

Universidad de la República
Maestría en Bioinformática

Análisis espacio-temporal de las redes de transmisión del subtipo B de VIH-1 en Uruguay

Tesis de Maestría presentada al programa de posgrado PEDECIBA (Programa de Desarrollo de las Ciencias Básicas), para la obtención del título de Magister en Bioinformática.

Estudiante: Lic. Agustina Añasco Podchibiakin

Orientadora: Dra. Daiana Mir

Co-orientador: Dr. Gonzalo Bello

Salto, Uruguay
Mayo, 2026



A mamá, Patricia.

Agradecimientos

A PEDECIBA y la Maestría en Bioinformática, a la comisión, los docentes y mis compañeros, por hacerme sentir un par más en esta trayectoria y darme la oportunidad de llegar a donde nunca imaginé.

A la ANII y CAPES por otorgar el apoyo financiero que me permitió llevar adelante este proyecto.

A mamá y a mi abuela, Patricia y Adelia, que fueron mi apoyo en todo momento, se alegraron por cada logro conseguido y supieron acompañarme cuando todo parecía más grande que yo. Gracias a ambas por ser el abrazo más fuerte de mi vida.

A Daiana, por ir más allá del gran trabajo que conlleva ser orientador académico y ser guía en muchos ámbitos más de la vida. Gracias por considerarme en todo momento y hacerme sentir una parte importante en este equipo de investigación. Por todos los momentos compartidos como amigas y por todo el conocimiento que transmitiste en todos estos años.

A Gonzalo e Ighor. Gracias por ser grandes mentores en el laboratorio y más allá. A Zezé, por ser el cariño personificado y preocuparse incluso estando lejos.

Y a todas las personas que conocí en Río, por ser como una familia y mostrarme lo mejor de ese lugar.

A Ana, por ser el cable a tierra en incontables casos. Por darme un lugarcito en su casa y hacerme sentir parte de su familia desde el primer día, en cada mimo y reto que ligué en estos años, gracias.

A Facundo, por todo su amor, contención y paciencia. Por las buenas ideas y la mejor compañía. Y a su familia, por darme ese lugarcito lleno de cariño.

Y mis amigos, por mirarme orgullosos, confiar y recordarme que siempre me puedo levantar de nuevo.

A todos ellos, gracias por cada abrazo, cada sonrisa y por acompañarme en este proceso de crecimiento.

Índice

Resúmen.....	10
1.Introducción.....	11
1.1. VIH-1: panorama general.....	11
1.2. Situación epidemiológica.....	12
1.2.1. Estado actual de la pandemia del VIH.....	12
1.2.2. Epidemia de VIH/SIDA en Uruguay.....	13
1.3. Biología del VIH.....	14
1.3.1. Estructura y genoma.....	14
1.3.2. Ciclo replicativo.....	16
1.3.3. Dinámica de replicación, intervención terapéutica y resistencia.....	16
1.4. Mecanismos generadores de diversidad.....	17
1.5. Clasificación molecular.....	18
1.6. Distribución geográfica de los subtipos y recombinantes.....	20
1.7. Subtipo B: origen y diseminación.....	22
1.8. Herramientas y marco conceptual.....	26
1.8.1. Filogenia molecular.....	26
1.8.2. Modelos probabilísticos de sustitución nucleotídica.....	26
1.8.3. Métodos de inferencia filogenética.....	27
1.8.4. Evaluación del soporte estadístico.....	29
1.8.5. Reloj molecular.....	30
1.8.6. Filogeografía.....	31
1.8.7. Filodinámica.....	33
1.9. Justificación del estudio.....	36
2. Hipótesis.....	38
3. Objetivos.....	38
3.1. Objetivo general.....	38
3.2. Objetivos específicos.....	38
4. Metodología.....	39
4.1. Aspectos éticos.....	39
4.2. Datos, muestreo y secuencias virales.....	39
4.3. Asignación filogenética de clados.....	40
4.4. Análisis filogenético por máxima verosimilitud.....	41
4.5. Caracterización epidemiológica y demográfica de los metadatos asociados al dataset analizado.....	42
4.6. Análisis filogeográfico por máxima verosimilitud.....	43
4.7. Inferencia bayesiana de la escala temporal.....	44
4.8. Reconstrucciones demográficas bayesianas.....	45
5. Resultados.....	47
5.1. Introducciones de VIH-1B y clados de transmisión local.....	47
5.2. Caracterización epidemiológica y demográfica del dataset analizado.....	52
5.3. Análisis evolutivos e inferencia demográfica.....	56

5.4. Origen geográfico de la epidemia BPAN en Uruguay.....	62
6. Discusión.....	64
7. Limitaciones.....	69
8. Conclusiones.....	71
9. Bibliografía.....	72
ANEXO 1. Manuscrito asociado a la presente tesis.....	86
ANEXO 2. Manuscrito derivado de la presente tesis.....	108

Lista de figuras

Figura 1. Estructura y organización genómica del VIH-1	15
Figura 2. Origen y clasificación del VIH	19
Figura 3. Distribución de los distintos subtipos y FRCs de VIH-1 prevalentes en el mundo	21
Figura 4. Origen y diseminación geográfica del VIH-1B en el mundo.....	23
Figura 5. Proporción estimada de los clados B_{CAR} y B_{PAN} de VIH-1 en distintos países de América Latina.....	25
Figura 6. Representación de la teoría coalescente.....	34
Figura 7. Esquema conceptual del marco de inferencia bayesiana implementado en los análisis realizados en BEAST y BEAST 2.....	46
Figura 8. Árbol filogenético de máxima verosimilitud utilizado para la asignación filogenética de las secuencias uruguayas de VIH-1 subtipo B.....	48
Figura 9. Árboles filogenéticos de máxima verosimilitud de clados B_{PAN} previamente descritos en América Latina con presencia de secuencias uruguayas.....	50
Figura 10. Subclados B_{PAN-UY} e introducciones esporádicas dentro de la diversidad B_{PAN} no asociada a clados latinoamericanos conocidos.....	51
Figura 11. Distribución proporcional de las introducciones independientes de variantes B_{PAN} en Uruguay.....	52
Figura 12. Proporción de clados pequeños, medios y grandes según el sexo predominante en cada uno.....	54
Figura 13. Distribución temporal de los clados de transmisión de B_{PAN} según su tiempo estimado al ancestro común más reciente (tMRCA).....	57
Figura 14. Filogenia calibrada en el tiempo y reconstrucción filodinámica conjunta de los principales subclados uruguayos de B_{PAN}	61
Figura 15. Distribución de los eventos de transmisión secundaria inferidos según el tiempo estimado desde la infección.....	62

Lista de tablas

Tabla 1. Clasificación filogenética de las secuencias del gen pol de VIH-1B de Uruguay...	49
Tabla 2. Distribución por sexo de los pacientes según los sublinajes de VIH-1 B _{PAN} identificados a partir de los metadatos de la plataforma GENOUY-VIH.....	53
Tabla 3. Distribución según vía de transmisión de las secuencias pertenecientes a los sublinajes de VIH-1 B _{PAN} identificados a partir de los metadatos asociados a las secuencias de la base de datos GENOUY-VIH.....	55
Tabla 4. Tiempo del ancestro común más reciente (<i>tMRCA</i>), origen geográfico estimado y número de secuencias (<i>N</i>) de los clados uruguayos identificados dentro de B _{PAN-SAM} , B _{PAN-AR} y B _{PAN-BR}	58
Tabla 5. Tiempo del ancestro común más reciente (<i>tMRCA</i>) y número de secuencias (<i>N</i>) de los clados identificados como B _{PAN-UY}	59

Lista de siglas y abreviaturas

ADN	Ácido Desoxirribonucleico
aLRT	Aproximate Likelihood Ratio Test
ARN	Ácido Ribonucleico
B_{CAR}	Subtipo B Caribeño
BDSKY	Birth-Death Skyline
B_{PAN-AR}	Subtipo B Pandémico Argentino
B_{PAN-BR}	Subtipo B Pandémico Brasileiro
B_{PAN-CAM}	Subtipo B Pandémico Centroamericano
B_{PAN-MX}	Subtipo B Pandémico Mexicano
B_{PAN-PE}	Subtipo B Pandémico Peruano
B_{PAN-SAM}	Subtipo B Pandémico Sudamericano
B_{PAN}	Subtipo B Pandémico
BSKG	Bayesian Skygrid
BSP	Bayesian Skyline Plot
CDC	Centros para el Control y la Prevención de Enfermedades
CENUR	Centro Universitario Regional
CSIC	Comisión Sectorial de Investigación Científica
CTMC	Continuous-Time Markov Chain
DLSP	Departamento de Laboratorios de Salud Pública
EPA	Evolutionary Placement Algorithm
ESS	Effective Sample Size
F81	Felsenstein model
FRC	Forma Recombinante Circulante
FRU	Forma Recombinante Única
GTR	General Time Reversible model
HKY85	Hasegawa-Kishino-Yano model
HSB	Hombres que tienen Sexo con Hombres
HTX	Transmisión Heterosexual
ICTV	Comité Internacional de Taxonomía de Virus
IN	Integrasa
JC69	Jukes y Cantor model
K80	Kimura 2 Parameters model
LAV	Virus Asociado a la Linfadenopatía
LTR	Long Terminal Repeats

MAP	Maximum a posteriori
MCCT	Maximum Clade Credibility Tree
MCMC	Markov Chain Monte Carlo
ML	Maximum Likelihood (Máxima Verosimilitud)
MRCA	Most Recent Common Ancestor
MSP	Ministerio de Salud Pública
ND	No data/Sin dato
Ne	Tamaño Poblacional Efectivo
NJ	Neighbor-Joining
NNI	Nearest Neighbor Interchange
ONU	Organización de las Naciones Unidas
ONUSIDA	Programa Conjunto de las Naciones Unidas sobre el VIH/SIDA.
PR	Proteasa
RDC	República Democrática del Congo
Re	Número Reproductivo Efectivo
RT-PCR	Reverse Transcription Polymerase Chain Reaction
RT	Retrotranscriptasa
SH-aLRT	Simodara-Hasegawa Approximate Likelihood Ratio Test
SIDA	Síndrome de Inmunodeficiencia Adquirida
SMS	Smart Model Selection
SPR	Subtree Pruning and Regrafting
TAR	Terapia Antirretroviral
TARV	Terapia Antirretroviral
TBR	Tree Bisection and Reconnection
tMRCA	Time to the Most Recent Common Ancestor
TN93	Tamura-Nei model
UDI	Usuarios de drogas intravenosas
VIH-1	Virus de Inmunodeficiencia Humana tipo 1
VIH-1B	Virus de Inmunodeficiencia Humana tipo 1, subtipo B
VIH-2	Virus de Inmunodeficiencia Humana tipo 2
VIH	Virus de Inmunodeficiencia Humana
VIS	Virus de Inmunodeficiencia Simia

Resumen

El subtipo B del VIH-1 (VIH-1B) es el linaje mejor caracterizado de la pandemia de VIH. En las Américas, VIH-1B constituye la variante genética predominante y se encuentra representado por dos grupos genéticamente diferenciados: el clado pandémico (B_{PAN}), responsable de la mayoría de las infecciones, y los clados caribeños parafiléticos (B_{CAR}). En Uruguay, el subtipo B representa el 39,4% de los casos de VIH; sin embargo, la contribución relativa y la dinámica espacio-temporal de los clados pandémicos y no pandémicos permanecen insuficientemente caracterizadas. En este trabajo se analizaron un total de 491 secuencias del gen *pol* de VIH-1B y sus metadatos asociados, obtenidos a partir de la vigilancia nacional en Uruguay entre 2007 y 2021. Los análisis filogenéticos permitieron identificar secuencias pertenecientes a los clados B_{PAN} y B_{CAR} , así como reconstruir eventos de introducción y dispersión viral. El análisis filogeográfico se realizó mediante reconstrucción de estados ancestrales. La dinámica poblacional y epidemiológica de los principales clados uruguayos B_{PAN} fue evaluada mediante los modelos Bayesian Skygrid y Birth–Death Skyline. Los resultados evidenciaron que la epidemia uruguaya de VIH-1B estuvo dominada por el clado B_{PAN} (96,3%), con una contribución minoritaria de los clados B_{CAR} (3,7%). Se identificaron 277 introducciones independientes de B_{PAN} a Uruguay, cerca del 80% de ellas provenientes de Brasil. Dos grandes linajes específicos de Uruguay ($B_{PAN-UY-I}$ y $B_{PAN-UY-II}$), emergidos entre mediados y finales de la década de 1980, representaron una proporción importante de las infecciones locales. Los análisis filodinámicos indicaron un crecimiento epidémico sostenido de $B_{PAN-UY-I}$ y $B_{PAN-UY-II}$ hasta aproximadamente 2005, seguido de una estabilización parcial (2005–2015) y una posterior disminución de la transmisión en los años más recientes. Cerca de la mitad de los eventos de transmisión inferidos para $B_{PAN-UY-I}$ y $B_{PAN-UY-II}$ ocurrieron dentro de los primeros dos años posteriores a la infección primaria. En conjunto, los resultados indican que la epidemia uruguaya de VIH-1B se caracteriza por frecuentes introducciones, pero con un establecimiento limitado de linajes, donde la expansión epidémica fue impulsada por un pequeño número de linajes pandémicos específicos del país. Los patrones filodinámicos indican una epidemia madura en transición desde una fase de expansión hacia una estabilización, seguida por una reciente disminución en la transmisión. La concentración de eventos de transmisión durante la infección temprana y la fuerte conectividad transfronteriza resaltan la necesidad de diagnóstico precoz y vigilancia molecular sostenida para interrumpir las redes activas de transmisión.

Palabras clave: VIH-1B; vigilancia genómica; Uruguay.

1.Introducción

1.1. VIH-1: panorama general

El reconocimiento del Virus de la Inmunodeficiencia Humana (VIH) como agente etiológico del síndrome de inmunodeficiencia adquirida (SIDA) fue el resultado de un proceso progresivo de caracterización clínica, epidemiológica y virológica iniciado a comienzos de la década de 1980. Los primeros casos posteriormente reconocidos como SIDA fueron reportados en Estados Unidos a comienzos de la década de 1980. En 1981 se describieron cinco casos de neumonía atípica por *Pneumocystis* en hospitales de Los Ángeles, todos ellos en hombres homosexuales previamente sanos, quienes además presentaban infecciones oportunistas como candidiasis mucocutánea y neoplasias como sarcoma de Kaposi (Gottlieb et al., 1981; Hymes et al., 1981).

Posteriormente, se notificaron casos similares en otros grupos poblacionales, incluidos usuarios de drogas intravenosas (Quagliarello et al., 1982), personas con hemofilia (CDC, 1982), receptores de transfusiones sanguíneas (CDC, 1982), niños (CDC, 1982), parejas heterosexuales de personas afectadas (CDC, 1983) y personas privadas de libertad (CDC, 1983), tanto en Estados Unidos como en Europa. Estos hallazgos permitieron caracterizar progresivamente el nuevo síndrome desde el punto de vista clínico y epidemiológico, y mostraron que el patrón de aparición de los casos era compatible con la transmisión de un agente infeccioso por vía sexual, sanguínea y perinatal. En este contexto, se consolidó la denominación SIDA, en referencia a un compromiso inmunológico severo que no era congénito, sino adquirido durante la vida.

En 1983, el grupo liderado por Luc Montagnier en el Instituto Pasteur de París logró aislar un retrovirus asociado al cuadro de inmunodeficiencia observado, a partir de células de un paciente con linfadenopatía, al que denominó virus asociado a linfadenopatía (LAV, por sus siglas en inglés) (Barre-Sinoussi & Chermann, 1983). Posteriormente, y tras la confirmación de que diferentes aislamientos virales correspondían al mismo agente, en 1986 se propuso unificar su denominación bajo el término VIH, en el marco del Comité Internacional de Taxonomía de Virus (ICTV) (Coffin et al., 1986; 2021). En la clasificación taxonómica vigente del ICTV, los virus denominados VIH-1 y VIH-2 se incluyen en las especies *Lentivirus humimdef1* y *Lentivirus humimdef2*, respectivamente, dentro del género *Lentivirus* (Zerbini et al., 2025).

Esta clasificación taxonómica aporta información relevante sobre sus características biológicas. El prefijo *retro*- alude al flujo inverso de la información genética característico de estos virus, en el cual el genoma de ARN es retrotranscrito a ADN bicatenario por acción de la transcriptasa reversa. Ese ADN viral puede luego integrarse en el genoma de la célula hospedadora, formando el provirus. Por su parte, el componente *lenti* (del latín *lentus*, lento), presente en la denominación del género *Lentivirus*, alude al curso clínico prolongado característico de estos virus, con un extenso período asintomático o una progresión lenta entre la infección inicial y la aparición de manifestaciones clínicas graves (Gifford et al., 2008).

1.2. Situación epidemiológica

1.2.1. Estado actual de la pandemia del VIH

En 2024, se estimó que aproximadamente 91,4 millones de infecciones por VIH [73,4-116,4 millones] se han producido desde el inicio de la pandemia. En ese mismo año, alrededor de 40,8 millones de personas [37,0-45,6] vivían con VIH y se registraron 630.000 muertes [490.000-820.000] relacionadas con SIDA.

Históricamente, África Subsahariana ha sido la región más afectada por la pandemia. En particular, en África oriental y África meridional concentran el 52% de las personas que viven con VIH a nivel global, lo que equivale a aproximadamente 21,1 millones de personas [18,9-23,0 millones]. Dentro de esta región, países como Sudáfrica y Mozambique presentan las mayores tasas de prevalencia.

En términos de impacto, le siguen Asia y el Pacífico, que concentran el 17% de las personas que viven con VIH a nivel global, con aproximadamente 6,9 millones de personas [4,9-7,2 millones], destacándose India como el país con mayor número de personas viviendo con VIH. Por su parte, África central y occidental concentra el 13% de las personas viviendo con VIH, con aproximadamente 5,2 millones [4,5-5,6 millones], destacándose la República Democrática del Congo (RDC) y Camerún como los países más afectados.

América Latina y el Caribe representan aproximadamente el 7% de las personas viviendo con VIH a nivel global, con alrededor de 2,8 millones [1,7–3,1 millones], siendo Brasil, México, Colombia y Argentina los países más afectados.

América del Norte y Europa central/occidental representan conjuntamente el 6% de las personas viviendo con VIH a nivel mundial, con aproximadamente 2,5 millones [1,9-2,6

millones], siendo Estados Unidos, Francia e Italia los países con mayor número de personas viviendo con VIH. Finalmente, Europa oriental y Asia central concentran el 5%, con aproximadamente 2,1 millones [1,7-2,1 millones], mientras que Medio Oriente y África del Norte representan el 0,6%, con aproximadamente 240.000 personas [150.000-210.000] viviendo con VIH (ONUSIDA, 2025. <http://aidsinfo.unaids.org/>).

Entre 2010 y 2024, el número de nuevas infecciones por VIH disminuyó de forma considerable en varias regiones del mundo, particularmente en África oriental y meridional (56%), África occidental y central (55%) y Asia y el Pacífico (17%). Sin embargo, esta tendencia no ha sido uniforme a nivel global. En algunas regiones, como Medio Oriente y África del Norte, el número de nuevas infecciones aumentó un 94% durante el mismo período (ONUSIDA, 2025; <http://aidsinfo.unaids.org/>).

En este contexto, América Latina se aparta de la tendencia mundial agregada de reducción de nuevas infecciones, con un aumento del 13% entre 2010 y 2024, mientras que el Caribe registró una reducción del 21% (ONUSIDA, 2025; <http://aidsinfo.unaids.org/>). Esta heterogeneidad en las tendencias epidemiológicas refleja, en parte, la influencia de factores estructurales (como la desigualdad, el estigma y la discriminación asociados al VIH y al SIDA) que continúan limitando el acceso al diagnóstico, la prevención y el tratamiento, y obstaculizan una respuesta sanitaria sostenida y equitativa frente a la pandemia (Conferencia Internacional sobre el SIDA–ONU, 2016).

1.2.2. Epidemia de VIH/SIDA en Uruguay

En Uruguay, el Ministerio de Salud Pública (MSP) estima que viven alrededor de 17.000 personas con VIH, es decir aproximadamente 0,6% del total de personas con VIH de América Latina y el Caribe en conjunto. En nuestro país, se registraron en promedio 900 nuevos casos de VIH anuales en el último quinquenio. Las tasas de notificación de nuevos diagnósticos en todas las franjas etarias son considerablemente más altas en el sexo masculino, con una relación de 2,4 hombres por cada mujer. Las personas de entre 25 y 44 años son el grupo etario con mayor tasa de notificación. A su vez, la vía sexual es la principal forma de transmisión. En relación con la transmisión vertical, el uso de terapia antirretroviral (TAR) y las políticas de salud pública han contribuido a que, durante el último quinquenio, esta tasa se haya mantenido por debajo del 2% (Situación Epidemiológica de VIH/SIDA en Uruguay, 2025).

Si bien en los últimos años la prevalencia de VIH en la población general se ha mantenido por debajo del 1%, la epidemia no se distribuye de manera homogénea entre los distintos

grupos poblacionales. Uruguay presenta un patrón epidemiológico concentrado en poblaciones clave, con prevalencias superiores al 5% en algunos grupos. Estas poblaciones incluyen mujeres trans, con prevalencias reportadas cercanas al 21% y valores de hasta 26% en estudios específicos que involucran particularmente a trabajadoras sexuales trans; hombres que tienen sexo con hombres (HSH), con prevalencias cercanas al 10%; y trabajadores sexuales masculinos, con valores alrededor del 8%.

Otras poblaciones también muestran valores elevados, como las personas usuarias de pasta base de cocaína y aquellas con antecedente de uso de drogas intravenosas (6%-10%), así como las personas privadas de libertad (2,4%). Por su parte, las trabajadoras sexuales mujeres presentan una prevalencia considerablemente menor (0,7%) dentro de las poblaciones clave relevadas, posiblemente asociada a estrategias de control sanitario. Estos datos evidencian la distribución heterogénea de la epidemia y subrayan la importancia de establecer estrategias de prevención, diagnóstico y tratamiento focalizadas en poblaciones de mayor vulnerabilidad (MSP, 2015).

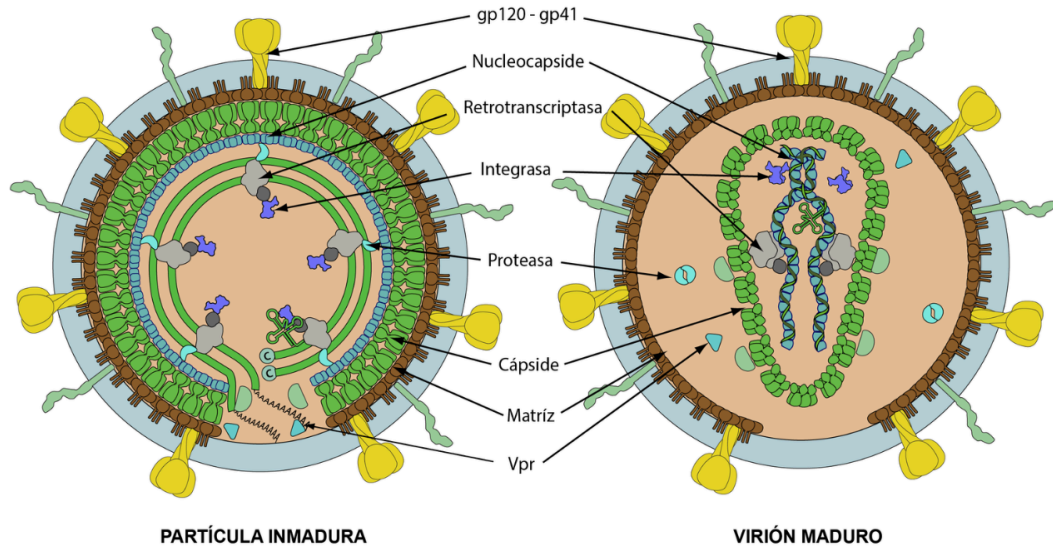
1.3. Biología del VIH

1.3.1. Estructura y genoma

La partícula viral madura del VIH presenta forma esférica con un diámetro de 80 a 120 nm. Está envuelta por una bicapa lipídica que se forma durante el proceso de gemación de la célula hospedera. En su superficie pueden observarse espículas formadas por glicoproteínas virales gp120, que intervienen en el reconocimiento del receptor celular, y por glicoproteínas transmembrana gp41, encargadas de la fusión de membranas (Knipe & Howley, 2013). Por debajo de la envoltura se encuentra la proteína de matriz p17 y, hacia el interior, la cápside viral interna (core) que presenta una estructura de cono truncado y está compuesta por la proteína p24. En el interior de esta estructura se aloja el genoma viral, asociado a proteínas de la nucleocápside que contribuyen a su organización y estabilidad.

El genoma del VIH está constituido por dos copias de ARN monocatenario de sentido positivo, idénticas y no complementarias, de aproximadamente 9,75 kb. En la partícula viral también se encuentran tres enzimas esenciales para la replicación: la retrotranscriptasa (RT), la integrasa (IN) y la proteasa (PR) (Carballal & Oubiña, 2001; Rosas et al., 2013) (Figura 1a).

a.



b.

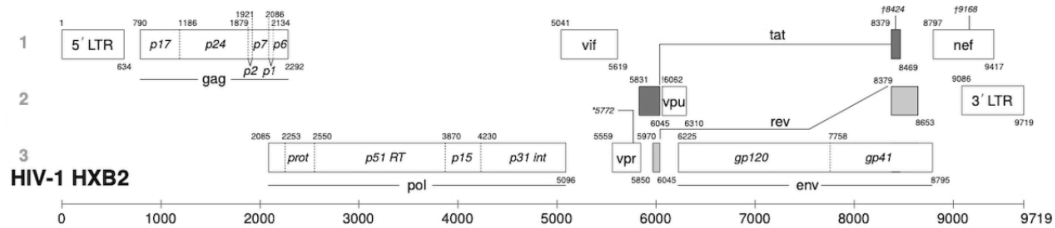


Figura 1. Estructura y organización genómica del VIH-1. a) Esquema comparativo de la partícula viral inmadura y del virión maduro de VIH-1, ilustrando los cambios estructurales asociados al proceso de maduración viral (adaptado de ViralZone, 2025). b) Mapa del genoma proviral de VIH-1, representado según la cepa de referencia HXB2 (extraído de Los Alamos HIV Sequence Database, <https://www.hiv.lanl.gov>).

A nivel genómico, el virus presenta tres regiones principales que codifican proteínas estructurales y enzimas necesarias para la infección viral productiva (*gag*, *pol* y *env*). La región *gag* codifica proteínas estructurales, incluyendo la proteína de matriz p17 y la proteína de la cápside p24; la región *env* codifica las glicoproteínas de la envoltura viral gp120 y gp41; mientras que la región *pol* codifica la RT, la IN y la PR virales. Adicionalmente, el genoma viral incluye genes regulatorios y accesorios, tales como *tat*, *rev*, *nef*, *vif*, *vpr* y *vpu*, que codifican proteínas que desempeñan funciones clave en la regulación de la replicación viral, la infectividad y la interacción con la célula hospedera. En estado de provirus, el genoma está flanqueado por secuencias largas repetidas (LTR, por sus siglas en inglés) que contienen regiones necesarias para la integración del ADN viral en el genoma de la célula hospedera y participan de la regulación de la transcripción viral (Figura 1b) (Knipe & Howley, 2013; Foley et al., 2016).

1.3.2. Ciclo replicativo

El ciclo replicativo del VIH comienza con la unión de la glicoproteína viral gp120 al receptor CD4 presente en la superficie de las células diana, principalmente linfocitos T CD4+, macrófagos y células dendríticas. Esta interacción induce cambios conformacionales que permiten la unión a los correceptores de quimiocinas CCR5 o CXCR4, lo que facilita la fusión de la envoltura viral con la membrana celular. Este proceso es mediado por la glicoproteína viral gp41 y permite la entrada del core viral al citoplasma celular.

Una vez en el interior de la célula, el ARN viral monocatenario es convertido en ADN de doble cadena mediante la acción de la enzima RT. Posteriormente, este ADN viral es transportado al núcleo e integrado en el genoma de la célula hospedadora por la enzima IN. A partir de este provirus integrado, la maquinaria celular transcribe nuevos ARN virales que sirven tanto como genomas virales como ARN mensajeros para la síntesis de proteínas virales. Estas proteínas y genomas se ensamblan en la membrana plasmática formando partículas virales inmaduras que, tras su liberación de la célula por gemación, maduran mediante la acción de la PR viral, dando lugar a viriones capaces de iniciar nuevos ciclos de infección (Knipe & Howley, 2021)

1.3.3. Dinámica de replicación, intervención terapéutica y resistencia

La infección por VIH se caracteriza por una fase aguda inicial, asociada a niveles elevados de viremia y a una disminución marcada de los linfocitos T CD4+. Aunque se desencadena una respuesta inmunitaria robusta, esta no logra erradicar el virus, lo que permite el establecimiento de una infección crónica persistente. Durante esta fase crónica, en ausencia de tratamiento, el virus continúa replicándose a niveles variables, lo que conduce a una progresiva depleción de linfocitos T CD4+. Tras un periodo asintomático de duración variable, esta pérdida sostenida de células inmunitarias culmina en el desarrollo del SIDA, definido por inmunosupresión avanzada y la aparición de infecciones oportunistas y neoplasias asociadas (Deeks et al., 2015).

El desarrollo de la TAR ha modificado significativamente la evolución clínica de la infección al intervenir de manera específica en distintas etapas del ciclo replicativo viral. Los diferentes grupos de antirretrovirales inhiben procesos como la entrada del virus a la célula, la retrotranscripción del ARN viral, la integración del ADN viral en el genoma celular y la maduración de nuevas partículas virales.

La utilización combinada de fármacos con mecanismos de acción complementarios permite una supresión sostenida de la replicación viral, lo que se traduce en una reducción significativa de la carga viral plasmática (frecuentemente hasta niveles indetectables por técnicas convencionales) y en una recuperación parcial del recuento de linfocitos T CD4+. En consecuencia, la infección por VIH ha pasado de ser una enfermedad inexorablemente progresiva y letal a una condición crónica controlable bajo tratamiento sostenido y adecuada adherencia terapéutica (Menendez-Arias, 2010; Sierra & Hauke, 2012).

