UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales.
Departamento de Ecología, Genética y Evolución

Modelos de epidemiología molecular para el estudio de la variabilidad y rutas de transmisión en virus de interés pecuario.

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en el área de Ciencias Biológicas.

Médica Veterinaria Zootecnista
Laura Camila Lozano Calderón

Director de tesis: Dr. Guido Alberto König
Directora Adjunta: Dra. Andrea Verónica Peralta
Consejero de Estudios: Dra. Verónica Lia

Instituto de Agrobiotecnología y Biología Molecular
(IABIMO- INTA- CONICET)
Fecha de defensa: 22 de marzo de 2024
Buenos Aires, Argentina.

Dedicatoría

La escritura de mi tesis está dedicada a:

- A Papá por ser todo.
- A mis papitos Martha y Andrés por ser mis grandes amores.
- A mi pariente uterina Diana por ser mi compañera de vida.
- A Jhon por darnos su amor siempre.
- A mis tíos, en especial a mi tia Lumiere por haber sido la causa de no desfallecer.
- A mi familia por ser mi motor.
- A mis plumas blancas por ser un ejemplo de vida, enseñándome no solo ciencia, si no tambien compromiso, trabajo en equipo, amistad y empatía.

Agradecimientos

Agradezco inmensamente:

- Al CONICET, por haberme dado los recursos para formarme como Doctora.
- Al INTA por ser un lugar de infinito aprendizaje.
- Al grupo Virus por su ayuda incondicional.
- Al grupo Bioenergía por toda la buena onda.
- A mis amiguitos Mechita y Juanito por ser los favoritos.
- A Escobar, a Julián, a Ale y a Jona por tanto amor.
- A la hermosa Argentina, por ser un lugar maravilloso.
- A todas las personas que fueron parte de mi proceso ¡MUCHAS GRACIAS!

Tabla de contenido

RESUMEN 8
ABSTRACT ..... **¡Error! Marcador no definido.**
INTRODUCCIÓN 11
Epidemiología Molecular 11
Métodos y herramientas utilizados en trabajos de Epidemiología Molecular 12
Recolección de datos epidemiológicos 12
Toma de muestras y recolección de datos de laboratorio 13
Secuenciación 13
Filogenias Moleculares 14
Modelos de sustitución 14
Metodología de Máxima Verosimilitud 15
Metodología Bayesiana 15
Filogeografía 16
Teoría de la Coalescencia 16
Filodinamia 16
Integración de Análisis Bayesianos con Filodinámica 17
Importancia de la Epidemiologia Molecular en la Salud Pública 17
Casos estudiados 18
Virus de la Fiebre Aftosa 19
Enfermedad Infecciosa de la Bursa 21
Síndrome Reproductivo y Respiratorio Porcino 23
Ectima Contagioso 25
PROBLEMÁTICA, HIPÓTESIS Y OBJETIVOS 28
Problemática 28
Hipótesis 28
Objetivos 28
Objetivo General 28
Objetivos Específicos 29
CAPITULO 1 - VFA 31
Materiales y Métodos 31
Selección de muestras, extracción de material genético y síntesis de ADNc., amplificación por PCR y secuenciación 31
Construcción de árboles y redes filogenéticas 37
Reconstrucción filogeografíca y análisis evolutivos 37
Resultados 38

Resultados del análisis de la epizootia en el partido Mar Chiquita 38
Resultados de la epizootia en la cuenca lechera..... 45
Discusión y conclusiones..... 47
Partido de Mar Chiquita 47
Cuenca lechera..... 49
Conclusiones generales del capítulo 1 50
CAPITULO 2 - OTROS VIRUS DE ARN..... 53
Sección 1 IBDV 53
Materiales y Métodos 53
Obtención de secuencias..... 53
Curado de secuencias y diseño del dataset enfocado a la determinación del genogrupo 53
Análisis filogenéticos enfocados a determinación de genogrupo..... 54
Diseño del dataset enfocado a conocer la historia evolutiva 54
Análisis de secuencias recombinantes..... 54
Análisis filogenéticos enfocados a conocer la historia evolutiva 54
Análisis filogeográficos y evolutivos 54
Resultados 55
Análisis filogenéticos enfocados a determinación de genogrupo..... 55
Análisis filogenéticos enfocados a conocer la historia evolutiva 56
Análisis filogeográficos y evolutivos 57
Discusión y conclusiones..... 58
Sección 2 PRRSV 60
Materiales y métodos 60
Datos de secuencia 60
Análisis filogenéticos enfocados a la determinación del genotipo 60
Identificación de secuencias recombinantes 60
Análisis filogenéticos enfocados a conocer la historia evolutiva 61
Reconstrucción de los estados ancestrales 61
Análisis evolutivos enfocados a conocer la historia y ancestria por ramas individuales 61
Resultados 62
Análisis filogenéticos enfocados a la determinación del genotipo 62
Análisis filogenéticos enfocados a conocer la historia evolutiva 62
Reconstrucción del estado ancestral más probable 69
Análisis evolutivos enfocados a conocer la historia y ancestria por ramas individuales 69
Discusión y conclusiones..... 70

Conclusiones generales del capítulo 2 73
CAPITULO 3 - ORFV 75
Materiales y métodos 75
Selección de muestras 75
Análisis filogenéticos, filogeográficos y evolutivos 78
RESULTADOS 79
Estadísticas de secuenciación y alineamiento 79
Análisis filogenéticos 79
Análisis evolutivos 86
Análisis filogeográficos 86
Discusión y conclusiones capítulo 3 87
Conclusiones generales de la tesis 90
Referencias bibliográficas 93

RESUMEN/ ABSTRACT

RESUMEN

Modelos de epidemiología molecular para el estudio de la variabilidad y rutas de transmisión en virus de interés pecuario.

La epidemiología molecular es una disciplina que utiliza herramientas de biología molecular con enfoques epidemiológicos. Esta disciplina emplea diversas metodologías, como la utilización de datos genómicos, taxonomía, filogenia y filogeografía, con el objetivo de comprender más a fondo la propagación y los determinantes de la transmisión de enfermedades infecciosas en poblaciones específicas. Actualmente, existen diversas problemáticas asociadas a la vigilancia y control de las enfermedades infecciosas presentes en campo que impiden proponer medidas eficaces que puedan hacer frente tanto a las endemias como a las epidemias presentes y futuras.

Basándose en este contexto, se planteó este trabajo de tesis centrado en el diseño de modelos de epidemiología molecular con el objetivo de identificar la dinámica y las vías de transmisión viral de enfermedades como la fiebre aftosa, la enfermedad infecciosa de la bursa, el síndrome reproductivo y respiratorio porcino, y el ectima contagioso. La investigación parte de la hipótesis de que la integración de datos epidemiológicos con las secuencias virales respectivas, analizadas a partir de modelos de epidemiología molecular, proporciona una comprensión más completa de la dinámica de transmisión y evolución de estos virus en diversos contextos epidemiológicos.

Este trabajo de tesis concluyó que los modelos de epidemiologia molecular permitieron identificar momentos importantes para intervenir durante una epidemia y mejorar así la eficacia de las medidas de control. Además, se evaluó la capacidad para identificar los factores que influyen en la transmisión y diseminación de las enfermedades mencionadas, incluyendo el análisis de vías de contagio por fómites. Se examinó el efecto de la variabilidad genética de las variantes virales circulantes y su relación con medidas preventivas como la vacunación. .

La integración de datos epidemiológicos con secuencias virales permitió una comprensión más profunda de la dinámica de propagación y evolución de los virus estudiados y esto posibilitó la formulación de nuevas hipótesis de dispersión espacial y de dinámica viral en el desarrollo de las epidemias.

Se sugiere la aplicación de estas herramientas en la toma de decisiones en salud pública veterinaria y humana para obtener resultados más efectivos y rápidos contra las enfermedades infecciosas, contribuyendo así a prevenir y controlar posibles brotes y proteger la economía y el desarrollo social de los países y las poblaciones afectadas. Se propone incorporar herramientas de epidemiología molecular en las operaciones rutinarias de los organismos de control encargados de preservar la salud pública veterinaria, con el fin de mitigar y reducir los riesgos sanitarios y biológicos.

ABSTRACT

Molecular epidemiology models for the study of variability and transmission routes in livestock viruses.

Molecular epidemiology is a discipline that utilizes molecular biology tools with epidemiological approaches. This discipline employs various methodologies, such as the use of genomic data, taxonomy, phylogeny, and phylogeography, aiming to gain a deeper understanding of the spread and determinants of transmission of infectious diseases in specific populations. Currently, there are various challenges associated with surveillance and control of infectious diseases in the field, hindering the proposal of effective measures to address both present and future endemic and epidemic situations.

Building upon this context, this thesis work was focused on designing molecular epidemiology models to identify the dynamics and viral transmission routes of diseases such as foot-and-mouth disease, infectious bursal disease, porcine reproductive and respiratory syndrome, and contagious ecthyma. This study is based on the hypothesis that integrating epidemiological data with respective viral sequences, analyzed in molecular epidemiology models, provides a more comprehensive understanding of the transmission dynamics and evolution of these viruses in various epidemiological contexts.

This thesis concluded that molecular epidemiology models allowed for identifying key intervention moments during an epidemic, thereby improving the effectiveness of control measures. Additionally, their capacity to identify factors influencing the transmission and spread of the mentioned diseases was evaluated, including the analysis of different transmission forms such as fomites. The effect of viral genetic variability and its relationship with preventive measures, such as vaccination. was also examined.
The integration of epidemiological data with viral sequences enabled a deeper understanding of the spread and evolution dynamics of the studied viruses, and thereof, enabling the formulation of new hypotheses regarding spatial dispersion and viral dynamics in an epidemic progress.

The application of these tools is suggested in the decision-making processes in veterinary and human public health to achieve more effective and rapid results against infectious diseases, thereby contributing to preventing and controlling potential outbreaks and safeguarding the economy and social development of affected countries and populations. We propose the incorporation of molecular epidemiology tools into the routine operations of control agencies responsible for preserving veterinary public health to mitigate and reduce sanitary and biological risks.

1 INTRODUCCIÓN

INTRODUCCIÓN

Epidemiología Molecular

La epidemiología, es una disciplina científica dedicada al análisis de la distribución y los factores determinantes de la transmisión de enfermedades en poblaciones específicas (CDC, 2018, disponible en https://www.cdc.gov/training/publichealth101/epidemiology.html). Se ha identificado que la epidemiología engloba más de 24 subespecialidades, siendo una de ellas la epidemiología molecular (Frérot et al., 2018).

La epidemiología molecular se define como una disciplina que utiliza las herramientas de la biología molecular para abordar casos dentro del contexto de la epidemiología tradicional. En este sentido, implica la utilización de material genómico, la taxonomía, la filogenia y los datos epidemiológicos para comprender la distribución y los factores determinantes que influyen en la aparición de enfermedades infecciosas en una población (Foxman and Riley, 2001).

El término "epidemiología molecular" fue descrito por primera vez en 1973 por Kilbourne y se utilizó inicialmente para analizar la distribución global del virus de la influenza (Kilbourne, 1973). Desde entonces, numerosos estudios han desempeñado un papel crucial en la resolución de interrogantes relacionados con la etiología y la patogénesis de diversas enfermedades. Estos trabajos han permitido desarrollar estrategias de tratamiento y medidas de prevención aplicables a brotes de diversos agentes infecciosos (Eybpoosh et al., 2017).

En términos generales, la epidemiología molecular responde a una serie de preguntas fundamentales, que van desde cómo y cuándo surgió una enfermedad en un lugar específico, cuáles son las vías de transmisión mediante las cuales se propaga la enfermedad, cuáles son los posibles agentes causales de una enfermedad y si distintos linajes virales pueden tener preferencias por diferentes hospedadores. También se investiga qué factores ambientales están relacionados con la infección, cómo influyen las características del huésped en la transmisión y, en el caso de enfermedades desconocidas, qué similitudes genéticas presenta el agente causal con otros ya estudiados. Todo esto contribuye a la comprensión de la enfermedad y apunta hacia la búsqueda de maneras de reducir la probabilidad de propagación de la misma. Esto es esencial para garantizar la salud tanto de seres humanos como de animales.

Las enfermedades virales juegan un papel importante en la epidemiología molecular, ya que los virus poseen una extraordinaria tasa de mutación y una rápida capacidad de replicación. Esto facilita un seguimiento preciso de las modificaciones en el genoma viral, que, junto con los datos epidemiológicos, permite una estimación más cercana a la realidad de la evolución temporal del virus. Además, estos estudios permiten realizar inferencias sobre aspectos espaciales, temporales y demográficos (Volz et al., 2013).

Un ejemplo destacado de la aplicación de la epidemiología molecular en el estudio de epidemias virales, se refleja en los múltiples trabajos generados para la compresión de la pandemia del SARS-CoV2, donde fue posible inferir que el origen del agente casual del Covid-19 fue resultado de eventos de recombinación de ancestros del SARS-CoV en murciélagos (Luk et al., 2019). Con respecto a la utilización de la epidemiologia molecular en

otros virus, la investigación realizada por Dellicour y sus colegas en 2019 permitió esclarecer la historia de dispersión y la dinámica del virus de la rabia (RABV) en Irán, confirmar la coexistencia de varios clados de virus circulantes en el país, calcular la velocidad de dispersión de estos linajes e identificar una tendencia de propagación hacia áreas cercanas con alta densidad de población humana (Dellicour et al., 2019).

Otro ejemplo destacado es el estudio realizado por (Rich et al., 2022) en la epidemia presente de HIV en los EEUU, donde fue posible identificar nueve grandes grupos de circulación del virus en Florida, que evolucionaron independientemente a los grupos identificados en el resto del país. Este estudio ayudó además a conocer la dinámica intraestatal que tiene el HIV en tiempo real.

Todos los avances descritos anteriormente han sido posibles gracias a la disponibilidad de modelos evolutivos diseñados para comprender la historia de las variaciones en las secuencias moleculares. Estos modelos analizan la tasa de sustitución de las secuencias de estudio, así como la inserción o eliminación de bases en función de la historia evolutiva. Además, consideran distintos modelos demográficos que una población viral pudo haber experimentado en relación con su entorno y su hospedero. Así, estos modelos permiten determinar el reloj molecular implícito en dichas secuencias, estableciendo la tasa de sustitución nucleotídica e identificando si la misma es constante o no en la población estudiada. Cuando se combinan todas estas herramientas con la información epidemiológica recopilada durante un brote en el campo, se obtiene una comprensión más completa de los patrones de dispersión y el comportamiento de las enfermedades causadas por virus.

Métodos y herramientas utilizados en trabajos de Epidemiología Molecular

En esta sección, nos enfocaremos exclusivamente en describir las técnicas más comúnmente utilizadas e implementadas para el desarrollo de esta tesis.

Recolección de datos epidemiológicos

La recopilación de datos representa el punto de partida en cualquier investigación epidemiológica. La identificación de los datos más pertinentes es esencial para generar una combinación de información efectiva. En el caso de enfermedades de declaración obligatoria, rastrear brotes y las características de las enfermedades que afectan a una región, se simplifica, dado que los informes deben mantenerse en resguardo en los departamentos de salud designados para esa área. Sin embargo, la recopilación de datos de enfermedades no sujetas a declaración obligatoria presenta desafíos adicionales ya que la información debe obtenerse a través de fuentes secundarias como historias clínicas, informes de veterinarios en granjas y la vigilancia epidemiológica por parte de la comunidad.

Acceder a datos confiables constituye el primer desafío para iniciar una investigación epidemiológica puesto que la calidad de los datos desempeña un papel crucial en el éxito del análisis epidemiológico (Rasmussen and Goodman, 2018).

Toma de muestras y recolección de datos de laboratorio

La selección del tipo de muestra a tomar en cada situación se rige por diversos factores vinculados a la enfermedad y los objetivos de investigación, como el tipo de tejido, el momento de la recolección, los recipientes, conservantes, medios de transporte y la duración del tránsito. Estos aspectos inciden directamente en la calidad de la información obtenida luego de realizar las correspondientes pruebas de biología molecular (Holland et al., 2003).

En estudios de dispersión viral, resulta importante examinar los distintos niveles en los que el virus puede actuar (tejido, individuo, rodeo, región, país, etc.) de manera independiente debido a las posibles variaciones en las tasas de evolución de las poblaciones virales (Belshaw et al., 2011). En distintos tipos de tejidos dentro de un animal, es posible identificar diferencias, no solo en la carga viral, sino también en los tipos de variantes virales presentes. Estas divergencias pueden vincularse al grado de exposición al sistema inmunológico al que la partícula viral ha estado expuesta. Por ejemplo, las partículas que ingresan al sistema circulatorio y generan viremia deben evadir al sistema inmunológico durante todo su tránsito por el torrente sanguíneo. Esto podría implicar que las variantes que colonicen futuras lesiones sean distintas entre sí y diferentes de las que ingresaron inicialmente al hospedador.

Este ciclo se reproduce en la población viral al transmitirse de un animal a otro. Por lo tanto, es importante identificar las variantes que pueden encontrarse a nivel intrahospedador, intrarrodeo y dentro de una epizootia, ya que de esta manera se genera una cantidad significativa de información filogenética que facilita la identificación de patrones de diseminación y la determinación de ancestros en común (Ramirez-Carvajal et al., 2018).

Secuenciación

La secuenciación del genoma del agente causal es el principio de los análisis en epidemiología molecular y es un componente más de todos los datos que se deben analizar. Uno de los desafíos de la disciplina radica en la selección del segmento del genoma que resultará más informativo para el estudio en cuestión. Es fundamental comprender el contenido genómico y las preguntas que se pueden abordar antes de emprender la secuenciación, por ejemplo, seleccionar una parte del genoma que contenga moderados caracteres variables permitirá una comparación entre diversas variantes del agente etiológico de estudio (Brown, 2002). Aunque la secuenciación del genoma completo (WGS) se ha convertido en la norma para los estudios filogenéticos y filogeográficos en las epidemias modernas, no se puede subestimar la importancia de analizar genes individuales. Aunque es innegable que la WGS brinda una visión más detallada de la estructura genética del agente causal, ofreciendo ventajas como una mayor resolución en los análisis. En los casos de enfermedades desatendidas que no reciben la financiación adecuada y según el objetivo del estudio, la selección de uno o más genes informativos es suficiente para poder comenzar a hipotetizar sobre la dinámica de transmisión y evolución de la enfermedad en tiempo real y, posteriormente, generar medidas efectivas de control (Holtz et al., 2023).

La elección de la metodología de secuenciación también se convierte en un factor de gran relevancia que requiere un análisis cuidadoso. Existen diversas metodologías disponibles y la selección apropiada dependerá de los objetivos específicos del estudio. Aunque técnicas como Sanger y NGS han demostrado una alta correlación en los resultados, la metodología

de NGS brinda la capacidad de detectar con mayor detalle mutaciones mayoritarias y minoritarias puntuales (Arias et al., 2018) que pueden quedar enmascaradas al emplear la metodología de Sanger. Si el objetivo es realizar un estudio sobre la dinámica de transmisión de un virus a nivel local, que implica principalmente una cantidad baja de mutaciones (Cottam et al., 2008), la metodología de NGS es la opción más idónea para el análisis.

Filogenias Moleculares

La filogenia molecular es el estudio de las relaciones evolutivas entre organismos o genes. En el pasado, el análisis filogenético se basaba exclusivamente en la comparación morfológica de fósiles, pero esta fuente de información tenía limitaciones evidentes. Sin embargo, con el descubrimiento del ADN y el desarrollo de metodologías como la secuenciación, el análisis filogenético molecular se volvió posible y aportó mucho a la comprensión de las relaciones evolutivas.

En el análisis filogenético molecular, los cambios en las secuencias nucleotídicas desempeñan un papel fundamental. Estos cambios se utilizan para agrupar genes relacionados y reconstruir la historia evolutiva del agente en estudio. Una filogenia es esencialmente un modelo de la historia genealógica representado en forma de árboles. Estos árboles contienen nodos conectados por ramas, donde cada rama representa una secuencia genética y cada nodo simboliza el punto de divergencia o nacimiento de una nueva línea evolutiva. (Yang and Rannala, 2012).

Los métodos de reconstrucción filogenética se dividen en dos categorías: aquellos basados en distancias y los basados en caracteres. En los métodos basados en distancias, se calcula la distancia entre cada par de secuencias, y la matriz de distancias resultante se emplea para la reconstrucción del árbol filogenético. Por otro lado, los métodos basados en caracteres evalúan carácter por carácter de todas las secuencias en el alineamiento y calculan un puntaje para cada árbol construido. En teoría, se debería seleccionar el árbol con el mejor puntaje como el resultado final (Brinkman and Leipe, 2001).

Por ser mucho más precisa, en el desarrollo de esta tesis se optó por utilizar exclusivamente metodologías basadas en caracteres. Entre estas, las metodologías de Máxima Verosimilitud y Bayesiana fueron las más empleadas, ya que están respaldadas por un cálculo de modelo evolutivo que permite estimar estadísticamente la probabilidad de mutaciones específicas y rastrear con mayor detalle la historia evolutiva.

Modelos de sustitución

Los modelos de sustitución plantean una suposición específica referente a las tasas de mutación de las secuencias alineadas. La elección del modelo adecuado es esencial, ya que los modelos básicos pueden vulnerarse con facilidad, mientras que los modelos sobrestimados pueden resultar en un costo bioinformático innecesario (Sullivan and Joyce, 2005). Se ha observado que una especificación incorrecta del modelo de sustitución puede sesgar las inferencias filogenéticas (Arenas, 2015) y es por este motivo que este paso se convirtió en necesario para construir los árboles filogenéticos basados en la metodología de Máxima Verosimilitud y Estadística bayesiana.

Los modelos de sustitución consideran parámetros como la probabilidad de cada transición o transversión, la distribución de la probabilidad de cambio a lo largo de la secuencia, la cantidad de sitios sin variaciones, la frecuencia relativa de cada base, la posición de la variación dentro de la secuencia, etc.

Es importante señalar que, además de lo expuesto anteriormente, el proceso evolutivo de un virus puede variar entre regiones genómicas debido a la naturaleza heterogénea de la evolución molecular (Shapiro et al., 2006). Como resultado, es necesario modelar la evolución del genoma considerando sustituciones y modelos específicos para cada región.

Basándonos en la información previamente expuesta, existen modelos fundamentales y básicos, de los cuales se desprenden los más actuales y perfeccionados. Entre los modelos fundamentales podemos destacar, el modelo básico que es el de Jukes Cantor (JC) reportado en 1969, que considera una tasa de cambio uniforme entre todos los nucleótidos y las frecuencias de los nucleótidos son iguales (Jukes and Cantor, 1969). En un nivel de complejidad superior se encuentra el modelo K80, donde las tasas de cambio difieren entre transiciones y transversiones (Kimura, 1980). Por último, el modelo más complejo es el modelo (GTR) descrito por Tavaré, 1986, que incorpora diferentes velocidades para cada tipo de cambio y frecuencias de nucleótidos diversas (Tavaré, 1986). Una vez estimado el modelo de sustitución que más se adecua al set de datos de trabajo, es posible inferir el árbol filogenético (King and Van Doorslaer, 2018)

Metodología de Máxima Verosimilitud

El análisis de Máxima Verosimilitud implica la formulación de diversas hipótesis sobre una misma historia evolutiva. A partir de esta premisa, se evalúan los árboles en función de la probabilidad de que la hipótesis (es decir, el árbol) haya generado los datos de secuencia observados. El árbol de Máxima Verosimilitud corresponde a la hipótesis con la probabilidad más alta de haber observado dichos datos (Strimmer and Von Haeseler, 1997). Luego de que se estimaron los árboles posibles, se realiza una búsqueda en todo el espacio de árboles y se selecciona la estimación con el mejor puntaje de MLE (Estimación de Máxima Verosimilitud) (Yang and Rannala, 2012).

Metodología Bayesiana

La inferencia Bayesiana es una formalización matemática de un proceso de toma de decisiones basado en un análisis de probabilidad. Puede ser mucho más preciso que los métodos estadísticos clásicos, ya que a medida que la información disponible cambia, es más probable que cambie la evaluación de las probabilidades finales. Por lo cual, antes de que comience el análisis, se deben especificar las creencias previas que se tienen sobre estas relaciones, lo que se conoce comúnmente en estadística bayesiana como *priors*. Utilizando esta metodología estadística en filogenética, datos como el modelo evolutivo o topologías construidas anteriormente mediante otra metodología, pueden funcionar como *priors* para lograr estimar el árbol final seleccionado. El árbol elegido, al igual que el de Máxima Verosimilitud será seleccionado por medio de un puntaje, el mismo estará designado por la probabilidad posterior (PP) que especifica la probabilidad de cada árbol dados los *priors*. Los árboles con los valores posteriores más altos, serán seleccionados (Ronquist et al., 2009).

Filogeografía

Los análisis filogeográficos representan un enfoque esencial en el campo de la ecología molecular, ya que intentan establecer conexiones entre los acontecimientos históricos de la evolución y la distribución geográfica de los virus. Estos análisis están estrechamente ligados a la reconstrucción de la filogenia, ya que las inferencias filogeográficas se basan, en gran medida, en esta estructura. En el contexto de la virología, su objetivo principal es trazar el patrón de dispersión viral a lo largo del espacio y el tiempo. El propósito fundamental de un análisis filogeográfico es inferir las ubicaciones ancestrales en un árbol filogenético, aprovechando la información sobre las ubicaciones geográficas asociadas a cada secuencia. De esta manera, se permite extraer conclusiones acerca de cómo las poblaciones virales se han diseminado y adaptado a lo largo de su historia evolutiva en diversas regiones geográficas (Lemey et al., 2009).

Teoría de la Coalescencia

La teoría de coalescencia es un enfoque que vincula la estructura genealógica y la topología de un árbol que representa una muestra aleatoria de individuos, como en el caso de secuencias virales, con el tamaño de la población de la que provienen (Kingman, 1982). Este enfoque modela el tiempo de coalescencia de dos linajes (t) en función del tamaño de la población (N). Se centra en una pequeña muestra de una población infectada y describe cómo los linajes retroceden en el tiempo, fusionándose hasta llegar al ancestro común más reciente de todos (Volz et al., 2009). La combinación del modelo coalescente de Kingman con la inferencia bayesiana es especialmente valiosa para estimar tanto la filogenia de las secuencias virales como la historia demográfica de la población viral de manera simultánea (Drummond et al., 2005).

Filodinamia

El término filodinamia viral se emplea para referirse al análisis de los procesos epidemiológicos y evolutivos que tienen influencia en la formación de las filogenias virales. La investigación relacionada con la filodinamia viral se ha centrado en comprender la dinámica de la transmisión, con el objetivo de aclarar cómo estas interacciones afectan la variabilidad genética de los virus (Magiorkinis et al., 2013). A su vez, su estudio se basa la teoría de la coalescencia y la filogeografía, entre otras metodologías.

Hay aspectos clave utilizados para analizar los patrones de comportamiento de los virus mediante árboles filogenéticos. Uno de ellos es la longitud de las ramas del árbol. En un contexto donde un virus se propaga rápidamente en una población, el árbol filogenético puede mostrar una estructura "en forma de estrella" en la que las ramas exteriores son significativamente más largas que las ramas internas. En contraste, en una población viral que mantiene un tamaño constante, las ramas externas serán más cortas en comparación con las ramas internas del árbol (Volz et al., 2013).

Otro punto importante a considerar es cómo la agrupación de los taxones en una filogenia viral puede estar influenciada por la estructura de la población del hospedador. En este sentido, es probable que tanto los virus que residen en hospedadores similares como aquellos que se encuentran en la misma región geográfica estén genéticamente más relacionados entre sí y se observen en lugares cercanos en la filogenia. Por último, se destaca que el

equilibrio del árbol filogenético puede ser afectado por la presión de selección. Por ejemplo, un árbol filogenético en forma de escalera puede reflejar características de una fuerte selección direccional, mientras que un árbol más equilibrado puede ser indicativo de un virus que no está experimentando una presión de selección significativa (Volz et al., 2013).

Actualmente, se han desarrollado métodos estadísticos avanzados que facilitan la obtención de filogenias al incorporar datos adicionales además de las secuencias genéticas como, por ejemplo, la información temporal (Rambaut, 2000). Estos métodos han sido combinados con modelos coalescentes provenientes de la genética de poblaciones (Pybus et al., 2000) con el fin de esclarecer los patrones de dinámica viral (Lemey et al., 2009).

Integración de Análisis Bayesianos con Filodinámica

En la actualidad, existen herramientas informáticas que posibilitan la fusión de análisis estadísticos bayesianos para poner a prueba hipótesis de modelos evolutivos basándose en datos epidemiológicos y en secuencias moleculares. El software más ampliamente utilizado para esta tarea es el "Bayesian Evolutionary Analysis by Sampling Trees" (BEAST). Esta herramienta no solo facilita la construcción de árboles filogenéticos, sino que también, a través de la teoría de coalescencia y el algoritmo Metropolis-Hastings (MCMC), integra datos temporales y geográficos. Esto posibilita la obtención de estimaciones filogenéticas datadas, así como de parámetros y estados ancestrales, permitiendo probar hipótesis evolutivas sin limitarse únicamente a una topología del árbol filogenético (Drummond and Rambaut, 2007).

Importancia de la Epidemiologia Molecular en la Salud Pública

El concepto de "Una sola Salud" tuvo su origen en 2004, cuando la Wildlife Conservation Society (WCS) convocó un congreso dedicado a la medicina de la conservación. Este encuentro se denominó "Un Mundo, Una Salud" y tuvo como objetivo central la formulación de una estrategia global que se centrara en la integración de tres componentes fundamentales para salvaguardar la salud a nivel mundial. En primer lugar, se contempló la salud humana; en segundo lugar, la salud animal; y, por último, la salud ambiental. Estos componentes se identificaron como elementos cruciales para la comprensión y la gestión de los riesgos sanitarios a nivel global (Canals and Cáceres, 2020).

“Una salud” se basa en los pilares conocidos como los "Principios de Berlín", que en términos generales buscan fomentar un enfoque integral y proactivo en el abordaje de enfermedades emergentes y remergentes. Esto implica considerar las interrelaciones entre las diferentes especies y buscar la convergencia entre la conservación de la biodiversidad, el bienestar animal y las necesidades humanas (Gruetzmacher et al., 2021). Cuando alguno de estos componentes se ve afectado, la salud de los demás integrantes también corre peligro. En otras palabras, la salud humana depende de la salud de los animales y del medio ambiente, ya que se requieren recursos naturales saludables para la producción de alimentos y el bienestar de la población (Al-Eitan et al., 2023). A su vez, la salud animal depende de la inocuidad en los alimentos, el correcto manejo de las producciones y la ausencia de antropozoonosis.

Si la salud de los animales se ve comprometida, pueden surgir graves consecuencias, como la propagación de enfermedades zoonóticas, que representan más del 60% de las

enfermedades infecciosas que afectan a los seres humanos y pueden tener un impacto devastador a nivel mundial, como se evidenció durante la pandemia de COVID-19 (Esposito et al., 2023). Por otro lado, si la salud animal se deteriora, la seguridad alimentaria se ve amenazada, ya que se reduce la disponibilidad y la distribución de alimentos, lo que contribuye directa e indirectamente a la inseguridad alimentaria a nivel global. Además, es importante destacar que muchas poblaciones en regiones con inseguridad alimentaria dependen directamente de la agricultura para su sustento (Cunningham et al., 2017). Esto también tiene un impacto significativo en el entorno social y económico de las comunidades, ya que puede dar lugar a aumentos desproporcionados en los precios de los alimentos, la aparición de la pobreza e incluso conflictos sociales como guerras (Ejeromedoghene et al., 2020).

Por los motivos expuestos, es de vital importancia que la medicina en salud pública tenga como prioridad la preservación de la sanidad animal. En el cumplimiento de esta responsabilidad, resulta esencial implementar medidas efectivas de control y prevención de enfermedades infecciosas de origen veterinario para responder ágilmente ante la aparición de brotes y minimizar así el impacto negativo en el conjunto de la población.

Disciplinas como la epidemiología, son de suma importancia en el ámbito de la salud pública, ya que la misma, se apoya en herramientas matemáticas y predictivas para prevenir y controlar enfermedades en poblaciones humanas y animales. Su capacidad de discernir patrones específicos en la propagación de enfermedades, analizar factores de riesgo y evaluar la eficacia de las decisiones adoptadas durante una epidemia la convierte en una herramienta fundamental a la hora de pensar en salvaguardar la sanidad de una comunidad. (Bartlett and Judge, 1997).

No obstante, a pesar de los notables beneficios que los estudios epidemiológicos tradicionales han aportado en la prevención y control de muchas enfermedades, persisten importantes incógnitas relacionadas con los mecanismos biológicos y genéticos que pueden influir en el desarrollo y evolución de enfermedades infecciosas. Como respuesta a estas interrogantes, la epidemiologia molecular, que integra datos generados por otras disciplinas como la biología molecular y la genética, da lugar a la formulación de hipótesis más profundas sobre la dinámica de las enfermedades infecciosas en el campo.

Casos estudiados

En el proceso de desarrollo de esta tesis, se eligieron diferentes enfermedades veterinarias en las cuales se hipotetiza que los modelos de epidemiología molecular pueden arrojar luz sobre interrogantes que hasta el momento permanecen sin respuesta. A continuación, se proporcionará una introducción de cada una de las enfermedades seleccionadas.

Virus de la Fiebre Aftosa

Agente Etiológico

La Fiebre Aftosa (FA) es ocasionada por el virus de la Fiebre Aftosa (VFA), perteneciente a la familia *Picornaviridae* y particularmente al género Aphthovirus. Este virus, desprovisto de envoltura, se compone de cuatro proteínas estructurales que se ensamblan en una cápside icosaédrica: VP4, VP2, VP3 y VP1. Su genoma es una molécula de ARN monocatenario de polaridad positiva con una longitud aproximada de 8,500 nucleótidos. En su extremo 3', presenta un tramo de poli(A) que permite su traducción sin la necesidad de transcripción previa. Además, incorpora una pequeña proteína llamada "VPg", la cual se une covalentemente a su extremo 5', que cumple la función de reemplazar la presencia del cap en el genoma (Lei et al., 2015).

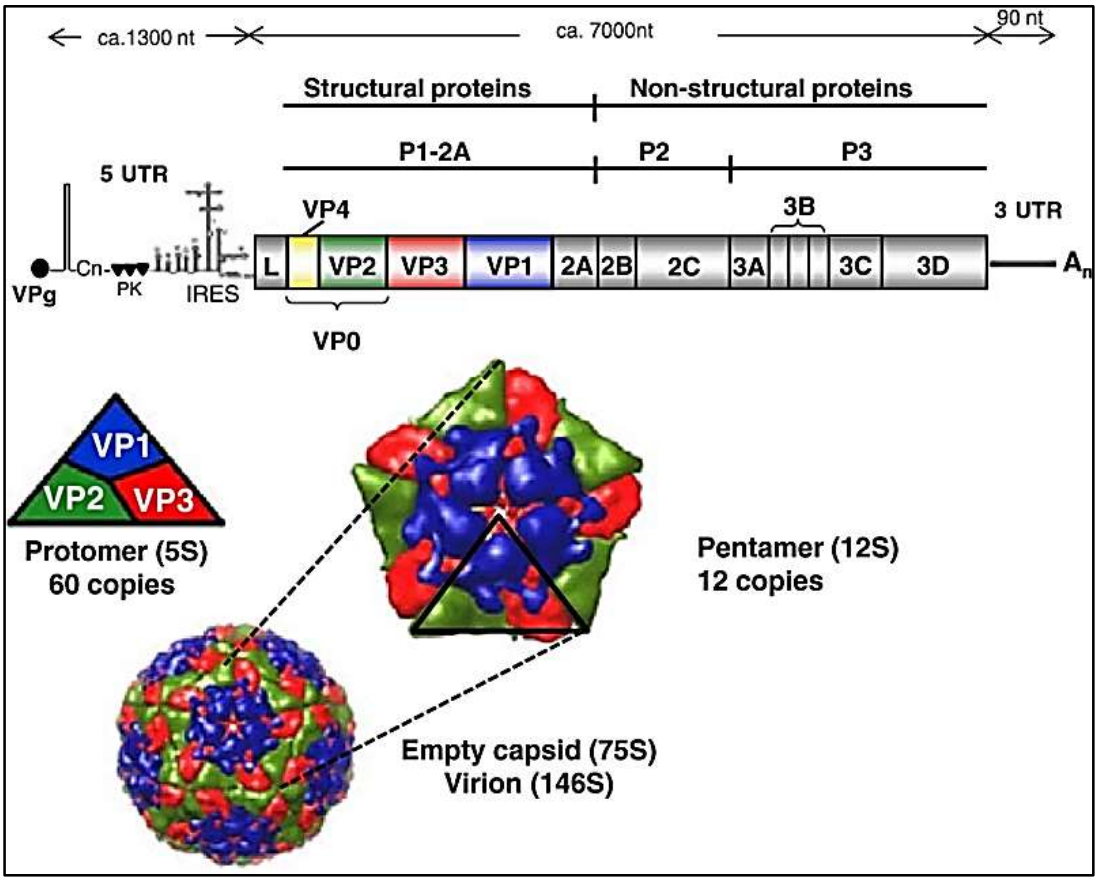

Figura 1. Representación gráfica del genoma y la estructura del virus de la Fiebre Aftosa (Jamal and Belsham, 2013)

Serotipos

El VFA se distingue por su destacada diversidad serológica, con la identificación de siete serotipos distintos (A, O, C, SAT-1, SAT-2, SAT-3 y Asia-1). Todos estos serotipos generan una enfermedad que, desde el punto de vista clínico es indistinguible, pero desde una perspectiva inmunológica presenta diferencias significativas. Es importante destacar que la infección con un serotipo no confiere inmunidad frente a otro (Rweyemamu et al., 2008).

