la síntesis y aprovechamiento de productos naturales.

8. DISCUSIÓN

Nocardia brasiliensis es una actinobacteria patógena que afecta a hombres y mujeres en edad productiva, residentes o trabajadores en zonas tropicales rurales, donde el acceso a los servicios de salud es escaso. Las estadísticas sobre la incidencia, prevalencia y años vida ajustados por discapacidad se encuentran subrepresentados, puesto que no se cuentan con programas de vigilancia epidemiológica en la mayoría de los países afectados, incluyendo a México.^{3,94}

En este estudio, se reactivaron 65 cepas de *N. brasiliensis* provenientes de lesiones de casos clínicos de actinomicetoma en humanos, obtenidos en México entre 1944 y 1999 y conservadas previamente a -20°C. La inspección de las cepas demostró que, si bien la morfología microscópica era consistente a filamentos fragmentados, la morfología macroscópica era variable en cuanto al patrón de crecimiento, color de las colonias, pigmentación difusible en el medio de cultivo y presencia de micelio aéreo. Estos resultados, coinciden con otros previamente reportados sobre la caracterización morfológica de la especie.^{107,108} Estos datos representan diversidad en la expresión genética, pero también podrían estar relacionados con la expresión de factores de virulencia, adaptación al huésped y diversidad genética.

Los resultados obtenidos en la identificación genética por medio del gen 16S rRNA y la concatenación de los genes *hsp65*-16S rRNA, corroboraron que el uso del gen ribosomal por sí mismo es insuficiente para diferenciar a *N. brasiliensis* de otras especies estrechamente cercanas como *N. vulneris* y *N. nova*. Sin embargo, añadir genes constitutivos puede mejorar la certeza en la identidad. Esto coincide con lo reportado previamente por Kosova-Maali y col.,¹⁰⁵ Sánchez-Herrera y col.¹⁰² y Carrasco y col. (103),¹⁰⁰ donde se mostró la utilidad de los genes *hsp65*, *sodA* y *gyrB* para *Nocardia* spp.

El análisis de polimorfismos en las secuencias del gen *hsp65* reveló que las cepas de *N. brasiliensis* divergen entre sí, mostrando 12 haplotipos con un valor de diversidad nucleotídica alto para la misma especie. El valor obtenido en la prueba de Tajima, indica eventos de selección negativa en la población analizada. Estudios anteriores,

como el de Kosova-Maali y col.,¹⁰⁵ reportó 22 haplotipos para el gen *hsp65* (401pb) y 16 para el gen 16S rRNA (569 pb) en 36 cepas clínicas originarias de Francia; mientras que Carrasco y col.¹⁰³ reportaron 10 haplotipos para el gen 16S rRNA en 10 cepas provenientes de España. La diferencia en los resultados podría estar relacionada con el tamaño y origen de muestra. Sin embargo, demuestran la variabilidad que existe dentro de los genes *hsp65* y 16S rRNA en la especie *N. brasiliensis*. Debido a esto, se sugiere la realización de estudios sobre la variabilidad genética de las cepas de *N. brasiliensis* de acuerdo con su distribución geográfica.

Además, las cepas clínicas aisladas en México divergen con las cepas de referencia de forma importante, mostrando diferenciación entre ambas poblaciones. También, se encontraron 6 eventos de inserción presentes en las cepas de México, pero no en las cepas de referencia, los cuales pueden tener implicación en las regiones codificantes y los aminoácidos resultantes en el gen. Por lo tanto, las cepas mexicanas podrían presentar estas diferencias derivado de la adaptación a su ambiente. Estos resultados son los primeros en presentarse en cepas de *N. brasiliensis*, por lo cual se sugiere ampliar el análisis de divergencia poblacional a cepas provenientes de otras regiones geográficas.

El análisis de la estructura poblacional con los genes *hsp65*-16S rRNA, mostró que las cepas se distribuyen en clados. Esto, sustenta el hecho de que las cepas de *N. brasiliensis* analizadas poseen suficientes variaciones a nivel intraespecie que permiten asignarlas en clústeres distintos. El resultado coincide con estudios anteriores, donde se identificó la distribución de cepas de *N. brasiliensis* en genotipos mediante secuencias de *sodA-hsp65*¹⁰⁵. Por lo tanto, se refuerza la propuesta de la consideración de *N. brasiliensis* como un complejo de especies.

El análisis filogenético con los genes *hsp65*-16S rRNA mostró que todas las cepas analizadas son más cercanas evolutivamente a las cepas tipo de *N. brasiliensis*, pero difieren entre sí al distribirse en clados. El primer clado es más cercano a las cepas tipo, mientras que el último clado muestra mayor divergencia entre sí, sugiriendo que pueda tratarse de subespecies o genotipos que no han sido reportados anteriormente.

La secuenciación genómica de las cuatro cepas representativas de la colección clínica mexicana, permitieron obtener genomas completos de alta calidad, lo cual no solo es útil en este estudio, sino que cobra relevancia por estar disponibles en bases de acceso público para su consulta y futuros estudios de *N. brasiliensis*.

Al obtener la identificación genómica de las cepas mexicanas y las de acceso público, se pudo distinguir que el uso de los parámetros ANI, DDH, AAI y distancia Mash en conjunto fue útil para diferenciar a *N. brasiliensis* de *N. nova*, *N. vulneris* y *Nocardia* sp. Estos resultados confirman el hecho de que la secuenciación genómica es la técnica ideal para obtener una identificación precisa, especialmente entre especies estrechamente relacionadas. Además, sugiere que los análisis genómicos son necesarios para poder realizar la detección de variaciones genéticas interespecie e intraespecie con mayor exactitud.

El análisis del pangenoma de *N. brasiliensis* constituye el primer reporte de la especie. En este estudio se encontró que la mayoría de los genes se consideran únicos, lo cual indica alta variabilidad entre las cepas. Además, la curva de acumulación de genes mostró por primera vez una tendencia hacia el crecimiento continuo, lo cual puede indicar que el pangenoma de *N. brasiliensis* es de tipo abierto.

El análisis filogenómico realizado con base en los ortogrupos encontrados generó un árbol donde se pudieron diferenciar a las cepas que fueron identificadas como *N. vulneris*, *N. nova* y *Nocardia* sp. Además, las cepas identificadas como *N. brasiliensis* se distribuyeron en cuatro principales clados, los cuales corresponden a 1) cepa de referencia y cepas tipo, 2) cepas clínicas aisladas en México, 3) cepas clínicas aisladas de Israel y 4) cepa ambiental de Estados Unidos junto a una cepa de colección genérica. Se sugiere ampliar el estudio de las relaciones evolutivas del grupo de *N. brasiliensis*, así como incluir cepas pertenecientes a otras regiones geográficas.

La detección de grupos de genes biosintéticos mostró que los genomas de *N. brasiliensis* contienen genes involucrados en la biosíntesis de moléculas con potencial función antibiótica, protectores de membrana celular, sideróforos, antioxidantes, antifúngicos, antitumorales y fertilizantes de suelos. Los tres estudios genómicos de

N. brasiliensis reportados hasta el momento describen el hallazgo de sideróforos y brasiliquinonas, mientras que este estudio detectó, además, e-Poli-L-lisina, antramicina, terpenibactina y [S,S]-EDDS.