Sin embargo, la elevada tasa de mutación del VIH y su intensa dinámica replicativa favorecen la emergencia de variantes virales con mutaciones asociadas a resistencia a antirretrovirales. Estas mutaciones pueden comprometer la eficacia de los tratamientos y limitar las opciones terapéuticas disponibles, constituyendo un desafío relevante para el control de la infección (Clavel et al., 2004). En este contexto, los estudios de genotipado de resistencia se han convertido en una herramienta clave para el manejo clínico de las personas con VIH y para la vigilancia de la resistencia a antirretrovirales (Günthard et al., 2019).

1.4. Mecanismos generadores de diversidad

La elevada variabilidad genética del VIH constituye un determinante clave de su capacidad adaptativa. Varios mecanismos generadores de diversidad hacen que el VIH sea una de las entidades virales de evolución más rápida estudiadas hasta el momento. Uno de ellos es su alta tasa de mutación debido a la baja fidelidad de su RT durante el proceso de replicación. Los retrovirus, así como otros virus ARN, presentan tasas de mutación elevadas debido a la falta de actividad correctora de errores de sus polimerasas virales. Particularmente, el VIH presenta una tasa de mutación de aproximadamente $3,4 \times 10^{-5}$ mutaciones/sitio/ciclo de replicación. Dado que su genoma tiene una longitud de 9,75 kb y que se producen aproximadamente 10^{10} viriones por día, podrían generarse millones de variantes en una persona infectada en tan solo un día (Morrondo et al., 2005; Taylor et al., 2008).

Otro mecanismo que contribuye significativamente a la amplia diversidad genética del VIH, es la recombinación. En este proceso se generan genomas recombinantes ya sea a partir de una infección secuencial (reinfección) o por una infección simultánea (coinfección) con linajes virales distintos. En estas condiciones, pueden generarse viriones que encapsidan dos genomas del mismo linaje (homodímeros) o genomas provenientes de linajes distintos (heterodímeros). Estos últimos, pueden originar genomas recombinantes durante la retrotranscripción, como producto de los saltos alternativos que realiza la RT entre ambos

genomas encapsidados. El mecanismo de recombinación genética media la reparación de genomas retrovirales defectivos, incrementa la diversidad viral y acelera la propagación de mutaciones ventajosas (Alexiev & Dimitrova, 2025; Onafuwa-Nuga & Telesnitsky, 2009).

La corta duración del ciclo replicativo del VIH constituye un factor que amplifica la generación de diversidad genética. Este ciclo puede completarse en aproximadamente 1–2 días, lo que permite numerosas rondas de replicación viral en un corto periodo de tiempo y, por lo tanto, múltiples oportunidades para la aparición de mutaciones y eventos de recombinación. Como consecuencia, la población viral dentro de un mismo hospedero puede diversificarse rápidamente, favoreciendo la selección de variantes con ventajas adaptativas frente a la respuesta inmune del hospedero o la presión selectiva ejercida por los fármacos antirretrovirales (Perelson et al., 1996; Santoro & Perno, 2013).

En conjunto, las altas tasas de mutación y recombinación, sumadas a los tiempos de replicación cortos, llevan a la producción de poblaciones altamente diversas. Este conjunto de variantes estrechamente relacionadas dentro de un mismo hospedero se denomina *cuasiespecie viral*. Estas poblaciones están compuestas por un espectro de genomas estrechamente relacionados, generados por mutación, recombinación y selección (Domingo et al., 1998; Schneider & Roossinck, 2001; Gartner et al., 2020). Esta alta diversidad le confiere al VIH la capacidad de adaptarse a la respuesta inmune del hospedero y adquirir resistencia a la TAR (Castelbranco et al., 2010; Delatorre et al., 2017).

1.5. Clasificación molecular

La elevada variabilidad genética del VIH ha dado lugar a una compleja clasificación molecular, que permite distinguir dos tipos principales, VIH-1 y VIH-2 (Figura 2). El VIH-1 se encuentra ampliamente distribuido a nivel mundial y se subdivide en cuatro grupos: M, N, O y P, cada uno originado a partir de eventos independientes de transmisión interespecie desde primates no humanos para humanos (eventos zoonóticos) (Sharp & Hahn, 2011).

Análisis filogenéticos indican que los grupos M y N surgieron a partir de linajes del Virus de la Inmunodeficiencia Simia (VIS) de chimpancés (*Pan troglodytes troglodytes*) que habitan el sureste y centro de Camerún, denominada VIScpz-Ptt (Gao et al., 1996). Por otra parte, los grupos O y P resultaron ser más cercanos a linajes de VISgor que infectan gorilas (*Gorilla gorilla*), principalmente de las regiones centro y sur de Camerún (Plantier et al., 2009; D'arc et al., 2015). Mientras que los grupos N, O y P se restringen principalmente al continente africano, el grupo M fue capaz de propagarse con éxito en la población humana,

siendo el responsable de la actual pandemia y de más del 99% de las infecciones por VIH a nivel global. Hasta el momento, el grupo M se subdivide en diez subtipos (A–D, F–H y J–L). Dentro de los subtipos A y F se han descrito diversos sub-subtipos (Knipe & Howley, 2013; Gartner et al., 2020). La divergencia genética dentro de un subtipo es de 15 a 20%, mientras que entre subtipos suele ser de 25 a 35% dependiendo del subtipo y la región del genoma analizada (Taylor et al., 2008).

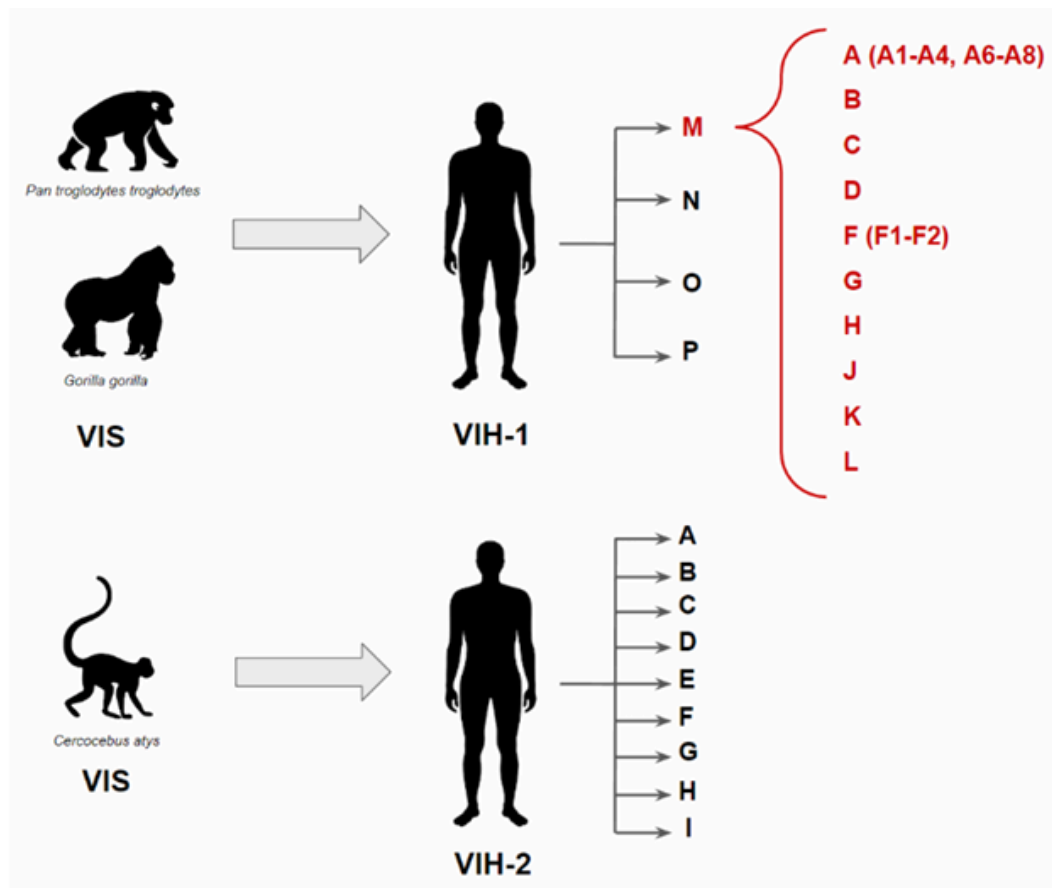


Figura 2. Origen y clasificación del VIH. Los distintos grupos de VIH-1 y VIH-2 son el resultado de al menos 13 eventos independientes de transmisión interespecie de linajes de VIS de primates no humanos a humanos (marcados con flechas). Se resalta en rojo el grupo M de VIH-1 y los distintos subtipos y sub-subtipos que lo componen.

Las formas recombinantes representan nuevos linajes con genoma en mosaico que surgen a partir de eventos de recombinación entre dos o más linajes parentales previamente identificados. Cuando se identifica un recombinante con el mismo patrón en al menos tres personas epidemiológicamente no vinculadas entre sí, se le denomina forma recombinante circulante (FRC; CRF, por sus siglas en inglés). A medida que se van descubriendo nuevas FRCs, se les asigna un número consecutivo seguido de los linajes parentales que la conforman (Ej. 02_AG, 12_BF, 31_BC) (Seitz, 2016).

La nomenclatura correspondiente se encuentra disponible en la base de datos Los Alamos HIV Database (<https://www.hiv.lanl.gov/components/sequence/HIV/crfsdb/crfs.com>), donde se listan las FRCs identificadas hasta la fecha. Por otro lado, los recombinantes entre subtipos o FRCs que se han encontrado en menos de tres individuos o en grupos de personas infectadas relacionadas epidemiológicamente, se denominan formas recombinantes únicas (FRU; URF, por sus siglas en inglés). Finalmente, cuando se trata de un recombinante con más de dos linajes parentales, se lo denomina “cpx” (complex, en inglés) (Taylor et al., 2008; Hemelaar et al., 2020). Hasta la fecha, se han identificado más de 170 FRCs.

El VIH-2 fue descubierto en 1985 y se considera endémico de la región de África Occidental. Hasta el momento, se han identificado nueve grupos de este tipo (A-I), cada uno asociado a la transmisión de linajes de VISmm provenientes de los monos mangabey (*Cercocebus atys*) a humanos (Reeves & Doms, 2002; Ayoub et al., 2013). El grupo A es el más frecuente, seguido por el grupo B. Debido a su naturaleza esporádica, los grupos C a H representan transmisiones aisladas, que se asocian a eventos zoonóticos esporádicos vinculados al contacto con monos mangabey (Visseaux et al., 2019; de Mendoza et al., 2021).

1.6. Distribución geográfica de los subtipos y recombinantes

La distribución geográfica de los subtipos y FRCs del VIH-1 identificados hasta el momento varía notablemente a nivel global (Figura 3). El subtipo C es el predominante y responsable de la mitad de las infecciones en el mundo (50,4%). Este subtipo presenta una marcada prevalencia en países del sur de África, India y Etiopía, regiones donde representa más del 97% de las infecciones. El subtipo A ocupa el segundo lugar en términos de prevalencia global (12,4%) y predomina en Europa oriental y Asia, donde representa más del 75,6% de los casos, así como en África occidental (51,7%). Por otro lado, el subtipo B es responsable del 11,3% de las infecciones a nivel mundial y predomina en el Caribe (91,0%), en América del Sur (76,1%) y Europa occidental y central junto con América del Norte (83,2%). Su prevalencia se ha mantenido relativamente estable durante los últimos 15 años, sin embargo, históricamente, siempre ha sido poco frecuente en el continente africano. A su vez, los subtipos G (2,9%) y D (2,6%), son prevalentes principalmente en África central y occidental. Los subtipos F, H, J, K y L presentan frecuencias globales muy bajas, cercanas o inferiores al 0,1% cada uno, y se concentran principalmente en el continente africano (Hemelaar et al., 2019; Nair et al., 2024).

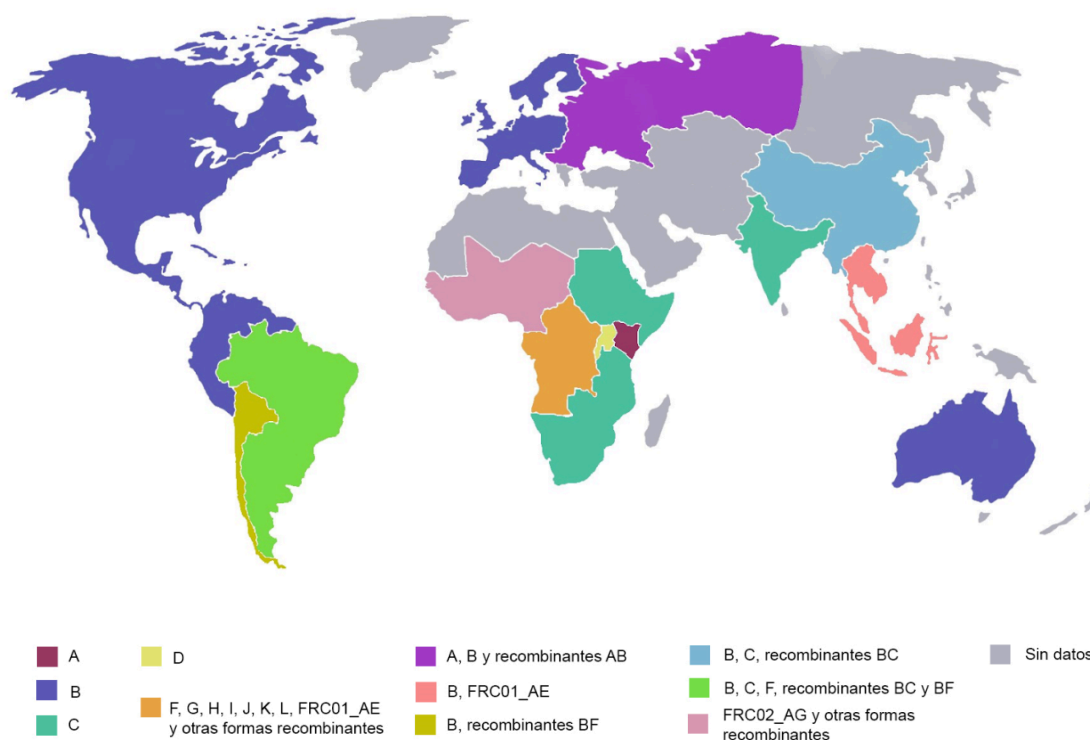


Figura 3. Distribución de los distintos subtipos y FRCs de VIH-1 prevalentes en el mundo (Abecasis & Vandamme, 2015; Bbosa et al., 2019; Hemelaar et al., 2019).

Por último, aproximadamente el 4% de las infecciones son causadas por linajes recombinantes (Hemelaar et al., 2019; Nair et al., 2024). A nivel mundial, las FRCs de mayor prevalencia son la 02_AG y la 01_AE, que predominan en África Central y Asia, respectivamente (Hemelaar et al., 2020). En el cono sur de América Latina, debido a la co-circulación principalmente de los subtipos B y C, y del sub-subtipo F1, han surgido numerosas FRCs (Abecasis & Vandamme, 2015). Entre las de mayor relevancia epidemiológica se destacan la 31_BC, la 28_BF1 y la 29_BF1, originadas en Brasil, así como la 12_BF descrita en Argentina (Bello et al., 2010; Gräf et al., 2021). Además, en esta región se han descrito otras FRCs, tales como la 39_BF1, 40_BF1, 71_BF1, 72_BF1, 81_cpx, 90_BF1 y 99_BF1 en Brasil, la 38_BF1 en Uruguay, la 44_BF1 en Chile y la 17_BF1 en Argentina, entre otras (Carr et al., 2001; Ríos et al., 2005; De Sá Filho et al., 2006; Santos et al., 2006; Guimarães et al., 2008; Ruchansky et al., 2009; Aulicino et al., 2012; Pessoa et al., 2014; Reis et al., 2017; 2019).

1.7. Subtipo B: origen y diseminación

El subtipo B del VIH-1 (VIH-1B) fue responsable de las primeras epidemias de gran escala establecidas fuera del continente africano, las cuales marcaron una etapa clave en la expansión global de la pandemia (Gilbert et al., 2007; Faria et al., 2014). Aunque representa aproximadamente el 11,3% de las infecciones a nivel mundial, constituye la variante de VIH con mayor dispersión geográfica y ha ocupado un lugar central como referencia en la investigación molecular, filogenética y farmacológica del virus (Junqueira & Almeida, 2016).

La evidencia histórica y filogenética sitúa el origen del subtipo B en África Central, particularmente en Kinshasa (actual RDC), a mediados del siglo XX. Posteriormente, se habría producido una única introducción en Haití entre 1962 y 1970. La hipótesis más aceptada vincula este evento con el retorno de trabajadores haitianos que residieron durante años en el antiguo Congo Belga (Pepin, 2011; Piot et al., 1984). Desde Haití se originaron dos procesos epidemiológicos y evolutivos diferenciados que estructuran la diversidad actual del VIH-1B (Gilbert et al., 2007; Faria et al., 2014; Worobey et al., 2016).

Por un lado, la dispersión regional temprana dentro del Caribe dio lugar a los clados denominados *subtipo B Caribeño* (B_{CAR}), derivados de la diversificación inicial del virus en la cuenca caribeña (Cabello et al., 2014). Estos clados dominan varias epidemias insulares, constituyendo un componente persistente de la transmisión local. Por otro lado, una introducción desde Haití hacia Estados Unidos entre 1966 y 1972 dio origen al clado denominado *subtipo B pandémico* (B_{PAN}), que posteriormente impulsó la expansión global del subtipo B hacia el resto del mundo (Robbins et al., 2003; Gilbert et al., 2007) (Figura 4).

Secuencias archivadas de VIH permitieron rastrear la diversificación temprana del clado B_{PAN} en Estados Unidos hasta la ciudad de Nueva York. El ancestro común más reciente fue estimado alrededor de 1972 (1970–1974). Esta reconstrucción temporal sugiere aproximadamente una década de circulación críptica del virus antes del reconocimiento clínico del VIH/SIDA en 1981 (Worobey et al., 2016). Durante los primeros años de la epidemia en Estados Unidos, aunque se registraron transmisiones heterosexuales (HTX), la expansión estuvo predominantemente asociada a redes de transmisión entre HSH. Se ha sugerido que la introducción inicial del virus desde Haití pudo haber estado vinculada al turismo sexual, dado que Haití constituía un destino frecuente para hombres homosexuales de Nueva York en ese período. La intensa movilidad internacional y los flujos migratorios hacia y desde Estados Unidos favorecieron la rápida diseminación internacional del clado B_{PAN} , consolidando su papel como linaje pandémico (Pepin 2011; Junqueira et al., 2011).

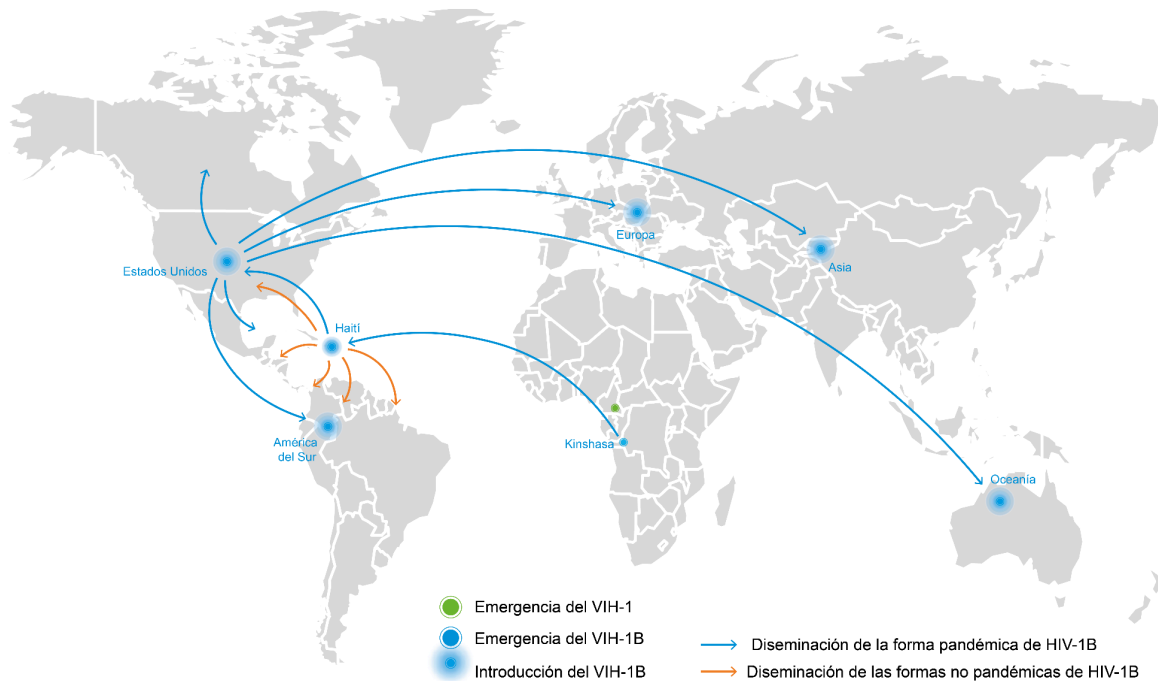


Figura 4. Origen y diseminación geográfica del VIH-1B en el mundo. Las líneas azules representan las principales rutas de dispersión de la forma pandémica del subtipo B del VIH-1 y los principales epicentros de la enfermedad. Las líneas naranjas representan la dirección de la dispersión del linaje B no pandémico (*extraído y modificado de Junqueira & Almeida, 2016*).

En América Latina, el clado B_{PAN} constituye el linaje predominante dentro del subtipo B. Este patrón es particularmente marcado en América del Sur, donde B_{PAN} representa más del 90% de las infecciones por VIH-1B en la mayoría de los países estudiados (Cabello et al., 2015; Worobey et al., 2016). En América Central, se ha propuesto que una única introducción de B_{PAN} habría dado origen a aproximadamente el 62% de las infecciones actuales por VIH-1B (Murillo et al., 2013). Los clados B_{CAR} , en cambio, predominan en varias islas del Caribe, pero también han sido detectados en la mayoría de los países de América Latina, generalmente con baja frecuencia relativa (1–10%) (Junqueira et al., 2011; Cabello et al., 2015; Arantes et al., 2019) (Figura 5).

Tras su establecimiento en la región, el clado B_{PAN} dio lugar a múltiples linajes de transmisión en América Latina, algunos de alcance regional y otros restringidos a países específicos. En un estudio filodinámico basado en 6789 secuencias del gen *pol* de VIH-1B provenientes de siete países latinoamericanos, se identificaron 12 clados principales del linaje B_{PAN} que, en conjunto, representaron aproximadamente el 36% de las secuencias analizadas de la región (Mir et al., 2015). Entre ellos se reconocieron dos linajes regionales principales: el Sudamericano ($B_{PAN-SAM}$), integrado por secuencias de Argentina y Brasil, y el

Centroamericano ($B_{\text{PAN-CAM}}$), integrado por secuencias de Honduras y El Salvador. Además, se identificaron clados predominantemente específicos de país en Argentina, Brasil, México y Perú. La mayoría de estos clados se habría originado entre finales de la década de 1960 y fines de la década de 1980, con una fase inicial de crecimiento exponencial seguida de una desaceleración reciente de su expansión. Estos resultados sugieren que una proporción importante de las infecciones por VIH-1B en América Latina derivó de la expansión de un número limitado de linajes fundadores, en un contexto de transmisión predominantemente intra-país y entre países vecinos, con limitada mezcla regional (Mir et al., 2015).

Dentro de este marco regional, la historia evolutiva del VIH-1B en Uruguay permanece en gran medida inexplorada. La vigilancia molecular a nivel nacional indica que el VIH-1B predomina en Uruguay (39,3%) y está fuertemente asociado con transmisión entre HSH (Podchibiakin et al., en revisión, Anexo 1). Sin embargo, Uruguay no estuvo representado en los estudios regionales previos sobre la diversificación del VIH-1B en América Latina, debido a la limitada disponibilidad de secuencias nacionales en ese momento (Cabello et al., 2015; Mir et al., 2015). En consecuencia, la contribución relativa de los clados B_{PAN} y B_{CAR} a la epidemia nacional, su frecuencia de introducción y la persistencia a largo plazo de las redes de transmisión local aún no han sido caracterizadas.

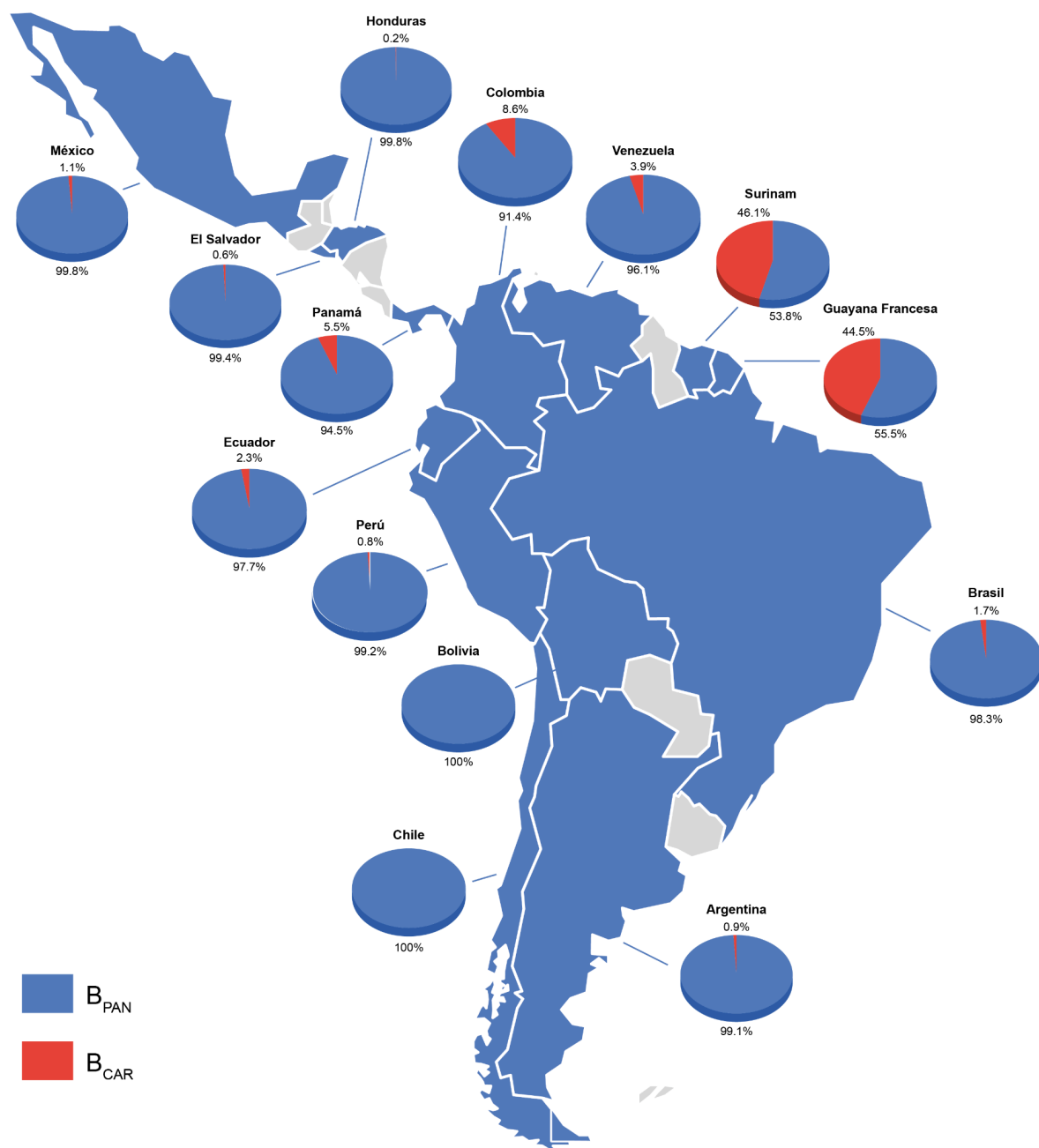


Figura 5. Proporción estimada de los clados B_{CAR} (rojo) y B_{PAN} (azul) de VIH-1B en distintos países de América Latina. Los países en azul representan aquellos con datos disponibles de secuencias de VIH-1B, mientras que los países en gris indican ausencia de datos publicados hasta 2015 (extraído y modificado de *Cabello et al., 2015*).

1.8. Herramientas y marco conceptual

1.8.1. Filogenia molecular

La filogenia molecular es una disciplina que utiliza secuencias genéticas para inferir relaciones evolutivas entre organismos, genes o poblaciones. Su fundamento es que las secuencias acumulan cambios a lo largo del tiempo como resultado de procesos evolutivos tales como mutación, recombinación, selección y deriva génica. A partir de la comparación de estas secuencias es posible reconstruir hipótesis sobre su historia evolutiva, usualmente representadas en forma de árboles filogenéticos. En este contexto, las diferencias observadas entre secuencias se interpretan como el resultado de cambios evolutivos acumulados a lo largo del tiempo, modelados principalmente como sustituciones nucleotídicas, bajo un determinado modelo de evolución molecular (Lemey et al., 2009).

1.8.2. Modelos probabilísticos de sustitución nucleotídica

La reconstrucción filogenética requiere un modelo explícito que describa cómo las secuencias cambian a lo largo del tiempo. En el marco de la inferencia filogenética moderna, el proceso de sustitución nucleotídica suele modelarse como una cadena de Markov en tiempo continuo, donde las sustituciones entre nucleótidos (A, C, G y T) ocurren según tasas definidas por una matriz de tasas instantáneas Q . Esta matriz resume el proceso evolutivo subyacente y permite calcular la probabilidad de observar las secuencias actuales dado un árbol filogenético y un conjunto de parámetros (Felsenstein, 2004).