Hospederos y signos clínicos

El VFA tiene la capacidad de infectar a una amplia variedad de animales, habiéndose documentado la presencia de la infección en más de 70 especies caracterizadas por poseer pezuñas hendidas. Entre los animales afectados se incluyen cerdos, vacas, ovejas, cabras y búfalos. La Fiebre Aftosa se manifiesta en estos animales mediante la aparición de un cuadro clínico que inicialmente desencadena fiebre. Posteriormente, se observa la formación de lesiones vesiculares, que tienden a desarrollarse principalmente en las mucosas del hocico, lengua, labios, cavidad oral, glándulas mamarias y espacios interdigitales. La gravedad de los signos clínicos depende de la cepa viral, la edad de los animales y la especie afectada (Li et al., 2021).

Transmisión

La infección natural generalmente se produce por el contacto directo entre animales infectados, siendo este el método de transmisión más significativo. La transmisión indirecta a través de fómites, como alimentos y maquinaria, también desempeña un papel epidemiológico considerable. En el caso del ganado porcino está documentada la propagación mediante la ingestión de alimentos contaminados. Además, se debe tener en cuenta la posibilidad de propagación aérea de partículas virales cuya extensión, dependiendo de factores ambientales, podría alcanzar hasta 60 km (Björnham et al., 2020).

Distribución Geográfica

La Fiebre Aftosa fue identificada por primera vez en el año 1870 y fue introducida en Sudamérica por inmigrantes ganaderos europeos cuando establecieron sus rebaños (Samuel and Knowles, 2001). En la actualidad, Australia, Europa y América del Norte se mantienen libres de la FA. En contraposición, la situación epidemiológica en Sudamérica presenta una complejidad mayor, a pesar de los esfuerzos que cada país ha realizado en términos de vigilancia y control. Desde el año 2000, se han registrado varios episodios notificados a la Organización Mundial de Sanidad Animal (OIE). Países como Argentina, Bolivia, Brasil, Colombia, Ecuador, Paraguay, Uruguay y Venezuela han reportado casos hasta la fecha (OIE, 2024).

Actualmente, Argentina es libre de la enfermedad (sin vacunación en la Patagonia y con vacunación en el resto del territorio). Sin embargo, esto contrasta con su pasado ya que, desde su introducción, esta enfermedad ha tenido un impacto significativo en la producción animal en la región y en el país en particular. En 1999, luego de emprender una campaña de erradicación, Argentina logró ser reconocida como un país libre de la enfermedad sin necesidad de aplicar vacunas. Sin embargo, en el año 2000, esta situación cambió drásticamente. Entre los años 2000 y 2001, se produjo una de las epidemias más devastadoras de FA, causada principalmente por un virus de serotipo A.

La enfermedad tuvo un impacto significativo en áreas con una alta concentración de ganado, afectando especialmente aquellos animales menores de dos años. Las primeras medidas de control fueron implementadas por el Servicio Nacional de Sanidad y Calidad Agroalimentaria (SENASA) e incluyeron sacrificios, vacunación en anillo, restricciones en el movimiento de animales y otras medidas preventivas. En un principio, se inició una campaña de vacunación de emergencia con las vacunas disponibles. Posteriormente, se desarrolló y aplicó una vacuna tetravalente que paulatinamente incorporó las nuevas cepas circulantes

A/Argentina/00 y A/Argentina/01 (Mattion et al., 2004). Durante un período que abarcó desde el 15 de mayo al 15 de julio de 2001, se reportaron 1549 granjas sospechosas de tener animales contagiados con FA en la provincia de Buenos Aires. De estas, 66 pertenecían al partido de Mar Chiquita. Por otra parte, la provincia de Santa Fe notificó un total de 197 casos sospechosos desde el 15 de marzo hasta el 25 de abril de 2001, mientras que Córdoba registró 97 granjas sospechosas desde el 14 de marzo hasta el 29 de noviembre del mismo año.

A pesar de que en la actualidad la Argentina y la mayoría de los países de América del Sur están libres de FA, el riesgo de futuras epidemias persiste debido al comercio legal e ilegal de animales y productos derivados, entre otros factores.

Importancia Económica y Epidemiológica

La FA es catalogada como la enfermedad animal de mayor relevancia económica (Domingo et al., 2002). Esta patología conlleva a una disminución en la productividad que puede llegar hasta el 25%. Además de esto, se deben tener en cuenta los considerables gastos asociados con la implementación de estrategias de control y erradicación. Las repercusiones económicas más significativas se derivan de las restricciones impuestas por naciones libres de la enfermedad a la importación de carne y productos derivados de animales provenientes de áreas afectadas.

El avance en la comprensión de la epidemiología de la Fiebre Aftosa es de suma importancia ya que proporciona la base para implementar medidas efectivas de control y erradicación de futuros brotes.

Enfermedad Infecciosa de la Bursa

Agente Etiológico

La Enfermedad Infecciosa de la Bursa (IBD) es causada por un virus no envuelto de estructura icosaédrica, perteneciente al género Avibirnavirus de la familia *Birnaviridae.* El genoma de este virus está compuesto por dos segmentos de ARN bicatenario: el fragmento A (con una longitud aproximada de 3200 nucleótidos) y el fragmento B (con una longitud aproximada de 2800 nucleótidos) que codifican posteriormente para cinco proteínas virales: VP1, VP2, VP3, VP4 y VP5 (Xu et al., 2019).

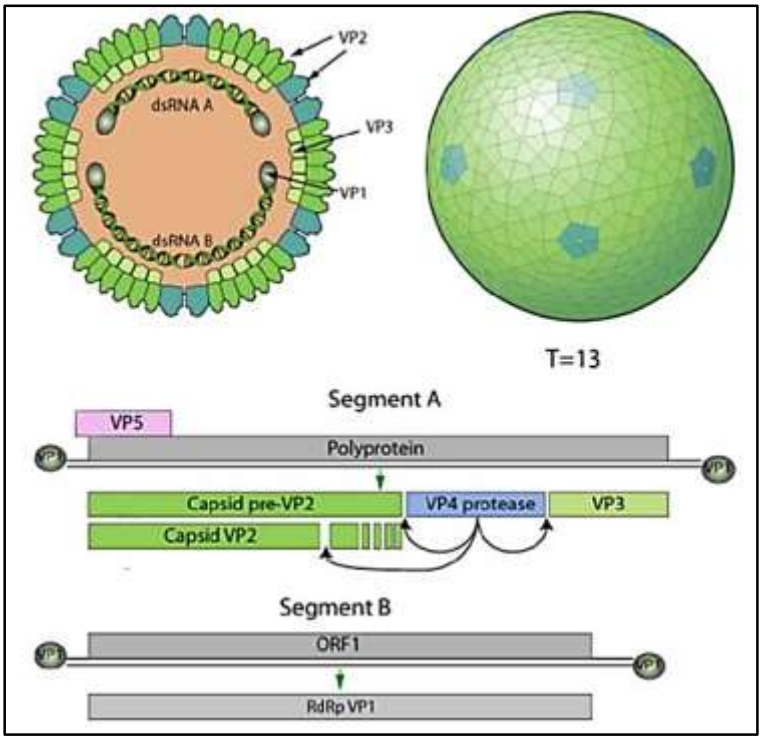

Figura 2. Representación gráfica del genoma y la estructura del virus de IBDV, imagen tomada de ViralZone Website (www.http://viralzone.expasy.org/viralzone/all_by_species/572.html).

Serotipos y Genotipos

Existen dos serotipos de IBDV. El serotipo 1 es patogénico para los pollos, mientras que el serotipo 2 no lo es. Para el caso del serotipo 1 los primeros aislamientos de IBDV se conocieron como "virus clásicos" (Cosgrove, 1962). Sin embargo, en 1985, en Delaware, Estados Unidos, surgieron cepas capaces de evadir la protección proporcionada por las vacunas. A estas cepas se les denominó "variantes antigénicas" (avIBDV). Posteriormente, con el acceso a la información genética, las cepas de IBDV se han clasificado en siete genogrupos determinados por la diversidad de la secuencia de la región hipervariable (hvVP2) (Michel and Jackwood, 2017). Las avIBDV se encuentran en el genogrupo 2.

Hospederos y signos clínicos

Aunque los pollos son los portadores más habituales del IBDV, el virus también se ha encontrado en pavos, aves acuáticas y aves rapaces. Se ha identificado una mayor susceptibilidad a la enfermedad en las pollitas Leghorn blancas (Eterradossi and Saif, 2013). Las aves mayores a 6 semanas de edad raramente muestran signos clínicos cuando se infectan. En contraste, los pollos jóvenes pueden experimentar síntomas como depresión, plumaje erizado, pérdida de apetito, diarrea, temblores y deshidratación. En galpones afectados por IBDV, la morbilidad puede alcanzar hasta el 100% y las tasas de mortalidad pueden ascender al 90% (Mahgoub, 2012).

Transmisión

La vía de transmisión fecal-oral representa la forma de infección más comúnmente identificada. En aves silvestres, se ha documentado que la infección a través de la ingestión

de carroña de pollos muertos infectados posiblemente tenga importancia epidemiológica (Orakpoghenor et al., 2020).

Distribución Geográfica

El virus de IBDV tiene una distribución global, impactando significativamente a los países con extensas áreas de producción avícola. Un estudio filogenético reciente, basado en aislamientos de campo de IBDV en Argentina, ha identificado que la variante sudamericana perteneciente al genogrupo 4 es responsable de la mayoría de los casos de la enfermedad en el país (Vera et al., 2015).

Importancia Económica y Epidemiológica

La bronquitis infecciosa, listada por la OIE en las enfermedades de declaración obligatoria, se reconoce a nivel global como una de las principales patologías que afecta la producción avícola. La fácil propagación del virus entre países constituye una de las causas del impacto económicos asociado a la enfermedad (Colvero et al., 2015). IBD produce pérdidas económicas directas debido a la mortalidad que ocasiona, mientras que entre las indirectas se encuentran las infecciones secundarias derivadas de la inmunosupresión generada por el virus, lo que propicia una mayor tasa de mortalidad y un retraso en la ganancia de peso (Zachar et al., 2016).

Síndrome Reproductivo y Respiratorio Porcino

Agente Etiológico

El Síndrome Reproductivo y Respiratorio Porcino (PRRS, por sus siglas en inglés) es ocasionado por el virus del PRRS (PRRSV), que pertenece a la familia *Arteriviridae* y al género Arterivirus según el Comité Internacional de Taxonomía de Virus (https://ictv.global/report/chapter/arteriviridae/arteriviridae). Este virus posee un genoma de ARN de sentido positivo de cadena sencilla (ssRNA(+)) con una longitud aproximada de 15 Kb (Montaner-Tarbes et al., 2019a).

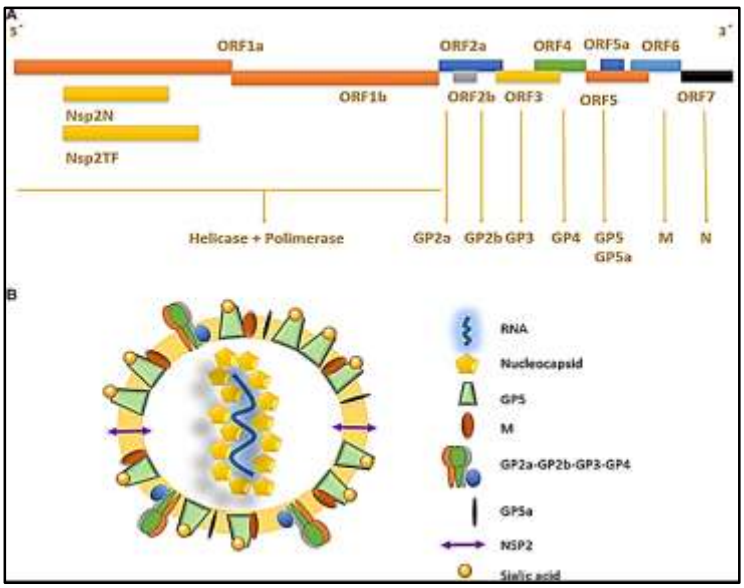

Figura 3. Representación gráfica de: A) la estructura del virus. B) el genoma de PRRSV, (Montaner-Tarbes et al., 2019b)

Genotipos

Actualmente, PRRSV se ha clasificado en dos genotipos virales: tipo 1 (PRRSV1), distribuido principalmente en Europa (Stadejek et al., 2008) y tipo 2 (PRSSV2), que ha tenido un mayor impacto en América (Shi et al., 2013), compartiendo entre ellos solo un 60% de identidad nucleotídica.

Hospederos y signos clínicos

El virus afecta tanto a las poblaciones de cerdos domésticos como a los cerdos salvajes (Shi et al., 2010). Sus manifestaciones clínicas en los animales afectados la han clasificado como una de las enfermedades más destructivas en la industria porcina (Neumann et al., 2005). Esta reputación está bien fundamentada, dada su capacidad para aumentar las tasas de aborto y fallos reproductivos en las cerdas, disminuir las tasas de supervivencia de los lechones y presentar signos clínicos respiratorios graves en cerdos de todas las edades (Chadha et al., 2023).

Transmisión

El virus se propaga, principalmente, a través del contacto directo, aunque se ha observado también la posibilidad de contagio indirecto mediante la exposición a fómites, vías iatrogénicas o aerosoles (Pileri and Mateu, 2016). Estudios experimentales también han documentado la propagación a través de aire entre granjas cercanas infectadas (Arruda et al., 2019). La importación de semen o animales positivos a PRRSV es una forma de contagio muy importante a largas distancias (Nathues et al., 2016).

Distribución Geográfica

La primera descripción del PRRS tuvo lugar a fines de la década de 1980 en los Estados Unidos y los Países Bajos, donde se observaron brotes que causaron trastornos reproductivos en los rebaños de cerdas (Zimmerman et al., 2019). A partir de ese momento, el PRRS se propagó rápidamente a diferentes países de todo el mundo y ahora es endémico en la mayoría de los países productores de cerdos (Valdes-Donoso and Jarvis, 2022). En Sudamérica, países como Colombia, Bolivia, Perú, Chile, Venezuela, Ecuador y Uruguay han informado de la presencia del virus en sus territorios. En el caso de Chile, hubo dos introducciones distintas, entre 2000 y 2006 y entre 2010 y 2015 (Neira et al., 2017a). Ecuador reportó la presencia de la enfermedad desde el 2017 hasta el 2019, luego durante tres semestres la enfermedad estuvo ausente, pero en 2021 fue posible identificar nuevamente PRRSV en los animales y desde entonces no ha cambiado el estado sanitario del país frente a la enfermedad (OIE, WAHIS 2024) (https://wahis.woah.org/#/event-management). Con respecto a Uruguay, se informó presencia de la enfermedad desde 2017 hasta 2021 (OIE, WAHIS 2024). En el caso de Perú, se reportó de la presencia de anticuerpos de la enfermedad en 1998, pero no fue hasta 2019 que se obtuvo secuencias del gen ORF5 de las cepas circulantes (Ramírez et al., 2019a). Curiosamente, aunque el virus es endémico en Colombia, las secuencias más recientes disponibles en NCBI (https://www.ncbi.nlm.nih.gov/nuccore) datan de 1998 hasta el 2000. Por otro lado, Brasil, Argentina, Paraguay, Guyana y Surinam nunca han informado de la presencia del virus en sus territorios.

Importancia Económica y Epidemiológica

PRRS ha sido catalogada como la enfermedad que mayores pérdidas económicas produce en la industria porcina. Se ha estimado que el costo de las pérdidas en EEUU asciende a más de $664 millones de dólares por año. Adicionalmente a lo expuesto, PRRSV representa un desafío significativo para los sistemas de salud animal, ya que tiende a eludir las medidas de bioseguridad tradicionales como la vacunación y el escenario se agrava con la aparición de nuevas cepas, algunas de las cuales pueden mostrar una mayor patogenicidad (Tan et al., 2019). La convergencia de estos factores constituye una restricción sustancial para la producción porcina y el comercio en países donde la enfermedad es endémica (VanderWaal and Deen, 2018).

Ectima Contagioso

Agente Etiológico

El agente causante del Ectima Contagioso (EC) es el virus ORF (ORFV), que pertenece al género Parapoxvirus de la familia *Poxviridae*. Su genoma está compuesto por ADN doble cadena lineal de aproximadamente 135.000 pares de bases. Posee una región central conservada común a varios Parapoxvirus, así como dos regiones terminales variables que albergan genes relacionados con virulencia, patogénesis, tropismo y respuesta inmunológica (Damon, 2011).

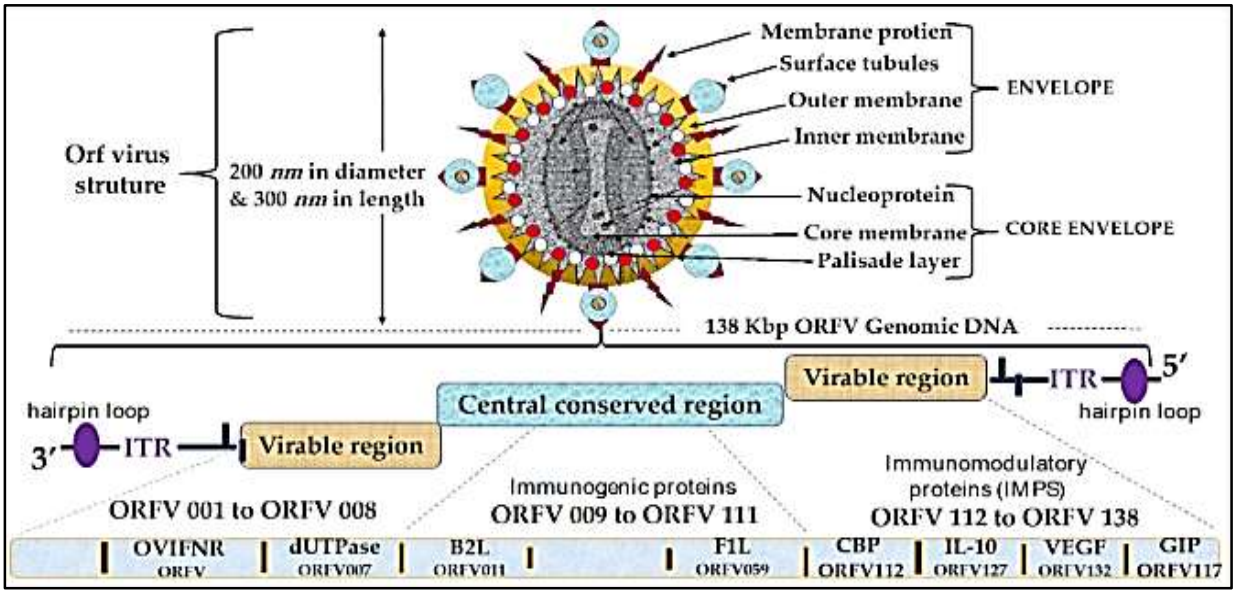

Figura 4. Representación gráfica de la estructura del virus y del genoma de ORFV (Bukar et al., 2021)

Genotipos

Hasta el momento, el virus no ha sido clasificado ni en serotipos ni en genotipos.

Hospederos y signos clínicos

El EC es una afección dermatológica severa que afecta principalmente a pequeños rumiantes domésticos, como ovinos y caprinos, aunque también se ha documentado su presencia en especies silvestres. Esta enfermedad se caracteriza por la formación de pústulas altamente infecciosas alrededor de la boca y los labios que evolucionan y dejan lesiones costrosas en los animales afectados (Damon, 2007). El EC es considerado una zoonosis ocupacional, ya que puede afectar a trabajadores rurales, empleados de plantas industriales o veterinarios que tienen contacto con cueros y lanas procedentes de animales infectados (Groves et al., 1990).

Transmisión

La infección por el virus ORF ocurre mediante el contacto directo con animales afectados o a través de objetos contaminados (Raele et al., 2021).

Distribución Geográfica

El EC se presenta de manera enzoótica y tiene una distribución global, sin estar influenciado por factores geográficos o climáticos. En Sudamérica, se han registrado casos de EC en Chile, Uruguay, Argentina y Brasil. En Brasil, la enfermedad es común en los rebaños de ovejas y cabras, afectando alrededor del 60% de las explotaciones (Abrahão et al., 2009). También se han reportado casos en humanos (Kassa, 2021). En Argentina, se han registrado casos de EC en el ganado ovino y caprino, identificados por veterinarios que observaron los síntomas clínicos (Robles et al., 2017).

Importancia Económica y Epidemiológica

El EC, generalmente tiene un impacto económico significativo en comunidades rurales que dependen de los animales para obtener carne, leche, cueros y lana. La prevención de la enfermedad se lleva a cabo mediante la vacunación con cepas atenuadas y es esencial que las cepas utilizadas en las vacunas estén relacionadas con las cepas circulantes para lograr una protección efectiva (Musser et al., 2012). Sin embargo, en la Argentina, se comercializa una vacuna viva atenuada obtenida a partir de una cepa de ORFV uruguaya, no caracterizada genéticamente. Esto representa un riesgo ya que se desconoce la relación filogenética con las cepas argentinas y ya se han reportado informalmente fallos en la protección proporcionada.

PROBLEMÁTICA, HIPÓTESIS & OBJETIVOS

PROBLEMÁTICA, HIPÓTESIS Y OBJETIVOS

Problemática

La problemática asociada a las enfermedades infecciosas radica en la necesidad de establecer medidas de vigilancia y control eficaces que puedan hacer frente tanto a las endemias como a las epidemias presentes y futuras. El objetivo final es reducir tanto el impacto económico como el social que puede resultar de la aparición de dichas enfermedades en la sociedad. A pesar de los notables avances en la investigación, persisten interrogantes metodológicos y epidemiológicos de gran relevancia que requieren respuestas concretas. Por ejemplo, ¿cuál es el comportamiento y la dinámica viral de las enfermedades en estudio en entornos específicos? ¿Cómo puede la influencia de factores externos, como la presencia de fómites, afectar la evolución y transmisión de la enfermedad? ¿Cuál es la cantidad óptima de secuencias a analizar según el objeto de estudio? ¿Cómo utilizar correctamente los datos epidemiológicos y genómicos en conjunto para informar datos específicos que puedan influir en la toma de decisiones en Salud Pública? ¿Cuál es la mejor elección de técnicas de secuenciación en diferentes situaciones epidemiológicas? ¿Cuál es el impacto de la vacunación en la dinámica evolutiva de una enfermedad durante una epidemia?

Para responder algunas de estas preguntas, se ha propuesto el diseño de modelos de epidemiología molecular, que tengan el potencial de proporcionar una comprensión más profunda de los parámetros y condiciones generales de dispersión que inciden en la dinámica evolutiva y en los patrones de transmisión de diversas enfermedades. Se espera que estos modelos permitan identificar características y dinámicas específicas de dispersión en diferentes escalas geográficas y temporales, aportando información que ayude a responder las preguntas anteriormente mencionadas y de la mano contribuya significativamente a la formulación de estrategias efectivas para la prevención, control y gestión de enfermedades en el ámbito de Una Salud.

Hipótesis

La integración de datos epidemiológicos con secuencias virales analizados en modelos de epidemiología molecular ayuda a brindar una comprensión más exhaustiva de la dinámica de transmisión y evolución de virus en diversos contextos epidemiológicos.

Objetivos

Objetivo General

Identificar rutas de transmisión viral durante un brote en campo, a nivel local, nacional e inter o intra continental de las enfermedades descriptas usando herramientas de análisis evolutivo y epidemiológico.

Objetivos Específicos

1. Realizar la secuenciación del genoma completo de distintos brotes de VFA y parcial (genes de interés) del virus ORF.

2. Realizar reconstrucciones filogenéticas y análisis de dinámica de dispersión de los cuatro virus de estudio.

3. Realizar análisis filogeográficos de los cuatro virus de estudio comparando parámetros claves y rutas de transmisión viral.

Para abordar los objetivos mencionados previamente, la metodología y los resultados de la tesis se estructuraron en tres capítulos, donde cada uno se enfoca en distintos aspectos que responden a los tres objetivos específicos planteados. En el primer capítulo se diseñaron modelos de epidemiología molecular continuos a una escala local, utilizando datos de la epidemia de FA ocurrida en Argentina en el año 2001. El segundo capítulo se centró en la formulación de modelos discretos a una escala intra o intercontinental, tomando como ejemplo a PRRSV e IBDV. Finalmente, el tercer capítulo se dedicó a la evaluación del alcance de los modelos a escala nacional, abordando la endemia de Ectima Contagioso presente en Argentina. Mientras que en los primeros dos capítulos se analizaron virus de ARN de rápida evolución, en el último se estudió un virus de ADN con tasas de sustitución menores. Cada uno de estos capítulos contribuye de manera integral al alcance de los objetivos generales y específicos planteados en este trabajo.

CAPITULO 1

MODELOS CONTINUOS A UNA ESCALA LOCAL

CON UN ENFOQUE EN VIRUS DE ARN

Dinámica del virus de la Fiebre Aftosa

CAPITULO 1 - VFA

Materiales y Métodos

Este capítulo de la tesis se centró en realizar un análisis utilizando datos recopilados durante la epidemia de Fiebre Aftosa que afectó a Argentina en el año 2001. El objetivo principal de este análisis fue evaluar si los modelos de epidemiología molecular continuos a una escala local pueden proporcionar una comprensión más profunda de la dinámica de contagio y la propagación entre establecimientos productores.

Para abordar estas cuestiones, se utilizaron datos procedentes de dos regiones distintas del país. La primera región estudiada corresponde al partido de Mar Chiquita, en la provincia de Buenos Aires, una zona dedicada a producir carne. La segunda región de interés se encuentra en la cuenca lechera argentina, abarcando algunos partidos de las provincias de Santa Fe y Córdoba.

Selección de muestras, extracción de material genético y síntesis de ADNc., amplificación por PCR y secuenciación

Selección y rastreo de muestras

Durante el brote de Fiebre Aftosa en Argentina en los años 2000 y 2001, los veterinarios de SENASA visitaron las instalaciones de producción ganadera que reportaron signos clínicos compatibles con la enfermedad. En estas visitas, se tomaron muestras de lesiones de animales enfermos, las cuales fueron posteriormente almacenadas a una temperatura de -80°C en el laboratorio de referencia nacional de bioseguridad NSB4OIE de SENASA en Buenos Aires.

Partido de Mar Chiquita

En colaboración con el Departamento de Epidemiología y el Laboratorio de Diagnóstico de SENASA, se llevó a cabo un rastreo epidemiológico en el partido de Mar Chiquita, donde se habían reportado 66 brotes sospechosos de FA. De estos, 41 muestras fueron confirmadas por diagnóstico de laboratorio. Para esta tesis, se seleccionaron 21 muestras de campo que contaban con datos epidemiológicos trazables y suficiente material biológico para su procesamiento, además de contar con resultados positivos en las pruebas de PCR y serología para VFA serotipo A2001 realizadas en su momento por el SENASA (Tabla 1 y Figura 5).

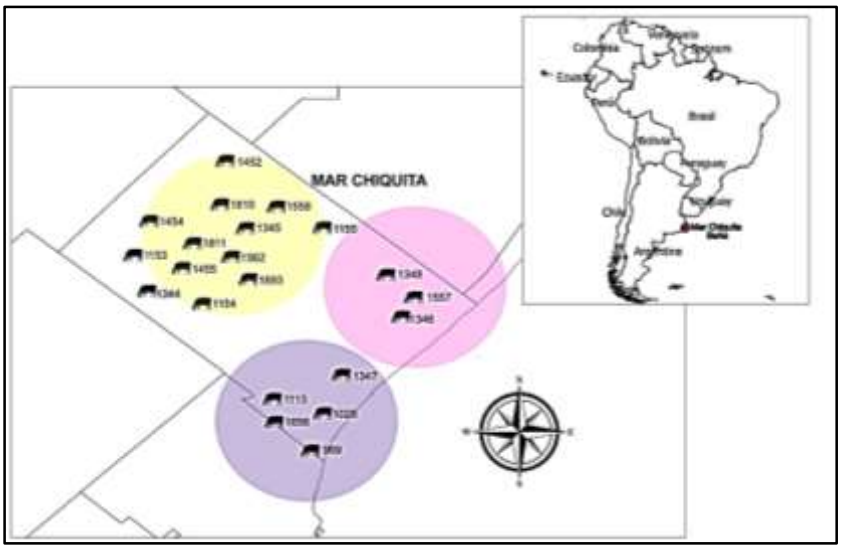

Figura 5. Ubicación del partido de Mar Chiquita en América del Sur y coordenadas geográficas aproximadas de las muestras de FA analizadas en este trabajo. Las ubicaciones dentro del partido se indican de la siguiente manera: la parte Oeste del partido en amarillo, el sur en morado y el noreste en rosa

Número de la Muestra	Número de acceso en GenBank	Fecha de reporte del caso
969	OR338601	19/5/2001
1028	OR338602	23/5/2001
1113	OR338603	24/5/2001
1153	OR338604	28/5/2001
1154	OR338605	29/5/2001
1155	OR338606	29/5/2001
1344	OR338607	2/6/2001
1345	OR338608	5/6/2001
1346	OR338609	6/6/2001
1347	OR338610	9/6/2001
1348	OR338611	9/6/2001
1452	OR338612	12/6/2001
1454	OR338613	12/6/2001
1455	OR338614	19/6/2001
1557	OR338615	23/6/2001
1558	OR338616	22/6/2001
1562	OR338617	26/6/2001
1656	OR338618	27/6/2001
1810	OR338619	3/7/2001
1811	OR338620	4/7/2001
1893	OR338621	14/7/2001

Tabla 1. Identificación, números de acceso en GenBank y fechas de muestreo de las secuencias analizadas del partido de Mar Chiquita en este trabajo.

Extracción de ARN y síntesis de ADNc

En el laboratorio de bioseguridad NSB4 del SENASA en Martínez, se procedió a la inactivación del ARN viral, proveniente de los tejidos, utilizando Tri-Reagent® (MRC). Para este procedimiento, se tomaron 300 µl de la muestra, generalmente macerado de afta, y se mezcló en una proporción de 3:1 con el agente desnaturalizante. Luego, todas las muestras se trasladaron refrigeradas al laboratorio de Epidemiología Molecular y Evolución Viral del IABIMO-INTA, donde se añadieron 200 µl de cloroformo por cada ml de Trizol, seguido de una centrifugación a 11.500 g durante 15 minutos. La fase acuosa resultante se combinó con 500 µl de isopropanol por cada ml Trizol, junto con 1 µl de glucógeno a una concentración de 20 µg/µl (Boeringher Mannheim). La mezcla se dejó precipitar durante 15 minutos a temperatura ambiente y luego se centrifugó a 11.500 g durante 10 minutos a 4°C. El precipitado se lavó con etanol al 70% y se secó a 37°C. Finalmente, el ARN se disolvió en 14 µl de agua DEPC a 60°C durante 10 minutos. Posteriormente, se obtuvo el ADNc mediante una reacción de transcripción reversa (RT). En este procedimiento, se añadieron 500 ng de hexámeros al azar (Promega) y los oligonucleótidos F3Rev a la solución de ARN obtenida, y se incubó a 70°C durante 10 minutos. Luego, Se preparó una mezcla con un volumen total de 20 µl, que contenía 4 µl de *buffer* de RT 5X, 1 µl de deoxinucleótidos trifosfato (dNTP) 10mM y 2µl de DTT 100mM. Después, se mezcló suavemente y se incubó a 37°C durante 2 minutos. Finalmente, se agregaron 200 unidades de la enzima transcriptasa reversa Super Script™ III de Invitrogen. La reacción fue incubada durante 1 hora a 37°C y luego se inactivó a 85°C durante 5 minutos.

Amplificación por PCR

Se realizaron siete PCR con productos de diferentes tamaños a fin de amplificar el genoma completo del virus (Tabla 2). Para este procedimiento, se utilizó la enzima de alta fidelidad Q5 de NEB y se emplearon entre 2 y 5 µl de ADNc como templado. Se añadieron 5 µl de Reaction Buffer Q5 (5x), 0,5 µl de dNTP 10 mM, 10 µM de oligonucleótidos sentido y antisentido, agua libre de nucleasas hasta llegar a un volumen final de 24,8 µl y 0,25 µl de Q5 High-Fidelity DNA Polymerase. Esta mezcla fue homogenizada con micropipeta. Los perfiles de ciclado se ajustaron al fragmento que se estaba amplificando, y para determinar la temperatura óptima de *anneling* se empleó la herramienta "The NEB Tm Calculator".

PASO	TEMPERATURA	TIEMPO	CICLOS
Desnaturalización inicial	98°C	30 segundos	
Desnaturalización Hibridación Extensión	98°C 45-72ºC 72ºC	30 segundos 30 segundos 30 segundos/kb	X 35
Extensión final	72ºC	10 minutos	

Tabla 2. Condiciones de ciclado utilizados para la amplificación del genoma completo del VFA.

Nombre	Orientación	Secuencia 5 'a 3'	Posición	Tm (ºC)
F1BFor	sentido	TTGAAAGGGGGCGCTAGGGTTTC	1-23	76ºC
F1BRev	antisentido	GGGGTGAAAGGCGGRCTTCGGG	348-369	67ºC
F2For	sentido	TTACGTCTCC23C	363-386	98ºC
F2Rev	antisentido	GGCCTCCGGTCACCTATT	1042-1059	63ºC
F3For	sentido	CCTGGTCTTTCCAGGTCTAGAGG	674-696	67ºC
F3 rev	antisentido	ATGTAGGCGTATGAGTCCACC	2249-2229	58ºC
F4 for	sentido	CCGACAAGAAGACGGAGGAGACG	1995-2017	74ºC
F4 rev	antisentido	GGCCCAGGGTTGGACTCAACG	3933-3953	74ºC
F5 for	sentido	AGGCACCGTTCACGAGACTTGC	3639-3660	72ºC
F5 rev	antisentido	TCCTTGATGGAGTCATGCACCA	5453-5474	69ºC
F6 for	sentido	CCATGGCATCACTCGAGGAC	5019-5039	67ºC
F6 rev	antisentido	TTCCATCGCATCGAGACCGTCG	7042-7062	76ºC
F7 for	sentido	CACGAAACACTTTCGTGACACAGC	6381-6404	70ºC
F7 rev	antisentido	GCTCTTTCGAGCTTTTCACTCCAGC	8200-8225	73ºC

Tabla 3. *Primers* utilizados para la amplificación del genoma completo del VFA.

Electroforesis en gel de agarosa

Los productos resultantes de la PCR se combinaron con un *buffer* de siembra concentración 10X, compuesto por un 50% de glicerol, TAE 10X y un 1% de azul de bromofenol. Posteriormente, las muestras se sembraron en geles de agarosa junto con marcadores de peso molecular seleccionados en función del tamaño de los productos. Con el propósito de visualizar las bandas de ADN se emplearon geles de agarosa con concentraciones que variaron entre 1 y 2% suspendidos en un *buffer* TAE 1X (que contiene 10 mM de Tris-acetato y 1 mM de EDTA) y fueron mezclados con el agente intercalante Bromuro de Etidio a una concentración de 0,5 µg/ml. Los geles fueron sometidos a una corriente de 90V durante 50 minutos. Finalmente, para la captura de imágenes de los geles fue utilizado el equipo Molecular Imager® Gel Doc™ XR+ de Bio-Rad, con el programa Image Lab™ de la misma marca.

Con el objetivo de garantizar la especificidad de los fragmentos generados por la PCR y descartar la posibilidad de contaminaciones, se incorporó un control negativo en cada etapa de extracción de ARN y síntesis de ADNc. En este control negativo se siguieron todos pasos mencionados anteriormente, incluyendo las reacciones de PCR.

Purificación de productos de PCR

Fue utilizando el kit de purificación de PCR QIAquick de Qiagen, el cual emplea resinas de purificación basadas en membranas de sílice para eliminar las impurezas del ADN en las muestras siguiendo el protocolo sugerido por el fabricante.

