



UNIVERSIDAD INTERNACIONAL SEK

FACULTAD DE CIENCIAS DE LA SALUD

Trabajo de Fin de Máster Titulado:

Desarrollo de un pipeline bioinformático para el análisis filogenético de *Babesia bovis* y *Babesia bigemina* en *Rhipicephalus microplus* (Ixodidae) mediante secuenciación con nanoporos

Realizado por:

WENDY ELIZABETH TIPANTA FLORES

Director del proyecto:

Dr. Juan Carlos Navarro C., Ph.D.

Como requisito para la obtención del título de:
MAGISTER EN BIOMEDICINA

Quito, 18 de septiembre de 2025

DECLARACION JURAMENTADA

Yo, **Wendy Tipanta**, con cédula de identidad 1723901094 declaro bajo juramento que el trabajo aquí desarrollado es de mi autoría, que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración, cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la UNIVERSIDAD INTERNACIONAL SEK, según lo establecido por la Ley de Propiedad Intelectual, por su reglamento y por la normativa institucional vigente.

FIRMA Y CÉDULA

1723901094

DECLARACIÓN DEL DIRECTOR DE TESIS

Declaro haber dirigido este trabajo a través de reuniones periódicas con el estudiante, orientando sus conocimientos y competencias para un eficiente desarrollo del tema escogido y dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación.

Dr Juan Carlos Navarro C, Ph.D

LOS PROFESORES INFORMANTES

Los Profesores Informantes:

MANUEL ANDRÉS HERRERA YELA

JOSÉ RUBÉN RAMÍREZ-IGLESIAS

Después de revisar el trabajo presentado,
lo han calificado como apto para su defensa oral ante
el tribunal examinador

FIRMA

FIRMA

Quito, 18 de septiembre de 2025.

DECLARACIÓN DE AUTORÍA DEL ESTUDIANTE

Declaro que este trabajo es original, de mi autoría, que se han citado las fuentes correspondientes y que en su ejecución se respetaron las disposiciones legales que protegen los derechos de autor vigentes.

WENDY ELIZABETH TIPANTA FLORES
C.I.: 1723901094

Quito, 18 de septiembre de 2025

DEDICATORIA

*A mis Padres, Julio y
Consuelo, mi hermano
Andrés.*

AGRADECIMIENTO

*A mis padres Julio y Consuelo,
por su apoyo incondicional en cada paso de mi vida tanto personal como
estudiantil, que con sus consejos y guía me motivaron día a día para
convertirme en un gran ser humano y una buena profesional. A mi hermano
Julio Andrés, quien sigue mis pasos, siendo mi mayor motivación, por darme el
empuje de salir adelante. A mis abuelitos, tíos, primos por estar presente en todo
momento de mi vida. A Carlos por brindarme su amor, comprensión cuidado y
ayuda en buenos y malos momentos que he vivido.*

*A mis profesores del programa de
Maestría, por compartir sus conocimientos.*

Artículo de tesis

Desarrollo de un pipeline bioinformático para el análisis filogenético de *Babesia bovis* y *Babesia bigemina* en *Rhipicephalus microplus* mediante secuenciación con nanoporos

Tipanta Wendy¹, Andrés Herrera^{1,2}, Juan-Carlos Navarro^{1,2}

- 1 Maestría en Biomedicina, Facultad de Ciencias de la Salud, UISEK; wendy.tipanta@uisek.edu.ec; manuel.herrera@uisek.edu.ec; juancarlos.navarro@uisek.edu.ec
- 2 Grupo de Investigación en Enfermedades Emergentes y Desatendidas, Ecoepidemiología y Biodiversidad, Facultad de Ciencias de la Salud, UISEK.

Resumen

La babesiosis bovina es una enfermedad parasitaria transmitida por garrapatas del género *Rhipicephalus*, que representa una amenaza significativa para la salud animal y la economía agropecuaria en regiones tropicales y subtropicales. En América Latina, las especies *Babesia bovis* y *Babesia bigemina* son las más prevalentes, por lo que su caracterización genética resulta esencial para comprender su diversidad, evolución y dinámica epidemiológica.

El objetivo de este estudio fue desarrollar un pipeline bioinformático estandarizado para el análisis filogenético de *B. bovis* y *B. bigemina* a partir de datos de secuenciación generados mediante tecnología Oxford Nanopore. El flujo de trabajo fue implementado en la plataforma Galaxy Australia e incluyó control de calidad con FastQC, recorte de adaptadores con Porechop, alineamiento genómico mediante Bowtie2 y Samtools, ensamblaje con SPAdes y filtrado de secuencias específicas a través de una base de datos BLAST personalizada y comandos AWK. Las secuencias seleccionadas correspondieron a los genes RAP-1a (para *B. bovis*) y SpeI-AvaI (para *B. bigemina*).

Los análisis filogenéticos se realizaron en PAUP v4.0b169*, bajo criterios de Máxima Parsimonia, y se evaluó la solidez estadística mediante 1.000 pseudoréplicas bootstrap. Los alineamientos se llevaron a cabo en MacVector v18.7.6 utilizando ClustalW (GOP: 30.0; GEP: 10.0). Los árboles obtenidos revelaron diferencias notables entre las dos especies: *B. bovis* presentó un clado ecuatoriano bien soportado y genéticamente homogéneo, con escasa diferenciación entre cepas de Napo, Sucumbíos y Orellana, lo que sugiere una expansión clonal reciente. En contraste, *B. bigemina* mostró una mayor estructuración filogenética, con cuatro subclados diferenciados y valores bootstrap entre 88% y 100%, reflejando una diversidad genética más alta y una posible estructuración geográfica regional.

Estos resultados evidencian la coexistencia de linajes locales y su divergencia frente a cepas de referencia internacionales. Asimismo, destacan la importancia de incorporar datos genéticos regionales en estrategias diagnósticas y de control. El pipeline desarrollado demostró ser eficaz y reproducible, aportando una base sólida para estudios moleculares y epidemiológicos sobre *Babesia spp.* en el Ecuador. Este trabajo contribuye significativamente al conocimiento de la filogeografía y diversidad genética de estos hemoparásitos, elementos clave para el diseño de programas de vigilancia y prevención de la babesiosis bovina.

Palabras clave: *Babesia bovis*, *Babesia bigemina*, *Rhipicephalus*, pipeline, filogenética, RAP-1a, SPE1-AVA1.

Abstract

Bovine babesiosis is a parasitic disease transmitted by *Rhipicephalus* ticks, posing a significant threat to animal health and agricultural productivity in tropical and subtropical regions. In Latin America, *Babesia bovis* and *Babesia bigemina* are the most prevalent species; therefore, precise genetic characterization is essential to understand their diversity, evolution, and epidemiological dynamics.

This study aimed to develop a standardized bioinformatic pipeline for the phylogenetic analysis of *B. bovis* and *B. bigemina* using sequencing data generated by Oxford Nanopore technology. The workflow was implemented in the Galaxy Australia platform and included quality control with FastQC, adapter trimming with Porechop, genome alignment using Bowtie2 and Samtools, de novo assembly with SPAdes, and specific sequence selection through a custom BLAST database and AWK filtering commands. The selected sequences corresponded to RAP-1a (for *B. bovis*) and SpeI-AvaI (for *B. bigemina*).

Phylogenetic analyses were performed using PAUP v4.0b169* under a Maximum Parsimony framework, and clade support was assessed using 1,000 bootstrap replicates. Multiple sequence alignments were conducted in MacVector v18.7.6 with ClustalW (GOP: 30.0; GEP: 10.0). The resulting trees revealed contrasting patterns between species: *B. bovis* displayed a highly supported and genetically homogeneous Ecuadorian clade, suggesting a recent clonal expansion across Napo, Sucumbíos, and Orellana provinces. In contrast, *B. bigemina* exhibited greater phylogenetic structure, with four well-supported subclades (bootstrap: 88–100%), reflecting higher genetic diversity and possible regional geographic differentiation.