Es de relevancia la detección del 100% de los genes involucrados en la codificación del antimicrobiano e-Poli-L-lisina; ¹⁶⁷ el ácido etilendiaminoisuccínico ([S,S]-EDDS) que participa como abono de hierro para suelos calcáreos, con la característica de ser biodegradable; ¹⁶⁸ las brasiliquinonas A y B, con actividad antibiótica contra bacterias Gram positivas; ¹²⁴ la ectoína que tiene un papel como protector de membrana celular, especialmente en la piel; ¹⁶⁹ y las rizomidas A, B y C con actividad antitumoral débil hacia las líneas celulares MGC-803 de cáncer gástrico y MCF-7 de cáncer de mama.¹⁷⁰

El genoma que presentó la mayor cantidad de BGCs fue el correspondiente a la cepa clínica N-100, aislada en México. Estos resultados demuestran que los genomas de *N. brasiliensis* contienen clústeres de genes biosintéticos con la capacidad de codificar a metabolitos secundarios que poseen actividad de importancia biomédica, por lo cual se sugiere continuar su análisis hacia la búsqueda de la expresión de estos genes y la producción de los metabolitos secundarios en pruebas *in vitro*.

9. CONCLUSIONES

A partir de los resultados obtenidos en la identificación de las 65 cepas clínicas de *N. brasiliensis*, se puede reafirmar que el uso de solo el gen universal 16S rRNA no es suficiente para lograr una identidad confiable a nivel especie, debido a que un porcentaje de similitud menor al 98% no es apto para una caracterización certera.

Por otro lado, se encontró que la secuenciación de los genes housekeeping *sodA* y *hsp65*, con alta capacidad de discriminación para *N. brasiliensis*, permitió mejorar la identificación de las 44 cepas estudiadas. También se encontró que la identificación con los genes *sodA*, *hsp65* y 16S rRNA es altamente similar a la obtenida con los genes *hsp65* y 16S rRNA, significando una alternativa viable para mejorar la identificación de las cepas. Por lo tanto, se proponen estos dos genes como parte del perfil MLST para la especie.

Los hallazgos en los haplotipos, los sitios polimórficos y la diversidad nucleotídica en el gen *hsp65*, demuestran la existencia de variabilidad genética dentro de la especie *N. brasiliensis*. Además, el valor negativo obtenido de la prueba de Tajima sugiere que la población estudiada se halla bajo selección negativa, presentando mutaciones deletéreas, y no está bajo un equilibrio neutral.

Las diferencias detectadas por medio de polimorfismos InDel, el índice de fijación entre poblaciones y la distribución filogenética, indican divergencia entre las cepas clínicas de México y las cepas tipo. Por lo tanto, se muestra evidencia de que las cepas mexicanas han pasado por cambios genéticos que las diferencian de las cepas tipo.

La inferencia poblacional y la divergencia en las relaciones evolutivas halladas en el análisis filogenético con los genes *hsp65* y 16S rRNA, sugieren la diferenciación de las cepas mexicanas en cuatro clústeres.

La identificación realizada con los genes *hsp65*-16S rRNA, las características morfológicas, los patrones de restricción en el gen 16S rRNA y la presencia de sistemas biosintéticos PKS y NRPS en el DNA de las cepas, permitieron elegir la muestra de cuatro cepas representantes para la secuenciación genómica.

De acuerdo con los parámetros evaluados en los análisis de calidad de las lecturas genómicas, se muestra que la secuenciación mediante plataformas híbridas (Nanopore/Illumina) es adecuada para las cepas de *N. brasiliensis*, resultando en genomas de buena calidad para realizar análisis bioinformáticos de genómica comparativa.

Las secuencias genómicas de *N. brasiliensis* de acceso público obtenidas muestran que, desde el año 2012 hasta la actualidad, solo se han depositado 22 genomas, enfatizando la necesidad de incrementar el número de secuenciaciones y estudios genómicos de esta especie. A su vez, solo dos genomas han sido aislados en México, lo cual evidencia que es necesario ampliar la información sobre las cepas mexicanas.

La identificación genómica con distintos parámetros (ANI, DDH, AAI y distancia Mash) permitió diferenciar al grupo de *N. brasiliensis* de dos especies estrechamente relacionadas como lo son *N. vulneris* y *N. nova*, lo cual indica que los análisis genómicos con distintos parámetros son necesarios para lograr un adecuado discernimiento a nivel especie. De igual forma, los resultados sugieren que es necesario ampliar los estudios de genómica en cepas identificadas como *N. brasiliensis* a nivel genético para detectar la diversidad de especies relacionadas, posibles especies nuevas o subespecies dentro de este grupo que, potencialmente, sería un complejo de especies.

La composición del pangenoma indica que los genomas del grupo de estudio comparten un porcentaje menor de clústeres y, en cambio, el porcentaje mayor de sus genes se pueden considerar como únicos.

La curva de acumulación de genes del pangenoma de *N. brasiliensis* muestra un crecimiento continuo conforme se van agregando los genomas analizados, lo cual sugiere que se trata de un pangenoma abierto.

El análisis filogenómico mostró que el sitio geográfico y el origen de aislamiento de las cepas influye en la relación evolutiva de la especie. Se sugiere ampliar el estudio de las relaciones evolutivas, incluyendo estas variables.

Las cepas de *N. brasiliensis* presentaron diversidad de BGCs (PKS, NRPS, NAPAA, terpenos, ácido aminopolicarboxílico, ectoína, latipéptidos e híbridos) que codifican para metabolitos secundarios con actividad antimicrobiana (ϵ -Poli-L-lisina, brasiliquinonas A-B, Planosporicina, E-837), antitumoral (rizomidas A-C, antramicina), antioxidante (isorenierateno), anticoagulante (antipaina), estabilizador celular (ectoína, hopeno), sideróforo (heterobactina A, B, S2, terpenibactina A-C) y un fertilizante para suelos calcáreos ([S,S]-EDDS). Finalmente, se demostró que las cepas de *N. brasiliensis* aisladas de actinomicetoma en México tienen potencial para producir moléculas de importancia biomédica aprovechables para el ser humano.

10. REFERENCIAS BIBLIOGRÁFICAS

1. Lin S, Capstick R, Ng S. Actinomycetoma of the hand: a case report. *ANZ J Surg.* 2020 Mar;90(3):E44–5.
2. Ramos-E-Silva M, Lopes RS, Trope BM. Cutaneous nocardiosis: A great imitator. *Clin Dermatol.* 2020;38(2):152–9.
3. Musa HH, Musa TH, Musa IH, Musa IH. Global scientific research progress in mycetoma: a bibliometric analysis. *Trans R Soc Trop Med Hyg.* 2021 Dec;115(12):1414–26.
4. Engelbrecht A, Saad H, Gross H, Kaysser L. Natural Products from *Nocardia* and Their Role in Pathogenicity. *Microb Physiol.* 2021;31(3):217–32.
5. Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J, Schloss JA, et al. DNA sequencing at 40: past, present and future. *Nature.* 2017 Oct;550(7676):345–53.
6. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science.* 1995 Jul;269(5223):496–512.
7. Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, et al. The minimal gene complement of *Mycoplasma genitalium*. *Science.* 1995 Oct;270(5235):397–403.
8. Lorenz SC, Gonzalez-Escalona N, Kotewicz ML, Fischer M, Kase JA. Genome sequencing and comparative genomics of enterohemorrhagic *Escherichia coli* O145:H25 and O145:H28 reveal distinct evolutionary paths and marked variations in traits associated with virulence & colonization. *BMC Microbiol.* 2017 Aug;17(1):183.
9. diCenzo GC, Finan TM. The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiol Mol Biol Rev.* 2017 Sep;81(3):e00019-17.
10. Patané JSL, Martins JJ, Setubal JC. Phylogenomics. *Methods Mol Biol.* 2018;1704:103–87.
11. Gonzalez-Nava, Janette Berenice; Alonso-Carmona, Scarlett; Manzanares-Leal, Gaudy Lizeth; Sandoval-Trujillo, Horacio; Ramírez-Durán Ninfa. Current Techniques For The Search For Natural Products In Actinobacteria. *Records of Natural Products.* 2022;16(4):274–92.
12. Lee JK, Kalia VC. Mapping Microbial Capacities for Bioremediation: Genes to Genomics. *Indian J Microbiol.* 2020 Mar;60(1):45–53.
13. Prihoda D, Maritz JM, Klempir O, Dzamba D, Woelk CH, Hazuda DJ, et al. The application potential of machine learning and genomics for understanding natural product diversity, chemistry, and therapeutic translatability. *Nat Prod Rep.* 2021;38(6):1100–8.