Cuantificación

Para determinar la concentración de ADN se empleó un espectrofotómetro NanoDrop ND-1000 (Thermo Scientific).

Secuenciación

La secuenciación se llevó a cabo en la Unidad de Genotipificación y Secuenciación Automatizada (UGB) del IABIMO en el INTA, utilizando el método de Sanger (Sanger et al., 1977). Se emplearon terminadores fluorescentes (kit Big Dye Terminator v3.1, Applied Biosystems) en secuenciadores automáticos de capilares Genetic Analyzer 3500xl y 3130xl (Applied Biosystems). Para el análisis de los datos, se utilizó el programa Sequencing Analysis v5.4 (Applied Biosystems).

Curado de secuencias

Los productos de secuencia fueron chequeados utilizando el programa BioEdit (Hall, 1999). Cada archivo fue examinado visualmente y editado para resolver cualquier ambigüedad. Se realizaron recortes en los extremos de los fragmentos donde se identificó una calidad de lectura baja. Para el ensamblaje final, se utilizó como referencia una secuencia conocida del VFA correspondiente a la cepa A24Cruzeiro (AN:AJ251476.1). Las extracciones del material genético, las reacciones de RT-PCR y el ensamblado de las secuencias fueron realizadas por el Doctor Cabanne, mientras que la secuenciación se realizó en la UGB.

Cuenca Lechera

Selección y rastreo de muestras

Con el respaldo del Departamento de Epidemiología y el Laboratorio de Diagnóstico de SENASA, se rastrearon 65 focos epidemiológicamente significativos provenientes de las provincias de Santa Fe y Córdoba, donde predominan las explotaciones de tipo lechero. Para la selección de estas muestras, se consideraron varios criterios, como el tipo de producción de la explotación, la confirmación del agente etiológico y su serotipo (cepa A/Argentina/2001), la concordancia entre los datos obtenidos en el área de epidemiología y el área de diagnóstico, y, por último, la disponibilidad de una cantidad suficiente de muestra para su posterior análisis. Se procesaron dichas muestras y, tras la clasificación, se eligieron finalmente 15 que destacaron por su preservación en la integridad y calidad del ARN. Las muestras seleccionadas provenían de los partidos de Iriondo, Castellanos, Las Colonias (Santa Fe) y San Justo (Córdoba).

Extracción de ARN y síntesis de ADNc

En el laboratorio de bioseguridad NSB4 del SENASA en Martínez, se llevó a cabo la inactivación del ARN viral. Posteriormente, en el laboratorio de Epidemiología Molecular y Evolución Viral del INTA, se realizó la extracción de ARN y la obtención de ADNc, siguiendo el protocolo previamente descrito en la sección correspondiente a la metodología usada en el partido de Mar Chiquita.

Síntesis de la segunda hebra de ADN

(Primera metodología)

El ADNc se procesó utilizando dos metodologías diferentes. Con la intención de evitar la introducción de errores en la secuencia nucleotídica, debido a posibles fallos en la amplificación por PCR, se optó primeramente por sintetizar directamente la segunda hebra del ADNc. Esta técnica se llevó a cabo mediante el uso de la enzima Second Strand cDNA Synthesis de NEB. Para ello, se preparó una mezcla que consistió en la adición de 48 µl de agua DEPC, 8 µl del Second Strand Synthesis Reaction Buffer 10x y 4 µl de la enzima Second Strand Synthesis a 20 µl de ADNc. Luego, se homogeneizó cuidadosamente la mezcla con micropipeta y se incubó durante una hora a 16°C.

A continuación, se procedió a cuantificar la carga viral en todas las muestras mediante PCR en tiempo real utilizando el termociclador stepOne™ de Applied Biosystems™, con su correspondiente *software*. Se empleó una curva estándar basada en clones de una secuencia viral, previamente caracterizada. En este paso, también se verificó que las muestras incógnita estudiadas presentaran una concentración superior a 1×10^8 copias/µl, de acuerdo con el protocolo establecido previamente por Logan y colaboradores (Logan et al., 2014).

Las muestras que cumplían con el requisito mencionado anteriormente fueron sometidas a un proceso de purificación utilizando el kit PCR QIAquick® de Qiagen, siguiendo las instrucciones del fabricante.

Amplificación de productos de PCR

(Segunda metodología)

En el segundo método, se aplicó la misma metodología utilizada en el partido de Mar Chiquita, que implica realizar PCR de diferentes longitudes para amplificar el genoma completo en siete fragmentos distintos (Tabla 2 y 3). La elección de esta metodología se fundamentó en la efectividad que demostró para amplificar muestras cuya carga viral había disminuido considerablemente debido al prolongado período de almacenamiento. Las siete PCR positivas que abarcaban el genoma completo del VFA se sometieron a un proceso de purificación mediante el kit de purificación de PCR QIAquick® de Qiagen.

Secuenciación

Después de obtener el ADNc de doble cadena o los productos de las PCR y purificar las muestras, se procedió a cuantificarlas utilizando el fluorómetro Qubit® (ThermoFisher Scientific), kit dsDNA HS Assay. Las bibliotecas fueron preparadas con el kit Illumina DNA Prep (Illumina, Inc) y luego los productos fueron secuenciados en un equipo MiSeq® (Illumina, Inc) con lecturas PE (2 x 250 pb) (El objetivo fue alcanzar una cobertura aproximada de 1000x). La secuenciación se realizó en la UGB.

Curado de secuencias

Se obtuvo las secuencias crudas en formato FASTQ y su calidad fue chequeada a través del programa FastQC (versión 0.12.1) (Brown et al., 2017). Las lecturas con una calidad promedio menor a q20 y de una longitud menor a 50 bp fueron descartadas utilizando el programa BBDuk (versión 39.01) disponible en https://sourceforge.net/projects/bbmap/ (Bushnell, 2020). Además, se eliminaron las secuencias de adaptadores remanentes. A continuación,

se utilizó el programa Bowtie 2 (versión 2.5.1) (Langmead and Salzberg, 2012) para alinear las lecturas de los productos de secuenciación con una secuencia de referencia correspondiente al genoma completo del VFA. Esta referencia fue previamente construida por el grupo durante la epidemia del 2001 y proviene del partido de Mar Chiquita. El programa SAMtools (versión 1.17) (Danecek et al., 2021) se utilizó para manipular los archivos generados en formato SAM (Sequence Alignment Map) y convertirlos a su formato binario BAM, así como para descartar las lecturas no mapeadas. Los alineamientos fueron visualizados con el programa Tablet (versión 1.21.02.08) (Milne et al., 2013). La detección de variantes fue realizada con el programa Lofreq (versión 2.1.5) (Wilm et al., 2012). La información de cobertura se obtuvo con el programa BEDTools (versión 2.31.0) (Quinlan and Hall, 2010). Por último, las secuencias consenso se determinaron con el programa BCFTools (Danecek et al., 2021) (versión 1.17) incluyendo todas aquellas variantes con una frecuencia mayor al 50 % y se enmascararon aquellas posiciones con una cobertura menor a 20x.

Construcción de árboles y redes filogenéticas

Alineamiento de secuencias

Una vez obtenidas las secuencias consenso de ambas regiones de estudio (cuenca lechera y producción de carne) se alinearon utilizando el paquete de software MAFFT versión 7 (disponible en https://mafft.cbrc.jp/alignment/server/) (Katoh and Standley, 2014) y se curaron manualmente. Para generar el archivo Nexus necesario para producir la red filogenética, se utilizó el paquete DNASequence Polymorphism (DnaSP) versión 5 (http://www.ub.edu/dnasp/) (Rozas and Rozas, 1999).

Árboles filogenéticos

Los árboles filogenéticos se estimaron utilizando el método de Máxima Verosimilitud implementado en el programa IQ-Tree versión 1.4.2 (http://iqtree.cibiv.univie.ac.at) (Trifinopoulos et al., 2016). El modelo de sustitución fue seleccionado dentro del mismo programa a través del módulo ModelFinder. El árbol se estimó de acuerdo con el Criterio de Información Bayesiano (BIC) con 10.000 réplicas de Bootstrap Ultrarrápido (UFB) como test de soporte. Los resultados se visualizaron con la herramienta FigTree v.1.4.0 (disponible en http://tree.bio.ed.ac.uk/software/figtree) (Rambaut, 2009).

Construcción de redes filogenéticas

Se realizaron análisis de redes filogenéticas utilizando el software de Análisis de Poblaciones con Árboles Reticulados (PopART) (disponible en http://popart.otago.ac.nz/index.shtml) (Leigh and Bryant, 2015) utilizando el algoritmo “median-joining network”, considerando *gaps*, sitios invariables y configurando el parámetro de épsilon en cero (este procedimiento solo fue realizado en las secuencias de Mar Chiquita)

Reconstrucción filogeografíca y análisis evolutivos

Se evaluó la señal temporal utilizando el software TempEst (disponible en https://beast.community/tempest (Rambaut et al., 2016). Posteriormente, se llevaron a cabo análisis filogenéticos y filodinámicos bayesianos utilizando el programa “Bayesian Evolutionary Analysis Sampling Trees” (BEAST versión 1.10.4), disponible en (https://beast.community/, (Suchard et al., 2018) para inferir el Árbol de Máxima Credibilidad

(MCC), la tasa evolutiva media, la fecha del ancestro común más reciente (ACMR), y la reconstrucción filogeográfica viral en un espacio continuo. Se probaron diferentes combinaciones de modelos coalescentes. Fueron testeados dos modelos de dinámica poblacional (Bayesian Skyride y Crecimiento Exponencial) y dos modelos de reloj molecular (Relajado Lognormal No Correlacionado y Estricto). Las fechas y ubicaciones de muestreo se especificaron como rasgos continuos. Las georreferencias de las explotaciones no se presentan en esta tesis para preservar el anonimato de las granjas, pero los análisis se realizaron con alta resolución espacial. Se utilizó el método de Bayes Factor (BF) para probar y seleccionar el modelo que más se adapta a la realidad. Se ejecutaron dos cadenas independientes de MCMC, cada una con 15 millones de estados, con la implementación de stepping-stone sampling. Se utilizó el programa Tracer v.1.7 (Rambaut et al., 2018) para verificar la convergencia, es decir, cuando el tamaño efectivo de la muestra (ESS) fue > 200 para parámetros relevantes. Para estos análisis y subsiguientes, se descartó el 10% inicial de la cadena MCMC conocido como "Burn-in" ya que en este período el algoritmo se encuentra aún en fase exploratoria. El árbol de máxima credibilidad se visualizó con la herramienta Figtree, y las diferencias en las tasas evolutivas específicas de las ramas se probaron estadísticamente, según si las explotaciones fueron vacunadas (o no) antes del brote. Este análisis se llevó a cabo utilizando el test estadístico no paramétrico de Wilcoxon (Derryberry et al., 2010) implementado en R Studio a través del programa R v.3.4.1 en el paquete ggpubr. Finalmente, se utilizó el programa SpreaD3 (Bielejec et al., 2011) para analizar y visualizar las reconstrucciones filogeográficas (https://rega.kuleuven.be/cev/ecv/software/SpreaD3). El archivo de salida se utilizó para generar un video en Google Earth®. (Este procedimiento solo fue realizado en las secuencias de Mar Chiquita)

Resultados

Resultados del análisis de la epizootia en el partido Mar Chiquita

Estadísticas de secuenciación

De las 21 secuencias analizadas, 20 de ellas presentaron una longitud que osciló entre 8122 y 8124 bases. El promedio de caracteres nucleotídicos ambiguos por genoma fue de 1,5, con valores que variaban desde 0 hasta 13. En el caso de la muestra 1454, no fue posible amplificar uno de los fragmentos de PCR (correspondiente a las regiones VP3, VP2 y VP1 del genoma) lo que resultó en un gap entre las posiciones 2315 y 3595. A pesar del mismo, la muestra fue analizada debido a que todavía conservaba más del 80% de la secuencia viral.

Análisis Filogenético

Las 21 secuencias de Mar Chiquita A/Arg/01 se alinearon con secuencias de genoma completo previamente publicadas de la epizootia de 2000/1 (Brito et al., 2016), demostrando que forman un grupo monofilético con un soporte de 97% de UFB. La secuencia de genoma completo de FA de General López (GL), provincia de Santa Fe, Argentina, fue utilizada como grupo externo. El modelo de sustitución mejor ajustado para el conjunto de datos fue (TN+F+I) el cual, se utilizó para estimar el árbol de Máxima Verosimilitud (Figura 6). Este análisis filogenético reveló tres grupos locales de transmisión en el partido (A, B y C), cada uno con un fuerte soporte de UFB (A=98%, B=95%, C=97%). Adicionalmente, los grupos B y C

presentaron valores de soporte altos para para la formación de grupos dentro de cada uno de ellos.

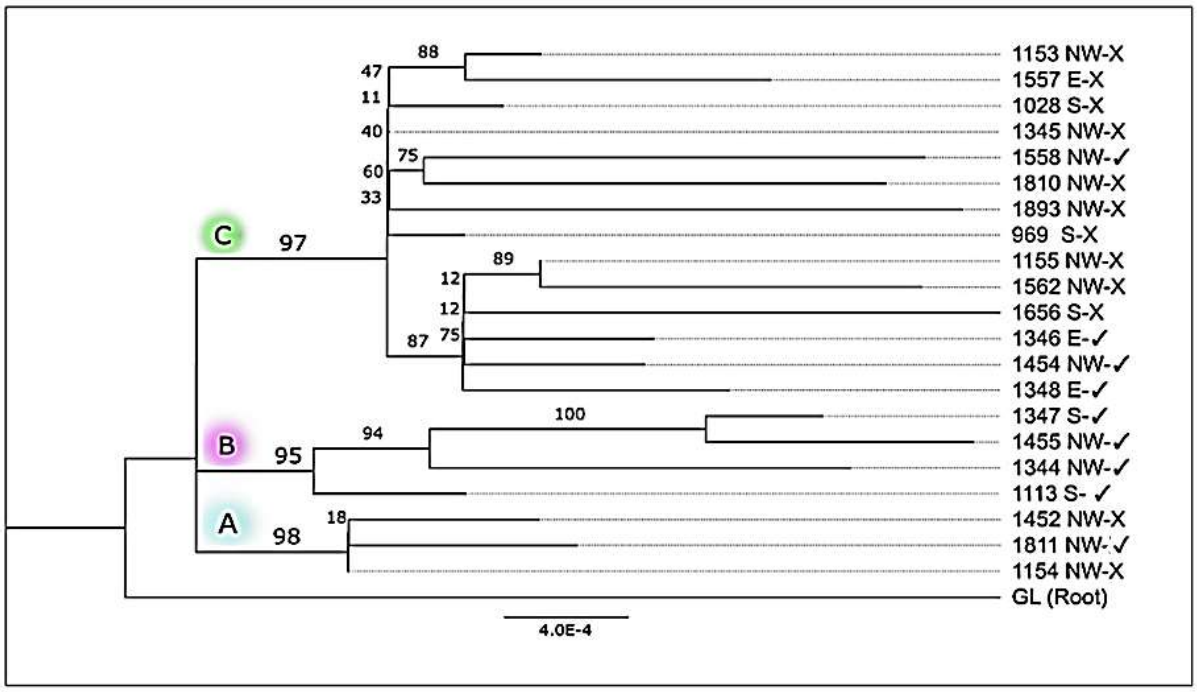

Figura 6. Árbol filogenético de máxima verosimilitud estimado a partir de 21 secuencias genómicas de Mar Chiquita y el grupo externo (GL). Los números sobre las ramas muestran los valores de UFB. El modelo de sustitución utilizado fue (TN+F+I). Los tres principales grupos genéticos se indican con las letras A, B y C. El estado de vacunación del rebaño de los aislados se muestra junto al nombre de la secuencia: ✓ indica animales vacunados y ✗ indica animales no vacunados. NW: Noroeste, S: Sur, E: Este. La barra de escala muestra la distancia genética en unidades de sustituciones por sitio. Este árbol se construyó utilizando el programa IQTree v. 1.4.2.

Las muestras de las tres subregiones (Noroeste, Sur y Este) se tomaron de animales vacunados y no vacunados. El grupo A comprende dos muestras de hatos no vacunados ubicados en el Noroeste del partido. Por otro lado, el grupo B estuvo constituido por hatos vacunados ubicados en el Noroeste y el Sur. Finalmente, el grupo C contenía el mayor número de focos muestreados (14) y comprendía, mayormente, rodeos no vacunados.

Para facilitar la visualización de las relaciones filogenéticas, se construyó una red en base al algoritmo “median-joining network”, que indica el número de cambios nucleotídicos entre secuencias (muestras reales y/o ancestros teóricos en la red). La red tiene una estructura similar a la del árbol de Máxima Verosimilitud (Figura 6) con tres grupos locales de transmisión (A, B y C). Algunos aislamientos, específicamente 1154, 1345 y 1155, se indicaron como predecesores de otros aislamientos. En algunas producciones incluidas en el mismo grupo (y pertenecientes, probablemente, a la misma cadena de transmisión) con menos de siete cambios nucleotídicos entre ellas, se midió la distancia geográfica lineal para un análisis posterior.

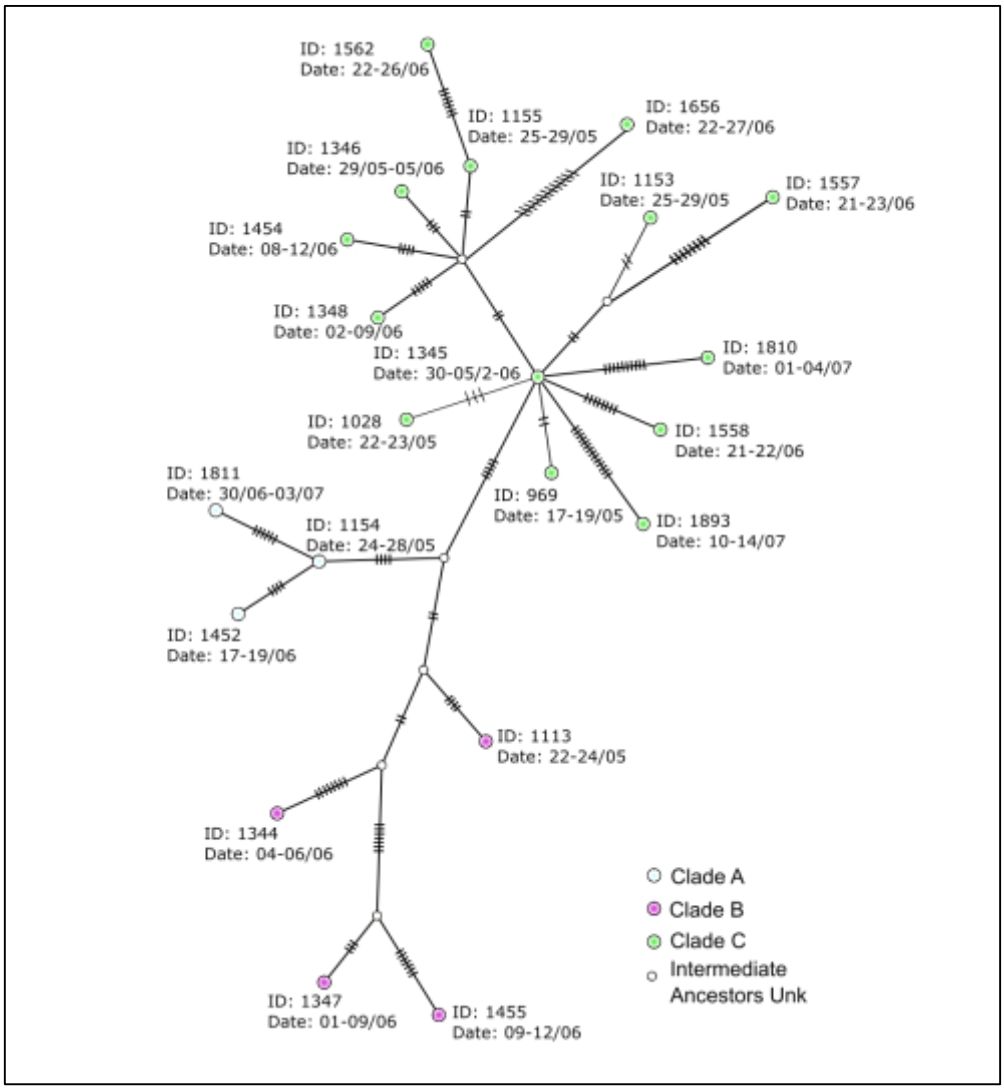

Figura 7. Red filogenética construida mediante el algoritmo de "median joining network". Las líneas transversales indican el número de cambios nucleotídicos entre dos secuencias de aislamientos o ancestros intermediarios potenciales. El clado filogenético se indicó con círculos de colores. Los ancestros intermedios desconocidos (Unk) se marcan con un punto más pequeño. Esta red se construyó utilizando el programa PopART.

Análisis Filodinámico

Con base al árbol de Máxima Verosimilitud y la datación temporal de las muestras, se llevó a cabo un análisis de señal temporal mediante el programa Tempest. Aquí, se observó una correlación positiva entre la divergencia genética y el tiempo, con un valor de R-cuadrado de 0.61, lo que indica la viabilidad de aplicar un modelo de reloj molecular a este conjunto de datos. Se probaron dos modelos coalescentes (Bayesian Skyride y Crecimiento Exponencial) en combinación con dos modelos de reloj molecular (Relajado Lognormal No Correlacionado y Estricto). Todas las corridas de MCMC mostraron convergencia. La prueba BF indicó que la mejor combinación de modelos fue el modelo Bayesian Skyride con un reloj molecular relajado. El árbol MCC obtenido (Figura 8) fue similar al árbol de Máxima Verosimilitud (Figura 6). La tasa de evolución molecular se estimó en 1.17×10^{-2} sustituciones/sitio/año (intervalo HPD del 95%: $6.7\text{x}10^{-3}$ - $1.68\text{x}10^{-2}$), la cual fue consistente con la obtenida mediante Tempest (1.24×10^{-2} sustituciones/sitio/año). La fecha estimada del Ancestro común más reciente de

las secuencias de Mar Chiquita fue abril de 2001 (mediana 2001.31, con intervalos HPD del 95% de 2001.22 a 2001.36), apenas un mes antes del primer brote reportado en el distrito.

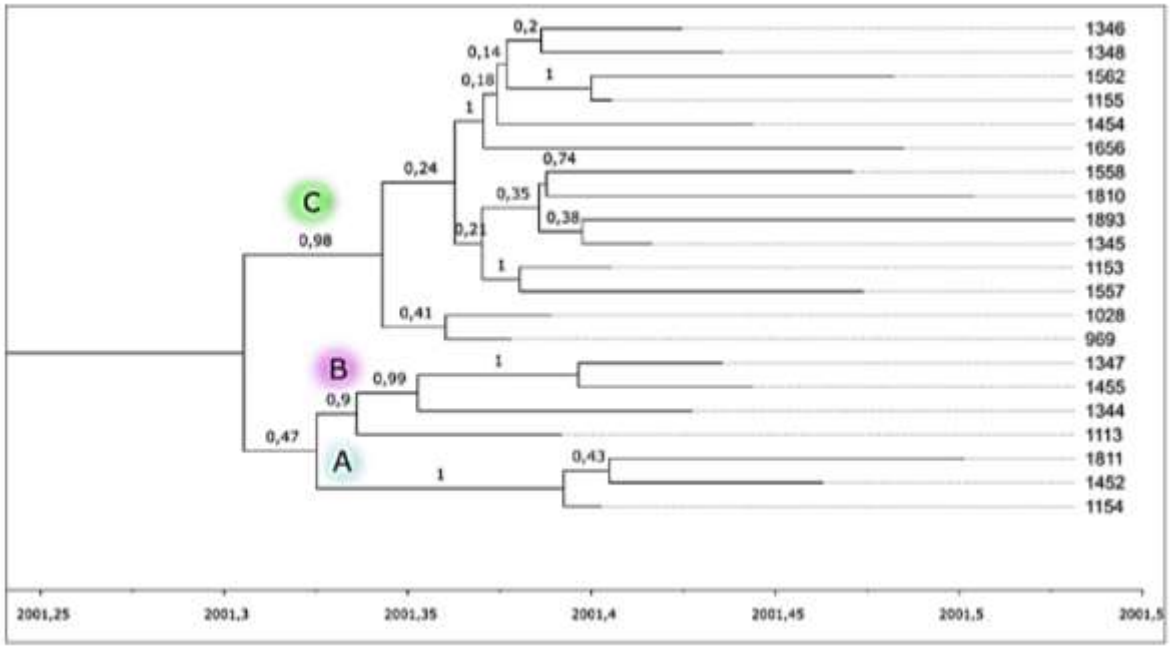

Figura 8. Árbol de máxima credibilidad estimado a partir de las 21 secuencias de Mar Chiquita. Los clusters genéticos se indican con las letras A-C. Los números sobre las ramas muestran los valores de probabilidad posterior. El tiempo se muestra en el eje X. Este árbol fue construido utilizando el programa BEAST v. 1.10.4.

Dado que antes de reportar sospechas de FA, algunas granjas habían sido inmunizadas con la vacuna inactivada, se llevó a cabo un análisis para determinar si el estado de vacunación tenía impacto en la evolución molecular del virus. Para ello, se calculó la tasa de evolución en las ramas terminales, con el objetivo de identificar posibles diferencias en la presión de selección entre los rebaños vacunados y no vacunados. Este análisis se repitió tanto para las secuencias del genoma completo como para las secuencias de la región de P1, la cual comprende todas las proteínas de la cápside viral y que podría verse particularmente afectada por la presencia de anticuerpos específicos (Figura 9 y 10).

Se realizaron test estadísticos mediante un análisis de Wilcoxon, y los resultados mostraron que no hubo diferencias significativas entre los grupos vacunados y no vacunados tanto para el genoma completo (P=0.14) como para la región P1 (P=0.31), Figura 11.

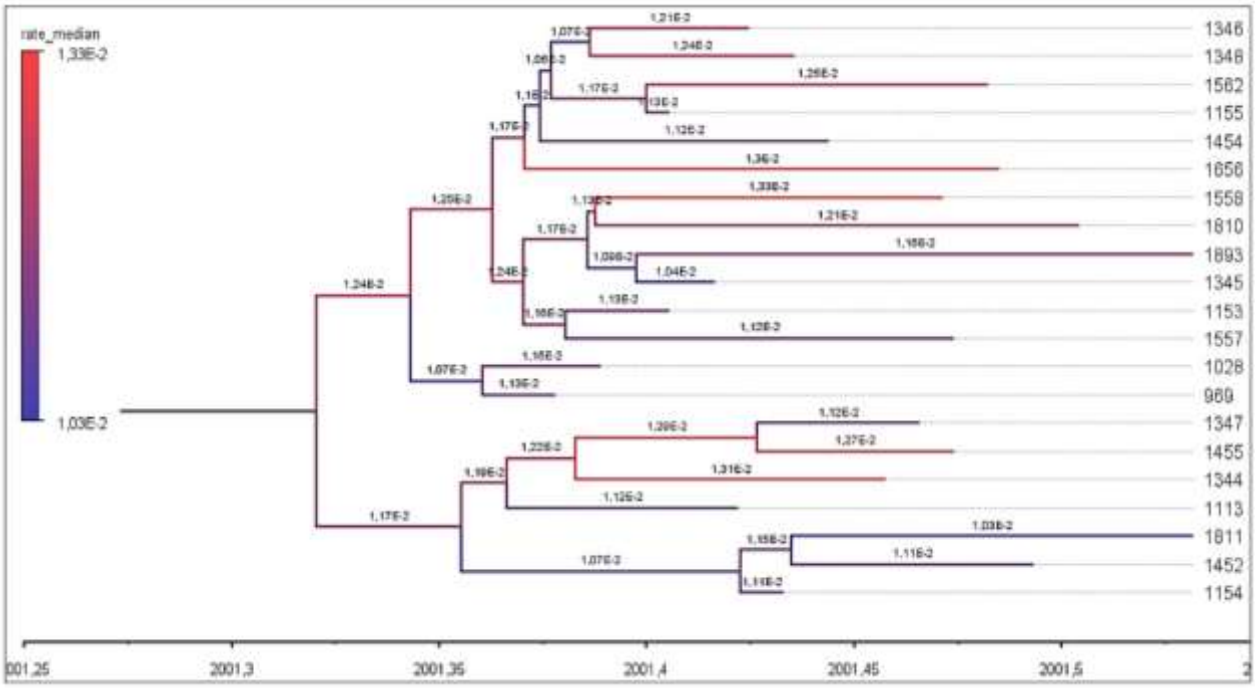

Figura 9. Árbol MCC estimado a partir de secuencias del genoma completo; los números sobre las ramas indican la tasa evolutiva específica de cada una y el color es dependiente del valor de la tasa evolutiva (rojo para las ramas más rápidas y azul para las más lentas). Este árbol se construyó utilizando el programa BEAST v. 1.10.4.

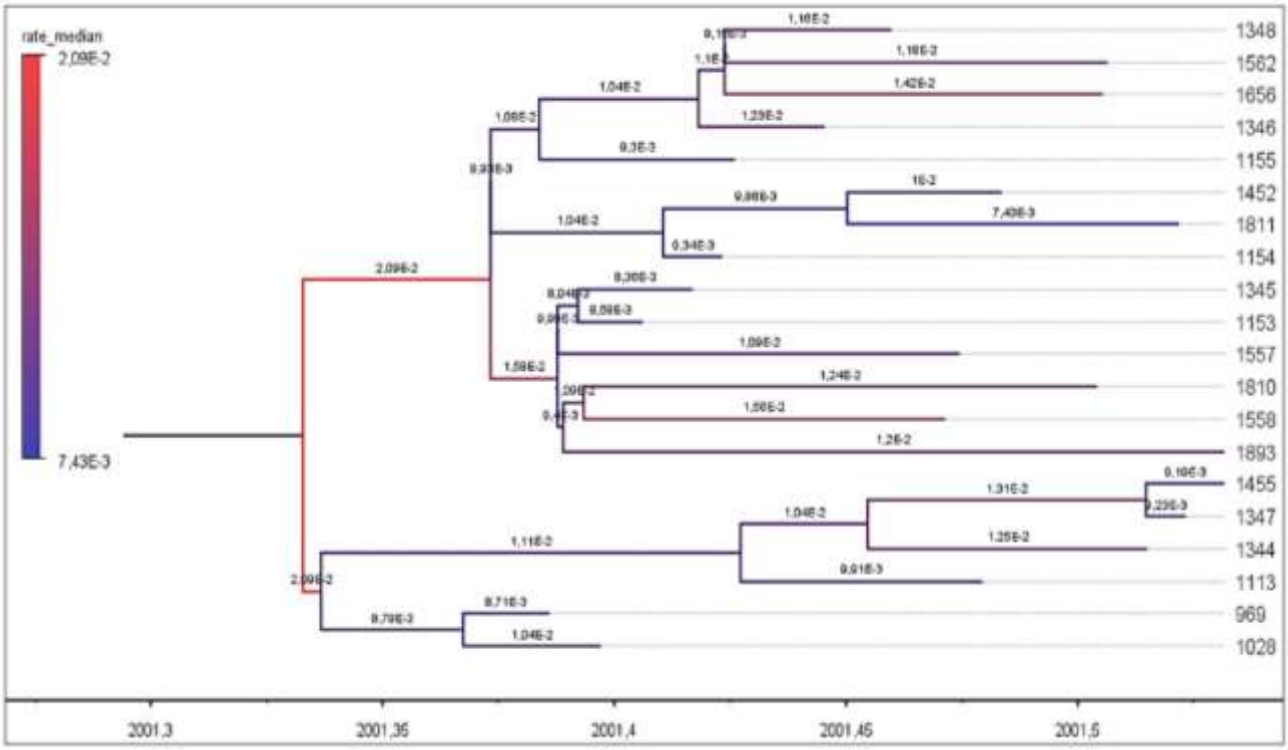

Figura 10. Árbol MCC estimado a partir de secuencias de la región P1; los números sobre las ramas indican la tasa evolutiva específica de cada una y el color es dependiente del valor de la tasa evolutiva (rojo para las ramas más rápidas y azul para las más lentas). Este árbol se construyó utilizando el programa BEAST v. 1.10.4.

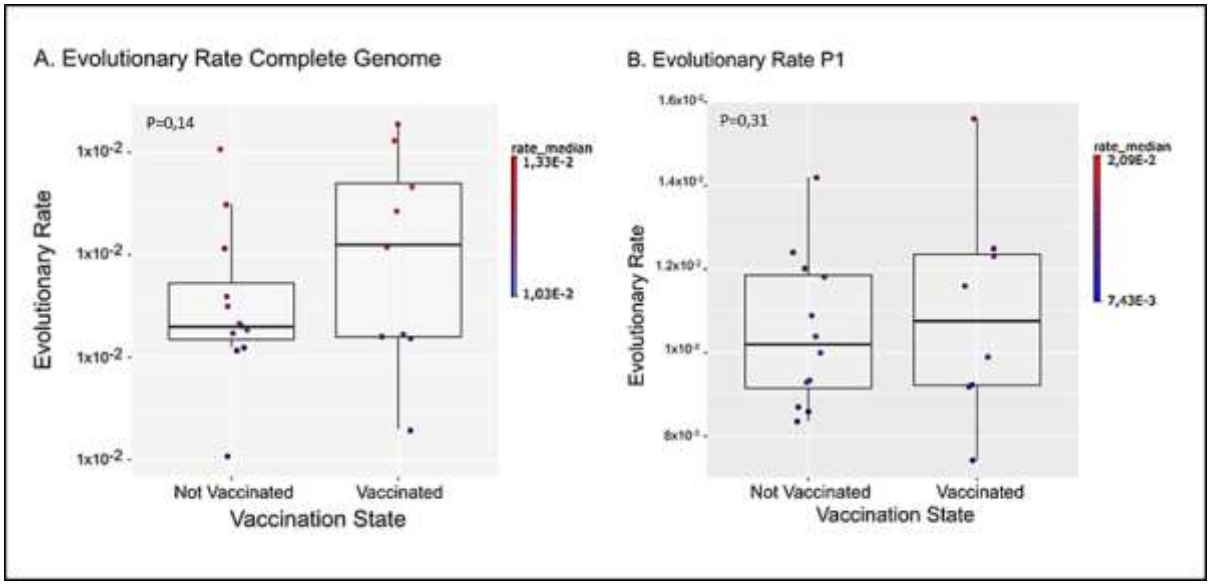

Figura 11. Gráfico de dispersión (Jitter plot) y análisis estadístico de Wilcoxon de las tasas evolutivas específicas de la rama terminal, calculadas para A) secuencias completas del genoma y B) secuencias de la región P1. Cada punto está coloreado según el valor de la tasa evolutiva para esa rama (rojo para las ramas más rápidas y azul para las ramas más lentas).

Por otro lado, se reconstruyó la dinámica del tamaño efectivo de la población durante el brote utilizando un gráfico Skyride (Figura 12). Este gráfico muestra un aumento progresivo de la población viral durante el período estudiado. A modo de comparación, también se graficó el número total de casos reportados de FA en el partido (Figura 13).

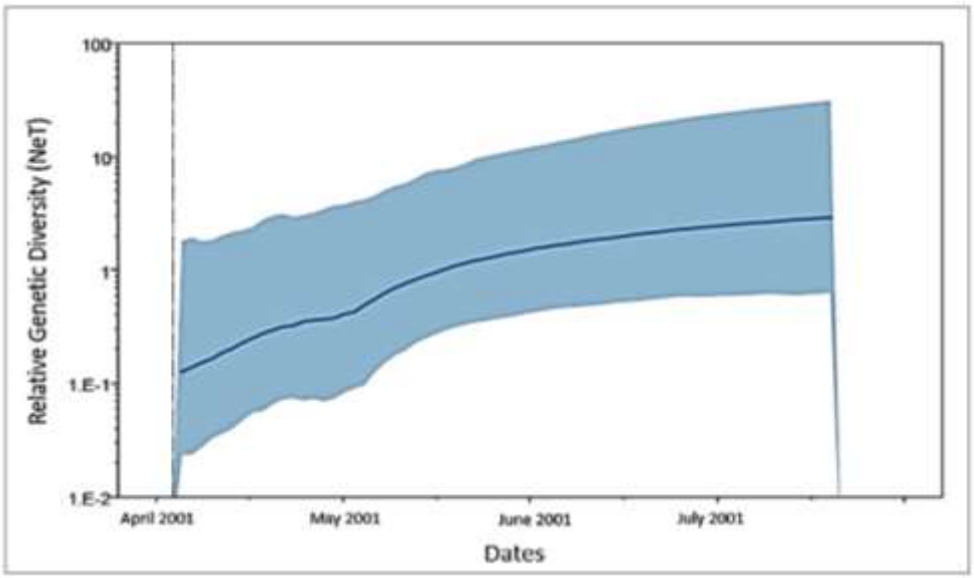

Figura 12. Gráfico Skyride bayesiano que muestra el tamaño efectivo estimado (Ne) de la población a lo largo del tiempo, basado en las secuencias del VFA de Mar Chiquita, Argentina.