Se han propuesto diversos modelos de sustitución, que difieren principalmente en el número de parámetros y supuestos utilizados para describir el proceso evolutivo. El modelo más simple propuesto es el modelo de Jukes y Cantor (JC69), que asume que los cuatro nucleótidos presentan frecuencias de equilibrio iguales y que todas las sustituciones ocurren con la misma tasa (Jukes & Cantor, 1969). Posteriormente, el modelo de Kimura de dos parámetros (K80) introdujo una distinción biológicamente más realista al reconocer que las transiciones suelen ocurrir con mayor frecuencia que las transversiones (Kimura, 1980). Modelos más complejos permiten estimar frecuencias nucleotídicas desiguales y tasas diferenciadas entre tipos de sustitución, ofreciendo representaciones más realistas del proceso evolutivo. Entre ellos se encuentran el modelo de Hasegawa-Kishino-Yano (HKY85) (Hasegawa et al., 1985) y el modelo de Tamura-Nei (TN93) (Tamura & Nei, 1993). Por último, el modelo GTR constituye la formulación reversible más general, ya que estima

libremente las frecuencias de equilibrio y permite seis tasas independientes para los distintos tipos de sustitución entre pares de nucleótidos (Tavare, 1986).

Además de las tasas relativas de sustitución, un aspecto fundamental del proceso evolutivo es la heterogeneidad en la tasa evolutiva entre sitios. Aunque los modelos más simples asumen que todas las posiciones del alineamiento evolucionan a la misma velocidad, en la mayoría de los genes coexisten regiones altamente conservadas (debido a restricciones estructurales o funcionales) con posiciones que acumulan cambios más rápidamente. Esta heterogeneidad puede modelarse mediante una distribución Gamma ($+\Gamma$), que asume que la tasa de sustitución de cada sitio es una variable aleatoria extraída de dicha distribución. Su parámetro de forma (α) determina el grado de variación entre sitios: valores bajos indican mayor heterogeneidad, con muchos sitios de evolución lenta y unos pocos sitios de evolución rápida. De manera complementaria, puede incorporarse un parámetro de sitios invariantes ($+I$), que asume que una fracción de las posiciones presenta una tasa de cambio igual a cero a lo largo del período evolutivo considerado (Yang, 1994; 1996). Por ello, la elección del modelo de sustitución es un paso crítico en los análisis filogenéticos, ya que modelos demasiado simplificados pueden sesgar las estimaciones de topología, longitudes de rama y tiempos de divergencia.

1.8.3. Métodos de inferencia filogenética

En el marco de la inferencia filogenética, la topología del árbol, las longitudes de rama y los parámetros que describen el proceso de sustitución nucleotídica se estiman bajo un modelo evolutivo especificado. La elección del marco inferencial determina cómo se explora el espacio de árboles y parámetros, cómo se estiman estos componentes y cómo se cuantifica la incertidumbre asociada a las estimaciones (Lemey et al., 2009).

Los métodos de inferencia filogenética pueden clasificarse en dos grandes enfoques: métodos basados en distancias y métodos basados en caracteres. Los métodos basados en distancias transforman la información del alineamiento en una matriz de distancias genéticas entre pares de secuencias, usualmente corregidas mediante modelos de sustitución nucleotídica, y a partir de esta matriz construyen un árbol que resume dichas distancias. En este enfoque, el modelo evolutivo se utiliza de manera indirecta, únicamente para estimar las distancias, y no forma parte del proceso de inferencia del árbol. En contraste, los métodos basados en caracteres utilizan directamente la información sitio a sitio del alineamiento y modelan explícitamente el proceso evolutivo mediante un modelo de sustitución nucleotídica. En estos enfoques, la topología del árbol, las longitudes de rama y

los parámetros del modelo se estiman de manera conjunta, evaluando la probabilidad de los datos bajo diferentes hipótesis filogenéticas. Dentro de este grupo se incluyen los métodos de máxima verosimilitud y los enfoques bayesianos, que difieren en su fundamento estadístico, complejidad computacional y forma de cuantificar la incertidumbre.

El método Neighbor-Joining (NJ), propuesto por Saitou y Nei (1987), es un algoritmo basado en distancias que construye un árbol no enraizado a partir de una matriz de distancias genéticas derivada del alineamiento, usualmente calculada bajo un modelo de sustitución nucleotídica. Su objetivo es identificar iterativamente pares de secuencias que minimicen la longitud total del árbol, siguiendo el principio de mínima evolución (Saitou & Nei, 1987). Al resumir la información del alineamiento en una matriz de distancias, el método reduce la complejidad computacional, sin embargo, esta reducción de la información puede traducirse en una menor precisión al inferir relaciones filogenéticas finas (Robertson, 1995).

El método de máxima verosimilitud (ML, por sus siglas en inglés) se basa en un marco estadístico explícito en el cual se estima la topología y el conjunto de parámetros que maximizan la probabilidad de observar el alineamiento dado un modelo de sustitución nucleotídica (Felsenstein, 1981). La función de verosimilitud puede expresarse formalmente como:

$$L(g, \theta) = P(D|g, \theta)$$

Donde D representa el alineamiento de secuencias, g la topología del árbol filogenético con sus longitudes de rama y θ representa los parámetros del modelo evolutivo. El objetivo es encontrar la combinación de estos elementos que maximice la verosimilitud. A diferencia de los métodos basados en distancias, ML utiliza directamente la información sitio por sitio del alineamiento y modela el proceso evolutivo mediante una matriz de tasas definida, permitiendo la estimación conjunta de la topología, las longitudes de rama y los parámetros del modelo. Dado que el número de topologías posibles crece de manera exponencial con el número de secuencias, una búsqueda exhaustiva del árbol óptimo resulta computacionalmente inviable para conjuntos de datos moderados o grandes. Por ello, los programas de ML implementan algoritmos heurísticos que exploran el espacio de árboles mediante operaciones locales como *Nearest Neighbor Interchange* (NNI), *Subtree Pruning and Regrafting* (SPR) o *Tree Bisection and Reconnection* (TBR), generando iterativamente nuevas topologías cuya verosimilitud es evaluada hasta alcanzar un óptimo local (Lemey et al., 2009).

Por su parte, la inferencia bayesiana constituye un marco probabilístico en el cual se estima la distribución posterior de árboles y parámetros evolutivos dado el alineamiento observado. A diferencia de los métodos de máxima verosimilitud, que buscan una única solución óptima, este enfoque permite obtener una distribución de soluciones ponderadas por su probabilidad posterior, incorporando explícitamente la incertidumbre asociada al proceso evolutivo. Esta distribución puede expresarse formalmente mediante el teorema de Bayes:

$$P(g, \theta|D) \propto P(D|g, \theta) P(g, \theta)$$

donde $P(g, \theta|D)$ representa la distribución posterior de la topología del árbol y los parámetros del modelo dado el alineamiento D , $P(D|g, \theta)$ corresponde a la verosimilitud de los datos, y $P(g, \theta)$ representa la distribución a priori sobre árboles y parámetros.

En este marco, la probabilidad posterior resulta de la combinación entre la evidencia aportada por los datos y el conocimiento previo incorporado mediante las distribuciones a priori, permitiendo una estimación conjunta de la genealogía y los parámetros evolutivos. La verosimilitud constituye el mismo componente utilizado en inferencia por máxima verosimilitud, pero es integrada junto con la información a priori para estimar la distribución completa de soluciones posibles.

Los parámetros evolutivos se tratan como variables aleatorias con distribuciones a priori, y el objetivo es aproximar la distribución posterior conjunta de topologías y parámetros. Dado el enorme espacio de topologías y combinaciones de parámetros posibles, el cálculo exacto de la distribución posterior es computacionalmente intratable. Por ello, la exploración del espacio de árboles se realiza mediante algoritmos de Monte Carlo por cadenas de Markov (MCMC), comúnmente implementados a través del algoritmo de Metropolis–Hastings. Este procedimiento genera una cadena estocástica que converge hacia la distribución posterior, realizando modificaciones aleatorias en la estructura del árbol o en los valores de los parámetros. Cada propuesta es aceptada o rechazada de acuerdo con una probabilidad de aceptación que depende de la razón entre las probabilidades posteriores de los estados propuesto y actual (Huelsenbeck et al., 2001).

1.8.4. Evaluación del soporte estadístico

Dado que las inferencias se basan en datos finitos y modelos aproximados, es fundamental cuantificar la incertidumbre asociada a los agrupamientos inferidos en el árbol.

En métodos como NJ y ML, la estrategia más utilizada de evaluación ha sido el bootstrap (Felsenstein, 1985), que consiste en generar múltiples pseudo-réplicas del alineamiento mediante remuestreo con reemplazo y reconstruir un árbol para cada una. El porcentaje de réplicas en las que aparece un clado determinado refleja su robustez frente a la variabilidad de los datos.

Debido al costo computacional del bootstrap bajo modelos complejos, se han desarrollado alternativas más eficientes, como el approximate Likelihood Ratio Test (aLRT) (Anisimova & Gascuel, 2006), que evalúa el soporte de cada rama comparando la verosimilitud del árbol óptimo con la de topologías alternativas locales. Una variante aún más conservadora es el SH-aLRT, basado en el SH-test (Shimodaira & Hasegawa, 1999), que incorpora correcciones para reducir la sobreestimación del soporte.

En el marco bayesiano, el soporte se evalúa como la probabilidad posterior del clado, calculada como la proporción de árboles muestreados por MCMC en los que la partición está presente (Huelsenbeck et al., 2001). A diferencia del bootstrap, este valor representa directamente la probabilidad del clado bajo el modelo y los datos observados.

1.8.5. Reloj molecular

El concepto de reloj molecular, propuesto por Zuckerkandl y Pauling en la década de 1960, establece que las diferencias genéticas entre secuencias se acumulan aproximadamente de manera proporcional al tiempo transcurrido desde que dos linajes divergieron a partir de un ancestro común. Bajo este principio, la cantidad de sustituciones observadas entre secuencias puede utilizarse como una medida del tiempo evolutivo, permitiendo estimar fechas de divergencia cuando la tasa de sustitución es conocida o puede ser calibrada mediante información temporal (Zuckerkandl & Pauling, 1962; 1965).

En el contexto de la inferencia filogenética, este principio permite incorporar explícitamente la dimensión temporal en los árboles evolutivos mediante modelos que vinculan la acumulación de sustituciones nucleotídicas con el tiempo. En su formulación más simple, el *reloj molecular estricto* asume que todas las ramas del árbol evolucionan a una misma tasa de sustitución. Bajo este modelo, la longitud de cada rama es proporcional al tiempo transcurrido, lo que permite inferir fechas de divergencia cuando se dispone de calibraciones temporales adecuadas, como fechas de muestreo o información externa.

Sin embargo, en muchos sistemas biológicos (particularmente en virus de ARN) la tasa evolutiva puede variar entre linajes debido a diferencias en presiones selectivas, dinámica de transmisión o contextos epidemiológicos. Para capturar esta heterogeneidad, se han

desarrollado modelos de *reloj relajado*, que permiten que cada rama del árbol tenga su propia tasa de sustitución, generalmente modelada como una variable aleatoria extraída de una distribución subyacente (Drummond, 2006).

En análisis de virus como el VIH-1, el uso de un reloj relajado resulta frecuentemente más apropiado que un reloj estricto, ya que distintos clados pueden experimentar variaciones en su tasa evolutiva asociadas a diferencias en la dinámica de transmisión, las redes epidemiológicas o la presión terapéutica (Drummond et al., 2006; Lemey et al., 2009; Pybus & Rambaut, 2009). La elección del modelo de reloj impacta directamente en la estimación de tiempos de divergencia, en la datación de eventos fundadores y en la reconstrucción de cambios demográficos mediante modelos coalescentes o modelos de nacimiento-muerte (birth-death).

La adecuación del modelo de reloj puede evaluarse mediante la presencia de señal temporal en los datos, por ejemplo, a través de la relación entre divergencia genética y fecha de muestreo. Cuando esta relación es débil o inexistente, las estimaciones temporales pueden ser poco informativas o depender fuertemente de los supuestos del modelo.

1.8.6. Filogeografía

La filogeografía estudia la distribución geográfica de linajes y poblaciones en relación con su historia evolutiva, con el objetivo de reconstruir patrones históricos de diversificación, dispersión y estructuración espacial. En su aplicación molecular, este enfoque combina información genética, geográfica y, frecuentemente temporal, para inferir la distribución espacial de los ancestros, y reconstruir la dinámica de dispersión entre regiones a lo largo del tiempo (Knowles & Maddison, 2002).

Desde el punto de vista metodológico, los enfoques filogeográficos aplicados a datos moleculares pueden modelarse principalmente bajo dos tipos de modelos: discretos y continuos, que difieren en la forma en que representan el espacio geográfico y en los supuestos sobre el movimiento espacial.

En los modelos discretos, las localizaciones geográficas se tratan como estados categóricos (regiones, países, estados o ciudades) y la evolución espacial se modela como un proceso de Markov en tiempo continuo (CTMC, por sus siglas en inglés). En este marco, los cambios entre estados geográficos a lo largo de las ramas del árbol están definidos por una matriz de tasas de migración entre localizaciones geográficas. Este enfoque permite estimar la

localización más probable del ancestro común más reciente (MRCA, por sus siglas en inglés) y reconstruir las transiciones espaciales más probables.

Por su parte, los modelos continuos representan la localización geográfica mediante coordenadas espaciales (usualmente latitud y longitud) y modelan la dispersión como un proceso de difusión en el espacio, generalmente bajo esquemas de caminata aleatoria relajada (*relaxed random walk*). Este enfoque es particularmente útil para estimar velocidades de dispersión cuando se dispone de datos geográficos densos y precisos (Lemey et al., 2006).

En el estudio de la epidemiología molecular de virus como el VIH, la elección del modelo filogeográfico depende de la resolución de los datos y la naturaleza del proceso de transmisión. Suele optarse por el uso de modelos discretos. Esto se debe, por un lado, a que los metadatos disponibles suelen estar restringidos a estados o países, y a que los procesos de transmisión humana suelen estructurarse en redes de movilidad y contacto entre localizaciones discretas más que como procesos continuos de difusión espacial (Lemey et al., 2009; Pybus & Rambaut, 2009).

Dentro de este marco, la reconstrucción de estados ancestrales pueden realizarse mediante enfoques de máxima verosimilitud (como los implementados en herramientas como PastML [Ishikawa et al., 2019]) o mediante marcos bayesianos, en los cuales los estados ancestrales se estiman a partir de distribuciones posteriores obtenidas mediante MCMC. En ambos casos, los resultados suelen resumirse identificando el estado geográfico con mayor probabilidad para cada nodo, por ejemplo, mediante criterios de máximo a posteriori (MAP).

Este tipo de análisis ha sido ampliamente utilizado para estudiar la propagación regional y global del VIH-1, identificando eventos de introducción entre países, rutas de dispersión y la expansión de clados epidémicos. Estudios sobre la expansión temprana del VIH-1 han demostrado cómo la integración de información temporal y espacial permite reconstruir eventos fundadores y fases de expansión epidémica, estableciendo un puente entre evolución molecular y epidemiología (Delatorre et al., 2013; Faria et al., 2014; Cabello et al., 2014; Mir et al., 2015; Arantes et al., 2019).

Sin embargo, uno de los principales desafíos de los análisis filogeográficos es el sesgo de muestreo. Cuando algunas regiones están sobrerrepresentadas en el conjunto de datos, pueden aparecer artificialmente como fuentes más probables de dispersión simplemente debido a su mayor tamaño muestral. Para mitigar este efecto, es habitual aplicar estrategias

de submuestreo estratificado o equilibrar el número de secuencias por región (Liu et al., 2022).

1.8.7. Filodinámica

La filodinámica constituye un marco conceptual que integra evolución, dinámica poblacional y epidemiología a partir de la información contenida en genealogías inferidas a partir de datos moleculares (Grenfell et al., 2004). En este marco, el árbol filogenético no se interpreta únicamente como una representación de parentesco, sino como la realización estocástica de un proceso de transmisión y crecimiento poblacional subyacente. En consecuencia, la topología del árbol, la distribución temporal de los nodos y las longitudes de las ramas contienen información sobre la dinámica histórica de la epidemia, tales como cambios en la tasa de transmisión, expansiones poblacionales o fases de estabilización (Pybus & Rambaut, 2009).

Teoría coalescente

La teoría coalescente, formalizada por Kingman (Kingman, 1982a; 1982b) y posteriormente desarrollada por Tavaré (Tavaré, 1986) y Hudson (Hudson, 1990), proporciona un marco probabilístico retrospectivo para describir la genealogía de una muestra. Bajo este modelo, los linajes muestreados se rastrean hacia atrás en el tiempo hasta converger en ancestros comunes. Cada evento de coalescencia representa la unión de dos linajes en un ancestro común y, si se retrocede lo suficiente, se alcanza el MRCA de la muestra (Figura 6).

En este contexto, la estructura temporal de la genealogía está directamente relacionada con la dinámica poblacional subyacente. En particular, la tasa de coalescencia es inversamente proporcional al tamaño efectivo poblacional (N_e), lo que puede expresarse como:

$$\lambda(t) \propto \frac{1}{N_e(t)}$$

Por lo tanto, en poblaciones grandes los eventos coalescentes tenderán a ubicarse más alejados en el tiempo, mientras que en poblaciones pequeñas o en fases de crecimiento reciente se concentran hacia el presente. Esta relación permite inferir cambios en la dinámica poblacional a partir de la distribución de los tiempos de coalescencia a lo largo del árbol (Hudson, 1990; Pybus & Rambaut, 2009).

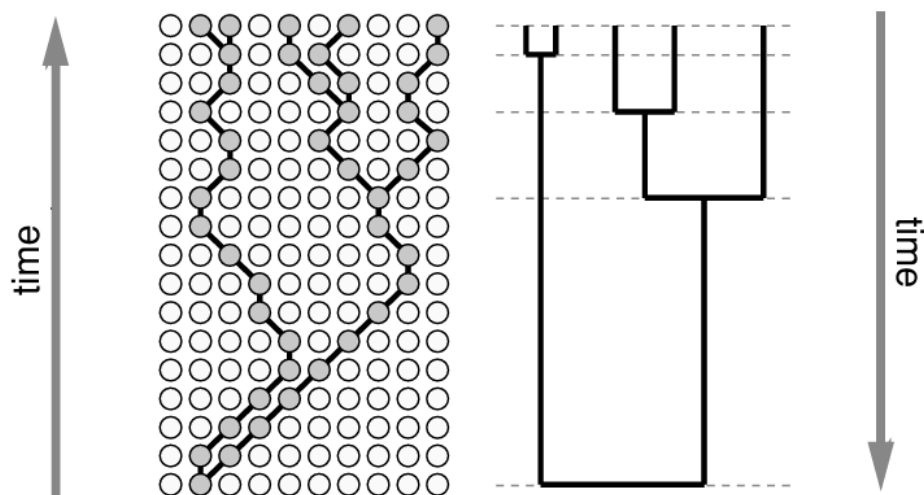


Figura 6. Representación de la teoría coalescente. A la izquierda, se muestra una genealogía de cinco secuencias, muestreadas en el presente a partir de una población de tamaño constante de 10 individuos. A la derecha, los linajes en la genealogía se representan como líneas continuas (extraído de *The Phylogenetic Handbook*–2da edición, 2009).

A partir de este marco se han desarrollado distintos modelos demográficos para reconstruir cambios temporales en N_e , con distintos grados de flexibilidad y diferentes supuestos sobre la dinámica poblacional. Los modelos paramétricos asumen una forma matemática específica para describir cómo cambia el N_e en el tiempo. El modelo de tamaño constante sirve como referencia teórica, aunque raramente describe adecuadamente epidemias virales. El modelo de crecimiento exponencial resulta apropiado para fases tempranas de expansión epidémica, mientras que el modelo logístico incorpora una capacidad de carga, permitiendo capturar escenarios donde la epidemia experimenta una fase de crecimiento inicial seguida de estabilización.

Cuando no se desea imponer una trayectoria demográfica rígida, se utilizan modelos no paramétricos que permiten reconstruir la historia demográfica de manera más flexible. Entre ellos, los modelos tipo skyline estiman cambios temporales en N_e a partir de la distribución de los eventos coalescentes en la genealogía. El Bayesian Skyline Plot (BSP) estima cambios escalonados en N_e a partir de los tiempos de coalescencia (Drummond et al., 2005). Extensiones posteriores, como Skyride y Skygrid, introducen esquemas de suavizado que mejoran la estabilidad y resolución temporal de las estimaciones (Minin et al., 2008; Gill et al., 2013). En particular, el modelo BSKG (Bayesian Skygrid) permite estimar el N_e en una grilla de intervalos temporales predefinidos, reduciendo la dependencia de particiones arbitrarias y facilitando el análisis de conjuntos de datos grandes o densamente muestreados.

En estudios filodinámicos del VIH-1, estos enfoques han permitido reconstruir fases de expansión en décadas tempranas, estabilización posterior a la implementación de la TAR y cambios asociados a transformaciones en las redes de transmisión (Pybus & Rambaut, 2009; Faria et al., 2014; Arantes et al., 2022). De este modo, el marco coalescente constituye el vínculo formal entre la genealogía viral y la dinámica epidemiológica subyacente.

Modelos Birth–Death

A diferencia de la teoría coalescente, que describe retrospectivamente la genealogía condicionada a un proceso demográfico, los modelos *birth-death* modelan explícitamente el proceso de transmisión hacia adelante en el tiempo, estableciendo un vínculo directo entre la filogenia y los procesos epidemiológicos subyacentes. En este marco, la tasa de “birth” (λ) corresponde a la generación de nuevas infecciones, la tasa de “death” (μ) a la pérdida de infectividad (ya sea por tratamiento, muerte o cambios conductuales) y se incorpora además una tasa de muestreo (ψ) que refleja la probabilidad de que una infección sea incluida en el conjunto de datos (Nee et al., 1994).

Esto permite interpretar la genealogía en términos directamente vinculados a la dinámica de transmisión, ya que esta se modela como el resultado de un proceso estocástico de infección, remoción y muestreo. En este contexto, uno de los parámetros clave es el número reproductivo efectivo (Re), que describe el número promedio de infecciones secundarias generadas por un individuo infectado en un momento dado y puede expresarse como:

$$Re = \frac{\lambda}{(\mu + \psi)}$$

Una extensión particularmente relevante de este marco es el modelo *birth–death skyline* (BDSKY), que permite estimar la variación temporal de estos parámetros en intervalos discretos, sin asumir que la dinámica epidemiológica permanece constante a lo largo del tiempo (Stadler et al., 2013; Stadler et al., 2014). En particular, BDSKY permite reconstruir cambios temporales en Re , proporcionando una aproximación más directa a la dinámica de transmisión de la epidemia. Este enfoque resulta especialmente adecuado para el estudio de epidemias virales donde factores como la introducción de intervenciones sanitarias, cambios en las redes de transmisión o la aparición de nuevos clados pueden generar variaciones sustanciales en la dinámica de propagación.

Desde el punto de vista epidemiológico, la estimación de $Re(t)$, permite caracterizar distintas fases de la epidemia: valores de Re mayores a uno indican una expansión activa del

número de infecciones, mientras que valores menores a uno reflejan una disminución en la transmisión. A diferencia de N_e , que no representa directamente el número de individuos infectados sino una medida relacionada con la estructura genealógica y la diversidad genética de la población viral, R_e proporciona una medida más directa de la intensidad de transmisión en el tiempo, facilitando la interpretación de los procesos epidemiológicos subyacentes (Volz & Frost, 2014).

Los enfoques coalescentes y birth–death proporcionan perspectivas complementarias sobre la dinámica epidémica: mientras los primeros permiten inferir cambios históricos en la diversidad genética, los segundos modelan de forma más explícita los procesos de transmisión y muestreo.

En el contexto del VIH-1, los modelos *birth-death* han sido ampliamente utilizados para evaluar la dinámica de transmisión y el impacto de intervenciones como la TAR, así como para identificar diferencias en la propagación entre distintos clados o grupos de riesgo (Stadler et al., 2013; Volz et al., 2013; Arantes et al., 2022).

1.9. Justificación del estudio

Debido a su rápida evolución y la disponibilidad de secuencias virales muestreadas a lo largo del tiempo, el VIH-1 constituye un sistema particularmente adecuado para el estudio de epidemias mediante enfoques filogenéticos y filodinámicos (Grenfell et al., 2004; Pybus & Rambaut, 2009). En particular, el VIH-1B presenta una historia evolutiva compleja caracterizada por múltiples eventos de introducción, expansión y diversificación en distintos países de América Latina. Estudios realizados en otros países de la región han permitido identificar redes de transmisión, reconstruir patrones históricos de dispersión y caracterizar la dinámica temporal de distintos linajes virales mediante análisis filogenéticos, filogeográficos y filodinámicos (Mir et al., 2015; Cabello et al., 2015; Arantes et al., 2019).

Sin embargo, a pesar de que el VIH-1B es una de las variantes de mayor prevalencia en Uruguay, la historia evolutiva y epidemiológica de este linaje en el país permanece escasamente caracterizada. En consecuencia, aún existen vacíos de conocimiento sobre las introducciones independientes de VIH-1B en Uruguay, el origen geográfico probable de los linajes circulantes, la composición de los clados de transmisión local y su dinámica evolutiva y demográfica.

Desde el año 2007, la creciente disponibilidad de secuencias virales generadas a partir de estudios de genotipado de resistencia representa una oportunidad valiosa para abordar

estas cuestiones desde una perspectiva evolutiva y epidemiológica integrada. En este contexto, el presente trabajo busca caracterizar la dinámica evolutiva del VIH-1B en Uruguay mediante el análisis de secuencias virales caracterizadas previamente por nuestro grupo de investigación (Podchibiakin et al., en revisión, Anexo 1) y la aplicación de herramientas de inferencia filogenética, filogeográfica y filodinámica. Se espera que los resultados contribuyan a una mejor comprensión de la estructura, el origen y la persistencia de los principales clados de transmisión local de VIH-1B en Uruguay, aportando información relevante para la vigilancia molecular y epidemiológica del VIH-1 en el país.

2. Hipótesis

La epidemia de VIH-1B en Uruguay se originó a partir de múltiples introducciones independientes, que contribuyeron de manera desigual a la transmisión nacional. La estructura actual de la epidemia estaría determinada principalmente por un número reducido de linajes que lograron establecerse y persistir localmente, dando origen a clados uruguayos diferenciados por su origen geográfico, antigüedad y dinámica temporal de transmisión.

3. Objetivos

3.1. Objetivo general

Caracterizar la historia evolutiva y la dinámica de transmisión del VIH-1B en Uruguay mediante un abordaje filogenético, filogeográfico y filodinámico, con énfasis en los procesos de introducción, establecimiento, expansión y persistencia de los linajes virales que conforman la epidemia nacional.

3.2. Objetivos específicos

- a) Caracterizar la estructura filogenética del VIH-1B en Uruguay, diferenciando introducciones independientes y clados de transmisión local.
- b) Inferir el probable origen geográfico de las introducciones independientes y de los clados de transmisión local de VIH-1B en Uruguay.
- c) Estimar la dinámica evolutiva y epidemiológica de los principales clados locales de VIH-1B en Uruguay.

4. Metodología

4.1. Aspectos éticos

Este trabajo forma parte del proyecto “Vigilancia genómica del Virus de la Inmunodeficiencia Humana Tipo 1 (VIH-1) en Uruguay”, cuya responsable es la Dra. Daiana Mir. Este proyecto fue financiado por la Comisión Sectorial de Investigación Científica (CSIC) y aprobado por el Comité de Ética del CENUR Litoral Norte (Exp. N° 311170-500476-21). Los datos se recopilaron en 2022 y se analizaron entre 2024 y 2025. Todos los métodos se realizaron de acuerdo con las pautas y regulaciones pertinentes.

4.2. Datos, muestreo y secuencias virales

El conjunto de datos de partida comprende 1622 secuencias del gen *pol* generadas mediante secuenciación Sanger de fragmentos que abarcan las regiones PR y RT, obtenidas a partir de ARN viral mediante RT-PCR seguida de PCR anidada. Estas secuencias forman parte de la base de datos nacional GENOUY-VIH, construida a partir de estudios de resistencia a antirretrovirales realizados en el DLSP-MSP. La base integra secuencias virales generadas en el marco del diagnóstico rutinario de resistencia junto con variables clínicas y epidemiológicas asociadas, vinculadas de forma anonimizada. En este contexto, se trata de un muestreo de conveniencia, basado en individuos para los cuales se indicó el test de resistencia en distintos escenarios clínicos (p.ej., falla virológica, primoinfección sintomática, embarazo, entre otros). Los procedimientos experimentales, así como la construcción de la base de datos y la vinculación con variables clínicas y epidemiológicas, han sido descritos previamente (Podchibiakin et al., en revisión, Anexo 1).

A partir de este conjunto inicial, las secuencias fueron sometidas a procesos de curación, control de calidad y filtrado, incluyendo la eliminación de redundancias por paciente, la exclusión de secuencias incompletas y la verificación de la cobertura de la región analizada. En los casos en que existía más de una secuencia por paciente, se conservó la primera secuencia disponible. En una etapa previa de este trabajo, se realizó la caracterización filogenética y subtipificación de las secuencias uruguayas, lo que permitió identificar los linajes virales circulantes y definir de manera robusta el conjunto correspondiente a VIH-1B.

El subconjunto analizado en esta tesis incluyó 491 secuencias del gen *pol* de VIH-1B correspondientes a muestras recolectadas en Uruguay entre los años 2007 y 2021, que cubren las posiciones 2358–3251 con respecto a la secuencia de referencia HXB2.

4.3. Asignación filogenética de clados

Con el objetivo de determinar la posición filogenética de las secuencias uruguayas dentro de la diversidad global del VIH-1B, se realizó un análisis de asignación filogenética utilizando secuencias de referencia representativas de los principales linajes descritos para este subtipo, incluyendo los linajes B_{PAN} y B_{CAR}, así como clados regionales latinoamericanos previamente caracterizados.