These findings highlight the coexistence of local lineages and their divergence from international reference strains. Furthermore, they emphasize the importance of incorporating regional genetic data into diagnostic and control strategies. The developed pipeline proved effective and reproducible, offering a robust foundation for future molecular and epidemiological studies on *Babesia spp.* in Ecuador. This research contributes significantly to the understanding of the phylogeography and genetic diversity of these hemoparasites, which are key components in the design of surveillance and prevention programs for bovine babesiosis.

Keywords: *Babesia bovis*, *Babesia bigemina*, *Rhipicephalus*, pipeline, phylogenetics, RAP-1a, SPE1-AVA1.

1. Introducción

La babesiosis es una enfermedad emergente transmitida por garrapatas (Kumar et al., 2021), provocada por protozoos intracelulares del género *Babesia*, que infectan a glóbulos rojos de sus hospedadores mamíferos (Schäfer et al., 2023). La babesiosis tiene una gran relevancia a nivel mundial, ya que afecta tanto a seres humanos como a animales domésticos (Bajer et al., 2022), siendo una enfermedad de gran impacto económico para la industria ganadera debido a la morbilidad y mortalidad que provoca.

La transmisión de *Babesia* en el ganado se produce principalmente a través de la picadura de garrapatas del género *Rhipicephalus*, como *R. microplus*. Durante la alimentación, las garrapatas ingieren eritrocitos infectados con *Babesia*. En el intestino medio de la garrapata, los parásitos se

diferencian en gametos, los cuales se fusionan para formar cigotos. Estos cigotos se desarrollan en cinetos, que migran a la hemolinfa e invaden diversos tejidos, incluyendo los ovarios, permitiendo la transmisión vertical a través de los huevos. No obstante, sólo una fracción de los parásitos sobreviven a este proceso de transmisión transovárica, lo que sugiere la existencia de un cuello de botella poblacional de *Babesia* durante la transmisión (Vimonish et al., 2025).

El ciclo de vida de *Babesia* comprende una serie de etapas y transiciones críticas para su replicación y propagación en el hospedador y vector. Durante estas fases, se observa una regulación de la expresión génica que modula proteínas implicadas en diversos procesos celulares y mecanismos moleculares, los cuales se encuentran integrados en múltiples vías de señalización intracelular. No obstante, el conocimiento actual de estos eventos biológicos es limitado, y las interacciones moleculares entre *Babesia*, el hospedador y el vector permanecen parcialmente caracterizadas. Esta brecha de conocimiento restringe la comprensión integral de los mecanismos de transmisión y patogénesis de *Babesia*, así como de las respuestas inmunes del hospedador. Por consiguiente, se dificulta el desarrollo de estrategias de control efectivas, incluyendo terapias farmacológicas y el diseño de vacunas basadas en blancos moleculares. Resulta crucial, por tanto, profundizar en la caracterización de los mecanismos moleculares que rigen la interacción hospedador-vector-*Babesia*, con el fin de identificar nuevas intervenciones terapéuticas y preventivas contra la babesiosis (Capelli-Peixoto et al., 2022).

Las especies de *Babesia* más comunes que infectan al ganado bovino son *B. bovis*, *B. bigemina* y *B. divergens* (Gohil et al., 2013). *B. bovis* y *B. bigemina* se encuentran en África, Australia, América Central y del Sur, siendo sus principales vectores *R. (Boophilus) microplus*, *R. annulatus* y *R. geigyi*. Por otro lado, *B. divergens* está distribuido en el noroeste de Europa, incluyendo España, Gran Bretaña e Irlanda, y su principal vector es *Ixodes ricinus* (Bock et al., 2004). El curso de la babesiosis puede variar considerablemente, desde infecciones completamente asintomáticas, pasando por manifestaciones leves no específicas, hasta una enfermedad multisistémica mortal (Bock et al., 2004).

Gran parte de estudios realizados en el Ecuador sobre *Babesia* se han centrado en la ganadería y en caninos, debido a la relevancia de estas especies en la salud pública y la economía del país, siendo las especies más relevantes: *B. bovis*, *B. bigemina*, *B. canis*, *B. vogeli*, *B. gibsoni* y *B. vulpes* (Chávez-Larrea et al., 2021).

Los marcadores moleculares son secuencias de ADN específicas que representan herramientas fundamentales para la identificación, diferenciación y análisis de la variabilidad genética en organismos, posibilitan la identificación de loci asociados a características fenotípicas de interés, facilitando estudios de genética de poblaciones, filogenia y mejoramiento genético (Mischkulnig et al., 2022). El marcador RAP-1a en *B. bovis* y el marcador SPE1-AVA1 en *B. bigemina* han

demostrado ser efectivos para discriminar entre estas especies, ya que se encuentran en regiones del genoma que presentan una alta variabilidad, lo que permite distinguir las de otras especies de *Babesia* y también evaluar su diversidad genética.

La detección de apicomplejos, incluyendo piroplásmidos transmitidos por garrapatas, ha dependido tradicionalmente de la microscopía de frotis sanguíneo, considerada como el gold standard. Sin embargo, la sensibilidad limitada de esta técnica y la dificultad en la diferenciación morfológica de especies estrechamente relacionadas, como las variantes de *Babesia*, plantean desafíos significativos. Los ensayos serológicos, aunque capaces de superar algunas de estas limitaciones, no permiten discriminar entre infecciones activas y pasadas. En contraste, las técnicas de PCR y qPCR han demostrado alta especificidad y sensibilidad en la detección de hemoparásitos, convirtiéndose en herramientas diagnósticas habituales. No obstante, su incapacidad para la detección simultánea de múltiples patógenos representa una limitación.

La secuenciación de nueva generación (NGS) ha emergido como una alternativa robusta a las técnicas moleculares convencionales, al permitir la secuenciación profunda de regiones genéticas conservadas en grupos de organismos objetivo. Esta aproximación genera un 'metacódigo de barras' que facilita la identificación simultánea de todos los patógenos presentes en una muestra.

Dentro de las plataformas NGS, la tecnología Oxford Nanopore se destaca por su capacidad para la secuenciación rápida y a gran escala, sin requerir infraestructura costosa o especializada. Esta tecnología ofrece una identificación precisa y exhaustiva de apicomplejos, incluyendo piroplásmidos como *Babesia*, mediante la secuenciación directa de ADN, su versatilidad permite no solo la detección simultánea de múltiples patógenos, sino también la caracterización de variaciones genéticas y linajes de *Babesia*, aspectos cruciales para estudios epidemiológicos y filogenéticos (Huggins et al., 2024).

Basándonos en los resultados obtenidos por Muñoz Luzuriaga (2023), quienes exploraron la variabilidad genotípica de *Babesia spp.* mediante secuenciación de tercera generación, detectaron inicialmente el parásito mediante PCR, realizando amplificaciones específicas de regiones conservadas, para la identificación de *B.bovis*, se amplifica la región del gen RAP-1a (*Rhoptry-Associated Protein 1a*), utilizado como marcador molecular en estudios filogenéticos, y considerado un candidato potencial para el desarrollo de vacunas debido a su papel inmunogénico en el hospedador (Niu et al., 2016).; por otro lado para la identificación de la especie *B. bigemina*, los genes SPE1 (*Spherical Body Protein 1*) y AVA1 (*Apical Variable Antigen 1*) muestran variabilidad genética, lo que refleja la capacidad de *B. bigemina* para evadir la respuesta inmune del hospedador. SPE1 es un marcador para la caracterización de cepas, mientras que AVA1 es importante para entender los mecanismos de evasión inmunológica y persistencia del parásito (Mosqueda et al., 2023).