14. Abdelghani Z, Hourani N, Zaidan Z, Dbaibo G, Mrad M, Hage-Sleiman R. Therapeutic applications and biological activities of bacterial bioactive extracts. *Arch Microbiol.* 2021 Oct;203(8):4755–76.
15. Sekurova ON, Schneider O, Zotchev SB. Novel bioactive natural products from bacteria via bioprospecting, genome mining and metabolic engineering. *Microb Biotechnol.* 2019 Sep;12(5):828–44.
16. Barak D, Zongqiang W, F. BS, Edouard J, Eddie C. Elucidating the Diversity and Potential Function of Nonribosomal Peptide and Polyketide Biosynthetic Gene Clusters in the Root Microbiome. *mSystems* [Internet]. 2020 Dec 22;5(6):e00866-20. Available from: <https://doi.org/10.1128/mSystems.00866-20>
17. Borsetto C, Amos GCA, da Rocha UN, Mitchell AL, Finn RD, Laidi RF, et al. Microbial community drivers of PK/NRP gene diversity in selected global soils. *Microbiome.* 2019 May;7(1):78.
18. Moreira SM, de Oliveira Mendes TA, Santanta MF, Huws SA, Creevey CJ, Mantovani HC. Genomic and gene expression evidence of nonribosomal peptide and polyketide production among ruminal bacteria: a potential role in niche colonization? *FEMS Microbiol Ecol.* 2020 Feb;96(2):fiz198.
19. Tyc O, Song C, Dickschat JS, Vos M, Garbeva P. The Ecological Role of Volatile and Soluble Secondary Metabolites Produced by Soil Bacteria. *Trends Microbiol.* 2017 Apr;25(4):280–92.
20. Nivina A, Yuet KP, Hsu J, Khosla C. Evolution and Diversity of Assembly-Line Polyketide Synthases. *Chem Rev.* 2019 Dec;119(24):12524–47.
21. Singh M, Chaudhary S, Sareen D. Non-ribosomal peptide synthetases: Identifying the cryptic gene clusters and decoding the natural product. *J Biosci.* 2017 Mar;42(1):175–87.
22. Sabatini M, Comba S, Altabe S, Recio-Balsells AI, Labadie GR, Takano E, et al. Biochemical characterization of the minimal domains of an iterative eukaryotic polyketide synthase. *FEBS J.* 2018 Dec;285(23):4494–511.
23. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, et al. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 2021 Jul;49(W1):W29–35.
24. Gherbawy YA, El-Dawy EGAE. Conventional detection and quantification real-time PCR of the pks-1 gene of *Chaetomium globosum*. *J Basic Microbiol.* 2020 May;60(5):407–14.
25. Liu SH, Wang W, Wang KB, Zhang B, Li W, Shi J, et al. Heterologous Expression of a Cryptic Giant Type I PKS Gene Cluster Leads to the Production of Ansaseomycin. *Org Lett.* 2019 May;21(10):3785–8.
26. Robertson AW, MacLeod JM, MacIntyre LW, Forget SM, Hall SR, Bennett LG, et al. Post Polyketide Synthase Carbon-Carbon Bond Formation in Type-II PKS-

- Derived Natural Products from *Streptomyces venezuelae*. *J Org Chem*. 2018 Feb;83(4):1876–90.
27. Sun L, Zeng J, Cui P, Wang W, Yu D, Zhan J. Manipulation of two regulatory genes for efficient production of chromomycins in *Streptomyces roseiscleroticus*. *J Biol Eng*. 2018;12:9.
 28. Hug JJ, Panter F, Krug D, Müller R. Genome mining reveals uncommon alkylnones as type III PKS products from myxobacteria. *J Ind Microbiol Biotechnol*. 2019 Mar;46(3–4):319–34.
 29. Čihák M, Kameník Z, Šmídová K, Bergman N, Benada O, Kofroňová O, et al. Secondary Metabolites Produced during the Germination of *Streptomyces coelicolor*. *Front Microbiol*. 2017;8:2495.
 30. Bonhomme S, Dessen A, Macheboeuf P. The inherent flexibility of type I non-ribosomal peptide synthetase multienzymes drives their catalytic activities. *Open Biol*. 2021 May;11(5):200386.
 31. Jaremko MJ, Davis TD, Corpuz JC, Burkart MD. Type II non-ribosomal peptide synthetase proteins: structure, mechanism, and protein-protein interactions. *Nat Prod Rep*. 2020 Mar;37(3):355–79.
 32. Arthur RA, Nicholson AC, Humrighouse BW, McQuiston JR, Lasker BA. Draft Genome Sequence of *Kroppenstedtia sanguinis* X0209(T), a Clinical Isolate Recovered from Human Blood. *Microbiol Resour Announc*. 2019 Jun;8(24).
 33. Zin PPK, Williams GJ, Ekins S. Cheminformatics Analysis and Modeling with MacrolactoneDB. *Sci Rep*. 2020;10(1):6284.
 34. Yang Z, Shao L, Wang M, Rao M, Ge M, Xu Y. Two novel quinomycins discovered by UPLC-MS from *Streptomyces* sp. HCCB11876. *J Antibiot (Tokyo)*. 2019 Mar;72(3):164–8.
 35. Kim HY, Jin H, Bae J, Choi HK. Metabolic and lipidomic investigation of the antiproliferative effects of coronatine against human melanoma cells. *Sci Rep*. 2019 Feb;9(1):3140.
 36. Bauman KD, Butler KS, Moore BS, Chekan JR. Genome mining methods to discover bioactive natural products. *Nat Prod Rep*. 2021;38(11):2100–29.
 37. Klau LJ, Podell S, Creamer KE, Demko AM, Singh HW, Allen EE, et al. The Natural Product Domain Seeker version 2 (NaPDoS2) webtool relates ketosynthase phylogeny to biosynthetic function. *J Biol Chem*. 2022 Sep;102480.
 38. Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR, van der Hooft JJJ, et al. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res*. 2020 Jan 8;48(D1):D454–8.
 39. Skinnider MA, Johnston CW, Edgar RE, Dejong CA, Merwin NJ, Rees PN, et al. Genomic charting of ribosomally synthesized natural product chemical