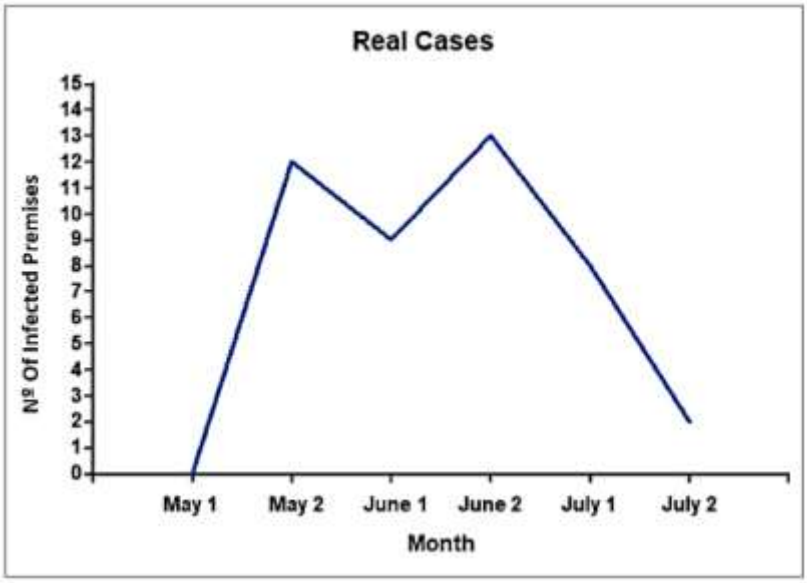

Figura 13. Gráfico de los casos reportados de FA en Mar Chiquita en 2001 agrupados en quincenas.

Por último, la reconstrucción filogeográfica (Figura 14) revela que el VFA se propagó inicialmente hacia el sur, donde se localizaba la primera granja que informó signos clínicos compatibles con el virus en Mar Chiquita. Posteriormente, se dispersó rápidamente hacia el noroeste y, finalmente, hacia el este, posiblemente desde granjas ubicadas en la región noroeste (video disponible en el Material Suplementario).

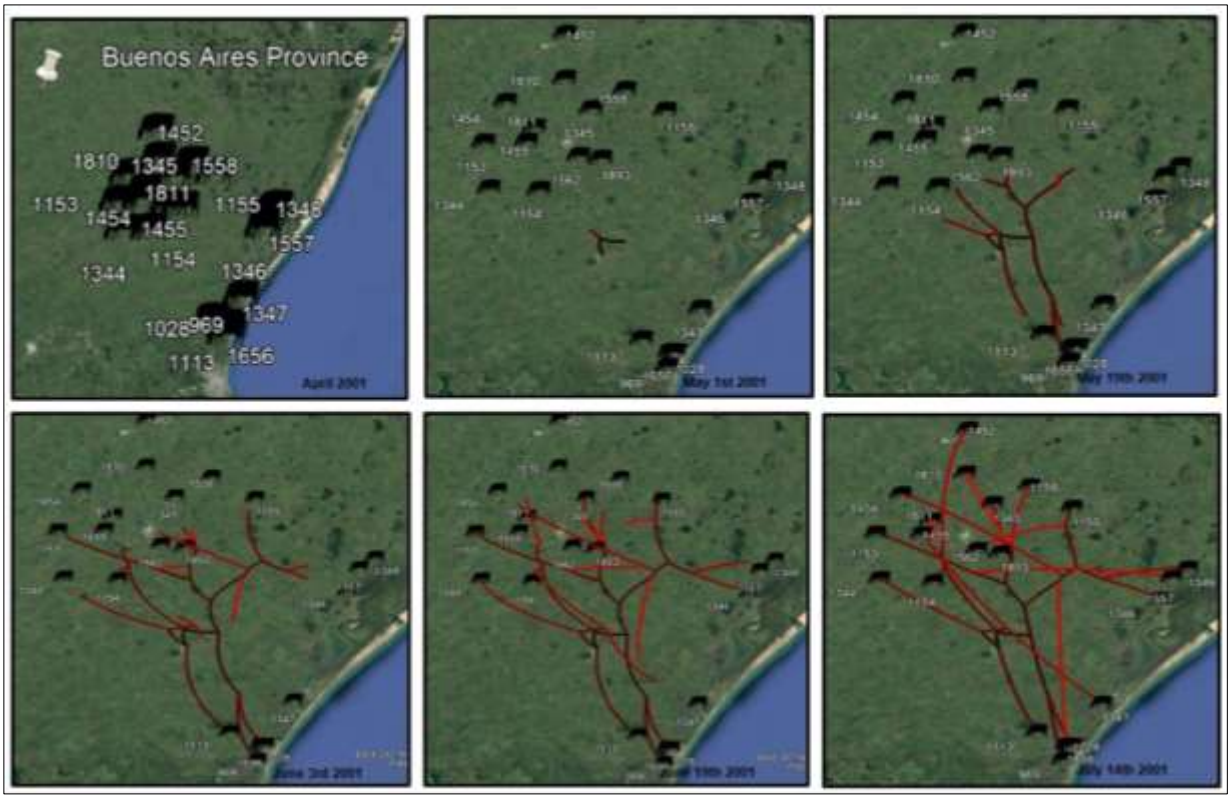

Figura 14. Reconstrucción filogeográfica de la dispersión del VFA en Mar Chiquita, Argentina. Las infecciones más antiguas se marcan en negro, degradándose a las más recientes en rojo. También se indican las fechas de las capturas.

Resultados de la epizootia en la cuenca lechera

Estadísticas de secuenciación

Las 15 secuencias analizadas, tuvieron una cantidad de aproximadamente 145.000 lecturas, con una longitud entre 45 y 251 pb. La mayoría exhibió una calidad mayor a 30 (con excepción de los extremos que fueron eliminados del análisis de forma manual).

Análisis filogenético

Las 15 secuencias procedentes de la cuenca lechera fueron alineadas con secuencias de genoma completo previamente publicadas durante la epizootia de 2000/1 (Brito et al., 2016). En el árbol filogenético resultante fue posible observar una distinción notable en comparación con las secuencias de Mar Chiquita. En lugar de formar un clúster privado, las secuencias de la cuenca lechera se distribuyen heterogéneamente a lo largo del árbol (Figura 15 A y B). La figura también revela que la secuencia 467, objeto de análisis en este estudio, presenta una sólida relación filogenética y una posible ascendencia con el clúster formado por las secuencias de Mar Chiquita. Esta relación tiene un muy buen soporte estadístico (UFB del 100%, rama A.1). Adicionalmente, en la misma rama y siguiendo las secuencias más cercanas a la raíz, es posible observar la muestra 1165 (también secuenciada en este trabajo), que muestra una relación filogenética importante con las secuencias pertenecientes a la rama A.1 (100% UFB), dicha rama la componen secuencias de La Pampa y de Buenos Aires. Con respecto a la rama A.2 es importante destacar que fue el lugar donde se ubicaron la mayor parte de las secuencias aportadas por este trabajo, exponiendo una relación filogenética cercana de las muestras analizadas en esta tesis provenientes de los partidos de la provincia de Santa Fe (Castellanos, Las Colonias, San Justo, Iriondo) y Córdoba (San Justo), sin embargo las mismas también comparten rama con secuencias de Santiago del Estero, Entre Ríos, Tucumán, Formosa, San Luis y Buenos Aires. Este resultado, refleja la similitud genética de las secuencias de la cuenca lechera con varias secuencias provenientes de diferentes lugares a lo largo de Argentina. Para terminar, en este análisis se siguió observando la formación de un pequeño grupo de secuencias filogenéticamente más distantes (Rama B) como lo observado en el trabajo de Brito et al., 2016.

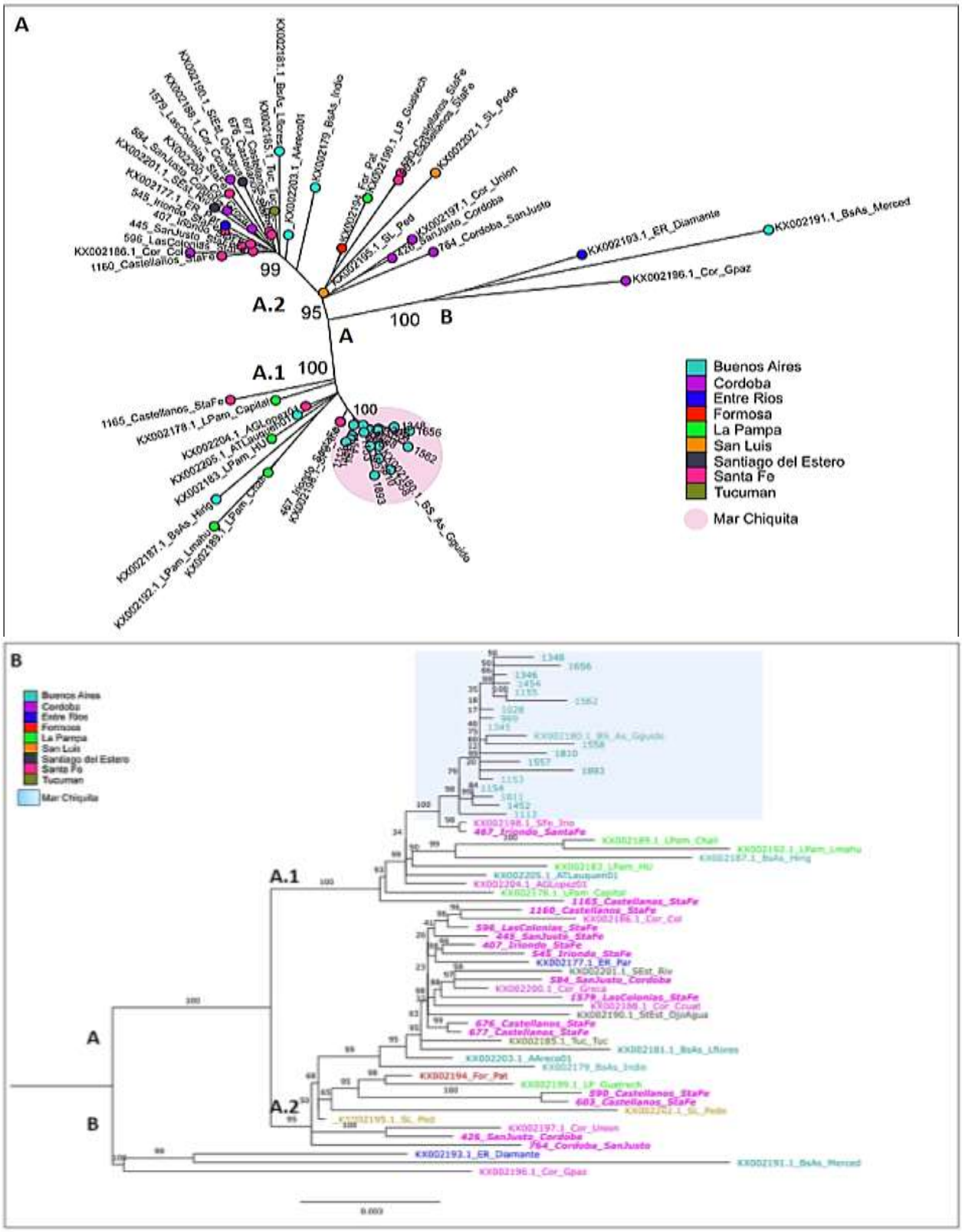

Figura 15. Árbol filogenético de Máxima Verosimilitud estimado a partir de las 15 secuencias analizadas en este trabajo provenientes de la cuenca lechera. 15 A es la representación radial de la filogenia y 15 B la representación rectangular de la misma. En negrita e itálica están señaladas las muestras analizadas en este trabajo. La provincia de origen de todas las secuencias esta detallada por colores y los números junto a las ramas representan el valor de UFB. La filogenia fue construida utilizando el programa IQTree v. 1.4.2

Discusión y conclusiones

Partido de Mar Chiquita

La integración de datos epidemiológicos con secuencias virales realizada en este estudio permitió analizar en detalle la dinámica de transmisión y evolución del virus de la Fiebre Aftosa en Argentina durante el brote de 2001 a nivel local. A partir de esta integración se pudieron desarrollar hipótesis sobre la dispersión espacial y el crecimiento epidémico de la población viral en el partido. Las hipótesis acerca de la transmisión del virus se basaron en tres conceptos generales derivados de la literatura científica. En primer lugar, se estimó que el número promedio de cambios nucleotídicos en un evento de contagio entre granjas es de 4.3 sustituciones nucleotídicas, con un desvío estándar de 2.1, según el estudio previo de (Cottam et al., 2008). En segundo lugar, se consideró que el VFA puede propagarse por transmisión aérea hasta una distancia de 60 km (Brown et al., 2022). Por último, se tomó en cuenta que el período promedio que toma el virus para manifestar la infección en los animales es de aproximadamente 1.7 días, basado en datos de Charleston et al., 2011. Nuestro estudio indica que todas las secuencias virales analizadas provienen de un ancestro común datado, aproximadamente, apenas un mes antes del primer informe de la enfermedad en el partido. Debido a la poca información sobre secuencias virales en los partidos vecinos, resultó imposible determinar si hubo más de una introducción del virus en Mar Chiquita. Sin embargo, nuestro análisis sí indica que, en caso de haber más de una, los virus deben haber sido genéticamente muy similares entre sí. Al tratarse de una zona geográfica aislada, es posible que el virus haya experimentado un único evento de transmisión con efecto fundador. Este fenómeno podría haber reducido la variabilidad genómica de las variantes circulantes, explicando así la formación del clúster privado evidenciado en el árbol filogenético (Bonneau et al., 2001).

El origen de la transmisión del virus dentro del partido sigue siendo incierto. Aunque el modelo filogeográfico sugirió una dispersión radial desde el área central de Mar Chiquita, este resultado podría haber sido significativamente influenciado por la distribución espacial de los brotes estudiados y las secuencias muestreadas (Kalkauskas et al., 2021). Varios estudios han informado que el virus puede sobrevivir en el medio ambiente durante largos períodos. Por ejemplo, la mediana de supervivencia es de 28.5 días en agua y aproximadamente 23.5 días en suelo y vegetación (Mielke and Garabed, 2020). Es posible que la propagación efectiva del virus no haya comenzado hasta que encontró un número suficiente de animales susceptibles. También es posible que el virus se haya transportado desde su ubicación original mediante el contacto entre animales o fómites hacia las primeras explotaciones que informaron signos clínicos en la zona sur del partido.

En cuanto a la transmisión dentro del partido, pudimos identificar al menos tres grupos independientes de transmisión. No podemos descartar la existencia de otros grupos, ya que solo analizamos 21 de las 42 explotaciones en la región que reportaron diagnóstico de laboratorio positivo de VFA. El primer grupo en el árbol filogenético (A) tuvo un buen respaldo estadístico y se dispersó solo en el noroeste en animales no vacunados. El grupo, B, también tenía un buen respaldo en los árboles filogenéticos, tanto en el valor de Ultra Fast Bootstrap como en la Probabilidad Posterior. La ruta de dispersión del grupo B fue primero hacia el sur y luego regresó al noroeste. Sin embargo, la presencia de ancestros no muestreados también

podría indicar una dispersión paralela en ambas direcciones. Un alto número de diferencias nucleotídicas entre explotaciones (hasta diecinueve) y la forma de la red filogenética indican la presencia de varios ancestros intermedios desconocidos. Finalmente, el grupo C presentó la mayoría de las secuencias estudiadas en este trabajo y probablemente fue el grupo más exitoso en dispersase en el partido. El mismo estuvo presente en el noroeste, este y sur en animales vacunados y no vacunados. Al igual que los otros dos grupos, es un grupo filogenético bien respaldado.

Si bien la observación de la red filogenética puede llevar a hacer suposiciones sobre las transmisiones del virus entre rodeos de la región, esto tiene sus limitaciones. Un ejemplo claro es el caso índice. A pesar de que el primer caso reportado en Mar Chiquita fue la muestra 969, en la red filogenética parece derivar de la secuencia (1345), de la cual se desprenden varios otros brotes. Esta información contradictoria no puede resolverse sin la disponibilidad de información epidemiológica o filogenética adicional. Además, existen otras secuencias relacionadas con la muestra 1345 que difieren significativamente en términos de un elevado número de cambios nucleotídicos (>7) y fechas de muestreo distintas (>20 días), lo que sugiere la existencia de explotaciones intermedias infectadas en la ruta de transmisión entre estos sitios. Por otro lado, menores diferencias nucleotídicas (<4) entre otras granjas sugieren que las conexiones entre ellas podrían ser más directas, como se observa en los casos 969-1345 y 1028-1345. Por su parte, factores como la prohibición de movimiento de animales y la gran distancia geográfica (>50 km) entre estas producciones indican que la transmisión entre ellas podría haber ocurrido a través de fómites, como por ejemplo maquinaria, agua o transporte humano (Auty et al., 2019). En estos casos, la transmisión aérea también podría haber tenido lugar, pero solo si el virus hubiera tenido condiciones ideales para su supervivencia (Brown et al., 2022). Aunque esta hipótesis de contacto directo entre explotaciones resulta atractiva, no puede confirmarse sin contar con información adicional, como secuencias de brotes no analizados, la variabilidad de secuencias virales dentro de una explotación, los movimientos de animales y personas en la zona, el rastreo de contactos y las condiciones climáticas, entre otros. Es importante señalar que el análisis filogenético por sí solo, rara vez proporciona suficiente evidencia para afirmar la transmisión directa y puede llevar a conclusiones erróneas en ausencia de información complementaria y verificable (Vandamme and Pybus, 2013).

La tasa de sustitución nucleotídica calculada en este estudio ($1{,}17\times10^{-2}$ sustituciones por sitio por año) fue coherente con estimaciones previas (Di Nardo et al., 2021). Dado que algunas explotaciones ganaderas habían sido previamente vacunadas con una vacuna que contenía cepas de serotipo A en su composición, pero no la A2001, se realizaron pruebas para detectar posibles diferencias en la tasa de sustitución entre estas explotaciones y las no vacunadas. A pesar de la dificultad de predecir la evolución viral (Zwart et al., 2023), se evaluó la presencia de un posible cuello de botella (Markov et al., 2023) causado por una vacunación, que solo protegía parcialmente contra la cepa A2001 (Mattion et al., 2004). Sin embargo, no se encontraron diferencias significativas entre las granjas vacunadas y las no vacunadas, ni en el genoma completo ni en la región P1, que expresa las proteínas de la cápside de la vacuna aplicada. Una posible explicación para este resultado podría ser que la fecha de vacunación fue muy cercana al momento de la infección (con un promedio de seis días), lo que no permitió tiempo suficiente para que se generara una respuesta inmunológica que ejerciera una presión de selección significativa y afectara la velocidad de evolución del virus.

Otra explicación podría ser que el estado de vacunación de las explotaciones no se informó adecuadamente en todos los casos.

En cuanto a la dinámica poblacional de este grupo de secuencias, se puede concluir que las curvas son bastante consistentes entre el modelo filodinámico de Skyride y la representación del número de casos totales reportados a lo largo del tiempo, lo que muestra un crecimiento progresivo de la población viral. Los datos epidemiológicos indican que hubo un crecimiento constante del virus hasta que las medidas de control lograron contener los brotes.

A pesar de que se lograron esclarecer parcialmente las rutas y los pasos de los eventos de transmisión, también se identificaron algunas lagunas y contradicciones en la información, marcando limitaciones en las conclusiones del trabajo. No obstante, es importante destacar que estas cuestiones podrían abordarse de manera más efectiva en un estudio de casos en tiempo real mediante un nuevo muestreo, la secuenciación de más muestras, la verificación de las condiciones climáticas o la investigación de posibles formas de contacto entre las granjas. Lamentablemente, estas acciones no pueden llevarse a cabo en un estudio retrospectivo de un caso que tiene más de 20 años.

En caso de querer analizar un hipotético futuro brote de la enfermedad, realizando un estudio de transmisión, sería recomendable secuenciar todas las explotaciones afectadas y tomar más de una muestra de animal por explotación. Esta metodología aplicada para el VFA también podría utilizarse para investigar las cadenas de transmisión de muchos virus de ARN que evolucionan rápidamente, como, por ejemplo, el virus de la influenza aviar o humana.

Cuenca lechera.

Ya que el análisis filogenético realizado rechaza la hipótesis inicial planteada de que las secuencias de la cuenca lechera formaban un clúster privado debido a su similitud en la producción y cercanía en su ubicación geográfica, no fue posible realizar un estudio de transmisión entre rodeos como se efectuó para el caso de Mar Chiquita. Sin embargo, nuestros análisis revelaron una fuerte relación filogenética entre las secuencias de Santa Fe/Córdoba y varias secuencias de diferentes lugares del país. Este hallazgo nos lleva a proponer la hipótesis de que Santa Fe y Córdoba podrían haber recibido variantes de la cepa A/Argentina 2001 desde distintos lugares y que, a su vez, estas provincias podrían ser puntos clave desde los cuales el virus se dispersó a otros lugares. Esta hipótesis se respalda epidemiológicamente por la ubicación geográfica central de la cuenca lechera, lo cual facilitaría una mayor circulación de animales y personas, propiciando la replicación y dispersión del virus. Este evento también fue identificado en el trabajo de Munsey et al., 2019, donde los autores señalaron zonas centrales con alta comercialización de ganado como un factor de riesgo para la presentación de FA. Adicionalmente, en el trabajo publicado por Brito y colaboradores en 2016, los modelos de epidemiologia molecular implementados a una escala nacional para estudiar la epidemia de FA en Argentina, identifican un virus ancestral en la zona central del país, que posteriormente se diseminó en tres clados filogenéticos centrales. Los hallazgos de estas tesis, respaldan la hipótesis que la cuenca lechera pudo haber sido un punto clave desde donde se dispersó el VFA a diferentes provincias argentinas. Al ser esta una región con una ganadería intensiva, se pudieron crear condiciones propicias para un aumento exponencial en el número de casos que podrían haberse mantenido en niveles más bajos desde la introducción del virus procedente del norte del país (Gotte 2022).

Este patrón contrasta con el observado en Mar Chiquita, dónde se plantea la posibilidad de que, al tratarse de una zona geográfica más aislada, el virus haya experimentado un único evento de efecto fundador o la introducción de pocos eventos de virus muy similares. Al profundizar en la dinámica de transmisión y evolución del virus a nivel nacional, con la información adquirida al incorporar nuevas secuencias, se plantea la posibilidad de que algunas de las muestras virales de Santa Fe, como la 467 y la 1165, sean antecesoras del brote en Mar Chiquita, ya que temporal y filogenéticamente son muy similares y las mismas están más cercanas a la raíz del árbol filogenético.

En este análisis, se evidenció la formación de dos clados filogenéticos sólidamente respaldados. Además, se observa una diversificación de la rama A en dos subramas, mientras que en el trabajo de Brito et al, 2016, las subramas son consideradas como ramas independientes. De todas formas, tanto en el estudio previo como en el nuestro, las ramas identificadas como A y B por Brito y las ramas A.1 y A.2 en nuestro análisis, están filogenéticamente más cercanas entre sí que la rama C en el trabajo de Brito y la rama B en nuestro análisis. Este grupo parece haber tenido menos éxito en su dispersión, ya que solo agrupa tres secuencias provenientes de las provincias de Buenos Aires, Entre Ríos y Córdoba.

Por otro lado, la rama A.2 en nuestro trabajo es la que sugiere haber tenido mayor éxito en su replicación ya que agrupa a la mayoría de secuencias aportadas por este trabajo y además involucra secuencias provenientes de las provincias de Entre Ríos, Santiago del Estero, Tucumán, Buenos Aires, Formosa, San Luis y La Pampa. La rama A.1 muestra una diseminación más acotada, que solo involucra secuencias de Buenos Aires (dentro de este el clado las secuencias de Mar Chiquita), La Pampa y Santa Fe.

Para concluir, es importante destacar la importancia de incorporar nuevas secuencias en los análisis, ya que esto posibilita una comprensión más detallada de la dinámica de dispersión del VFA. Asi mismo, pone de manifiesto la relevancia de mantener una actualización constante de secuencias y datos epidemiológicos y desarrollar nuevos análisis y plantear nuevas hipótesis a partir de ellos.

Conclusiones generales del capítulo 1

Los análisis de epidemiologia molecular han proporcionado una compresión más detallada sobre la dinámica que tuvo el virus de la Fiebre Aftosa durante la epidemia en Argentina en el año 2001. Fue posible identificar diferentes patrones de dispersión que posiblemente fueron influenciados por el lugar geográfico en el cual se desarrollaron. Se exploró la dinámica de transmisión local y evolución viral en el partido de Mar Chiquita, proponiendo la ocurrencia de un evento de efecto fundador en esta zona geográfica más aislada. También fue posible esclarecer y refinar más los análisis que se tenían previamente de la transmisión a nivel nacional. Por otra parte, si bien no se pudieron sacar conclusiones acerca del efecto de la vacunación sobre la evolución viral, sí se pudieron realizar hipótesis acerca de la transmisión del virus a corta y larga distancia y la importancia que pudieron tener los fómites durante la epizootia. Todos estos aportes permiten tener datos y metodologías de laboratorio y de

análisis disponible para ser utilizado en una posible epidemia futura. Finalmente, se enfatiza sobre la necesidad de recopilar y analizar más muestras representativas a nivel de granja, partido y provincia en el futuro, con el fin de obtener una visión más detallada de la dinámica de transmisión del virus. Además, se destaca la importancia de continuar sumando más información a los conjuntos de datos previamente analizados para refinar, confirmar o descartar nuevas hipótesis. Esto permitirá que los responsables del sector de salud pública y animal tomen decisiones cada vez más informadas para el manejo y control de la enfermedad en situaciones futuras.

CAPITULO 2

MODELOS DISCRETOS A UNA ESCALA INTER E INTRA CONTINENTAL

CON UN ENFOQUE EN VIRUS DE ARN

Dinámica de los virus IBDV y PRRSV

CAPITULO 2 - OTROS VIRUS DE ARN

Sección 1 IBDV

Materiales y Métodos

Obtención de secuencias

Se obtuvieron 11 secuencias recolectadas en la República Argentina que corresponden al fragmento que codifica para la proteína VP2 de IBDV y provienen de muestras tomadas de la Bolsa de Fabricio (BF) de pollos que presentaban signos clínicos compatibles con la infección por el virus.

Para obtener las secuencias, se llevó a cabo la extracción del ARN total a partir de los tejidos de la BF, utilizando TransZol (TransGen Biotech, Beijing, China) siguiendo el protocolo establecido por el fabricante. Posteriormente, se verificó la concentración y la pureza del ARN mediante el uso de Nanodrop N100 (Thermo ScientificTM, Wilmington, EE.UU.) y electroforesis en geles de agarosa. Una vez obtenido el ARN, se procedió a la síntesis de ADNc utilizando un microgramo de muestra de ARN y la enzima MMLV (Promega, EE.UU.) en una reacción de 20 µL totales. Esta reacción de síntesis de ADNc se realizó en un termociclador Biometra, siguiendo las recomendaciones del fabricante.

La amplificación de los productos de IBDV se obtuvo mediante una reacción de PCR utilizando la enzima Taq DNA Polymerase (Promega, Madison, WI, EE. UU.) con cebadores específicos para la región hipervariable de VP2, cuyas secuencias fueron Fw: GCCCAGAGTCTACACCAT y Rv: ATGGCTCCTGGGTCAAATCG (5' a 3'). El ciclo de amplificación constó de una fase de desnaturalización (5 min a 95 °C), seguida por 35 ciclos de desnaturalización (30 s a 95 °C), hibridación (1 min a 55 °C) y extensión (2 min a 72 °C) y finalmente una fase adicional de extensión de 10 min a 72ºC. Los productos de amplificación resultantes se visualizaron en un gel de agarosa al 1.5% teñido con Bromuro de Etidio, buscando un producto específico de 579 bases. Una vez obtenidos los productos de amplificación, estos fueron purificados y enviados para su secuenciación. Este paso fue llevado a cabo por la empresa HIPRA.

Curado de secuencias y diseño del dataset enfocado a la determinación del genogrupo

Una vez obtenidos los datos de secuenciación correspondientes al fragmento del genoma de IBDV que codifica para la proteína VP2, se procedió a realizar una revisión empleando el *software* BioEdit (Hall, 1999). Cada archivo de secuencia fue inspeccionado visualmente y se realizaron las ediciones necesarias para resolver cualquier ambigüedad presente. Se efectuaron recortes en las regiones de los extremos de los fragmentos en los que se detectó una calidad de lectura deficiente. Además, se diseñó un dataset que incorporó 91 secuencias de referencia previamente clasificadas por genogrupo (Michel and Jackwood, 2017), con el objetivo de determinar a cuál pertenecían las 11 secuencias argentinas que fueron objeto de

nuestros análisis. Todas estas secuencias fueron alineadas utilizando el paquete R DECIPHER v2.0 (Wright, 2016).

Análisis filogenéticos enfocados a determinación de genogrupo

La estimación de los árboles filogenéticos se llevó a cabo empleando el enfoque de Máxima Verosimilitud, el cual se aplicó utilizando el software IQ-Tree versión 1.4.2 (Trifinopoulos et al., 2016). La selección del modelo de sustitución correspondiente se realizó dentro de la misma plataforma, mediante el uso del módulo ModelFinder. La construcción del árbol se basó en el Criterio de Información Bayesiano (BIC) y se realizaron 10,000 réplicas de UFB como medida de respaldo. Los resultados obtenidos se representaron visualmente utilizando la herramienta FigTree v.1.4.0 (Rambaut, 2009).

Diseño del dataset enfocado a conocer la historia evolutiva

Para determinar la historia evolutiva que tuvieron las secuencias de estudio, fue construido un dataset que incluyó aproximadamente 1200 secuencias pertenecientes al genogrupo 2, todas disponibles en el GenBank, más las secuencias argentinas que fueron el foco principal de la investigación. Las secuencias fueron alineadas utilizando el paquete R DECIPHER v2.0 (Wright, 2016). Finalmente se excluyeron del análisis aquellas secuencias que presentaban información epidemiológica incompleta, longitud insuficiente o nucleótidos indeterminados.

Análisis de secuencias recombinantes

Siguiendo el protocolo establecido, se procedió a realizar una identificación preliminar de posibles eventos de recombinación en el conjunto de datos utilizando la herramienta bioinformática RDP4 (Martin et al., 2017) y aplicando métodos que incluyeron el chequeo por BOOTSCAN41, MAXCHI42, CHIMAERA43, 3SEQ44, GENECONV45 y SISCAN46. En caso de detección de señal de recombinación mediante alguno de estos métodos, la o las secuencias se excluyeron del conjunto de datos con el fin de garantizar la coherencia en las historias evolutivas representadas por cada secuencia y evitar cualquier ambigüedad en los resultados.

Análisis filogenéticos enfocados a conocer la historia evolutiva

La inferencia de los árboles filogenéticos se realizó mediante el método de Máxima Verosimilitud utilizando el software IQ-Tree versión 1.4.2 (Trifinopoulos et al., 2016). La elección del modelo de sustitución se efectuó en la misma plataforma utilizando el módulo ModelFinder. La construcción del árbol se basó en la aplicación del Criterio de Información Bayesiano (BIC) y se realizaron 10,000 réplicas de Ultra Fast Bootstrap como medida de soporte. La representación gráfica de los resultados se llevó a cabo utilizando la herramienta FigTree v.1.4.0 (Rambaut, 2009).

Análisis filogeográficos y evolutivos

Se empleó el software TempEst (Rambaut et al., 2016) para evaluar la señal temporal del dataset. Posteriormente, se llevaron a cabo análisis filogenéticos y filodinámicos Bayesianos utilizando el programa BEAST versión 1.10.4 (Suchard et al., 2018) que permitieron determinar la tasa evolutiva viral, inferir un Árbol de Maxima Credibilidad, establecer la fecha

del ancestro común más reciente y reconstruir la dispersión filogeográfica del virus. En este análisis se utilizó un modelo de difusión discreta, combinando un reloj molecular relajado Lognormal No Correlacionado con un modelo poblacional coalescente SkyRide. Asi mismo, se incorporaron las fechas de muestreo y el país de origen de cada secuencia. Se ejecutaron dos cadenas independientes de MCMC, cada una con un total de 300 millones de generaciones. La convergencia se evaluó utilizando el programa Tracer v.1.7 (Rambaut et al., 2018) y únicamente fueron tenidos en cuenta los parámetros que alcanzaron un tamaño de muestra efectiva mayor a 200. El 10% de los resultados iniciales de la MCMC se descartó como "Burn-in", y posteriormente, se determinó el Árbol de Credibilidad Máxima utilizando la herramienta TreeAnnotator (disponible en el paquete de BEAST). La visualización del árbol de MCC se realizó mediante la herramienta Figtree. Finalmente, para la visualización de la reconstrucción filogeográfica, se empleó el programa SpreaD4 (Kanika D Nahata et al., 2022).

Resultados

Análisis filogenéticos enfocados a determinación de genogrupo

El análisis filogenético que abarcó nuestras once secuencias junto con 91 secuencias de referencia asignadas a los genogrupos G1 a G7, utilizando la metodología de Máxima Verosimilitud (figura 16), identificó que todas las secuencias argentinas pertenecen al genogrupo 2 (G2). Además, se agrupan de manera conjunta en un nodo monofilético con un soporte máximo del 100% de UFB.

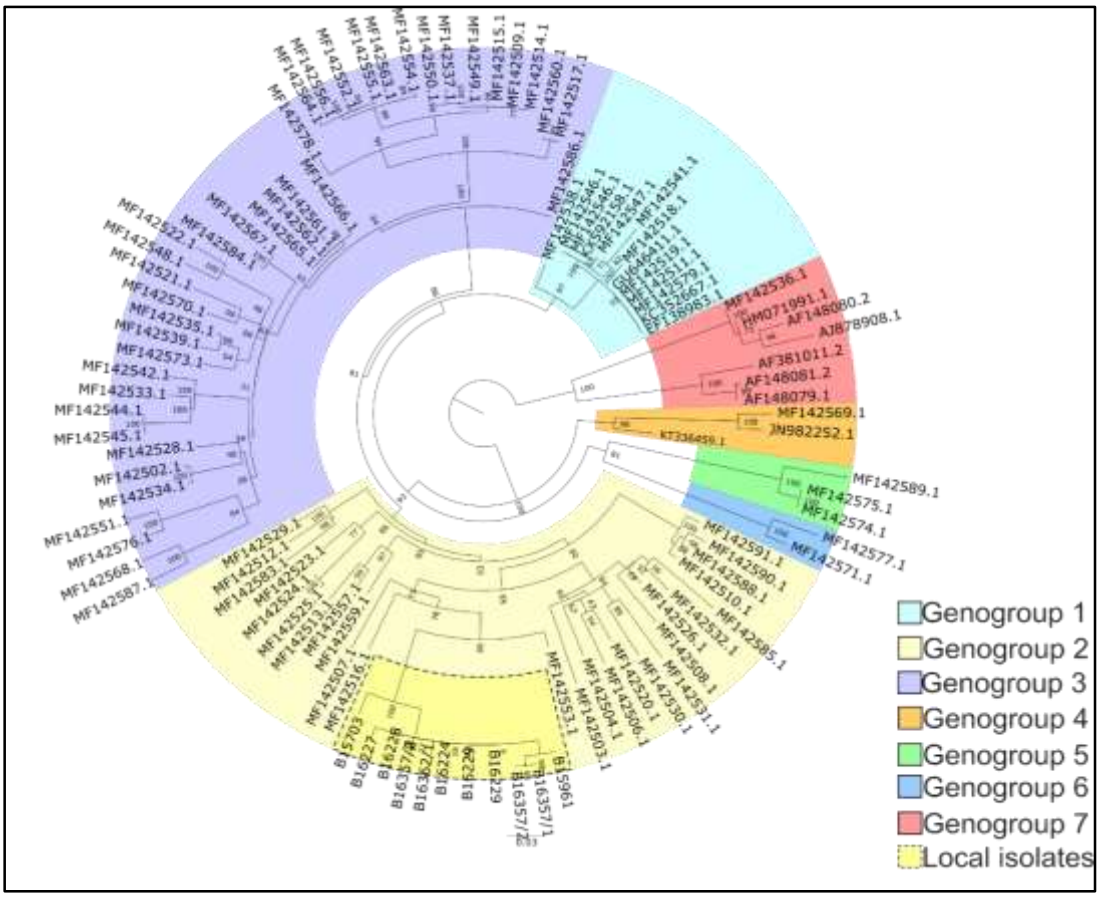

Figura 16. Árbol filogenético circular obtenido mediante el método de Máxima Verosimilitud. Cada color indica un genogrupo diferente, las secuencias argentinas de esta investigación están señaladas en línea punteada.