Algunos trabajos han analizado la caracterización molecular y diversidad genética a través de la técnica de PCR, con una prevalencia de infecciones por *B.bovis* del 2.58% y la de *B.bigemina* del 5.80% en muestras de sangre de ganado. El análisis filogenético identificó cinco clados para *B.bovis* y tres para *B.bigemina*, mostrando una estrecha relación genética con cepas de otros países (Srionrod et al., 2022).

Otros se han centrado en la identificación de genes diferencialmente expresados y su implicación en el ciclo de vida de *Babesia spp* (Capelli-Peixoto et al., 2022).

Así como se han utilizado herramientas como PiroplasmaDB y OrthoMCL para el análisis de genes específicos y alineamientos genómicos, complementadas con Tandem Repeats Finder para la identificación de regiones repetitivas. Este enfoque permitió detectar genes conservados y sus variantes, así como la realización de análisis funcionales detallados de las secuencias, ofreciendo una base sólida para el desarrollo de pipelines bioinformáticos orientados a la caracterización genética y funcional de *B. bigemina* y *B. bovis* (Ueti et al., 2021).

En Ecuador, no se han identificado trabajos que describan específicamente el desarrollo de pipelines bioinformáticos avanzados aplicados a *B. bovis* y *B. bigemina*. Sin embargo, existen investigaciones relacionadas que utilizan enfoques moleculares y bioinformáticos básicos; por lo que en la presente investigación se pretende realizar un procesamiento de datos crudos (NGS) y alineación al genoma, análisis estadístico para identificar diferencias significativas en expresión génica, comparaciones funcionales para inferir el rol biológico de los genes diferencialmente expresados. Siendo el objetivo final el análisis filogenético comparativo entre *B. bovis* y *B. bigemina*, usando marcadores genéticos clave. *La hipótesis que proponemos se basa en que las secuencias de B. bovis y B. bigemina extraídas por el pipeline desarrollado, mostrarán una estructura filogenética correlacionada con variantes intraespecie y con un patrón geográfico asociado a las tres zonas muestreadas.*

2. Materiales y Métodos:

Descarga de Datos

Se descargaron archivos FASTQ provenientes de los resultados obtenidos por Muñoz Luzuriaga (2023), correspondiente al estudio CT-BIO M971d/2023. Estos archivos contienen datos de secuenciación generados mediante la plataforma Oxford Nanopore (ONT) y corresponden a muestras de *B. bovis* y *B. bigemina*, que fueron encontradas en ganado bovino en 3 provincias de la Amazonía ecuatoriana.

Tabla 1. Clasificación de muestras con códigos de barra por provincia

Barcode	<i>Babesia bovis</i>	<i>Babesia bigemina</i>	Localidad
BC1	18	12	NAPO
BC2	4	21	NAPO
BC3	190	248	NAPO
BC4	595	17	ORELLANA
BC5	354	200	SUCUMBIOS
BC6	279	232	SUCUMBIOS
BC7	147	171	SUCUMBIOS
BC8	331	168	SUCUMBIOS

Control de calidad y trimado de los archivos de secuenciación

El procesamiento y análisis de los datos se realizaron en la plataforma Galaxy Australia (Abueg et al., 2024), la cual proporciona un entorno bioinformático accesible y reproducible. El control de calidad de los datos de secuenciación se llevó a cabo utilizando FastQC (Andrews, S., 2023) para evaluar la calidad de las lecturas. Además, Porechop (Wick, R., 2017) se empleó para eliminar y recortar adaptadores, mejorando la calidad de las secuencias. Finalmente, FastQC se utilizó nuevamente para verificar la calidad de las lecturas tras el procesamiento, asegurando datos óptimos mediante configuraciones de parámetros específicos.

Mapeo y alineamiento frente al genoma de referencia

Se seleccionaron las secuencias de referencia en formato FASTA para *B. bovis* (CM031752.1) y *B. bigemina* (LK391707.1) desde la base de datos del National Center for Biotechnology Information (NCBI). Las secuencias en formato FASTQ fueron alineadas con sus respectivos genomas de referencia utilizando el alineador Bowtie2 (Langmead & Salzberg, 2012). Esta herramienta permitió un mapeo preciso y eficiente de las lecturas, optimizando la identificación de regiones homólogas y facilitando el análisis posterior de la ubicación y detección de variantes genéticas. Tras la alineación de las secuencias, se utilizó Samtools view (Danecek et al., 2021) para convertir el archivo de alineación en formato SAM a BAM, reduciendo el tamaño del archivo y optimizando su manejo. Posteriormente, con Samtools sort, se ordenaron las lecturas en función de su posición genómica, mejorando la eficiencia en los análisis posteriores. Finalmente, Samtools view se empleó nuevamente para filtrar las lecturas alineadas de acuerdo con criterios específicos, asegurando que solo se conservarán aquellas con calidad suficiente.

Ensamblaje de fragmentos de genomas

Para el ensamblaje del genoma de referencia y de las secuencias analizadas, se utilizó el ensamblador SPAdes (Bankevich et al., 2012), como resultado se generaron scaffolds, los cuales

representan fragmentos ordenados y orientados del genoma reconstruido. SPAdes empleó gráficos de Bruijn para mejorar la organización y continuidad de estos scaffolds, asegurando una representación precisa tanto del genoma de referencia como de las secuencias procesadas. Se aplicaron parámetros específicos para optimizar la calidad del ensamblaje.

Implementación de Herramientas Bioinformáticas mediante Anaconda

Anaconda (Anaconda, Inc., 2020) permitió la instalación controlada de las herramientas necesarias, asegurando la compatibilidad entre paquetes y evitando conflictos en las versiones de software utilizadas en el análisis.

Curaduría de Secuencias de Referencia para el Análisis

Las secuencias de referencia de *B. bigemina* y *B. bovis* fueron obtenidas de la base de datos del National Center for Biotechnology Information (NCBI). Para *B. bigemina*, se utilizaron las siguientes secuencias correspondientes al fragmento génico **SPE 1-AVA 1**.

Tabla 2. Secuencias de referencia para *B. bigemina*

Acceso NCBI	Aislado	Descripción	Enzimas de restricción
OP174630.1	MASP	Fragmento genómico SpeI-AvaI	SPE 1-AVA 1
OP174627.1	SAI 2017 atenuado	Fragmento genómico SpeI-AvaI (cepa atenuada)	SPE 1-AVA 1
OP174626.1	SAI 2017 virulento	Fragmento genómico SpeI-AvaI (cepa virulenta)	SPE 1-AVA 1
ON862162.1	Veracruz buffalo 100.1	Fragmento genómico SpeI-AvaI aislado de búfalo	SPE 1-AVA 1
ON862158.1	Tabasco buffalo 57.1	Fragmento genómico SpeI-AvaI aislado de búfalo	SPE 1-AVA 1

mientras que para *B. bovis*, se seleccionó la secuencia del gen **Rap1a**

Tabla 3. Secuencias de referencia para *B. bovis*

Acceso NCBI	Aislado	Descripción	Enzimas de restricción
JF279443.1	Cuba	Aislado	rap1a
JX090146.1	Rodonia Cacoal	Aislado	rap1a
MT914605.1	Bb13 LE MAF 2017 Bb254	Aislado	rap1a
JX177355.1	Placido de Castro	Cepa	rap1a
JX177354.1	Nova Uniao	Cepa	rap1a
JX177353.1	Mancio Lima	Cepa	rap1a

Estas secuencias fueron empleadas como referencia en los análisis bioinformáticos posteriores.