- space facilitates targeted mining. *Proc Natl Acad Sci U S A*. 2016 Oct;113(42):E6343–51.
40. Palaniappan K, Chen IMA, Chu K, Ratner A, Seshadri R, Kyrpides NC, et al. IMG-ABC v.5.0: an update to the IMG/Atlas of Biosynthetic Gene Clusters Knowledgebase. *Nucleic Acids Res*. 2020 Jan;48(D1):D422–30.
 41. van Heel AJ, de Jong A, Song C, Viel JH, Kok J, Kuipers OP. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res*. 2018 Jul;46(W1):W278–81.
 42. van Santen JA, Poynton EF, Iskakova D, McMann E, Alsup TA, Clark TN, et al. The Natural Products Atlas 2.0: a database of microbially-derived natural products. *Nucleic Acids Res*. 2022 Jan 7;50(D1):D1317–23.
 43. Fritz S, Rajaonison A, Chabrol O, Raoult D, Rolain JM, Merhej V. Full-length title: NRPPUR database search and in vitro analysis identify an NRPS-PKS biosynthetic gene cluster with a potential antibiotic effect. *BMC Bioinformatics*. 2018 Dec;19(1):463.
 44. Zeng X, Zhang P, He W, Qin C, Chen S, Tao L, et al. NPASS: natural product activity and species source database for natural product research, discovery and tool development. *Nucleic Acids Res*. 2018 Jan;46(D1):D1217–22.
 45. Mungan MD, Alanjary M, Blin K, Weber T, Medema MH, Ziemert N. ARTS 2.0: feature updates and expansion of the Antibiotic Resistant Target Seeker for comparative genome mining. *Nucleic Acids Res*. 2020 Jul;48(W1):W546–52.
 46. Navarro-Muñoz JC, Selem-Mojica N, Mallowney MW, Kautsar SA, Tryon JH, Parkinson EI, et al. A computational framework to explore large-scale biosynthetic diversity. *Nat Chem Biol*. 2020 Jan;16(1):60–8.
 47. Blum M, Chang HY, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res*. 2021 Jan;49(D1):D344–54.
 48. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. 2021 Jan;49(D1):D480–9.
 49. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990 Oct;215(3):403–10.
 50. Alser M, Bingöl Z, Cali DS, Kim J, Ghose S, Alkan C, et al. Accelerating Genome Analysis: A Primer on an Ongoing Journey. *IEEE Micro*. 2020;40(5):65–75.
 51. Avalon NE, Murray AE, Baker BJ. Integrated Metabolomic-Genomic Workflows Accelerate Microbial Natural Product Discovery. *Anal Chem*. 2022 Sep;94(35):11959–66.
 52. Ashraf N, Zafar S, Makitrynsky R, Bechthold A, Spiteller D, Song L, et al. Revealing Genome-Based Biosynthetic Potential of *Streptomyces* sp. BR123 Isolated from Sunflower Rhizosphere with Broad Spectrum Antimicrobial Activity. *Antibiotics (Basel)*. 2022 Aug;11(8):1057.

53. Yi JS, Kim JM, Kang MK, Kim JH, Cho HS, Ban YH, et al. Whole-genome sequencing and analysis of *Streptomyces* strains producing multiple antineoplastic drugs. *BMC Genomics*. 2022 Aug;23(1):610.
54. Lau NS, Heng WL, Miswan N, Azami NA, Furusawa G. Comparative Genomic Analyses of the Genus *Photobacterium* Illuminate Biosynthetic Gene Clusters Associated with Antagonism. *Int J Mol Sci*. 2022 Aug;23(17):9712.
55. Xu S, Zhang Z, Xie X, Shi Y, Chai A, Fan T, et al. Comparative genomics provides insights into the potential biocontrol mechanism of two *Lysobacter* enzymogenes strains with distinct antagonistic activities. *Front Microbiol*. 2022;13:966986.
56. Babadi ZK, Garcia R, Ebrahimipour GH, Risdian C, Kämpfer P, Jarek M, et al. *Coralococcus soli* sp. Nov., a Soil Myxobacterium Isolated from Subtropical Climate, Chalus County, Iran, and Its Potential to Produce Secondary Metabolites. *Microorganisms*. 2022 Jun;10(7):1262.
57. Jo HG, Adidjaja JJ, Kim DK, Park BS, Lee N, Cho BK, et al. Comparative genomic analysis of *Streptomyces rapamycinicus* NRRL 5491 and its mutant overproducing rapamycin. *Sci Rep*. 2022 Jun;12(1):10302.
58. Pooalai R, Khongfak S, Leungtongkam U, Thummeepak R, Kunthalert D, Sitthisak S. Genomic analysis uncovers laccase-coding genes and biosynthetic gene clusters encoding antimicrobial compounds in laccase-producing *Acinetobacter baumannii*. *Sci Rep*. 2022 Jul;12(1):11932.
59. Schultz J, Parise MTD, Parise D, Medeiros LG, Sousa TJ, Kato RB, et al. Unraveling the Genomic Potential of the Thermophilic Bacterium *Anoxybacillus flavithermus* from an Antarctic Geothermal Environment. *Microorganisms*. 2022 Aug;10(8):1673.
60. Ostos-Ortiz OL, Rosas-Arango SM, González-Devia JL. Aplicaciones biotecnológicas de los microorganismos . Vol. 17, Nova . scieloco ; 2019. p. 129–63.
61. Melander RJ, Basak AK, Melander C. Natural products as inspiration for the development of bacterial antibiofilm agents. *Nat Prod Rep*. 2020 Nov;37(11):1454–77.
62. Newman DJ, Cragg GM. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J Nat Prod*. 2020 Mar;83(3):770–803.
63. Wencewicz TA. Crossroads of Antibiotic Resistance and Biosynthesis. *J Mol Biol*. 2019 Aug;431(18):3370–99.
64. World Health Organization. Global action plan on antimicrobial resistance. Ginebra PP - Ginebra: Organización Mundial de la Salud; 2016. 45 p.
65. Rossiter SE, Fletcher MH, Wuest WM. Natural Products as Platforms To Overcome Antibiotic Resistance. *Chem Rev*. 2017 Oct;117(19):12415–74.

66. Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet*. 2022 Feb;399(10325):629–55.
67. Pham J V, Yilma MA, Feliz A, Majid MT, Maffetone N, Walker JR, et al. A Review of the Microbial Production of Bioactive Natural Products and Biologics. *Front Microbiol*. 2019;10:1404.
68. Komaki H, Ichikawa N, Hosoyama A, Takahashi-Nakaguchi A, Matsuzawa T, Suzuki K ichiro, et al. Genome based analysis of type-I polyketide synthase and nonribosomal peptide synthetase gene clusters in seven strains of five representative *Nocardia* species. *BMC Genomics*. 2014 Apr;15(1):323.
69. Alberti F, Leng DJ, Wilkening I, Song L, Tosin M, Corre C. Triggering the expression of a silent gene cluster from genetically intractable bacteria results in scleric acid discovery. *Chem Sci*. 2019 Jan;10(2):453–63.
70. Tan LTH, Chan CK, Chan KG, Pusparajah P, Khan TM, Ser HL, et al. *Streptomyces* sp. MUM256: A Source for Apoptosis Inducing and Cell Cycle-Arresting Bioactive Compounds against Colon Cancer Cells. *Cancers (Basel)*. 2019 Nov;11(11).
71. Dhaneesha M, Benjamin Naman C, Krishnan KP, Sinha RK, Jayesh P, Joseph V, et al. *Streptomyces artemisiae* MCCB 248 isolated from Arctic fjord sediments has unique PKS and NRPS biosynthetic genes and produces potential new anticancer natural products. *3 Biotech*. 2017 May;7(1):32.
72. Dan VM, Muralikrishnan B, Sanawar R, J S V, Burkul BB, Srinivas KP, et al. *Streptomyces* sp metabolite(s) promotes Bax mediated intrinsic apoptosis and autophagy involving inhibition of mTOR pathway in cervical cancer cell lines. *Sci Rep*. 2018 Feb;8(1):2810.
73. Inahashi Y, Shiraishi T, Také A, Matsumoto A, Takahashi Y, Ōmura S, et al. Identification and heterologous expression of the actinoallolide biosynthetic gene cluster. *J Antibiot (Tokyo)*. 2018 Aug;71(8):749–52.
74. Zhang W, Che Q, Tan H, Qi X, Li J, Li D, et al. Marine *Streptomyces* sp. derived antimycin analogues suppress HeLa cells via depletion HPV E6/E7 mediated by ROS-dependent ubiquitin-proteasome system. *Sci Rep*. 2017 Feb;7:42180.
75. Deng Q, Zhou L, Luo M, Deng Z, Zhao C. Heterologous expression of Avermectins biosynthetic gene cluster by construction of a Bacterial Artificial Chromosome library of the producers. *Synth Syst Biotechnol*. 2017 Mar;2(1):59–64.
76. Larson CB, Crüseemann M, Moore BS. PCR-Independent Method of Transformation-Associated Recombination Reveals the Cosmomycin Biosynthetic Gene Cluster in an Ocean *Streptomyces*. *J Nat Prod*. 2017 Apr;80(4):1200–4.

77. Zhang B, Wang KB, Wang W, Bi SF, Mei YN, Deng XZ, et al. Discovery, Biosynthesis, and Heterologous Production of Streptoseomycin, an Anti-Microaerophilic Bacteria Macrodilactone. *Org Lett*. 2018 May;20(10):2967–71.
78. Ibrahim AY, Youness ER, Mahmoud MG, Asker MS, El-Newary SA. Acidic Exopolysaccharide Produced from Marine *Bacillus amyloliquefaciens* 3MS 2017 for the Protection and Treatment of Breast Cancer. *Breast Cancer (Auckl)*. 2020;14:1178223420902075.
79. Faramarzian Azimi Maragheh B, Fatourachi P, Mohammadi SM, Valipour B, Behtari M, Dehnad A, et al. *Streptomyces Levis* ABRIINW111 Inhibits SW480 Cells Growth by Apoptosis Induction. *Adv Pharm Bull*. 2018 Nov;8(4):675–82.
80. Valipour B, Mohammadi SM, Abedelahi A, Faramarzian Azimi Maragheh B, Naderali E, Dehnad A, et al. Culture filtrate ether extracted metabolites from *Streptomyces levis* ABRIINW111 increased apoptosis and reduced proliferation in acute lymphoblastic leukemia. *Biomed Pharmacother*. 2018 Dec;108:216–23.
81. Fu G, Wang R, Ding J, Qi H, Zhao Z, Chen C, et al. *Micromonospora zhangzhouensis* sp. nov., a Novel Actinobacterium Isolated from Mangrove Soil, Exerts a Cytotoxic Activity in vitro. *Sci Rep*. 2020 Mar;10(1):3889.
82. Jahanshah G, Yan Q, Gerhardt H, Pataj Z, Lämmerhofer M, Pianet I, et al. Discovery of the Cyclic Lipopeptide Gacamide A by Genome Mining and Repair of the Defective GacA Regulator in *Pseudomonas fluorescens* Pf0-1. *J Nat Prod*. 2019 Feb;82(2):301–8.
83. Breyner NM, Michon C, de Sousa CS, Vilas Boas PB, Chain F, Azevedo VA, et al. Microbial Anti-Inflammatory Molecule (MAM) from *Faecalibacterium prausnitzii* Shows a Protective Effect on DNBS and DSS-Induced Colitis Model in Mice through Inhibition of NF- κ B Pathway. *Front Microbiol*. 2017;8:114.
84. Liu L, Zhu H, Wu W, Shen Y, Lin X, Wu Y, et al. Neoantimycin F, a *Streptomyces*-Derived Natural Product Induces Mitochondria-Related Apoptotic Death in Human Non-Small Cell Lung Cancer Cells. *Front Pharmacol*. 2019;10:1042.
85. An BC, Ryu Y, Yoon YS, Choi O, Park HJ, Kim TY, et al. Colorectal Cancer Therapy Using a *Pediococcus pentosaceus* SL4 Drug Delivery System Secreting Lactic Acid Bacteria-Derived Protein p8. *Mol Cells*. 2019 Nov;42(11):755–62.
86. Kim K, Kwon O, Ryu TY, Jung CR, Kim J, Min JK, et al. Propionate of a microbiota metabolite induces cell apoptosis and cell cycle arrest in lung cancer. *Mol Med Rep*. 2019 Aug;20(2):1569–74.
87. Huang F, Tang J, He L, Ding X, Huang S, Zhang Y, et al. Heterologous expression and antitumor activity analysis of syringolin from *Pseudomonas syringae* pv. *syringae* B728a. *Microb Cell Fact*. 2018 Feb;17(1):31.

88. Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, Reimer LC, Göker M. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol*. 2020 Nov;70(11):5607–12.
89. Lindenberg A. Un nouveau mycétome. *Arch Parasit*. 1909;13:265–82.
90. GORDON RE, MIHM JM. A comparison of *Nocardia asteroides* and *Nocardia brasiliensis*. *J Gen Microbiol*. 1959 Feb;20(1):129–35.
91. Martínez-Barricarte R. Isolated Nocardiosis, an Unrecognized Primary Immunodeficiency? *Front Immunol*. 2020;11:590239.
92. Dumic I, Brown A, Magee K, Elwasila S, Kaljevic M, Antic M, et al. Primary Lymphocutaneous *Nocardia brasiliensis* in an Immunocompetent Host: Case Report and Literature Review. Vol. 58, *Medicina (Kaunas, Lithuania)*. 2022.
93. Herisse M, Ishida K, Porter JL, Howden B, Hertweck C, Stinear TP, et al. Identification and Mobilization of a Cryptic Antibiotic Biosynthesis Gene Locus from a Human-Pathogenic *Nocardia* Isolate. *ACS Chem Biol*. 2020 May;15(5):1161–8.
94. Emery D, Denning DW. The global distribution of actinomycetoma and eumycetoma. *PLoS Negl Trop Dis*. 2020 Sep;14(9):e0008397.
95. Olsen GJ, Woese CR. Ribosomal RNA: a key to phylogeny. *The FASEB Journal [Internet]*. 1993 Jan 1;7(1):113–23. Available from: <https://doi.org/10.1096/fasebj.7.1.8422957>
96. Conville PS, Brown-Elliott BA, Smith T, Zelazny AM. The Complexities of *Nocardia* Taxonomy and Identification. *J Clin Microbiol [Internet]*. 2017 Jan 1 [cited 2023 Nov 22];56(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/29118169/>
97. Traxler RM, Bell ME, Lasker B, Headd B, Shieh WJ, McQuiston JR. Updated Review on *Nocardia* Species: 2006–2021. *Clin Microbiol Rev [Internet]*. 2022 Dec 1 [cited 2024 Jun 10];35(4). Available from: [/pmc/articles/PMC9769612/](https://pubmed.ncbi.nlm.nih.gov/39769612/)
98. Vautrin F, Pujic P, Paquet C, Bergeron E, Mounié D, Marchal T, et al. Microbial risk assessment of *Nocardia cyriacigeorgica* in polluted environments, case of urban rainfall water. *Comput Struct Biotechnol J [Internet]*. 2020 Jan 1 [cited 2023 Nov 22];19:384–400. Available from: <https://pubmed.ncbi.nlm.nih.gov/33489008/>
99. Gnanam H, Rajapandian SGK, Gunasekaran R, Prithviraj SR, Ravindran RS, Sen S, et al. Molecular identification of *Nocardia* species causing endophthalmitis using multilocus sequence analysis (MLSA): a 10-year perspective. *J Med Microbiol [Internet]*. 2020 [cited 2023 Nov 22];69(5):728–38. Available from: <https://pubmed.ncbi.nlm.nih.gov/32310744/>
100. Carrasco G, Valdezate S, Garrido N, Villalón P, Medina-Pascual MJ, Sáez-Nieto JA. Identification, typing, and phylogenetic relationships of the main clinical *Nocardia* species in Spain according to their *gyrB* and *rpoB* genes. *J Clin*

