

Tesina para optar al título de Licenciada en Bioinformática

*“Identificación de especies bacterianas del género
Leptospira y análisis preliminar de patogenicidad a partir de
datos de secuenciación en muestras ambientales del río
Salado, río Coronda y río Paraná, y lagunas asociadas, en la
ciudad de Santa Fe”*

- **Estudiante:** Mohamed, Gianella Tamara.
- **Director:** Dr. Cabrera, Juan Manuel.
- **Co-directora:** Dra. Rueda, Eva Carolina.

“Caminante, no hay camino, se hace camino al andar.”

Antonio Machado, Proverbios y cantares XXIX.

AGRADECIMIENTOS

La realización de esta tesina representa el cierre de una etapa muy importante de mi vida y el comienzo de nuevos desafíos. Ningún logro se construye en soledad, y por eso quiero agradecer a todas las personas que me acompañaron durante este camino.

A Juan y Eva, gracias por ser los directores que cualquier tesinista desearía tener. Por acompañarme con paciencia, entusiasmo y una enorme calidad humana; por responder dudas, compartir conocimientos y hacerme sentir apoyada en cada etapa del proceso. Haber contado con ustedes fue, sin dudas, una de las mayores fortalezas de este trabajo.

A Bruno, mi compañero de vida, gracias por estar a mi lado en cada momento. Por escucharme hablar de este proyecto una y otra vez, por celebrar cada pequeño avance, por sostenerme en los días difíciles y por recordarme que podía seguir cuando yo misma dudaba. Tu amor, tu paciencia y tu apoyo incondicional fueron fundamentales para llegar hasta acá. Esta tesina también lleva un pedacito de vos, porque sin tu compañía este camino habría sido mucho más difícil.

A mi familia, gracias por acompañarme siempre, por confiar en mí incluso cuando las cosas parecían complicarse, y por brindarme el amor y el apoyo que hicieron posible cada paso de esta carrera. Saber que siempre estuvieron ahí me dio la fuerza necesaria para seguir adelante.

A mis amigos, colegas y futuros colegas, Valentina, Luisina, Priscila, Bianca y Dardo, gracias por compartir esta etapa conmigo. Por las horas de estudio, las charlas, las risas, los consejos, los mates compartidos, el compañerismo y el apoyo mutuo. Haber recorrido este camino junto a ustedes hizo que cada desafío fuera más llevadero y cada logro mucho más especial.

A todos mis profesores, gracias por transmitir sus conocimientos, despertar mi curiosidad y contribuir a mi formación profesional. Cada enseñanza recibida dejó una huella que me acompañará en el futuro.

A mis abuelos, Ana y Pancho, quienes ya no están físicamente conmigo, pero cuya presencia sentí durante toda esta carrera. Gracias por acompañarme de la manera más hermosa que conocían: encendiendo una vela antes de cada examen, confiando siempre en mí y deseándome lo mejor en cada paso que daba. Hoy me gusta pensar que esa luz siguió iluminando mi camino incluso después de su partida. Este logro también es para ustedes.

Finalmente, agradezco a todas aquellas personas que, de una u otra manera, formaron parte de este proceso. A quienes me alentaron, me escucharon, me enseñaron, me acompañaron o simplemente estuvieron presentes cuando lo necesité. Cada gesto, por pequeño que pareciera, contribuyó a que hoy pueda alcanzar esta meta.

Gracias.

INTRODUCCIÓN

La leptospirosis como problema de salud pública

La leptospirosis es una enfermedad infecciosa aguda y febril causada por espiroquetas patógenas del género *Leptospira* (Morey *et al.*, 2006; Ministerio de Salud de la Nación, 2025b). Se trata de la zoonosis bacteriana de mayor distribución geográfica a nivel global, presentándose en todos los continentes con excepción de la Antártida (Adler & de la Peña Moctezuma, 2010; Ministerio de Salud de la Nación, 2025b). Esta patología se considera un paradigma de la interfaz "Una Salud" (*One Health*), ya que su ciclo de transmisión involucra de manera crítica la interacción entre seres humanos, diversos reservorios animales y el entorno ambiental, particularmente las matrices hídricas (Ministerio de Salud de la Nación, 2025a; Costa *et al.*, 2015; Warnasekara *et al.*, 2022).

Debido a su impacto sanitario, la Organización Mundial de la Salud (OMS) la ha clasificado como una enfermedad tropical desatendida de importancia mundial (Lau *et al.*, 2010; Ministerio de Salud de la Nación, 2025b; Lotto Batista, 2023). Se estima que anualmente ocurren aproximadamente 1,03 millones de casos severos y cerca de 58.900 muertes en todo el mundo (Ministerio de Salud de la Nación, 2025b; Lotto Batista, 2023; Agudelo-Flórez *et al.*, 2014; Costa *et al.*, 2015). La carga de la enfermedad es desproporcionadamente alta en regiones tropicales y subtropicales de países en desarrollo, donde factores como la urbanización desorganizada, la falta de saneamiento básico y los eventos climáticos extremos favorecen la transmisión (Adler, 2014; Gómez *et al.*, 2022; Picardeau, 2017). La mortalidad por formas graves, como el síndrome de Weil (caracterizado por ictericia, falla renal y hemorragias) o el síndrome de hemorragia pulmonar grave, puede superar el 10% y el 50%, respectivamente (Lotto Batista, 2023; Bierque *et al.*, 2020; McBride *et al.*, 2005).

En la República Argentina, la leptospirosis es un evento de notificación obligatoria y representa un problema de salud pública reemergente (Ministerio de Salud de la Nación, 2025b). La provincia de Santa Fe presenta históricamente una de las mayores tasas de incidencia anual del país, concentrando una parte sustancial de las notificaciones nacionales (Ricardo *et al.*, 2018; Gómez *et al.*, 2022). De hecho, el primer caso documentado en humanos en Argentina fue reportado precisamente en la provincia de Santa Fe (Rosario) en el año 1915 (Ricardo *et al.*, 2018; Ministerio de Salud de la Nación, 2025b). La vulnerabilidad de esta región está intrínsecamente ligada a su configuración hidrológica; la ciudad de Santa

Fe, situada en la confluencia de los ríos Paraná y Salado, está expuesta a recurrentes riesgos hídricos por crecidas fluviales y lluvias intensas (Ricardo *et al.*, 2018). Grandes brotes epidémicos se han registrado localmente en asociación directa con inundaciones, como los ocurridos en 1998, 2003, 2007, 2010 y 2016 (Vanasco *et al.*, 2000; López *et al.*, 2019; Lotto Batista, 2023).

La bacteria se mantiene en la naturaleza mediante la infección crónica de los túbulos renales de reservorios animales, que actúan como portadores asintomáticos y excretan el agente a través de la orina durante períodos prolongados (Bierque *et al.*, 2020; Picardeau, 2017; Ministerio de Salud de la Nación, 2025b). Los roedores, particularmente el género *Rattus*, son los reservorios más importantes y distribuidos a nivel mundial (Adler, 2014; Ministerio de Salud de la Nación, 2025b; Lotto Batista, 2023). No obstante, una amplia variedad de animales domésticos (perros, ganado vacuno, porcino y equino) y silvestres también desempeñan un papel crucial en la epidemiología local (Ministerio de Salud de la Nación, 2025b; Ministerio de Salud de la Nación, 2025a; Lau *et al.*, 2010). En Argentina, los perros son considerados centinelas clave de la contaminación ambiental en áreas urbanas, habiéndose aislado serovariedades como *Canicola* en casos humanos graves de la región (Ministerio de Salud de la Nación, 2025a; Ricardo *et al.*, 2018).

La transmisión al ser humano es principalmente indirecta, a través del contacto de la piel (especialmente si presenta lesiones) o mucosas con agua, suelo o barro contaminados con orina de animales infectados (Lotto Batista, 2023; Bierque *et al.*, 2020; Bilung *et al.*, 2026; Ministerio de Salud de la Nación, 2025b). Los ambientes acuáticos son el vehículo principal de diseminación, ya que las leptospiras dependen de la humedad para sobrevivir fuera del hospedador y pueden persistir en el agua dulce durante varios meses (Picardeau, 2017; Vanasco *et al.*, 2000; Vijayachari *et al.*, 2008). Las inundaciones actúan como un factor de riesgo crítico al forzar la salida de los roedores de sus madrigueras hacia asentamientos humanos y al dispersar la carga bacteriana en aguas superficiales de contacto recreativo u ocupacional (Gómez *et al.*, 2022; Ministerio de Salud de la Nación, 2025a; Lau *et al.*, 2010; Wynwood *et al.*, 2014).

Dada la complejidad epidemiológica y las limitaciones de los métodos diagnósticos convencionales para detectar el agente en matrices ambientales, es imperativo profundizar en el estudio tanto de la ecología así como de la estructura biológica de estas bacterias. El conocimiento de las especies circulantes en los recursos hídricos es el primer paso para

evaluar el riesgo sanitario real. Por ello, es necesario abordar el análisis detallado del agente etiológico responsable de esta dinámica: el género *Leptospira*.

El género *Leptospira*

El agente etiológico de la leptospirosis es el género *Leptospira*, un grupo de espiroquetas morfológicamente distintivas que pertenecen a la familia *Leptospiraceae* (Bilung *et al.*, 2026; Adler & de la Peña Moctezuma, 2010). Estas bacterias se caracterizan por ser extremadamente finas, con un diámetro de aproximadamente 0.1 μm y una longitud que oscila entre 6 y 20 μm (Bilung *et al.*, 2026; Adler, 2014). Su estructura es helicoidal y presentan extremos curvados en forma de gancho o interrogante, una particularidad que les otorgó su nombre específico original, *Spirochaeta interrogans* (Figura 1) (Bilung *et al.*, 2026; Ministerio de Salud de la Nación, 2025b). Son organismos aerobios obligados y poseen una movilidad excepcional gracias a dos endoflagelos periplásmicos insertados subterminalmente en cada polo, lo que les permite desplazarse con rapidez incluso en medios altamente viscosos (Bilung *et al.*, 2026; Adler, 2014; Picardeau, 2017). Su envoltura celular es similar a la de las bacterias Gram-negativas, con una membrana externa que contiene lipopolisacárido (LPS), aunque este presenta una toxicidad significativamente menor en comparación con el de otras bacterias Gram-negativas (Adler & de la Peña Moctezuma, 2010; Picardeau, 2017).

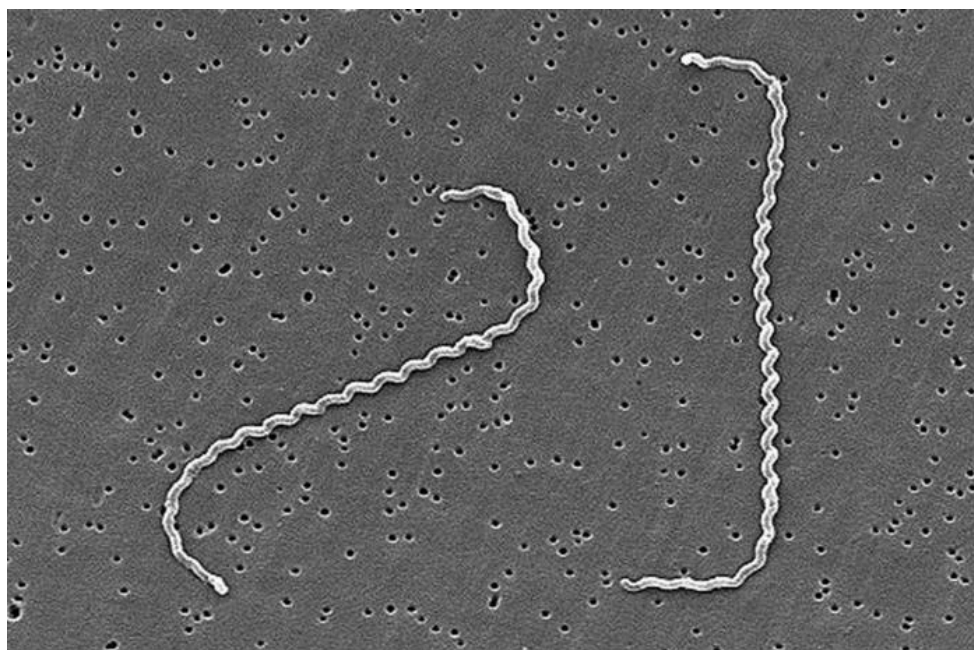


Figura 1. Micrografía electrónica de barrido de *Leptospira interrogans*. **Nota.** Se observan espiroquetas con morfología helicoidal y extremos curvados en forma de gancho, características del género. Imagen obtenida de Public Health Image Library (PHIL), por Centers for Disease Control and Prevention, ID 1220.

La clasificación del género ha experimentado una profunda transformación con el advenimiento de la era genómica. Históricamente, se dividía de manera simplificada en dos especies basadas en el fenotipo: la patógena *Leptospira interrogans* y la saprófita *Leptospira biflexa* (Bilung *et al.*, 2026; Picardeau, 2017). Sin embargo, la taxonomía actual, fundamentada en la homología de ADN y análisis filogenéticos, reconoce una diversidad taxonómica mucho mayor, con más de 66 especies descritas (Caimi & Ruybal, 2020; Picardeau, 2017). Estas se organizan principalmente en tres grandes grupos o clados definidos por su virulencia y nicho ecológico: especies patógenas (Clado P1), que incluyen los agentes de alta virulencia responsables de la mayoría de los casos graves en humanos y animales; especies intermedias (Clado P2), que presentan una patogenicidad oportunista y suelen causar cuadros clínicos más leves; y especies saprófitas (Clados S1 y S2), microorganismos de vida libre que habitan el agua y el suelo sin causar enfermedad (Figura 2) (Bilung *et al.*, 2026; Picardeau, 2017).

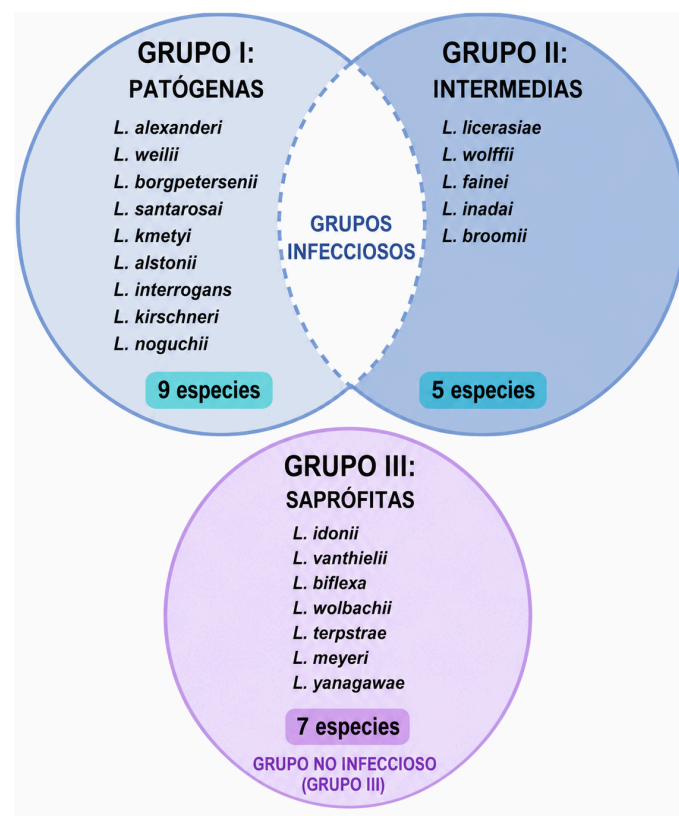


Figura 2. Clasificación tradicional de las especies del género *Leptospira* basada en las secuencias del gen 16S *rRNA*. Las especies se agrupan en tres categorías principales: patógenas (Grupo I), intermedias (Grupo II) y saprófitas (Grupo III). Los grupos I y II conforman las leptospirosis infecciosas, mientras que el Grupo III corresponde a especies no infecciosas. Imagen modificada de Bilung *et al.* (2026).

La relevancia epidemiológica de las especies patógenas, como *Leptospira interrogans*, *Leptospira kirschneri* y *Leptospira santarosai*, radica en su capacidad para colonizar crónicamente los túbulos renales de los reservorios animales (Bilung *et al.*, 2026; Adler, 2014; Picardeau, 2017). Desde este nicho, las bacterias son excretadas continuamente a través de la orina, permitiendo su persistencia y diseminación en el ambiente (Caimi & Ruybal, 2020; Lotto Batista, 2023). La capacidad de estos linajes para evadir el sistema inmune del hospedador y su adaptación genética —que incluye la adquisición de factores de virulencia específicos como el gen *LipL32*— son determinantes críticos para el éxito de la infección (Picardeau, 2017; Bilung *et al.*, 2026; Adler, 2014).