Creación de la Base de Datos BLAST a partir de Secuencias de Referencia y Ensambladas

Tras la obtención de las secuencias de referencia, estas fueron integradas en un único archivo en formato FASTA, junto con las secuencias ensambladas previamente en Galaxy. Para la creación de la base de datos, se empleó la herramienta makeblastdb de BLAST+ (Camacho et al., 2009), incorporando tanto las secuencias de referencia de *B. bigemina* y *B. bovis* como los ensamblajes generados. Este procedimiento permitió estructurar una base de datos personalizada, optimizando la búsqueda y alineación de secuencias, lo que facilitó el análisis comparativo y la identificación de similitudes entre los genomas ensamblados y las referencias.

Análisis de Similitud de Secuencias con BLASTn

Para la identificación de similitudes entre las secuencias ensambladas y las referencias, se realizó un análisis utilizando BLASTn. Las secuencias de las muestras fueron consultadas contra la base de datos previamente creada, mediante la herramienta blastn de BLAST + (Camacho et al., 2009). Se establecieron criterios estrictos de búsqueda, configurando un umbral del 95% de identidad y 98% de cobertura. Esta configuración permitió detectar con alta precisión aquellas secuencias que presentaban un alto grado de similitud con las secuencias de referencia de *B. bigemina* y *B. bovis*, asegurando la fiabilidad de los resultados en el análisis filogenético posterior.

Optimización de Alineamientos BLASTn mediante AWK

Tras la ejecución de BLASTn, los resultados obtenidos fueron procesados utilizando la herramienta AWK. Mediante expresiones condicionales y patrones de búsqueda, se realizó un filtrado selectivo para extraer únicamente los alineamientos más relevantes, priorizando aquellas secuencias con la mayor identidad y cobertura. Este procedimiento optimizó la selección de datos, asegurando que solo las coincidencias de alta calidad fueran consideradas.

Alineamiento de secuencias y construcción de árboles filogenéticos

Se realizaron análisis filogenéticos para cada especie de *Babesia*, mediante matrices de alineamiento (MacVector, 18.7.6, ClustalW; GOP:30.0, GEP:10.0), utilizando 51 secuencias de ADN de *B. bovis* (31 secuencias) y *B. bigemina* (20 secuencias), de una longitud de 290 pb (*bovis*) y 170 pb (*bigemina*).

Para *B. bigemina*, se obtuvo un alineamiento en formato MSF (Multiple Sequence Format) con secuencias de 200 pb. Estas secuencias fueron comparadas con *B. bovis* como grupo externo para evaluar las relaciones filogenéticas entre los aislamientos locales.

En el caso de *B. bovis*, el alineamiento incluyó 280 pb por secuencia, con muestras locales principalmente de Napo, Sucumbíos y Orellana. Todas las secuencias fueron tratadas con el mismo peso (*Weight: 1.0*), garantizando una contribución equitativa en los análisis filogenéticos. La mayor longitud de estas secuencias permitió una mejor resolución en la inferencia de relaciones evolutivas dentro de *B. bovis*.

Estos alineamientos proporcionaron una base sólida para la reconstrucción de árboles filogenéticos, permitiendo evaluar la diversidad genética y las relaciones evolutivas entre las poblaciones locales de *B. bigemina* y *B. bovis*.

Para el análisis filogenético de *Babesia bovis*, se emplearon como secuencias externas (outgroup) dos secuencias del gen *rap-1* de *Babesia bigemina*: AB617644.1 y KF192811.1. En contraste, para el análisis de *B. bigemina*, debido a la ausencia de secuencias del marcador SpeI-AvaI correspondientes a *B. bovis*, el árbol fue enraizado mediante el método del punto medio (midpoint rooting), con el objetivo de identificar posibles patrones de divergencia entre las ramas principales.

Las construcciones filogenéticas se realizaron con Paup 4.05 build 169 (Swofford, 2021) con optimizaciones mediante Neighbor-Joining y Máxima Parsimonia, con 1.000 réplicas-adición al azar, TBR para barrido de ramas y búsqueda de islas de árboles, y repesado mediante índice de consistencia recalculado. El apoyo estadístico de ramas o clados se calculó por Bootstrapping de 1.000 pseudoréplicas.

3. Resultados

Tabla 4. Cambios en el número de lecturas de *B. bovis* y *B. bigemina* antes y después del filtrado y análisis bioinformático.

Barcode	<i>B. bovis</i>	<i>B. bigemina</i>		<i>B. bovis</i>	<i>B. bigemina</i>	Localidad
BC1	18	12		4	5	NAPO
BC2	4	21		1	5	NAPO
BC3	190	248		5	1	NAPO
BC4	595	17	F I L T R O	9	2	ORELLANA
BC5	354	200		3	3	SUCUMBIOS
BC6	279	232		1	1	SUCUMBIOS
BC7	147	171		4	1	SUCUMBIOS
BC8	331	168		4	2	SUCUMBIOS

Árbol o hipótesis filogenética de B. bigemina

El filograma fue construido utilizando secuencias del gen SpeI-AvaI, y enraizado mediante el método del punto medio (*midpoint rooting*), debido a la ausencia de secuencias de *Babesia bovis* para este marcador específico. Este tipo de enraizamiento permite una inferencia balanceada de divergencia entre las ramas, sin requerir un grupo externo definido.

El árbol resultante evidencia la presencia de dos clados principales, bien diferenciados:

Clado I

Este clado agrupa secuencias de referencia de *B. bigemina* procedentes de Portugal (FJ939724.1) e India (AB922127.1). Aunque no se utilizó como outgroup formal, este clado presenta un soporte bootstrap del 73% y se comporta filogenéticamente como un grupo externo relativo, separado de las secuencias de Ecuador. Esto sugiere una divergencia genética significativa, posiblemente relacionada con una evolución geográfica o presión de selección diferencial.

Clado II

Incluye exclusivamente secuencias de origen americano, incluyendo Ecuador y secuencias de México (bovinos y de búfalo), así como la secuencia de referencia *B. bigemina* SpeI-AvaI strain BOND. Este clado presenta una estructura interna compleja y está subdividido en cuatro subclados principales, cada uno con características filogenéticas particulares:

Clado II.1 – Bootstrap: 94%

- Presenta conexiones geográficas mixtas entre Ecuador y México.
- El alto soporte estadístico indica que la agrupación de estas secuencias es altamente confiable.
- Subdivisiones:
 - **II.1.1:** Agrupa la cepa mexicana OP174626.1 y la cepa de referencia BOND.
 - **II.1.2:** Incluye las secuencias de Napo y Sucumbíos (BC2.2, BC2.3, BC2.4 y BC8.1).
 - La homogeneidad genética entre estas secuencias sugiere un origen evolutivo reciente o compartido, posiblemente con intercambio genético regional.

Clado II.2 - Bootstrap: 100%

- Soporte estadístico máximo, lo cual refleja una agrupación altamente robusta.
- Agrupa secuencias de México (ON862154.1, OP174630.1) y de Orellana (BC4, BC4.1).
- Subdivisiones:
 - **II.2.1:** secuencias de México OP174630.1 (bovino) y ON862154.1 (búfalo).
 - **II.2.2:** Secuencias de Orellana (BC4, BC4.1).

Este subclado sugiere un posible linaje compartido entre Ecuador y México, aunque con una mayor variabilidad genética respecto al Clado II.1, lo que podría reflejar una diferenciación más temprana o eventos de recombinación.