Análisis filogenéticos enfocados a conocer la historia evolutiva

Del dataset inicial que se componía de aproximadamente 1200 secuencias, el alineamiento final incluyó un total de 436 secuencias pertenecientes al G2 del IBDV, luego de aplicar los filtros de calidad de secuencia, de la información asociada a la misma y de la detección de posibles eventos de recombinación. El análisis filogenético bayesiano realizado mediante el programa BEAST (Figura 17) y el árbol generado con la metodología de Máxima Verosimilitud usando el programa IQTREE arrojaron resultados consistentes y concordantes.

En estos análisis, se pudo observar que la mayoría de las secuencias argentinas (10 de ellas) presentan una relación filogenética cercana con secuencias de China, respaldada por un sólido soporte de probabilidad posterior (PP:1). Sin embargo, una de las once secuencias parece estar filogenéticamente más relacionada con secuencias provenientes de Corea del Sur (Figura 17).

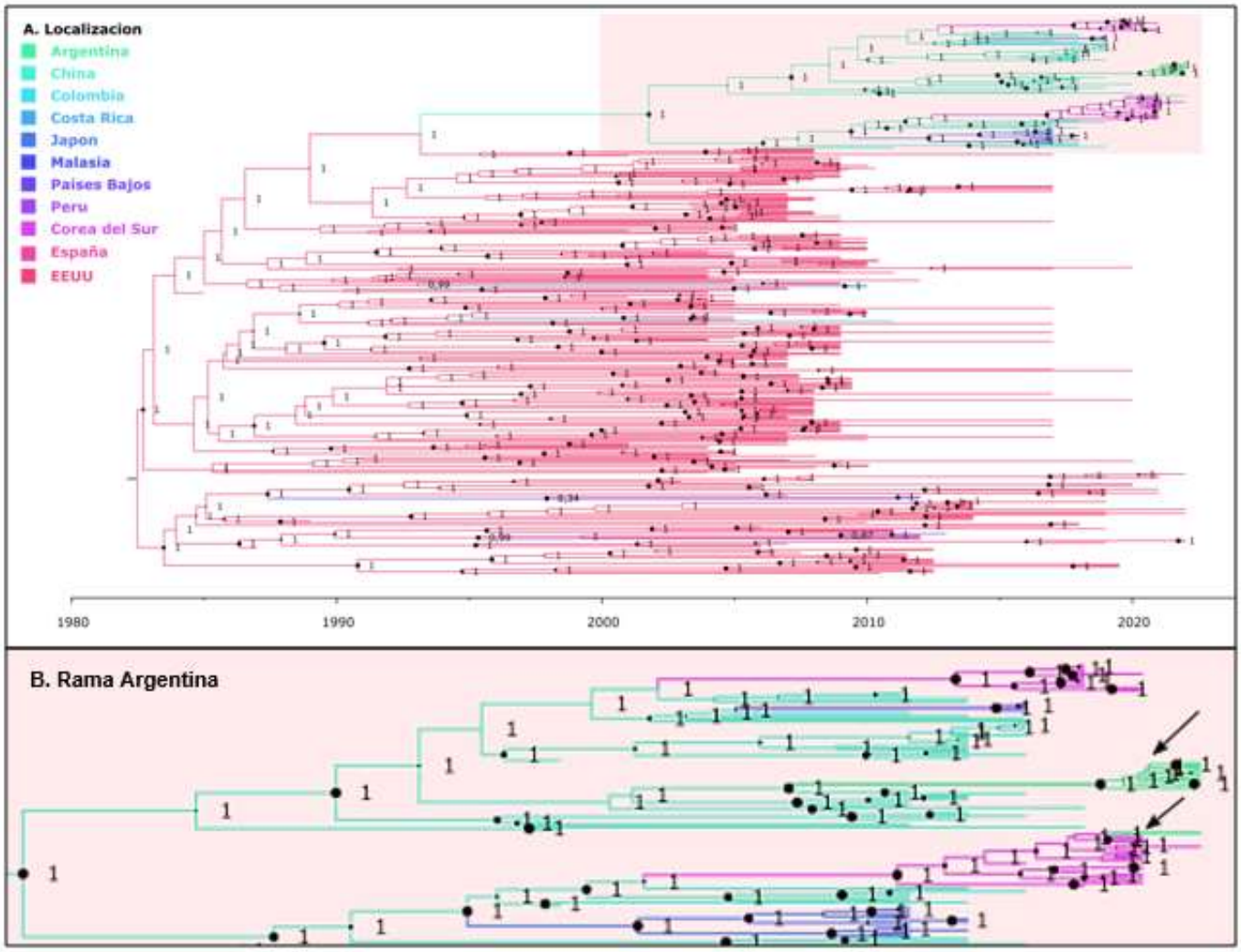

Figura 17. Árbol MCC datado, que muestra secuencias del genotipo 2 de IBDV. Los colores representan el país de origen de las secuencias, en el margen inferior se indican los años de presentación del foco, los números junto a los nodos indican la probabilidad de la localización, siendo 1 la máxima y 0 la mínima, El tamaño de los círculos en los nodos es dependiente a la probabilidad posterior del nodo, los más grandes representan PP de 1, el nodo que abarca las secuencias argentinas están sombreadas con color rosa y en la parte B de la figura se hace un Zoom de las mismas.

Análisis filogeográficos y evolutivos

Se identificó una correlación positiva entre el tiempo y la divergencia genética, con un valor de R-cuadrado de 0.59, lo que respalda la idoneidad del conjunto de datos para análisis posteriores en un contexto evolutivo. Como consecuencia, se llevaron a cabo dos corridas del programa BEAST y se verificó la convergencia de los parámetros, con todos los tamaños efectivos de muestra superando el umbral de 200. Nuestro análisis revela que el ancestro común más reciente de todas las muestras se originó en los Estados Unidos, aproximadamente en el año 1983 (Figura 18) con un intervalo de credibilidad del 95% (HPD) que abarca desde 1982 hasta 1984. Además, se evidencia la subsiguiente propagación del virus hacia China, Japón y Corea del Sur (todas con un soporte de probabilidad de localización geográfica del nodo de 1)

Según nuestro modelo, la mayoría de las secuencias argentinas (10) comenzaron a propagarse dentro del país alrededor de 2014, con un intervalo de HPD del 95% que abarca desde 2013 hasta 2015. Sin embargo, la secuencia restante sigue una ruta de transmisión distinta, en la que se destaca Corea del Sur como antecesor. Esta trayectoria sugiere una posible introducción alternativa en Argentina alrededor de 2020, con un intervalo de HPD del 95% que va desde 2019 hasta 2021. A pesar de pertenecer al mismo clado, esta última secuencia no comparte una relación directa con las otras diez.

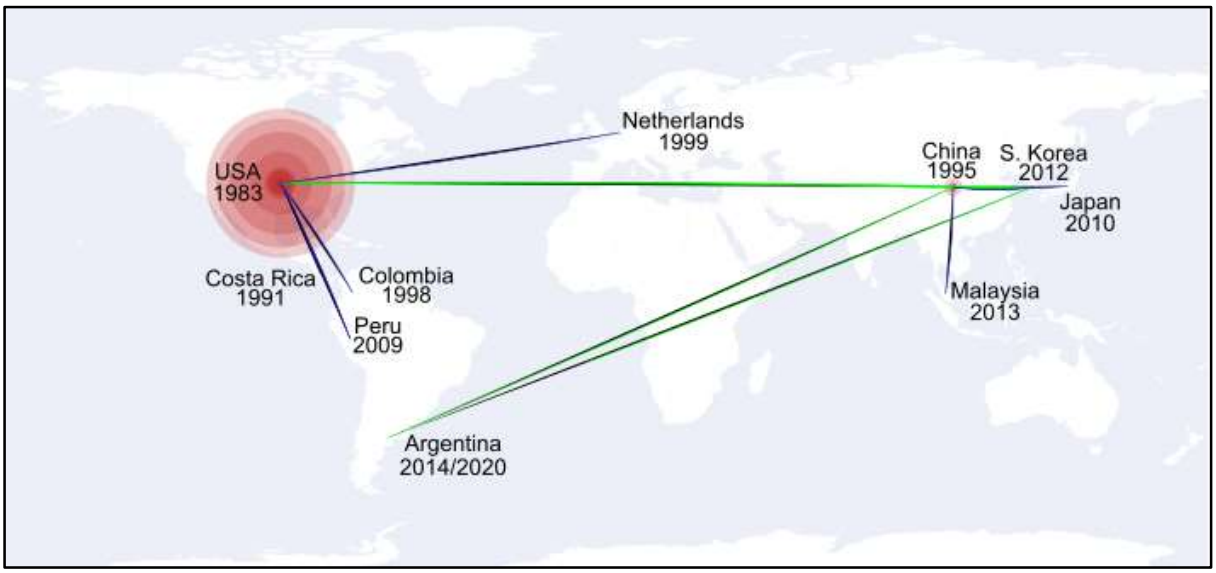

Figura 18. Reconstrucción filogeográfica de la dispersión del IBDV. Las líneas verdes indican las posibles rutas de entrada del genotipo 2 en Argentina. Se indica el año de la posible divergencia del ACRM para cada país.

Finalmente, se obtuvo una tasa de evolución de las secuencias de 1.74×10^{-3} sustituciones por sitio por año, con un intervalo de probabilidad del 95% que se sitúa entre 1.5×10^{-3} y 1.9×10^{-3}.

Discusión y conclusiones

En este trabajo, presentamos el primer informe molecular sobre la presencia del genotipo 2 del IBDV en Argentina. El análisis de secuencias, junto con sus datos epidemiológicos, nos permitió trazar la historia evolutiva parcial del IBDV a una escala intercontinental. Esta aproximación nos brindó la oportunidad de generar hipótesis sobre el posible recorrido del virus hasta su posterior introducción en Argentina.

Nuestros modelos de epidemiología molecular concuerdan con investigaciones previas que sugieren que los Estados Unidos podrían haber sido el origen del brote del genotipo 2 a nivel mundial entre 1982 a 1984 (Wang et al., 2022) Sin embargo, resulta sorprendente que China se presente como el ancestro en común más reciente de las secuencias argentinas, dado que existe una gran distancia geográfica entre Argentina y China, además de la falta de un comercio activo avícola, con dirección de Asia-Argentina. Aunque esto no descarta completamente la posibilidad de una transmisión directa, sí la reduce considerablemente. Por lo tanto, una hipótesis alternativa que podría plantearse es que existen eslabones intermedios

en la cadena de transmisión que podrían no haber sido muestreados. Actualmente, solo podemos obtener una visión parcial de la dinámica del genotipo basándonos en los datos estudiados. Para una comprensión más completa del escenario, es fundamental continuar con el análisis de los casos reportados y compartir esta información en bases de datos públicas. Esto podría contribuir significativamente a mejorar nuestro conocimiento sobre este virus en general y, conocer las rutas de transmisión en nuestro continente en particular. Como consecuencia, también es importante para obtener detalles más precisos sobre la introducción y distribución de este genotipo en Argentina.

Cabe destacar que este estudio no solo confirma la presencia del genotipo 2 en Argentina, sino que brinda las primeras evidencias de una posible doble introducción del genotipo en nuestro país. A pesar de que solo contamos con el respaldo de una secuencia que sugiere esta hipótesis, su presencia resulta altamente informativa y destaca la importancia de incrementar los esfuerzos para implementar un muestreo activo en las granjas afectadas que incluya el análisis molecular de las secuencias virales. Además, el estudio pone de manifiesto la complejidad de la dinámica de transmisión viral y permite tener una imagen más completa y precisa de las formas de introducción y propagación del virus en la Argentina.

La tasa evolutiva del virus de 1.74×10^{-3} ($1.5\text{-}1.9 \times 10^{-3}$) se encuentra en línea con lo reportado en la literatura (Yilmaz et al., 2019) y sugiere que el virus no experimentó una presión de selección significativa, sino que mantuvo su tasa evolutiva constante a lo largo del tiempo. En cuanto al grupo de las diez secuencias argentinas, observamos una tasa de evolución ligeramente superior, calculada en 1.95×10^{-3} sustituciones por sitio por año. Sin embargo, este fenómeno no es infrecuente al analizar muestras de brotes estrechamente relacionados producidos por virus de ARN.

En resumen, los modelos discretos aplicados a una escala intercontinental han proporcionado una visión parcial de la dinámica y evolución del genotipo 2 del IBDV en Argentina. A pesar de contar con un conjunto limitado de datos de secuencias, al ser combinados con las secuencias disponibles en bases de datos públicas, se logró confirmar la presencia del genotipo en el país además de alertar sobre un posible cambio en el genotipo predominante en esta región (el último reporte indicaba que era el genotipo 4). Esto ha permitido lanzar alertas sobre la posible evasión de la vacunación e incluso la aparición de brotes con diferentes niveles de patogenicidad en las aves. Además, los resultados del estudio permiten comenzar a realizar hipótesis sobre el origen de las rutas de transmisión y ofrecen información para analizar posibles fuentes de infección. Este estudio refuerza la importancia de mantener y fortalecer las medidas de vigilancia activa para preservar la salud avícola en Argentina, destacando la necesidad de un monitoreo continuo para abordar los desafíos sanitarios del país.

Sección 2 PRRSV

Materiales y métodos

Datos de secuencia

Para llevar a cabo este trabajo se obtuvieron secuencias de referencia del ORF5 del PRRSV de la base de datos del Centro Nacional de Información Biotecnológica (NCBI GenBank). En primer lugar, se seleccionaron secuencias originarias de países de América del Sur y luego estas secuencias fueron comparadas con otras secuencias públicas en la misma base de datos mediante el uso de la herramienta BLAST. Entonces, se agregaron al estudio las cinco secuencias que mostraron mayor similitud con aquellas provenientes de América del Sur, sin importar su lugar de origen. Además, a partir de una colaboración con el Laboratorio de Diagnóstico Veterinario de la Universidad de Minnesota (UMN) se enriqueció el conjunto de datos con aproximadamente 8000 secuencias de origen estadounidense y no disponibles públicamente. Por otra parte, se obtuvieron dos secuencias colombianas del año 2022, a través de un acuerdo con el Laboratorio de Diagnóstico de la Universidad CES, en una colaboración voluntaria con productores que utilizan de forma rutinaria los servicios de diagnóstico y secuenciación de la universidad.

Análisis filogenéticos enfocados a la determinación del genotipo

Se procedió a examinar el conjunto de secuencias de ORF5 mediante el *software* BioEdit (Hall, 1999) con el fin de corregir manualmente los errores de edición presentes en las secuencias. Se llevaron a cabo recortes en las regiones de los extremos de los fragmentos donde se identificó una calidad de lectura deficiente. Posteriormente, se seleccionaron las secuencias que cumplían con los criterios necesarios de calidad y longitud para el análisis, incluyendo aquellas previamente identificadas como pertenecientes al genotipo 1. Seguidamente, el conjunto de datos fue alineado mediante el software MAFFT versión 7 (Katoh y Standley, 2014). Por último, se llevó a cabo la construcción de árboles filogenéticos para confirmar que todas las secuencias estudiadas en este trabajo pertenecían al genotipo 2 del PRRSV. Para ello, se utilizó el método de Máxima Verosimilitud implementado en el programa IQ-Tree versión 1.4.2 (Trifinopoulos et al., 2016). La elección del modelo de sustitución se determinó utilizando el módulo ModelFinder dentro del mismo programa. El árbol se estimó basado en el Criterio de Información Bayesiano (BIC), con 10000 réplicas de UFB como método de soporte. Los resultados obtenidos se visualizaron utilizando la herramienta FigTree v.1.4.0 (Rambaut, 2009).

Identificación de secuencias recombinantes

El análisis de recombinación en el conjunto de datos de muestras del genotipo 2 se realizó utilizando el programa RDP4 (Martin et al., 2017), empleando una combinación de métodos que incluyeron BOOTSCAN41, MAXCHI42, CHIMAERA43, 3SEQ44, GENECONV45 y SISCAN46. Cuando se detectó alguna señal de recombinación mediante cualquiera de estos métodos, se procedió a eliminar la secuencia del conjunto de datos.

Análisis filogenéticos enfocados a conocer la historia evolutiva

Con un total de 3732 secuencias completas del gen ORF5, libres de señales de recombinación y exclusivamente pertenecientes al genotipo 2 del PRRSV, se llevaron a cabo análisis filogenéticos a través del programa IQTree (Trifinopoulos et al., 2016) siguiendo la metodología previamente descrita en este capítulo.

Reconstrucción de los estados ancestrales

La reconstrucción de los estados ancestrales se llevó a cabo utilizando el programa Past ML v. 1.9.7 (Ishikawa et al., 2019). Para la construcción de los árboles, aplicamos la metodología de Máxima Verosimilitud utilizando el programa IQTree2 (Minh et al., 2020), con cada rama datada, y registrando la información del país en las ramas terminales de los árboles. La reconstrucción y visualización de los estados ancestrales en Past ML se realizó mediante la interfaz gráfica disponible en el sitio web del Instituto Pasteur (https://pastml.pasteur.fr/).

Para la reconstrucción de los estados ancestrales se utilizó el modelo F81, que permite cuatro estados para la sustitución de nucleótidos (Felsenstein, 1981). Además, Past ML simula un modelo de reloj molecular estricto que se aplica de manera uniforme a todas las ramas del árbol. Para evaluar la robustez de la reconstrucción de los estados ancestrales, realizamos análisis aleatorios que se repitieron en tres ocasiones.

Análisis evolutivos enfocados a conocer la historia y ancestria por ramas individuales

A partir del árbol general, se identificaron todas las ramas de interés que contenían secuencias sudamericanas. A continuación, se creó un nuevo conjunto de datos específicos, para cada una de estas ramas, con el propósito de estudiar la evolución de cada una en particular. En primer lugar, se construyó un árbol filogenético preliminar mediante la metodología de Máxima Verosimilitud utilizando el programa IQTree (Trifinopoulos et al., 2016) y se evaluó la señal temporal empleando el software TempEst (Rambaut et al., 2016). Luego, se llevaron a cabo análisis filodinámicos Bayesianos utilizando el programa BEAST versión 1.10.4 (Suchard et al., 2018) para determinar la tasa evolutiva promedio de la rama y establecer la fecha del ancestro común más reciente.

Para este análisis se incorporaron las fechas de muestreo y se utilizó un modelo de reloj molecular Relajado Lognormal No Correlacionado junto con un modelo poblacional coalescente SkyRide. Se ejecutaron dos cadenas independientes de MCMC, cada una con un total de 50 millones de generaciones. La convergencia se evaluó mediante el programa Tracer v.1.7 (Rambaut et al., 2018) y solo se consideraron los parámetros que alcanzaron una muestra efectiva superior a 200. Se descartó el 10% inicial de la cadena MCMC como "Burn-in".

Resultados

Análisis filogenéticos enfocados a la determinación del genotipo

Fue posible observar una diferencia filogenética clara entre las secuencias pertenecientes a PRRSV1 y las secuencias pertenecientes a PRRSV2 en el árbol filogenético (Figura 19). Ambas ramas, que representan estos dos genotipos, están respaldadas por un sólido soporte de UFB. El 100% de las nuevas muestras analizadas para este estudio pertenecieron al genotipo 2. Cabe destacar que, para los análisis posteriores, únicamente se utilizaron las secuencias PRRSV2, ya que el interés epidemiológico estaba focalizado en América.

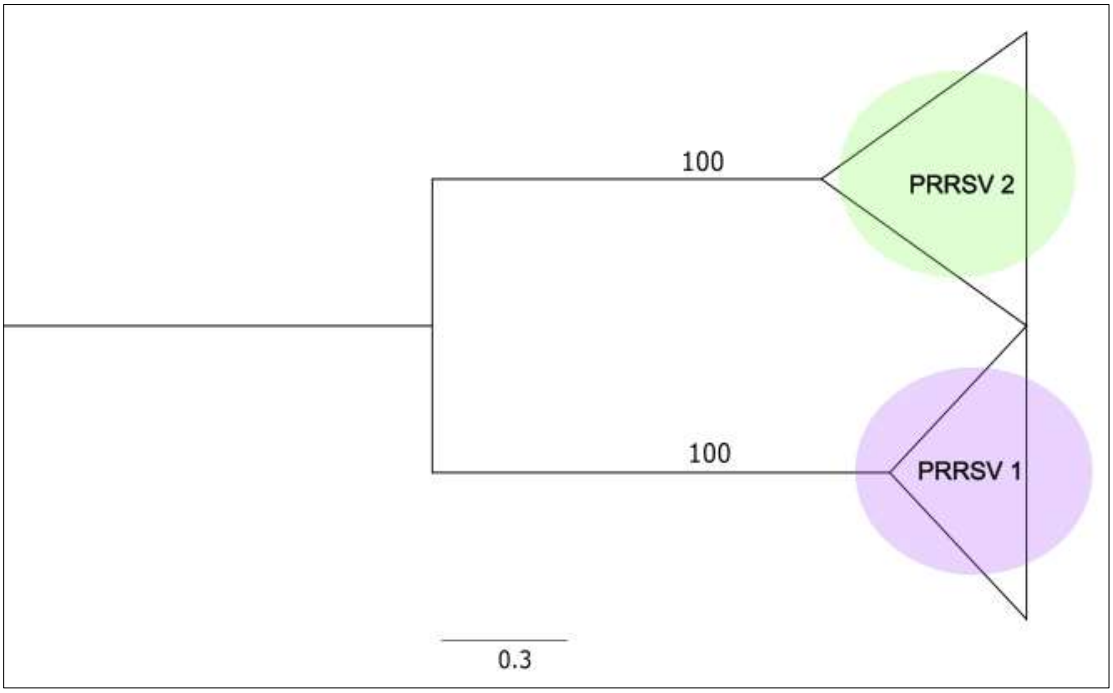

Figura 19. Árbol filogenético construido mediante la metodología de Máxima Verosimilitud, con sus ramas colapsadas para una mejor visualización. Se destacan en verde todas las secuencias pertenecientes al genotipo 2 de PRRSV, mientras que en la rama violeta se agrupan las secuencias correspondientes al genotipo 1 de PRRSV.

Análisis filogenéticos enfocados a conocer la historia evolutiva

Tras la aplicación de los filtros, el alineamiento resultante abarcó un total de 3732 secuencias correspondientes al genotipo 2 de PRRSV, libres de sospechas de recombinación. En este análisis, se observó que la mayoría de las secuencias sudamericanas no forman parte de un único clúster; sino que, por el contrario, cada una se posiciona en una rama individual compartida con diferentes secuencias estadounidenses (Figura 20). Además, las ramas principales están respaldadas por un valor de UFB mayor a 90%

El árbol resultante fue subdividido en rama A y B para su mejor comprensión. A partir de la inspección de cada rama se analizó la dinámica que tuvo el virus en cada una de ellas. En términos generales, se observó que únicamente las secuencias de Uruguay se posicionaron

en la rama A, la cual cuenta con un soporte de UFB del 95%, evidenciando una destacada divergencia filogenética frente a las demás muestras sudamericanas.

Aunque las secuencias representativas de Uruguay del año 2017 (3 secuencias) y del año 2020 (2 secuencias) son parte de una misma gran rama con soporte de UFB bajo (24%), cada grupo tiene mayor similitud filogenética con secuencias provenientes de EEUU que entre ellos. Por ejemplo las secuencias de Uruguay del 2020, se relacionan más cercanamente con un clado de secuencias estadounidenses con un soporte de UFB de 87%. Por otro lado, las secuencias de Uruguay del 2017 se relacionan más cercanamente con otro clado de secuencias estadounidenses con un soporte de UFB del 50% (Figura 21).

En cuanto a la rama B (Figura 20), esta incluye las secuencias de Chile, Colombia y Perú. Aunque las mismas comparten un ancestro común antiguo, muestran una significativa divergencia filogenética, situándose cada una en una rama considerablemente diferenciada y la mayoría bien soportada (secuencias de Chile de 2000-2006 con UFB del 100% y del 2010-2015 con 100%; secuencias colombianas de 1998-2000 con 96% y del 2022 con 100%; secuencias peruanas de 2015 a 2017 con 56% de UFB).

Analizando más en detalle la estructura interna del árbol, las dos ramas que componen las secuencias de Chile (dos introducciones independientes, en el 2000 y en el 2010) están bien diferenciadas entre ellas (Figura 22). Además, todas las variantes virales, en ambas introducciones, parecen derivar de un mismo ancestro común, sugiriendo que sus mutaciones fueron resultado de un proceso evolutivo dentro del territorio y no de múltiples eventos de introducción de virus al país.

Por otro lado, en el caso de Perú (Figura 23), a pesar de que muchas secuencias corresponden al mismo año, se observa que no derivan de un único ancestro común cercano; se distribuyen a lo largo de la rama y hay una clara ruptura de la monofilia con secuencias estadounidenses. Aunque algunas relaciones carecen de un respaldo sólido, como evidencian valores de UFB de 65%, 57%, 49%, entre otros, existen casos en los que se puede apreciar un respaldo del 100% en la relación de secuencias peruanas con secuencias estadounidenses. Es importante resaltar que existe una secuencia peruana (MH791380) que, en nuestros análisis filogenéticos, muestra una notable distancia con respecto al resto de las muestras provenientes de Perú.

En relación a Colombia, se observa una marcada divergencia filogenética entre las muestras de dominio público de los años 1998 y 2000 en comparación con las muestras actuales recolectadas en 2022 (Figura 24). Al examinar el caso específico del brote de 2022, destaca que, a pesar de la similitud temporal y geográfica de las secuencias, se evidencia una divergencia filogenética considerable. En particular, una de las muestras exhibe una estrecha relación con una secuencia estadounidense (con un respaldo de UFB del 98 %), mientras que la otra se ubica en una rama hermana, agrupándose con otras secuencias estadounidenses diferentes.

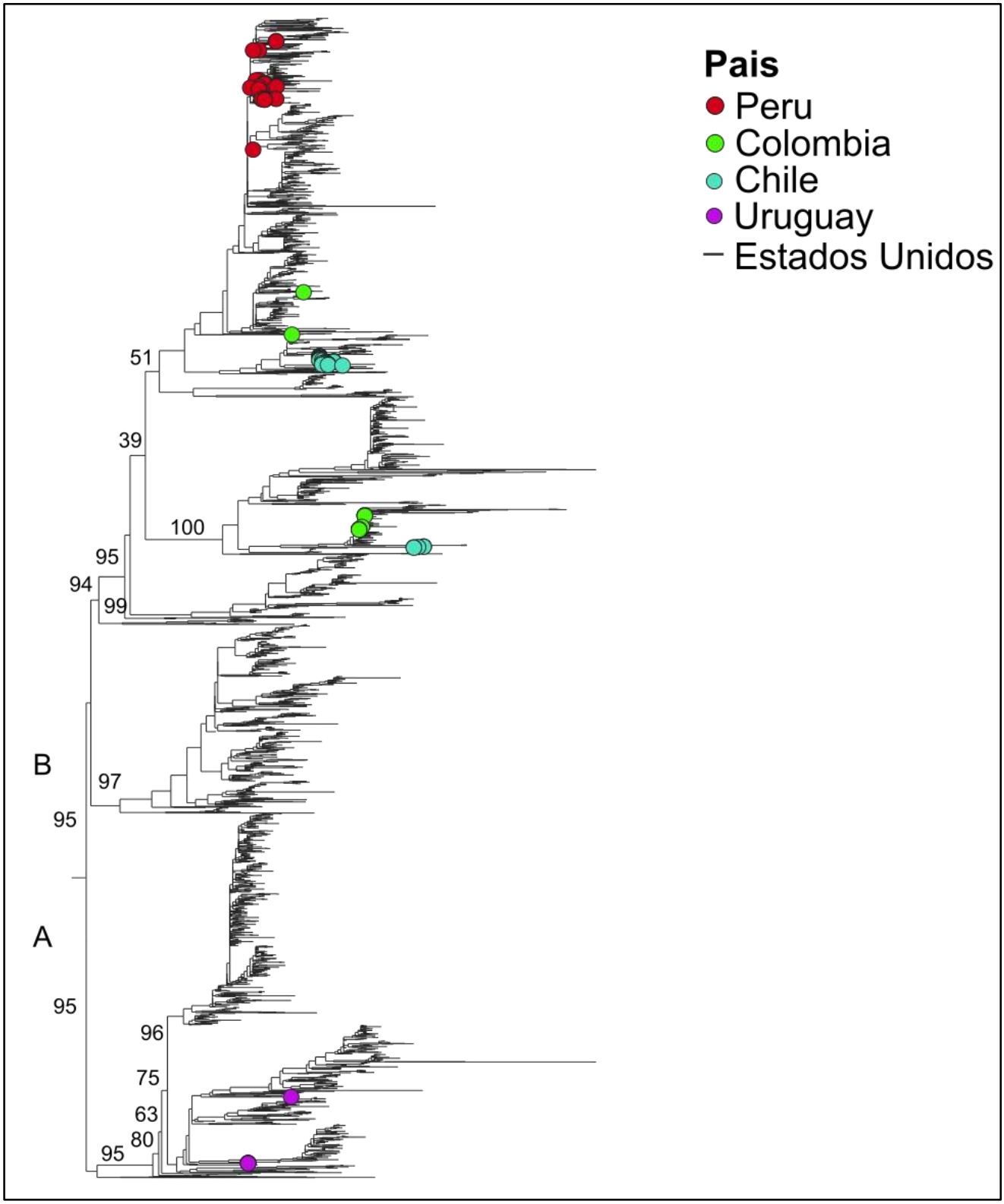

Figura 20. Árbol filogenético construido a través de la metodología de Máxima Verosimilitud mediante el programa IQTREE. Se pueden observar la formación de dos grandes ramas, identificadas en la figura como A y B. El país de origen de cada secuencia está indicado por los colores señalados en la leyenda.

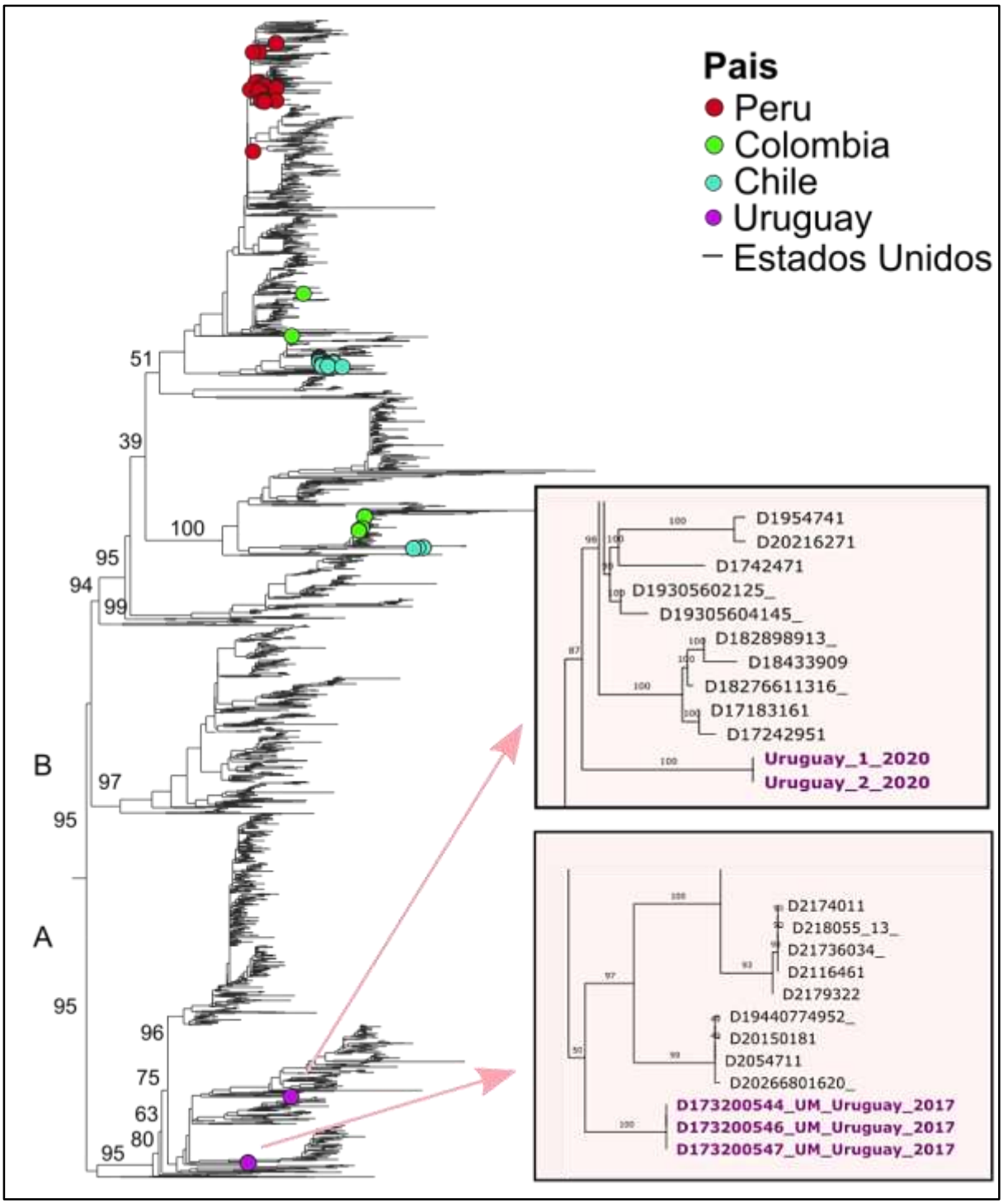

Figura 21. Árbol filogenético construido utilizando la metodología de Máxima Verosimilitud mediante el programa IQTREE. La figura se centra en las ramas que contienen las secuencias de Uruguay, destacadas en color violeta.

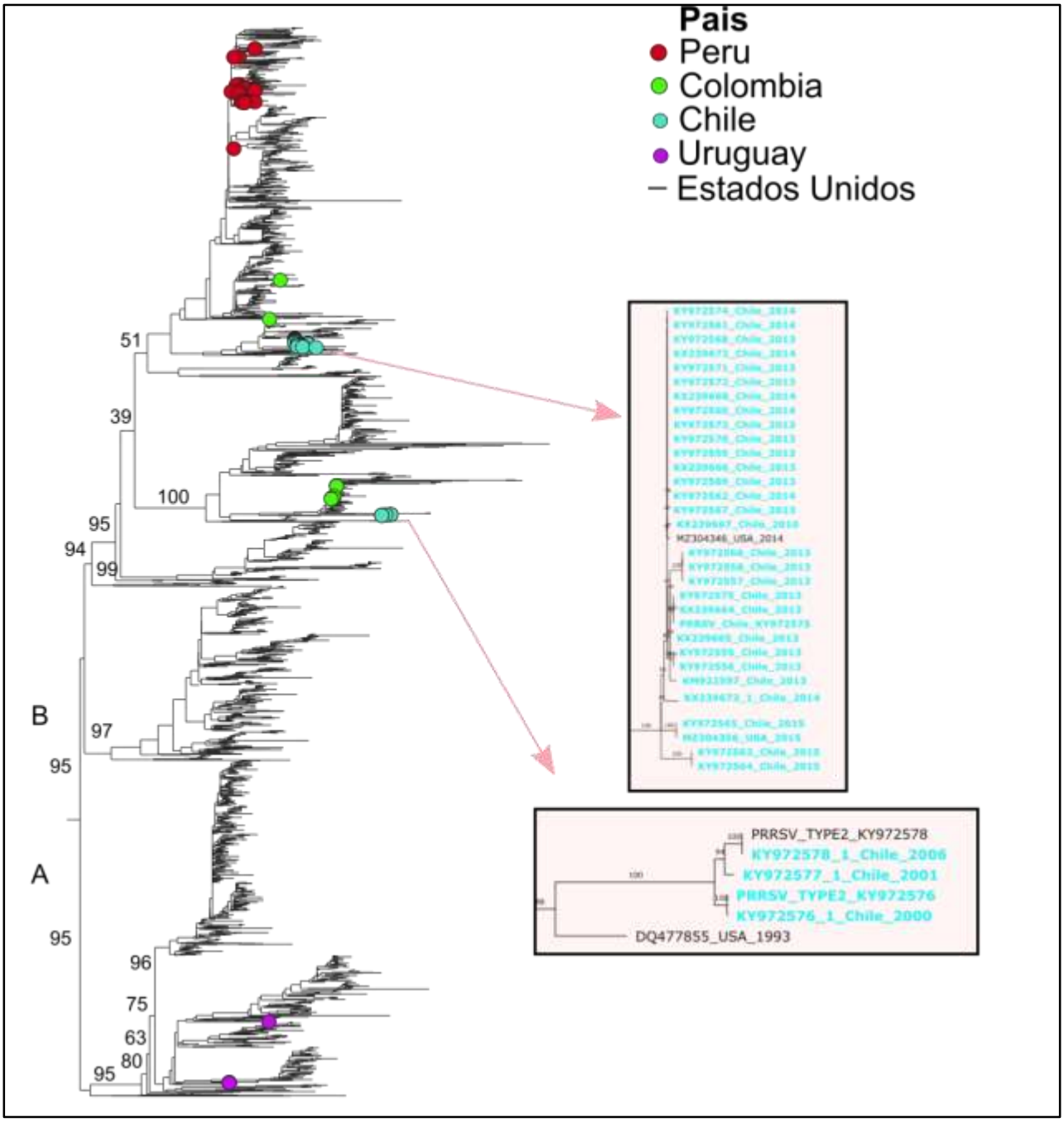

Figura 22. Árbol filogenético construido utilizando la metodología de Máxima Verosimilitud a través del programa IQTREE. La figura se enfoca en las ramas que contienen las secuencias de Chile, resaltadas en color azul.