- Microbiol [Internet]. 2013 Nov [cited 2023 Nov 22];51(11):3602–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/23966490/>
101. Wei M, Wang P, Yang C, Gu L. Molecular identification and phylogenetic relationships of clinical *Nocardia* isolates. *Antonie Van Leeuwenhoek* [Internet]. 2019 Dec 1 [cited 2023 Nov 22];112(12):1755–66. Available from: <https://pubmed.ncbi.nlm.nih.gov/31350617/>
 102. Sánchez-Herrera K, Sandoval H, Mouniee D, Ramírez-Durán N, Bergeron E, Boiron P, et al. Molecular identification of *Nocardia* species using the sod A gene: Identificación molecular de especies de *Nocardia* utilizando el gen sodA. *New Microbes New Infect* [Internet]. 2017 Sep 1 [cited 2023 Nov 22];19:96–116. Available from: <https://pubmed.ncbi.nlm.nih.gov/28794885/>
 103. Carrasco G, Valdezate S, Garrido N, Medina-Pascual MJ, Villalón P, Sáez-Nieto JA. GyrB analysis as a tool for identifying *Nocardia* species and exploring their phylogeny. *J Clin Microbiol* [Internet]. 2015 Mar 1 [cited 2024 Mar 18];53(3):997–1001. Available from: <https://journals.asm.org/doi/10.1128/jcm.03072-14>
 104. Chen KW, Lu CW, Huang TC, Lu CF, Liao YL, Lin JF, et al. Cutaneous manifestations of *Nocardia brasiliensis* infection in Taiwan during 2002-2012-clinical studies and molecular typing of pathogen by gyrB and 16S gene sequencing. *Diagn Microbiol Infect Dis*. 2013 Sep;77(1):74–8.
 105. Kosova-Maali D, Bergeron E, Maali Y, Durand T, Gonzalez J, Mouniée D, et al. High Intraspecific Genetic Diversity of *Nocardia brasiliensis*, a Pathogen Responsible for Cutaneous Nocardiosis Found in France: Phylogenetic Relationships by Using sod and hsp65 Genes. *Biomed Res Int*. 2018;2018:7314054.
 106. Rodríguez-Nava V, Couble A, Devulder G, Flandrois JP, Boiron P, Laurent F. Use of PCR-restriction enzyme pattern analysis and sequencing database for hsp65 gene-based identification of *Nocardia* species. *J Clin Microbiol* [Internet]. 2006 Feb [cited 2024 Mar 13];44(2):536–46. Available from: <https://pubmed.ncbi.nlm.nih.gov/16455910/>
 107. Brown-Elliott BA, Brown JM, Conville PS, Wallace RJ. Clinical and Laboratory Features of the *Nocardia* spp. Based on Current Molecular Taxonomy. *Clin Microbiol Rev* [Internet]. 2006 Apr [cited 2025 May 18];19(2):259–82. Available from: <https://journals.asm.org/doi/10.1128/cmr.19.2.259-282.2006>
 108. Salinas-Carmona MC. *Nocardia brasiliensis*: from microbe to human and experimental infections. *Microbes Infect*. 2000 Sep 1;2(11):1373–81.
 109. McTaggart LR, Doucet J, Witkowska M, Richardson SE. Antimicrobial Susceptibility among Clinical *Nocardia* Species Identified by Multilocus Sequence Analysis. *Antimicrob Agents Chemother* [Internet]. 2015 Jan 1 [cited 2024 Jun 11];59(1):269. Available from: [/pmc/articles/PMC4291361/](https://pubmed.ncbi.nlm.nih.gov/25411111/)

110. Wang H, Zhu Y, Cui Q, Wu W, Li G, Chen D, et al. Epidemiology and Antimicrobial Resistance Profiles of the *Nocardia* Species in China, 2009 to 2021. *Microbiol Spectr* [Internet]. 2022 Apr 27 [cited 2024 Jun 10];10(2). Available from: /pmc/articles/PMC9045199/
111. Cruz Medrano MG. Estudio de variaciones genéticas de cepas de *Nocardia brasiliensis*, basado en técnicas de restricción enzimática [Internet] [Tesis de Maestría]. [Toluca, México]: Universidad Autónoma del Estado de México; 2022 [cited 2025 Oct 22]. Available from: <http://ri.uaemex.mx/handle/20.500.11799/112279>
112. World Health Organization. Neglected tropical diseases [Internet]. [cited 2022 Aug 14]. Available from: https://www.who.int/health-topics/neglected-tropical-diseases#tab=tab_1
113. Arenas R, Fernandez Martinez RF, Torres-Guerrero E, Garcia C. Actinomycetoma: An update on diagnosis and treatment. *Cutis*. 2017;99(2):E11–5.
114. Hao X, Cognetti M, Burch-Smith R, Mejia EO, Mirkin G. Mycetoma: Development of Diagnosis and Treatment. *J Fungi (Basel)*. 2022 Jul;8(7).
115. Reis CMS, Reis-Filho EG de M. Mycetomas: an epidemiological, etiological, clinical, laboratory and therapeutic review. *An Bras Dermatol*. 2018;93(1):8–18.
116. Cárdenas-de la Garza JA, Welsh O, Cuéllar-Barboza A, Suarez-Sánchez KP, De la Cruz-Valadez E, Cruz-Gómez LG, et al. Clinical characteristics and treatment of actinomycetoma in northeast Mexico: A case series. *PLoS Negl Trop Dis*. 2020 Feb;14(2):e0008123.
117. Zhu JW, Zhou H, Jia WQ, You J, Xu RX. A clinical case report of brain abscess caused by *Nocardia brasiliensis* in a non-immunocompromised patient and a relevant literature review. *BMC Infect Dis*. 2020 May;20(1):328.
118. Robinson SL, Terlouw BR, Smith MD, Pidot SJ, Stinear TP, Medema MH, et al. Global analysis of adenylate-forming enzymes reveals β -lactone biosynthesis pathway in pathogenic *Nocardia*. *J Biol Chem*. 2020 Oct;295(44):14826–39.
119. Itoh J, Miyadoh S. SF2457, A NEW ANTIBIOTIC RELATED TO AMICETIN. *J Antibiot (Tokyo)*. 1992 Jun 25;45(6):846–53.
120. Ishibashi M, Kobayashi J, Mikami Y. THE PRODUCER AND BIOLOGICAL ACTIVITIES OF SO-075R1, A NEW MUTACTIMYCIN GROUP ANTIBIOTIC. *J Antibiot (Tokyo)*. 1992 Dec 25;45(12):1848–52.
121. Kiyoshi T, Akio F, Shigeyasu N, Masatomo F, Masaru F, Kazuo K. PC-766B, A NEW MACROLIDE ANTIBIOTIC PRODUCED BY *Nocardia brasiliensis* I. TAXONOMY, FERMENTATION AND BIOLOGICAL ACTIVITY. *J Antibiot (Tokyo)*. 1993 Jun 25;46(6):972–8.