Clado II.3 – Subclados con soporte mixto

- Este clado se subdivide en dos grupos con diferentes niveles de soporte:
 - **II.3.1 - Bootstrap: 55%**

- Incluye las secuencias BC3 (Napo), BC6, BC7 y BC8 (Sucumbíos).
- El soporte bajo - moderado sugiere que esta agrupación debe ser interpretada con precaución.
- La baja robustez podría deberse a una alta diversidad intra-regional o a posiciones filogenéticas ambiguas.
- **II.3.2 - Bootstrap: 96%:**
- Agrupa las secuencias BC5, BC5.1 y BC5.2 (todas de Sucumbíos).
- El alto soporte indica un linaje claramente establecido, posiblemente reflejando una expansión clonal en esta zona.

Este subclado sugiere una conexión evolutiva común entre cepas del oriente ecuatoriano, con especial énfasis en Sucumbíos.

Clado II.4 - Bootstrap: 88%

- Agrupa exclusivamente secuencias de Napo (BC1, BC1.1, BC1.2, BC1.3 y BC1.4).
- El alto valor de soporte indica un linaje genéticamente homogéneo y bien definido, lo cual podría reflejar una transmisión local mantenida o procesos de adaptación regional.

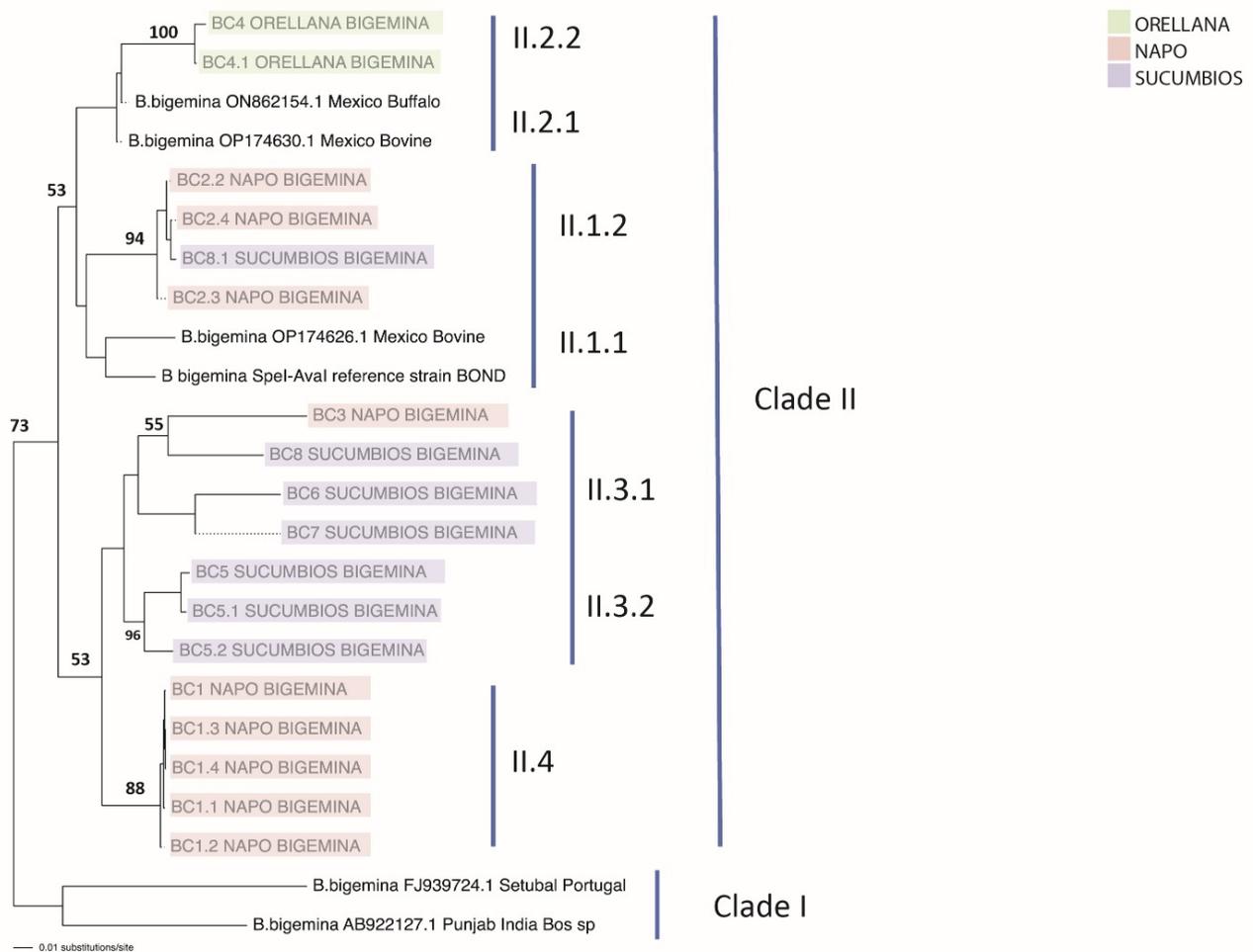


Fig.1: Árbol o hipótesis filogenética de *Babesia bigemina*

Árbol o hipótesis filogenética de *B. bovis*

La hipótesis filogenética muestra una topología estructurada de la siguiente forma: (FIG 2)

El árbol filogenético fue construido a partir de secuencias del gen RAP-1 de *Babesia bovis*, utilizando un modelo de sustitución apropiado y una topología rectangular para su visualización. El árbol fue enraizado utilizando secuencias de *Babesia bigemina* (AB617644.1 y KF192811.1) como grupo externo (*outgroup*), permitiendo establecer la serie de transformación evolutiva de las secuencias de *B. bovis*. La escala del árbol representa aproximadamente 0.005 sustituciones por sitio, lo cual indica una baja divergencia entre las secuencias analizadas.

Estructura general del árbol

El árbol evidencia un Clado principal de *B. bovis* con soporte estadístico del 100%, lo que confirma la monofilia de las secuencias dentro de la especie. Dentro de este clado, se identifican dos grandes agrupaciones:

- Subclado I: compuesto por secuencias de referencia y secuencias fuera de Ecuador.
- Subclado II: compuesto exclusivamente por secuencias de Ecuador.

Aunque estas subdivisiones presentan valores de soporte <50%, se observan tendencias filogenéticas marcadas, especialmente en la agrupación geográfica de las secuencias de Ecuador.

Subclado I – secuencias generales

- Agrupa secuencias de referencia y de diferentes regiones del mundo:
 - Egipto, Sudáfrica, México, Brasil, Paraguay, Cuba, Sri Lanka, Mongolia, Yak, entre otros.
- Incluye secuencias ampliamente utilizadas como controles:
 - *B. bovis* RAP-1 T2Bo (cepa de referencia)
 - OP174624.1 (cepa virulenta SAI 2017 – México)
- Presenta una diversidad genética mayor, con longitudes de rama ligeramente más extensas, lo que sugiere divergencia evolutiva acumulada.
- Este subclado cumple la función de comparativo filogenético, demostrando la cohesión de las cepas ecuatorianas como un grupo bien definido y separado.

Subclado II – secuencias de Ecuador

- Agrupa todas las secuencias provenientes de las provincias de Napo, Sucumbíos y Orellana.
- Este agrupamiento presenta un soporte interno del 100%, lo cual indica una agrupación robusta y confiable.
- Las ramas son cortas, lo que indica una alta homogeneidad genética entre las secuencias locales.
- A pesar del bajo valor de soporte entre ciertas bifurcaciones internas (como 60), se pueden observar microagrupaciones geográficas, por ejemplo:
 - Subagrupaciones de Napo (BC1, BC3, BC3.1, BC3.2, etc.)
 - secuencias de Sucumbíos (BC5, BC5.1, BC5.2, BC6, BC7.1, etc.)
 - Agrupaciones mixtas Napo–Sucumbíos–Orellana, posiblemente por circulación compartida.