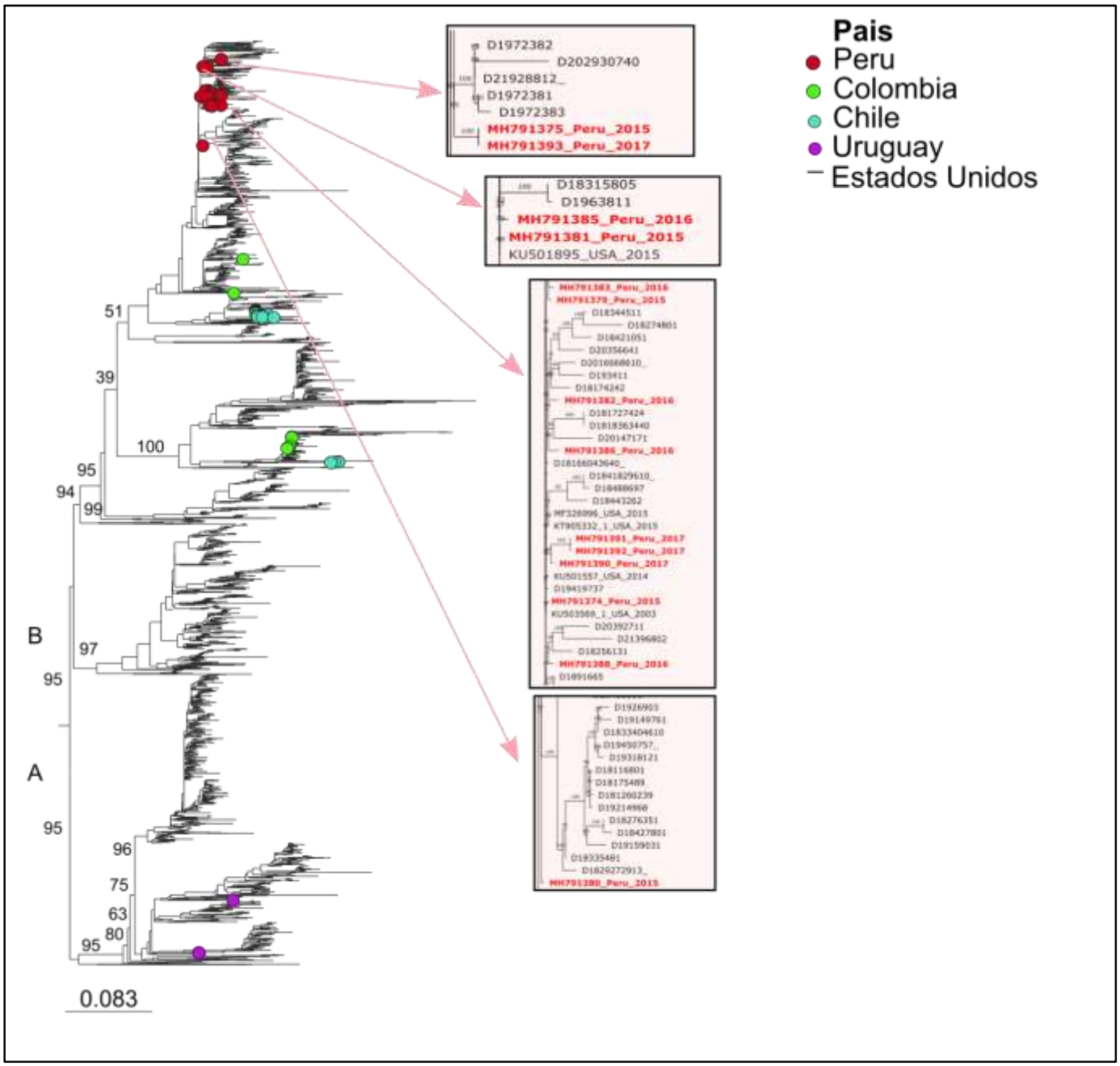

Figura 23. Árbol filogenético construido utilizando la metodología de Máxima Verosimilitud mediante el programa IQTREE. La figura se enfoca en las ramas que contienen las secuencias de Perú, resaltadas en color rojo.

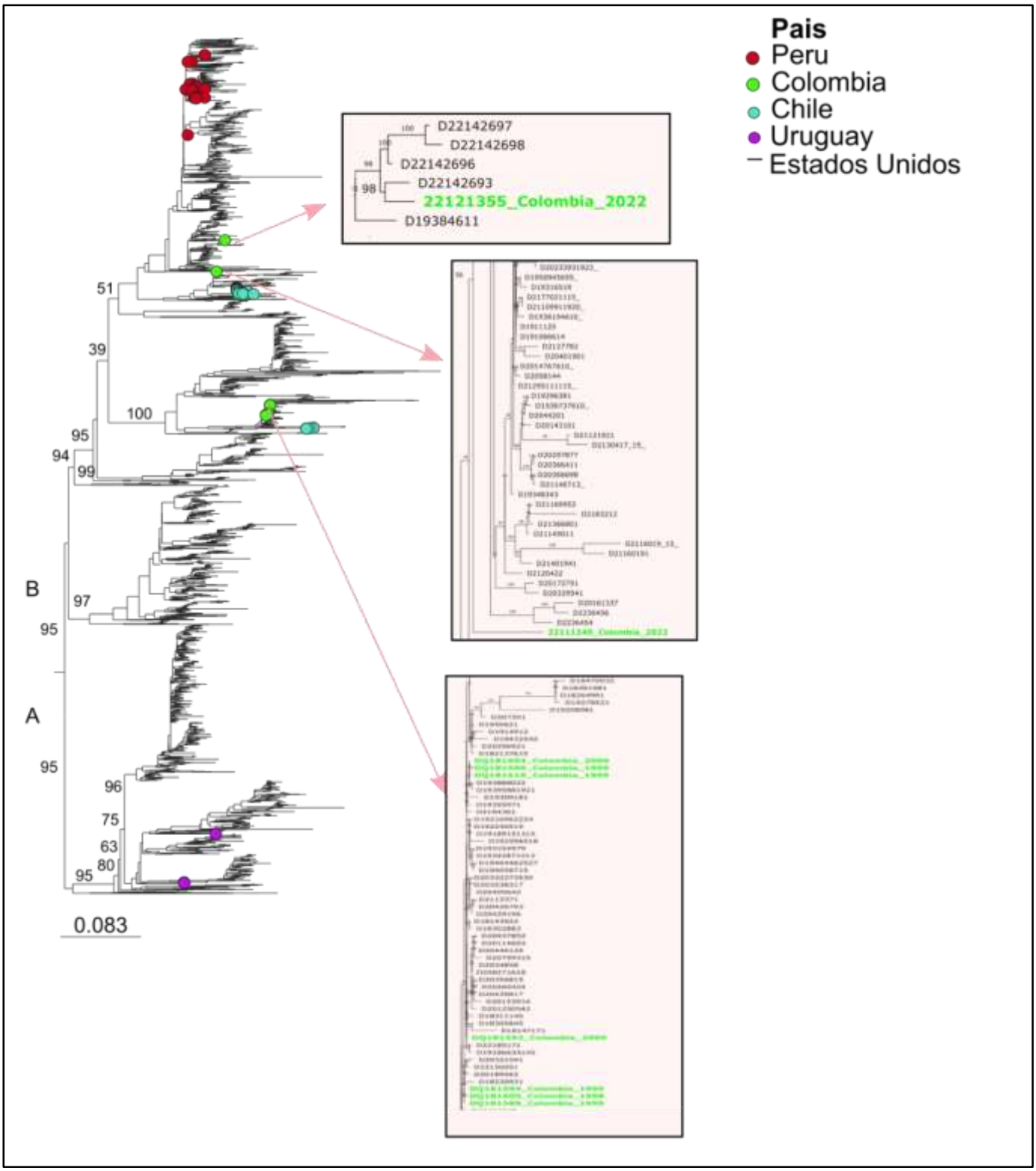

Figura 24. Árbol filogenético construido utilizando la metodología de Máxima Verosimilitud a través del programa IQTREE. La figura se enfoca en las ramas que contienen las secuencias de Colombia, resaltadas en color verde.

Reconstrucción del estado ancestral más probable

Tras el análisis del árbol filogenético resultante, se identificaron y marcaron las ramas terminales según el país de origen para inferir los estados ancestrales (Figura 25). El análisis de nuestras secuencias reveló, con una probabilidad muy alta que los Estados Unidos son el lugar de origen de los virus analizados en este trabajo. Además, se observó que cada país sudamericano experimentó introducciones independientes del virus en su territorio. En todos los casos, se evidenciaron indicios de más de una introducción del virus, siendo el origen de cada uno de ellos los Estados Unidos. Los resultados fueron consistentes en cada repetición de los análisis y confirmaron la persistencia de este patrón de distribución.

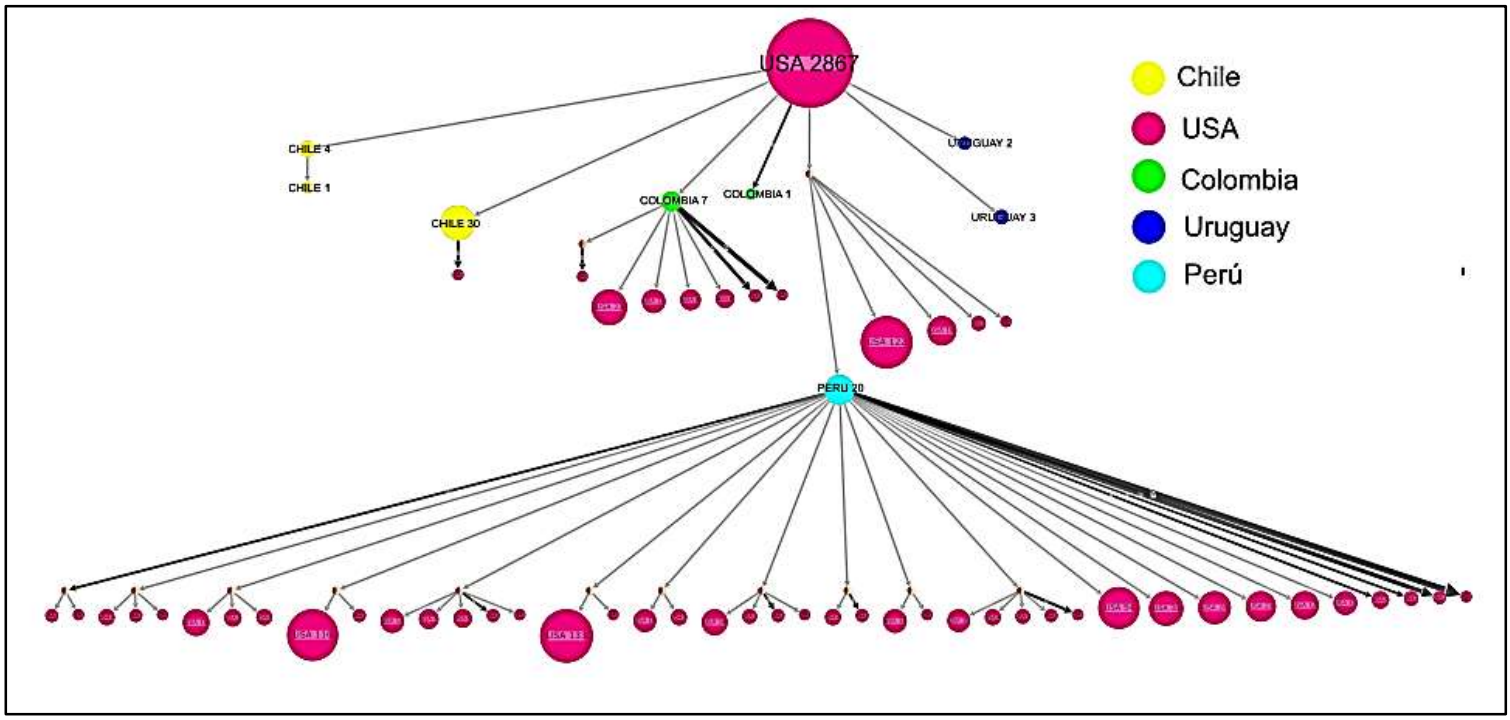

Figura 25. Reconstrucción del estado ancestral del *dataset* de secuencias enfocada a comprender la evolución del PRRSV2 en Sudamérica. Los círculos representan grupos genéticos de muestras y su color refleja su país de origen. El tamaño de cada círculo es proporcional al tamaño de la muestra. Las flechas que conectan dos círculos indican eventos de transmisión desde el grupo superior al grupo inferior.

Análisis evolutivos enfocados a conocer la historia y ancestria por ramas individuales

Con el fin de determinar la tasa de evolución y el ACMR de las ramas de interés en este estudio, se llevó a cabo un análisis en BEAST, proporcionando un promedio de la tasa de evolución general para cada rama examinada. Como resultado general, se identificó que las secuencias de Uruguay correspondientes al año 2017 exhibieron la tasa evolutiva más rápida; no obstante, los resultados del intervalo de probabilidad del 95% para esta rama se mantuvieron dentro del promedio. Chile y Colombia compartieron el ancestro común más antiguo, y las secuencias de Uruguay en 2020 presentaron un ancestro más antiguo en comparación con las del mismo país analizadas en 2017 (Tabla 4).

RAMA	TASA DE EVOLUCION sustituciones/sitio/año	HPD 95% sustituciones/sitio/año	AÑO DE ACMR	HPD 95%
Chile 2000-6	1.3x10-3	7.4x10-3 - 2.11×10 −2	1993	1993-1993
Chile 2010-5	1.6x10-3	1x10-7 – 3.3×10 −2	1976	1876-2010
Colombia 1998-2000	1.5x10-3	1.1x10-3 - 1.8×10 −3	1996	1995-1997
Colombia 2022	7.9 x10-3	5.9x10-3 – 1.1×10 −2	2017	2016-2017
Perú 2015-7	3.3 x10-3	2.9x10-3 - 4×10 −3	2002	2002-2003
Perú 2 MH791380	6.9 x10-3	5x10-3 – 7.5×10 −3	2014	2013-2014
Uruguay 2017	1.14 x10-2	8x10-3 – 1.4×10 −2	2016	2015-2016
Uruguay 2020	5.7x10-3	2x10-3 – 9×10 −3	2007	1998-2014

Tabla 4: Análisis evolutivo de las ramas individuales con secuencias de PRRSV provenientes de Sudamérica.

Discusión y conclusiones

La combinación de datos epidemiológicos con secuencias virales ha permitido obtener una visión integral de la dinámica de transmisión y evolución del PRRSV en Sudamérica. Es relevante señalar que, para llevar a cabo este estudio, se analizó exclusivamente la secuencia del gen ORF5 que, aunque es el más empleado en estudios filogenéticos y de epidemiología molecular (Mateu et al., 2003), representa tan solo el 3% del genoma viral. Por otro lado, la falta de disponibilidad general de secuencias del virus en los países sudamericanos también limita el alcance de sus conclusiones.

Nuestros hallazgos sugieren que la dispersión del virus PRRS en el continente sudamericano no se debe, principalmente, a procesos transfronterizos sino que indica eventos independientes de introducción en cada uno de los países estudiados, posiblemente relacionados con importaciones de animales o material genético desde los Estados Unidos. También pueden existir eslabones intermedios desconocidos no identificados, justamente por la falta de secuencias representativas del subcontinente. A pesar que la gran cantidad de las secuencias analizadas provienen de los Estados Unidos y esto puede crear un sesgo sobre el origen ancestral de los virus, la datación de los informes de enfermedades en los diversos países respalda esta hipótesis (Ramos et al., 2022) (Ramírez et al., 2019b) (Neira et al., 2017b).

Nuestros resultados también respaldan y amplían los hallazgos previos publicados por Neira y sus colaboradores en 2017, quienes afirmaron que la epidemia de 2010 en Chile fue causada por una variante de PRRSV genéticamente diferente de la detectada y eliminada del país en el 2006. En nuestro árbol filogenético, hemos identificado ambas introducciones, cada

una originada por una aparente variante viral que luego continuó su evolución en Chile siendo responsable de los brotes posteriores. Además, nuestros modelos refuerzan la teoría de que los virus PRRSV chilenos de 2010 comparten un ancestro en común muy cercano con un PRRSV de origen estadounidense. La utilización de secuencias no disponibles públicamente corrobora la teoría de doble introducción propuesta por Neira y colaboradores a partir de un número menor de secuencias. De igual forma la tasa de evolución que tuvo el virus obtenida en nuestro trabajo es congruente con la reportada ($1{,}0x10^{-2}$).

Dado que en Chile, las áreas destinadas a la producción porcina están geográficamente aisladas de las fronteras y los animales silvestres transmisores, como los jabalíes, no se encuentran en la región norte del país, se hace improbable una introducción desde Perú. Además, se sugiere una forma de transmisión diferente a la trasfronteriza y se descartan los animales silvestres como hospederos epidemiológicamente representativos en la zona (Neira et al., 2017). Nuestro árbol filogenético respalda estas observaciones al agrupar las secuencias de ambos países en ramas distantes (a pesar de su proximidad temporal) lo que hace altamente improbable, al menos con las secuencias reportadas hasta la fecha, un intercambio de virus entre los dos países.

Con respecto a la hipótesis planteada, referente a la transmisión del virus desde Estados Unidos a Perú, la misma está respalda desde una perspectiva epidemiológica, ya que en ese período Perú importó alrededor de 148 toneladas de animales vivos de raza pura con el objetivo de mejorar la genética de las producciones porcinas (Ramírez et al., 2019c). En este trabajo los investigadores realizaron un estudio que incorporó las secuencias peruanas en otro conjunto de datos y los resultados obtenidos también indicaron una posible descendencia desde Estados Unidos.

En cuanto a la tasa de evolución, ambas ramas peruanas analizadas concordaron con el promedio estimado en la literatura para las sustituciones por sitio por año que puede presentar el PRRSV. La fecha aproximada del ACMR para el grupo mayoritario es alrededor de 2002, lo cual coincide con los datos epidemiológicos, ya que el PRRSV ha sido diagnosticado e identificado en Perú desde 1998 (Alegría et al., 1998). En el análisis evolutivo para obtener el ACMR y la tasa de evolución, la secuencia peruana MH791380 fue evaluada junto con otras secuencias estadounidenses. Sin embargo, se subraya que cuenta con un respaldo de UFB relativamente bajo en el árbol filogenético (56%) y que los resultados derivados de este análisis deben interpretarse con precaución.

En relación a Uruguay, los resultados de las cinco secuencias analizadas muestran una clara divergencia dependiendo del año del aislamiento. La formación de estos grupos (secuencias del 2017 y del 2020), al no ser filogenéticamente cercanos y agrupar con otras muestras intermediarias de los Estados Unidos, sugiere un proceso de introducciones múltiples a este país. Este análisis respalda el trabajo previo de Ramos y colaboradores en 2022, en el cual se propone que el ancestro en común más reciente de las muestras uruguayas del 2020 proviene de los Estados Unidos. Otro dato interesante entre los virus que componen las dos ramas con secuencias uruguayas es que tienen una divergencia en la antigüedad del ACMR muy marcada. Para la rama que enmarca los brotes del año 2017, el ACMR data de solo un año antes, mientras que la de los brotes del año 2020 puede remontarse hasta 13 años en el tiempo. Este resultado refuerza la hipótesis de evolución independiente y eventos de introducción diferentes. Sin embargo, estos resultados deben tomarse con precaución ya que

son pocas las secuencias uruguayas analizadas en este estudio y es posible que falten virus intermediarios que darían mayor precisión al análisis.

Con respecto a Colombia, se evidenció una notable divergencia filogenética entre las secuencias de virus informadas entre 1998 y 2000 y las reportadas en 2022. Sin embargo, aún más relavante es la marcada divergencia genética identificada entre las secuencias obtenidas en 2022, a pesar de que ambas provienen del mismo departamento y están separadas temporalmente por apenas un mes.

En un informe citado de la Asociación Colombiana de Porcicultores en 2013 por (Castillo Espinoza and Ramírez Velásquez, 2021), se afirmó que la variabilidad del virus en Colombia había permanecido baja hasta ese momento. No obstante, en este trabajo fue posible identificar una nueva introducción de secuencias que coalescen hacia el año 2017, las cuales seguramente cambiaron el panorama e incrementaron la presencia de variantes virales en el país. Al igual que se observó en Uruguay, los ACMR estimados para los dos grupos de secuencias de Colombia son muy divergentes en el tiempo, reforzando la idea de múltiples introducciones.

Por último, a pesar de que nuestro análisis se basa solo en dos secuencias actuales, nos permite identificar que existe una marcada variabilidad de los virus circulantes en el departamento colombiano analizado. Estos resultados también concuerdan con los informes epidemiológicos proporcionados por los productores y veterinarios de la región, quienes han documentado olas de brotes de PRRSV que han afectado gravemente la producción porcina en Colombia y esto podría deberse a la introducción de diferentes variantes virales a la región. En este contexto, es crucial destacar la importancia de promover la realización de pruebas y el intercambio de datos de secuencias en Colombia. Esto no solo contribuirá a una comprensión más profunda de la dinámica de la enfermedad, sino que también permitirá el desarrollo de medidas efectivas de control para proteger la industria porcina de ese país.

Concluyendo, en todos los casos estudiados, las secuencias sudamericanas guardan una estrecha relación con las secuencias reportadas de los Estados Unidos e indican introducciones múltiples a cada uno de ellos desde el origen de los virus en Norte América o desde algún intermediario no muestreado. Además, en el análisis de estados ancestrales y para los casos de Perú y Colombia, se puede observar un intercambio bidireccional de material genético viral. Para validar esta hipótesis de transmisión, resulta esencial la inclusión en el futuro de datos epidemiológicos y análisis de riesgo que posibiliten la identificación de potenciales transmisores de virus entre estos países sudamericanos y los Estados Unidos. También será fundamental incorporar más secuencias, de aislamientos actuales e históricos, del gen ORF5 y también de genoma completo.

Conclusiones generales del capítulo 2

Los modelos de epidemiología molecular discretos, con un enfoque en virus de ARN, permitieron obtener una visión general de la transmisión y la dinámica evolutiva de los virus de IBDV y PRRSV. En general, se identificaron las rutas de transmisión a nivel continental e intercontinental y determinaron la fecha aproximada de introducción de nuevos genotipos en lugares específicos. También se identificaron y documentaron nuevas variantes circulantes en algunos países particulares.

Este capítulo de la tesis aportó respuestas sobre cómo impacta la incorporaron de nuevas secuencias en diferentes conjuntos de datos utilizados en estudios previos y como la vigilancia epidemiológica molecular puede revelar una inesperada relación en cuanto al flujo de virus entre países.

Este trabajo además confirmó o refutó hipótesis previamente planteadas sobre la dinámica viral de los virus estudiados. Asimismo, resaltó la importancia de complementar los análisis con la mayor cantidad posible de datos y secuencias informativas. Por otro lado, también resalta la necesidad de realizar curados previos sobre la calidad de los datos, evitando así la inclusión de información redundante o incompleta que pudiera distorsionar los resultados finales y, por ende, las conclusiones.

Finalmente, en este capítulo de la tesis se dejó en evidencia la necesidad constante de desafiar los resultados obtenidos por modelos de epidemiología molecular con los datos epidemiológicos como una herramienta esencial para llegar a conclusiones más cercanas a la realidad. Además, desde una perspectiva sanitaria y productiva se enfatizó la importancia de mantener una vigilancia epidemiológica molecular en el campo para comprender cómo están evolucionando las enfermedades virales en las producciones animales. Finalmente, se recalcó la importancia de abordar las epidemias con medidas efectivas de vigilancia y control para contribuir directamente a la salud y la sostenibilidad de la industria pecuaria.

CAPITULO 3

MODELOS DISCRETOS A UNA ESCALA NACIONAL

CON UN ENFOQUE EN VIRUS DE ADN

Dinámica de los virus ORFV

CAPITULO 3 - ORFV

Materiales y métodos

Selección de muestras

Las muestras utilizadas en este trabajo se obtuvieron a través de una colaboración con una red de veterinarios distribuidos en todo el país. Estos veterinarios voluntariamente proporcionaron costras de ovinos y caprinos sospechosos de estar infectados con Ectima Contagioso. Se realizó una selección en la que se eligieron 34 muestras de diferentes provincias del país (tabla 5). Estas muestras se seleccionaron siguiendo dos criterios fundamentales: en primer lugar, debían contener suficiente cantidad de material para permitir la secuenciación de los cuatro genes de interés (orf011, orf020, orf109 y orf127); en segundo lugar, fue importante que las muestras estuvieran acompañadas de información epidemiológica trazable, que incluyera al menos la provincia de origen, la especie del paciente del cual se obtuvo la muestra y el año de obtención. Los médicos veterinarios recolectaron las muestras usando bisturí y pinzas. Posteriormente, las mismas se mantuvieron refrigeradas en tubos Falcón hasta su envío al laboratorio, donde se almacenaron a -80ºC hasta el momento de su procesamiento.

MUESTRA	PROVINCIA	HOSPEDERO	AÑO
sCh97	Buenos_Aires	Ovino	1997
gYa11	Jujuy	Caprino	2011
sPi13	Rio_Negro	Ovino	2013
sAz13	Buenos_Aires	Ovino	2013
gCom14	Rio_Negro	Caprino	2014
gLag14	Salta	Caprino	2014
gNeu14	Neuquen	Caprino	2014
sPil14	Rio_Negro	Ovino	2014
sChu14	Chubut	Ovino	2014
sSju14	Santa_Cruz	Ovino	2014
gCnc15	San_Luis	Caprino	2015
gEma15	Salta	Caprino	2015
sBal15	Buenos_Aires	Ovino	2015
sAca15	Cordoba	Ovino	2015
sHre15	Cordoba	Ovino	2015
sPua15	Buenos_Aires	Ovino	2015
sMbu15	Buenos_Aires	Ovino	2015
gCrl16	San_Luis	Caprino	2016
sChb16	Buenos_Aires	Ovino	2016
sAng16	La_Pampa	Ovino	2016
gMor17	Salta	Caprino	2017
sOve17	Salta	Ovino	2017
gCha17	La_Rioja	Caprino	2017

gEqn18	La_Rioja	Caprino	2018
gTum18	Jujuy	Caprino	2018
gMor18	Salta	Caprino	2018
gTuc18	Tucuman	Caprino	2018
gVol18	Jujuy	Caprino	2018
gApe18	Chubut	Caprino	2018
sCor18	Cordoba	Ovino	2018
gVol18	Jujuy	Caprino	2018
sBav18	Buenos_Aires	Ovino	2018
sInd19	Buenos_Aires	Ovino	2019
sEsq20	Chubut	Ovino	2020
sMer21	San_Luis	Ovino	2021

Tabla 5: Información epidemiológica de las muestras de ORFV utilizadas en el presente análisis.

Extracción de ADN

El ADN total se extrajo de las costras obtenidas utilizando el kit QIAamp DNA Mini (QIAGEN), siguiendo las indicaciones del fabricante. En primer lugar, las costras se disgregaron en nitrógeno líquido, y el polvo resultante se procesó en un tubo eppendorf de 1,5 ml junto con solución tamponada de lisis y proteinasa K. Posteriormente, se incubó a 56°C hasta que los tejidos se lisaron por completo. Luego, se extrajo el ADN según las instrucciones del fabricante. El ADN se resuspendió en 0,2 ml de *buffer* de elución y se almacenó a -20°C hasta su posterior procesamiento.

Diagnóstico por PCR

La confirmación del agente etiológico ORFV se llevó a cabo mediante una PCR diagnóstica sugerida en el trabajo de (Kottaridi et al., 2006). Esta PCR utiliza los *primers* 045F y 045R para amplificar un fragmento de 392 pares de bases del gen orf045, el cual se encuentra altamente conservado entre las diferentes cepas de ORFV. La PCR se realizó utilizando la enzima Taq DNA Polymerase de Promega (Madison, WI, EE. UU.) con cebadores específicos diseñados para la región del gen orf045. Las secuencias de estos cebadores fueron las siguientes: Fw: 5' CCTACTTCTCGGAGTTCAGC 3' y Rv: 5' GCAGCACTTCTCCTCGTAG 3'. El ciclo de amplificación consistió en una fase de desnaturalización de 5 minutos a 95°C, seguida por 40 ciclos con fases de desnaturalización de 10 segundos a 95°C, apareamiento de cebadores de 10 segundos a 47°C y extensión de 10 segundos a 72°C. Finalmente, se realizó una fase adicional de extensión de 10 minutos a 72°C. Los productos de amplificación resultantes se visualizaron en un gel de agarosa al 1,5% teñido con Bromuro de Etidio, buscando el producto específico de 392 pares de bases.

Amplificación por PCR de los genes de interés

Se seleccionaron los genes orf011, orf020, orf109 y orf127 debido a que el gen orf011 codifica para una proteína que induce una marcada respuesta de anticuerpos, el gen orf109 codifica para una proteína de superficie del virus, que estaría involucrada en la inducción de anticuerpos neutralizantes y es una secuencia variable y los genes orf020 y orf127 están involucrados en la modulación de la respuesta inmune del hospedador (Li et al 2023).

Además, estos genes están distribuidos a lo largo del genoma, lo que podría ofrecer una visión general de la dinámica y evolución del virus.

La amplificación de estos genes se llevó a cabo mediante una PCR, utilizando la enzima Taq DNA Polymerase de Promega (Madison, WI, EE. UU.), junto con cebadores y condiciones específicas diseñadas para cada región de amplificación (Tablas 5 y 6). Cada reacción se ajustó a un volumen total de 25 µL, que incluyó 5 µL de Buffer 5X GoTaq, 0,5 µL de dNTPs de 10 mM, 0,2 µL de ADN polimerasa GoTaq (Promega), 100 ng de ADN y 25 pmol de los cebadores específicos.

Gen	***Primers***	**Tamaño del producto**
orf011	PP1: 5' GTCGTCCACGATGAGCAGCT 3' PP4: 5' TACGTGGGAAGCGCCTCGCT 3'	594pb
orf020	VIR1: 5' ACAATGGCCTGCGAGTG 3' VIR2: 5' TTAGAAGCTGATGCCGCAG 3'	552pb
orf109	109F: 5' TATAGAAATTAGAGCCTCCTA 3' 109R: 5' CAGCAACCCATTTATTAGA 3'	550pb
orf127	127F: 5' ATGTCGAAGAACAAAATTCT 3' 127R: 5' TTTAGTAGTCATGTATGATT 3'	560pb

Tabla 6. *Primers* utilizados para la amplificación parcial de los genes de ORFV.

Gen	**Condiciones de PCR**
orf011	Desnaturalización: 94ºC por 30 seg, Anneling: 62ºC por 30 seg, Extensión: 72ºC por 30 seg. X 35 ciclos
orf020	Desnaturalización: 94ºC por 60 seg, Anneling: 54ºC por 30 seg, Extensión: 72ºC por 30 seg. X 35 ciclos
orf109	Desnaturalización: 94ºC por 30 seg, Anneling: 52ºC por 30 seg, Extensión: 72ºC por 30 seg. X 35 ciclos
orf127	Desnaturalización: 94ºC por 30 seg, Anneling: 50ºC por 30 seg, Extensión: 72ºC por 30 seg. X 35 ciclos

Tabla 7. Condiciones de ciclado para cada una de las PCR.

Visualización en geles de agarosa

Los productos resultantes de la PCR se combinaron con un buffer de siembra de concentración 10X, compuesto por un 50% de glicerol, TAE 10X y un 1% de Azul de Bromofenol. Posteriormente las muestras, se sembraron en geles de agarosa junto con marcadores de peso molecular, seleccionados en función del tamaño de los productos. En general, se utilizó el marcador de 100 bp DNA Ladder® de Invitrogen. Con el propósito de

visualizar las bandas de ADN, se emplearon geles de agarosa con concentraciones que variaron desde 1% hasta 1,5%, suspendidos en un buffer TAE 1X mezclado con el agente intercalante Bromuro de Etidio a una concentración de 0,5 µg/ml. Los geles fueron sometidos a una corriente de 90V durante 50 minutos. Finalmente, para la captura de las imágenes de los geles, se utilizó el equipo Molecular Imager® Gel Doc™ XR+ de Bio-Rad, con el programa Image Lab™ de la misma marca.

Purificación de productos de PCR

Las PCR positivas correspondientes a cada gen del virus ORFV se sometieron a un proceso de purificación mediante el kit de purificación de PCR QIAquick de Qiagen.

Secuenciación

Los fragmentos de PCR se secuenciaron en el Departamento de UGB del INTA, utilizando la metodología de Sanger (Sanger, et al. 1977). Se emplearon terminadores fluorescentes (kit Big Dye Terminator v3.1, Applied Biosystems) en secuenciadores automáticos de capilares Genetic Analyzer 3500xl (Applied Biosystems) la secuenciación se llevó a cabo en ambas orientaciones utilizando los *primers* de amplificación anteriormente descritos.

Curado de secuencias

Los productos de secuencia obtenidos fueron chequeados utilizando el programa BioEdit. (Hall et al. 2011). Cada archivo de secuencia fue examinado visualmente y editado para resolver cualquier ambigüedad. Se realizaron recortes en los extremos de los fragmentos donde la calidad de la lectura fue baja.

Alineación de secuencias

Una vez se obtuvo las secuencias consenso de las muestras, se procedió a alinearlas utilizando el paquete de software MAFFT versión 7 (Katoh y Standley, 2014) y posteriormente se realizaron ajustes manuales en la alineación. Este proceso se llevó a cabo de manera independiente para cada uno de los genes de estudio y además se creó un concatenado de los genes orf011, orf020 y orf127.

Análisis filogenéticos, filogeográficos y evolutivos

Se llevaron a cabo análisis filogenéticos y filodinámicos bayesianos utilizando el programa BEAST versión 1.10.4 (Suchard et al., 2018) que permitió inferir el Árbol de Máxima Credibilidad, la tasa evolutiva media, la fecha del Ancestro Común Más Reciente y la reconstrucción filogeográfica viral en un espacio discreto para cada gen y el concatenado de los mismos. Se utilizó el modelo coalescente Bayesian Skyride junto con un modelo de reloj molecular Relajado Lognormal No Correlacionado y un modelo evolutivo GTR. Las fechas, el hospedador y las ubicaciones de muestreo se especificaron como *traits*. Se ejecutaron dos cadenas independientes de MCMC para cada análisis, cada una con 50 millones de generaciones y para verificar la convergencia fue utilizado el programa Tracer v.1.7 (Rambaut et al., 2014). El árbol MCC se visualizó con la herramienta Figtree. Por último para analizar y visualizar las reconstrucciones filogeográficas fue utilizado el programa SpreaD4 (Kanika D. Nahata et al., 2022).

RESULTADOS

Estadísticas de secuenciación y alineamiento

Gen orf011

Las 34 secuencias analizadas presentaron una longitud promedio de 558 pares de bases. Se identificó un patrón de alineamiento uniforme en todas las muestras.

Gen orf020

Las 34 secuencias analizadas tuvieron una longitud promedio de 517 pares de bases. Se destacó un patrón de alineamiento uniforme en todas las muestras.

Gen orf109

Las 34 secuencias analizadas tuvieron una longitud promedio de 487 pares de bases. Se observó un patrón de alineamiento notablemente divergente entre todas las muestras, con múltiples inserciones y deleciones en diferentes sitios del alineamiento.

Gen orf127

Las 34 secuencias analizadas presentaron una longitud promedio de 505 pares de bases. Se identificó un patrón de alineamiento uniforme en todas las muestras.

Concatenado de Genes

Dado que el gen 109 exhibió una considerable hipervariabilidad en todas las secuencias analizadas, se optó por realizar el concatenado únicamente con los genes 011, 020 y 127 de las 34 muestras a fin de evitar cualquier interferencia en el alineamiento que pudiera distorsionar los resultados finales. El concatenado resultante de estos genes tuvo una longitud aproximada de 1578 pares de bases y mostró un patrón de alineamiento uniforme.