El conjunto de secuencias uruguayas fue alineado junto a un panel de secuencias de referencia curadas utilizando MAFFT v7 (Katoh & Standley, 2013). El alineamiento fue posteriormente analizado y editado manualmente en Aliview (Larsson, 2014) para eliminar los sitios de resistencia a fármacos. Esta etapa resulta fundamental ya que la región *pol* constituye el principal blanco de Terapia Antirretroviral (TARV), y la presión selectiva ejercida por los fármacos puede favorecer eventos de convergencia evolutiva hacia caracteres específicos que no son resultado de los eventos de la evolución natural del virus, pudiendo sesgar la inferencia evolutiva.

El panel de inferencia incluyó: (i) secuencias B_{PAN} provenientes de Estados Unidos y Francia (n = 300), (ii) secuencias B_{CAR} del Caribe (n = 200), y (iii) los principales clados B_{PAN} previamente caracterizados en América Latina (B_{PAN-SAM}, B_{PAN-AR-II}, B_{PAN-AR-III}, B_{PAN-BR-II}, B_{PAN-BR-III}, B_{PAN-BR-IV}, B_{PAN-PE}, B_{PAN-CAM}, B_{PAN-MX-I} y B_{PAN-MX-II}) (Mir et al., 2015). Estos conjuntos de referencia permiten representar la estructura filogenética conocida del subtipo B a escala global y regional.

La asignación de las secuencias a los distintos linajes de VIH-1B se realizó mediante el Evolutionary Placement Algorithm (EPA) implementado en RAxML v8.2.4 (Stamatakis, 2014), el cual permite ubicar secuencias consulta (*query sequences*) sobre un árbol filogenético de referencia previamente inferido, reduciendo en gran medida la complejidad de reconstruir *de novo* la topología completa del árbol para cada conjunto de datos nuevo.

El programa se utilizó mediante el siguiente comando:

```
raxmlHPC -f v -s alineamiento.fasta -t arbol-referencia.nwk -m GTRCAT -n output
```

donde:

- **-f v** indica la ejecución del EPA
- **-s** corresponde al alineamiento de secuencias como input
- **-t** al árbol de referencia previamente inferido

- **-m GTRCAT** al modelo de sustitución utilizado, que incorpora el modelo GTR con una aproximación CAT, que representa una aproximación computacionalmente eficiente de la heterogeneidad de tasas entre sitios.
- **-n** permite nombrar el archivo de salida (*output*)

Una vez asignadas a los principales linajes, las secuencias clasificadas dentro de la diversidad B_{PAN} fueron analizadas con mayor resolución para refinar su posición dentro de los clados regionales latinoamericanos y evaluar la presencia de posibles clados de transmisión local. Para ello, se estratificaron en función de su afinidad filogenética inicial y se re-analizaron junto a sus paneles de referencia correspondientes.

Específicamente:

- Secuencias clasificadas dentro de la diversidad sudamericana fueron analizadas junto a las secuencias de referencia de los clados $B_{PAN-SAM}$, $B_{PAN-AR-II}$, $B_{PAN-AR-III}$, $B_{PAN-BR-II}$, $B_{PAN-BR-III}$, $B_{PAN-BR-IV}$ y B_{PAN-PE} .
- Secuencias vinculadas a la diversidad centroamericana y mexicana se analizaron junto a los clados $B_{PAN-CAM}$, $B_{PAN-MX-I}$ y $B_{PAN-MX-II}$ respectivamente
- Secuencias asignadas a B_{PAN} pero no pertenecientes a clados latinoamericanos previamente descritos fueron analizadas junto a referencias B_{PAN} de Estados Unidos y Francia.

4.4. Análisis filogenético por máxima verosimilitud

Las inferencias filogenéticas se realizaron mediante ML utilizando el programa PhyML (Guindon et al., 2005; 2010). El modelo de sustitución nucleotídica más apropiado para cada conjunto de datos fue seleccionado mediante SMS (Smart Model Selection) implementado en PhyML (Lefort et al., 2017). La búsqueda de la topología óptima se llevó a cabo mediante un algoritmo heurístico basado en SPR. El soporte estadístico de las ramas fue evaluado mediante el approximate likelihood-ratio test (aLRT) bajo el procedimiento tipo Shimodaira-Hasegawa (SH-aLRT) (Anisimova & Gascuel, 2006), considerándose valores ≥ 0.85 como evidencia de alto soporte.

Las inferencias se realizaron de manera independiente para cada uno de los subconjuntos definidos en la etapa de asignación filogenética, incluyendo los clados latinoamericanos previamente descritos y las secuencias no asociadas a clados conocidos. En total, se llevaron a cabo 11 inferencias independientes por ML, lo que permitió delimitar clados y

refinar la asignación filogenética de las secuencias uruguayas dentro de la estructura del subtipo B.

Los árboles resultantes fueron visualizados y editados con el programa FigTree v1.4.4. (Rambaut, 2018).

4.5. Caracterización epidemiológica y demográfica de los metadatos asociados al dataset analizado

La información demográfica y clínico-epidemiológica de los pacientes se obtuvo a partir de los formularios clínicos anexos a las secuencias virales en la plataforma de vigilancia GENOUY-VIH, cumplimentados al momento de realizar el test de resistencia a ARV. Con el objetivo de caracterizar la muestra bajo estudio y contextualizar la estructura de los clados de transmisión viral dentro del escenario epidemiológico nacional, los metadatos fueron integrados, curados y analizados estadísticamente utilizando el software R v4.5.0 (R Core Team, 2025). Es importante destacar que, dado que este dataset proviene de un muestreo de conveniencia ligado a la vigilancia de resistencias de pacientes en seguimiento clínico, las variables reflejan el perfil de la población bajo control médico y no necesariamente la dinámica demográfica irrestricta de la epidemia nacional. Durante la etapa de curación de datos, se aplicó un criterio de exclusión estricto por el cual no se tuvieron en cuenta para los análisis específicos aquellas secuencias cuyos metadatos presentaban registros ambiguos o inconsistentes en alguna de las variables consideradas.

Para el análisis descriptivo, se clasificaron las secuencias de acuerdo con las siguientes variables discretas:

- **Sexo:** categorizado en masculino, femenino, transgénero y no disponible (ND) para aquellos registros sin dato asociado.
- **Vía de transmisión reportada:** clasificada en las categorías de hombres que tienen sexo con hombres (HSH), heterosexual (HTX), usuarios de drogas inyectables (UDI), accidente laboral y ND.
- **Lugar de adquisición de la infección:** categorizado según la procedencia geográfica reportada, discriminando por departamento del territorio uruguayo, país extranjero (en caso de infecciones importadas) o ND.

4.6. Análisis filogeográfico por máxima verosimilitud

La señal temporal de los árboles ML fue evaluada mediante regresión *root-to-tip* utilizando TempEst v1.5 (Rambaut, 2016), determinando el mejor enraizamiento y verificando la adecuación del modelo de reloj molecular. Posteriormente, los árboles fueron escalados en el tiempo mediante TreeTime v0.8.6 (Sagulenko et al., 2018). El análisis se llevó a cabo mediante el siguiente comando:

```
treetime --tree arbol.nwk --aln alineamiento.fasta --dates fechas-de-muestreo.tsv --clock-rate 0.0019 --gtr infer --confidence --stochastic-resolve --keep-root --clock-filter 0
```

donde:

- **--tree** corresponde al árbol filogenético de entrada
- **--aln** al alineamiento de secuencias correspondiente al árbol
- **--dates** al archivo que contiene las fechas de muestreo de las secuencias (información disponible en la base de datos GENOUY-VIH)
- **--clock-rate** define una tasa de sustitución inicial
- **--gtr infer** indica la inferencia del modelo de sustitución
- **--confidence** permite estimar intervalos de confianza para los nodos
- **--stochastic-resolve** resuelve politomias
- **--keep-root** conserva la raíz previamente definida
- **--clock-filter 0** evita el filtrado de secuencias en función de desviaciones del reloj molecular

Se utilizó una tasa de sustitución inicial de 0.0019 sustituciones/sitio/año, consistente con estimaciones previas para el gen *pol* del VIH-1. Este valor fue utilizado como punto de partida para la inferencia, siendo posteriormente optimizado durante el análisis por TreeTime.

La reconstrucción geográfica ancestral de los nodos internos se realizó con PastML (Ishikawa et al., 2019) a partir de árboles enraizados y de la información geográfica asociada a las secuencias terminales. Se empleó un enfoque de reconstrucción ancestral por máxima verosimilitud utilizando el algoritmo MAP para asignar a los nodos internos los estados geográficos con mayor probabilidad posterior, bajo un modelo tipo F81 para caracteres discretos. Este modelo asume que las transiciones entre estados ocurren con probabilidades proporcionales a sus frecuencias de equilibrio, lo cual constituye una

aproximación adecuada en contextos donde la información disponible por carácter es limitada, permitiendo una estimación robusta sin sobreparametrización. En este análisis, los árboles previamente calibrados temporalmente con TreeTime fueron utilizados como insumo para la reconstrucción ancestral, lo que permitió mantener la coherencia entre la escala temporal del árbol y la interpretación de los patrones de dispersión geográfica inferidos.

Los clados de transmisión uruguayos fueron definidos como clados monofiléticos altamente soportados ($aLRT \geq 0.85$), cuyo nodo ancestral fue inferido geográficamente en Uruguay y compuestos predominantemente ($>90\%$) por secuencias uruguayas. A partir de estos criterios, fueron clasificados según su tamaño en: grandes: $n \geq 30$, medianos: $10 \leq n < 30$ y pequeños: $2 \leq n < 10$.

4.7. Inferencia bayesiana de la escala temporal

La estimación del tiempo al ancestro común más reciente (tMRCA) de los clados B_{PAN} uruguayos se realizó mediante un enfoque bayesiano basado en cadenas MCMC, implementado en BEAST v1.10.4 (Drummond et al., 2002; Suchard et al., 2018), utilizando la biblioteca BEAGLE para optimizar el rendimiento computacional (Suchard & Rambaut, 2009).

Las filogenias fueron inferidas bajo el modelo coalescente Bayesian Skyline (Drummond, 2005), el modelo de sustitución nucleotídica GTR + I + Γ_4 (Drummond et al., 2006) y un reloj molecular relajado no correlacionado lognormal (Drummond et al., 2006). Se especificó un prior uniforme para la tasa de sustitución ($1.7\text{--}2.1 \times 10^{-3}$ sustituciones/sitio/año), basado en estimaciones previas reportadas en la literatura (Figura 7) (Mendoza et al., 2014; Divino et al., 2016).

Las cadenas MCMC fueron ejecutadas durante $10\text{--}30 \times 10^7$ generaciones, asegurando valores de tamaño efectivo de muestra (ESS, por sus siglas en inglés) superiores a 200 para todos los parámetros relevantes. Las distribuciones posteriores y los intervalos de mayor densidad posterior al 95% (95% HPD) fueron evaluados mediante Tracer v1.7.1 (Rambaut et al., 2018). Los árboles de máxima credibilidad (Maximum Clade Credibility Trees, MCCT) fueron resumidos con TreeAnnotator v1.10.4 y visualizados en FigTree v1.4.4.

4.8. Reconstrucciones demográficas bayesianas

La historia demográfica de los principales clados B_{PAN} uruguayos fue reconstruida mediante enfoques bayesianos MCMC. Los cambios temporales en el N_e fueron estimados utilizando el modelo coalescente BSKG, implementado en BEAST v1.10.4 (Drummond et al., 2002; Suchard et al., 2018). Los cambios en el R_e fueron estimados mediante el modelo BDSKY, implementado en BEAST v2.1. (Russel et al., 2019).

Las configuraciones a priori del modelo BDSKY fueron las siguientes: para la tasa de pérdida de infectividad (*becoming non-infectious rate*) se especificó una distribución lognormal (media = 0,25; desviación estándar = 0,5); para el R_e se utilizó una distribución lognormal (media = 0,0; desviación estándar = 1,0); y para la proporción de muestreo se definió una distribución beta ($\alpha = 1,0$; $\beta = 25,0$).

El parámetro de origen fue condicionado a la altura de la raíz del árbol (*root height*), y el R_e se estimó de manera segmentada (*piecewise*) a lo largo de ocho intervalos temporales definidos desde la fecha de la muestra más reciente hasta la raíz del árbol. La cadena MCMC se ejecutó hasta que todos los parámetros relevantes alcanzaron valores de tamaño efectivo de muestra (ESS) superiores a 200.

La evolución temporal estimada de R_e se representó utilizando el paquete bdskytools en R (<https://github.com/laduplessis/bdskytools>). Las distancias temporales entre eventos filogenéticos sucesivos se estimaron mediante el software TreeStat (Suchard et al., 2018), a partir de las distancias internodales de los árboles de MCCT, y se utilizaron como una aproximación indirecta a la temporalidad de posibles eventos de transmisión secundaria. Estas distancias se agruparon en cuatro intervalos: <0,5; 0,5–2; 2–5 y >5 años.

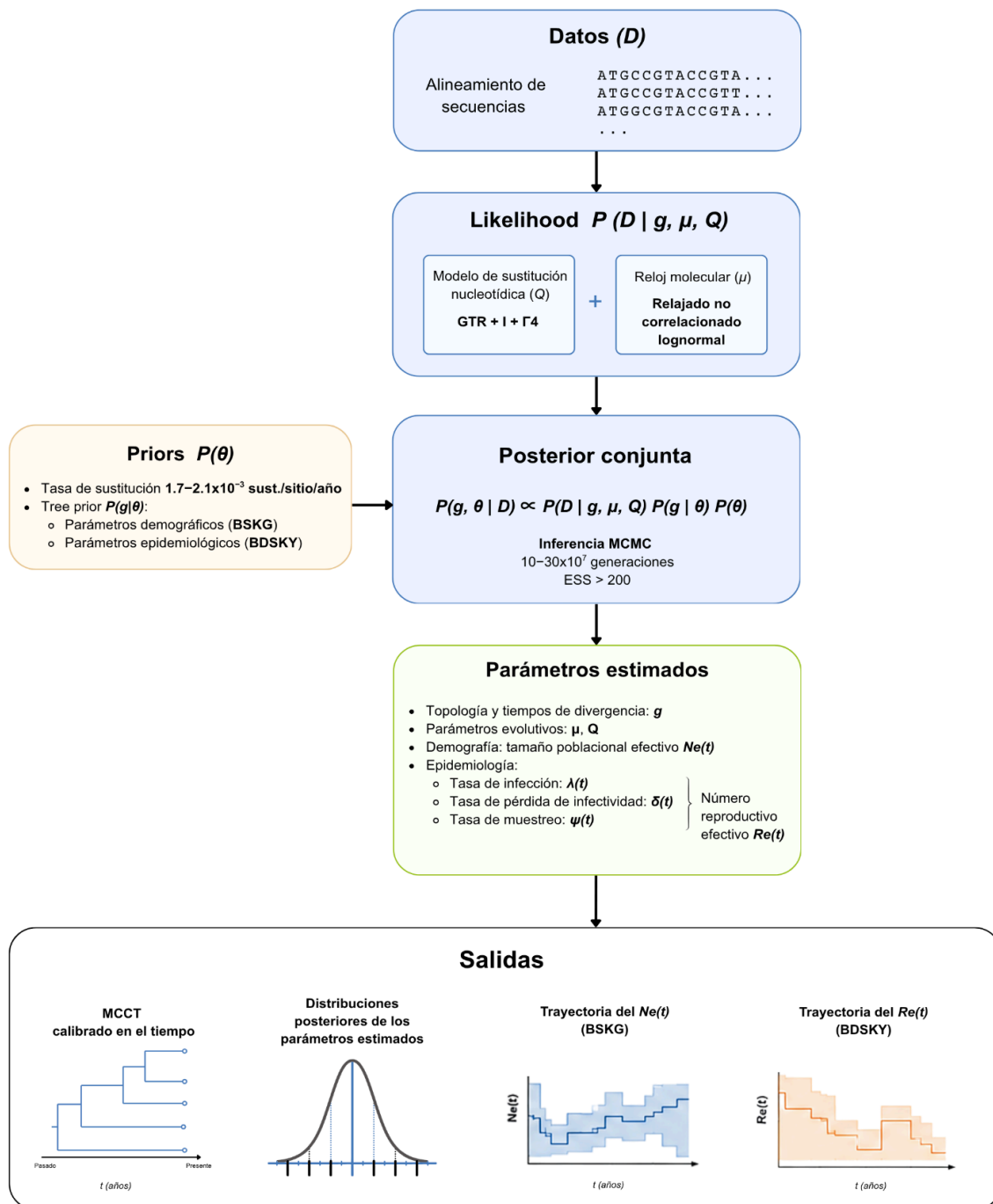


Figura 7. Esquema conceptual del marco de inferencia bayesiana implementado en los análisis realizados en BEAST y BEAST 2. El modelo integra la información contenida en el alineamiento de secuencias (D) mediante la combinación de la likelihood $P(D|g, \mu, Q)$, definida por el modelo de sustitución nucleotídica (Q) y el reloj molecular (μ), junto con el tree prior $P(g|\theta)$, especificado mediante modelos coalescentes (Bayesian Skygrid, BSKG) o birth–death (Birth–Death Skyline, BDSKY), y los priors $P(\theta)$ sobre los parámetros evolutivos y epidemiológicos. La inferencia bayesiana permite estimar la distribución posterior conjunta $P(g, \theta | D)$, a partir de la cual se obtienen la genealogía calibrada en el tiempo (g), las trayectorias temporales del tamaño efectivo poblacional $N_e(t)$, el número reproductivo efectivo $Re(t)$, y otros parámetros evolutivos y epidemiológicos. Los outputs finales incluyen los MCCT, distribuciones posteriores de los parámetros de interés (tMRCA; parámetros generales del MCMC; parámetros del reloj molecular; del modelo de sustitución, entre otros) y reconstrucciones demográficas.

5. Resultados

5.1. Introducciones de VIH-1B y clados de transmisión local

Como era esperado, los análisis filogenéticos de ML mostraron que, del total de secuencias uruguayas de VIH-1B analizadas ($n = 491$), la gran mayoría se agrupó dentro del linaje B_{PAN} ($n = 473$; 96,3%), y una pequeña proporción al linaje B_{CAR} ($n = 18$; 3,7%) (Figura 8). Entre las secuencias B_{PAN} , el 37,0% ($n = 175$) se agrupó dentro de clados previamente descritos en América Latina por Mir et al. (2015), mientras que una gran proporción (63,0%; $n = 298$) no se asoció con ningún clado B_{PAN} descrito hasta la fecha (Tabla 1). Estas secuencias fueron analizadas mediante una inferencia filogenética independiente (B_{PAN-UY}), que permitió identificar múltiples subclados de transmisión local, así como introducciones filogenéticamente aisladas.

Se realizaron 11 inferencias filogenéticas por ML, correspondientes a los principales clados B_{PAN} (Figura 9) y para el conjunto de secuencias B_{PAN-UY} (Figura 10). A partir del conjunto de árboles inferidos, se identificaron 229 introducciones independientes de variantes B_{PAN} en Uruguay. De estas, 52 dieron lugar a clados de transmisión local ($aLRT \geq 0,85$), mientras que las 177 restantes corresponden a introducciones esporádicas sin evidencia de diseminación local posterior (*singletons*).

Los 52 clados de transmisión presentaron una marcada heterogeneidad en tamaño: dos fueron clasificados como grandes (3,8%; $n \geq 30$), nueve como medianos (17,3%; $10 \leq n < 30$) y 41 como pequeños (78,9%; $2 \leq n < 10$). Los dos clados grandes correspondieron a subclados del grupo B_{PAN-UY} y fueron designados $B_{PAN-UY-I}$ ($n = 37$) y $B_{PAN-UY-II}$ ($n = 30$).

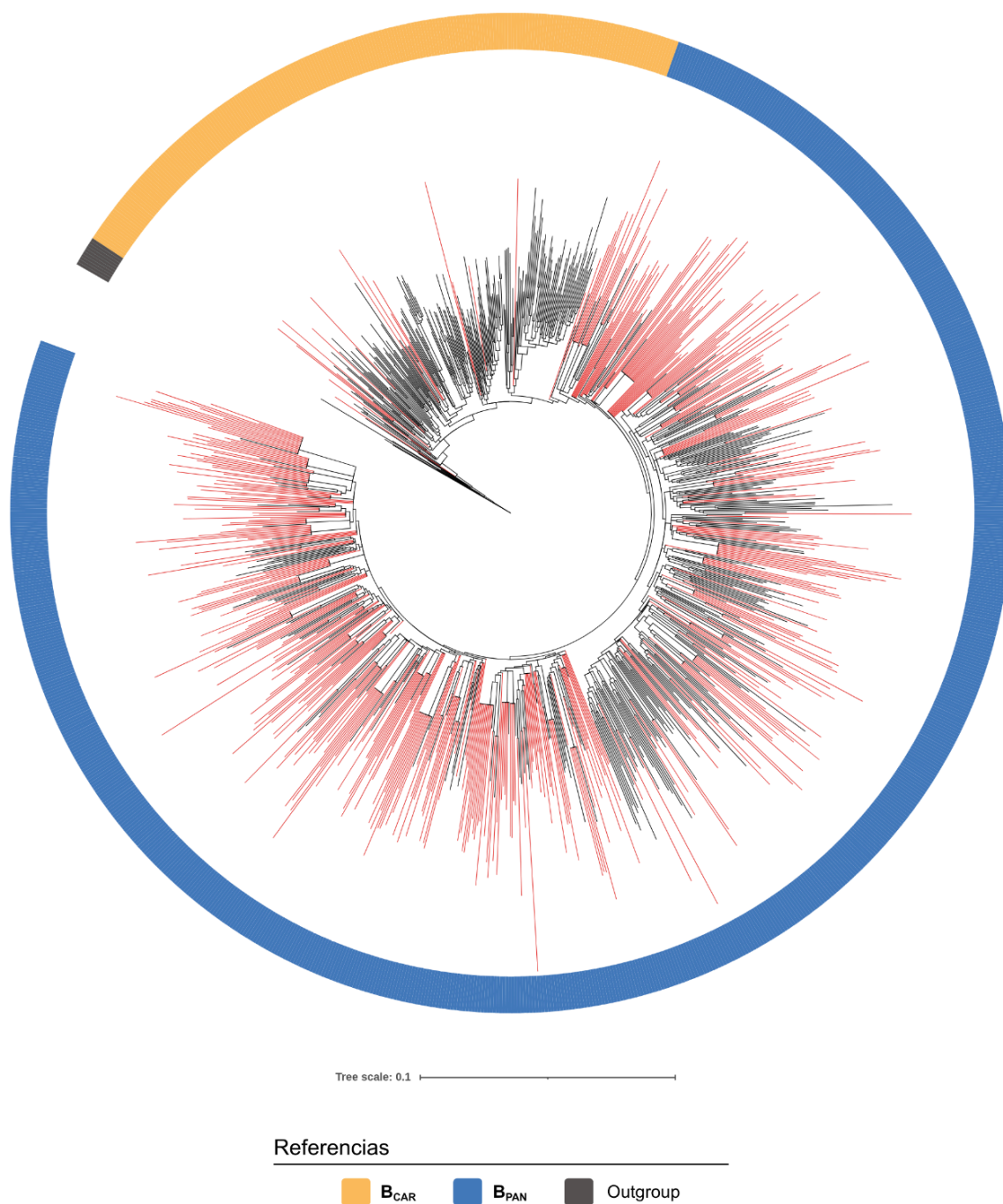


Figura 8. Árbol filogenético de máxima verosimilitud utilizado para la asignación filogenética de las secuencias uruguayas de VIH-1 subtipo B. Filogenia de máxima verosimilitud de secuencias del gen *pol* del VIH-1 subtipo B, utilizada como árbol de referencia para el análisis de asignación mediante EPA/RAXML. Las secuencias uruguayas provenientes de la base de datos GENOUY-VIH se muestran en rojo, mientras que las secuencias de referencia correspondientes a los linajes pandémico (B_{PAN}, n = 300) y caribeño (B_{CAR}, n = 200) del subtipo B se muestran en negro. El anillo externo indica la asignación de clados, con B_{CAR} en amarillo y B_{PAN} en azul; el clúster gris corresponde a secuencias del subtipo D utilizadas como grupo externo (*outgroup*) (n = 10). La longitud de las ramas es proporcional al número de sustituciones de nucleótidos por sitio.

Tabla 1. Clasificación filogenética de las secuencias del gen *pol* de VIH-1B de Uruguay.

Linaje	Sub-linaje	N (%)	Rango de muestreo
B_{PAN}		473 (100%)	2007 – 2021
	B _{PAN-UY}	298 (63.0%)	2007 – 2021
	B _{PAN-SAM}	94 (19.8%)	2007 – 2021
	B _{PAN-AR-II}	21 (4.4%)	2008 – 2021
	B _{PAN-BR-III}	15 (3.2%)	2007 – 2021
	B _{PAN-MX-II}	13 (2.7%)	2008 – 2021
	B _{PAN-BR-II}	10 (2.1%)	2010 – 2019
	B _{PAN-MX-I}	8 (1.7%)	2013 – 2015
	B _{PAN-AR-III}	4 (0.8%)	2012 – 2015
	B _{PAN-CAM}	4 (0.8%)	2008 – 2015
	B _{PAN-BR-IV}	3 (0.6%)	2015 – 2019
	B _{PAN-PE}	3 (0.6%)	2009 – 2018
B_{CAR}		18 (100%)	2007 – 2021

Los clados medianos y pequeños se identificaron tanto en subclados B_{PAN-UY} como en siete clados regionales previamente reportados (B_{PAN-SAM}, B_{PAN-AR-II}, B_{PAN-AR-III}, B_{PAN-BR-II}, B_{PAN-BR-III}, B_{PAN-MX-I} y B_{PAN-MX-II}), lo que evidencia múltiples eventos de introducción con distinta capacidad de generar transmisión secundaria detectable. En contraste, los clados B_{PAN-BR-IV}, B_{PAN-CAM} y B_{PAN-PE} estuvieron representados únicamente por introducciones esporádicas, sin formación de clados de transmisión local detectables (Figura 9–11).

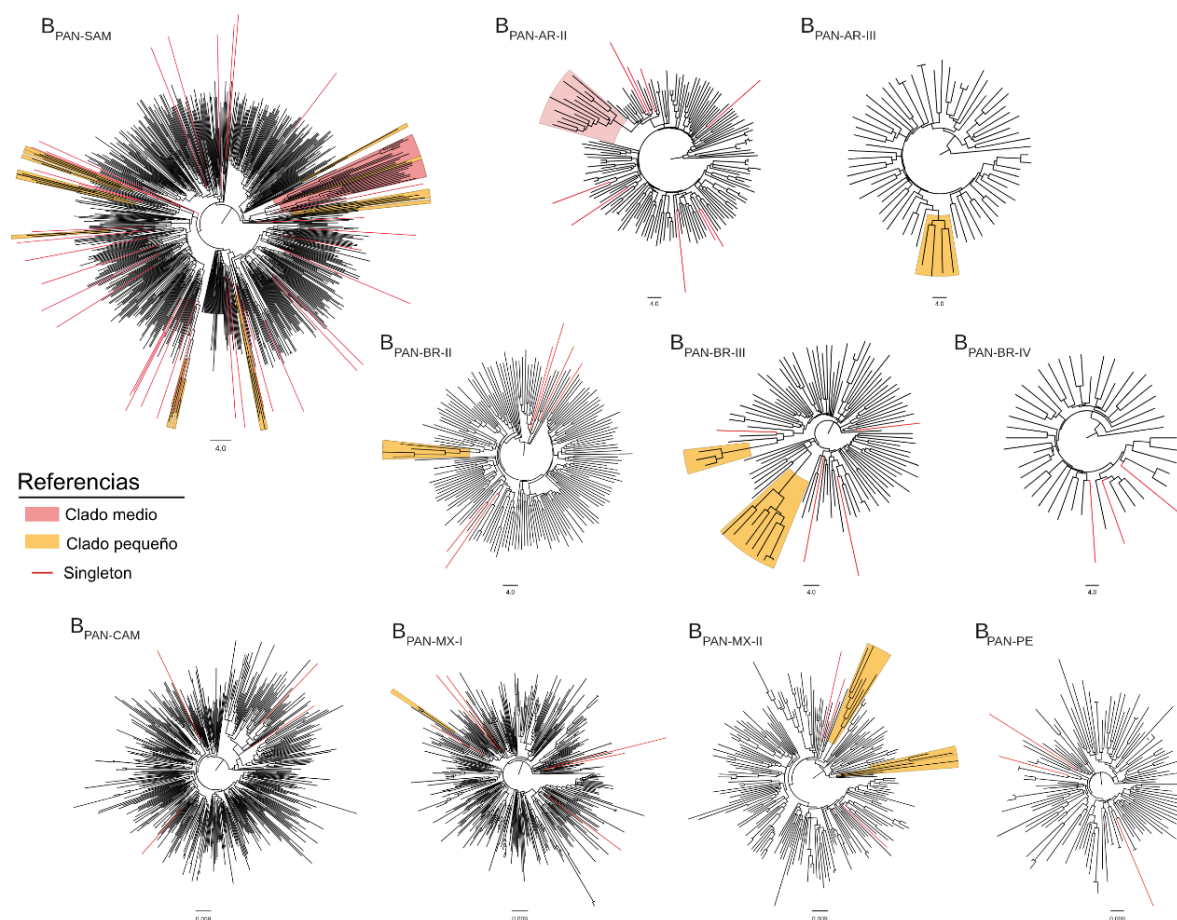


Figura 9. Árboles filogenéticos de máxima verosimilitud de clados B_{PAN} previamente descritos en América Latina con presencia de secuencias uruguayas. Para cada clado B_{PAN} detectado en Uruguay se inferió un árbol independiente mediante PhyML. Los clados de transmisión local uruguayos se resaltan según su tamaño: clados medianos en rosa y clados pequeños en amarillo. Las ramas terminales en rojo corresponden a introducciones esporádicas, representadas por secuencias uruguayas sin evidencia de transmisión local posterior (singletons). Los clados de transmisión destacados presentan alto soporte estadístico ($aLRT \geq 0,85$). Los árboles fueron enraizados en el punto medio y las longitudes de las ramas son proporcionales al número de sustituciones nucleotídicas por sitio.