Este subclado evidencia una baja variabilidad genética intraespecífica, lo que sugiere una expansión clonal local o circulación reciente del mismo linaje entre las tres provincias.

Bootstrap y distancias evolutivas

- Los valores bootstrap del árbol indican:
 - 100% para el clado *B. bovis*: soporte total.
 - 60% y <50% en algunos nodos internos: soporte moderado o bajo, lo que sugiere precaución en la interpretación de microclados.
- Las distancias genéticas (0.005 sustituciones/sitio) son muy pequeñas, lo cual respalda la baja diversidad molecular entre las cepas ecuatorianas.
- La separación entre subclados I y II se interpreta más como una tendencia topológica que una diferenciación estadísticamente fuerte, debido al bajo soporte en la rama que los divide.

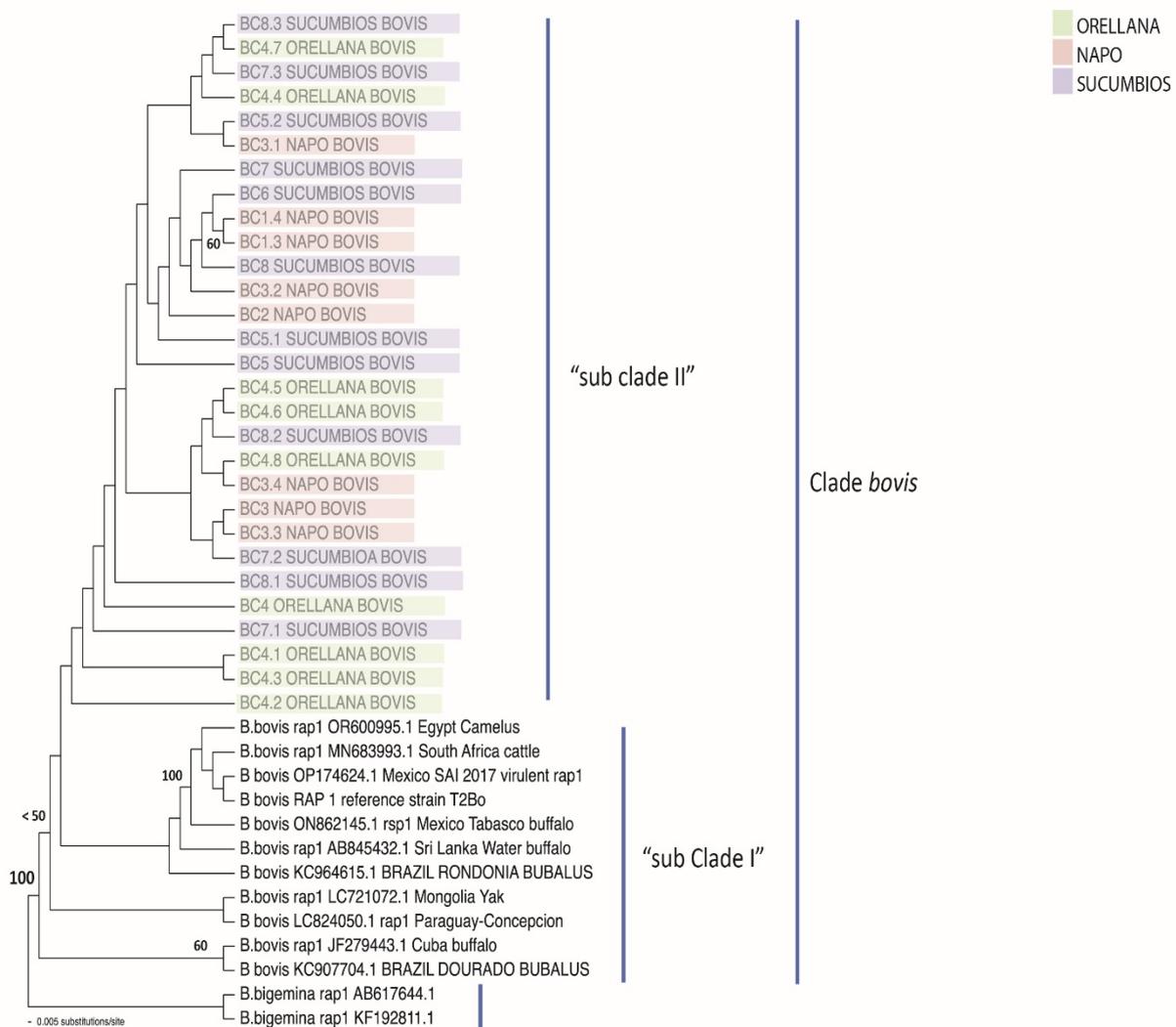


Fig.2: Árbol o hipótesis filogenética de *Babesia bovis*

4. Discusión

El desarrollo del pipeline bioinformático fue fundamental para asegurar la calidad de los datos utilizados en el análisis filogenético. A través de una secuencia estructurada de pasos, se logró procesar adecuadamente las lecturas crudas obtenidas por secuenciación con Oxford Nanopore, una tecnología que, aunque poderosa por su capacidad de lectura larga, requiere una depuración rigurosa por su tasa de error inicial.

El control de calidad con FastQC y el recorte de adaptadores con Porechop permitió conservar únicamente las lecturas de alta calidad. Posteriormente, el mapeo con Bowtie2 frente a genomas de referencia y el procesamiento con Samtools garantizaron una alineación precisa. El ensamblaje genómico con SPAdes generó secuencias representativas que fueron clave para los análisis posteriores.

Además, la curaduría de secuencias mediante la creación de una base de datos BLAST personalizada y el filtrado con AWK permitió seleccionar únicamente las secuencias con alta similitud a los genes diana (RAP-1a y SPE1-AVA 1). Este proceso aseguró que las secuencias utilizadas en la inferencia filogenética fueran específicas, completas y comparables entre muestras.

En conjunto, el pipeline implementado no solo permitió extraer información genética de alta calidad, sino que también estableció una metodología robusta y reproducible para futuros estudios filogenéticos y epidemiológicos de *Babesia spp.*

El análisis filogenético constituye una herramienta esencial para la comprensión de la diversidad genética y la dinámica evolutiva de parásitos hemoparásitos como *Babesia bovis* y *Babesia bigemina*, agentes etiológicos responsables de la babesiosis bovina en América Latina. En este estudio se generaron hipótesis filogenéticas a partir de secuencias parciales de los genes RAP-1 (*B. bovis*) y SpeI-AvaI (*B. bigemina*), obtenidas mediante secuenciación y análisis bioinformático, con el fin de evaluar la diversidad molecular de secuencias de Ecuador y su relación con cepas de referencia internacionales.

El árbol filogenético de *B. bovis*, enraizado con secuencias del gen *rap-1* de *B. bigemina* como outgroup, reveló un clado principal bien soportado (bootstrap: 100%), en el cual se observaron dos subclados claramente distinguibles: uno compuesto exclusivamente por cepas ecuatorianas (Subclado II), y otro integrado por cepas de origen internacional (Subclado I). A pesar de que el valor de soporte para la separación entre estos dos grupos fue inferior al 50%, la agrupación de las secuencias ecuatorianas fue altamente robusta y consistente.

El Subclado Ecuador (II) mostró una homogeneidad genética notable, con ramas cortas y valores bootstrap internos elevados, lo que sugiere una baja variabilidad genética entre cepas de las provincias de Napo, Sucumbíos y Orellana. Esta baja diversidad puede explicarse por una expansión clonal local o una reciente dispersión de un linaje común, probablemente impulsada por la movilidad del ganado y la distribución del vector *Rhipicephalus microplus* en la región amazónica del Ecuador.