122. Maeda A, Nagai H, Yazawa K, Tanaka Y, Imai T, Mikami Y, et al. THREE NEW REDUCED ANTHRACYCLINE RELATED COMPOUNDS FROM PATHOGENIC *Nocardia brasiliensis*. *J Antibiot (Tokyo)*. 1994 Sep 25;47(9):976–81.
123. Tanaka Y, Komaki H, Yazawa K, Mikami Y, Nemoto A, Tojyo T, et al. Brasilinolide A, a New Macrolide Antibiotic Produced by *Nocardia brasiliensis*: Producing Strain, Isolation and Biological Activity. *J Antibiot (Tokyo)*. 1997 Dec 25;50(12):1036–41.
124. Nemoto A, Tanaka Y, Karasaki Y, Komaki H, Yazawa K, Mikami Y, et al. Brasiliquinones A, B and C, new benz[alpha]anthraquinone antibiotics from *Nocardia brasiliensis*. I. Producing strain, isolation and biological activities of the antibiotics. *J Antibiot (Tokyo)* [Internet]. 1997 [cited 2024 Oct 28];50(1):18–21. Available from: <https://pubmed.ncbi.nlm.nih.gov/9066761/>
125. Kobayashi J, Tsuda M, Nemoto A, Tanaka Y, Yazawa K, Mikami Y. Brasilidine A, a new cytotoxic isonitrile-containing indole alkaloid from the actinomycete *Nocardia brasiliensis*. *J Nat Prod* [Internet]. 1997 [cited 2024 Oct 28];60(7):719–20. Available from: <https://pubmed.ncbi.nlm.nih.gov/9249977/>
126. Momose I, Kinoshita N, Sawa R, Naganawa H, Inuma H, Hamada M, et al. Nothramicin, a New Anthracycline Antibiotic from *Nocardia* sp. MJ896-43F17. *J Antibiot (Tokyo)*. 1998 Feb 25;51(2):130–5.
127. Tsuda M, Nemoto A, Komaki H, Tanaka Y, Yazawa K, Mikami Y, et al. Nocarasins A–C and Brasiliquinone D, New Metabolites from the Actinomycete *Nocardia brasiliensis*. *J Nat Prod* [Internet]. 1999 Dec [cited 2025 Oct 12];62(12):1640–2. Available from: <https://pubs.acs.org/doi/abs/10.1021/np990265v>
128. Komaki H, Nemoto A, Tanaka Y, Takagi H, Yazawa K, Mikami Y, et al. Brasilicardin A, a new terpenoid antibiotic from pathogenic *Nocardia brasiliensis*: fermentation, isolation and biological activity. *J Antibiot (Tokyo)* [Internet]. 1999 [cited 2024 Oct 28];52(1):13–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/10092191/>
129. Mikami Y, Yazawa Y, Tanaka Y, Ritzau M, Gräfe U. Isolation and Structure of Nocardiolactone, A New Dialkyl-Substituted β -Lactone From Pathogenic *Nocardia* Strains. *Nat Prod Lett* [Internet]. 1999 [cited 2025 Oct 12];13(4):277–84. Available from: <https://www.tandfonline.com/doi/abs/10.1080/10575639908048798>
130. Mikami Y, Yazawa K, Nemoto A, Komaki H, Tanaka Y, Grafe U. Production of Erythromycin E by Pathogenic *Nocardia brasiliensis*. *J Antibiot (Tokyo)*. 1999 Feb 25;52(2):201–2.
131. Mikami Y, Komaki H, Imai T, Yazawa K, Nemoto A, Tanaka Y, et al. A New Antifungal Macrolide Component, Brasilinolide B, Produced by *Nocardia brasiliensis*. *J Antibiot (Tokyo)*. 2000 Jan 25;53(1):70–4.

132. Komatsu K, Tsuda M, Shiro M, Tanaka Y, Mikami Y, Kobayashi J. Brasilicardins B–D, new tricyclic terpenoids from actinomycete *Nocardia brasiliensis*. *Bioorg Med Chem*. 2004 Nov 1;12(21):5545–51.
133. Tsuda M, Yamakawa M, Oka S, Tanaka Y, Hoshino Y, Mikami Y, et al. Brasilibactin A, a cytotoxic compound from actinomycete *Nocardia brasiliensis*. *J Nat Prod* [Internet]. 2005 Mar [cited 2025 Oct 8];68(3):462–4. Available from: <https://pubs.acs.org/doi/abs/10.1021/np0496385>
134. Fahal AH. Mycetoma: A global medical and socio-economic dilemma. *PLoS Negl Trop Dis*. 2017 Apr;11(4):e0005509.
135. National Library of Medicine (US) NC for BI. Genome [Internet]. Bethesda (MD). 2022 [cited 2022 Sep 20]. Available from: <https://www.ncbi.nlm.nih.gov/genome/>
136. Schatz A, Bugie E, Waksman SA. Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. 1944. *Clin Orthop Relat Res* [Internet]. 2005 [cited 2025 Oct 12];437:3–6. Available from: https://journals.lww.com/clinorthop/fulltext/2005/08000/the_classic__streptomycin,_a_substance_exhibiting.2.aspx
137. Junior JCR, Tamanini R, Soares BF, Oliveira AM de, Silva F de G, Silva FF da, et al. Efficiency of boiling and four other methods for genomic DNA extraction of deteriorating spore-forming bacteria from milk. *Semin Cienc Agrar* [Internet]. 2016 [cited 2024 Mar 31];37(5):3069–78. Available from: <https://www.redalyc.org/articulo.oa?id=445748363019>
138. Hall TA. BioEdit: A user-friendly biological sequence alignment program for Windows 95/98/NT. *Nucleic Acids Symp Ser*. 1999;41(41):95–8.
139. Madeira F, Madhusoodanan N, Lee J, Eusebi A, Niewielska A, Tivey ARN, et al. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res* [Internet]. 2024 Jul 5 [cited 2025 May 18];52(W1):W521–5. Available from: <https://dx.doi.org/10.1093/nar/gkae241>
140. Castresana J. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Mol Biol Evol* [Internet]. 2000 Apr 1 [cited 2025 May 18];17(4):540–52. Available from: <https://dx.doi.org/10.1093/oxfordjournals.molbev.a026334>
141. LastNameBethesda (MD): National Library of Medicine (US) NC for BI 2004. ORFfinder [Internet]. [cited 2025 May 18]. Available from: <https://www.ncbi.nlm.nih.gov/orffinder/>
142. Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol Biol Evol* [Internet]. 2017 Dec 1 [cited 2025 May 18];34(12):3299–302. Available from: <https://dx.doi.org/10.1093/molbev/msx248>
143. Pritchard JK, Stephens M, Donnelly P. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* [Internet]. 2000 Jun 1 [cited 2025 May