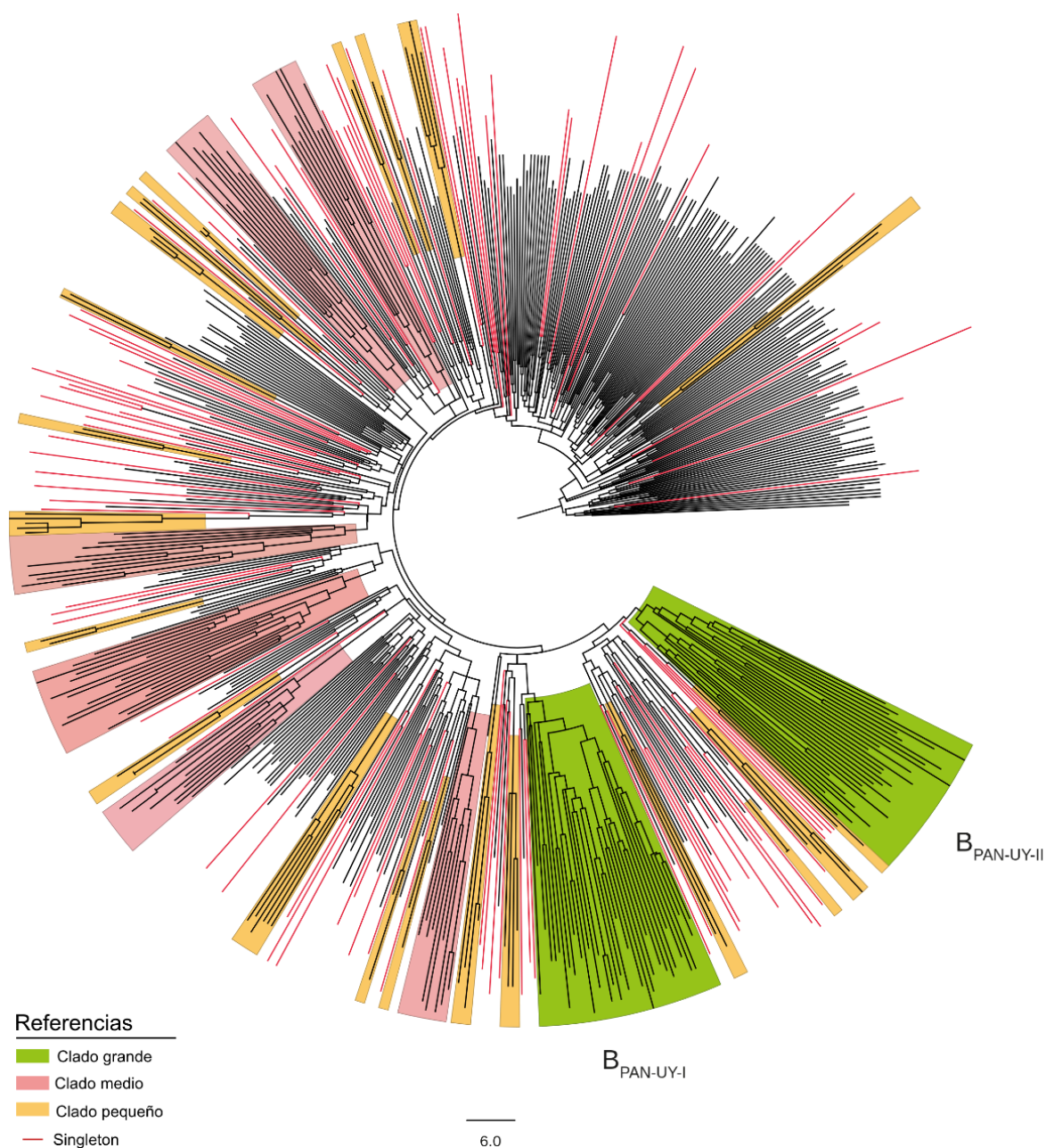


Figura 10. Subclados B_{PAN-UY} e introducciones esporádicas dentro de la diversidad B_{PAN} no asociada a clados latinoamericanos conocidos. Árbol filogenético de máxima verosimilitud inferido mediante PhyML para las secuencias uruguayas de VIH-1 subtipo B asignadas al linaje B_{PAN} , pero no agrupadas dentro de clados B_{PAN} latinoamericanos previamente descritos. Los clados de transmisión uruguayos se resaltan según su tamaño: grandes en verde, medianos en rosa y pequeños en amarillo. Las ramas terminales en rojo corresponden a introducciones esporádicas sin evidencia de transmisión local posterior (*singletons*). Todos los clados destacados presentan valores de aLRT $\geq 0,85$. El árbol fue enraizado por punto medio y las longitudes de las ramas son proporcionales al número de sustituciones de nucleótidos por sitio.

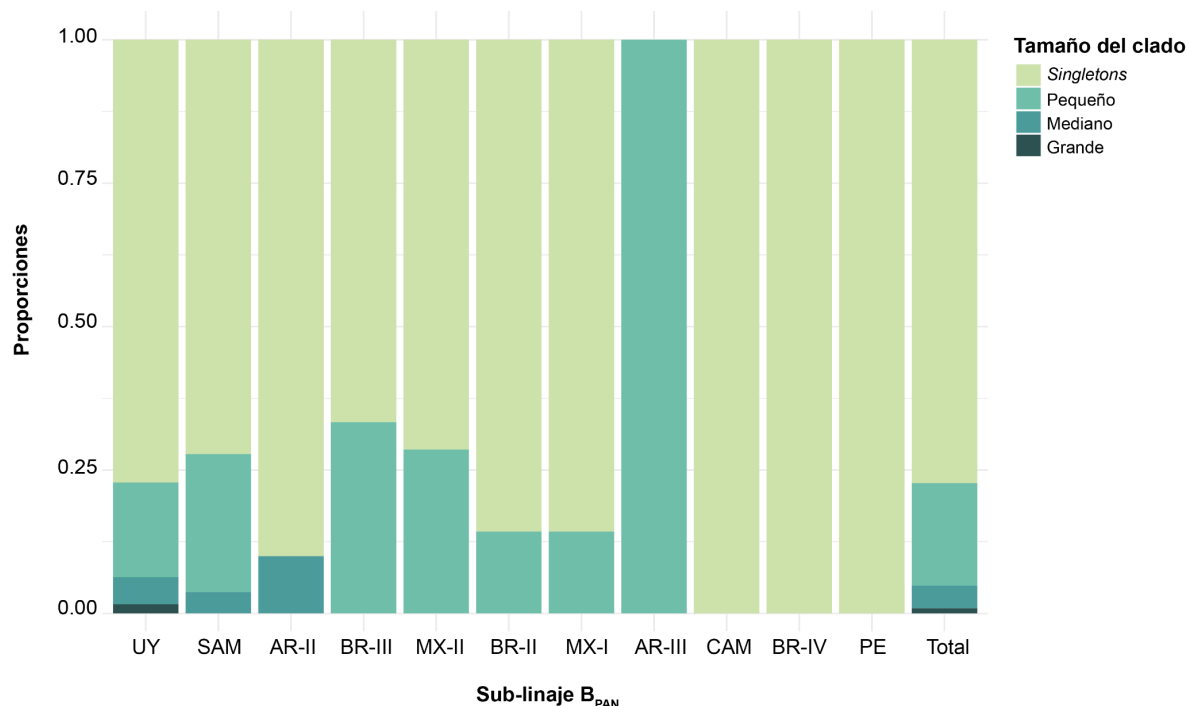


Figura 11. Distribución proporcional de las introducciones independientes de variantes B_{PAN} en Uruguay. Para cada sublinaje B_{PAN} se muestra la proporción de introducciones esporádicas (*singletons*) y de introducciones que dieron lugar a clados pequeños ($2 \leq n < 10$), medianos ($10 \leq n < 30$) o grandes ($n \geq 30$). Las proporciones totales se muestran en la última barra.

5.2. Caracterización epidemiológica y demográfica del dataset analizado

La base de datos analizada está compuesta por un total de 491 secuencias, de las cuales 473 corresponden al clado B_{PAN} y 18 al clado B_{CAR}.

Características demográficas y epidemiológicas asociadas a las secuencias B_{PAN}

Sexo

En el clado B_{PAN}, y de acuerdo con los metadatos obtenidos a partir de los formularios asociados a las secuencias virales (test de resistencia a antirretrovirales), se observó que 305 de las 473 secuencias (64,5%) corresponden a infecciones en pacientes de sexo masculino, 160 (33,8%) corresponden a pacientes de sexo femenino, cuatro (0,8%) pertenecen a personas transgénero y otras cuatro (0,8%) no presentan información disponible. Esta predominancia del sexo masculino se observa también en la mayoría de los

sublinajes B_{PAN} identificados, al considerar el total de secuencias incluidas en cada uno de ellos (Tabla 2).

Tabla 2. Distribución por sexo de los pacientes según los sublinajes de VIH-1 B_{PAN} identificados a partir de los metadatos de la plataforma GENOUY-VIH.

Sub-linaje B _{PAN}	Femenino	Masculino	Transgénero	ND
UY	106	189	2	1
SAM	25	67	1	1
AR-II	7	13	0	1
AR-III	3	1	0	0
BR-II	4	5	1	0
BR-III	6	9	0	0
BR-IV	1	2	0	0
CAM	2	1	0	1
MX-I	3	5	0	0
MX-II	3	10	0	0
PE	0	3	0	0
TOTAL	160	305	4	4

En cuanto a la estructura filogenética, del total de secuencias B_{PAN} (n = 473), 295 (62,4%) se agrupan en alguno de los 52 clados identificados dentro de los sublinajes B_{PAN}, incluyendo B_{PAN-UY}, conformando clados de tamaño pequeño, medio y grande. De estos 52 clados, 20 (38,5%) están predominantemente conformados por individuos de sexo masculino ($\geq 80\%$ de las secuencias), de los cuales 17 corresponden a clados pequeños y tres a clados de tamaño medio. En contraste, únicamente tres clados (5,8%) presentan predominancia de individuos de sexo femenino, todos ellos de tamaño pequeño. Los 29 clados restantes (55,8%) presentan una composición mixta (es decir, que presenta individuos de ambos sexos pero sin que uno de ellos se encuentre en una proporción mayor al 80%). Cabe destacar que la mayoría de los clados de tamaño medio y grande muestran una distribución mixta en términos de género (Figura 12).

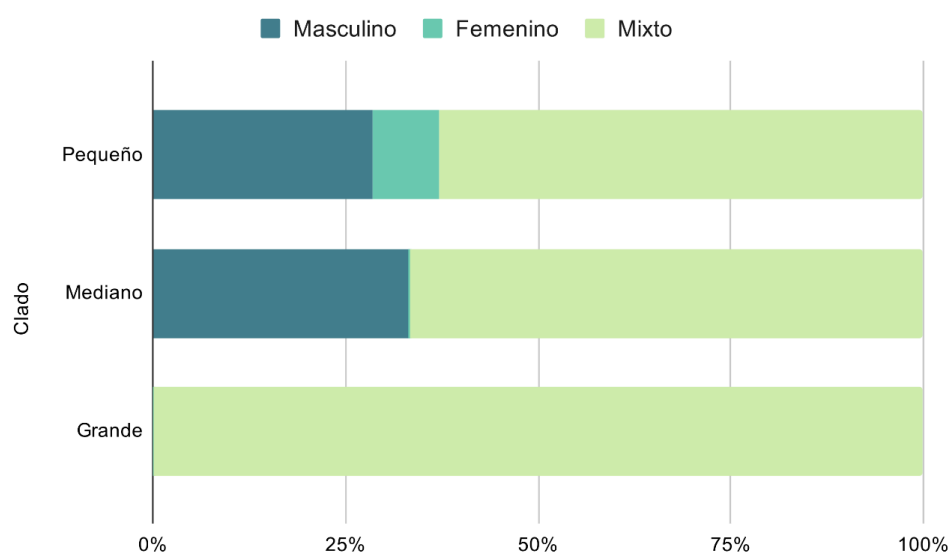


Figura 12. Proporción de clados pequeños, medios y grandes según el sexo predominante en cada uno.

Vía de Transmisión

La información sobre vía de transmisión estuvo disponible con distinto grado de resolución entre las 473 secuencias B_{PAN} analizadas. De estas, 103 secuencias (21,8%) correspondieron a transmisión en HSH, 210 (44,4%) a HTX, nueve (1,9%) a UDI y una (0,2%) a accidente laboral. Además, 128 secuencias (27,1%) fueron clasificadas únicamente como “sexual”, sin discriminación entre HSH y HTX, y 22 secuencias (4,7%) no contaron con información disponible sobre la vía de transmisión. Por este motivo, los análisis que requerían una clasificación específica de la vía de transmisión se realizaron sobre 323 secuencias (68,3%), correspondientes a aquellas con información discriminada en las categorías HSH, HTX, UDI o accidente laboral (Tabla 3).

Al considerar la distribución de la vía de transmisión en los clados identificados ($n = 52$), se observó que solo ocho clados presentaron predominio de secuencias asociadas a transmisión HTX ($\geq 80\%$ de las secuencias) y un clado presentó predominio de secuencias asociadas a transmisión HSH. En todos los casos, se trató de clados pequeños. En los clados de tamaño medio y grande no se observaron patrones claros de predominio de una vía de transmisión específica.

Tabla 3. Distribución según vía de transmisión de las secuencias pertenecientes a los sublinajes de VIH-1 B_{PAN} identificados a partir de los metadatos asociados a las secuencias de la base de datos GENOUY-VIH.

Sub-linaje B _{PAN}	HTX	HSH	UDI	Accidente	Sexual	ND
UY	124	63	8	1	90	12
SAM	46	21	0	0	22	5
AR-II	10	4	1	0	4	2
AR-III	2	0	0	0	1	1
BR-II	5	1	0	0	4	0
BR-III	7	4	0	0	2	2
BR-IV	2	1	0	0	0	0
CAM	2	0	0	0	2	0
MX-I	4	4	0	0	0	0
MX-II	7	4	0	0	2	0
PE	1	1	0	0	1	0
TOTAL	210	103	9	1	128	22

Cabe señalar que no se realizó una representación gráfica de las proporciones de las vías de transmisión por clados debido a la elevada proporción de secuencias con información incompleta o no discriminada.

Lugar de adquisición de la infección

En relación con el lugar de adquisición del virus, de las 473 secuencias B_{PAN}, 309 (65,3%) contaron con información geográfica específica, mientras que 121 (25,6%) no presentaron datos disponibles y 43 (9,1%) consignaron únicamente “Uruguay”, sin especificación departamental. Entre las secuencias con información específica, 277 correspondieron a infecciones adquiridas en Uruguay. De estas, la mayoría se concentra en la región metropolitana (Canelones, Montevideo y San José), con 218 secuencias (78,7%), seguida por la región noreste (Artigas, Rivera, Tacuarembó, Cerro Largo y Treinta y Tres) con 26 secuencias (9,4%); la región oeste (Salto, Paysandú, Río Negro, Soriano y Colonia) con 19 (6,9%); la región este (Maldonado y Rocha) con 12 (4,3%) y, finalmente, la región centro (Flores, Florida, Durazno y Lavalleja) con 2 secuencias (0,7%).

El resto de las secuencias con información geográfica correspondió a infecciones adquiridas fuera de Uruguay: 12 en Argentina, 3 en Brasil, 3 en Norteamérica (Estados Unidos y

Canadá), 9 de otros países de América Latina y el Caribe (Paraguay, Ecuador, Venezuela, Colombia y República Dominicana), y 5 en Europa (España, Italia, Países Bajos y Francia).

La marcada sobrerrepresentación de secuencias provenientes de la región metropolitana, en conjunto con la menor disponibilidad de datos para el resto de las regiones del país y la proporción no menor de secuencias sin información geográfica precisa, limita la posibilidad de realizar análisis robustos sobre la distribución espacial de los linajes virales dentro de Uruguay. En este sentido, la heterogeneidad en la cobertura geográfica del muestreo impide establecer inferencias concluyentes acerca de patrones de dispersión o predominio regional de los distintos sublinajes.

Características demográficas y epidemiológicas asociadas a las secuencias B_{CAR}

En cuanto al clado B_{CAR} (n = 18), el análisis de los metadatos disponibles muestra que 11 secuencias (61,1%) corresponden a pacientes de sexo masculino y 7 (38,9%) a pacientes de sexo femenino. Respecto a la vía de transmisión, 10 secuencias (55,6%) se asocian a transmisión HTX, tres (16,7%) HSH, cuatro (22,2%) se encuentran categorizadas como “sexual” sin mayor especificación y una (5,6%) no presenta información disponible.

En relación con el lugar de adquisición del virus, ocho secuencias (44,4%) corresponden a la región metropolitana, mientras que dos (11,1%) se ubican en la región noreste y una (5,6%) en la región este. Asimismo, dos secuencias (11,1%) corresponden a infecciones adquiridas en Argentina, y cinco (27,8%) no cuentan con información geográfica disponible.

Dado el bajo número de secuencias disponibles para este clado, estos resultados deben interpretarse con cautela y no permiten realizar inferencias robustas sobre patrones epidemiológicos o de distribución geográfica.

5.3. Análisis evolutivos e inferencia demográfica

Se infirieron filogenias bayesianas calibradas en el tiempo para los clados B_{PAN} de América del Sur que involucran redes de transmisión uruguayas (B_{PAN-UY}, B_{PAN-SAM}, B_{PAN-AR-II}, B_{PAN-AR-III}, B_{PAN-BR-II} y B_{PAN-BR-III}). La tasa media de sustitución estimada para el conjunto de análisis fue de $1,8 \times 10^{-3}$ sustituciones/sitio/año (HPD 95%: $1,7 \times 10^{-3} - 2,1 \times 10^{-3}$). El coeficiente de variación de las tasas específicas por rama fue de 0,36 (HPD 95%: 0,26–0,47), lo que es consistente con una heterogeneidad sustancial en las tasas evolutivas

entre linajes y respalda la utilización de un modelo de reloj molecular relajado no correlacionado con distribución lognormal.

Los resultados obtenidos sugieren que las introducciones de linajes B_{PAN} en Uruguay sucedieron entre 1978 (HPD 95%: 1972–1985) y 2013 (HPD 95%: 2008–2016). La mayoría de las introducciones inferidas entre 1995 y 2015 no se asociaron con el establecimiento de clados de transmisión local sostenidos, en concordancia con lo observado en el análisis filogenético de máxima verosimilitud. El clado más antiguo inferido (tMRCA medio: 1978; HPD 95%: 1972–1985) correspondió a un clado de transmisión mediano dentro de la filogenia B_{PAN-UY} . Las estimaciones medias del tMRCA para los dos principales subclados uruguayos, $B_{PAN-UY-I}$ y $B_{PAN-UY-II}$, fueron 1985 (HPD 95%: 1978–1989) y 1987 (HPD 95%: 1978–1988), respectivamente, lo que es consistente con la emergencia independiente de ambos subclados hacia mediados y fines de la década de 1980 (Figura 13, Tablas 4 y 5).

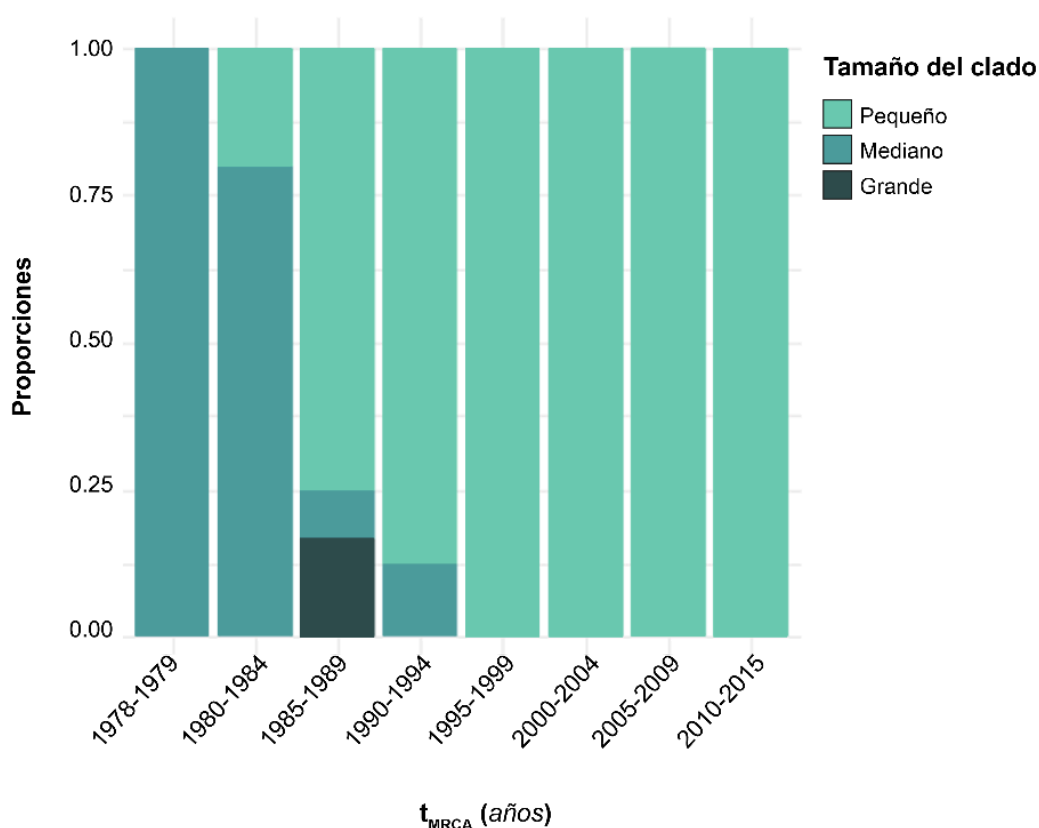


Figura 13. Distribución temporal de los clados de transmisión de B_{PAN} según su tiempo estimado al ancestro común más reciente (tMRCA). Los clados fueron agrupados en intervalos discretos de acuerdo con las estimaciones bayesianas de tMRCA. Las barras apiladas muestran la proporción de clados pequeños ($2 \leq n < 10$), medianos ($10 \leq n < 30$) y grandes ($n \geq 30$).

Tabla 4. Tiempo del ancestro común más reciente (tMRCA), origen geográfico estimado y número de secuencias (N) de los clados uruguayos identificados dentro de B_{PAN-SAM}, B_{PAN-AR} y B_{PAN-BR}.

Sub-linaje	Clado	N	t _{MRCA}	Origen geográfico
B_{PAN-SAM}				
	B _{PAN-SAM-I}	4	1987 (1981 – 1994)	Brasil
	B _{PAN-SAM-II}	2	2004 (1997 – 2011)	Brasil
	B _{PAN-SAM-III}	3	1990 (1981 – 1998)	Brasil
	B _{PAN-SAM-IV}	2	2002 (1996 – 2008)	Brasil
	B _{PAN-SAM-V}	2	2005 (1998 – 2012)	Brasil
	B _{PAN-SAM-VI}	2	1996 (1987 – 2003)	Brasil
	B _{PAN-SAM-VII}	10	1982 (1975 – 1988)	Brasil
	B _{PAN-SAM-VIII}	2	2013 (2008 – 2016)	Argentina
	B _{PAN-SAM-IX}	2	2005 (1999 – 2012)	Argentina
	B _{PAN-SAM-X}	2	1984 (1974 – 1993)	Brasil
	B _{PAN-SAM-XI}	2	1997 (1989 – 2005)	Brasil
	B _{PAN-SAM-XII}	13	1980 (1973 – 1987)	Brasil
	B _{PAN-SAM-XIII}	2	1990 (1982 – 1999)	Brasil
	B _{PAN-SAM-XIV}	5	1989 (1982 – 1994)	Brasil
	B _{PAN-SAM-XV}	2	1986 (1978 – 1993)	Brasil
B_{PAN-AR-II}				
	B _{PAN-AR-II}	12	1990 (1984 – 1997)	Argentina
B_{PAN-AR-III}				
	B _{PAN-AR-III}	4	1988 (1982 – 1995)	Argentina
B_{PAN-BR-II}				
	B _{PAN-BR-II}	4	1988 (1981 – 1995)	Brasil
B_{PAN-BR-III}				
	B _{PAN-BR-III-I}	8	1990 (1983 – 1996)	Brasil
	B _{PAN-BR-III-II}	3	2005 (1999 – 2010)	Brasil

Tabla 5. Tiempo del ancestro común más reciente (tMRCA) y número de secuencias (N) de los clados identificados como B_{PAN-UY} .

Sub-linaje	Clado	N	t _{MRCA}
BPAN-UY			
	$B_{PAN-UY-I}$	37	1985 (1978 – 1989)
	$B_{PAN-UY-II}$	30	1987 (1978 – 1988)
	$B_{PAN-UY-III}$	4	1993 (1986 – 1999)
	$B_{PAN-UY-IV}$	2	2010 (2009 – 2011)
	$B_{PAN-UY-V}$	3	1986 (1978 – 1993)
	$B_{PAN-UY-VI}$	4	1988 (1982 – 1995)
	$B_{PAN-UY-VII}$	4	1987 (1980 – 1994)
	$B_{PAN-UY-VIII}$	10	1982 (1974 – 1988)
	$B_{PAN-UY-IX}$	2	1998 (1991 – 2004)
	$B_{PAN-UY-X}$	2	2002 (1998 – 2006)
	$B_{PAN-UY-XI}$	6	1992 (1985 – 1998)
	$B_{PAN-UY-XII}$	10	1989 (1983 – 1995)
	$B_{PAN-UY-XIII}$	3	1999 (1992 – 2006)
	$B_{PAN-UY-XIV}$	18	1980 (1972 – 1987)
	$B_{PAN-UY-XV}$	2	2011 (2006 – 2016)
	$B_{PAN-UY-XVI}$	12	1978 (1972 – 1984)
	$B_{PAN-UY-XVII}$	5	1998 (1991 – 2006)
	$B_{PAN-UY-XVIII}$	2	2001 (1995 – 2008)
	$B_{PAN-UY-XIX}$	2	2005 (1993 – 2013)
	$B_{PAN-UY-XX}$	3	2005 (1999 – 2010)
	$B_{PAN-UY-XXI}$	2	1994 (1986 – 2002)
	$B_{PAN-UY-XXII}$	2	2009 (2007 – 2011)
	$B_{PAN-UY-XXIII}$	10	1979 (1973 – 1986)
	$B_{PAN-UY-XXIV}$	10	1978 (1972 – 1985)
	$B_{PAN-UY-XXV}$	2	2003 (1996 – 2009)
	$B_{PAN-UY-XXVI}$	2	2002 (1996 – 2009)
	$B_{PAN-UY-XXVII}$	4	1999 (1992 – 2006)
	$B_{PAN-UY-XXVIII}$	3	1988 (1981 – 1993)
	$B_{PAN-UY-XXIX}$	2	1992 (1985 – 2000)

Con el fin de caracterizar en mayor profundidad la historia evolutiva y demográfica de la epidemia B_{PAN} en Uruguay, se realizó un análisis filodinámico conjunto de $B_{PAN-UY-I}$ y $B_{PAN-UY-II}$. Estos sublinajes fueron seleccionados por ser los de mayor tamaño y densidad de muestreo, lo que permite estimar con mayor sustento los cambios temporales en parámetros poblacionales y epidemiológicos (Figura 14a).

La dinámica temporal del N_e fue inferida mediante BSKG, mientras que la variación temporal del R_e se estimó mediante un modelo BDSKY. El análisis BSKG mostró una fase sostenida de crecimiento del conjunto $B_{PAN-UY-I}/B_{PAN-UY-II}$, iniciada a mediados de la década de 1980 y extendida aproximadamente hasta 2010, reflejada en un incremento marcado del N_e . Por su parte, el modelo BDSKY estimó el R_e en ocho intervalos temporales equidistantes entre 1981 y 2021. Los valores de R_e fueron mayores a 1 durante las etapas iniciales de la epidemia, correspondientes al período comprendido entre fines de la década de 1980 y comienzos de los años 2000. Posteriormente, descendieron hacia valores cercanos al umbral epidémico entre mediados de la década de 2000 y 2015, y se mantuvieron por debajo de 1 en los intervalos mas recientes, lo que sugiere una reducción en la intensidad de transmisión de estos sublinajes (Figura 14b).

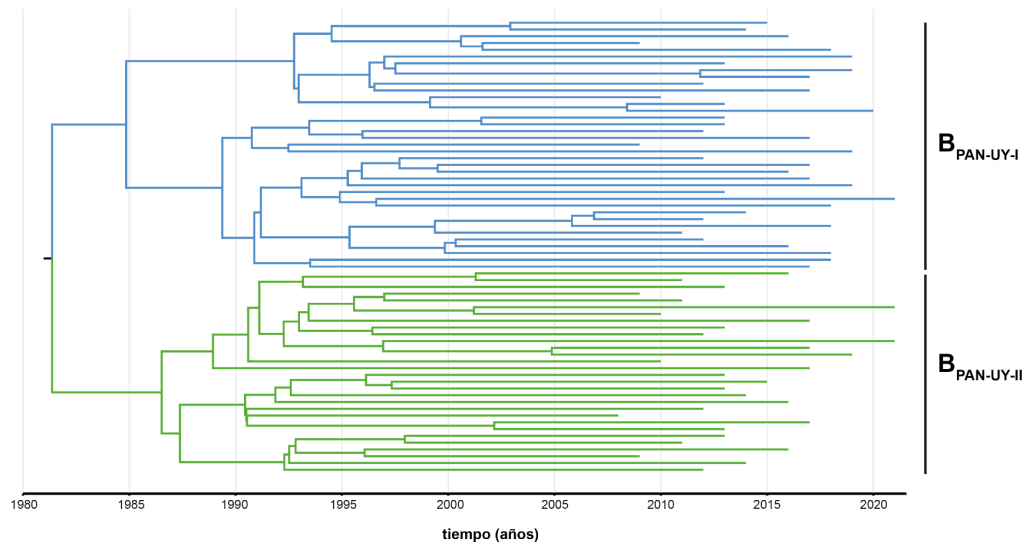
En conjunto, los análisis BSKG y BDSKY muestran un patrón temporal concordante, compatible con tres fases principales en la dinámica de B_{PAN-UY} . La primera corresponde a una etapa de expansión sostenida entre aproximadamente 1985 y 2005, caracterizada por valores de R_e consistentemente superiores a 1 y coincidente con el aumento del N_e . La segunda fase, entre mediados de la década de 2000 y 2015, se caracteriza por valores de R_e cercanos a 1, lo que sugiere una estabilidad de la transmisión. Finalmente, desde 2015 hasta al menos 2022, los valores de R_e inferiores a 1 son compatibles con una disminución sostenida de la transmisión de estos sublinajes.