Dada esta complejidad, los métodos tradicionales de identificación fenotípica y serológica resultan insuficientes para discriminar la vasta diversidad de especies y serovariedades presentes en entornos naturales (Bilung *et al.*, 2026). Por lo tanto, el estudio del género ha girado hacia la caracterización molecular y genómica, utilizando herramientas bioinformáticas para analizar genes marcadores y genomas completos. Este enfoque permite una delimitación taxonómica precisa y es fundamental para comprender la ecología y el riesgo sanitario que representan estas bacterias en matrices ambientales (Picardeau, 2017; Bilung *et al.*, 2026; Caimi & Ruybal, 2020).

Caracterización molecular y genómica de *Leptospira*

La arquitectura genómica del género *Leptospira* es notablemente compleja y distintiva dentro del orden *Spirochaetales*. El tamaño del genoma oscila entre 3,8 y 5,1 megabases (Mb), con un contenido de G+C que varía entre el 33% y el 43% según la especie (Bilung *et al.*, 2026; Adler, 2014; Vijayachari *et al.*, 2008; Picardeau, 2017). Una de sus características más inusuales es la organización en dos cromosomas circulares (Bilung *et al.*, 2026; Adler, 2014; Bharti *et al.*, 2003). El cromosoma I (CI), de mayor tamaño (~3,6 a 4,4 Mb), alberga la mayoría de los genes de mantenimiento o *housekeeping* y factores de virulencia esenciales (Bilung *et al.*, 2026; Adler, 2014). Por su parte, el cromosoma II (CII), que funciona como un cromidio (~278 a 400 kb), porta genes vinculados a funciones accesorias, detección ambiental y adaptación al hospedador (Bilung *et al.*, 2026; Adler, 2014). Además, algunas especies como *Leptospira biflexa* poseen un tercer replicón circular denominado *p74* (Adler, 2014; Levett, 2001). Esta plasticidad genómica, impulsada por procesos de duplicación génica y transferencia horizontal, ha permitido al género colonizar nichos ecológicos tan

diversos como los túbulos renales de mamíferos y los sedimentos acuáticos (Picardeau, 2017; Bharti *et al.*, 2003).

La identificación molecular ha cobrado una importancia capital debido a que los métodos tradicionales presentan limitaciones severas (Chiani *et al.*, 2016). El cultivo bacteriológico es extremadamente lento, pudiendo requerir de 6 a 13 semanas, y es frecuentemente ineficaz en muestras hídricas debido a la competencia de la microbiota acompañante (Picardeau, 2012; Ministerio de Salud de la Nación, 2025b). Asimismo, la clasificación serológica mediante la Prueba de Aglutinación Microscópica (MAT) es laboriosa y no permite una discriminación precisa a nivel de especie genómica (McBride *et al.*, 2005; Picardeau, 2017). En este contexto, las herramientas moleculares permiten una detección directa y específica, optimizando la vigilancia epidemiológica y la comprensión de la diversidad del género (McBride *et al.*, 2005).

Para los estudios taxonómicos y filogenéticos, se emplean diversos genes marcadores. El gen *16S rRNA* (*rrs*) constituye el estándar universal para la asignación de clados; sin embargo, su alta conservación impide, en ocasiones, discriminar entre especies estrechamente relacionadas, separadas a veces por apenas dos o tres nucleótidos (Bilung *et al.*, 2026; Chiani *et al.*, 2016; Adler, 2014; Morey *et al.*, 2006). Para superar esta limitación, se integran marcadores *housekeeping* con mayor polimorfismo, como el gen *secY*, que codifica una preproteína translocasa y ofrece un poder discriminatorio superior para la identificación intra-género (Bierque *et al.*, 2020; Adler, 2014; Caimi & Ruybal, 2020). Complementariamente, se utiliza el gen *LipL32*, que codifica la lipoproteína de membrana externa más abundante del género (Adler, 2014). Dado que este gen está presente exclusivamente en linajes patógenos e intermedios y ausente en saprófitos, su detección funciona como un marcador molecular crítico de potencial patogenicidad (Bierque *et al.*, 2020; Agudelo-Flórez *et al.*, 2014; Vital-Brazil *et al.*, 2010).

La aplicación de la secuenciación de ADN ha revolucionado este campo. Mientras que la secuenciación Sanger sigue siendo valiosa para la caracterización de alta fidelidad de blancos genéticos individuales (como *secY* o *LipL32*), las tecnologías de secuenciación masiva (NGS), como Illumina, permiten un análisis profundo de la diversidad microbiana mediante la secuenciación de amplicones a gran escala (Picardeau, 2017). El procesamiento de este volumen masivo de datos depende estrictamente de las herramientas bioinformáticas. La implementación de *pipelines* que incluyen el control de calidad (FastQC), la limpieza de

secuencias (*trimming*), el ensamblaje *de novo* (SPAdes) y la posterior asignación taxonómica mediante algoritmos de alineamiento local como BLASTn, permite comparar los hallazgos contra bases de datos globales como GenBank (Agudelo-Flórez *et al.*, 2014; Pavan *et al.*, 2010).

Bajo este marco técnico y conceptual, el presente trabajo aborda la caracterización de secuencias obtenidas a partir de una exhaustiva campaña de muestreo ambiental en la región de Santa Fe. El análisis integrado de los marcadores *16S rRNA*, *secY* y *LipL32* permitirá descifrar la compleja estructura poblacional de *Leptospira* en los recursos hídricos locales, aportando información esencial sobre la ecología del patógeno y el riesgo sanitario en zonas vulnerables.

Contexto del estudio, justificación y pregunta de investigación

La ciudad de Santa Fe constituye uno de los principales escenarios epidemiológicos de leptospirosis en Argentina, debido a su ubicación en un valle de inundación y a la frecuente ocurrencia de crecidas, precipitaciones intensas y anegamientos. Sin embargo, a pesar de la importancia de los ambientes acuáticos en la transmisión de la enfermedad, aún existe un conocimiento limitado acerca de la diversidad de especies de *Leptospira* que habitan los cuerpos de agua de la región y de su potencial patogénico. Históricamente, gran parte de las investigaciones se han centrado en los casos clínicos humanos y en los reservorios animales, mientras que la caracterización de las poblaciones bacterianas presentes en el ambiente ha recibido menor atención.

El presente estudio se enmarca en el proyecto de investigación “Determinantes socio-ambientales de leptospirosis y dengue en la ciudad de Santa Fe”, dirigido por la Dra. Previtali, cuyo principal objetivo consiste en comprender las múltiples dimensiones biomédicas, ambientales, epidemiológicas y socioculturales asociadas a estas zoonosis. Para ello, se analizaron secuencias de ADN obtenidas a partir de muestras ambientales de agua dulce recolectadas en distintos puntos de los ríos Salado, Coronda y Paraná, así como en lagunas asociadas a estos sistemas fluviales. La selección de estos ambientes resulta particularmente relevante debido a la compleja dinámica hídrica de la región y a las condiciones ambientales que favorecen la persistencia y dispersión de bacterias del género *Leptospira*.

El desarrollo de esta tesina adquiere especial relevancia considerando que el contacto con agua contaminada constituye la principal vía de transmisión para el ser humano y que, las limitaciones inherentes a los métodos de cultivo tradicionales dificultan la detección y diferenciación de especies patógenas, intermedias y saprófitas presentes en muestras ambientales. En este contexto, las herramientas moleculares y bioinformáticas son clave para obtener información taxonómica y funcional directamente a partir del material genético recuperado del ambiente, contribuyendo a una mejor comprensión de la ecología del patógeno y de los factores que influyen en su persistencia en los ecosistemas acuáticos regionales.

La generación de conocimiento sobre las especies circulantes en los cuerpos de agua santafesinos resulta fundamental para fortalecer la vigilancia epidemiológica, identificar ambientes de mayor riesgo sanitario y aportar evidencia para el diseño de estrategias de prevención. En función de estos antecedentes, surge la siguiente pregunta de investigación: ¿qué especies del género *Leptospira* se encuentran presentes en muestras ambientales de agua dulce de la región de Santa Fe y cuál es su potencial patogénico?

OBJETIVOS

Objetivo general

- Identificar y caracterizar las especies bacterianas del género *Leptospira* presentes en muestras ambientales de cuerpos de agua dulce de la región de Santa Fe (ríos Salado, Coronda, Paraná y sus lagunas asociadas), mediante el análisis bioinformático de datos de secuenciación de ADN, con el fin de evaluar su potencial patogénico en el marco de la salud pública local.

Objetivos específicos

- Procesar y evaluar la calidad de los datos crudos de secuenciación obtenidos mediante las tecnologías Illumina (gen *16S rRNA*) y Sanger (genes *secY* y *LipL32*), realizando la inspección de electroferogramas y aplicando herramientas de control de calidad y recorte (*trimming*).
- Ensamblar las lecturas procesadas para reconstruir contigs y generar secuencias consenso de alta precisión representativas de los fragmentos amplificados.
- Determinar la identidad taxonómica de los linajes de *Leptospira* presentes en los recursos hídricos santafesinos mediante el alineamiento local y la comparación por similitud con bases de datos públicas utilizando el algoritmo BLASTn.
- Evaluar el potencial patogénico de las variantes ambientales identificadas mediante un análisis integrado de los marcadores moleculares *16S rRNA*, *secY* y *LipL32*, considerando la detección del gen *LipL32* como principal indicador molecular de patogenicidad.
- Construir árboles filogenéticos a partir de las secuencias alineadas y filtradas para inferir las relaciones evolutivas entre las muestras ambientales y cepas de referencia del género *Leptospira*.
- Delimitar clados, identificar posibles linajes patogénicos y saprófitos, y contextualizar filogenéticamente las variantes detectadas en los cuerpos de agua estudiados.

MATERIALES Y MÉTODOS

Diseño del estudio

El presente trabajo se fundamenta en un diseño de investigación descriptivo–analítico centrado en la caracterización taxonómica de bacterias del género *Leptospira* y el análisis de marcadores de patogenicidad mediante herramientas computacionales de alta resolución. El estudio se define por un abordaje exclusivamente bioinformático, utilizando como insumo primario datos de secuenciación previamente generados a partir de muestras ambientales, sin que exista una intervención experimental directa o manipulación física de los ejemplares originales durante el desarrollo de esta fase de la investigación. La integridad de las secuencias originales se mantuvo sin modificaciones manuales, optando por procesos automatizados para garantizar la objetividad y reproducibilidad de los resultados.

La arquitectura metodológica se apoya en una estrategia multi-marcador que integra los genes *16S rRNA*, *secY* y *LipL32*, permitiendo una evaluación complementaria de la biodiversidad y el riesgo sanitario en los cuerpos de agua analizados. En este contexto, un marcador molecular se define como una secuencia de ADN específica que permite inferir la identidad taxonómica o propiedades biológicas de un organismo debido a su variabilidad o conservación evolutiva (Chiani *et al.*, 2016; Caimi & Ruybal, 2020). El gen *16S rRNA* se empleó como un marcador universal para la caracterización taxonómica bacteriana global, permitiendo situar a los integrantes de la muestra dentro del marco filogenético del género. Por su parte, el gen *secY* se seleccionó por su condición de gen *housekeeping* o de mantenimiento, lo que implica que es esencial para funciones celulares básicas y presenta una tasa de evolución que otorga un poder discriminatorio superior al del *16S rRNA* para diferenciar especies estrechamente relacionadas dentro del género *Leptospira*. Finalmente, el marcador *LipL32* se utilizó específicamente para evaluar la patogenicidad latente, dado que codifica la lipoproteína de membrana externa más abundante en las especies patógenas y de patogenicidad intermedia, encontrándose ausente en las variantes saprófitas o no patógenas (Bierque *et al.*, 2020; Picardeau, 2017).

La integración tecnológica del estudio combinó la secuenciación masiva de amplicones (NGS) mediante plataforma Illumina con la secuenciación capilar por el método de Sanger, aprovechando las fortalezas de ambas para obtener un panorama detallado de las muestras. La tecnología Illumina, basada en la secuenciación por síntesis, permitió la generación de millones de lecturas cortas y pareadas (*paired-end*) del gen *16S rRNA*,

facilitando un análisis profundo de la diversidad microbiana ambiental con una alta cobertura estadística (Khan *et al.*, 2021). En contraste, la secuenciación Sanger se aplicó de forma dirigida a los genes *secY* y *LipL32* para obtener lecturas individuales de alta fidelidad y mayor longitud, fundamentales para la identificación precisa a nivel de especie y la confirmación de factores de virulencia (Deharvengt *et al.*, 2020). Esta complementariedad técnica permitió que el volumen de datos masivos de Illumina fuera validado por la especificidad y precisión de las secuencias obtenidas por Sanger.

El procesamiento de la información se estructuró a través de un *pipeline* bioinformático robusto, entendido como una sucesión automatizada de programas y algoritmos diseñados para transformar datos crudos en conocimiento taxonómico (Criscuolo & Gribaldo, 2010). Este flujo de trabajo comenzó con un riguroso control de calidad de los archivos FASTQ y electroferogramas .ab1 mediante herramientas como FastQC (Andrews, 2010) y el paquete Staden, seguido de un proceso de *trimming* automático para eliminar adaptadores y regiones de baja calidad o ruido basal en los extremos de las lecturas (Khan *et al.*, 2021). El ensamblaje de las secuencias y la generación de contigs representativos permitieron realizar búsquedas por similitud utilizando el algoritmo BLASTn contra bases de datos internacionales curadas, como las del NCBI (*National Center for Biotechnology Information*) (Khan *et al.*, 2021).

Para asegurar la fiabilidad de los hallazgos, se implementó una estrategia de validación cruzada y corroboración taxonómica. Las secuencias del gen *16S rRNA* fueron sometidas a comparaciones independientes tanto en bases de datos específicas de secuencias ribosomales como en la base de datos no redundante (nr), priorizando los niveles taxonómicos donde existiera concordancia plena entre ambos resultados. Esta corroboración se vio reforzada por el análisis filogenético de las secuencias mediante MAFFT (Katoh & Standley, 2013), lo que permitió contextualizar a las secuencias obtenidas de las muestras ambientales dentro de clados de referencia (patógenos, intermedios o saprófitos) previamente establecidos en la literatura científica. La asignación taxonómica final se realizó con la interpretación integrada de los resultados de los tres genes analizados, garantizando una resolución taxonómica robusta y una evaluación precisa del potencial patogénico de las bacterias detectadas.

Muestras ambientales y metadatos

El componente empírico de este estudio se sustenta en el análisis bioinformático de un conjunto de 76 muestras ambientales (Tabla 1 - Tabla 2 - Tabla 3) recolectadas en diversos puntos estratégicos de la cuenca hídrica de la ciudad de Santa Fe y gran Santa Fe (Figura 3). La procedencia de estos registros se circunscribe a cuerpos de agua superficiales de gran relevancia regional, incluyendo los cursos principales del río Salado, el río Coronda y el río Paraná, así como las lagunas asociadas a cada uno de estos sistemas fluviales. La selección de estos sitios respondió a la necesidad de evaluar la presencia de bacterias del género *Leptospira* en ambientes con diferentes dinámicas hidrológicas y grados de intervención antrópica, abarcando desde áreas de cauce abierto hasta cuerpos de agua lénticos donde la acumulación de sedimentos y materia orgánica podría favorecer la supervivencia de la bacteria. Todas las muestras fueron procesadas inicialmente mediante protocolos estandarizados de filtración y extracción de ADN genómico total a partir de biomasa retenida en filtros de nitrocelulosa de 0.2 μm de tamaño de poro, proporcionando el material genético necesario para las posteriores amplificaciones dirigidas.