En contraste, el Subclado I, conformado por cepas provenientes de Egipto, Sudáfrica, México, Brasil, Cuba, Paraguay, Sri Lanka, entre otros, presentó una mayor diversidad genética, evidenciada por la longitud de las ramas y la dispersión filogenética. Este patrón concuerda con estudios previos que reportan una elevada diversidad de *B. bovis* en regiones con una circulación prolongada del parásito y presencia de múltiples reservorios y vectores.

El árbol filogenético de *B. bigemina*, construido a partir del marcador SpeI-AvaI y enraizado mediante el método del punto medio, mostró también una clara estructuración evolutiva. Se identificaron dos clados principales: Clado I, conformado por cepas de Portugal e India, y Clado II, que incluyó a las cepas ecuatorianas y algunas cepas de México. La separación de estos clados, aunque respaldada por distancias genéticas, presentó soporte estadístico limitado (<70%), lo cual requiere una interpretación cautelosa.

Dentro del Clado II, se observaron cuatro subclados bien definidos con valores bootstrap entre 88% y 100%, lo que indica una alta confiabilidad estadística.

En comparación con *B. bovis*, las cepas de *B. bigemina* mostraron una mayor diferenciación intraespecífica y un patrón de agrupación más segmentado por provincia, lo cual podría estar relacionado con la evolución independiente de linajes locales o con la influencia de diferentes presiones ambientales y vectoriales.

Ambos árboles filogenéticos muestran que las cepas ecuatorianas de *Babesia spp.* forman agrupaciones definidas, bien soportadas y, en muchos casos, geográficamente coherentes. Sin embargo, existen diferencias clave entre las dos especies.

Los resultados filogenéticos aquí presentados aportan evidencia sólida sobre la existencia de microlinajes locales de *Babesia spp.* en Ecuador, con implicaciones relevantes para el control de la enfermedad. La baja variabilidad genética de *B. bovis* podría facilitar el diseño de estrategias diagnósticas y vacunales homogéneas, mientras que la mayor diversidad de *B. bigemina* podría representar un reto para la eficacia de vacunas y el monitoreo molecular.

Asimismo, la identificación de posibles relaciones evolutivas con cepas mexicanas en *B. bigemina* refuerza la necesidad de ampliar la vigilancia genética transfronteriza, considerando el comercio de ganado y la dispersión de vectores.

5. Conclusiones

El pipeline bioinformático desarrollado permitió establecer una metodología robusta, reproducible y específica para el procesamiento, ensamblaje y curaduría de datos genómicos de *B. bovis* y *B. bigemina* a partir de secuenciación con nanoporos. Cada etapa del flujo de trabajo, desde la descarga de datos hasta la selección de secuencias candidatas para análisis filogenético, fue optimizada para garantizar la calidad y confiabilidad de los resultados.

El uso de herramientas como FastQC y Porechop aseguró la depuración efectiva de las lecturas crudas, mientras que el mapeo con Bowtie2 y el procesamiento con Samtools permitió un alineamiento preciso con genomas de referencia. La implementación de SPAdes para el ensamblaje genómico aportó scaffolds consistentes que reflejan con mayor fidelidad las regiones de interés del genoma parasitario.

La integración de secuencias de referencia, junto con las ensambladas, en una base de datos personalizada mediante makeblastdb, seguida de un análisis de similitud con BLASTn y posterior filtrado con AWK, permitió seleccionar fragmentos genéticos altamente específicos y conservados, correspondientes a los genes RAP-1a y SPE1-AVA 1. Esto garantiza la selección de secuencias comparables y relevantes para el análisis evolutivo posterior.

Los análisis filogenéticos realizados a partir de los genes RAP-1 para *Babesia bovis* y SpeI-AvaI para *Babesia bigemina* permitieron identificar patrones evolutivos diferenciados entre ambas especies en cepas procedentes del nororiente ecuatoriano. Mientras que *B. bovis* mostró una homogeneidad genética notable, con un subclado de Ecuador bien definido y con bajo nivel de diferenciación interna, *B. bigemina* presentó una mayor diversidad filogenética, estructurada en varios subclados con soporte estadístico elevado y correlación geográfica marcada.

La existencia de clados bien soportados, tanto para *B. bovis* como para *B. bigemina*, respalda la hipótesis de la circulación de linajes locales y la posible expansión clonal regional, particularmente en las provincias de Napo, Orellana y Sucumbíos. La diferenciación parcial de algunas cepas ecuatorianas con respecto a secuencias internacionales, especialmente en *B. bigemina*, sugiere

procesos evolutivos activos y potenciales rutas de transmisión compartida entre Ecuador y otros países, como México.

Estos hallazgos refuerzan la importancia del monitoreo genético y la caracterización molecular de *Babesia spp.* como herramientas fundamentales para la vigilancia epidemiológica, el diseño de estrategias diagnósticas, y el desarrollo de vacunas más eficaces adaptadas al contexto local.

Referencia Citadas:

1. Abueg, L., Afgan, E., Allart, O., Awan, A., Bacon, W., Baker, D., Bassetti, M., Batut, B., Bernt, M., Blankenberg, D., Bombarely, A., Bretaudeau, A., Bromhead, C. J., Burke, M. L., Capon, P. K., Čech, M., Chavero-Díez, M., Chilton, J. M., Collins, T. J., ... Zoabi, R. (2024). The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic Acids Research*, 52(W1), W83–W94. <https://doi.org/10.1093/nar/gkae410>.
2. Anaconda, Inc. (2020). *Anaconda Software Distribution*. Version 2020.02. <https://www.anaconda.com/>.
3. Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
4. Bajer, A., Beck, A., Beck, R., Behnke, J., Dwuznik-Szarek, D., Eichenberger, R., Farkas, R., Fuehrer, H., Heddergott, M., Jokelainen, P., Leschnik, M., Oborina, V., Paulauskas, A., Radzijevskaia, J., Ranka, R., Schnyder, M., Springer, A., Strube, C., Tolkacz, K., & Walochnik, J. (2022). Babesiosis in Southeastern, Central and Northeastern Europe: An Emerging and Re-Emerging Tick-Borne Disease of Humans and Animals. In *Microorganisms* (Vol. 10, Issue 5). MDPI. <https://doi.org/10.3390/microorganisms10050945>.
5. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A., Dvorkin, M., Kulikov, A., Lesin, V., Nikolenko, S., Pham, S., Prjibelski, A., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, 19(5), 455–477. <https://doi.org/10.1089/cmb.2012.0021>.
6. Bock, R., Jackson, L., De Vos, A., & Jorgensen, W. (2004). Babesiosis of cattle. In *Parasitology* (Vol. 129, Issue SUPPL.). <https://doi.org/10.1017/S0031182004005190>.
7. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, 10(1), 421. <https://doi.org/10.1186/1471-2105-10-421>.
8. Capelli-Peixoto, J., Saelao, P., Johnson, W., Kappmeyer, L., Reif, K., Masterson, H., Taus, N., Suarez, C., Brayton, K., & Ueti, M. (2022). Comparison of high throughput RNA sequences between *Babesia bigemina* and *Babesia bovis* revealed consistent differential gene expression that is required for the *Babesia* life cycle in the vertebrate and invertebrate hosts. *Frontiers in Cellular and Infection Microbiology*, 12. <https://doi.org/10.3389/fcimb.2022.1093338>.
9. Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2), giab008. <https://doi.org/10.1093/gigascience/giab008>.
10. Gohil, S., Herrmann, S., Günther, S., & Cooke, B. (2013). Bovine babesiosis in the 21st century: Advances in biology and functional genomics. In *International Journal for Parasitology* (Vol. 43, Issue 2, pp. 125–132). <https://doi.org/10.1016/j.ijpara.2012.09.008>.
11. Huggins, L., Colella, V., Young, N., & Traub, R. (2024). Metabarcoding using nanopore long-read sequencing for the unbiased characterization of apicomplexan haemoparasites. *Molecular Ecology Resources*, 24(2). <https://doi.org/10.1111/1755-0998.13878>.
12. Kumar, A., O'bryan, J., & Krause, P. (2021). The global emergence of human babesiosis. In *Pathogens* (Vol. 10, Issue 11). MDPI. <https://doi.org/10.3390/pathogens10111447>.