Cabe destacar que las estimaciones de N_e obtenidas mediante BSKG se interrumpen alrededor de 2010. Esto se debe a la naturaleza retrospectiva de los modelos coalescentes, cuya capacidad para estimar cambios recientes en el N_e depende de la densidad de eventos de coalescencia en el árbol. En este caso, la baja densidad de nodos internos en las ramas más recientes reduce la señal filogenética disponible para inferir el N_e durante el tramo final del período analizado.

Finalmente, el análisis de las distancias internodales en los árboles de MCCT indicó que aproximadamente la mitad de los eventos filogenéticamente inferidos, compatibles con posibles eventos de transmisión secundaria (49,3%) ocurrieron dentro de los primeros dos

años posteriores a la infección. Un 29,2% adicional se produjo entre los 2 y 5 años post-infección, mientras que el 21,5% restante tuvo lugar más de 5 años después de la infección inicial estimada (Figura 15).

a.



b.

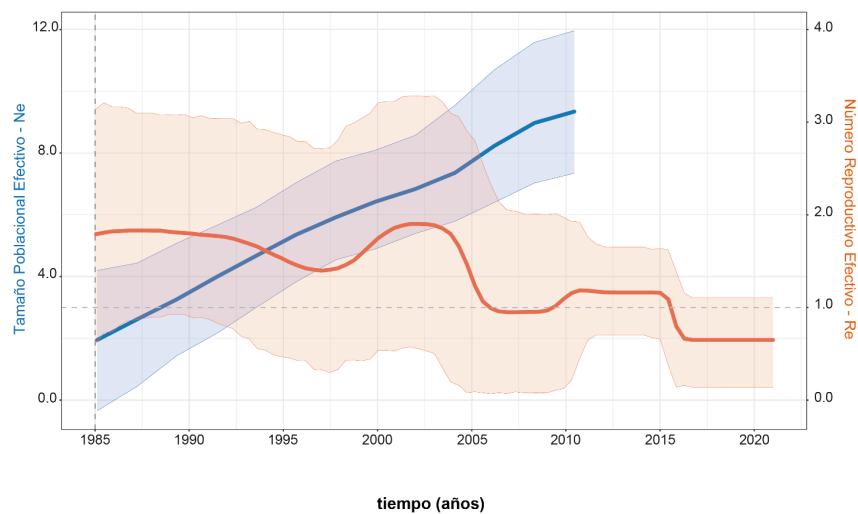


Figura 14. Filogenia calibrada en el tiempo y reconstrucción filodinámica conjunta de los principales subclados uruguayos de B_{PAN} . a) Filogenia bayesiana calibrada en el tiempo de los dos subclados de B_{PAN} , $B_{PAN-UY-I}$ y $B_{PAN-UY-II}$, inferida en BEAST. La longitud de las ramas es proporcional al tiempo (años), según se indica en el eje horizontal, y los colores distinguen ambos subclados. b) Reconstrucción filodinámica conjunta de $B_{PAN-UY-I}$ y $B_{PAN-UY-II}$. La dinámica temporal del tamaño poblacional efectivo (N_e) se estimó mediante el modelo Bayesian SkyGrid (BSKG) (línea azul, con su HPD del 95% sombreado; eje y izquierdo), mientras que la variación temporal del Re se estimó mediante el modelo Birth–Death Skyline (BDSKY) (línea naranja, con su HPD del 95% sombreado; eje y derecho). La línea horizontal discontinua indica el umbral epidémico ($Re = 1$).

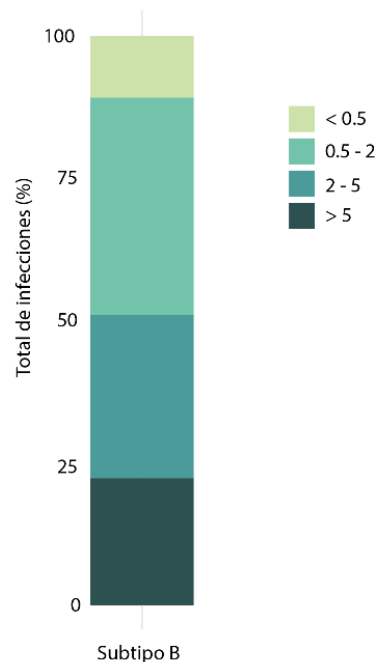


Figura 15. Distribución de los eventos filogenéticamente inferidos, compatibles con posibles eventos de transmisión secundaria, según el tiempo aproximado desde la infección. Los eventos fueron estimados a partir de las distancias internodales en los árboles de máxima credibilidad de clado (MCCT) de clados de transmisión uruguayos de B_{PAN} más grandes. La barra apilada muestra la proporción de eventos asignados a distintos intervalos aproximados al tiempo posterior a la infección estimada: <0,5; 0,5–2; 2–5 y >5 años.

5.4. Origen geográfico de la epidemia B_{PAN} en Uruguay

La inferencia del origen geográfico probable de los clados de transmisión uruguayos se realizó considerando los principales sublinajes sudamericanos de B_{PAN} que contribuyen a la epidemia nacional. En el caso de $B_{PAN-SAM}$, que incluye secuencias de distintos países de la región, la reconstrucción del estado geográfico ancestral se realizó mediante PastML. Para los sublinajes definidos previamente como de origen mayoritariamente argentino o brasileño ($B_{PAN-AR-II}$, $B_{PAN-AR-III}$, $B_{PAN-BR-II}$, $B_{PAN-BR-III}$ y $B_{PAN-BR-IV}$), el origen probable de los clados uruguayos se asignó a partir de su inserción filogenética dentro del sublinaje regional correspondiente y de la composición geográfica de las secuencias de referencia.

Bajo este criterio, Brasil fue identificado como la principal fuente regional de los clados B_{PAN} presentes en Uruguay. De los 20 clados de transmisión uruguayos identificados dentro de estos linajes, 16 (80,0%) fueron asignados a un origen brasileño, mientras que 4 (20,0%) fueron asignados a un origen argentino (Tabla 4).

Un patrón similar se observó en las introducciones independientes que no dieron lugar a clados de transmisión local detectables. Entre las 61 secuencias uruguayas

correspondientes a este tipo de eventos, la reconstrucción del estado geográfico ancestral asignó 45 (73,8%) a una localización ancestral brasileña y 16 (26,2%) a una localización argentina.

En conjunto, estos resultados indican que Brasil representa la principal fuente regional de introducciones de linajes B_{PAN} hacia Uruguay, tanto entre aquellas que dieron lugar a clados de transmisión local como entre aquellas que permanecieron como eventos esporádicos.

6. Discusión

En América Latina, las epidemias de VIH-1B no se distribuyen como una población viral homogénea, sino que muestran una marcada estructuración geográfica y epidemiológica. Diversos estudios han evidenciado múltiples introducciones virales en sistemas de transmisión parcialmente aislados, con compartimentalización geográfica a nivel de redes y una importante heterogeneidad en la capacidad de establecimiento y expansión local de los distintos linajes (Pagán & Holguín, 2013; Mir et al., 2015). Los hallazgos de este trabajo confirman que la epidemia de VIH-1B en Uruguay también está fuertemente condicionada por introducciones múltiples y por diferencias en el éxito de expansión local de los linajes introducidos. Un aspecto central de los resultados obtenidos es la elevada frecuencia de introducciones independientes que no logran consolidar una transmisión local sostenida. A pesar de la introducción recurrente de una amplia diversidad de linajes B_{PAN} circulantes en la región, con una fuerte influencia de variantes provenientes de Argentina y Brasil, la mayoría de estos eventos correspondió a introducciones esporádicas o a clados de pequeño tamaño, lo que sugiere una capacidad limitada de establecimiento y expansión dentro de la epidemia nacional.

Este patrón fue particularmente evidente en linajes que han experimentado una expansión epidémica considerable en otros países de América Latina, como $B_{PAN-BR-IV}$, $B_{PAN-CAM}$ y B_{PAN-PE} , los cuales en Uruguay fueron detectados exclusivamente como introducciones esporádicas sin evidencia de transmisión secundaria. Incluso entre los clados sudamericanos más frecuentemente introducidos ($B_{PAN-SAM}$, $B_{PAN-AR-II}$, $B_{PAN-AR-III}$, $B_{PAN-BR-II}$ y $B_{PAN-BR-III}$) y los linajes mexicanos ($B_{PAN-MX-I}$ y $B_{PAN-MX-II}$), la amplificación local fue, en general, limitada. Este hallazgo sugiere la existencia de restricciones estructurales o epidemiológicas que condicionan el establecimiento exitoso de nuevas cadenas de transmisión tras su introducción en el país.

Por su parte, los linajes B_{CAR} contribuyeron de manera marginal a la epidemia uruguaya y no mostraron evidencia de transmisión local sostenida en el conjunto de secuencias analizado. Este patrón es consistente con lo reportado para otros países del Cono Sur, donde la contribución epidemiológica de B_{CAR} ha sido limitada en comparación con B_{PAN} (Cabello et al., 2015). La escasa expansión de B_{CAR} en Uruguay podría estar relacionada con su baja circulación en las regiones de mayor intercambio epidemiológico con el país. En efecto, la circulación de B_{CAR} se ha descrito principalmente en Surinam, Guayana Francesa, República Dominicana y otros países de América Central y el Caribe (Cabello et al., 2014; Bello et al., 2019), regiones geográficamente distantes de Uruguay y con menor

conectividad directa con la epidemia local. Si bien B_{CAR} también ha sido identificado en Brasil, país con el que Uruguay comparte una extensa frontera, su distribución dentro del territorio brasileño es marcadamente heterogénea. Este linaje se concentra principalmente en estados del norte y noreste de Brasil, donde se han reportado frecuencias elevadas, como 41% en Roraima y 14% en Amazonas y Maranhão. En cambio, en las regiones Sur, Sudeste y Centro-Oeste, más próximas a los principales circuitos de intercambio regional con Uruguay, la frecuencia de B_{CAR} desciende a valores inferiores al 10% (Divino et al., 2016; Arantes et al., 2019, 2022). Por lo tanto, la baja frecuencia de B_{CAR} en las regiones brasileñas más conectadas con Uruguay, junto con la distancia geográfica respecto de sus principales focos de circulación, podría haber limitado su introducción recurrente y su establecimiento sostenido en la epidemia uruguaya.

En este contexto de introducciones frecuentes, pero mayormente limitadas en términos de expansión local, $B_{PAN-UY-I}$ y $B_{PAN-UY-II}$ se destacan como los únicos subclados uruguayos con evidencia de persistencia epidémica a largo plazo. Ambos constituyen los clados de transmisión más grandes y duraderos identificados en Uruguay y cumplen un papel comparable, a escala nacional, al de los grandes clados fundadores B_{PAN} que sostienen epidemias prolongadas en otros países de América Latina. Las estimaciones temporales derivadas de los análisis filogenéticos indican que ambos linajes emergieron de manera independiente entre mediados y fines de la década de 1980. Este período coincide con la fase inicial de expansión del subtipo B en el Cono Sur y con la diversificación regional de los linajes B_{PAN} (Bello et al., 2006; Mir et al., 2015; Junqueira et al., 2016; Mendoza et al., 2014). En este sentido, su establecimiento temprano, previo a la disponibilidad de intervenciones terapéuticas efectivas, pudo haber favorecido su expansión y persistencia a largo plazo.

En contraste, las introducciones virales ocurridas después de mediados de la década de 1990 mostraron una probabilidad considerablemente menor de establecer transmisión sostenida, dando lugar predominantemente a clados pequeños o a infecciones esporádicas. Esta asimetría temporal es consistente con lo descrito para otras epidemias sudamericanas de subtipo B, en las cuales los clados más antiguos presentan fases de crecimiento más prolongadas y mayores tamaños poblacionales efectivos en comparación con linajes introducidos más recientemente (Mir et al., 2015; Mendoza et al., 2014). Estos patrones respaldan un régimen filodinámico en el cual el establecimiento temprano de determinados linajes y los efectos de contingencia histórica desempeñan un papel central en la configuración de la estructura epidémica a largo plazo.

Los análisis filodinámicos de los dos linajes uruguayos dominantes de B_{PAN} indican que la epidemia de VIH-1B en el país ha alcanzado una fase de madurez. Las trayectorias

inferidas se caracterizan por una expansión sostenida desde mediados de la década de 1980 hasta aproximadamente 2010, seguida por una etapa de estabilización y, más recientemente, señales consistentes con una disminución en la intensidad de la transmisión. Este patrón es congruente con lo reportado para epidemias de subtipo B de larga data en América Latina (Mendoza et al., 2014; Mir et al., 2015; Arantes et al., 2022). En este contexto, las estimaciones del R_e se mantuvieron persistentemente por encima del umbral epidémico durante varias décadas, comenzando a aproximarse a valores críticos entre mediados de los años 2000 y mediados de la década de 2010, y descendiendo posteriormente a valores subcríticos ($R_e < 1$) en los últimos años. Esta transición temporal coincide con el período posterior a la universalización del acceso a la TAR en Uruguay, iniciada en 1996–1997 (MSP, 2001), y sugiere un impacto gradual del tratamiento sobre la transmisión a nivel poblacional, en concordancia con el marco conceptual de tratamiento como prevención (Granich et al., 2009; Montaner et al., 2010). En este sentido, la disminución del R_e podría reflejar que se alcanzó progresivamente una masa crítica de pacientes con supresión viral.

Asimismo, la divergencia parcial observada entre las reconstrucciones basadas en modelos coalescentes (BSKG) y aquellas basadas en procesos de nacimiento–muerte (BDSKY) en los períodos más recientes, resulta esperable en escenarios epidémicos avanzados. Este desacople refleja la menor sensibilidad de los enfoques coalescentes para capturar dinámicas de transmisión cercanas al presente, particularmente cuando la incidencia disminuye y los tiempos de coalescencia se vuelven más prolongados (Mir et al., 2018; Arantes et al., 2022).

El análisis de las distancias internodales en los MCCT, utilizadas como una aproximación indirecta a la separación temporal entre eventos de diversificación filogenéticamente inferidos, mostró una distribución heterogénea de los intervalos entre eventos sucesivos dentro de las filogenias inferidas. Casi la mitad (49,3%) de los intervalos estimados fueron inferiores a dos años, y 78,5% no superó los cinco años. Estos resultados indican la coexistencia de intervalos cortos y prolongados entre eventos filogenéticamente inferidos, lo que sugiere una dinámica evolutiva caracterizada tanto por episodios de diversificación relativamente rápida como por la permanencia durante períodos prolongados de determinadas ramas o redes filogenéticas en la epidemia uruguaya.

La predominancia de intervalos cortos podría ser compatible con períodos de diversificación viral relativamente rápida dentro de algunas cadenas. En términos epidemiológicos, este patrón podría vincularse con la heterogeneidad característica de la transmisión del VIH, en la que determinados períodos o contextos concentran una mayor cantidad de eventos de

transmisión (Pilcher et al., 2004; Hollingsworth et al., 2008). Sin embargo, dado que este análisis no reconstruye pares de transmisión ni estima el momento de infección de los individuos, los intervalos observados no deben interpretarse como el tiempo transcurrido desde la infección del caso fuente hasta un evento de transmisión secundaria. La presencia de una proporción relevante de intervalos prolongados (>5 años) podría reflejar la persistencia temporal de determinadas ramas filogenéticas, pero también el efecto de un muestreo incompleto, la existencia de intermediarios no muestreados o una baja densidad de secuencias en ciertos períodos. En otros contextos epidemiológicos, la persistencia prolongada de linajes del subtipo B se ha relacionado con una combinación de procesos epidemiológicos, asistenciales y de muestreo (Arantes et al., 2021). No obstante, este análisis no permite discriminar entre estos mecanismos ni atribuir los intervalos observados a alguno de ellos en particular.

En conjunto, estos resultados sugieren que la epidemia uruguaya combina períodos de diversificación relativamente rápida con la persistencia temporal de determinadas ramas o agrupamientos filogenéticos. En el contexto de una epidemia madura y de amplia disponibilidad de TAR, este patrón podría reflejar la interacción entre períodos de mayor concentración de eventos de diversificación y la permanencia prolongada de algunos linajes, modulada por factores epidemiológicos, estructurales y de muestreo, hipótesis que deberá ser evaluada mediante estudios que integren información clínica, epidemiológica y de seguimiento longitudinal.

Debe considerarse, además, que estas estimaciones derivan de reconstrucciones filogenéticas realizadas a partir de secuencias obtenidas en el contexto de estudios de resistencia a TARV, y no de un muestreo poblacional diseñado específicamente para reconstruir cadenas de transmisión. Este sesgo de selección, junto con la cobertura incompleta de la epidemia y la incertidumbre inherente a la reconstrucción filogenética, debe tenerse en cuenta al interpretar estos intervalos. Finalmente, las reconstrucciones filogeográficas indican que la epidemia uruguaya de VIH-1B está fuertemente influida por introducciones transfronterizas recurrentes, principalmente desde Brasil y, en menor medida, desde Argentina. Este patrón es consistente con un régimen de dispersión característico de América Latina, dominado por la propagación transfronteriza de corta distancia, la fuerte compartimentalización geográfica y un limitado intercambio viral entre regiones distantes (Pagán & Holguín, 2013; Mir et al., 2015). A su vez, contrasta con lo observado en epidemias europeas de subtipo B, donde redes transnacionales densamente interconectadas y frecuentes intercambios virales entre países generan una estructura espacial más débil y filogenias altamente entremezcladas (Paraskevis et al., 2009).

Más allá de la dinámica filogenética y filogeográfica, la composición demográfica de la muestra aporta una dimensión epidemiológica relevante para interpretar la circulación del VIH-1B en Uruguay. La predominancia de individuos de sexo masculino en la muestra de VIH-1B analizada es particularmente relevante, considerando que esta muestra proviene de la base GENOUY-VIH, la cual presenta sesgos metodológicos de muestreo que, en principio, podrían favorecer una sobrerrepresentación de mujeres, de acuerdo con los criterios discutidos en la sección 4.2. de este trabajo. La asociación del sexo masculino a este linaje se ve respaldada por análisis comparativos previos del dataset global (Podchibiakin et al., en revisión, Anexo 1) donde mediante pruebas de Fisher y test de proporciones (Z) se demostró una asociación significativa ($p < 0,001$). Específicamente, el VIH-1B exhibe una proporción significativamente mayor de hombres y una proporción significativamente menor de mujeres en comparación con otras variantes de relevancia epidemiológica para el país como los FRC12_BF y FRC38_BF1, que muestran perfiles más desplazados hacia poblaciones femeninas y transgénero. Por lo tanto, la predominancia masculina distintiva del subtipo B refuerza la hipótesis de que este linaje circula predominantemente en redes de transmisión conformadas mayoritariamente por hombres, en concordancia con escenarios internacionales donde este subtipo se asocia históricamente a poblaciones de HSH (Volz et al., 2018; Le Vu et al., 2019; Shiino et al., 2020). No obstante, esta interpretación debe tomarse con cautela, dado que los metadatos disponibles no permiten caracterizar con suficiente detalle las vías de transmisión de todos los individuos incluidos en la muestra.

7. Limitaciones

Las limitaciones de este estudio se relacionan principalmente con la calidad y representatividad de los datos disponibles. Si bien el formulario asociado al genotipado de resistencia a antirretrovirales constituye una herramienta valiosa para la recolección de información demográfica y epidemiológica, presenta desafíos importantes. En particular, la falta de completitud en ciertas variables y la ambigüedad en la definición de otras afectan la consistencia de los datos y limitan la profundidad de los análisis.

En relación con la vía de transmisión, la elevada proporción de registros categorizados como “sexual” sin discriminación entre HSH y HTX, así como la ausencia de información en una fracción no despreciable de las secuencias, impiden caracterizar de manera robusta la composición epidemiológica de las redes de transmisión en Uruguay. En consecuencia, aunque la distribución observada por sexo es consistente con patrones epidemiológicos descritos para el subtipo B en otros contextos, no es posible confirmar ni cuantificar la contribución relativa de las distintas vías de transmisión en la epidemia local.

De manera similar, las limitaciones en la información geográfica condicionan la interpretación de la estructura espacial de la epidemia. La marcada concentración de secuencias en la región metropolitana, junto con la escasa representación de otras regiones del país y la proporción de registros sin información detallada, introduce un sesgo de muestreo que impide realizar inferencias robustas sobre la distribución de los linajes virales a nivel nacional.

Adicionalmente, el conjunto de datos analizado deriva del genotipado rutinario de resistencia a antirretrovirales en el marco del sistema nacional de vigilancia del VIH, lo que implica que no constituye una muestra aleatoria de la población infectada. Está enriquecido en individuos vinculados al sistema de atención y con viremia detectable, ya que las pruebas de resistencia se realizan principalmente bajo indicaciones clínicas específicas (Flieller & Cabrera, 2019). Además, debido a que el muestreo suele ocurrir años después de la infección, la incertidumbre temporal puede afectar las estimaciones del origen de los linajes y de las dinámicas recientes de transmisión. La ausencia de individuos con infecciones recientes por el VIH, también impone una limitación al momento de capturar cambios epidemiológicos más cercanos en tiempo.

Desde el punto de vista metodológico, los análisis se restringieron a una región parcial del gen *pol*, lo que limita la resolución filogenética y puede conducir a la subestimación del tamaño de los clados, a la clasificación errónea de eventos de introducción y a una

detección incompleta de recombinación. A esto se suma el muestreo incompleto de las cadenas de transmisión, que puede generar fragmentación filogenética, sesgar la interpretación del número de introducciones independientes y sobreestimar los intervalos entre eventos de diversificación filogenética inferidos.

En conjunto, estas limitaciones no solo afectan la interpretación de los resultados obtenidos, sino que también ponen de manifiesto desafíos relevantes para la implementación de estrategias de vigilancia genómica en el país. La disponibilidad de datos incompletos o poco resolutivos, tanto a nivel epidemiológico como geográfico, restringe la capacidad de integrar información genética con variables contextuales clave, lo que dificulta la identificación precisa de redes de transmisión, la detección temprana de nuevos linajes y la caracterización de patrones de dispersión. En este sentido, el fortalecimiento de los sistemas de recolección de datos, la estandarización de los metadatos asociados a las secuencias virales, y la inclusión de secuencias provenientes de individuos diagnosticados recientemente, resultan fundamentales para potenciar el uso de la vigilancia genómica como herramienta para el monitoreo y control de la epidemia de VIH en Uruguay.

8. Conclusiones

Los análisis realizados en este trabajo muestran que la epidemia de VIH-1 subtipo B en Uruguay se configura como un sistema filodinámico maduro y altamente estructurado, caracterizado por una elevada frecuencia de introducciones virales independientes, la mayoría de las cuales no logra establecer transmisión sostenida. En este contexto, la persistencia epidémica a largo plazo no parece depender de un flujo continuo de nuevas introducciones, sino de la continua transmisión de un número reducido de linajes fundadores introducidos décadas atrás, particularmente los sublinajes B_{PAN-UY-I} y B_{PAN-UY-II}.

Desde una perspectiva espacial, la epidemia uruguaya se encuentra fuertemente influida por introducciones transfronterizas, principalmente desde Brasil y Argentina, en concordancia con un régimen regional de dispersión dominado por la propagación de corta distancia y la alta compartimentalización geográfica.

Los análisis temporales indican que la epidemia transitó desde una fase de expansión sostenida hacia un escenario más reciente de estabilización, con señales de disminución en la intensidad de la transmisión. No obstante, el hecho de que la transmisión no se haya interrumpido completamente sugiere que, a pesar de los avances en el acceso a la TAR, persisten brechas en el continuo diagnóstico–tratamiento que permiten la circulación y continua diseminación local del virus.

El análisis de los intervalos entre eventos filogenéticamente inferidos mostró la coexistencia de intervalos cortos y prolongados dentro de los linajes analizados, compatible con una dinámica de propagación caracterizada por episodios de diversificación relativamente rápida y una persistencia temporal de algunos linajes virales. Si bien esta aproximación indirecta debe interpretarse con cautela, el patrón observado resalta la necesidad de fortalecer tanto el diagnóstico precoz como la continuidad de la atención y el acceso efectivo a la TAR.

En términos demográficos, la predominancia de secuencias provenientes de individuos de sexo masculino es consistente con patrones epidemiológicos descritos para el subtipo B en otros contextos, y sugiere una circulación relevante en redes de transmisión conformadas mayoritariamente por hombres.

En conjunto, estos resultados constituyen una primera aproximación a la caracterización filodinámica de la epidemia de VIH-1 subtipo B en Uruguay, y proporcionan un marco de referencia para futuros estudios orientados a comprender con mayor profundidad la dinámica de transmisión y evolución viral a nivel local. Asimismo, estos resultados resaltan

el potencial de la vigilancia genómica como herramienta para complementar los sistemas tradicionales de vigilancia epidemiológica, monitorear la introducción de nuevos linajes y orientar estrategias de intervención más focalizadas.

9. Bibliografía

Abecasis, A. B., & Vandamme, A.-M. (2014). HIV-1 subtypes. En M. S. S. (Ed.), *Encyclopedia of AIDS*. Springer. https://doi.org/10.1007/978-1-4614-9610-6_130-2

Alexiev, I., & Dimitrova, R. (2025). The Origins and Genetic Diversity of HIV-1: Evolutionary Insights and Global Health Perspectives. *International journal of molecular sciences*, 26(22), 10909. <https://doi.org/10.3390/ijms262210909>

Anisimova, M., & Gascuel, O. (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Systematic biology*, 55(4), 539–552. <https://doi.org/10.1080/10635150600755453>

Arantes, I., Esashika Crispim, M. A., Nogueira da Guarda Reis, M., Martins Araújo Stefani, M., & Bello, G. (2019). Reconstructing the Dissemination Dynamics of the Major HIV-1 Subtype B Non-Pandemic Lineage Circulating in Brazil. *Viruses*, 11(10), 909. <https://doi.org/10.3390/v11100909>

Arantes, I., Bello, G., Darcissac, E., Lacoste, V., & Nacher, M. (2021). Using phylogenetic surveillance and epidemiological data to understand the HIV-1 transmission dynamics in French Guiana. *AIDS*, 35(6), 979–984. <https://doi.org/10.1097/QAD.0000000000002817>

Arantes, I., Gräf, T., Andrade, P., Oliveira Chaves, Y., Guimarães, M. L., & Bello, G. (2022). Dissemination Dynamics of HIV-1 Subtype B Pandemic and Non-pandemic Lineages Circulating in Amazonas, Brazil. *Frontiers in microbiology*, 13, 835443. <https://doi.org/10.3389/fmicb.2022.835443>

Aulicino, P. C., Gómez-Carrillo, M., Bello, G., Rocco, C., Mangano, A., Carr, J., Sen, L., & Foley, B. (2012). Characterization of full-length HIV-1 CRF17_BF genomes and comparison to the prototype CRF12_BF strains. *Infection, Genetics and Evolution*, 12(2), 443–447. <https://doi.org/10.1016/j.meegid.2012.01.003>

Ayoub, A., Akoua-Koffi, C., Calvignac-Spencer, S., Esteban, A., Locatelli, S., Li, H., Li, Y., Hahn, B. H., Delaporte, E., Leendertz, F. H., & Peeters, M. (2013). Evidence for continuing cross-species transmission of SIVsmm to humans: characterization of a new HIV-2 lineage in rural Côte d'Ivoire. *AIDS*, 27(15), 2488–2491. <https://doi.org/10.1097/01.aids.0000432443.22684.50>

Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W., & Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, 220(4599), 868–871. <https://doi.org/10.1126/science.6189183>

Bbosa, N., Kaleebu, P., & Ssemwanga, D. (2019). HIV subtype diversity worldwide. *Current opinion in HIV and AIDS*, 14(3), 153–160. <https://doi.org/10.1097/COH.0000000000000534>

Bello, G., Guimarães, M. L., & Morgado, M. G. (2006). Evolutionary history of HIV-1 subtype B and F infections in Brazil. *AIDS*, 20(5), 763–768. <https://doi.org/10.1097/01.aids.0000216377.84313.52>

Bello, G., Aulicino, P. C., Ruchansky, D., Guimarães, M. L., Lopez-Galindez, C., Casado, C., Chiparelli, H., Rocco, C., Mangano, A., Sen, L., & Morgado, M. G. (2010). Phylodynamics of HIV-1 circulating recombinant forms 12_BF and 38_BF in Argentina and Uruguay. *Retrovirology*, 7(22). <https://doi.org/10.1186/1742-4690-7-22>

Bello, G., Arantes, I., Lacoste, V., Ouka, M., Boncy, J., Césaire, R., Liautaud, B., Nacher, M., & Dos Santos, G. (2019). Phylogeographic Analyses Reveal the Early Expansion and Frequent Bidirectional Cross-Border Transmissions of Non-pandemic HIV-1 Subtype B Strains in Hispaniola. *Frontiers in microbiology*, 10, 1340. <https://doi.org/10.3389/fmicb.2019.01340>

Cabello, M., Junqueira, D. M., & Bello, G. (2015). Dissemination of nonpandemic Caribbean HIV-1 subtype B clades in Latin America. *AIDS*, 29(4), 483–492. <https://doi.org/10.1097/QAD.0000000000000552>

Cabello, M., Mendoza, Y., & Bello, G. (2014). Spatiotemporal dynamics of dissemination of non-pandemic HIV-1 subtype B clades in the Caribbean region. *PloS one*, 9(8), e106045. <https://doi.org/10.1371/journal.pone.0106045>

Carballal, G., & Oubiña, J. R. (2013). *Virología médica* (2.^a ed.). Corpus.