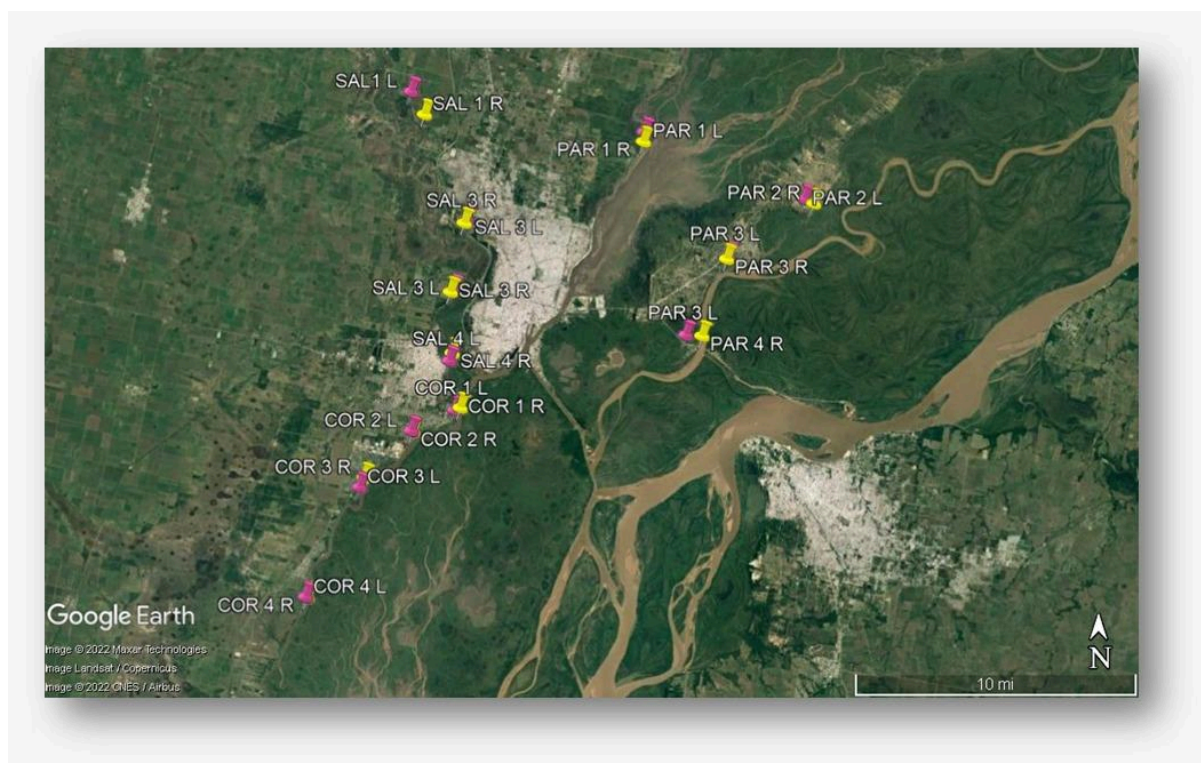


Figura 3. Localización de los sitios de muestreo ambiental en la ciudad de Santa Fe. **Nota.** Se presenta la distribución espacial de los sitios de muestreo en la ciudad de Santa Fe y áreas circundantes, incluyendo los sistemas fluviales del río Salado, río Coronda y río Paraná, así como lagunas asociadas y cuerpos de agua del

peridomicilio. Cada punto representa un sitio de muestreo, pudiendo corresponder a más de una muestra por ubicación.

ID muestra gen <i>16S rRNA</i>	Sitio de muestreo	Cuerpo de agua
1681	Salado	Río
1682	Paraná	Laguna
1683	Paraná	Río
1684	Coronda	Laguna
1685	Coronda	Laguna
1686	Paraná	Laguna
1687	Salado	Laguna
1688	Salado	Río
1689	Paraná	Río
1690	Coronda	Laguna
1691	Coronda	Río
1692	Paraná	Laguna
1721	Salado	Río
1722	Salado	Laguna
1723	Paraná	Laguna
1724	Paraná	Laguna
1725	Paraná	Laguna
1726	Paraná	Laguna
1727	Coronda	Laguna
1728	Paraná	Río
1812	Coronda	Río

Tabla 1. ID, sitio de muestreo y cuerpo de agua correspondientes a las muestras tomadas para el gen *16S rRNA*.

ID muestra gen <i>secY</i>	Sitio de muestreo	Cuerpo de agua
1	La vuelta del Paraguay	Río

3	Paraná	Río
5	Paraná	Río
6	Paraná	Río
10	Coronda	Río
10b	Coronda	Río
11	Coronda	Laguna
12	Coronda	Río
13	Coronda	Río
13b	Coronda	Río
15	Coronda	Río
17	Salado	Río
17b	Salado	Río
18	Salado	Río
18b	Salado	Río
19	Salado	Laguna
20	Salado	Río
20b	Salado	Río
21	Salado	Laguna
22	Salado	Río
23	Salado	Laguna
23PI	Salado	Laguna
26	Paraná	Río
27	Paraná	Laguna
28	Paraná	Río
28b	Paraná	Río
29	Paraná	Río
34	Coronda	Laguna
34b	Coronda	Laguna
36	Coronda	Río

36PI	Coronda	Río
37	Coronda	Laguna
38	Coronda	Río
40	Salado	Río
40b	Salado	Río
41	Salado	Río
43	Salado	Río
45	Salado	Río

Tabla 2. ID, sitio de muestreo y cuerpo de agua correspondientes a las muestras tomadas para el gen *secY*.

ID muestra gen <i>LipL32</i>	Sitio de muestreo	Cuerpo de agua
1	La vuelta del Paraguay	Río
2	La vuelta del Paraguay	Laguna
3	Paraná	Río
6	Paraná	Río
8	Paraná	Río
10	Coronda	Río
12	Coronda	Río
13	Coronda	Río
17	Salado	Río
19	Salado	Laguna
20	Salado	Río
22	Salado	Río
23	Salado	Laguna
27	Paraná	Laguna
28	Paraná	Río
34	Coronda	Laguna
41	Salado	Río

Tabla 3. ID, sitio de muestreo y cuerpo de agua correspondientes a las muestras tomadas para el gen *LipL32*.

La distribución del esfuerzo analítico se organizó de manera diferencial según los marcadores moleculares seleccionados, garantizando una cobertura metodológica complementaria para la identificación taxonómica y el análisis de patogenicidad. De las 76 muestras procesadas, un total de 21 se amplificaron con el gen del marcador universal *16S rRNA*, y fueron destinadas a la generación de bibliotecas para secuenciación masiva mediante plataforma Illumina. Por su parte, la caracterización de mayor resolución y el análisis de virulencia se centraron en 38 muestras para el gen *housekeeping secY* y 17 muestras para el marcador de patogenicidad *LipL32*, realizándose en ambos casos la secuenciación mediante el método de terminación de cadena de Sanger (Chao *et al.*, 2021). Para garantizar la integridad y el rigor científico durante todo el flujo de trabajo computacional, se implementó un sistema de identificación de muestras y trazabilidad bioinformática que vinculó de forma unívoca cada código de campo con su respectivo gen analizado y el archivo de secuenciación resultante. Este esquema permitió mantener la coherencia de los datos desde la obtención de las lecturas crudas hasta la asignación taxonómica final, evitando errores de cruce de información y facilitando la auditoría de los resultados.

Secuenciación del gen *16S rRNA* mediante Illumina

La caracterización molecular del gen *16S rRNA* se realizó mediante la secuenciación de amplicones en la plataforma Illumina, empleando una estrategia de lecturas pareadas (*paired-end*) para obtener una mayor cobertura y mejorar la resolución en el solapamiento de las secuencias (Illumina, 2013) (Figura 4). Los datos crudos se obtuvieron en formato FASTQ, un archivo de texto que almacena tanto la secuencia de nucleótidos como su correspondiente puntuación de calidad Phred. El *pipeline* bioinformático comenzó con un control de calidad exhaustivo utilizando el software FastQC (Figura 5), el cual genera reportes estadísticos sobre la distribución del contenido GC, la presencia de duplicados y, críticamente, la calidad por base a lo largo de la lectura (Khan *et al.*, 2021). La interpretación de estas métricas se basó en el umbral Q30, donde valores iguales o superiores indican una precisión en la llamada de bases del 99.9%, asegurando la confiabilidad de los datos para análisis posteriores (Pichler *et al.*, 2018).

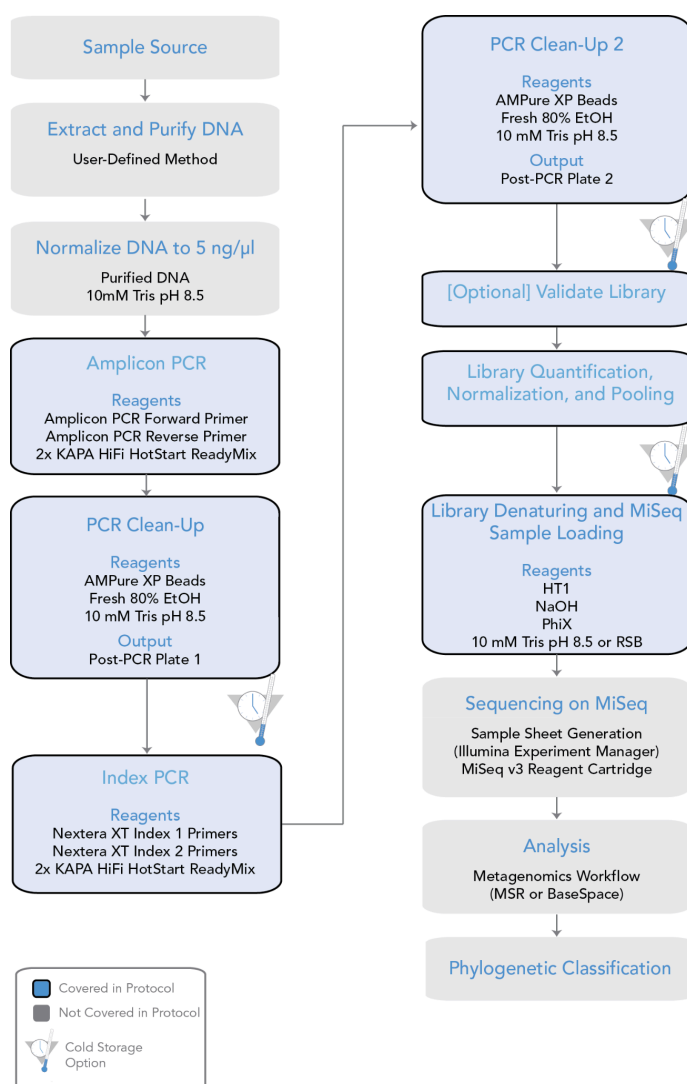


Figura 4. Flujo de trabajo para la preparación de bibliotecas de secuenciación del gen *16S rRNA* mediante el protocolo de Illumina para secuenciación de amplicones en MiSeq. Extraído de Illumina (2013).

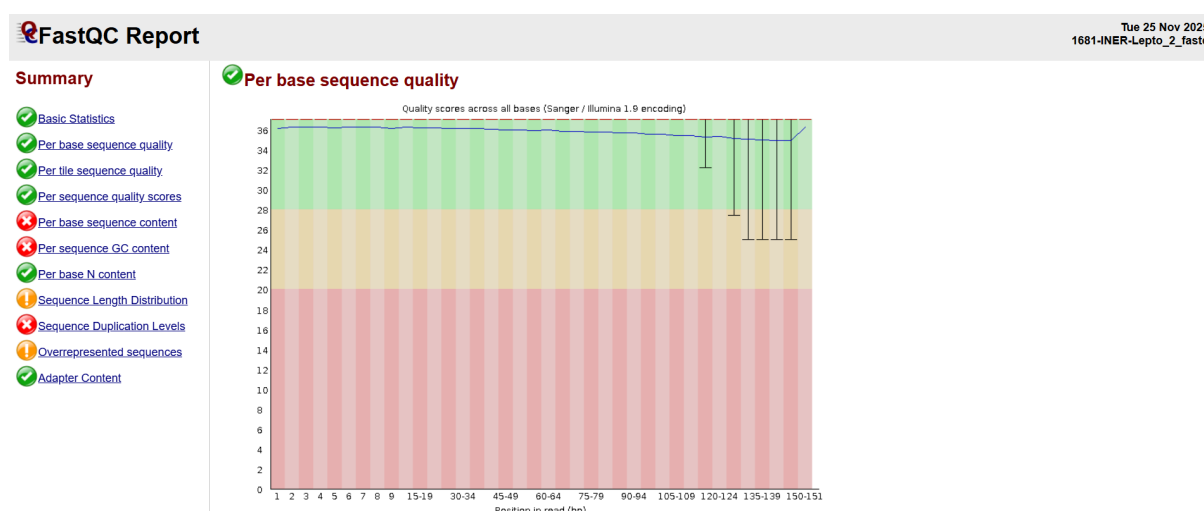


Figura 5. Reporte representativo del control de calidad de una muestra secuenciada mediante la plataforma Illumina, obtenido con el software FastQC. **Nota.** Se muestra el gráfico de calidad por base (Per base sequence

quality), donde la línea azul representa la calidad media (puntaje Phred) a lo largo de la lectura y las barras negras indican la distribución de los valores de calidad en cada posición. El fondo coloreado corresponde a los rangos de interpretación del software: verde (calidad alta), amarillo (calidad intermedia) y rojo (calidad baja). En la muestra analizada, la mayoría de las bases presentaron valores superiores a Q30 durante toda la lectura, con una leve disminución de la calidad hacia el extremo 3', característica de la secuenciación Illumina.

Tras el diagnóstico de calidad, se procedió al filtrado y *trimming* (recorte) de las secuencias mediante herramientas como TrimGalore, cuya función es remover secuencias de adaptadores y eliminar nucleótidos de los extremos 5' y 3' que presentan valores de calidad deficientes (Khan *et al.*, 2021). Durante este proceso, el ajuste de parámetros como el umbral de calidad mínima y la longitud de lectura remanente permitió descartar artefactos técnicos y lecturas excesivamente cortas que podrían generar ambigüedades en el ensamblaje (Khan *et al.*, 2021). El procesamiento de las lecturas limpias se realizó con el software SPAdes,, un ensamblador *de novo* diseñado para datos de Illumina que utiliza grafos de Bruijn y *k-mers* de distintos tamaños para reconstruir las secuencias originales en fragmentos continuos denominados contigs (Deurenberg *et al.*, 2017).

Secuenciación de *secY* y *LipL32* mediante Sanger

La caracterización molecular de los genes *secY* y *LipL32* se llevó a cabo mediante la secuenciación por el método de terminación de cadena de Sanger, empleando una estrategia de lecturas pareadas (*forward* y *reverse*) para garantizar la robustez del material genético analizado. El flujo de trabajo bioinformático comenzó con el procesamiento de los datos crudos obtenidos en formato binario ABIF (.ab1), los cuales almacenan tanto la secuencia de nucleótidos como los electroferogramas o cromatogramas generados por el secuenciador (Figura 6) (Chao *et al.*, 2021). La revisión y edición de estos archivos se realizó utilizando el software Staden Package, específicamente mediante el componente TREV, que permite una inspección visual exhaustiva de la definición de los picos y la evaluación de la relación señal-ruido (Chiani *et al.*, 2016; Staden, 1996; Staden *et al.*, 2000).

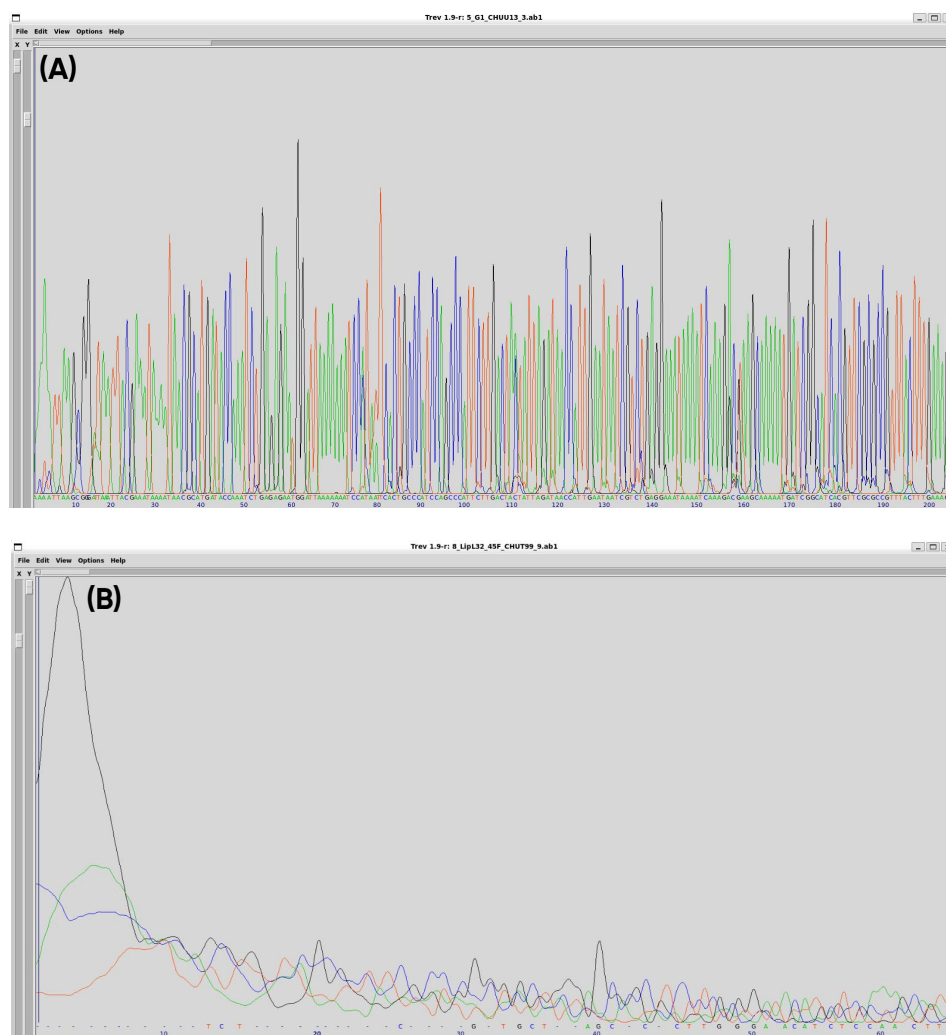


Figura 6. Electroferogramas representativos de secuenciación Sanger correspondientes a los genes *secY* y *LipL32*. **Nota.** (A) Electroferograma representativo de una muestra correspondiente al gen *secY*, que presenta picos bien definidos, adecuada resolución de señal y mínima superposición de nucleótidos, indicativo de secuencias de alta calidad. (B) Electroferograma representativo de una muestra correspondiente al gen *LipL32*, que evidencia baja calidad de señal, con picos solapados, elevado ruido de fondo y ambigüedad en la asignación de bases.