Análisis filogenéticos

Gen orf011

El análisis llego a convergencia en cada uno de los parámetros testeados. Se calculó el árbol de MCC, el cual evidenció la formación de dos clúster principales (Figuras 26 y 27). La rama A, con una PP de 0,14, está compuesta por 13 secuencias caprinas y 10 ovinas, con 9 de estas secuencias provenientes del norte, 8 del centro y 6 del sur. El clúster B resultó bien soportado con una PP de 1, compuesto por secuencias de 9 ovejas y 2 de cabras de las cuales 6 pertenecen a la zona sur, 4 a la zona centro y 1 a la zona norte. Algunas subramas están bien soportadas, como las que agrupan las secuencias sPil14, gNeu14 y sChu14 con PP de 0.99 y la que agrupa las secuencias gEma15, gEnq18, gMor17 y gMor18 con PP de 1.

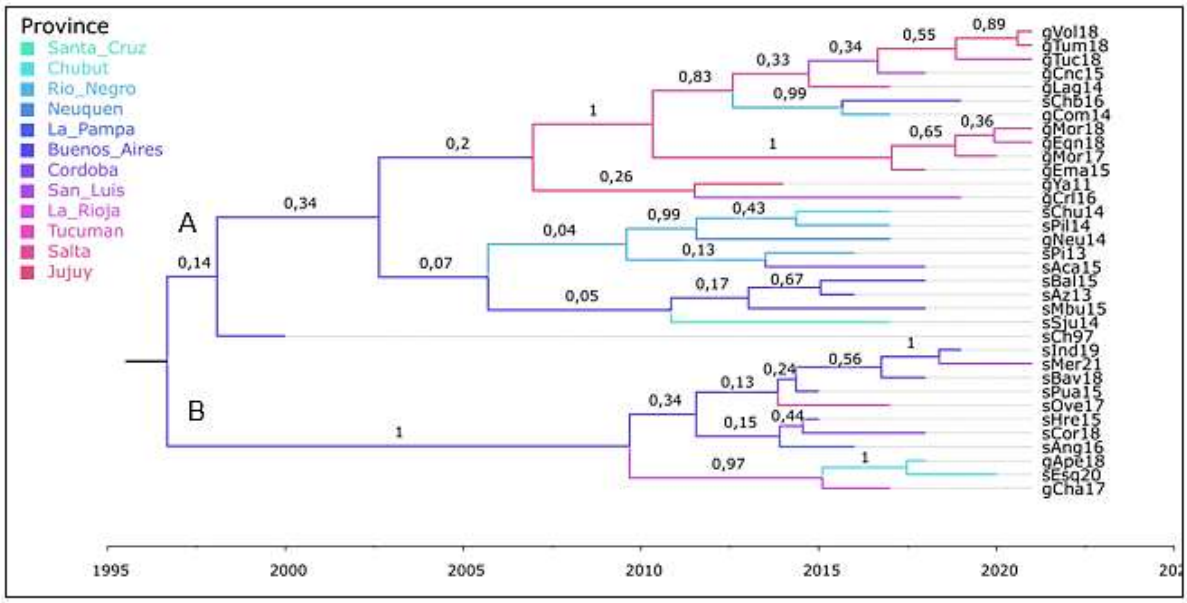

Figura 26. Árbol de MCC del gen 011, estimado a partir del programa BEAST. Las líneas de las ramas están coloreadas de acuerdo a la provincia en la cual fue tomada la muestra, los números sobre las ramas indican la probabilidad posterior estimada, en el margen inferior se indican los años de presentación del foco.

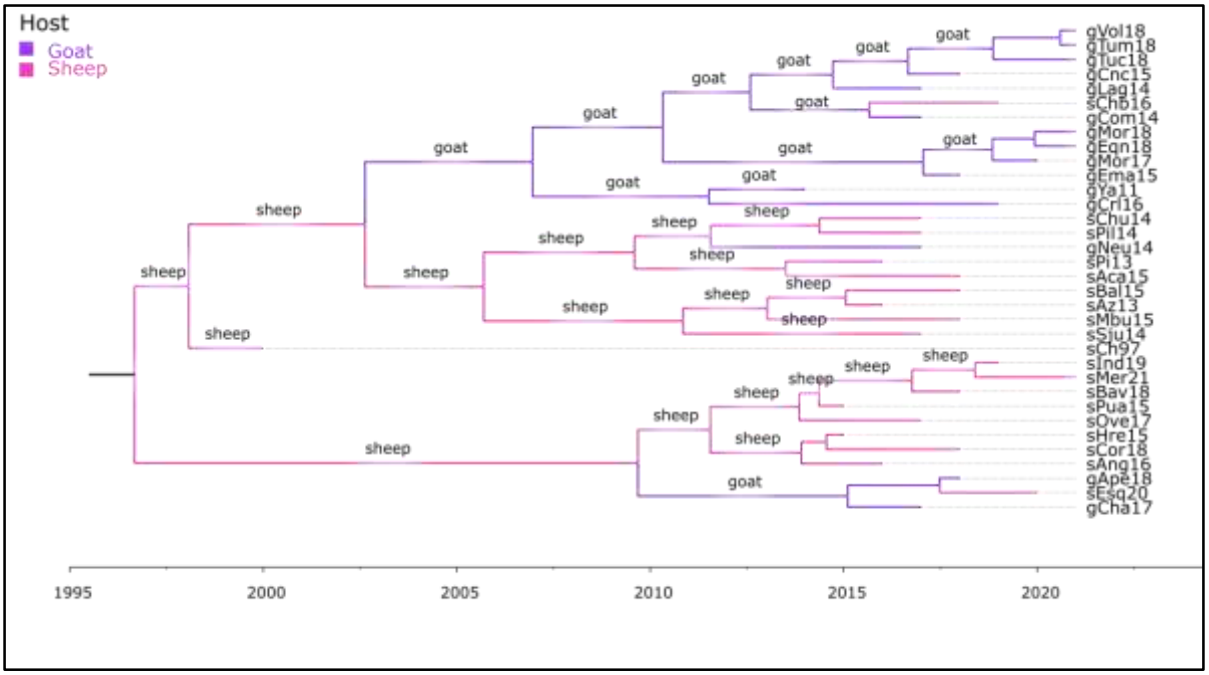

Figura 27. Árbol de MCC del gen 011, estimado a partir del programa BEAST. Las líneas de las ramas están coloreadas de acuerdo al hospedero del cual fue tomada la muestra y en el margen inferior se indican los años de presentación del foco.

Gen orf020

El análisis alcanzó la convergencia en todos los parámetros evaluados. Se generó el árbol de MCC, el cual reveló la formación de dos clúster principales. El clúster A, con una PP, de 0,49 comprende 20 secuencias, con 18 provenientes de ovinos y dos de caprinos. De estas, una se origina en la zona norte del país, 13 en el centro y 6 en el sur. Por otro lado, la rama B,

con una PP de 0,58, incluye 14 secuencias, siendo 13 de caprinos y una de ovino. Entre estas secuencias, 5 provienen del centro del país, 2 del sur y 7 del norte. En este análisis, las secuencias sPil14, gNeu14 y sChu14 continúan agrupadas con un sólido respaldo de PP de 0,99. Otra agrupación con un respaldo sólido de PP (0,98) la conforman las muestras gEma15, gLac14, gMor17 y gYa11.

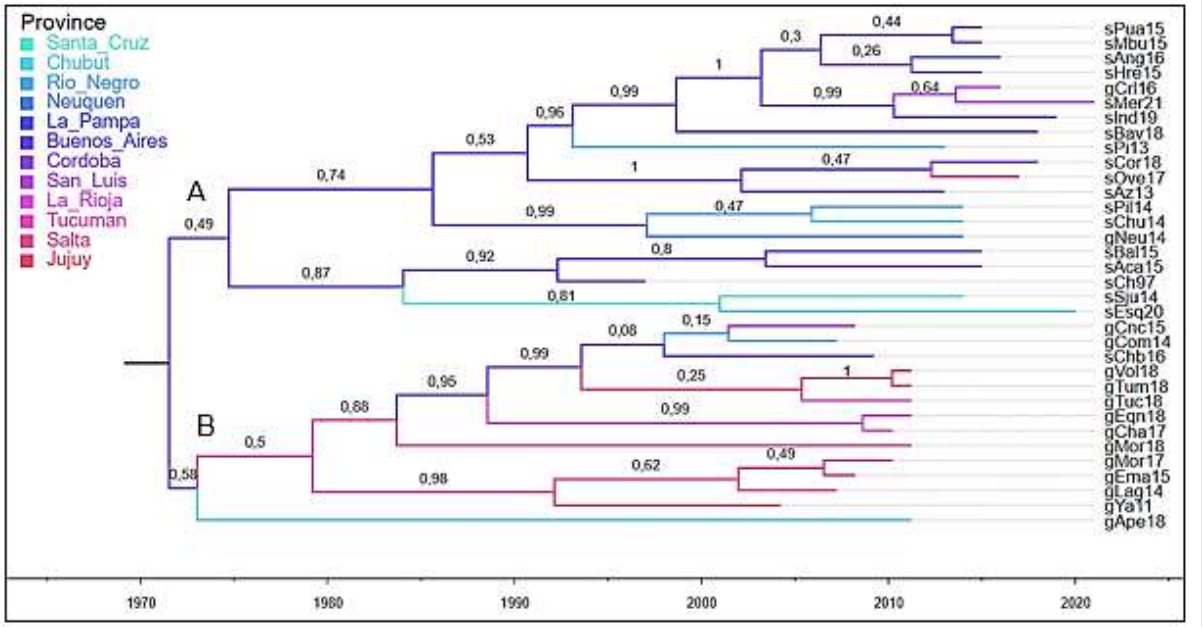

Figura 28. Árbol de MCC del gen 020, estimado a partir del programa BEAST. Las líneas de las ramas están coloreadas de acuerdo a la provincia en la cual fue tomada la muestra, los números sobre las ramas indican la probabilidad posterior estimada y en el margen inferior se indican los años de presentación del foco.

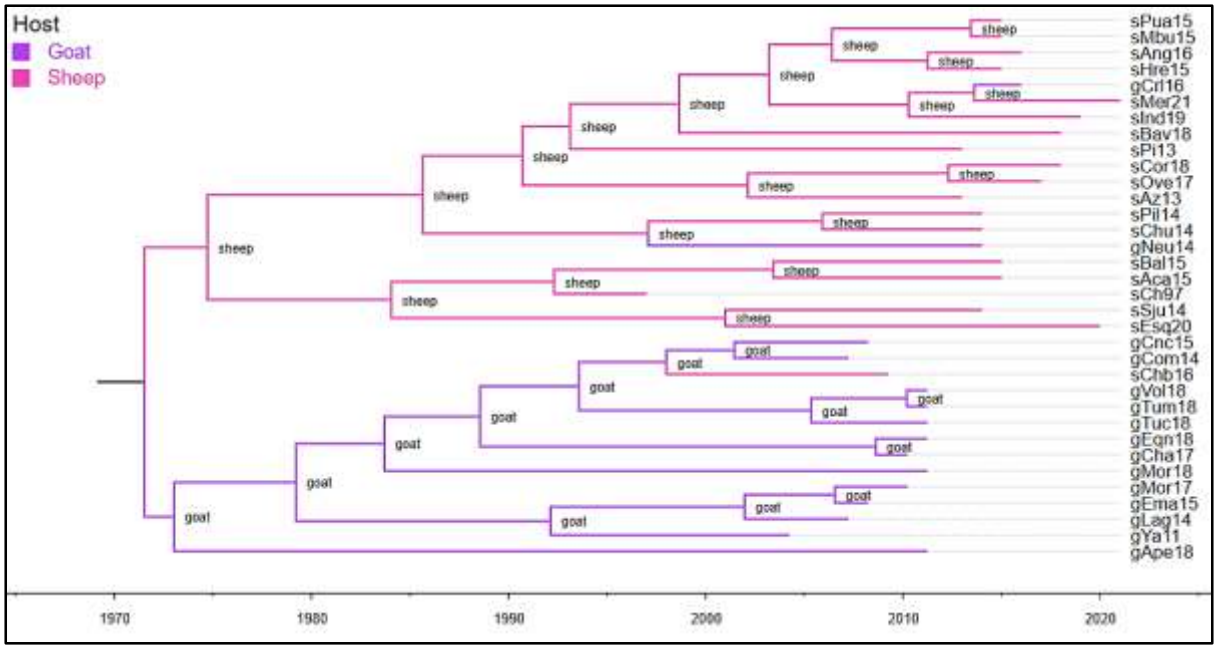

Figura 29. Árbol de MCC del gen 020, estimado a partir del programa BEAST. Las líneas de las ramas están coloreadas de acuerdo al hospedero del cual fue tomada la muestra. En el margen inferior se muestra una escala de tiempo para referencia de los años de presentación del foco.

Gen orf109

El análisis alcanzó la convergencia en todos los parámetros evaluados. Se generó el árbol de MCC, el cual reveló la formación de dos clúster principales. El grupo A, con una PP de 0,44, se compone de secuencias de 13 ovejas y 8 cabras, distribuidas de la siguiente manera: 3 de la zona sur, 13 de la zona centro y 5 de la zona norte. Por otro lado, la rama B, con una PP de 0,86, está compuesta por 6 secuencias de ovinos y 7 de cabra, de las cuales 5 proceden del sur, 4 del centro y 4 del norte. En este árbol, las secuencias sPil14, gNeu14 y sChu14 continúan agrupándose con una PP de 1, las secuencias gYa11, gEnq18, gMor17 y gMor18 también mantienen su agrupamiento con una PP de 1.

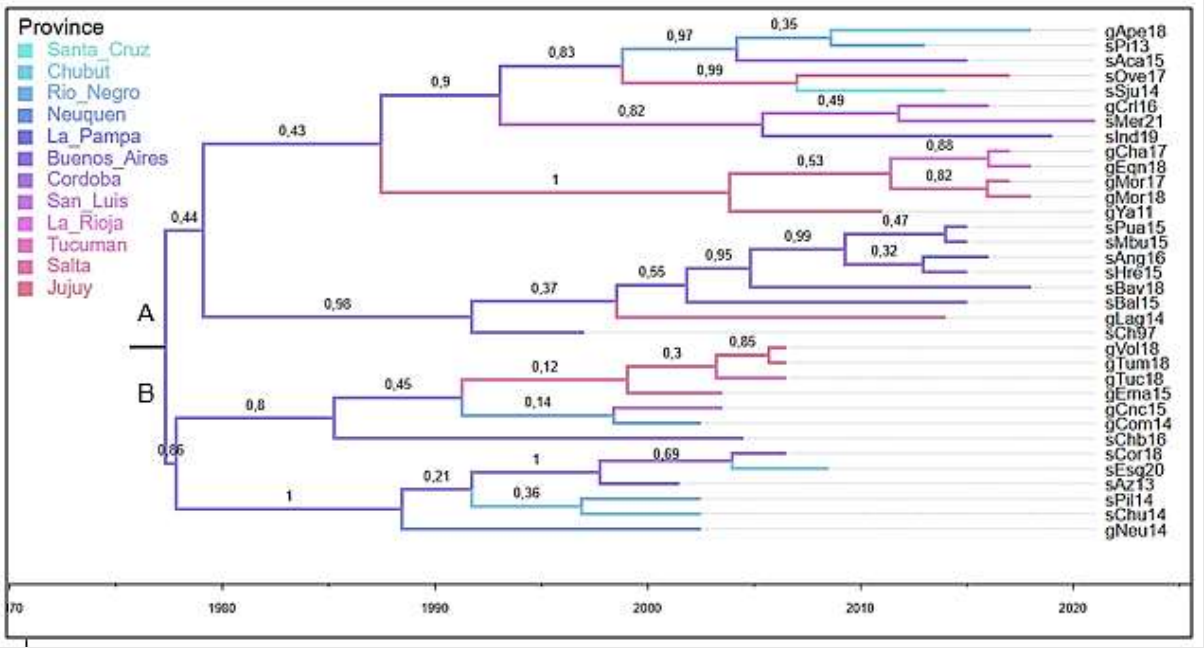

Figura 30. Árbol de MCC del gen 109, estimado a partir del programa BEAST. Las líneas de las ramas están coloreadas de acuerdo a la provincia en la cual fue tomada la muestra, los números sobre las ramas indican la probabilidad posterior estimada y en el margen inferior se muestra una escala de tiempo para referencia de los años de presentación del foco

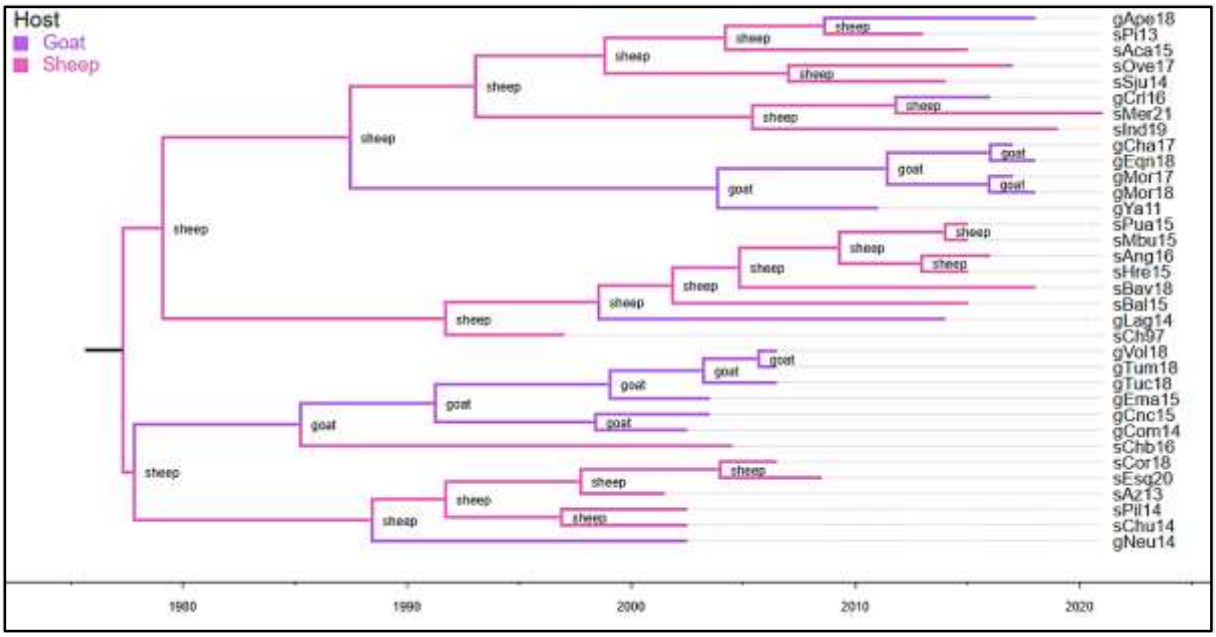

Figura 31. Árbol de MCC del gen 109, estimado a partir del programa BEAST. Las líneas de las ramas están coloreadas de acuerdo hospedero del cual fue tomada la muestra y en el margen inferior se muestra una escala de tiempo para referencia de los años de presentación del foco.

Gen orf127

El análisis alcanzó la convergencia en todos los parámetros evaluados. Se generó el árbol de MCC, el cual reveló la formación de dos clúster principales (Figura 32). La rama A, con una PP de 0,25, estaba compuesta por 23 secuencias, de las cuales 13 fueron aisladas de caprinos y 10 de ovinos (Figura 33). En cuanto a su procedencia geográfica, 6 provenían del sur, 7 del centro y 10 del norte. Por otro lado, la rama B, con una PP de 0,85, se formó con 11 secuencias, de las cuales 9 fueron aisladas de ovinos y 2 de caprinos. En términos de distribución geográfica, 2 secuencias tenían su origen en el sur, 8 en el centro y 1 en el norte. Cabe destacar la formación de algunas subramas bien respaldadas, como la que abarca las secuencias de sPil14, gNeu14 y sChu14, donde se repite el patrón de agrupamiento observado en los árboles anteriores, con una PP de 1 y la de las secuencias gYa11, gEnq18 y gMor17, con PP de 1.

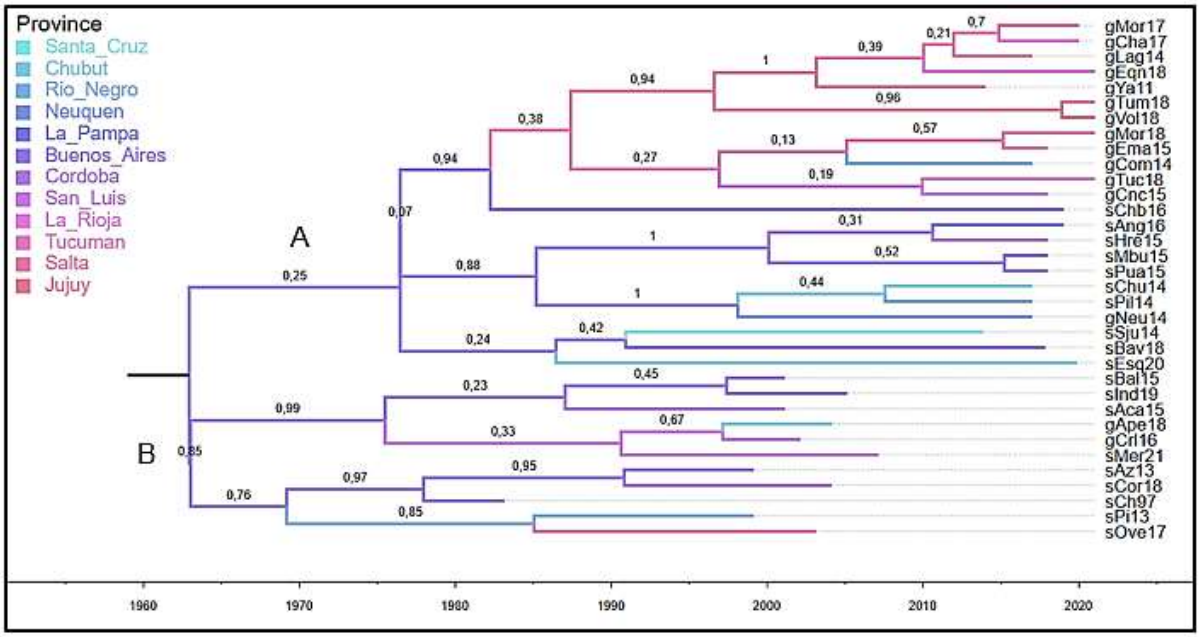

Figura 32. Árbol de MCC del gen 127, estimado a partir del programa BEAST. Las líneas de las ramas están coloreadas de acuerdo a la provincia en la cual fue tomada la muestra, los números sobre las ramas indican la probabilidad posterior estimada y en el margen inferior se muestra una escala de tiempo para referencia de los años de presentación del foco.

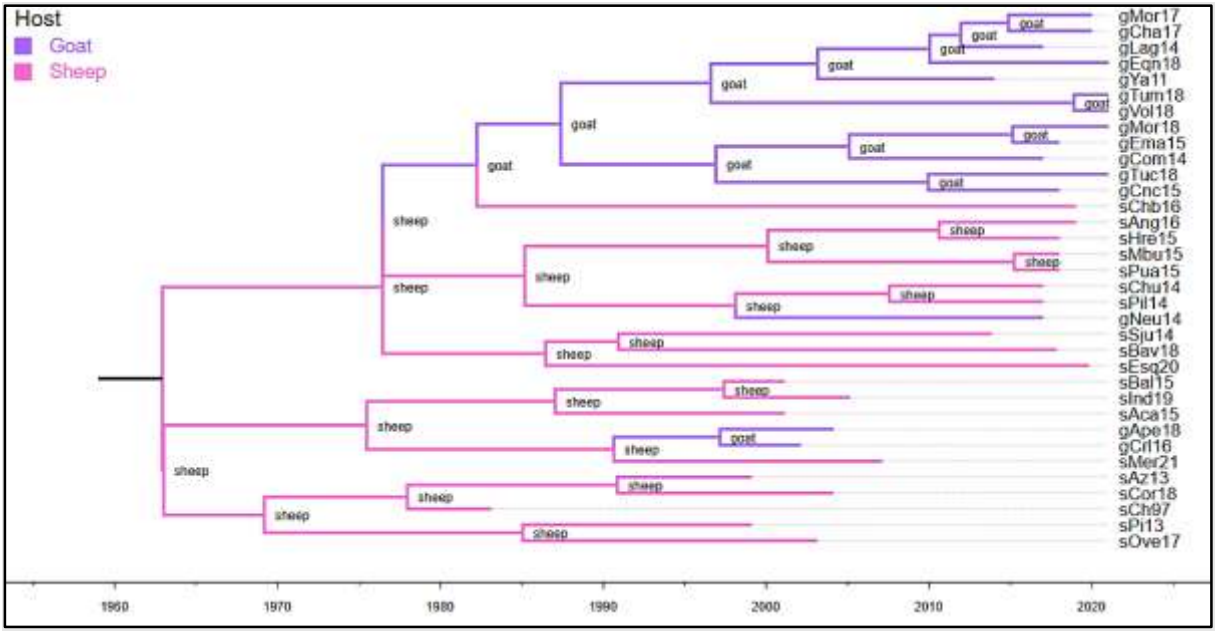

Figura 33. Árbol de MCC del gen 127, estimado a partir del programa BEAST. Las líneas de las ramas están coloreadas de acuerdo hospedero del cual fue tomada la muestra y en el margen inferior se indican los años de presentación del foco.

Concatenado de genes

El análisis alcanzó la convergencia en todos los parámetros evaluados. Se generó el árbol de MCC, el cual reveló la formación de dos clúster principales (Figura 34). Sin embargo, ninguno de los dos tiene buen soporte de PP (Rama A 0,29 y Rama B 0,77), Resulta notable que no

hubo agrupamiento por especie, ni por región. La rama A estuvo compuesta por 9 secuencias de ovejas y 12 de cabra, distribuidas en las tres diferentes regiones del país: 9 secuencias del norte, 6 del centro y 6 del sur. Por otro lado, la rama B incluyó 10 secuencias de ovejas y 3 de cabras, también provenientes de las tres diferentes zonas de estudio: 8 secuencias del centro, 2 del norte y 3 del sur. Las secuencias de sPil14, gNeu14 y sChu14, siguen permaneciendo unidas, formando un clúster con una PP de 1. Las secuencias gYa11, gEnq18 y gMor17 también permanecen unidas (con el anexo de otras secuencias en el clúster) y su PP es de 1.

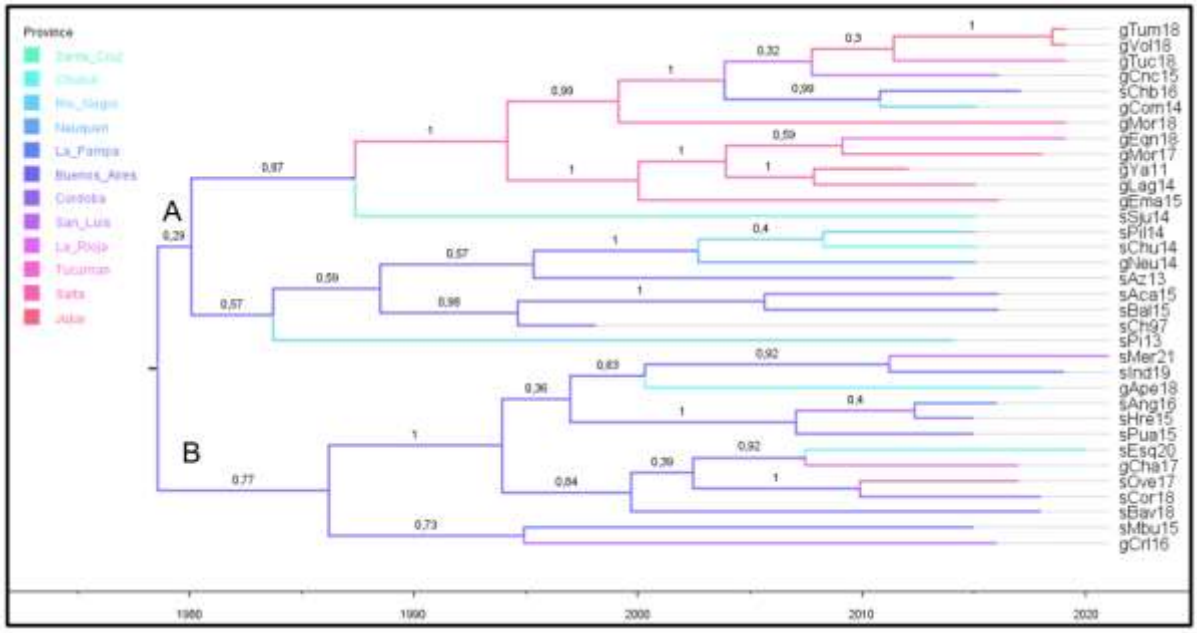

Figura 34. Árbol de MCC del concatenado de genes, estimado a partir del programa BEAST. Las líneas de las ramas están coloreadas de acuerdo a la provincia en la cual fue tomada la muestra, los números sobre las ramas indican la probabilidad posterior estimada y en el margen inferior se muestra una escala de tiempo para referencia de los años de presentación del foco.

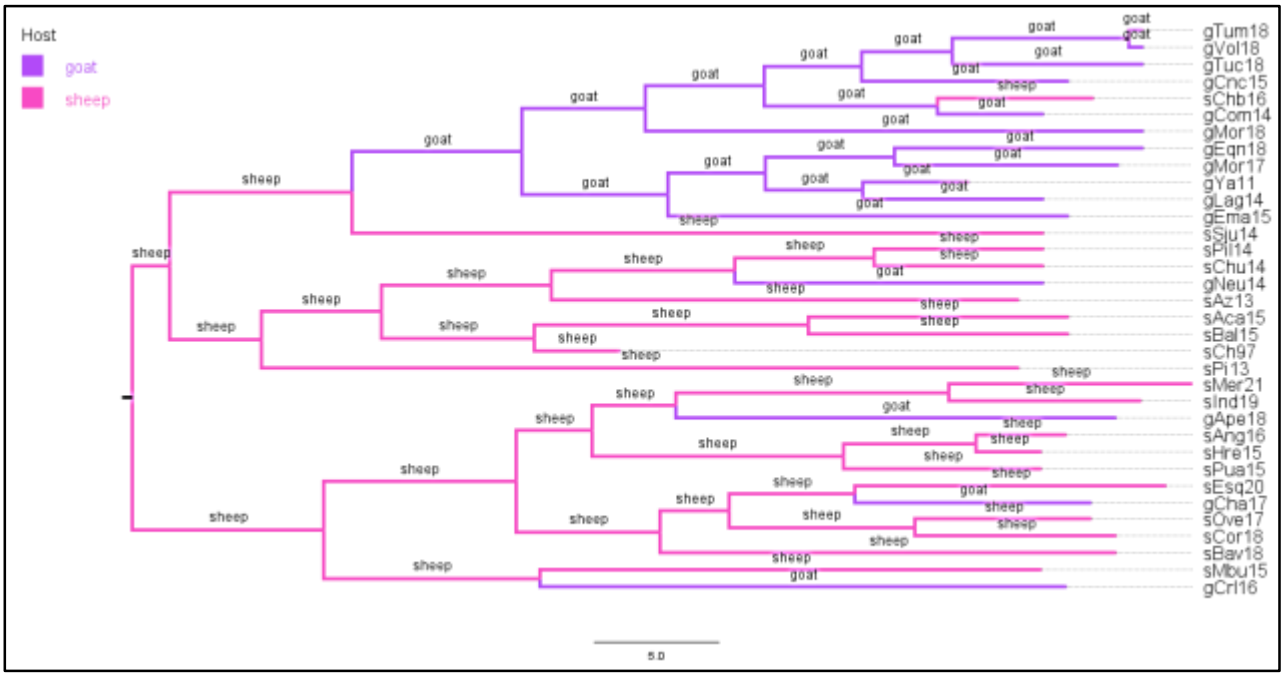

Figura 35. Árbol de MCC del concatenado de genes, estimado a partir del programa BEAST. Las líneas de las ramas están coloreadas de acuerdo hospedero del cual fue tomada la muestra. En el margen inferior se muestra una escala de tiempo para referencia de los años de presentación del foco.

Análisis evolutivos

La estimación de la tasa de evolución y los ACMR para cada gen y el concatenado se muestran en la tabla 8.

GEN	TASA DE EVOLUCION sustituciones/sitio/año	HPD 95% sustituciones/sitio/año	AÑO DE ACMR	HPD 95%
011	$7x10^{-3}$	1x10-3-2x10-^{2}	1994	1987-1996
020	$5.9x10^{-3}$	6x10-4-3x10^{-2}	1952	1866-1996
109	$3x10^{-3}$	7.7x10-4-6x10^{-3}	1967	1916-1996
127	$2.4x10^{-4}$	4x10-5-4x10^{-4}	1951	1881-1996
Concatenado	$1x10^{-3}$	3x10-4-2x10^{-3}	1971	1931-1996

Tabla 8: Análisis evolutivo de los genes individuales y del concatenado proveniente de secuencias de ORFV argentinas.

Análisis filogeográficos

La reconstrucción filogeográfica con la información suministrada por el concatenado de genes revela que el virus comenzó a dispersarse por el centro, y posteriormente hacia el norte y sur, culminando con la colonización de todas las provincias analizadas. La probabilidad de que la endemia haya iniciado en el centro del país es muy alta (0,94 para la Rama A y 0,97 para la rama B). Un patrón similar se pudo evidenciar en el análisis por cada uno de los genes individuales, aunque estos análisis presentaban un menor soporte en general y algunas diferencias en la escala de tiempo.

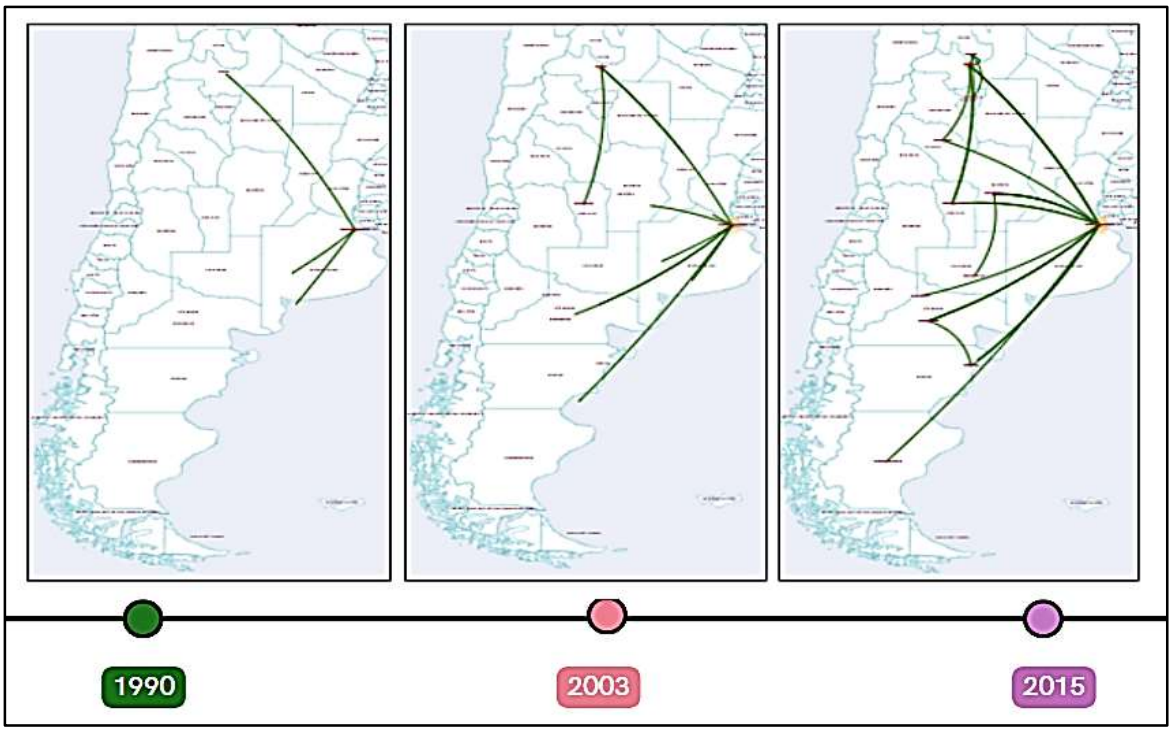

Figura 36. Reconstrucción filogeográfica del concatenado de genes estimada a partir del programa BEAST. En el margen inferior se muestra una escala de tiempo para referencia de los años de presentación de los focos

Discusión y conclusiones capítulo 3

Los modelos de epidemiología molecular son más comunes en el estudio de virus de ARN debido a que la naturaleza más variable de su genoma permite una comprensión detallada de su historia evolutiva en un período de tiempo más corto. Sin embargo, en este estudio logramos aplicar con éxito estos modelos que, utilizando datos epidemiológicos y secuencias virales, nos permitieron abordar preguntas relacionadas con la caracterización y dinámica viral durante el largo proceso de la endemia del virus ORF en Argentina.