Carr, J. K., Avila, M., Gomez Carrillo, M., Salomon, H., Hierholzer, J., Watanaveeradej, V., Pando, M. A., Negrete, M., Russell, K. L., Sanchez, J., Bix, D. L., Andrade, R., Vinales, J., & McCutchan, F. E. (2001). Diverse BF recombinants have spread widely since the introduction of HIV-1 into South America. *AIDS*, 15(15), F41–F47. <https://doi.org/10.1097/00002030-200110190-00002>

Castelbranco, E. P., da Silva Souza, E., Cavalcanti, A. M., Martins, A. N., de Alencar, L. C., & Tanuri, A. (2010). Frequency of primary resistance to antiretroviral drugs and genetic variability of HIV-1 among infected pregnant women recently diagnosed in Luanda-Angola. *AIDS research and human retroviruses*, 26(12), 1313–1316. <https://doi.org/10.1089/aid.2010.0111>

Centers for Disease Control (CDC). (1982, December 10). Epidemiologic notes and reports: Possible transfusion-associated acquired immune deficiency syndrome (AIDS)—California. *Morbidity and Mortality Weekly Report*, 31(48), 652–654. <https://www.cdc.gov/mmwr/preview/mmwrhtml/00001203.htm>

Centers for Disease Control (CDC). (1982, December 17). Unexplained immunodeficiency and opportunistic infections in infants—New York, New Jersey, California. *Morbidity and Mortality Weekly Report*, 31(49), 665–667. <https://www.cdc.gov/mmwr/preview/mmwrhtml/00001208.htm>

Centers for Disease Control (CDC). (1982, July 16). Epidemiologic notes and reports: Pneumocystis carinii pneumonia among persons with hemophilia A. *Morbidity and Mortality Weekly Report*, 31(27), 365–367. <https://www.cdc.gov/mmwr/preview/mmwrhtml/00001126.htm>

Centers for Disease Control (CDC). (1983, January 7). Epidemiologic notes and reports: Acquired immune deficiency syndrome (AIDS) in prison inmates—New York, New Jersey. *Morbidity and Mortality Weekly Report*, 31(52), 700–701. <https://www.cdc.gov/mmwr/preview/mmwrhtml/00001223.htm>

Centers for Disease Control (CDC). (1983, January 7). Epidemiologic notes and reports: Immunodeficiency among female sexual partners of males with acquired immune deficiency syndrome (AIDS)—New York. *Morbidity and Mortality Weekly Report*, 31(52), 697–698. <https://www.cdc.gov/mmwr/preview/mmwrhtml/00001221.htm>

Clavel, F., & Hance, A. J. (2004). HIV drug resistance. *The New England journal of medicine*, 350(10), 1023–1035. <https://doi.org/10.1056/NEJMra025195>

Coffin, J., Haase, A., Levy, J. A., Montagnier, L., Oroszlan, S., Teich, N., Temin, H., Toyoshima, K., Varmus, H., & Vogt, P. (1986). Human immunodeficiency viruses. *Science* (New York, N.Y.), 232(4751), 697. <https://doi.org/10.1126/science.3008335>

Coffin, J., Blomberg, J., Fan, H., Gifford, R., Hatzioannou, T., Lindemann, D., Mayer, J., Stoye, J., Tristem, M., Johnson, W., & Ictv Report Consortium (2021). ICTV Virus Taxonomy Profile: Retroviridae 2021. *Journal of general virology*, 102(12), 001712. <https://doi.org/10.1099/jgv.0.001712>

D'arc, M., Ayoub, A., Esteban, A., Learn, G. H., Boué, V., Liegeois, F., Etienne, L., Tagg, N., Leendertz, F. H., Boesch, C., Madinda, N. F., Robbins, M. M., Gray, M., Cournil, A., Ooms, M., Letko, M., Simon, V. A., Sharp, P. M., Hahn, B. H., Delaporte, E., ... Peeters, M. (2015). Origin of the HIV-1 group O epidemic in western lowland gorillas. *Proceedings of the National Academy of Sciences of the United States of America*, 112(11), E1343–E1352. <https://doi.org/10.1073/pnas.1502022112>

de Mendoza, C., Ramos, J. M., Caballero, E., Soriano, V., & en nombre del Grupo Español de Estudio del HTLV y VIH-2 (2021). Current epidemiological status of HIV-2 and HTLV-1 infection in Spain. *Medicina clínica*, 156(6), 290–296. <https://doi.org/10.1016/j.medcli.2020.05.029>

De Sá Filho, D. J., Sucupira, M. C., Caseiro, M. M., Sabino, E. C., Diaz, R. S., & Janini, L. M. (2006). Identification of two HIV type 1 circulating recombinant forms in Brazil. *AIDS*, 22(1), 1–13. <https://doi.org/10.1089/aid.2006.22.1>

- Deeks, S., Overbaugh, J., Phillips, A. & Buchbinder, S. (2015). HIV infection. *Nat Rev Dis Primers* 1, 15035. <https://doi.org/10.1038/nrdp.2015.35>
- Delatorre, E., & Bello, G. (2013). Phylodynamics of the HIV-1 epidemic in Cuba. *PloS one*, 8(9), e72448. <https://doi.org/10.1371/journal.pone.0072448>
- Delatorre, E., Silva-de-Jesus, C., Couto-Fernandez, J. C., Pilotto, J. H., & Morgado, M. G. (2017). High HIV-1 Diversity and Prevalence of Transmitted Drug Resistance Among Antiretroviral-Naive HIV-Infected Pregnant Women from Rio de Janeiro, Brazil. *AIDS research and human retroviruses*, 33(1), 68–73. <https://doi.org/10.1089/AID.2016.0159>
- Divino, F., de Lima Guerra Corado, A., Gomes Naveca, F., Stefani, M. M., & Bello, G. (2016). High Prevalence and Onward Transmission of Non-Pandemic HIV-1 Subtype B Clades in Northern and Northeastern Brazilian Regions. *PloS One*, 11(9), e0162112. <https://doi.org/10.1371/journal.pone.0162112>
- Domingo, E., Baranowski, E., Ruiz-Jarabo, C. M., Martín-Hernández, A. M., Sáiz, J. C., & Escarmís, C. (1998). Quasispecies structure and persistence of RNA viruses. *Emerging infectious diseases*, 4(4), 521–527. <https://doi.org/10.3201/eid0404.980402>
- Drummond, A. J., Nicholls, G. K., Rodrigo, A. G., & Solomon, W. (2002). Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*, 161(3), 1307–1320. <https://doi.org/10.1093/genetics/161.3.1307>
- Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, 22(5), 1185–1192. <https://doi.org/10.1093/molbev/msi103>
- Drummond, A. J., Ho, S. Y., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology*, 4(5), e88. <https://doi.org/10.1371/journal.pbio.0040088>
- Faria, N. R., Rambaut, A., Suchard, M. A., Baele, G., Bedford, T., Ward, M. J., Tatem, A. J., Sousa, J. D., Arinaminpathy, N., Pépin, J., Posada, D., Peeters, M., Pybus, O. G., & Lemey, P. (2014). HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science*, 346(6205), 56–61. <https://doi.org/10.1126/science.1256739>
- Felsenstein J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of molecular evolution*, 17(6), 368–376. <https://doi.org/10.1007/BF01734359>
- Felsenstein J. (1985). CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. *Evolution; international journal of organic evolution*, 39(4), 783–791. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>
- Felsenstein, J. (2004). Inferring phylogenies. Sinauer Associates.

Flieller, R., & Cabrera, S. (2019). Virus de inmunodeficiencia humana y resistencia a antirretrovirales. Cátedra de Enfermedades Infecciosas, Facultad de Medicina, Universidad de la República.

Foley, B. T., Leitner, T., Paraskevis, D., & Peeters, M. (2016). Primate immunodeficiency virus classification and nomenclature: Review. Infection, genetics and evolution. *Journal of molecular epidemiology and evolutionary genetics in infectious diseases*, 46, 150–158. <https://doi.org/10.1016/j.meegid.2016.10.018>

Gao, F., Bailes, E., Robertson, D. L., Chen, Y., Rodenburg, C. M., Michael, S. F., Cummins, L. B., Arthur, L. O., Peeters, M., Shaw, G. M., Sharp, P. M., & Hahn, B. H. (1999). Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature*, 397(6718), 436–441. <https://doi.org/10.1038/17130>

Gartner, M. J., Roche, M., Churchill, M. J., Gorry, P. R., & Flynn, J. K. (2020). Understanding the mechanisms driving the spread of subtype C HIV-1. *EBioMedicine*, 53, 102682. <https://doi.org/10.1016/j.ebiom.2020.102682>

Gifford, R. J., Katzourakis, A., Tristem, M., Pybus, O. G., Winters, M., & Shafer, R. W. (2008). A transitional endogenous lentivirus from the genome of a basal primate and implications for lentivirus evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 105(51), 20362–20367. <https://doi.org/10.1073/pnas.0807873105>

Gilbert, M. T., Rambaut, A., Wlasiuk, G., Spira, T. J., Pitchenik, A. E., & Worobey, M. (2007). The emergence of HIV/AIDS in the Americas and beyond. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47), 18566–18570. <https://doi.org/10.1073/pnas.0705329104>

Gill, M. S., Lemey, P., Faria, N. R., Rambaut, A., Shapiro, B., & Suchard, M. A. (2013). Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Molecular biology and evolution*, 30(3), 713–724. <https://doi.org/10.1093/molbev/mss265>

Gottlieb, M. S., Schroff, R., Schanker, H. M., Weisman, J. D., Fan, P. T., Wolf, R. A., & Saxon, A. (1981). Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. *The New England journal of medicine*, 305(24), 1425–1431. <https://doi.org/10.1056/NEJM198112103052401>

Gräf, T., Bello, G., Andrade, P., Arantes, I., Pereira, J. M., da Silva, A. B. P., Veiga, R. V., Mariani, D., Boullosa, L. T., Arruda, M. B., Fernandez, J. C. C., Dennis, A. M., Rasmussen, D. A., & Tanuri, A. (2021). HIV-1 molecular diversity in Brazil unveiled by 10 years of sampling by the national genotyping network. *Scientific reports*, 11(1), 15842. <https://doi.org/10.1038/s41598-021-94542-5>

Granich, R. M., Gilks, C. F., Dye, C., De Cock, K. M., & Williams, B. G. (2009). Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of

HIV transmission: a mathematical model. *The Lancet*, 373(9657), 48–57. [https://doi.org/10.1016/S0140-6736\(08\)61697-9](https://doi.org/10.1016/S0140-6736(08)61697-9)

Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L., Daly, J. M., Mumford, J. A., & Holmes, E. C. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*, 303(5656), 327–332. <https://doi.org/10.1126/science.1090727>

Griffiths, R. C., & Tavaré, S. (1994). Sampling theory for neutral alleles in a varying environment. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 344(1310), 403–410. <https://doi.org/10.1098/rstb.1994.0079>

Griffiths, R. C., & Marjoram, P. (1996). Ancestral inference from samples of DNA sequences with recombination. *Journal of computational biology: journal of computational molecular cell biology*, 3(4), 479–502. <https://doi.org/10.1089/cmb.1996.3.479>

Guimarães, M. L., Eyer-Silva, W. A., Couto-Fernandez, J. C., & Morgado, M. G. (2008). Identification of two new CRF_BF in Rio de Janeiro State, Brazil. *AIDS*, 22, 433–438. <https://doi.org/10.1097/QAD.0b013e3282f47ad0>

Guindon, S., Lethiec, F., Duroux, P., & Gascuel, O. (2005). PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic acids research*, 33(Web Server issue), W557–W559. <https://doi.org/10.1093/nar/gki352>

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology*, 59(3), 307–321. <https://doi.org/10.1093/sysbio/syq010>

Günthard, H. F., Calvez, V., Paredes, R., Pillay, D., Shafer, R. W., Wensing, A. M., Jacobsen, D. M., & Richman, D. D. (2019). Human Immunodeficiency Virus Drug Resistance: 2018 Recommendations of the International Antiviral Society-USA Panel. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 68(2), 177–187. <https://doi.org/10.1093/cid/ciy463>

Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of molecular evolution*, 22(2), 160–174. <https://doi.org/10.1007/BF02101694>

Hemelaar, J., Elangovan, R., Yun, J., Dickson-Tetteh, L., Fleminger, I., Kirtley, S., Williams, B., Gouws-Williams, E., Ghys, P. D., Abimiku, A. G., Agwale, S., Archibald, C., Avidor, B., Barbás, M. G., Barre-Sinoussi, F., Barugahare, B., Belabbes, E. H., Bertagnolio, S., Bix, D., & Zhang, R. (2019). Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *The Lancet Infectious Diseases*, 19(2), 143–155. [https://doi.org/10.1016/S1473-3099\(18\)30647-9](https://doi.org/10.1016/S1473-3099(18)30647-9)

Hemelaar, J., Loganathan, S., Elangovan, R., Yun, J., Dickson-Tetteh, L., Kirtley, S., & WHO-UNAIDS Network for HIV Isolation and Characterization (2020). Country Level

Diversity of the HIV-1 Pandemic between 1990 and 2015. *Journal of Virology*, 95(2), e01580-20. <https://doi.org/10.1128/JVI.01580-20>

Hollingsworth, T. D., Anderson, R. M., & Fraser, C. (2008). HIV-1 transmission, by stage of infection. *Journal of Infectious Diseases*, 198(5), 687–693. <https://doi.org/10.1086/590501>

Hudson, R. R. (1990). Gene genealogies and the coalescent process. *Oxford Surveys in Evolutionary Biology*, 7, 1–44.

Huelsenbeck, J. P., Ronquist, F., Nielsen, R., & Bollback, J. P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, 294(5550), 2310–2314. <https://doi.org/10.1126/science.1065889>

Hymes, K. B., Cheung, T., Greene, J. B., Prose, N. S., Marcus, A., Ballard, H., William, D. C., & Laubenstein, L. J. (1981). Kaposi's sarcoma in homosexual men—a report of eight cases. *The Lancet*, 2(8247), 598–600. [https://doi.org/10.1016/s0140-6736\(81\)92740-9](https://doi.org/10.1016/s0140-6736(81)92740-9)

Ishikawa, S. A., Zhukova, A., Iwasaki, W., & Gascuel, O. (2019). A Fast Likelihood Method to Reconstruct and Visualize Ancestral Scenarios. *Molecular biology and evolution*, 36(9), 2069–2085. <https://doi.org/10.1093/molbev/msz131>

Joint United Nations Programme on HIV/AIDS (UNAIDS). (2025). AIDSinfo. <http://aidsinfo.unaids.org/>

Jukes, T. H., & Cantor, C. R. (1969). Evolution of protein molecules. En H. N. Munro (Ed.), *Mammalian protein metabolism* (Vol. 3, pp. 21–132). Academic Press.

Junqueira, D. M., de Medeiros, R. M., Matte, M. C., Araújo, L. A., Chies, J. A., Ashton-Prolla, P., & Almeida, S. E. (2011). Reviewing the history of HIV-1: spread of subtype B in the Americas. *PloS one*, 6(11), e27489. <https://doi.org/10.1371/journal.pone.0027489>

Junqueira, D. M., & Almeida, S. E. (2016). HIV-1 subtype B: Traces of a pandemic. *Virology*, 495, 173–184. <https://doi.org/10.1016/j.virol.2016.05.003>

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4), 772–780. <https://doi.org/10.1093/molbev/mst010>

Kimura M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16(2), 111–120. <https://doi.org/10.1007/BF01731581>

Kingman, J. F. C. (1982). On the genealogy of large populations. *Journal of Applied Probability*, 19(A), 27–43. <https://doi.org/10.2307/3213548>

Knipe, D. M., & Howley, P. M. (Eds.). (2013). *Fields Virology* (6th ed.). Lippincott Williams & Wilkins.

- Knowles, L. L., & Maddison, W. P. (2002). Statistical phylogeography. *Molecular Ecology*, 11(12), 2623–2635. <https://doi.org/10.1046/j.1365-294X.2002.01625.x>
- Krone, S. M., & Neuhauser, C. (1997). Ancestral Processes with Selection. *Theoretical population biology*, 51(3), 210–237. <https://doi.org/10.1006/tpbi.1997.1299>
- Kuhner, M. K., Yamato, J., & Felsenstein, J. (1998). Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics*, 149(1), 429–434. <https://doi.org/10.1093/genetics/149.1.429>
- Larsson A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* (Oxford, England), 30(22), 3276–3278. <https://doi.org/10.1093/bioinformatics/btu531>
- Le Vu, S., Velter, A., Meyer, L., Peytavin, G., Guinard, J., Pillonel, J., Barin, F., Semaille, C., & Desenclos, J.-C. (2019). Biomarker-based HIV incidence in a population of men who have sex with men in France. *AIDS*, 33(6), 1087–1095. <https://doi.org/10.1097/QAD.0000000000002168>
- Lefort, V., Longueville, J. E., & Gascuel, O. (2017). SMS: Smart Model Selection in PhyML. *Molecular biology and evolution*, 34(9), 2422–2424. <https://doi.org/10.1093/molbev/msx149>
- Leitner, T., Escanilla, D., Franzén, C., Uhlén, M., & Albert, J. (1996). Accurate reconstruction of a known HIV-1 transmission history by phylogenetic tree analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 93(20), 10864–10869. <https://doi.org/10.1073/pnas.93.20.10864>
- Lemey, P., Rambaut, A., & Pybus, O. G. (2006). HIV Evolutionary Dynamics Within and Among Hosts. *AIDS*, 8, 125–140.
- Lemey, P., Rambaut, A., Drummond, A. J., & Suchard, M. A. (2009). Bayesian phylogeography finds its roots. *PLoS computational biology*, 5(9), e1000520. <https://doi.org/10.1371/journal.pcbi.1000520>
- Lemey, P., Salemi, M., & Vandamme, A.-M. (Eds.). (2009). *The phylogenetic handbook: A practical approach to phylogenetic analysis and hypothesis testing* (2nd ed.). Cambridge University Press.
- Liu, P., Song, Y., Colijn, C., & MacPherson, A. (2022). The impact of sampling bias on viral phylogeographic reconstruction. *PLOS global public health*, 2(9), e0000577. <https://doi.org/10.1371/journal.pgph.0000577>
- Mendoza, Y., Martínez, A. A., Castillo Mewa, J., González, C., García-Morales, C., Avila-Ríos, S., Reyes-Terán, G., Armién, B., Pascale, J. M., & Bello, G. (2014). Human immunodeficiency virus type 1 (HIV-1) subtype B epidemic in Panama is mainly driven by dissemination of country-specific clades. *PloS One*, 9(4), e95360. <https://doi.org/10.1371/journal.pone.0095360>

Menéndez-Arias L. (2010). Molecular basis of human immunodeficiency virus drug resistance: an update. *Antiviral research*, 85(1), 210–231. <https://doi.org/10.1016/j.antiviral.2009.07.006>

Minin, V. N., Bloomquist, E. W., & Suchard, M. A. (2008). Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Molecular biology and evolution*, 25(7), 1459–1471. <https://doi.org/10.1093/molbev/msn090>

Ministerio de Salud Pública. (2001). *Crecimiento del SIDA en Uruguay*. <https://archivo.presidencia.gub.uy/noticias/archivo/2001/noviembre/2001112612.htm>

Ministerio de Salud Pública. (2015). *Situación epidemiológica del VIH/SIDA en Uruguay*. Dirección General de Salud, Departamento de Vigilancia de Salud, Área Programática ITS-VIH/SIDA. <https://www.gub.uy/ministerio-salud-publica/comunicacion/publicaciones/boletin-epidemiologico-vih-2015-uruguay>

Ministerio de Salud Pública. (2025). *Situación epidemiológica del VIH/SIDA en Uruguay: diciembre 2025*. <https://www.gub.uy/ministerio-salud-publica/comunicacion/publicaciones/situacion-epidemiologica-vihsida-uruguay-diciembre-2025>

Mir, D., Cabello, M., Romero, H., & Bello, G. (2015). Phylodynamics of major HIV-1 subtype B pandemic clades circulating in Latin America. *AIDS*, 29(14), 1863–1869. <https://doi.org/10.1097/QAD.0000000000000770>

Mir, D., Gräf, T., Esteves De Matos Almeida, S., Pinto, A. R., Delatorre, E., & Bello, G. (2018). Inferring population dynamics of HIV-1 subtype C epidemics in Eastern Africa and Southern Brazil applying different Bayesian phylodynamics approaches. *Scientific Reports*, 8(1). <https://doi.org/10.1038/s41598-018-26824-4>

Montaner, J. S., Lima, V. D., Barrios, R., Yip, B., Wood, E., Kerr, T., Shannon, K., Harrigan, P. R., Hogg, R. S., Daly, P., & Kendall, P. (2010). Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. *The Lancet*, 376(9740), 532–539. [https://doi.org/10.1016/S0140-6736\(10\)60936-1](https://doi.org/10.1016/S0140-6736(10)60936-1)

Moriyama, E. N., Ina, Y., Ikeo, K., Shimizu, N., & Gojobori, T. (1991). Mutation pattern of human immunodeficiency virus gene. *Journal of molecular evolution*, 32(5), 360–363. <https://doi.org/10.1007/BF02101275>

Morrondo, N., Blanco, D., Álvarez, P., & Okatsu, T. (2005). El papel de la recombinación genética en el desarrollo de la pandemia del SIDA. *Investigación Clínica Farmacológica*, 2(1), 09-22.

Murillo, W., Veras, N., Prosperi, M., de Rivera, I. L., Paz-Bailey, G., Morales-Miranda, S., Juarez, S. I., Yang, C., DeVos, J., Marín, J. P., Mild, M., Albert, J., & Salemi, M. (2013). A

single early introduction of HIV-1 subtype B into Central America accounts for most current cases. *Journal of Virology*, 87(13), 7463–7470. <https://doi.org/10.1128/JVI.01602-12>

Nair, M., Gettins, L., Fuller, M., Kirtley, S., & Hemelaar, J. (2024). Global and regional genetic diversity of HIV-1 in 2010-21: systematic review and analysis of prevalence. *The Lancet. Microbe*, 5(11), 100912. [https://doi.org/10.1016/S2666-5247\(24\)00151-4](https://doi.org/10.1016/S2666-5247(24)00151-4)

Nee, S., May, R. M., & Harvey, P. H. (1994). The reconstructed evolutionary process. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 344(1309), 305–311. <https://doi.org/10.1098/rstb.1994.0068>

Onafuwa-Nuga, A., & Telesnitsky, A. (2009). The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. *Microbiology and molecular biology reviews*, 73(3), 451–480. <https://doi.org/10.1128/MMBR.00012-09>

Organización de las Naciones Unidas. (2016). Conferencia Internacional sobre el SIDA. <https://www.unaids.org/es/AIDS2016>

Organización Panamericana de la Salud. VIH/SIDA. <https://www.paho.org/es/temas/vihsida>

Pagán, I., & Holguín, Á. (2013). Reconstructing the Timing and Dispersion Routes of HIV-1 Subtype B Epidemics in The Caribbean and Central America: A Phylogenetic Story. *PLoS One*, 8(7). <https://doi.org/10.1371/journal.pone.0069218>

Paraskevis, D., Pybus, O., Magiorkinis, G., Hatzakis, A., Wensing, A. M. J., van de Vijver, D. A., Albert, J., Angarano, G., Åsjö, B., Balotta, C., Boeri, E., Camacho, R., Chaix, M. L., Coughlan, S., Costagliola, D., de Luca, A., de Mendoza, C., Derdelinckx, I., Grossman, Z., & Vandamme, A. M. (2009). Tracing the HIV-1 subtype B mobility in Europe: A phylogeographic approach. *Retrovirology*, 6(49). <https://doi.org/10.1186/1742-4690-6-49>

Pépin, J. (2021). *The Origins of AIDS* (2nd ed.). Cambridge: Cambridge University Press.

Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M., & Ho, D. D. (1996). HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*, 271(5255), 1582–1586. <https://doi.org/10.1126/science.271.5255.1582>

Pessôa, R., Watanabe, J. T., Calabria, P., Felix, A. C., Loureiro, P., Sabino, E. C., Busch, M. P., & Sanabani, S. S. (2014). Deep sequencing of HIV-1 near full-length proviral genomes identifies high rates of BF1 recombinants including two novel circulating recombinant forms (CRF) 70-BF1 and a disseminating 71-BF1 among blood donors in Pernambuco, Brazil. *PLoS ONE*, 9(11). <https://doi.org/10.1371/journal.pone.0112674>

Pilcher, C. D., Tien, H. C., Eron, J. J., Vernazza, P. L., Leu, S. Y., Stewart, P. W., Goh, L. E., & Cohen, M. S. (2004). Brief but efficient: Acute HIV infection and the sexual transmission of HIV. *Journal of Infectious Diseases*, 189(10), 1785–1792. <https://academic.oup.com/jid/article/189/10/1785/2191338>

- Piot, P., Quinn, T. C., Taelman, H., Feinsod, F. M., Minlangu, K. B., Wobin, O., Mbendi, N., Mazebo, P., Ndangi, K., & Stevens, W. (1984). Acquired immunodeficiency syndrome in a heterosexual population in Zaire. *The Lancet*, 2(8394), 65–69. [https://doi.org/10.1016/s0140-6736\(84\)90241-1](https://doi.org/10.1016/s0140-6736(84)90241-1)
- Plantier, J. C., Leoz, M., Dickerson, J. E., De Oliveira, F., Cordonnier, F., Lemée, V., Damond, F., Robertson, D. L., & Simon, F. (2009). A new human immunodeficiency virus derived from gorillas. *Nature medicine*, 15(8), 871–872. <https://doi.org/10.1038/nm.2016>
- Podchibiakin, A. A., Flieller, R., Ruchansky, D., Brasesco, M., Cortinas, M. N., Lizasoain, A., Chiparelli, H., Cabrera, S., Bello, G., Soler, A. M., & Mir, D. Nationwide molecular epidemiology of HIV-1 in Uruguay (2007–2021): lineage diversity, BF1 recombinant complexity, and epidemiological patterns. Submitted for publication.
- Posada, D., & Crandall, K. A. (2002). The effect of recombination on the accuracy of phylogeny estimation. *Journal of molecular evolution*, 54(3), 396–402. <https://doi.org/10.1007/s00239-001-0034-9>
- Pybus, O. G., & Rambaut, A. (2009). Evolutionary analysis of the dynamics of viral infectious disease. *Nature Reviews Genetics*, 10(8), 540–550. <https://doi.org/10.1038/nrg2583>
- Quagliarello V. (1982). The Acquired Immunodeficiency Syndrome: current status. *The Yale Journal of Biology and Medicine*, 55(5-6), 443–452.
- 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 Evolution*, 2(1), vew007. <https://doi.org/10.1093/ve/vew007>
- Rambaut A. (2018), FigTree. In (Version 1.4). Figtree. <http://tree.bio.ed.ac.uk/software/figtree/>. Accessed May 7 2025.
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic biology*, 67(5), 901–904. <https://doi.org/10.1093/sysbio/syy032>
- Reeves, J. D., & Doms, R. W. (2002). Human immunodeficiency virus type 2. *The Journal of General Virology*, 83(6), 1253–1265. <https://doi.org/10.1099/0022-1317-83-6-1253>
- Reis, M. N. D. G., Bello, G., Guimarães, M. L., & Stefani, M. M. A. (2017). Characterization of HIV-1 CRF90_BF1 and putative novel CRFs_BF1 in Central West, North and Northeast Brazilian regions. *PLoS ONE*, 12(6). <https://doi.org/10.1371/journal.pone.0178578>
- Reis, M. N. G., Guimarães, M. L., Bello, G., & Stefani, M. M. A. (2019). Identification of new HIV-1 circulating recombinant forms CRF81-cpx and CRF99-BF1 in central Western Brazil and of unique BF1 recombinant forms. *Frontiers in Microbiology*, 10(97). <https://doi.org/10.3389/fmicb.2019.00097>

- Rios, M., Fernandez, J., Jaramillo, P., Paredes, V., Sanchez, J. L., Laguna-Torres, V. A., Carr, J. K., & Ramirez, E. (2005). Molecular Epidemiology of HIV Type 1 in Chile: Differential Geographic and Transmission Route Distribution of B and F Subtypes. *AIDS*, 21(10). <https://doi.org/10.1089/aid.2005.21.835>
- Robbins, K. E., Lemey, P., Pybus, O. G., Jaffe, H. W., Youngpairoj, A. S., Brown, T. M., Salemi, M., Vandamme, A. M., & Kalish, M. L. (2003). U.S. Human immunodeficiency virus type 1 epidemic: date of origin, population history, and characterization of early strains. *Journal of Virology*, 77(11), 6359–6366. <https://doi.org/10.1128/jvi.77.11.6359-6366.2003>
- Robertson, D. L., Hahn, B. H., & Sharp, P. M. (1995). Recombination in AIDS viruses. *Journal of Molecular Evolution*, 40(3), 249–259. <https://doi.org/10.1007/BF00163230>
- Robertson, D. L., Anderson, J. P., Bradac, J. A., Carr, J. K., Foley, B., Funkhouser, R. K., Gao, F., Hahn, B. H., Kalish, M. L., Kuiken, C., Learn, G. H., Leitner, T., McCutchan, F., Osmanov, S., Peeters, M., Pieniazek, D., Salminen, M., Sharp, P. M., Wolinsky, S., & Korber, B. (2000). HIV-1 nomenclature proposal. *Science*, 288(5463), 55–56. <https://doi.org/10.1126/science.288.5463.55d>
- Rosas, N., Hernandez, P., Najar, I., Guzman, C., & Castañeda, D. (2013). Características estructurales y funcionales del Virus de la Inmunodeficiencia Humana. *Enfermedades Infecciosas y Microbiología*, 33(4), 163–173.
- Ruchansky, D., Casado, C., Russi, J. C., Arbiza, J. R., & Lopez-Galindez, C. (2009). Identification of a new HIV Type 1 Circulating Recombinant Form (CRF38_BF1) in Uruguay. *AIDS*, 23(3). <https://doi.org/10.1089/aid.2008.0248>
- Russel, P. M., Brewer, B. J., Klaere, S., & Bouckaert, R. R. (2019). Model Selection and Parameter Inference in Phylogenetics Using Nested Sampling. *Systematic Biology*, 68(2), 219–233. <https://doi.org/10.1093/sysbio/syy050>
- Sagulenko, P., Puller, V., & Neher, R. A. (2018). TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evolution*, 4(1), vex042. <https://doi.org/10.1093/ve/vex042>
- Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), 406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>
- Santoro, M. M., & Perno, C. F. (2013). HIV-1 Genetic Variability and Clinical Implications. *ISRN Microbiology*, 2013, 481314. <https://doi.org/10.1155/2013/481314>
- Santos, A. F., Sousa, T. M., Soares, E. A., Sanabani, S., Martinez, A. M., Sprinz, E., Silveira, J., Sabino, E. C., Tanuri, A., & Soares, M. A. (2006). Characterization of a new circulating recombinant form comprising HIV-1 subtypes C and B in southern Brazil. *AIDS*, 20(16), 2011–2019. <https://doi.org/10.1097/01.aids.0000247573.95880.db>

- Schneider, W. L., & Roossinck, M. J. (2001). Genetic diversity in RNA virus quasispecies is controlled by host-virus interactions. *Journal of Virology*, 75(14), 6566–6571. <https://doi.org/10.1128/JVI.75.14.6566-6571.2001>
- Seitz, R. (2016). Human Immunodeficiency Virus (HIV). *Transfusion Medicine and Hemotherapy*, 43(3), 203–222. <https://doi.org/10.1159/000445852>
- Sharp, P. M., & Hahn, B. H. (2011). Origins of HIV and the AIDS pandemic. *Cold Spring Harbor perspectives in medicine*, 1(1), a006841. <https://doi.org/10.1101/cshperspect.a006841>
- Shiino, T., Kondo, M., Sudo, K., et al. (2020). Molecular epidemiology of HIV-1 subtype B in men who have sex with men in Japan. *Journal of Medical Virology*, 92(6), 700–709. <https://doi.org/10.1002/jmv.25675>
- Shimodaira, H., & Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution*, 16(8), 1114–1116. <https://doi.org/10.1093/oxfordjournals.molbev.a026201>
- Sierra-Aragón, S., & Walter, H. (2012). Targets for inhibition of HIV replication: entry, enzyme action, release and maturation. *Intervirology*, 55(2), 84–97. <https://doi.org/10.1159/000331995>
- Slatkin, M., & Maddison, W. P. (1990). Detecting isolation by distance using phylogenies of genes. *Genetics*, 126(1), 249–260. <https://doi.org/10.1093/genetics/126.1.249>
- Stadler, T., Kühnert, D., Bonhoeffer, S., & Drummond, A. J. (2013). Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proceedings of the National Academy of Sciences of the United States of America*, 110(1), 228–233. <https://doi.org/10.1073/pnas.1207965110>
- Stamatakis A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Suchard, M. A., & Rambaut, A. (2009). Many-core algorithms for statistical phylogenetics. *Bioinformatics*, 25(11), 1370–1376. <https://doi.org/10.1093/bioinformatics/btp244>
- 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 Evolution*, 4(1), vey016. <https://doi.org/10.1093/ve/vey016>
- Suzuki, Y., Yamaguchi-Kabata, Y., & Gojobori, T. (2000). Nucleotide Substitution Rates of HIV-1. *AIDS*, 20(2), 39–47.
- Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 10(3), 512–526. <https://doi.org/10.1093/oxfordjournals.molbev.a040023>

Tavaré, S. (1986). *Some probabilistic and statistical problems in the analysis of DNA sequences*. En R. M. Miura (Ed.). *Some mathematical questions in biology—DNA sequence analysis*, 17, 57–86. American Mathematical Society.