Durante esta fase, se aplicó un proceso de *trimming* o recorte de calidad para eliminar las regiones terminales de los extremos 5' y 3' que fue realizado con el programa PREGAP4, perteneciente también al Staden Package (Staden *et al.*, 2000), donde habitualmente se observa un descenso en los valores de calidad y un incremento del ruido basal característico de la tecnología Sanger. La calidad de las secuencias se evaluó en función de la nitidez de los picos y la ausencia de señales ambiguas, descartando aquellas lecturas que, tras el recorte, no conservaran una longitud suficiente o una cobertura adecuada para un análisis confiable. Con el fin de minimizar errores individuales y aumentar la precisión, las lecturas *forward* y *reverse*

de cada muestra fueron alineadas y ensambladas para obtener una secuencia consenso representativa del fragmento amplificado (Chiani *et al.*, 2016; Staden *et al.*, 2000).

El software GAP4 facilitó este ensamblaje al proporcionar herramientas para localizar discrepancias entre las hebras y verificar la consistencia de los datos antes de la generación del consenso final (Staden, 1996; Staden *et al.*, 2000). No se realizaron ediciones directas de nucleótidos ambiguos ni correcciones manuales arbitrarias, manteniendo la integridad de la información original procesada bajo parámetros estandarizados.

Identificación taxonómica mediante BLASTn

La identificación taxonómica de las secuencias obtenidas se realizó mediante búsquedas por similitud utilizando el algoritmo BLASTn (*Basic Local Alignment Search Tool*) disponible en el servidor del NCBI. Este proceso metodológico consistió en comparar las secuencias consenso de los genes *secY* y *LipL32* (Sanger), así como los *contigs* ensamblados del gen *16S rRNA* (Illumina), contra bases de datos públicas de nucleótidos para determinar la especie o el linaje más probable dentro del género *Leptospira* (Figura 7). Para el marcador *16S rRNA*, se implementó una estrategia de validación cruzada mediante dos búsquedas independientes: una contra una base de datos especializada en secuencias ribosomales (RefSeq_RNA), que ofrece mayor precisión en la delimitación taxonómica bacteriana, y otra contra la base de datos no redundante (nr) de GenBank, que integra registros generales de múltiples estudios (Pichler *et al.*, 2018).

Job Title

CHUU13.p1t CHUU13.0.1

RID

48C12BU6016 Search expires on 07-02 05:07 am [Download All](#) ▼

Program

BLASTN ? [Citation](#) ▼

Database

core_nt [See details](#) ▼

Query ID

lcl|Query_6226713

Description

CHUU13.p1t CHUU13.0.1

Molecule type

dna

Query Length

259

Other reports

[Distance tree of results](#) [MSA viewer](#) ?

Filter Results

Organism

only top 20 will appear ☐ exclude

Type common name, binomial, taxid or group name

[+ Add organism](#)

Percent Identity

to

E value

to

Query Coverage

to

Filter

Reset

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download ▼ Select columns ▼ Show 100 ▼ ?

☒ select all 100 sequences selected

[GenBank](#)
[Graphics](#)
[Distance tree of results](#)
[MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☒	Leptospira interrogans serovar Hardjo-prajitno strain Hardjoprajtino SecY (secY) gene, partial cds	Leptospira interr...	438	438	98%	2e-118	98.03%	1386	EU357983.1
☒	Uncultured Leptospira sp. clone CH-27 preprotein translocase subunit SecY (secY) gene, partial cds	uncultured Lepto...	438	438	98%	2e-118	98.03%	1083	OK632495.1
☒	Leptospira interrogans serovar Hardjo str. Norma chromosome I, complete sequence	Leptospira interr...	438	438	98%	2e-118	98.03%	4406718	CP012603.1
☒	Leptospira interrogans strain R13 chromosome 1	Leptospira interr...	438	438	98%	2e-118	98.03%	4214175	CP047506.1
☒	Leptospira interrogans serovar Naam SecY (secY) gene, partial cds	Leptospira interr...	438	438	98%	2e-118	98.03%	1304	EU357999.1

Figura 7. Interfaz de resultados de BLASTn del NCBI para la identificación taxonómica de una secuencia del gen *secY*. Se observan los alineamientos significativos obtenidos y los principales parámetros de similitud empleados para la asignación taxonómica y la selección de secuencias de referencia para los análisis posteriores.

Para la interpretación de los resultados y la asignación taxonómica, se evaluaron de manera sistemática tres parámetros bioinformáticos críticos: el porcentaje de identidad, la cobertura de la consulta (*query cover*), y el *E-value*. El porcentaje de identidad cuantifica la similitud de nucleótidos entre la muestra y la referencia, utilizándose umbrales de $\geq 97\%$ para la asignación a nivel de especie en el gen *16S rRNA* (Agudelo-Flórez *et al.*, 2014; Pichler *et al.*, 2018). La cobertura de la consulta indica la proporción de la secuencia de entrada que se alinea con la secuencia de la base de datos, asegurando que la identificación no se base en fragmentos parciales irrelevantes.

El *E-value* representa la significancia estadística del alineamiento, indicando el número de coincidencias que podrían ocurrir por azar; valores cercanos a cero sugieren una homología biológica significativa. La identificación taxonómica final se definió priorizando aquellos resultados que presentaron la mayor concordancia entre bases de datos, con alta identidad porcentual y una cobertura extensa del fragmento amplificado. En casos de discrepancia entre marcadores o bases de datos, se adoptó el nivel taxonómico común más conservador para evitar sobreinterpretaciones, integrando la información de los tres marcadores moleculares para clasificar las secuencias en los clados patógenos, intermedios o saprófitos.

Alineamiento múltiple y análisis filogenético

El análisis de las relaciones evolutivas y la corroboración taxonómica de las secuencias obtenidas se basó en la recopilación de secuencias nucleotídicas de referencia del gen 16S rRNA en formato FASTA, correspondientes a cepas tipo de especies patógenas, intermedias y saprófitas de *Leptospira* obtenidas de bases de datos públicas (Morey *et al.*, 2006). Estas referencias se integraron con las secuencias consenso y contigs generados en el presente estudio para conformar los archivos de trabajo iniciales. El alineamiento múltiple de las secuencias se llevó a cabo utilizando el software MAFFT, empleando el algoritmo automático (`--auto`) que selecciona la estrategia de alineamiento más adecuada según las características de los datos y optimizando el tiempo de cómputo mediante procesamiento multihilo (Kato & Standley, 2013). Una vez obtenido el alineamiento, se procedió a su inspección manual con la herramienta AliView (Larsson, 2014) (Figura 8), con el fin de verificar la consistencia de las columnas y realizar el *trimming* o recorte de las regiones terminales. Específicamente, se eliminaron las posiciones correspondientes a los extremos 5' y 3' debido a la presencia frecuente de gaps y nucleótidos ambiguos, asegurando que solo las regiones de alta calidad y con información filogenética confiable fueran incluidas en la reconstrucción de los árboles.



Figura 8. Inspección manual del alineamiento múltiple de secuencias del gen 16S rRNA utilizando el software AliView. Se muestra el alineamiento de las secuencias correspondientes a las secuencias analizadas junto con secuencias de referencia del género *Leptospira*.

La reconstrucción filogenética se realizó mediante dos enfoques complementarios con el fin de evaluar la consistencia de las relaciones evolutivas inferidas. En primer lugar, se construyeron árboles utilizando el método de Neighbor-Joining (NJ) mediante el paquete PHYLIP. Para ello, los alineamientos fueron convertidos al formato compatible utilizando

herramientas de EMBOSS, posteriormente se calcularon las matrices de distancia genética con el programa dnadist, empleando el modelo de sustitución de Kimura de dos parámetros (K2P), y finalmente se generó la topología del árbol con el programa neighbor (Agudelo-Flórez et al., 2014). De manera complementaria, se realizaron análisis mediante el método de Maximum Likelihood (ML) utilizando IQ-TREE. Previamente a la reconstrucción del árbol, se empleó la función ModelFinder Plus (-m MFP) para identificar el modelo evolutivo que mejor se ajustaba a cada conjunto de datos. El árbol filogenético fue posteriormente inferido por IQ-TREE utilizando el modelo seleccionado, mediante un enfoque de máxima verosimilitud. Mientras que el método Neighbor-Joining reconstruye las relaciones filogenéticas a partir de una matriz de distancias genéticas entre las secuencias, Maximum Likelihood estima el árbol que maximiza la probabilidad de observar los datos bajo un modelo explícito de evolución molecular, proporcionando una inferencia filogenética basada en un marco estadístico.

La confiabilidad de las ramas y la consistencia de los clados obtenidos se validaron mediante el análisis de *bootstrap* con 1000 réplicas (Agudelo-Flórez et al., 2014). El *bootstrap* es un método de remuestreo estadístico donde valores cercanos a 100 indican que el agrupamiento de estos taxones es altamente robusto y se mantiene independientemente de las variaciones en el conjunto de datos analizado (Agudelo-Flórez et al., 2014). En el caso del análisis mediante Maximum Likelihood, se incorporó además el soporte SH-aLRT (*Shimodaira-Hasegawa approximate Likelihood-Ratio Test*) como medida adicional de confianza para evaluar la robustez de las ramas del árbol filogenético. El SH-aLRT proporciona una estimación estadística del soporte de cada nodo basada en una aproximación de la prueba de razón de verosimilitud y se encuentra implementado en IQ-TREE como una alternativa rápida y eficiente para la evaluación del soporte de ramas (Minh et al., 2020). Finalmente, la visualización, edición y anotación de los árboles filogenéticos resultantes se realizó en la plataforma web iTOL (*Interactive Tree Of Life*) (Letunic & Bork, 2024), lo que permitió ajustar gráficamente la representación, incorporar etiquetas de identificación y diferenciar claramente los clados patógenos, intermedios y saprófitos para su posterior interpretación.

RESULTADOS

Resultados del control de calidad

La evaluación de la calidad de los datos crudos mediante FastQC demostró que las secuencias del gen *16S rRNA* obtenidas por Illumina poseían una calidad Phred sobresaliente, con la gran mayoría de las bases situándose por encima del umbral Q30, lo cual asegura una precisión en la llamada de bases (*base calling*) superior al 99.9% (Khan *et al.*, 2021; Pichler *et al.*, 2018). Los gráficos de distribución de calidad por posición evidenciaron perfiles de alta fidelidad en los primeros 200 nucleótidos de los reads *forward* y *reverse*, observándose el descenso gradual de calidad hacia el extremo 3' que es característico de la química de secuenciación por síntesis (Khan *et al.*, 2021). El contenido GC se mantuvo dentro del rango esperado para el género *Leptospira*, que oscila entre el 35% y 41%, validando la consistencia biológica de las muestras ambientales analizadas (Adler & de la Peña Moctezuma, 2010). Aunque inicialmente se detectó la presencia de adaptadores y bases de baja confianza en las regiones terminales, el proceso de *trimming* y filtrado automatizado eliminó estos artefactos de manera efectiva. Como resultado de esta limpieza, las lecturas finales mantuvieron una longitud óptima que garantizó el solapamiento necesario para el ensamblaje de contigs robustos y de alta cobertura.

En cuanto a la secuenciación Sanger de los genes *secY*, la inspección de los electroferogramas confirmó una elevada resolución de los picos y una relación señal-ruido satisfactoria en la parte central de las lecturas. Las métricas de calidad revelaron un incremento de ruido basal hacia los extremos 5' y 3', lo que hizo necesario un recorte automático para excluir las zonas de baja confiabilidad y retener exclusivamente la región informativa. Tras el procesamiento, el gen *secY* conservó fragmentos útiles para la identificación en la totalidad de las muestras, mientras que el gen *LipL32* mostró una sensibilidad superior al ruido técnico, resultando en recortes más profundos que comprometieron la longitud efectiva en determinadas unidades.

Resultados del ensamblaje

El procesamiento de los datos de secuenciación masiva mediante el ensamblador SPAdes permitió la reconstrucción de contigs de alta fidelidad para las 21 muestras ambientales destinadas al marcador *16S rRNA*. Los resultados del ensamblaje arrojaron fragmentos con una longitud máxima de 2002 pb y una longitud mínima de 1144 pb para las

secuencias identificadas como pertenecientes al género *Leptospira*. El tamaño total del ensamblaje por muestra se concentró predominantemente en contigs únicos o mayoritarios que abarcaron la región casi completa del gen ribosomal, presentando valores de N50 consistentes con la longitud total de los fragmentos recuperados, lo que indica una baja fragmentación de las secuencias de interés. La cobertura obtenida para estos ensamblajes, correspondiente a la cobertura de *k-mers* reportada por SPAdes para el mayor valor de *k* utilizado durante el ensamblado, mostró una alta variabilidad entre muestras, registrándose valores sobresalientes de hasta 9417.8X y mínimos técnicos funcionales de 13.7X.

En cuanto a la secuenciación Sanger de los genes *secY* y *LipL32*, el ensamblaje de lecturas *forward* y *reverse* mediante GAP4 permitió obtener secuencias consenso robustas. Para el gen *secY*, se consolidaron fragmentos con longitudes situadas mayoritariamente entre los 200 y 300 pb. Las métricas obtenidas reflejan una calidad general del ensamblaje adecuada, caracterizada por la obtención de fragmentos contiguos que superan el umbral de información necesaria para la discriminación taxonómica. Los valores de longitud y cobertura indican que el *pipeline* bioinformático logró mitigar el ruido basal y los artefactos de secuenciación resultando en un dataset útil para las comparaciones mediante BLASTn y la posterior reconstrucción de las relaciones evolutivas. Las posibles limitaciones observadas se restringieron a muestras puntuales con coberturas menores o fragmentaciones parciales que, no obstante, conservaron la región informativa suficiente para su inclusión en los análisis de patogenicidad y diversidad.

Identificación taxonómica basada en 16S rRNA

La identificación taxonómica de las secuencias del gen *16S rRNA* mediante búsquedas por similitud con el algoritmo BLASTn permitió clasificar la mayoría de las muestras analizadas dentro del género *Leptospira*, con una marcada predominancia de linajes pertenecientes al clado de las especies saprófitas. Los resultados obtenidos mostraron métricas de alineamiento robustas, con porcentajes de identidad nucleotídica que se situaron mayoritariamente entre el 98.43% y el 100%, lo cual supera el umbral estándar del 97% aceptado para la asignación a nivel de especie (Bilung *et al.*, 2026). La cobertura de la consulta (*query cover*) fue igualmente elevada, oscilando predominantemente entre el 92% y el 99%, mientras que los valores de significancia estadística (*E-value*) fueron de 0.0 en la

totalidad de las coincidencias vinculadas al género, confirmando la validez biológica de los alineamientos hallados.

Las especies que presentaron la mayor similitud y recurrencia en el conjunto de datos fueron *Leptospira iowaensis*, *Leptospira mgodei* y *Leptospira abararensis*, todas clasificadas como saprófitas (Bilung *et al.*, 2026). Por ejemplo, la muestra 1682 alcanzó una identidad del 100% con *Leptospira iowaensis*, mientras que la serie de muestras 1688 a 1691 mostró una notable consistencia al compartir de manera sistemática a *Leptospira iowaensis* y *Leptospira abararensis* como sus principales coincidencias con identidades superiores al 99.5%. No obstante, el análisis permitió discriminar taxones no pertenecientes al género de interés, detectándose muestras con alta afinidad hacia los géneros *Lacunisphaera* (muestras 1686 y 1722), *Reyranella* (1728) y *Phenylobacterium* (1721 y 1812), lo que evidencia la diversidad microbiana ambiental capturada durante el muestreo (Tabla 4).