- 18];155(2):945–59. Available from: <https://dx.doi.org/10.1093/genetics/155.2.945>
144. Li YL, Liu JX. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol Ecol Resour* [Internet]. 2018 Jan 1 [cited 2025 May 18];18(1):176–7. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/1755-0998.12719>
 145. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol* [Internet]. 2020 May 1 [cited 2025 May 18];37(5):1530–4. Available from: <https://dx.doi.org/10.1093/molbev/msaa015>
 146. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol* [Internet]. 2018 Feb 1 [cited 2025 May 18];35(2):518–22. Available from: <https://dx.doi.org/10.1093/molbev/msx281>
 147. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* 2017 14:6 [Internet]. 2017 May 8 [cited 2025 May 18];14(6):587–9. Available from: <https://www.nature.com/articles/nmeth.4285>
 148. Rambaut A. FigTree. Tree Figure Drawing Tool Version 1.4.4 [Internet]. 2018 [cited 2025 Jun 18]. Available from: <http://tree.bio.ed.ac.uk/software/figtree/>
 149. Zhang J, Kapli P, Pavlidis P, Stamatakis A. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* [Internet]. 2013 Nov 15 [cited 2025 Jun 22];29(22):2869–76. Available from: <https://dx.doi.org/10.1093/bioinformatics/btt499>
 150. Spöri Y, Stoch F, Dellicour S, Birky CW, Flot JF. KoT: an automatic implementation of the K/θ method for species delimitation. *bioRxiv*. 2022 Feb 28;2021.08.17.454531.
 151. González-Nava JB, Manzanares-Leal GL, Zapi-Colín LÁ, Dávila-Ramos S, Sandoval-Trujillo H, Ramírez-Durán N. Antibacterial activity of *Nocardia* spp. and *Streptomyces* sp. on multidrug-resistant pathogens causing neonatal sepsis. *Rev Inst Med Trop Sao Paulo* [Internet]. 2024 [cited 2024 Nov 20];66:e42. Available from: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11295289/>
 152. Andrews S. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data [Internet]. 2010. Available from: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
 153. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUASt: quality assessment tool for genome assemblies. *Bioinformatics* [Internet]. 2013 Apr 15 [cited 2025 Oct 14];29(8):1072–5. Available from: <https://dx.doi.org/10.1093/bioinformatics/btt086>

154. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. Bakta: Rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb Genom* [Internet]. 2021 Nov 5 [cited 2025 Oct 14];7(11):000685. Available from: <https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000685>
155. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* [Internet]. 2014 Jul 15 [cited 2025 Oct 14];30(14):2068–9. Available from: <https://dx.doi.org/10.1093/bioinformatics/btu153>
156. Estrada K, Ayixon. kjestradag/RaPDTTool: 2.2.0 [Internet]. Zenodo; 2023. Available from: <https://doi.org/10.5281/zenodo.8327564>
157. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biol* [Internet]. 2016 Jun 20 [cited 2025 Oct 14];17(1):1–14. Available from: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-0997-x>
158. Kim D, Park S, Chun J. Introducing EzAAI: a pipeline for high throughput calculations of prokaryotic average amino acid identity. *Journal of Microbiology* [Internet]. 2021 May 1 [cited 2025 Oct 14];59(5):476–80. Available from: <https://link.springer.com/article/10.1007/s12275-021-1154-0>
159. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Analytical Methods* [Internet]. 2015 Dec 17 [cited 2025 Oct 14];8(1):12–24. Available from: <https://pubs.rsc.org/en/content/articlehtml/2016/ay/c5ay02550h>
160. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications* 2018 9:1 [Internet]. 2018 Nov 30 [cited 2025 Oct 14];9(1):1–8. Available from: <https://www.nature.com/articles/s41467-018-07641-9>
161. Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen CY, et al. Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Res* [Internet]. 2023 Jul 5 [cited 2025 Oct 14];51(W1):W484–92. Available from: <https://dx.doi.org/10.1093/nar/gkad326>
162. Contreras-Moreira B, Vinuesa P. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl Environ Microbiol* [Internet]. 2013 Dec 15 [cited 2025 Oct 14];79(24):7696–701. Available from: <https://journals.asm.org/doi/10.1128/AEM.02411-13>
163. Emms DM, Liu Y, Belcher L, Holmes J, Kelly S. OrthoFinder: scalable phylogenetic orthology inference for comparative genomics. *bioRxiv* [Internet].

- 2025 Jul 16 [cited 2025 Oct 14];2025.07.15.664860. Available from: <https://www.biorxiv.org/content/10.1101/2025.07.15.664860v1>
164. Emms DM, Kelly S. STAG: Species Tree Inference from All Genes. bioRxiv [Internet]. 2018 Feb 19 [cited 2024 Nov 11];267914. Available from: <https://www.biorxiv.org/content/10.1101/267914v1>
165. Lefort V, Desper R, Gascuel O. FastME 2.0: A Comprehensive, Accurate, and Fast Distance-Based Phylogeny Inference Program. *Mol Biol Evol* [Internet]. 2015 Oct 1 [cited 2025 Oct 14];32(10):2798–800. Available from: <https://dx.doi.org/10.1093/molbev/msv150>
166. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res* [Internet]. 2024 Jul 5 [cited 2025 Oct 14];52(W1):W78–82. Available from: <https://dx.doi.org/10.1093/nar/gkae268>
167. Purev E, Kondo T, Takemoto D, Niones JT, Ojika M. Identification of ϵ -Poly-L-lysine as an Antimicrobial Product from an *Epichloë* Endophyte and Isolation of Fungal ϵ -PL Synthetase Gene. *Molecules* 2020, Vol 25, Page 1032 [Internet]. 2020 Feb 25 [cited 2025 Oct 15];25(5):1032. Available from: <https://www.mdpi.com/1420-3049/25/5/1032/htm>
168. López-Rayó S, Sanchis-Pérez I, Ferreira CMH, Lucena JJ. [S,S]-EDDS/Fe: A new chelate for the environmentally sustainable correction of iron chlorosis in calcareous soil. *Science of The Total Environment*. 2019 Jan 10;647:1508–17.
169. Prabhu J, Schauwecker F, Grammel N, Keller U, Bernhard M. Functional expression of the ectoine hydroxylase gene (*thpD*) from *Streptomyces chrysomallus* in *Halomonas elongata*. *Appl Environ Microbiol* [Internet]. 2004 May [cited 2025 Oct 15];70(5):3130–2. Available from: <https://journals.asm.org/doi/10.1128/AEM.70.5.3130-3132.2004>
170. Wang X, Zhou H, Chen H, Jing X, Zheng W, Li R, et al. Discovery of recombinases enables genome mining of cryptic biosynthetic gene clusters in Burkholderiales species. *Proc Natl Acad Sci U S A* [Internet]. 2018 May 1 [cited 2025 Oct 15];115(18):E4255–63. Available from: <https://www.pnas.org/doi/abs/10.1073/pnas.1720941115>