En relación al objetivo específico de explorar la diversidad de las variantes circulantes del virus ORF en el país, observamos que existe una considerable diversidad de variantes del virus en el territorio. El uso de herramientas filogenéticas y análisis evolutivos nos permite llegar a esa conclusión a pesar de que en este estudio solo examinamos 4 de los 132 genes que componen el virus ORF.

Por otro lado, los resultados obtenidos no indican que estas variantes estén agrupadas en regiones geográficas específicas ni estén limitadas a hospedadores particulares sino más bien que se distribuyen de manera heterogénea en todo el territorio. Es importante destacar que estas conclusiones deben ser evaluadas con precaución ya que se basan en un conjunto de datos limitado. Para obtener resultados más sólidos desde el punto de vista estadístico, es fundamental continuar recopilando datos y muestras y extender las regiones del genoma secuenciadas para poder fortalecer los análisis y poner a prueba nuevamente estas hipótesis.

Tratándose de un virus de ADN, tanto la diversidad encontrada como la falta de agrupamiento de las secuencias analizadas, son indicativas de la ocurrencia de un proceso largo de diversificación y evolución viral en el país y/o de múltiples entradas de distintas variantes de ORFV a la Argentina.

Por otra parte, este trabajo ha proporcionado información que sugiere la identificación de posibles casos involucrados en una misma cadena de contagio. Un ejemplo de ello es el caso de las muestras sPil14, gNeu14 y sChu14, que en todas las reconstrucciones filogenéticas, ya sea en el análisis individual de los genes o a través del concatenado, siempre formaron un grupo conjunto respaldado por sólidos soportes en todos los árboles. Además, la cercanía geográfica y temporal de las muestras también respaldan esta hipótesis y completa aún más la cadena de transmisión sugerida anteriormente en (Peralta et al en 2018).

Continuando con los objetivos establecidos en este capítulo, fue de gran importancia explorar la posibilidad de identificar si alguno de los cuatro genes estudiados puede ser utilizado como marcador de especie. Anteriormente, Coradduzza y colaboradores en 2021 han sugerido al gen orf020 como marcador, aunque el análisis se realizó para un número limitado de secuencias y restringido a una zona geográfica. En este estudio, que comparte estas limitaciones, también logramos identificar agrupamientos por especie para este gen, aunque estos agrupamientos no son completamente excluyentes entre sí. Puntualmente, se observaron dos casos de secuencias caprinas incluidas en el clúster ovino (gNeu14 y gCrl16) y una secuencia de oveja (sChb16) en el grupo de cabras. Como esos eventos ocurren en ramas terminales del árbol y no se detectó que luego, este gen se mantuviera en el hospedador, se podría considerar que se trata de salto interespecie eventual del gen viral.

Tanto en el caso de la muestra sChb16 como en el caso de gNeu14, no solo el gen orf020, sino que los cuatro genes y el concatenado sugieren que su virus ancestral proviene de la especie contraria. Igualmente, cabe resaltar que los otros genes y el concatenado generan un patrón más difuso entre clúster que contienen tanto caprinos como ovinos y son peores indicadores que orf020.

En relación con la tasa de evolución inferida en este estudio, ya sea para los genes individuales como para el concatenado, se sitúa en el rango de 10^{-3} a 10^{-4} sustituciones por sitio por año. En un estudio realizado por (Li et al., 2023), se estimó que las secuencias de orf11, orf020, orf109 y orf127 tenían una tasa de evolución que oscilaba entre 10^{-4} y 10^{-5} sustituciones por sitio por año. Por otro lado, Coradduzza y colaboradores en 2021 (Coradduzza et al., 2021) estimaron que el gen orf020 presentaba una tasa de alrededor de 10^{-3} sustituciones por sitio por año. Nuestros resultados se asemejan más a estos últimos hallazgos. Es importante destacar que estos valores obtenidos son significativamente más altos en comparación con las tasas reportadas para el genoma completo del virus ($1{,}324 \times 10^{-6}$ sustituciones/sitio/año) (Coradduzza et al., 2022). Esto puede atribuirse al hecho de que los genes evaluados en este trabajo están directamente vinculados a la respuesta inmunológica del hospedador frente al virus y a la virulencia, lo que podría conducir a tasas de sustitución más elevadas.

Por otro lado, los análisis filogeográficos señalan una probabilidad significativa de que la dispersión del virus se haya originado en el centro del país, extendiéndose posteriormente hacia las regiones norte y sur. El Río de la Plata, situado en la provincia de Buenos Aires, funcionó como un corredor a través del cual se introdujeron varios ejemplares ovinos en el pasado, ya que el transporte fluvial era una de las principales vías para trasladar ganado y mercancías a Argentina (Montoya, 1984). Se plantea la hipótesis de que este podría haber sido un importante canal de introducción del virus al país y su posterior diversificación que condujo a las variantes posiblemente relacionadas con el ACMR de nuestras muestras, estimado en el año 1971, con un intervalo de confianza del 95% que abarca desde 1931 hasta 1996. Es importante destacar que estos resultados deben interpretarse con precaución debido a la limitada disponibilidad de información. Sin embargo, es notable que todos los genes exhibieron un patrón consistente de dispersión, respaldado por una sólida probabilidad de localización, independientemente de si se trató de la introducción de una o distintas variantes en el país.

CONCLUSIONES FINALES

Conclusiones generales de la tesis

En las últimas décadas, el acceso a la generación de datos de secuencias ha experimentado un notable incremento. La disponibilidad de este recurso no solo aporta beneficios en el área científica, sino que, bien aprovechada, también tiene un impacto muy relevante en ámbitos productivos y de salud pública. En esta tesis, exploramos el potencial de los modelos de epidemiología molecular aplicados a algunas enfermedades virales que en la actualidad generan tanto epidemias como endemias, para el desarrollo de estrategias destinadas a controlar y erradicar enfermedades relevantes en la salud pública veterinaria.

En este trabajo, nos enfocamos en determinar si estos modelos nos permiten identificar momentos óptimos de intervención durante el curso de una epidemia, mejorando la efectividad de las medidas de control. Además, buscamos comprender si estos modelos son capaces de identificar los determinantes de transmisión y diseminación de las enfermedades estudiadas. Se exploró además el papel de las formas de transmisión, como los fómites, en dichas epidemias y endemias. Asimismo, se pretendió evaluar la concordancia entre las medidas preventivas, como la vacunación, y la similitud genética de las variantes circulantes de virus, entre otros.

Esta tesis concluyó que la integración de datos epidemiológicos con secuencias virales, realizada en cada caso de estudio, proporcionó una exploración más profunda de las dinámicas de transmisión y evolución de los virus examinados. Mediante la conjunción de las diferentes fuentes de información, se logró formular hipótesis sobre la dispersión espacial y el desarrollo epidémico o endémico en cada contexto. Además, posibilitó la identificación de nuevas rutas de transmisión y permitió refinar aún más los análisis previos, dando lugar a la generación de nuevas hipótesis sobre la transmisión y de ancestria en los casos estudiados.

Los modelos de epidemiología molecular permitieron proponer la probable existencia de eventos de efecto fundador en zonas geográficas aisladas. También destacaron la importancia de los fómites en la diseminación de los posibles agentes causales y subrayaron la necesidad del testeo continuo en granjas para asegurar el estatus sanitario de las mismas. Estos modelos también facilitaron la comprensión de la genética del virus circulante, su dinámica de dispersión y sus resultados dieron soporte científico a propuestas de aumento en los esfuerzos de vigilancia y control en zonas geográficas críticas para la diseminación de la enfermedad estudiada.

Los modelos de epidemiologia molecular permitieron identificar cadenas locales de transmisión y confrontar hipótesis sobre posibles genes marcadores de especie, abriendo un espacio para profundizar en el estudio de los mismos y su compleja interacción con el hospedero.

Por lo tanto, con este cuerpo de evidencia, se recomienda la utilización de estas herramientas como apoyo en la toma de decisiones tanto en salud pública veterinaria como humana, a fin de posibilitar la obtención de resultados más eficaces y rápidos en la lucha contra las

enfermedades infecciosas. Esto, a su vez, facilitará el control de posibles epidemias que podrían afectar no solo la salud de los animales y los seres humanos, sino también la economía y el desarrollo social de los países involucrados.

Se propone la incorporación rutinaria de herramientas de epidemiología molecular en las operaciones de los organismos encargados de preservar la salud pública veterinaria. De este modo será posible prevenir, controlar y reducir riesgos sanitarios y biológicos y contribuir de manera significativa a una producción agropecuaria competitiva, sustentable y económicamente rentable en el continente sudamericano.

REFERENCIAS BIBLIOGRAFICAS

Referencias bibliográficas

- Abrahão, J.S., Campos, R.K., Trindade, G.S., Guedes, M.I., Lobato, Z.I., Mazur, C., Ferreira, P.C., Bonjardim, C.A., Kroon, E.G., 2009. Detection and phylogenetic analysis of Orf virus from sheep in Brazil: a case report. Virol. J. 6, 1–4.
- Alegría, M.E., Rivera, H., Manchego, A., 1998. Evidencia del virus del síndrome reproductivo y respiratorio porcino de crianza tecnificada. Rev Inv Pec IVITA 9, 53–58.
- Al-Eitan, L., Sendyani, S., Alnimri, M., 2023. Applications of the one health concept: current status in the Middle East. J. Biosaf. Biosecurity.
- Arenas, M., 2015. Trends in substitution models of molecular evolution. Front. Genet. 6, 319.
- Arias, A., López, P., Sánchez, R., Yamamura, Y., Rivera-Amill, V., 2018. Sanger and next generation sequencing approaches to evaluate HIV-1 virus in blood compartments. Int. J. Environ. Res. Public. Health 15, 1697.
- Arruda, A.G., Tousignant, S., Sanhueza, J., Vilalta, C., Poljak, Z., Torremorell, M., Alonso, C., Corzo, C.A., 2019. Aerosol detection and transmission of porcine reproductive and respiratory syndrome virus (PRRSV): What is the evidence, and what are the knowledge gaps? Viruses 11, 712.
- Auty, H., Mellor, D., Gunn, G., Boden, L.A., 2019. The risk of foot and mouth disease transmission posed by public access to the countryside during an outbreak. Front. Vet. Sci. 6, 381.
- Bartlett, P.C., Judge, L.J., 1997. The role of epidemiology in public health. Rev. Sci. Tech.-Off. Int. Epizoot. 16, 331–336.
- Belshaw, R., Sanjuan, R., Pybus, O.G., 2011. Viral mutation and substitution: units and levels. Curr. Opin. Virol. 1, 430–435.
- Bielejec, F., Rambaut, A., Suchard, M.A., Lemey, P., 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics 27, 2910–2912. doi:10.1093/bioinformatics/btr481
- Björnham, O., Sigg, R., Burman, J., 2020. Multilevel model for airborne transmission of foot-and-mouth disease applied to Swedish livestock. PloS One 15, e0232489.
- Bonneau, K., Mullens, B., MacLachlan, N., 2001. Occurrence of genetic drift and founder effect during quasispecies evolution of the VP2 and NS3/NS3A genes of bluetongue virus upon passage between sheep, cattle, and Culicoides sonorensis. J. Virol. 75, 8298–8305.
- Brinkman, F.S., Leipe, D.D., 2001. Phylogenetic analysis. Bioinforma. Pract. Guide Anal. Genes Proteins 2, 349.
- Brito, B., König, G., Cabanne, G.S., Beascoechea, C.P., Rodriguez, L., Perez, A., 2016. Phylogeographic analysis of the 2000–2002 foot-and-mouth disease epidemic in Argentina. Infect. Genet. Evol. 41, 93–99.
- Brown, J., Pirrung, M., McCue, L.A., 2017. FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinformatics 33, 3137–3139.
- Brown, T.A., 2002. Understanding a genome sequence, in: Genomes. 2nd Edition. Wiley-Liss.

- Bukar, A.M., Jesse, F.F.A., Abdullah, C.A.C., Noordin, M.M., Lawan, Z., Mangga, H.K., Balakrishnan, K.N., Azmi, M.-L.M., 2021. Immunomodulatory strategies for parapoxvirus: current status and future approaches for the development of vaccines against orf virus infection. Vaccines 9, 1341.
- Bushnell, B., 2020. BBDuk. Jt Genome Inst Available Online Httpsjgi Doe Govdata--Toolsbbtoolsbb-Tools-Userguidebbduk-Guid. 25 June 2020.
- Canals, M., Cáceres, D., 2020. Una Salud: conectando la salud humana, animal y ambiental. Cuad. Méd. Soc. 60, 9–18.
- Castillo Espinoza, A., Ramírez Velásquez, M., 2021. Síndrome Reproductivo y Respiratorio Porcino: Una revisión del agente etiológico y su influencia en el comportamiento actual de la enfermedad. Rev. Investig. Vet. Perú 32.
- Chadha, A., Dara, R., Pearl, D.L., Gillis, D., Rosendal, T., Poljak, Z., 2023. Classification of porcine reproductive and respiratory syndrome clinical impact in Ontario sow herds using machine learning approaches. Front. Vet. Sci. 10, 1175569.
- Charleston, B., Bankowski, B.M., Gubbins, S., Chase-Topping, M.E., Schley, D., Howey, R., Barnett, P.V., Gibson, D., Juleff, N.D., Woolhouse, M.E., 2011. Relationship between clinical signs and transmission of an infectious disease and the implications for control. Science 332, 726–729.
- Colvero, L.P., Villarreal, L.Y.B., Torres, C.A., Brañdo, P.E., 2015. Assessing the economic burden of avian infectious bronchitis on poultry farms in Brazil. Rev Sci Tech 34, 993–9.
- Coradduzza, E., Sanna, D., Rocchigiani, A.M., Pintus, D., Scarpa, F., Scivoli, R., Bechere, R., Dettori, M.A., Montesu, M.A., Marras, V., 2021. Molecular insights into the genetic variability of ORF Virus in a Mediterranean region (Sardinia, Italy). Life 11, 416.
- Coradduzza, E., Sanna, D., Scarpa, F., Azzena, I., Fiori, M.S., Scivoli, R., Rocchigiani, A.M., Bechere, R., Dettori, M.A., Pintus, D., 2022. A deeper insight into evolutionary patterns and phylogenetic history of ORF virus through the whole genome sequencing of the first Italian strains. Viruses 14, 1473.
- Cosgrove, A., 1962. An apparently new disease of chickens: avian nephrosis. Avian Dis. 6, 385–389.
- Cottam, E.M., Thébaud, G., Wadsworth, J., Gloster, J., Mansley, L., Paton, D.J., King, D.P., Haydon, D.T., 2008. Integrating genetic and epidemiological data to determine transmission pathways of foot-and-mouth disease virus. Proc. R. Soc. B Biol. Sci. 275, 887–895.
- Cunningham, A.A., Daszak, P., Wood, J.L., 2017. One Health, emerging infectious diseases and wildlife: two decades of progress? Philos. Trans. R. Soc. B Biol. Sci. 372, 20160167.
- Damon, I.K., 2011. Poxviruses. Man. Clin. Microbiol. 1647–1658.
- Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., 2021. Twelve years of SAMtools and BCFtools. Gigascience 10, giab008.
- Dellicour, S., Troupin, C., Jahanbakhsh, F., Salama, A., Massoudi, S., Moghaddam, M.K., Baele, G., Lemey, P., Gholami, A., Bourhy, H., 2019. Using phylogeographic approaches to analyse the dispersal history, velocity and direction of viral lineages—Application to rabies virus spread in Iran. Mol. Ecol. 28, 4335–4350.

- Derryberry, D.R., Schou, S.B., Conover, W., 2010. Teaching rank-based tests by emphasizing structural similarities to corresponding parametric tests. J. Stat. Educ. 18.
- Di Nardo, A., Ferretti, L., Wadsworth, J., Mioulet, V., Gelman, B., Karniely, S., Scherbakov, A., Ziay, G., Özyörük, F., Parlak, Ü., 2021. Evolutionary and ecological drivers shape the emergence and extinction of foot-and-mouth disease virus lineages. Mol. Biol. Evol. 38, 4346–4361.
- Domingo, E., Baranowski, E., Escarmís, C., Sobrino, F., 2002. Foot-and-mouth disease virus. Comp. Immunol. Microbiol. Infect. Dis. 25, 297–308.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 1–8.
- Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22, 1185–1192.
- Ejeromedoghene, O., Tesi, J.N., Uyanga, V.A., Adebayo, A.O., Nwosisi, M.C., Tesi, G.O., Akinyeye, R.O., 2020. Food security and safety concerns in animal production and public health issues in Africa: A perspective of COVID-19 pandemic era. Ethics Med. Public Health 15, 100600.
- Esposito, M.M., Turku, S., Lehrfield, L., Shoman, A., 2023. The Impact of Human Activities on Zoonotic Infection Transmissions. Animals 13, 1646.
- Eterradossi, N., Saif, Y.M., 2013. Infectious bursal disease. Dis. Poult. 219–246.
- Eybpoosh, S., Haghdoost, A.A., Mostafavi, E., Bahrampour, A., Azadmanesh, K., Zolala, F., 2017. Molecular epidemiology of infectious diseases. Electron. Physician 9, 5149.
- Foxman, B., Riley, L., 2001. Molecular epidemiology: focus on infection. Am. J. Epidemiol. 153, 1135–1141.
- Frérot, M., Lefebvre, A., Aho, S., Callier, P., Astruc, K., Aho Glélé, L.S., 2018. What is epidemiology? Changing definitions of epidemiology 1978-2017. PloS One 13, e0208442.
- Groves, R., MacDonald, D., Jones, E.W., 1990. Human orf. Arch. Dermatol. 126, 1649–1649.
- Gruetzmacher, K., Karesh, W.B., Amuasi, J.H., Arshad, A., Farlow, A., Gabrysch, S., Jetzkowitz, J., Lieberman, S., Palmer, C., Winkler, A.S., 2021. The Berlin principles on one health–Bridging global health and conservation. Sci. Total Environ. 764, 142919.
- Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Presented at the Nucleic Acids Symp. Ser., pp. 95–98.
- Holland, N.T., Smith, M.T., Eskenazi, B., Bastaki, M., 2003. Biological sample collection and processing for molecular epidemiological studies. Mutat. Res. Mutat. Res. 543, 217–234.
- Holtz, A., Baele, G., Bourhy, H., Zhukova, A., 2023. Integrating full and partial genome sequences to decipher the global spread of canine rabies virus. Nat. Commun. 14, 4247.
- Ishikawa, S.A., Zhukova, A., Iwasaki, W., Gascuel, O., 2019. A fast likelihood method to reconstruct and visualize ancestral scenarios. Mol. Biol. Evol. 36, 2069–2085.
- Jamal, S.M., Belsham, G.J., 2013. Foot-and-mouth disease: past, present and future. Vet. Res. 44, 1–14.

- Jukes, T.H., Cantor, C.R., 1969. Evolution of protein molecules. Mamm. Protein Metab. 3, 21–132.
- Kalkauskas, A., Perron, U., Sun, Y., Goldman, N., Baele, G., Guindon, S., De Maio, N., 2021. Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk. PLOS Comput. Biol. 17, e1008561.
- Kassa, T., 2021. A review on human orf: a neglected viral zoonosis. Res. Rep. Trop. Med. 153–172.
- Katoh, K., Standley, D.M., 2014. MAFFT: iterative refinement and additional methods. Mult. Seq. Alignment Methods 131–146.
- Kilbourne, E.D., 1973. The molecular epidemiology of influenza. J. Infect. Dis. 127, 478–487.
- Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.
- King, K.M., Van Doorslaer, K., 2018. Building (viral) phylogenetic trees using a maximum likelihood approach. Curr. Protoc. Microbiol. 51, e63.
- Kingman, J.F.C., 1982. The coalescent. Stoch. Process. Their Appl. 13, 235–248.
- Kottaridi, C., Nomikou, K., Lelli, R., Markoulatos, P., Mangana, O., 2006. Laboratory diagnosis of contagious ecthyma: comparison of different PCR protocols with virus isolation in cell culture. J. Virol. Methods 134, 119–124.
- Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.
- Lei, X., Cui, S., Zhao, Z., Wang, J., 2015. Etiology, pathogenesis, antivirals and vaccines of hand, foot, and mouth disease. Natl. Sci. Rev. 2, 268–284.
- Leigh, J.W., Bryant, D., 2015. POPART: full-feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110–1116.
- Lemey, P., Rambaut, A., Drummond, A.J., Suchard, M.A., 2009. Bayesian phylogeography finds its roots. PLoS Comput. Biol. 5, e1000520.
- Li, K., Wang, C., Yang, F., Cao, W., Zhu, Z., Zheng, H., 2021. Virus–host interactions in foot-and-mouth disease virus infection. Front. Immunol. 12, 571509.
- Li, S., Jing, T., Zhu, F., Chen, Y., Yao, X., Tang, X., Zuo, C., Liu, M., Xie, Y., Jiang, Y., 2023. Genetic Analysis of Orf Virus (ORFV) Strains Isolated from Goats in China: Insights into Epidemiological Characteristics and Evolutionary Patterns. Virus Res. 334, 199160.
- Logan, G., Freimanis, G.L., King, D.J., Valdazo-González, B., Bachanek-Bankowska, K., Sanderson, N.D., Knowles, N.J., King, D.P., Cottam, E.M., 2014. A universal protocol to generate consensus level genome sequences for foot-and-mouth disease virus and other positive-sense polyadenylated RNA viruses using the Illumina MiSeq. BMC Genomics 15, 1–10.
- Luk, H.K., Li, X., Fung, J., Lau, S.K., Woo, P.C., 2019. Molecular epidemiology, evolution and phylogeny of SARS coronavirus. Infect. Genet. Evol. 71, 21–30.
- Magiorkinis, G., Sypsa, V., Magiorkinis, E., Paraskevis, D., Katsoulidou, A., Belshaw, R., Fraser, C., Pybus, O.G., Hatzakis, A., 2013. Integrating phylodynamics and epidemiology to estimate transmission diversity in viral epidemics. PLoS Comput. Biol. 9, e1002876.
- Mahgoub, H.A., 2012. An overview of infectious bursal disease. Arch. Virol. 157, 2047–2057.

- Markov, P.V., Ghafari, M., Beer, M., Lythgoe, K., Simmonds, P., Stilianakis, N.I., Katzourakis, A., 2023. The evolution of SARS-CoV-2. Nat. Rev. Microbiol. 21, 361–379.
- Martin, D.P., Murrell, B., Khoosal, A., Muhire, B., 2017. Detecting and analyzing genetic recombination using RDP4. Bioinforma. Vol. Data Seq. Anal. Evol. 433–460.
- Mateu, E., Martín, M., Vidal, D., 2003. Genetic diversity and phylogenetic analysis of glycoprotein 5 of European-type porcine reproductive and respiratory virus strains in Spain. J. Gen. Virol. 84, 529–534.
- Mattion, N., König, G., Seki, C., Smitsaart, E., Maradei, E., Robiolo, B., Duffy, S., León, E., Piccone, M., Sadir, A., 2004. Reintroduction of foot-and-mouth disease in Argentina: characterisation of the isolates and development of tools for the control and eradication of the disease. Vaccine 22, 4149–4162.
- Michel, L.O., Jackwood, D.J., 2017. Classification of infectious bursal disease virus into genogroups. Arch. Virol. 162, 3661–3670.
- Mielke, S.R., Garabed, R., 2020. Environmental persistence of foot-and-mouth disease virus applied to endemic regions. Transbound. Emerg. Dis. 67, 543–554.
- Milne, I., Stephen, G., Bayer, M., Cock, P.J., Pritchard, L., Cardle, L., Shaw, P.D., Marshall, D., 2013. Using Tablet for visual exploration of second-generation sequencing data. Brief. Bioinform. 14, 193–202.
- 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. Mol. Biol. Evol. 37, 1530–1534.
- Montaner-Tarbes, S., Del Portillo, H.A., Montoya, M., Fraile, L., 2019a. Key gaps in the knowledge of the porcine respiratory reproductive syndrome virus (PRRSV). Front. Vet. Sci. 6, 38.
- Montaner-Tarbes, S., Del Portillo, H.A., Montoya, M., Fraile, L., 2019b. Key gaps in the knowledge of the porcine respiratory reproductive syndrome virus (PRRSV). Front. Vet. Sci. 6, 38.
- Montoya, A.J., 1984. Cómo evolucionó la ganadería en la época del virreinato.
- Munsey, A., Mwiine, F.N., Ochwo, S., Velazquez-Salinas, L., Ahmed, Z., Maree, F., Rodriguez, L.L., Rieder, E., Perez, A., VanderWaal, K., 2019. Spatial distribution and risk factors for foot and mouth disease virus in Uganda: Opportunities for strategic surveillance. Prev. Vet. Med. 171, 104766.
- Musser, J.M., Waldron, D.F., Taylor, C.A., 2012. Evaluation of homologous and heterologous protection induced by a virulent field strain of Orf virus and an Orf vaccine in goats. Am. J. Vet. Res. 73, 86–90.
- Nahata, Kanika D, Bielejec, F., Monetta, J., Dellicour, S., Rambaut, A., Suchard, M.A., Baele, G., Lemey, P., 2022. SPREAD 4: online visualisation of pathogen phylogeographic reconstructions. Virus Evol. 8, veac088.
- Nahata, Kanika D., Bielejec, F., Monetta, J., Dellicour, S., Rambaut, A., Suchard, M.A., Baele, G., Lemey, P., 2022. SPREAD 4: online visualisation of pathogen phylogeographic reconstructions. Virus Evol. 8, veac088.
- Nathues, C., Perler, L., Bruhn, S., Suter, D., Eichhorn, L., Hofmann, M., Nathues, H., Baechlein, C., Ritzmann, M., Palzer, A., 2016. An outbreak of porcine reproductive and respiratory syndrome virus in Switzerland following import of boar semen. Transbound. Emerg. Dis. 63, e251–e261.

- Neira, V., Brito, B., Mena, J., Culhane, M., Apel, M.I., Max, V., Perez, P., Moreno, V., Mathieu, C., Johow, M., 2017a. Epidemiological investigations of the introduction of porcine reproductive and respiratory syndrome virus in Chile, 2013-2015. PLoS One 12, e0181569.
- Neira, V., Brito, B., Mena, J., Culhane, M., Apel, M.I., Max, V., Perez, P., Moreno, V., Mathieu, C., Johow, M., 2017b. Epidemiological investigations of the introduction of porcine reproductive and respiratory syndrome virus in Chile, 2013-2015. PLoS One 12, e0181569.
- Neumann, E.J., Kliebenstein, J.B., Johnson, C.D., Mabry, J.W., Bush, E.J., Seitzinger, A.H., Green, A.L., Zimmerman, J.J., 2005. Assessment of the economic impact of porcine reproductive and respiratory syndrome on swine production in the United States. J. Am. Vet. Med. Assoc. 227, 385–392.
- Orakpoghenor, O., Oladele, S.B., Abdu, P.A., 2020. Infectious bursal disease: Transmission, pathogenesis, pathology and control-an overview. Worlds Poult. Sci. J. 76, 292–303.
- Pileri, E., Mateu, E., 2016. Review on the transmission porcine reproductive and respiratory syndrome virus between pigs and farms and impact on vaccination. Vet. Res. 47, 108.
- Pybus, O.G., Rambaut, A., Harvey, P.H., 2000. An integrated framework for the inference of viral population history from reconstructed genealogies. Genetics 155, 1429–1437.
- Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.
- Raele, D.A., Stoffolano Jr, J.G., Vasco, I., Pennuzzi, G., Nardella La Porta, M.C., Cafiero, M.A., 2021. Study on the role of the common house fly, Musca domestica, in the spread of ORF Virus (Poxviridae) DNA under laboratory conditions. Microorganisms 9, 2185.
- Rambaut, A., 2009. FigTree. Tree figure drawing tool. Httptree Bio Ed Ac Uksoftwarefigtree.
- Rambaut, A., 2000. Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. Bioinformatics 16, 395–399.
- Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904.
- Rambaut, A., Lam, T.T., Max Carvalho, L., Pybus, O.G., 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol. 2, vew007.
- Rambaut, A., Suchard, M., Xie, D., Drummond, A., 2014. Tracer v1. 6. Computer program and documentation distributed by the author. Tracer V1 6 Comput. Program Doc. Distrib. Author.
- Ramírez, M., Bauermann, F.V., Navarro, D., Rojas, M., Manchego, A., Nelson, E.A., Diel, D.G., Rivera, H., 2019a. Detection of porcine reproductive and respiratory syndrome virus (PRRSV) 1-7-4-type strains in Peru. Transbound. Emerg. Dis. 66, 1107–1113.
- Ramírez, M., Bauermann, F.V., Navarro, D., Rojas, M., Manchego, A., Nelson, E.A., Diel, D.G., Rivera, H., 2019b. Detection of porcine reproductive and respiratory

syndrome virus (PRRSV) 1-7-4-type strains in Peru. Transbound. Emerg. Dis. 66, 1107–1113.

- Ramírez, M., Bauermann, F.V., Navarro, D., Rojas, M., Manchego, A., Nelson, E.A., Diel, D.G., Rivera, H., 2019c. Detection of porcine reproductive and respiratory syndrome virus (PRRSV) 1-7-4-type strains in Peru. Transbound. Emerg. Dis. 66, 1107–1113.
- Ramirez-Carvajal, L., Pauszek, S.J., Ahmed, Z., Farooq, U., Naeem, K., Shabman, R.S., Stockwell, T.B., Rodriguez, L.L., 2018. Genetic stability of foot-and-mouth disease virus during long-term infections in natural hosts. PLoS One 13, e0190977.
- Ramos, N., Betancour, G., Puig, J., Arbiza, J., 2022. An update on genetic analysis of porcine reproductive and respiratory syndrome virus type 2 (PRRSV-2) in South America: identification of ORF5 sequences of lineage 1A, 1C and 1G. Arch. Microbiol. 204, 367.
- Rasmussen, S.A., Goodman, R.A., 2018. The CDC field epidemiology manual. Oxford University Press.
- Rich, S.N., Prosperi, M.C., Dellicour, S., Vrancken, B., Cook, R.L., Spencer, E.C., Salemi, M., Mavian, C., 2022. Molecular Epidemiology of HIV-1 Subtype B Infection across Florida Reveals Few Large Superclusters with Metropolitan Origin. Microbiol. Spectr. 10, e01889-22.
- Robles, C.A., Peralta, A.V., Alvarez, L.P., Martínez, A., 2017. Brote atípico de ectima contagioso en ovinos Merino de la Patagonia Argentina.
- Ronquist, F., van der Mark, P., Huelsenbeck, J.P., 2009. Bayesian phylogenetic analysis using MrBayes. Phylogenetic Handb. 2, 723.
- Rozas, J., Rozas, R., 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinforma. Oxf. Engl. 15, 174–175.
- Rweyemamu, M., Roeder, P., Mackay, D., Sumption, K., Brownlie, J., Leforban, Y., Valarcher, J.-F., Knowles, N.J., Saraiva, V., 2008. Epidemiological patterns of foot-and-mouth disease worldwide. Transbound. Emerg. Dis. 55, 57–72.
- Samuel, A., Knowles, N., 2001. Foot-and-mouth disease type O viruses exhibit genetically and geographically distinct evolutionary lineages (topotypes). J. Gen. Virol. 82, 609–621.
- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. 74, 5463–5467.
- Shapiro, B., Rambaut, A., Pybus, O.G., Holmes, E.C., 2006. A phylogenetic method for detecting positive epistasis in gene sequences and its application to RNA virus evolution. Mol. Biol. Evol. 23, 1724–1730.
- Shi, M., Lemey, P., Brar, M.S., Suchard, M.A., Murtaugh, M.P., Carman, S., D'Allaire, S., Delisle, B., Lambert, M.-È., Gagnon, C.A., 2013. The spread of type 2 Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) in North America: a phylogeographic approach. Virology 447, 146–154.
- Stadejek, T., Oleksiewicz, M.B., Scherbakov, A.V., Timina, A.M., Krabbe, J.S., Chabros, K., Potapchuk, D., 2008. Definition of subtypes in the European genotype of porcine reproductive and respiratory syndrome virus: nucleocapsid characteristics and geographical distribution in Europe. Arch. Virol. 153, 1479–1488.

- Strimmer, K., Von Haeseler, A., 1997. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc. Natl. Acad. Sci. 94, 6815–6819.
- Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. 4, vey016.
- Sullivan, J., Joyce, P., 2005. Model selection in phylogenetics. Annu Rev Ecol Evol Syst 36, 445–466.
- Tan, S., Dvorak, C.M., Murtaugh, M.P., 2019. Rapid, unbiased PRRSV strain detection using MinION direct RNA sequencing and bioinformatics tools. Viruses 11, 1132.
- Tavaré, S., 1986. Some probabilistic and statistical problems on the analysis of DNA sequence. Lect. Math. Life Sci. 17, 57.
- Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., Minh, B.Q., 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44, W232–W235.
- Valdes-Donoso, P., Jarvis, L.S., 2022. Combining epidemiology and economics to assess control of a viral endemic animal disease: Porcine Reproductive and Respiratory Syndrome (PRRS). PloS One 17, e0274382.
- Vandamme, A.-M., Pybus, O.G., 2013. Viral phylogeny in court: the unusual case of the Valencian anesthetist. BMC Biol. 11, 1–3.
- VanderWaal, K., Deen, J., 2018. Global trends in infectious diseases of swine. Proc. Natl. Acad. Sci. 115, 11495–11500.
- Vera, F., Craig, M.I., Olivera, V., Rojas, F., König, G., Pereda, A., Vagnozzi, A., 2015. Molecular characterization of infectious bursal disease virus (IBDV) isolated in Argentina indicates a regional lineage. Arch. Virol. 160, 1909–1921.
- Volz, E.M., Koelle, K., Bedford, T., 2013. Viral phylodynamics. PLoS Comput. Biol. 9, e1002947.
- Volz, E.M., Kosakovsky Pond, S.L., Ward, M.J., Leigh Brown, A.J., Frost, S.D., 2009. Phylodynamics of infectious disease epidemics. Genetics 183, 1421–1430.
- Wang, W., He, X., Zhang, Y., Qiao, Y., Shi, J., Chen, R., Chen, J., Xiang, Y., Wang, Z., Chen, G., 2022. Analysis of the global origin, evolution and transmission dynamics of the emerging novel variant IBDV (A2dB1b): The accumulation of critical aa-residue mutations and commercial trade contributes to the emergence and transmission of novel variants. Transbound. Emerg. Dis. 69, e2832–e2851.
- Wilm, A., Aw, P.P.K., Bertrand, D., Yeo, G.H.T., Ong, S.H., Wong, C.H., Khor, C.C., Petric, R., Hibberd, M.L., Nagarajan, N., 2012. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. 40, 11189–11201.
- Wright, E.S., 2016. Using DECIPHER v2. 0 to analyze big biological sequence data in R. R J. 8.
- Xu, G., Li, J., Shen, Q., Hou, L., Xia, Y., Li, Q., Xie, S., Sun, S., Ding, J., Mao, Y., 2019. Complete genome characterization of a novel infectious bursal disease virus strain isolated from a chicken farm in China. Microbiol. Resour. Announc. 8, 10.1128/mra. 00632-19.
- Yang, Z., Rannala, B., 2012. Molecular phylogenetics: principles and practice. Nat. Rev. Genet. 13, 303–314.

- Yilmaz, A., Turan, N., Bayraktar, E., Gurel, A., Cizmecigil, U.Y., Aydin, O., Bamac, O.E., Cecchinato, M., Franzo, G., Tali, H.E., 2019. Phylogeny and evolution of infectious bursal disease virus circulating in Turkish broiler flocks. Poult. Sci. 98, 1976–1984.
- Zachar, T., Popowich, S., Goodhope, B., Knezacek, T., Ojkic, D., Willson, P., Ahmed, K.A., Gomis, S., 2016. A 5-year study of the incidence and economic impact of variant infectious bursal disease viruses on broiler production in Saskatchewan, Canada. Can. J. Vet. Res. 80, 255.
- Zimmerman, J.J., Dee, S.A., Holtkamp, D.J., Murtaugh, M.P., Stadejek, T., Stevenson, G.W., Torremorell, M., Yang, H., Zhang, J., 2019. Porcine reproductive and respiratory syndrome viruses (porcine arteriviruses). Dis. Swine 685–708.
- Zwart, M.P., Kupczok, A., Iranzo, J., 2023. Predicting virus evolution: from genome evolution to epidemiological trends. Front. Virol. 3, 1215709.