ID muestra gen 16S rRNA	Especie	Patogenicidad	Top 3 % identidad nt	Top 3 Query cover
1681	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.80% <i>L. mgodei</i> 99.32% <i>L. abararensis</i> 99.25%	<i>L. iowaensis</i> 95% <i>L. mgodei</i> 95% <i>L. abararensis</i> 95%
1682-a	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 100% <i>L. abararensis</i> 99.20% <i>L. mgodei</i> 99.13%	<i>L. iowaensis</i> 92% <i>L. abararensis</i> 92% <i>L. mgodei</i> 92%
1682-b	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.84% <i>L. mgodei</i> 99.68% <i>L. abararensis</i> 99.68%	<i>L. iowaensis</i> 98% <i>L. mgodei</i> 98% <i>L. abararensis</i> 98%
1683-a	<i>L. mgodei</i>	Saprófita	<i>L. mgodei</i> 99.80% <i>L. abararensis</i> 99.80% <i>L. iowaensis</i> 99.60%	<i>L. mgodei</i> 99% <i>L. abararensis</i> 99% <i>L. iowaensis</i> 99%
1683-b	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.46% <i>L. abararensis</i> 99.29% <i>L. mgodei</i> 99.11%	<i>L. iowaensis</i> 98% <i>L. abararensis</i> 98% <i>L. mgodei</i> 98%
1684	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.66% <i>L. abararensis</i> 99.52% <i>L. mgodei</i> 99.45%	<i>L. iowaensis</i> 95% <i>L. abararensis</i> 95% <i>L. mgodei</i> 95%
1685	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.59%	<i>L. iowaensis</i> 97%

			<i>L. abararensis</i> 99.45% <i>L. mgodei</i> 99.38%	<i>L. abararensis</i> 97% <i>L. mgodei</i> 97%
1686-b	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.59% <i>L. mgodei</i> 99.52% <i>L. abararensis</i> 99.45%	<i>L. iowaensis</i> 97% <i>L. mgodei</i> 97% <i>L. abararensis</i> 97%
1687	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.66% <i>L. abararensis</i> 99.52% <i>L. mgodei</i> 99.45%	<i>L. iowaensis</i> 99% <i>L. abararensis</i> 99% <i>L. mgodei</i> 99%
1688	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.73% <i>L. abararensis</i> 99.59% <i>L. mgodei</i> 99.52%	<i>L. iowaensis</i> 93% <i>L. abararensis</i> 93% <i>L. mgodei</i> 93%
1689	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.73% <i>L. abararensis</i> 99.59% <i>L. mgodei</i> 99.52%	<i>L. iowaensis</i> 95% <i>L. abararensis</i> 95% <i>L. mgodei</i> 95%
1690	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.73% <i>L. abararensis</i> 99.59% <i>L. mgodei</i> 99.52%	<i>L. iowaensis</i> 98% <i>L. abararensis</i> 98% <i>L. mgodei</i> 98%
1691	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.66% <i>L. abararensis</i> 99.52% <i>L. mgodei</i> 99.45%	<i>L. iowaensis</i> 93% <i>L. abararensis</i> 93% <i>L. mgodei</i> 93%
1692	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.93% <i>L. abararensis</i> 99.39% <i>L. mgodei</i> 99.32%	<i>L. iowaensis</i> 94% <i>L. abararensis</i> 94% <i>L. mgodei</i> 94%
1721	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 98.43% <i>L. mgodei</i> 98.23% <i>L. abararensis</i> 98.16%	<i>L. iowaensis</i> 97% <i>L. mgodei</i> 97% <i>L. abararensis</i> 97%
1722-b	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.73% <i>L. abararensis</i> 99.59% <i>L. mgodei</i> 99.52%	<i>L. iowaensis</i> 97% <i>L. abararensis</i> 97% <i>L. mgodei</i> 97%
1723	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.73% <i>L. abararensis</i> 99.59% <i>L. mgodei</i> 99.52%	<i>L. iowaensis</i> 97% <i>L. abararensis</i> 97% <i>L. mgodei</i> 97%
1724	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.73% <i>L. abararensis</i> 99.59% <i>L. mgodei</i> 99.52%	<i>L. iowaensis</i> 98% <i>L. abararensis</i> 98% <i>L. mgodei</i> 98%
1725	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.66% <i>L. abararensis</i> 99.52% <i>L. mgodei</i> 99.45%	<i>L. iowaensis</i> 99% <i>L. abararensis</i> 99% <i>L. mgodei</i> 99%
1726	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.66% <i>L. mgodei</i> 99.59% <i>L. abararensis</i> 99.52%	<i>L. iowaensis</i> 97% <i>L. mgodei</i> 97% <i>L. abararensis</i> 97%
1727	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.73% <i>L. abararensis</i> 99.59% <i>L. mgodei</i> 99.52%	<i>L. iowaensis</i> 98% <i>L. abararensis</i> 98% <i>L. mgodei</i> 98%

1812-a	<i>L. iowaensis</i>	Saprófita	<i>L. iowaensis</i> 99.93% <i>L. abararensis</i> 99.39% <i>L. mgodei</i> 99.32%	<i>L. iowaensis</i> 97% <i>L. abararensis</i> 97% <i>L. mgodei</i> 97%
--------	---------------------	-----------	---	--

Tabla 4. Resultados de la identificación taxonómica de las secuencias del gen 16S rRNA mediante BLASTn. Para cada muestra se indica la especie de *Leptospira* con la mejor coincidencia en la base de datos consultada y su correspondiente clasificación en el grupo de patogenicidad (saprófita, intermedia o patógena), además de los top 3 hits con sus porcentajes de identidad y cobertura correspondientes.

Los elevados porcentajes de identidad observados indican una asignación confiable a nivel de género y permiten la delimitación de los grandes clados de patogenicidad, pero el nivel de resolución taxonómica alcanzado encontró limitaciones críticas para la discriminación precisa entre especies estrechamente relacionadas. Se identificaron numerosos casos ambiguos donde los mejores resultados (*top 3 hits*) presentaban porcentajes de identidad idénticos o separados por diferencias de apenas pocos nucleótidos, como ocurrió en la muestra 1683-a, que arrojó simultáneamente un 99.80% de identidad para *Leptospira mgodei* y *Leptospira abararensis*.

Identificación taxonómica basada en secY

La identificación taxonómica de las secuencias del gen *secY* mediante el algoritmo BLASTn permitió una caracterización detallada de las muestras, logrando una resolución significativamente superior a la obtenida con el marcador 16S rRNA. Los resultados revelaron una marcada presencia de especies pertenecientes al clado de las leptospiros patógenas, destacándose *Leptospira interrogans* como la especie con mayor similitud y frecuencia en muestras como los registros 5, 13, 19 y 28b. Los porcentajes de identidad nucleotídica para estas identificaciones fueron sobresalientes, alcanzando el 100% en muestras como 1 y 19. La cobertura de la consulta (*query cover*) se mantuvo predominantemente elevada, situándose entre el 81% y el 100%, mientras que los *E-values* obtenidos (en el rango de $1e^{-110}$ a $1e^{-143}$) confirmaron la alta significancia estadística de los alineamientos (Tabla 5).

ID muestra gen <i>secY</i>	Especie	Patogenicidad	Top 3 % identidad nt	Top 3 Query cover
1-a	<i>L. borgpetersenii</i>	Patógena	<i>L. borgpetersenii</i> str. 4E 92.91% <i>L. borgpetersenii</i> serovar Hardjo str. BK-30 93.39% <i>L. borgpetersenii</i> serovar Hardjo str. BK-6 93.39%	<i>L. borgpetersenii</i> str. 4E 99% <i>L. borgpetersenii</i> serovar Hardjo str. BK-30 95% <i>L. borgpetersenii</i> serovar Hardjo str. BK-6 95%
1-b	<i>L. borgpetersenii</i>	Patógena	<i>L. borgpetersenii</i> serovar	<i>L. borgpetersenii</i> serovar

			Hardjo str. BK-6 93.62% <i>L. borgpetersenii</i> serovar Hardjo str. NVSL S 1343 93.62% <i>L. borgpetersenii</i> serovar Hardjo str. NVSL S 818 93.62%	Hardjo str. BK-6 99% <i>L. borgpetersenii</i> serovar Hardjo str. NVSL S 1343 99% <i>L. borgpetersenii</i> serovar Hardjo str. NVSL S 818 99%
5	<i>L. interrogans</i>	Patógena	<i>L. interrogans</i> isolate 114 98.03% <i>L. interrogans</i> isolate 64 98.03% <i>L. interrogans</i> serovar Manilae str. LT 398 98.03%	<i>L. interrogans</i> isolate 114 98% <i>L. interrogans</i> isolate 64 98% <i>L. interrogans</i> serovar Manilae str. LT 398 98%
10	<i>L. borgpetersenii</i>	Patógena	<i>Leptospira</i> sp. str. EC14 92.63% <i>L. borgpetersenii</i> serovar Nyanza str. Kibos 91.58% <i>L. borgpetersenii</i> serovar Mini str. Sari 91.58%	<i>Leptospira</i> sp. str. EC14 99% <i>L. borgpetersenii</i> serovar Nyanza str. Kibos 99% <i>L. borgpetersenii</i> serovar Mini str. Sari 99%
10b	<i>L. borgpetersenii</i>	Patógena	<i>Leptospira</i> sp. str. EC14 92.31% <i>L. borgpetersenii</i> serovar Nyanza str. Kibos 91.26% <i>L. borgpetersenii</i> serovar Mini str. Sari 91.26%	<i>Leptospira</i> sp. str. EC14 99% <i>L. borgpetersenii</i> serovar Nyanza str. Kibos 99% <i>L. borgpetersenii</i> serovar Mini str. Sari 99%
11	<i>L. borgpetersenii</i>	Patógena	<i>L. borgpetersenii</i> str. 4E 81.15% <i>L. borgpetersenii</i> serovar Hardjo str. BK-30 80.74% <i>L. borgpetersenii</i> serovar Hardjo str. BK-6 80.74%	<i>L. borgpetersenii</i> str. 4E 78% <i>L. borgpetersenii</i> serovar Hardjo str. BK-30 78% <i>L. borgpetersenii</i> serovar Hardjo str. BK-6 78%
12	<i>L. tipperaryensis</i>	Patógena	<i>L. tipperaryensis</i> str. GWTS #1 93.68% <i>L. knetyi</i> str. LS 001/16 91.58% <i>L. gorisiae</i> str. WS92.C1 93.26%	<i>L. tipperaryensis</i> str. GWTS #1 71% <i>L. knetyi</i> str. LS 001/16 71% <i>L. gorisiae</i> str. WS92.C1 71%
13	<i>L. interrogans</i>	Patógena	<i>L. interrogans</i> serovar Manilae str. LT 398 99.65% <i>L. interrogans</i> serovar Haemolytica str. Marsh 99.65% <i>L. interrogans</i> serovar Birkin str. Birkin 99.65%	<i>L. interrogans</i> serovar Manilae str. LT 398 98% <i>L. interrogans</i> serovar Haemolytica str. Marsh 98% <i>L. interrogans</i> serovar Birkin str. Birkin 98%
13b-a	<i>L. interrogans</i>	Patógena	<i>L. interrogans</i> isolate 114 98.77% <i>L. interrogans</i> isolate 64 98.77% <i>L. interrogans</i> serovar Haemolytica str. Marsh 98.77%	<i>L. interrogans</i> isolate 114 98% <i>L. interrogans</i> isolate 64 98% <i>L. interrogans</i> serovar Haemolytica str. Marsh 98%

13b-b	<i>L. interrogans</i>	Patógena	<i>L. interrogans</i> isolate 1118 98.35% <i>L. interrogans</i> serovar <i>Manilae</i> str. LT 398 98.35% <i>L. interrogans</i> serovar <i>Haemolytica</i> str. Marsh 98.35%	<i>L. interrogans</i> isolate 1118 95% <i>L. interrogans</i> serovar <i>Manilae</i> str. LT 398 95% <i>L. interrogans</i> serovar <i>Haemolytica</i> str. Marsh 95%
15-b	<i>L. cinconiae</i>	Patógena	<i>L. cinconiae</i> str. WS58.C1 95.42% Uncultured <i>Leptospira</i> sp. clone S2 95.42% Uncultured <i>Leptospira</i> sp. clone S5 97.39%	<i>L. cinconiae</i> str. WS58.C1 96% Uncultured <i>Leptospira</i> sp. clone S2 96% Uncultured <i>Leptospira</i> sp. clone S5 85%
17-b	<i>L. interrogans</i>	Patógena	<i>L. interrogans</i> serovar Fugis str. Fudge 78.74% <i>L. interrogans</i> serovar <i>Autumnalis</i> str. Akiyami A 78.74% <i>L. interrogans</i> serovar <i>Djasiman</i> 78.74%	<i>L. interrogans</i> serovar Fugis str. Fudge 97% <i>L. interrogans</i> serovar <i>Autumnalis</i> str. Akiyami A 97% <i>L. interrogans</i> serovar <i>Djasiman</i> 97%
18	<i>L. borgpetersenii</i>	Patógena	<i>L. borgpetersenii</i> str. MN900 chromosome 1, complete sequence 91.78% <i>L. borgpetersenii</i> serovar <i>Hardjo</i> strain BK-30 chromosome 1, complete sequence 91.78% <i>L. noguchii</i> str. IP1705032 chromosome 1, complete sequence 91.78%	<i>L. borgpetersenii</i> str. MN900 chromosome 1, complete sequence 35% <i>L. borgpetersenii</i> serovar <i>Hardjo</i> strain BK-30 chromosome 1, complete sequence 35% <i>L. noguchii</i> str. IP1705032 chromosome 1, complete sequence 35%
18b	<i>Leptospira</i> sp.	Indeterminado	<i>Leptospira</i> sp. str. EC18 78.21% <i>Leptospira</i> sp. isolate SMG19185 80.29% <i>L. interrogans</i> serovar <i>Manilae</i> isolate POM23 76.26%	<i>Leptospira</i> sp. str. EC18 90% <i>Leptospira</i> sp. isolate SMG19185 68% <i>L. interrogans</i> serovar <i>Manilae</i> isolate POM23 69%
19	<i>L. interrogans</i>	Patógena	<i>L. interrogans</i> serovar <i>Haemolytica</i> str. Marsh 100% <i>L. interrogans</i> serovar <i>Birkini</i> str. Birkin 100% <i>L. interrogans</i> isolate B16 100%	<i>L. interrogans</i> serovar <i>Haemolytica</i> str. Marsh 100% <i>L. interrogans</i> serovar <i>Birkini</i> str. Birkin 100% <i>L. interrogans</i> isolate B16 100%
20	<i>L. cinconiae</i>	Patógena	<i>L. cinconiae</i> str. WS58.C1 98.36% <i>L. knetyi</i> str. LS 001/16 82.45% <i>L. gorisiae</i> str. WS92.C1 82.11%	<i>L. cinconiae</i> str. WS58.C1 97% <i>L. knetyi</i> str. LS 001/16 97% <i>L. gorisiae</i> str. WS92.C1 97%
21	<i>L. cinconiae</i>	Patógena	<i>L. cinconiae</i> str. WS58.C1	<i>L. cinconiae</i> str.