13. Langmead, B., & Salzberg, S. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>.
14. Mischkulnig, M., Kiesel, B., Rötzer-Pejrimovsky, Th., Borkovec, M., Lang, A., Millesi, M., Wadiura, L. I., Hervey-Jumper, Sh., Penninger, J. M., Berger, M. S., Widhalm, G., & Erhart, F. (2022). The impact of heme biosynthesis regulation on glioma aggressiveness: Correlations with diagnostic molecular markers. *Frontiers in Molecular Neuroscience*, 15. <https://doi.org/10.3389/fnmol.2022.928355>.
15. Mosqueda, J., Hernandez-Silva, D., Ueti, M. W., Cruz-Reséndiz, A., Marquez-Cervantez, R., Valdez-Espinoza, U., Dang-Trinh, M., Nguyen, T., Camacho-Nuez, M., Mercado-Uriostegui, M., Aguilar-Tipacamú, G., Ramos-Aragon, J. A., Hernandez-Ortiz, R., Kawazu, S. I., & Igarashi, I. (2023). Spherical Body Protein 4 from *Babesia bigemina*: A Novel Gene That Contains Conserved B-Cell Epitopes and Induces Cross-Reactive Neutralizing Antibodies in *Babesia ovata*. *Pathogens*, 12(3). <https://doi.org/10.3390/pathogens12030495>.
16. Niu, Q., Bonsergent, C., Rogniaux, H., Guan, G., Malandrino, L., & Moreau, E. (2016). RAP-1a is the main rhoptry-associated-protein-1 (RAP-1) recognized during infection with *Babesia* sp. BQ1 (Lintan) (*B. motasi*-like phylogenetic group), a pathogen of sheep in China. *Veterinary Parasitology*, 232, 48–57. <https://doi.org/10.1016/j.vetpar.2016.11.013>.
17. Schäfer, R. A., Rabsch, D., Scholz, G. E., Stadler, P. F., Hess, W. R., Backofen, R., Fallmann, J., & Voß, B. (2023). RNA Interaction Format—A General Data Format for RNA Interactions. *Bioinformatics*, btad665. <https://doi.org/10.1093/bioinformatics/btad665>.
18. Srionrod, N., Noorong, P., Poolsawat, N., Minsakorn, S., Watthanadirek, A., Junsiri, W., Sangchuai, S., Chawengkirttikul, R., & Anuracpreeda, P. (2022). Molecular characterization and genetic diversity of *Babesia bovis* and *Babesia bigemina* of cattle in Thailand. *Frontiers in Cellular and Infection Microbiology*, 12. <https://doi.org/10.3389/fcimb.2022.1065963>.
19. Ueti, M. W., Johnson, W. C., Kappmeyer, L. S., Herndon, D. R., Mousel, M. R., Reif, K. E., Taus, N. S., Ifeonu, O. O., Silva, J. C., Suarez, C. E., & Brayton, K. A. (2021). Comparative analysis of gene expression between *Babesia bovis* blood stages and kinetes allowed by improved genome annotation. *International Journal for Parasitology*, 51(2–3), 123–136. <https://doi.org/10.1016/j.ijpara.2020.08.006>.
20. Vimonish, R., Capelli-Peixoto, J., Johnson, W., Kappmeyer, L., Saelao, P., Taus, N., Chung, Ch., & Ueti, M. (2025). Transcriptomic analysis of *Rhipicephalus microplus* hemocytes from female ticks infected with *Babesia bovis* or *Babesia bigemina*. *Parasites & Vectors*, 18(1), 37. <https://doi.org/10.1186/s13071-025-06662-w>.
21. Wick, R. R. (2017). *Porechop: Adapter trimmer for Oxford Nanopore reads* [Software]. <https://github.com/rrwick/Porechop>.

ANEXOS

Tabla A1. Datos de alineamiento MSF de *B. bigemina*

Código de acceso y nombre de la secuencia	Longitud (pb)	Check (Checksum)	Peso (Weight)
JF279443.1_Babesia_bovis	200	8852	1
JX090146.1_Babesia_bovis	200	7632	1
MT914605.1_Babesia_bovis	200	7504	1
BC8SUCUMBIOSBIGEMINA	200	5024	1
BC8.1SUCUMBIOSBIGEMINA	200	7280	1
BC7SUCUMBIOSBIGEMINA	200	5904	1
BC6SUCUMBIOSBIGEMINA	200	1513	1
BC5SUCUMBIOSBIGEMINA	200	247	1

BC5.1SUCUMBIOSBIGEMINA	200	3864	1
BC5.2SUCUMBIOSBIGEMINA	200	4379	1
BC4ORELLANABIGEMINA	200	6427	1
BC4.1ORELLANABIGEMINA	200	2605	1
BC3NAPOBIGEMINA	200	6969	1
BC2.2NAPOBIGEMINA	200	3669	1
BC2.3NAPOBIGEMINA	200	9272	1
BC2.4NAPOBIGEMINA	200	8132	1
BC1NAPOBIGEMINA	200	1592	1
BC1.1NAPOBIGEMINA	200	2224	1
BC1.2NAPOBIGEMINA	200	8128	1
BC1.3NAPOBIGEMINA	200	5766	1
BC1.4NAPOBIGEMINA	200	4318	1

Tabla A2. Datos de alineamiento MSF de *B bovis*

Nombre de la secuencia	Longitud (pb)	Check (Checksum)	Peso (Weight)
BC8SUCUMBIOSBOVIS	280	8878	1
BC8.1SUCUMBIOSBOVIS	280	8167	1
BC8.2SUCUMBIOSBOVIS	280	4078	1
BC8.3SUCUMBIOSBOVIS	280	8350	1
BC7SUCUMBIOSBOVIS	280	5749	1
BC7.1SUCUMBIOSBOVIS	280	7325	1
BC7.2SUCUMBIOSBOVIS	280	9800	1
BC6SUCUMBIOSBOVIS	280	9389	1
BC7.3SUCUMBIOSBOVIS	280	8331	1
BC5SUCUMBIOSBOVIS	280	8514	1
BC5.1SUCUMBIOSBOVIS	280	3268	1
BC5.2SUCUMBIOSBOVIS	280	6544	1
BC4ORELLANABOVIS	280	1398	1
BC4.1ORELLANABOVIS	280	5340	1
BC4.2ORELLANABOVIS	280	1401	1
BC4.3ORELLANABOVIS	280	806	1
BC4.4ORELLANABOVIS	280	7389	1
BC4.5ORELLANABOVIS	280	8536	1
BC4.6ORELLANABOVIS	280	7183	1
BC4.7ORELLANABOVIS	280	1291	1
BC4.8ORELLANABOVIS	280	3131	1
BC3NAPOBOVIS	280	5695	1
BC3.1NAPOBOVIS	280	7346	1
BC3.2NAPOBOVIS	280	5202	1
BC3.3NAPOBOVIS	280	584	1
BC3.4NAPOBOVIS	280	8051	1
BC2NAPOBOVIS	280	4113	1
BC1.3NAPOBOVIS	280	2558	1
BC1.4NAPOBOVIS	280	8396	1
OP174630.1_bigemina	280	4192	1
OP174627.1_bigemina	280	3928	1
OP174626.1_bigemina	280	9099	1