			98.90% <i>L. kmetyi</i> str. LS 001/16 85.56% <i>L. gorisiae</i> str. WS92.C1 85.00%	WS58.C1 97% <i>L. kmetyi</i> str. LS 001/16 97% <i>L. gorisiae</i> str. WS92.C1 97%
22	<i>L. santarosai</i>	Patógena	<i>L. santarosai</i> str. ah2 94.30% <i>L. santarosai</i> str. U164 94.30% <i>L. santarosai</i> str. U160 94.30%	<i>L. santarosai</i> str. ah2 90% <i>L. santarosai</i> str. U164 90% <i>L. santarosai</i> str. U160 90%
23	<i>Leptospira</i> sp.	Indeterminado	<i>Leptospira</i> sp. str. EC7 89.17% <i>Leptospira</i> sp. str. EC9 87.50% <i>Leptospira</i> sp. str. EC19 87.50%	<i>Leptospira</i> sp. str. EC7 98% <i>Leptospira</i> sp. str. EC9 98% <i>Leptospira</i> sp. str. EC19 98%
23PI	<i>L. borgpetersenii</i>	Patógena	<i>L. borgpetersenii</i> str. MN900 chromosome 1, complete sequence 97.37% <i>L. borgpetersenii</i> serovar Hardjo str. BK-30 97.37% <i>L. noguchii</i> str. IP1705032 chromosome 1, complete sequence 97.37%	<i>L. borgpetersenii</i> str. MN900 chromosome 1, complete sequence 35% <i>L. borgpetersenii</i> serovar Hardjo str. BK-30 35% <i>L. noguchii</i> str. IP1705032 chromosome 1, complete sequence 35%
26	<i>L. cinconiae</i>	Patógena	<i>L. cinconiae</i> str. WS58.C1 97.54% <i>L. kmetyi</i> str. LS 001/16 81.63% <i>L. gorisiae</i> str. WS92.C1 81.71%	<i>L. cinconiae</i> str. WS58.C1 97% <i>L. kmetyi</i> str. LS 001/16 97% <i>L. gorisiae</i> str. WS92.C1 97%
28-a	<i>L. cinconiae</i>	Patógena	<i>L. cinconiae</i> str. WS58.C1 97.54%	<i>L. cinconiae</i> str. WS58.C1 95%
28-b	<i>L. cinconiae</i>	Patógena	<i>L. cinconiae</i> str. WS58.C1 97.45% <i>L. kmetyi</i> str. LS 001/16 86.03% <i>L. noguchii</i> str. IP1705032 85.47%	<i>L. cinconiae</i> str. WS58.C1 98% <i>L. kmetyi</i> str. LS 001/16 90% <i>L. noguchii</i> str. IP1705032 90%
28b	<i>L. interrogans</i>	Patógena	<i>L. interrogans</i> isolate 1118 99.59% <i>L. interrogans</i> serovar Manilae str. LT 398 99.59% <i>L. interrogans</i> serovar Haemolytica str. Marsh 99.59%	<i>L. interrogans</i> isolate 1118 100% <i>L. interrogans</i> serovar Manilae str. LT 398 100% <i>L. interrogans</i> serovar Haemolytica str. Marsh 100%
36	<i>L. santarosai</i>	Patógena	<i>Leptospira</i> sp. str. EC14 92.34%	<i>Leptospira</i> sp. str. EC14 97%

			<i>L. santarosai</i> serovar Canalzonae str. CZ 188 89.92% <i>L. santarosai</i> serovar <i>Brasiliensis</i> str. An 776 89.52%	<i>L. santarosai</i> serovar Canalzonae str. CZ 188 97% <i>L. santarosai</i> serovar <i>Brasiliensis</i> str. An 776 97%
36PI	<i>L. noguchii</i>	Patógena	<i>L. noguchii</i> str. barbudensis 97.60% <i>L. noguchii</i> str. IP1705032 97.52% <i>L. noguchii</i> str. IP1605021 97.52%	<i>L. noguchii</i> str. barbudensis 96% <i>L. noguchii</i> str. IP1705032 96% <i>L. noguchii</i> str. IP1605021 96%
38	<i>L. cinconiae</i>	Patógena	<i>L. cinconiae</i> str. WS58.C1 97.91% <i>L. kmetyi</i> str. LS 001/16 84.92% <i>L. noguchii</i> str. IP1705032 94.36%	<i>L. cinconiae</i> str. WS58.C1 97% <i>L. kmetyi</i> str. LS 001/16 91% <i>L. noguchii</i> str. IP1705032 91%
40b-b	<i>L. borgpetersenii</i>	Patógena	<i>Leptospira</i> sp. str. EC14 92.59% <i>L. borgpetersenii</i> serovar Mini str. Sari 89.47% <i>L. borgpetersenii</i> str. 411 89.71%	<i>Leptospira</i> sp. str. EC14 96% <i>L. borgpetersenii</i> serovar Mini str. Sari 98% <i>L. borgpetersenii</i> str. 411 96%
41	<i>L. tipperaryensis</i>	Patógena	<i>L. tipperaryensis</i> str GWTS #1 97.49% <i>L. santarosai</i> str. ah2 95.40% <i>L. kmetyi</i> str. LS 001/16 95.40%	<i>L. tipperaryensis</i> str GWTS #1 97% <i>L. santarosai</i> str. ah2 97% <i>L. kmetyi</i> str. LS 001/16 97%
43-a	<i>L. cinconiae</i>	Patógena	<i>L. cinconiae</i> str. WS58.C1 94.90%	<i>L. cinconiae</i> str. WS58.C1 80%
43-b	<i>L. cinconiae</i>	Patógena	<i>L. cinconiae</i> str. WS58.C1 96.43% <i>L. kmetyi</i> strain LS 001/16 84.92% <i>L. noguchii</i> str. IP1705032 84.36%	<i>L. cinconiae</i> str. WS58.C1 88% <i>L. kmetyi</i> strain LS 001/16 80% <i>L. noguchii</i> str. IP1705032 80%
45-a	<i>L. cinconiae</i>	Patógena	<i>L. cinconiae</i> str. WS58.C1 98.51%	<i>L. cinconiae</i> str. WS58.C1 95%
45-b	<i>L. cinconiae</i>	Patógena	<i>L. cinconiae</i> str. WS58.C1 85.47% <i>L. kmetyi</i> str. LS 001/16 85.47% <i>L. noguchii</i> str. IP1705032 84.92%	<i>L. cinconiae</i> str. WS58.C1 99% <i>L. kmetyi</i> str. LS 001/16 89% <i>L. noguchii</i> str. IP1705032 89%

Tabla 5. Resultados de la identificación taxonómica de las secuencias del gen *secY* mediante BLASTn. Para cada muestra se presenta el ID, la especie de *Leptospira* con la mejor coincidencia obtenida en la base de datos

consultada y su clasificación según el grupo de patogenicidad, además de los top 3 *hits* con sus porcentajes de identidad y cobertura correspondientes. Las muestras asignadas a *Leptospira* sp. corresponden a secuencias cuya resolución taxonómica no permitió la identificación a nivel de especie.

Además de *Leptospira interrogans*, el marcador secY permitió identificar otros linajes patógenos y de relevancia ambiental. Se detectó la presencia de *Leptospira borgpetersenii* en muestras como 1, 11 y 10b, con identidades que oscilaron entre el 81% y el 93%. Asimismo, se identificaron secuencias vinculadas a *Leptospira santarosai* (22 y 36) y *Leptospira noguchii* (36PI), reforzando la capacidad de este gen para detectar variantes de alta virulencia en matrices ambientales. Un hallazgo relevante fue la recurrencia de *Leptospira cinconiae* en una serie extensa de muestras (15-b, 20, 21, 22, 26, 28-a, 28-b, 38, 43-b, 45-a y 45-b), con identidades que alcanzaron hasta el 99%.

Resultados del análisis de LipL32

En lo que respecta al marcador molecular *LipL32*, los resultados indicaron que las secuencias obtenidas no pudieron ser integradas en el análisis taxonómico ni en la evaluación de patogenicidad posterior. Esta exclusión derivó de problemas críticos de calidad en las lecturas generadas mediante el método de secuenciación Sanger, las cuales presentaron electroferogramas con una relación señal-ruido deficiente y picos de intensidad poco definidos hacia los extremos de los fragmentos.

Como consecuencia de esta degradación de la señal, fue imposible realizar un ensamblaje confiable de las lecturas *forward* y *reverse*, impidiendo la generación de secuencias consenso robustas que permitieran ejecutar alineamientos estadísticamente significativos con las bases de datos de referencia.

Debido a que la reducción de la longitud alineable comprometía la precisión de la identificación por similitud, se determinó que la totalidad de las muestras destinadas a este marcador no cumplían con los estándares mínimos de confiabilidad. Por lo tanto, el gen *LipL32* no fue analizado en la fase final, explicitándose que su descarte constituye una limitación metodológica ligada a la calidad del material genético recuperado y no debe interpretarse como una ausencia biológica del gen en las muestras ambientales originales.

Resultados de BLASTn con asignaciones taxonómicas distintas al género *Leptospira*

La caracterización taxonómica de las secuencias obtenidas mediante búsquedas por similitud con el algoritmo BLASTn reveló la presencia de organismos no pertenecientes al género *Leptospira* en diversas muestras ambientales analizadas. En el análisis de los contigs del gen *16S rRNA*, se identificaron varios taxones bacterianos típicos de entornos hídricos y suelos que fueron discriminados durante el procesamiento bioinformático. Las muestras 1686 y 1722 presentaron como mejor coincidencia al género *Lacunisphaera*, registrándose identidades nucleotídicas del 98.86% para *Lacunisphaera anatis* y 95.48% para *Lacunisphaera limnophila*, respectivamente, con coberturas de consulta iguales o superiores al 97%. Asimismo, en la muestra 1723, se obtuvo una coincidencia del 99.24% de identidad y 99% de cobertura con *Cytobacillus kochii*. En la muestra 1728, se recuperaron dos secuencias con afinidades taxonómicas distintas: una vinculada a *Reyranella massiliensis* con un 98.41% de identidad y otra a *Phenylobacterium conjunctum* con un 99.93% de identidad. Esta última especie, *Phenylobacterium conjunctum*, también fue identificada en la muestra 1812 con una identidad del 99.65% y una cobertura del 97%. En todos los casos vinculados al marcador ribosomal, los E-valores registrados fueron de 0.0, reflejando una alta significancia estadística en los alineamientos locales (Tabla 6).

ID muestra gen <i>16S rRNA</i>	Especie	Patogenicidad	Top 3 % identidad nt	Top 3 Query cover
1686-a	<i>Lacunisphaera anatis</i>	-	<i>Lacunisphaera anatis</i> 98.86% <i>Lacunisphaera limnophila</i> 98.03% <i>Lacunisphaera limnophila</i> 97.99%	<i>Lacunisphaera anatis</i> 99% <i>Lacunisphaera limnophila</i> 97% <i>Lacunisphaera limnophila</i> 97%
1722-a	<i>Lacunisphaera limnophila</i>	-	<i>Lacunisphaera limnophila</i> 95.48% <i>Lacunisphaera anatis</i> 95.53% <i>Lacunisphaera limnophila</i> 95.39%	<i>Lacunisphaera limnophila</i> 99% <i>Lacunisphaera anatis</i> 97% <i>Lacunisphaera limnophila</i> 97%
1722-c	<i>Cytobacillus kochii</i>	-	<i>Cytobacillus kochii</i> 99.24% <i>Cytobacillus purgationiresistensis</i> 99.33% <i>Cytobacillus</i>	<i>Cytobacillus kochii</i> 99% <i>Cytobacillus purgationiresistensis</i> 98% <i>Cytobacillus</i>

			<i>horneckiae</i> 99.71%	<i>horneckiae</i> 97%
1728-a	<i>Reyranella massiliensis</i> 521	-	<i>Reyranella massiliensis</i> 521 98.41% <i>Reyranella soli</i> 98.74% <i>Reyranella graminifolii</i> 97.93%	<i>Reyranella massiliensis</i> 521 98% <i>Reyranella soli</i> 97% <i>Reyranella graminifolii</i> 98%
1728-b	<i>Phenylobacterium conjunctum</i>	-	<i>Phenylobacterium conjunctum</i> 99.93% <i>Phenylobacterium haematophilum</i> 98.33% <i>Phenylobacterium parvum</i> 97.76%	<i>Phenylobacterium conjunctum</i> 99% <i>Phenylobacterium haematophilum</i> 99% <i>Phenylobacterium parvum</i> 99%
1812-b	<i>Phenylobacterium conjunctum</i>	-	<i>Phenylobacterium conjunctum</i> 99.65% <i>Phenylobacterium haematophilum</i> 98.11% <i>Phenylobacterium parvum</i> 97.57%	<i>Phenylobacterium conjunctum</i> 97% <i>Phenylobacterium haematophilum</i> 97% <i>Phenylobacterium parvum</i> 98%

Tabla 6. Resultados de la identificación taxonómica mediante BLASTn de las secuencias del gen 16S rRNA con mejor coincidencia fuera del género *Leptospira*. Para cada muestra se presenta el ID, la especie con mayor similitud obtenida en la base de datos consultada y su clasificación de patogenicidad, además de los top 3 hits con sus porcentajes de identidad y query cover correspondientes. El símbolo (–) indica que no corresponde una clasificación de patogenicidad dentro del género *Leptospira*.

Por otro lado, el análisis de las secuencias del marcador *secY* obtenidas mediante el método Sanger también permitió identificar material genético no relacionado con el género de estudio. Específicamente, las muestras 17b y 20b presentaron una identidad nucleotídica del 100% con *Wigglesworthia glossinidia*. No obstante, estas coincidencias estuvieron asociadas a coberturas de consulta marcadamente bajas, situándose en el 11% y 13% respectivamente, con E-value de 0.008 y 0.007. Estos resultados se sistematizaron para asegurar la trazabilidad y la correcta asignación taxonómica de los fragmentos procesados antes de proceder a las reconstrucciones filogenéticas intragenéricas (Tabla 7).

ID muestra gen <i>secY</i>	Especie	Patogenicidad	Top 3 % identidad nt	Top 3 Query cover
17b	<i>Wigglesworthia glossinidia</i>	-	<i>Wigglesworthia glossinidia</i> 100%	<i>Wigglesworthia glossinidia</i> 11%
20b	<i>Wigglesworthia glossinidia</i>	-	<i>Wigglesworthia glossinidia</i> 100%	<i>Wigglesworthia glossinidia</i> 13%

Tabla 7. Resultados de la identificación taxonómica mediante BLASTn de las secuencias del gen *secY* con mejor coincidencia fuera del género *Leptospira*. Para cada muestra se presenta el ID, la especie con mayor similitud obtenida en la base de datos consultada y su clasificación de patogenicidad, además de los top 3 hits con sus porcentajes de identidad y cobertura correspondientes. El símbolo (–) indica que no corresponde una clasificación de patogenicidad dentro del género *Leptospira*.

Relaciones filogenéticas y corroboración evolutiva de secuencias 16S rRNA

La interpretación y organización de la arquitectura del árbol filogenético se realizó tomando como referencia el trabajo de Agudelo-Flórez *et al.* (2014), el cual sirvió como guía fundamental para establecer la jerarquía de los agrupamientos y las relaciones evolutivas entre los linajes principales del género, permitiendo comparar la topología obtenida con un esquema taxonómico validado. Con el fin de aumentar la precisión en la asignación evolutiva, se incorporó al análisis una secuencia de referencia de *Leptospira iowaensis*, debido a que los análisis previos mediante BLASTn evidenciaron una alta similitud nucleotídica de las muestras ambientales con dicha especie. El conjunto de datos final utilizado para la reconstrucción filogenética fue sometido a un proceso de depuración riguroso en el que se excluyeron aquellas secuencias identificadas inicialmente como géneros no pertenecientes a *Leptospira*, tales como *Lacunisphaera*, *Reyranella* y *Phenylobacterium*, asegurando así la especificidad del análisis intra-genérico. Asimismo, siguiendo criterios metodológicos de refinamiento del *dataset*, las muestras 1682 y 1683 fueron descartadas y no se incluyeron en la construcción de los árboles finales.

La reconstrucción de las relaciones evolutivas basada en el gen *16S rRNA*, ejecutada mediante los métodos de *Neighbor-Joining* (NJ - Figura 9) y *Maximum Likelihood* (ML - Figura 10), generó una topología robusta y coherente en la cual la totalidad de las secuencias ambientales analizadas se agruparon de manera consistente dentro del clado de las leptospiros saprófitas. Las muestras analizadas conformaron un clado monofilético compacto que mostró la mayor cercanía evolutiva con *Leptospira iowaensis*, lo cual presenta una concordancia absoluta con los resultados de identificación taxonómica por similitud obtenidos previamente. La estabilidad de estas relaciones evolutivas se encuentra respaldada por valores de *bootstrap* elevados en los nodos que definen la separación de los grupos de patogenicidad, confirmando que la ubicación de las muestras en el reservorio ambiental de vida libre es estadísticamente confiable.

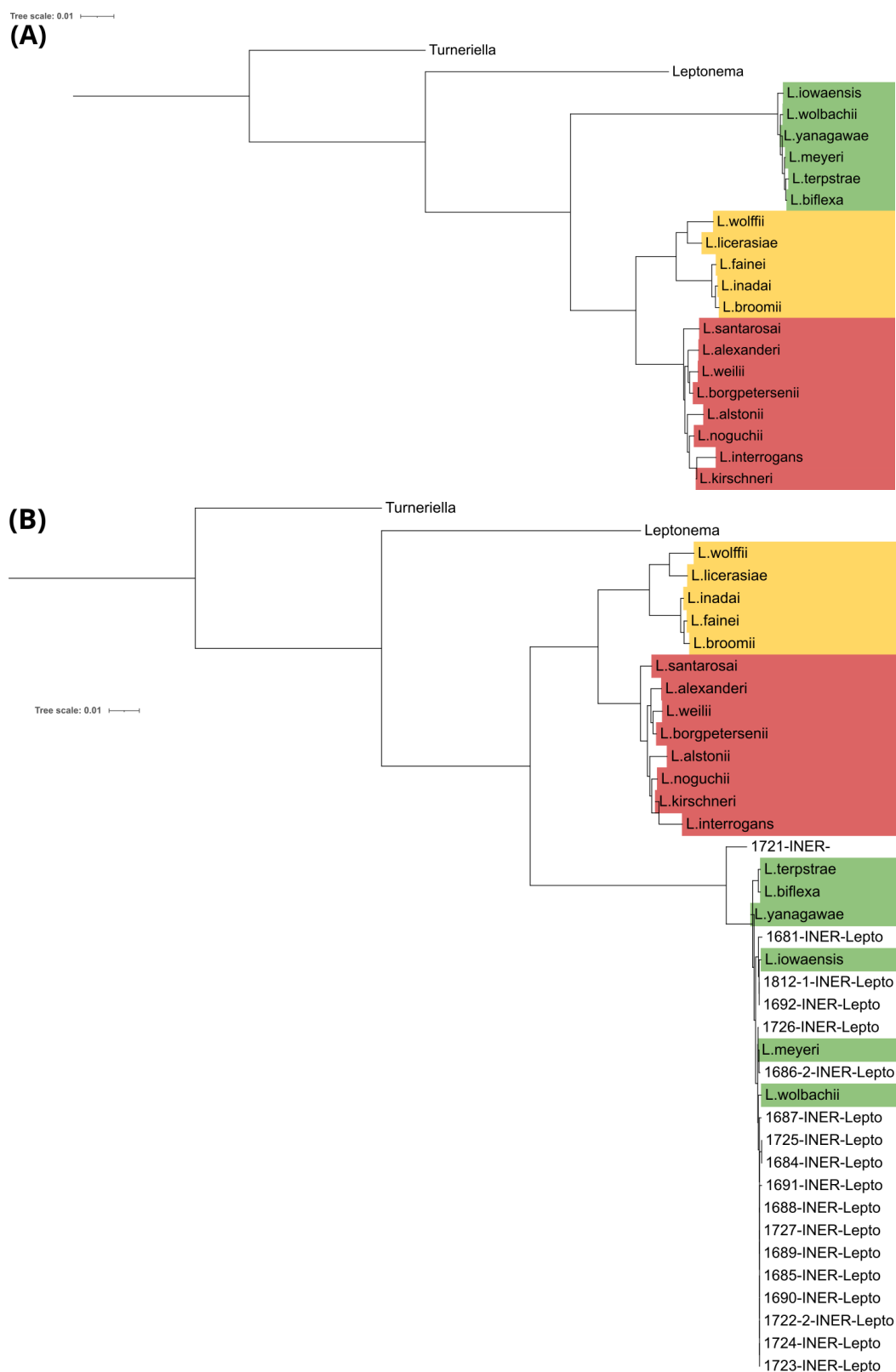


Figura 9. Árboles filogenéticos del gen 16S rRNA de *Leptospira* contruidos mediante el método de *Neighbor-Joining*. El árbol A contiene secuencias de referencia basadas en Agudelo-Flórez *et al.* (2014), mientras que el árbol B incorpora las secuencias obtenidas en este estudio. Los colores indican el grupo filogenético al que pertenece cada especie: rojo, especies patógenas; amarillo, especies de patogenicidad intermedia; y verde, especies saprófitas.

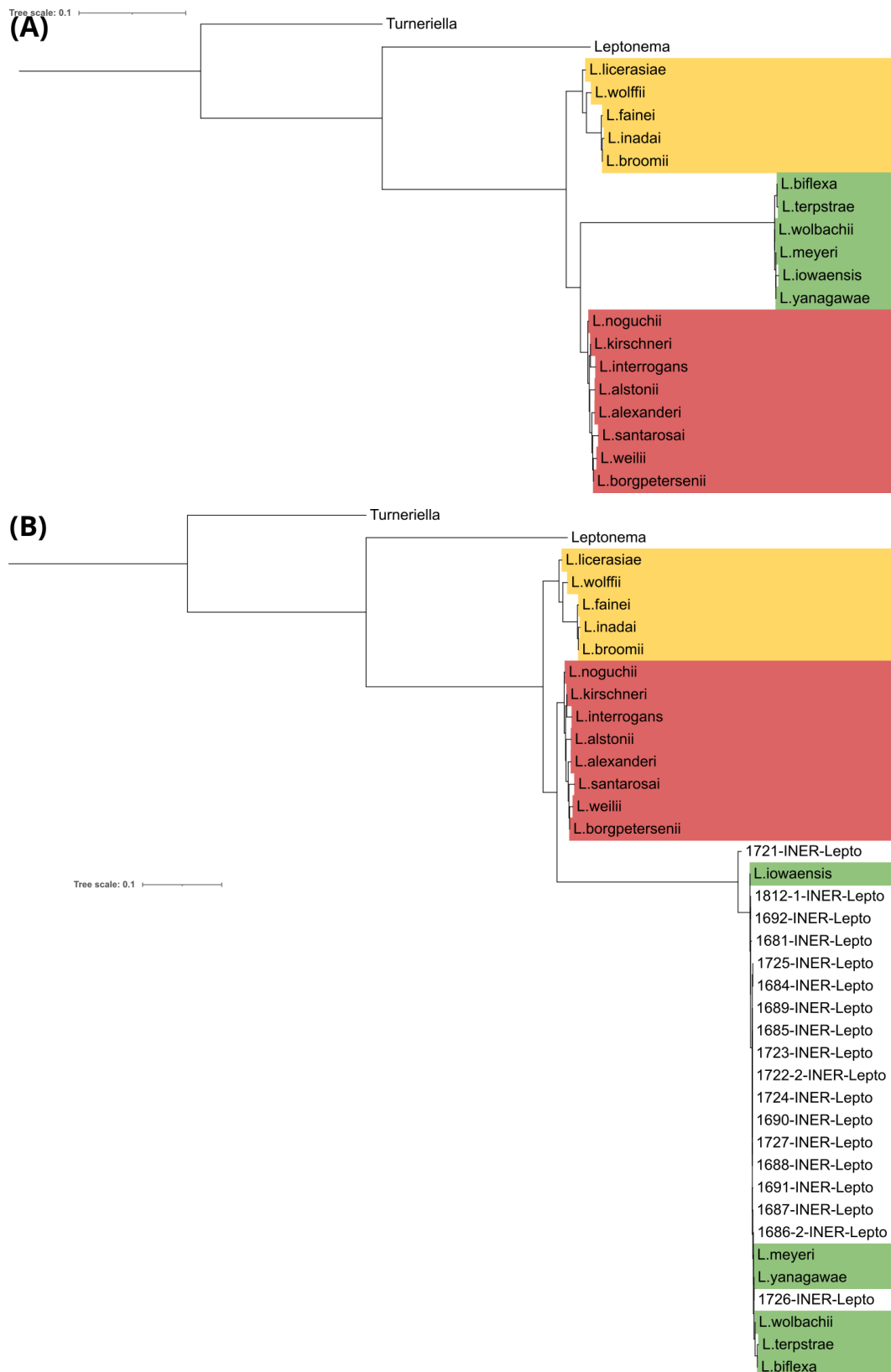


Figura 10. Árboles filogenéticos del gen 16S rRNA de *Leptospira* construidos mediante el método de Maximum Likelihood. El árbol A contiene secuencias de referencia basadas en Agudelo-Flórez *et al.* (2014), mientras que el árbol B incorpora las secuencias obtenidas en este estudio. Los colores indican el grupo filogenético al que pertenece cada especie: rojo, especies patógenas; amarillo, especies de patogenicidad intermedia; y verde, especies saprófitas.

No se detectaron discrepancias topológicas significativas ni agrupamientos inesperados entre los métodos filogenéticos aplicados, manteniendo tanto el NJ como el ML la misma jerarquía evolutiva para los clados mayoritarios, coincidiendo con la organización presentada en el trabajo de Agudelo-Flórez *et al.* (2014). La posición de la raíz en los géneros externos *Turneriella* y *Leptonema* validó la estructura global del árbol, presentando ramas largas que reflejan la distancia evolutiva característica respecto al género *Leptospira*. No obstante, el análisis reveló una resolución interna limitada dentro del subgrupo saprófito, donde la proximidad genética extrema entre especies generó agrupamientos con distancias de rama mínimas, lo que explica la dificultad para alcanzar una discriminación precisa a nivel de especie en este marcador altamente conservado. En conjunto, estos resultados permitieron corroborar la identificación taxonómica de las muestras ambientales procesadas, validando su pertenencia al linaje saprófito mediante un análisis filogenético robusto.

DISCUSIÓN

La caracterización molecular realizada en este estudio evidencia la eficacia de integrar diferentes plataformas de secuenciación para el estudio de *Leptospira* en matrices ambientales complejas, aunque también revela los desafíos técnicos asociados a la secuenciación dirigida. La calidad general de las secuencias de *16S rRNA* obtenidas mediante Illumina fue sobresaliente, con valores de Phred predominantemente superiores a Q30, lo que permitió obtener ensamblajes de alta fidelidad mediante SPAdes, reconstruyendo fragmentos de hasta 2002 pb con coberturas que alcanzaron valores de 9417.8X. En contraste, la variabilidad observada en los datos Sanger correspondientes a los genes *secY* y *LipL32* refleja la sensibilidad de esta tecnología a la calidad del ADN y a la concentración de los amplicones, factores especialmente críticos cuando se trabaja con muestras ambientales, donde la presencia de inhibidores puede comprometer la amplificación y la calidad de las lecturas obtenidas (Deurenberg *et al.*, 2017; Chiani *et al.*, 2016).

La identificación taxonómica mediante *16S rRNA* demostró ser una herramienta robusta para la asignación de las muestras a los grandes clados de patogenicidad, confirmando una presencia mayoritaria de linajes saprófitos asociados a cuerpos de agua (Morey *et al.*, 2006). Sin embargo, la alta conservación de este marcador universal impuso limitaciones críticas en la resolución taxonómica. Se observaron numerosos casos de igual porcentaje de identidad nucleotídica entre taxones distintos, fenómeno que coincide con lo reportado por Morey *et al.* (2006), quienes describen la dificultad de discriminar especies estrechamente relacionadas utilizando únicamente este gen. Esta falta de resolución se debe a la naturaleza altamente conservada del gen *16S rRNA* dentro del género *Leptospira*, lo que impide distinguir especies con secuencias ribosomales casi idénticas y limita la clasificación a niveles taxonómicos más finos, como serovariedades o cepas (Bilung *et al.*, 2026; Caimi & Ruybal, 2020; Morey *et al.*, 2006). Como consecuencia, diferentes especies presentaron porcentajes de identidad equivalentes en los análisis de similitud, dificultando la asignación inequívoca a nivel específico.

El nivel de resolución taxonómica alcanzado con *secY* constituyó la diferencia más relevante respecto al análisis basado en *16S rRNA*. Mientras que el gen ribosomal permitió identificar principalmente clados generales o especies saprófitas con frecuentes ambigüedades taxonómicas, *secY*, al ser un gen *housekeeping* con una tasa de divergencia mayor, proporcionó una resolución significativamente superior. Esta capacidad discriminatoria se

evidenció en la obtención de coincidencias unívocas para especies patógenas como *Leptospira interrogans* y *Leptospira borgpetersenii*, evitando el solapamiento de resultados observado en los *top hits* del *16S rRNA*. En términos de concordancia, los resultados obtenidos con *secY* validaron la presencia del género *Leptospira* previamente detectada mediante otros marcadores, pero aportaron la precisión necesaria para la clasificación intra-género. En conjunto, estos hallazgos indican que, aunque el marcador *16S rRNA* fue eficaz para caracterizar el reservorio ambiental saprófito, la incorporación de *secY* permitió confirmar de manera robusta la circulación de especies patógenas en los sitios estudiados, proporcionando información más precisa para la evaluación preliminar del riesgo sanitario. Asimismo, los resultados subrayan la necesidad de complementar el uso de *16S rRNA* con marcadores de mayor poder discriminatorio cuando se busca una identificación definitiva dentro del género *Leptospira* (Bilung *et al.*, 2026; Caimi & Ruybal, 2020).

Las relaciones filogenéticas obtenidas mediante los métodos *Neighbor-Joining* (Figura 8) y *Maximum Likelihood* (Figura 9) mostraron una elevada concordancia con los resultados de BLASTn, validando las relaciones evolutivas inferidas para la mayoría de las muestras analizadas. El agrupamiento de numerosas secuencias dentro de un clado monofilético compacto junto a referencias de *Leptospira iowaensis* refuerza la consistencia de los resultados obtenidos, especialmente considerando que esta especie fue identificada de manera recurrente como el principal hit en los análisis de similitud. Estos hallazgos son consistentes con la organización taxonómica propuesta por Agudelo-Flórez *et al.* (2014), utilizada como marco de referencia para interpretar la estabilidad de los linajes identificados y la correcta ubicación de la raíz mediante los géneros externos *Turneriella* y *Leptonema*. Asimismo, las ambigüedades taxonómicas observadas en el análisis del *16S rRNA*, donde múltiples especies saprófitas compartían porcentajes de identidad equivalentes, explican la formación de subgrupos separados por distancias genéticas mínimas dentro de los árboles filogenéticos, reflejando nuevamente las limitaciones inherentes de este marcador para resolver relaciones evolutivas recientes (Morey *et al.*, 2006).

Un aspecto crítico de esta discusión es la exclusión del marcador *LipL32* de los análisis finales debido a problemas severos de calidad en los electroferogramas obtenidos. La presencia de una baja relación señal-ruido y de picos ambiguos exigió la aplicación de procedimientos de trimming agresivos que resultaron en la pérdida de regiones alineables esenciales, impidiendo la generación de secuencias consenso confiables. Es importante

interpretar estos resultados como una limitación metodológica y no como evidencia de ausencia biológica del gen en las muestras analizadas. Factores técnicos como la degradación del ADN ambiental, la baja concentración de amplicones o limitaciones inherentes a la secuenciación Sanger en matrices complejas podrían explicar el deterioro observado en la calidad de las lecturas (Deurenberg *et al.*, 2017; Chiani *et al.*, 2016). Esta problemática ha sido señalada previamente como un obstáculo frecuente para la detección molecular de leptospiros patógenos en ambientes acuáticos, donde la dilución del material genético y la presencia de inhibidores dificultan la obtención de secuencias de alta calidad (Warnasekara *et al.*, 2022; Wynwood *et al.*, 2014).

Las fortalezas del *pipeline* bioinformático empleado radican en la validación cruzada entre múltiples marcadores moleculares, el uso de metodologías filogenéticas complementarias y la utilización de bases de datos curadas para la identificación taxonómica, lo que permitió minimizar posibles sesgos de asignación. No obstante, el estudio presenta limitaciones inherentes a la naturaleza intermitente de la detección bacteriana en ambientes acuáticos y a la dependencia de la calidad inicial del ADN extraído. Los resultados obtenidos demuestran que la combinación de marcadores como *16S rRNA* y *secY* constituye una estrategia efectiva para la detección y caracterización de *Leptospira* en muestras ambientales; sin embargo, también evidencian que la resolución taxonómica alcanzable mediante secuenciación dirigida continúa siendo limitada para abordar niveles inferiores de clasificación, como serovariedades o cepas. En este contexto, futuras investigaciones podrían beneficiarse de la implementación de estrategias basadas en secuenciación de genoma completo (WGS) o enfoques metagenómicos utilizando plataformas de alta capacidad como Illumina u Oxford Nanopore (Bilung *et al.*, 2026; Picardeau, 2017). Estas aproximaciones permitirían superar las limitaciones asociadas a la amplificación dirigida, alcanzar una resolución taxonómica más fina y proporcionar una visión integral de la diversidad del género *Leptospira* presente en reservorios ambientales, contribuyendo a una evaluación más precisa de los riesgos potenciales para la salud pública (Picardeau, 2017; Bierque *et al.*, 2020; Adler, 2014).

CONCLUSIONES

El trabajo realizado permitió concluir que la combinación de los marcadores moleculares *16S rRNA* y *secY* constituyó una estrategia complementaria para la caracterización taxonómica del género *Leptospira* en matrices ambientales. El análisis del *16S rRNA* permitió identificar principalmente especies saprófitas, con un claro predominio de *Leptospira iowaensis*, mientras que el marcador *secY* proporcionó una mayor resolución taxonómica, posibilitando la discriminación a nivel de especie e identificando diversas especies patógenas, entre ellas *Leptospira interrogans*, *Leptospira borgpetersenii*, *Leptospira santarosai*, *Leptospira noguchii*, *Leptospira tipperaryensis* y *Leptospira cinconiae*. En contraste, el marcador *LipL32* no pudo ser incorporado a los análisis finales debido a limitaciones en la calidad de las secuencias obtenidas.

Se observó concordancia entre las identificaciones obtenidas mediante búsquedas por similitud con BLASTn y las reconstrucciones filogenéticas realizadas mediante los métodos de Neighbor-Joining y Maximum Likelihood, lo que respalda la consistencia de la asignación taxonómica y de la ubicación evolutiva de las muestras analizadas dentro de los grupos filogenéticos esperados. Asimismo, la mayoría de las secuencias correspondientes al gen *16S rRNA* mostró una estrecha relación filogenética con *Leptospira iowaensis*, especie que representó la principal coincidencia en los análisis de identidad nucleotídica.

El *pipeline* bioinformático implementado demostró ser una herramienta eficiente para el control de calidad, el procesamiento de datos de secuenciación y la integración de análisis de identificación taxonómica y reconstrucción filogenética. Este flujo de trabajo permitió obtener secuencias de calidad adecuada para los análisis taxonómicos y filogenéticos en aquellos casos en que los datos de secuenciación cumplieron los criterios de calidad establecidos. No obstante, el estudio presentó limitaciones metodológicas asociadas al marcador *LipL32*, cuya baja calidad en las secuencias obtenidas mediante secuenciación Sanger impidió la generación de secuencias consenso confiables. En consecuencia, este marcador fue excluido de los análisis taxonómicos y filogenéticos. Es importante destacar que esta exclusión responde exclusivamente a una limitación técnica del proceso de secuenciación y no debe interpretarse como evidencia de la ausencia del gen en las muestras analizadas.

Por otra parte, la detección de secuencias correspondientes a organismos ajenos al género *Leptospira*, como *Lacunisphaera*, *Phenylobacterium*, *Cytobacillus*, *Reyranella* y *Wigglesworthia*, pone de manifiesto la complejidad inherente al análisis de muestras

ambientales. La presencia de estos microorganismos puede atribuirse tanto a la diversidad natural de la microbiota presente en agua y suelo como a posibles factores técnicos, entre ellos la amplificación inespecífica, la contaminación cruzada o errores ocurridos durante la manipulación de muestras y reactivos. Estos resultados resaltan la importancia de implementar controles de calidad rigurosos, realizar una revisión de los primers específicos diseñados para este experimento, y de complementar las búsquedas por similitud con análisis filogenéticos para validar la asignación taxonómica de las secuencias obtenidas.

En conjunto, los resultados obtenidos demuestran que la integración de herramientas bioinformáticas para el procesamiento de datos de secuenciación, la identificación mediante BLASTn y la reconstrucción filogenética constituye una estrategia robusta para el estudio de la diversidad de *Leptospira* en matrices ambientales. La combinación de un marcador conservado, como *16S rRNA*, con un marcador de mayor resolución, como *secY*, permitió obtener una caracterización taxonómica más completa de las secuencias analizadas y aporta una base metodológica útil para futuros estudios de vigilancia epidemiológica, monitoreo ambiental y comprensión de la diversidad evolutiva del género *Leptospira*.

BIBLIOGRAFÍA / REFERENCIAS

Adler, B. (2014). *Leptospira and leptospirosis*. Springer.
<https://doi.org/10.1007/978-3-662-45059-8>

Adler, B., & de la Peña Moctezuma, A. (2010). *Leptospira and leptospirosis*. Veterinary Microbiology, 140(3-4), 287-296. <https://doi.org/10.1016/j.vetmic.2009.03.012>

Agudelo-Flórez, P., Durango, H., Aranzazu, D., Rodas, J. D., & Travi, B. (2014). *Genotipificación y evaluación de la dinámica de infección de un aislamiento colombiano de Leptospira santarosai en el modelo experimental en hámster*. Biomédica, 34(3), 460–472.
<https://doi.org/10.7705/biomedica.v34i3.1703>

Amplicon, P. C. R., Clean-Up, P. C. R., & Index, P. C. R. (2013). *16s metagenomic sequencing library preparation*. Illumina: San Diego, CA, USA, 21.

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Babraham Bioinformatics. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Bharti, A. R., Nally, J. E., Ricaldi, J. N., Matthias, M. A., Diaz, M. M., Lovett, M. A., Levett, P. N., Gilman, R. H., Willig, M. R., Gotuzzo, E., & Vinetz, J. M. (2003). *Leptospirosis: A zoonotic disease of global importance*. The Lancet Infectious Diseases, 3(12), 757-771.
[https://doi.org/10.1016/S1473-3099\(03\)00830-2](https://doi.org/10.1016/S1473-3099(03)00830-2)

Bierque, E., Thibeaux, R., Giralt, D., Soupé-Gilbert, M.-E., & Goarant, C. (2020). *A systematic review of Leptospira in water and soil environments*. PLoS ONE, 15(1), e0227055.
<https://doi.org/10.1371/journal.pone.0227055>

Bilung, L. M., Tahar, A. S., Pui, C. F., Bakeri, M. K. S., Su'ut, L., Ngui, R., Kira, R., & Apun, K. (2026). *Leptospira and leptospirosis: A review of species classifications, genomes, morphological structures, antimicrobial resistances, transmissions, and clinical manifestations*. Current Microbiology, 83(2), 122.
<https://doi.org/10.1007/s00284-026-04722-7>

Caimi, K., & Ruybal, P. (2020). *Leptospira spp., a genus in the stage of diversity and genomic data expansion*. Infection, Genetics and Evolution. Advanced online publication. <https://doi.org/10.1016/j.meegid.2020.104241>

Centers for Disease Control and Prevention. (s.f.). *Leptospira interrogans*. [Micrografía electrónica]. Public Health Image Library (PHIL). <https://wwwn.cdc.gov/phil/details.aspx?pid=1220>

Chao, K.-H., Barton, K., Palmer, S., & Lanfear, R. (2021). *sangeranalyseR: Simple and interactive processing of Sanger sequencing data in R*. Genome Biology and Evolution, 13(3), evab028. <https://doi.org/10.1093/gbe/evab028>

Chiani, Y., Jacob, P., Varni, V., Landolt, N., Schmeling, M. F., Pujato, N., Caimi, K., & Vanasco, B. (2016). *Isolation and clinical sample typing of human leptospirosis cases in Argentina*. Infection, Genetics and Evolution, 37, 245-251. <https://doi.org/10.1016/j.meegid.2015.11.022>

Costa, F., Hagan, J. E., Calcagno, J., Kane, M., Torgerson, P., Martinez-Silveira, M. S., Stein, C., Abela-Ridder, B., & Ko, A. I. (2015). *Global morbidity and mortality of leptospirosis: A systematic review*. PLoS Neglected Tropical Diseases, 9(9), e0003898. <https://doi.org/10.1371/journal.pntd.0003898>

Criscuolo, A., & Gribaldo, S. (2010). *BMGE (Block Mapping and Gathering with Entropy): A new software for selection of phylogenetic informative regions from multiple sequence alignments*. BMC Evolutionary Biology, 10, 210. <https://doi.org/10.1186/1471-2148-10-210>

Deharvengt, S. J., Petersen, L. M., Jung, H.-S., & Tsongalis, G. J. (2020). *Nucleic acid analysis in the clinical laboratory*. En W. J. Marshall et al. (Eds.), *Contemporary practice in clinical chemistry* (4th ed., pp. 365-387). Elsevier. <https://doi.org/10.1016/B978-0-12-815499-1.00013-2>

- Deurenberg, R. H., Bathoorn, E., Chlebowicz, M. A., Couto, N., Ferdous, M., García-Cobos, S., Kooistra-Smid, A. M. D., Raangs, E. C., Rosema, S., Veloo, A. C. M., Zhou, K., Friedrich, A. W., & Rossen, J. W. A. (2017). *Application of next generation sequencing in clinical microbiology and infection prevention*. Journal of Biotechnology, 243, 16-24. <https://doi.org/10.1016/j.jbiotec.2016.12.022>
- Gómez, A. A., López, M. S., Müller, G. V., López, L. R., Sione, W., & Giovanini, L. (2022). *Modeling of leptospirosis outbreaks in relation to hydroclimatic variables in the northeast of Argentina*. Heliyon, 8(6), e09758. <https://doi.org/10.1016/j.heliyon.2022.e09758>
- Gozzi, A. C., Guichón, M. L., Benítez, V. V., Romero, G. N., Auteri, C., & Brihuega, B. (2013). *First isolation of Leptospira interrogans from the arboreal squirrel Callosciurus erythraeus introduced in Argentina*. Wildlife Biology, 19, 483-489. <https://doi.org/10.2981/13-051>
- Grune Löffler, S. (2017). *Leptospirosis en poblaciones urbanas y periurbanas de roedores en la Argentina*. Salud(i)Ciencia, 22, 668-671.
- Haake, D. A., & Levett, P. N. (2015). *Leptospirosis in humans*. En B. Adler (Ed.), *Leptospira and leptospirosis* (pp. 65-97). Springer. https://doi.org/10.1007/978-3-662-45059-8_5
- Johnson, M., Zaretskaya, I., Raytselis, Y., Merezuk, Y., McGinnis, S., & Madden, T. L. (2008). *NCBI BLAST: A better web interface*. Nucleic Acids Research, 36(Web Server issue), W5-W9. <https://doi.org/10.1093/nar/gkn201>
- Katoh, K., & Standley, D. M. (2013). *MAFFT multiple sequence alignment software version 7: Improvements in performance and usability*. Molecular Biology and Evolution, 30(4), 772–780. <https://doi.org/10.1093/molbev/mst010>
- Khan, M. A., Khan, S. T., Sequeira, M. C., Faheem, S. M., & Rais, N. (2021). *Illumina sequencing of 16S rRNA genes reveals a unique microbial community in three anaerobic sludge digesters of Dubai*. PLoS ONE, 16(4), e0249023. <https://doi.org/10.1371/journal.pone.0249023>

Ko, A. I., Reis, M. G., Dourado, C. M. R., Johnson, W. D., Jr., Riley, L. W., & Salvador Leptospirosis Study Group. (1999). *Urban epidemic of severe leptospirosis in Brazil*. The Lancet, 354(9181), 820-825. [https://doi.org/10.1016/S0140-6736\(99\)80012-9](https://doi.org/10.1016/S0140-6736(99)80012-9)

Larsson, A. (2014). *AliView: A fast and lightweight alignment viewer and editor for large datasets*. Bioinformatics, 30(22), 3276–3278.

Lau, C. L., Smythe, L. D., Craig, S. B., & Weinstein, P. (2010). *Climate change, flooding, urbanisation and leptospirosis: Fuelling the fire?* Transactions of the Royal Society of Tropical Medicine and Hygiene, 104(10), 631-638. <https://doi.org/10.1016/j.trstmh.2010.07.002>

Levett, P. N. (2001). *Leptospirosis*. Clinical Microbiology Reviews, 14(2), 296-326. <https://doi.org/10.1128/CMR.14.2.296-326.2001>

Letunic, I., & Bork, P. (2024). *Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool*. Nucleic Acids Research, 52(W1), W78–W82. <https://doi.org/10.1093/nar/gkae268>

López, M. S., Müller, G. V., Lovino, M. A., Gómez, A. A., Sione, W. F., & Aragonés Pomares, L. (2019). *Spatio-temporal analysis of leptospirosis incidence and its relationship with hydroclimatic indicators in northeastern Argentina*. Science of the Total Environment, 694, 133651. <https://doi.org/10.1016/j.scitotenv.2019.133651>

Lotto Batista, M., Rees, E. M., Gómez, A., López, S., Castell, S., Kucharski, A. J., Ghazzi, S., Müller, G. V., & Lowe, R. (2023). *Towards a leptospirosis early warning system in northeastern Argentina*. Journal of the Royal Society, Interface, 20, 20230069. <https://doi.org/10.1098/rsif.2023.0069>

McBride, A. J. A., Athanazio, D. A., Reis, M. G., & Ko, A. I. (2005). *Leptospirosis: Current Opinion in Infectious Diseases*, 18(5), 375-386. <https://doi.org/10.1097/01.qco.0000178824.05715.2c>

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). *IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era*. *Molecular Biology and Evolution*, 37(5), 1530–1534. <https://doi.org/10.1093/molbev/msaa015>

Ministerio de Salud de la República Argentina, Dirección de Epidemiología. (2025a). *Boletín Epidemiológico Nacional N° 771, SE 34*. (pp. 58-70).

Ministerio de Salud de la Nación (2025b). *Leptospirosis*.

Morey, R. E., Galloway, R. L., Bragg, S. L., Steigerwalt, A. G., Mayer, L. W., & Levett, P. N. (2006). *Species-specific identification of Leptospiraceae by 16S rRNA gene sequencing*. *Journal of Clinical Microbiology*, 44(10), 3510–3516. <https://doi.org/10.1128/JCM.00670-06>

Pavan, M. E., Cairó, F., Pettinari, M. J., Samartino, L., & Brihuega, B. (2011). *Genotyping of Leptospira interrogans strains from Argentina by multiple-locus-variable-number tandem repeat analysis (MLVA)*. *Comparative Immunology, Microbiology and Infectious Diseases*, 34(2), 135-141. <https://doi.org/10.1016/j.cimid.2010.06.002>

Picardeau, M. (2013). *Diagnosis and epidemiology of leptospirosis*. *Médecine et Maladies Infectieuses*, 43(1), 1-9. <https://doi.org/10.1016/j.medmal.2012.11.005>

Picardeau, M. (2017). *Virulence of the zoonotic agent of leptospirosis: Still terra incognita?* *Nature Reviews Microbiology*, 15(5), 297-307. <https://doi.org/10.1038/nrmicro.2017.5>

Pichler, M., Coskun, Ö. K., Ortega-Arbulú, A.-S., Conci, N., Wörheide, G., Vargas, S., & Orsi, W. D. (2018). *A 16S rRNA gene sequencing protocol for the Illumina MiniSeq platform*. *MicrobiologyOpen*, 7(6), e611. <https://doi.org/10.1002/mbo3.611>

Ricardo, T., Bergero, L. C., Bulgarella, E. P., & Previtali, M. A. (2018). *Knowledge, attitudes and practices (KAP) regarding leptospirosis among residents of riverside settlements of Santa Fe, Argentina*. PLoS Neglected Tropical Diseases, 12(5), e0006470. <https://doi.org/10.1371/journal.pntd.0006470>

Ricardo, T., Jacob, P., Chiani, Y., Schmeling, M. F., Cornejo, P., Ojeda, A. A., Teta, P. V., Vanasco, N. B., & Previtali, M. A. (2020). *Seroprevalence of leptospiral antibodies in rodents from riverside communities of Santa Fe, Argentina*. PLoS Neglected Tropical Diseases, 14(4), e0008222. <https://doi.org/10.1371/journal.pntd.0008222>

Staden, R. (1996). *The Staden sequence analysis package*. Molecular Biotechnology, 5(3), 233-241.

Staden, R., Beal, K. F., & Bonfield, J. K. (2000). *The Staden package*. En A. M. Griffin & H. G. Griffin (Eds.), *Methods in molecular biology* (Vol. 132, pp. 115-130). Humana Press. <https://doi.org/10.1385/1-59259-192-2:115>

Vanasco, N. B., Sequeira, G., Dalla Fontana, M. L., Fusco, S., Sequeira, M. D., & Enría, D. (2000). *Descripción de un brote de leptospirosis en la ciudad de Santa Fe, Argentina, marzo-abril de 1998*. Revista Panamericana de Salud Pública, 7(1), 35-40.

Vijayachari, P., Sugunan, A. P., & Shriram, A. N. (2008). *Leptospirosis: An emerging global public health problem*. Journal of Biosciences, 33(4), 557-569.

Vital-Brazil, J. M., Balassiano, I. T., Oliveira, F. S., Costa, A. D. S., Hillen, L., & Pereira, M. M. (2010). *Multiplex PCR-based detection of leptospira in environmental water samples obtained from a slum settlement*. Memórias do Instituto Oswaldo Cruz, 105(3), 353-355.

Warnasekara, J., Srimantha, S., Senavirathna, I., Kappagoda, C., Farika, N., Nawinna, A., & Agampodi, S. (2022). *The variable presence of Leptospira in the environment: An*

epidemiological explanation based on serial analysis of water samples. PLoS ONE, 17(2), e0263719.

Wynwood, S. J., Graham, G. C., Weier, S. L., Collet, T. A., McKay, D. B., & Craig, S. B. (2014). *Leptospirosis from water sources*. Pathogens and Global Health, 108(7), 334-338.
<https://doi.org/10.1179/2047773214Y.00000000156>