Taylor, B. S., Sobieszczyk, M. E., McCutchan, F. E., & Hammer, S. M. (2008). The challenge of HIV-1 subtype diversity. *The New England journal of medicine*, 358(15), 1590–1602. <https://doi.org/10.1056/NEJMra0706737>

Vrancken, B., Rambaut, A., Suchard, M. A., Drummond, A., Baele, G., Derdelinckx, I., Van Wijngaerden, E., Vandamme, A. M., Van Laethem, K., & Lemey, P. (2014). The genealogical population dynamics of HIV-1 in a large transmission chain: bridging within and among host evolutionary rates. *PLoS computational biology*, 10(4), e1003505. <https://doi.org/10.1371/journal.pcbi.1003505>

Visseaux, B., Le Hingrat, Q., Damond, F., Charpentier, C., & Descamps, D. (2019). Physiopathologie de l'infection par le VIH-2 [Physiopathology of HIV-2 infection]. *Virologie*, 23(5), 277–291. <https://doi.org/10.1684/vir.2019.0789>

Volz, E. M., Ionides, E., Romero-Severson, E. O., Brandt, M.-G., Mokotoff, E., Koopman, J. S., & Eisenberg, M. C. (2018). HIV-1 transmission during early infection in men who have sex with men: A phylodynamic analysis. *PLoS Medicine*, 15(1), e1002497. <https://doi.org/10.1371/journal.pmed.1002497>

Worobey, M., Watts, T. D., McKay, R. A., Suchard, M. A., Granade, T., Teuwen, D. E., Koblin, B. A., Heneine, W., Lemey, P., & Jaffe, H. W. (2016). 1970s and 'Patient 0' HIV-1 genomes illuminate early HIV/AIDS history in North America. *Nature*, 539(7627), 98–101. <https://doi.org/10.1038/nature19827>

Yang Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution*, 39(3), 306–314. <https://doi.org/10.1007/BF00160154>

Yang Z. (1996). Among-site rate variation and its impact on phylogenetic analyses. *Trends in ecology & evolution*, 11(9), 367–372. [https://doi.org/10.1016/0169-5347\(96\)10041-0](https://doi.org/10.1016/0169-5347(96)10041-0)

Zuckerkandl, E. and Pauling, L.B. (1962) *Molecular disease, evolution, and genetic heterogeneity*. In: Kasha, M. and Pullman, B., Eds., *Horizons in Biochemistry*, Academic Press, New York, 189-225.

Zuckerkandl, E., & Pauling, L. (1965). *Evolutionary divergence and convergence in proteins*. En V. Bryson & H. J. Vogel (Eds.), *Evolving genes and proteins* (pp. 97–166). Academic Press. <https://doi.org/10.1016/B978-1-4832-2734-4.50017-6>

Zerbini, F. M., Simmonds, P., Adriaenssens, E. M., Lefkowitz, E. J., Oksanen, H. M., Alfenas-Zerbini, P., Aylward, F. O., Freitas-Astúa, J., Hughes, H. R., Łobocka, M., Krupovic, M., Kuhn, J. H., Mushegian, A., Penzes, J. J., Reyes, A., Robertson, D. L., Roux, S., Rubino, L., Sabanadzovic, S., Smith, D. B., ... Varsani, A. (2025). Virus species names have been

standardized; virus names remain unchanged. *mSphere*, 10(5), e0002025.
<https://doi.org/10.1128/msphere.00020-25>

ANEXO 1. Manuscrito asociado a la presente tesis

Nationwide molecular epidemiology of HIV-1 in Uruguay (2007–2021): lineage diversity, BF1 recombinant complexity, and epidemiological patterns.

Revista: Journal of the International AIDS Society (JIAS)

Estado: Submitido el 21 de enero de 2026 (en revisión)

Authors: Podchibiakin, A. A.¹, Flieller, R.², Ruchansky, D.², Brasesco, M.², Cortinas, M.N.³, Lizasoain, A.⁴, Chiparelli, H.², Cabrera, S.⁵, Bello, G.⁶, Soler, A.M.⁷, Mir, D¹.

Affiliation:

¹Universidad de la República - CENUR, Department of Biological Sciences, Genomics and Bioinformatics Unit, Salto, Uruguay

²Ministry of Public Health, Department of Public Health Laboratories, Virology Unit, Montevideo, Uruguay

³Ministry of Public Health, Department of Public Health Laboratories, Genomics Unit, Montevideo, Uruguay

⁴Universidad de la República - CENUR, Department of Biological Sciences, Molecular Virology Laboratory, Salto, Uruguay

⁵Universidad de la República, Faculty of Medicine, Montevideo, Uruguay

⁶Fiocruz, Oswaldo Cruz Institute, Arbovirus and Hemorrhagic Virus Laboratory, Rio de Janeiro, Brazil

⁷Universidad de la República - CENUR, Department of Biological Sciences, Human Molecular Genetics Laboratory, Salto, Uruguay

***Corresponding author: Daiana Mir.**

Conflicts of Interest and Source of Funding: The authors have no conflict of interest. This work was supported by a research grant from the Comisión Sectorial de Investigación Científica (CSIC-Uruguay) by the code CSIC_VUSP_Modalidad-2-2021 and by L' Oreal Uruguay for Women in Science Fellowship Program. A. A. Podchibiakin is the recipient of a predoctoral scholarship (POS_NAC_2023_1_178397) from ANII (Uruguay).

ABSTRACT

Introduction: Uruguay exhibits one of the highest HIV incidence rates in Latin America, but its molecular landscape remains insufficiently studied. Understanding the diversity and dynamics of circulating HIV-1 lineages is essential for strengthening surveillance and informing public health responses.

Methods: A total of 1,268 HIV-1 *pol* sequences (PR/RT region) generated between 2007 and 2021 through the nationwide routine genotyping program were analyzed. Subtyping was performed with online tools, maximum-likelihood phylogenies and recombination analyses. Associations between patients' demographic data and HIV-1 lineages were assessed using Z-tests and Fisher's exact tests, while temporal trends in lineage prevalence were evaluated using Pearson correlations.

Results: Seventeen HIV-1 lineages were identified, including four pure subtypes (B, C, A1, F1) and twelve CRFs. Subtype B was the most prevalent (39.3%), followed by CRFs 12_BF (27.8%) and 38_BF1 (12.9%). An additional 8.5% of sequences formed a well-supported clade of B/F1 recombinants (designated BF1.UY), which could not be attributed to any known CRF. Subtype B showed a strong association with men and MSM transmission; 38_BF1 was enriched among women and individuals with injection-related exposure; and BF1.UY was predominantly linked to heterosexual transmission. Geographic distribution reflects regional connectivity: 12_BF predominated in western and eastern regions with strong Argentine influence, whereas Brazilian-associated variants (subtype C, 31_BC, F1) clustered in the Northeast and East. Temporal analyses showed a significant decline in subtype B and a gradual increase in the combined BF1 recombinant group over the study period.

Conclusions: Uruguay harbors a highly diverse, recombinant-enriched HIV-1 epidemic, largely dominated by BF1 recombinant forms and shaped by regional epidemics as well as extensive local diversification. The identification of the BF1.UY clade highlights ongoing recombination events arising within the country. These findings underscore the need to strengthen epidemiological and molecular surveillance, together with enhanced regional genomic cooperation to monitor emerging variants and guide public health strategies.

Keywords: HIV-1; molecular epidemiology; Uruguay.

INTRODUCTION

Over the past decade, the global HIV epidemic has undergone significant shifts. In 2024, an estimated 1.3 million people (1.0 - 1.7 million) were newly diagnosed with HIV, representing

a 40% decline in global incidence since 2010. However, this overall decline masks notable regional disparities such as in Latin America, which has experienced a 13% increase in new infections since 2010 [1]. This rise unfolds within a broader context of regional challenges, including limited access to antiretroviral therapy (ART), high prevalence of coinfections, socio-economic disparities, and insufficient investment in healthcare infrastructure and research [2]. Furthermore, political and economic instability in the region exacerbates population mobility, which disrupts continuity of care and undermines prevention strategies, thereby intensifying transmission dynamics [3].

In this context, Uruguay has one of the highest HIV incidence rates in Latin America, with 0.51 new cases per 1,000 adults (aged 15–49) in 2023 [4,5]. Notably, 83% of people diagnosed with HIV are currently receiving antiretroviral therapy (ART), underscoring national progress in expanding access and supporting adherence. Furthermore, vertical transmission has consistently remained below 2% in recent years, highlighting sustained success in prevention [6]. Despite these advances, a substantial proportion of new diagnoses, particularly among men, still occur at advanced stages, as reflected in low CD4+ T-cell counts at presentation. Moreover, while HIV prevalence in the general population remains relatively low at 0.6%, it exceeds 5% among key populations, including men who have sex with men (MSM), injecting drug users (IDUs), male sex workers (MSWs), transgender individuals, and incarcerated people [7].

Understanding the genetic diversity of HIV is essential for detecting emerging strains, improving diagnostic and therapeutic strategies, monitoring drug resistance, and tracing transmission networks. In Latin America, HIV-1 subtype B has historically dominated the epidemic. However, co-circulation with subtypes C and F1 has led to the emergence of multiple circulating recombinant forms (CRFs), such as 12_BF (Argentina and Uruguay), 17_BF1 (Argentina), 28_BF1, 29_BF1 and 31_BC (Brazil), 38_BF1 (Uruguay), 44_BF (Chile), and 89_BF1 (Bolivia and Peru), which together constitute a substantial component of the regional epidemic [8–13]. Despite shared overall patterns, HIV-1 variant distribution is highly heterogeneous across countries as exemplify by Brazil and Argentina. Brazil, which accounts for roughly 40% of HIV cases in the region [5], is characterised by a national predominance of subtype B, a high prevalence of subtype C and BC recombinants in the South region, and a wide distribution of subtype F1 and B/F recombinants across states [14]. In Argentina, subtype B remains the most prevalent variant, although B/F recombinants have been rising in frequency over time [15–17].

In Uruguay, the HIV epidemic has likely been shaped by cross-border mobility with Argentina and Brazil, driven by longstanding historical and cultural ties [18]. Molecular surveillance indicates predominance of subtype B and B/F recombinants, with 12_BF playing a major role [17,19]. Phylogenetic analyses further reveal a shared regional transmission network for 12_BF between Argentina and Uruguay [20], and the detection of subtype C strains closely related to those circulating in southern Brazil supports ongoing transnational viral flow [21,22]. Despite these insights, Uruguay's HIV molecular epidemiology remains insufficiently characterized and further research is needed to better understand the transmission patterns and evolutionary forces shaping the national epidemic.

To address this gap, we conducted the first nationwide, longitudinal molecular epidemiology study of HIV-1 in Uruguay, analyzing 1,598 *pol* sequences and associated patient metadata from routine genotyping conducted between 2007 and 2021. By characterizing viral lineages and their transmission dynamics, this study provides novel insights into the evolution and spread of HIV-1 within Uruguay and its relationship to broader regional dynamics.

METHODS

Routine genotyping data. This study analyzed 1,598 partial *pol* nucleotide sequences of HIV-1, obtained through routine genotyping conducted by the Department of Public Health Laboratories of the Uruguayan Ministry of Public Health, using samples collected between 2007 and 2021. The sequences spanned the protease and partial reverse transcriptase (PR/RT) regions (HXB2 positions 2,358–3,251) and were generated via Sanger sequencing, following the protocol described by Alemán et al. (2015) [23]. Epidemiological metadata, including sampling location, gender, and reported transmission route, were integrated with sequence data to build GENOUY-HIV, the first national HIV molecular database in Uruguay. To ensure patient confidentiality, all records were anonymized using unique alphanumeric identifiers. Sequences were excluded if they met any of the following criteria: (i) patients under 14 years of age, to minimize potential biases from vertical transmission; (ii) duplicate samples from the same individual, retaining only the earliest available sequence; or (iii) failure to meet quality standards, as assessed by the Los Alamos HIV Sequence Quality Assessment tool (<https://www.hiv.lanl.gov>), due to evidence of hypermutation or premature stop codons. After applying these filters, 1,268 high-quality sequences were retained for downstream analyses. The final dataset has been deposited in GenBank under accession numbers PZ351606–PZ352874.

HIV-1 subtype classification and phylogenetic analysis. HIV-1 *pol* sequences were initially assigned to subtypes or recombinant forms using three automated tools: REGA v3.0 [24], COMET [25], and RIP [26]. Only concordant results across all tools were considered

definitive. Sequences with discordant or unassigned outputs were further screened for mosaic structure using Bootscan in SimPlot package v3.5.1 [27] (300-nt sliding window, 10-nt step size). Mosaic consistent with known CRFs were annotated accordingly, while others were retained for further analyses.

Three targeted phylogenetic validations were then performed: (i) sequences classified as pure subtypes were co-analyzed with subtype-specific reference panels to confirm their assignments; (ii) sequences assigned to CRFs were evaluated against CRF-specific reference datasets to verify clustering consistency; and (iii) sequences not attributable to any known CRF were analyzed with a comprehensive panel including reference sequences from all recognized CRFs to resolve their relationships to existing forms. All reference panels (pure subtype, CRF-specific, and the comprehensive CRF panel for unassigned sequences) were curated from the Los Alamos HIV Sequence Database.

All alignments were generated using MAFFT [28], excluding codons associated with major drug resistance mutations (12 in PR, 21 in RT) with Aliview [29]. Phylogenetic trees were inferred using IQ-TREE v1.6.12 [30], with the best-fit substitution model selected by ModelFinder [31]. Node support was assessed with the SH-aLRT test [32], considering values ≥ 0.80 as strong evidence of clustering with subtype or CRF reference sequences.

Frequency and statistical analysis. Following lineage assignment, HIV-1 variant frequencies for Uruguay (2007–2021) were determined overall and by macro-region. Uruguay's first-level administrative divisions were grouped into five macro-regions: Northeast, West, Center, East, and the Metropolitan Area. Associations between HIV-1 variants (restricted to those with >40 sequences) and geographic region, gender, and transmission route were tested using Fisher's exact test and Z-tests for proportions (two-sided, $\alpha=0.05$) in IBM SPSS Statistics v22.0. Gender was categorized as male, female, or transgender. For the purposes of this analysis, we treated entries labelled as "transgender" as indicating a gender identity different from sex assigned at birth, although the underlying variable does not explicitly distinguish between these attributes. Transmission routes were classified as heterosexual (HTX), men who have sex with men (MSM), and injection drug use (IDU). Temporal trends for major lineages and recombinant forms were assessed using Pearson's correlation on triennial counts (non-overlapping 3-years periods).

Ethical aspects. This work was approved by the Ethics Committee of the Centro Universitario Regional (CENUR) Litoral Norte of the Universidad de la República (Udelar), Uruguay, under Registration No. 311170-500476-21. Data was compiled in 2022 and

analyzed between 2022 and 2025. All methods were performed in accordance with the relevant guidelines and regulations.

RESULTS

Molecular profile of HIV-1 lineages. A total of 1,268 HIV-1 *pol* sequences from samples collected between 2007 and 2021 were analyzed to characterize the molecular diversity of the epidemic in Uruguay. Based on an average of nearly 900 new HIV diagnoses per year in Uruguay during this period [6], these sequences represent approximately 9.4% of all newly diagnosed cases, providing substantial coverage of the epidemic, particularly in the metropolitan area.

Of these 1,268 sequences, 91.5% ($n = 1,160$) could be assigned to known HIV-1 lineages, revealing 17 distinct genetic forms circulating in the Uruguayan population. These included four pure subtypes and sub-subtypes (B, C, A1 and F1) and twelve CRFs: 01_AE, 12_BF, 17_BF1, 19_cpx, 20_BG, 24_BG, 28/29_BF1 (these two recombinants could not be distinguished due to identical recombination patterns in the analyzed region), 31_BC, 38_BF1, 41_CD, 46_BF1, and 60_BC (Figure 1). Subtype B was the predominant lineage (39.3%), followed by CRFs 12_BF (27.8%), 38_BF1 (12.9%) and 28/29_BF1 (3.4%). Subtype C accounted for 3.2%, while the remaining subtypes and recombinant forms together represented 5.1% of the dataset (Figure 1 and Supplementary Table 1).

A total of 108 *pol* sequences (8.5%) could not be assigned to any known CRF by automated subtyping or recombination analysis. Recombination analysis showed that all of these sequences exhibited B/F1 mosaic patterns within the analyzed region. All sequences harbored a primary recombination breakpoint whose position varied among sequences but consistently fell between nucleotides 2,400 and 2,600 relative to HXB2 (bp1), delineating two segments: an upstream F1-derived fragment and a downstream B-derived fragment. In most cases, the analyzed *pol* region consisted solely of this F1/B mosaic structure, whereas a subset of sequences exhibited two additional breakpoints (bp2: 2,600–2,700; bp3: 2,850–2,900), resulting in an F1/B/F1/B configuration (data not shown). In the inferred maximum-likelihood phylogeny, these B/F1 recombinants formed a strongly supported clade (SH-aLRT > 0.90) that did not include any CRF reference sequences. This clade of B/F1-like recombinants, hereafter designated BF1.UY, arose from a basal node as part of a well-supported polytomy (SH-aLRT > 0.90) with two additional lineages: one comprising 38_BF1 and 42_BF1 reference strains (SH-aLRT > 0.80) and another consisting of 17_BF1 (SH-aLRT > 0.90) (Figure 2).

Geographic distribution of HIV-1 lineages in Uruguay. Of the 1,268 pol sequences analyzed, 777 (61.3%) contained metadata on the reported location of HIV acquisition. Most originated from the Metropolitan Area (75.2%), followed by the Northeast (9.4%), West (9.1%), East (5.5%), and Center (0.8%) (Figure 3); results for the latter should be interpreted cautiously due to limited representation. Subtype B and the major BF recombinants 12_BF, 38_BF1, and BF1.UY were detected nationwide. Subtype B showed relatively uniform frequencies ($\approx 30\text{--}40\%$) and predominated in the Metropolitan, Northeast, and Central regions. In contrast, 12_BF displayed marked geographic heterogeneity, being most prevalent in the East and West but less frequent in the Northeast ($\sim 16\%$). Accordingly, pairwise Z-tests indicated that the West had a higher proportion of 12_BF and a lower proportion of subtype B than both the Northeast and Metropolitan regions. The 38_BF1 occurred countrywide at moderate frequencies (8–16%), while BF1.UY was found in all regions at generally lower levels ($<10\%$). Less frequent forms, including subtype C and 31_BC, were mainly restricted to the Northeast and East; consistent with this, Z-tests showed that the Northeast had a higher proportion of subtype C and lower proportions of 12_BF and 38_BF1 than the Metropolitan Area (Figure 3).

Association between HIV-1 lineages and gender. Gender information was available for 1,230 sequences (97.0% of those included in the analyses). Using Fisher's exact test to compare the gender distribution of each major lineage (>40 sequences) against all remaining lineages, significant associations were observed for subtype B ($p < 0.001$), 12_BF ($p = 0.005$), and 38_BF1 ($p = 0.036$). Z-tests for proportions showed that subtype B was overrepresented among men (65.2% in subtype B vs 50.5% in the rest of the dataset) and underrepresented among women (34.0% vs 47.3%). In contrast, 12_BF was relatively enriched among transgender individuals (3.3% vs 1.0%) and depleted among men (51.5% vs 58.2%), whereas 38_BF1 was enriched among women (51.3% vs 40.6%) and depleted among men (47.5% vs 57.7%) (Figure 4A and Supplementary Table 2).

Associations between HIV-1 lineages and transmission route. Transmission route data suitable for analysis were available for 854 sequences (67.4% of the 1,268 sequences). Fisher's exact tests revealed significant associations between transmission route and subtype B ($p < 0.001$), 38_BF1 ($p < 0.001$) and the BF1.UY clade ($p = 0.013$). Z-tests for proportions indicated that subtype B was more frequently associated with MSM transmission (31.1% of subtype B sequences vs 18.6% among all other lineages combined). In contrast, 38_BF1 showed a higher proportion of IDU transmission (12.5% vs 3.1% in the remaining lineages) and a lower proportion of MSM transmission than the rest of the dataset (12.5% vs

25.2%). The BF1.UY clade was predominantly associated with HTX transmission (86.3% vs 70.8% in the remaining lineages) (Figure 4B and Supplementary Table 3).

Temporal trends in major HIV-1 lineages and BF1 recombinants. Temporal patterns were evaluated in non-overlapping three-year periods (2007–2009, 2010–2012, 2013–2015, 2016–2018, 2019–2021) for the six most frequent lineages and for a composite group of BF1 recombinants. Pearson's correlation on triennial proportions showed a significant negative temporal trend for subtype B ($R^2 = 0.829$; $p = 0.033$), indicating a progressive decline in its relative frequency. In contrast, the BF1 recombinant group showed a positive correlation ($R^2 = 0.778$; $p = 0.0479$) which, although it should be interpreted with caution, is consistent with an increasing contribution to the epidemic over the study period (Figure 4C). The BF1 recombinant group included 12_BF, 17_BF1, 28/29_BF1, 38_BF1, 46_BF1 and BF1.UY.

DISCUSSION

This study provides the first nationwide, longitudinal characterization of HIV-1 molecular diversity in Uruguay, revealing a heterogeneous epidemic dominated by subtype B and BF1 recombinant lineages. Together, the high diversity and prevalence of BF1 forms, including the newly delineated BF1.UY clade, underscore that Uruguay harbours a recombination-driven, BF1-enriched epidemic landscape, with multiple related lineages undergoing ongoing local diversification.

In quantitative terms, subtype B remains the single most frequent lineage, accounting for 39.3% of characterized infections, but it is collectively outnumbered by BF1 recombinants, which together represent about 55% of the sample. This BF1 spectrum is dominated by 12_BF (27.8%) and 38_BF1 (12.9%), followed by the BF1.UY clade (8.5%), with other BF1 forms such as 28/29_BF1 and 46_BF1 contributing smaller proportions. Analysis of triennial proportions revealed a significant negative temporal trend for subtype B, indicating a progressive decline in its relative contribution over the study period, whereas the grouped BF1 recombinants showed a positive, albeit more modest and borderline, correlation, consistent with a gradual increase in their relative prevalence over the study period. Similar patterns, in which the relative contribution of “pure” subtype B decreases while CRFs and URFs gain prominence, have been reported in other settings [33]; however, in Uruguay this diversification is concentrated mainly within the BF1 recombinant spectrum rather than involving a broader range of non-B variants.

Geographically, the Uruguayan epidemic shows limited phylogeographic structuring, with most spatial signal confined to lineages that have clear epidemiological links to neighboring epidemics. Among BF1 recombinants, 12_BF, originally described in Argentina and Uruguay

[8], represents the second most common lineage in Uruguay after subtype B and is particularly frequent in the West and along the eastern Atlantic coast, consistent with intense cross-border population exchange with Argentina and seasonal touristic flows [18,34]. In Argentina, molecular analyses have shown that many BF1 recombinants are second-generation forms derived from 12_BF through successive rounds of recombination with locally circulating subtype B viruses [17,35,36]. A similar process may underlie the 12_BF-like diversity observed in Uruguay, although near full-length genome data will be required to determine whether these viruses represent prototypical 12_BF or related second-generation recombinants. Further investigation is also needed to disentangle the relative contributions of local onward transmission and repeated introductions to the persistence of this lineage in Uruguay. By contrast, lineages associated with the Brazilian epidemic, including 28/29_BF1, 31_BC, subtype C and sub-subtype F1 [14], are mainly concentrated in the East and Northeast, where cross-border mobility with southern Brazil is greatest [37,38]. Subtype B, in turn, is detected at similar frequencies across all macro-regions, consistent with early introduction and long-standing dissemination at the national scale. The other major BF1 recombinant lineages, 38_BF1 and BF1.UY, are also widely distributed throughout the country, supporting their consolidation as endemic components of the epidemic. Taken together, these spatial patterns indicate that Uruguay's current HIV-1 diversity reflects the combined effects of early nationwide dissemination of subtype B, multiple regionally constrained introductions from neighboring countries, and subsequent local amplification of successful recombinant BF1 lineages.

Beyond their geographic distribution, the major HIV-1 lineages circulating in Uruguay display distinct demographic and transmission profiles. Subtype B is strongly associated with men and with MSM transmission, consistent with multiple international studies describing the persistence of this subtype within MSM transmission networks [14, 39–41]. In contrast, 38_BF1 shows higher representation among women and a greater frequency of injection-related transmission compared with other lineages. This pattern is consistent with the original description of 38_BF1 by Ruchansky et al. (2009), who identified this recombinant as widely distributed in Uruguay and more frequently found among individuals with injection-related or high-risk sexual exposure, including vertically infected children born to mothers who used drugs. The persistence of these epidemiological associations over more than two decades reflects the continued circulation of 38_BF1 within population groups historically affected by parenteral exposure. Regarding other BF1 recombinants, 12_BF, a lineage of regional relevance, showed no statistically significant association with any single transmission category. Nevertheless, 74% of analysed 12_BF sequences were linked to heterosexual exposure, consistent with observations from Argentina where BF1

recombinants, including 12_BF, have been predominantly associated with heterosexual transmission networks [42,43]. Within this overall profile, 12_BF also appeared comparatively more frequent among individuals recorded as transgender, a pattern that has not, to date, been described in regional studies of 12_BF circulation and that points to transgender and other gender-diverse populations as potentially important nodes in the dissemination of BF1 lineages in Latin America.

Within the BF1-enriched background, the BF1.UY clade identified in this study represents one of the most distinctive findings of the Uruguayan HIV-1 epidemic. Phylogenetic and recombination analyses indicate that BF1.UY behaves as a family of closely related B/F1 recombinants rather than a single CRF. In addition, BF1.UY is predominantly associated with heterosexual transmission, consistent with the prominent role of heterosexual networks in the dissemination of BF1 recombinants described in the region [8, 43]. However, because our analyses are restricted to partial *pol* sequences, full genome sequencing will be required to determine whether BF1.UY corresponds to several novel CRFs or instead represents a locally diversified constellation of URFs within the broader BF1 recombinant spectrum.

Taken together, these results should be interpreted in light of several limitations. GENOUY-HIV, although by far the most extensive HIV-1 molecular dataset compiled in Uruguay, remains a convenience sample restricted to individuals for whom genotypic resistance testing was requested. Coverage is therefore incomplete and geographically unbalanced, with particularly sparse sampling from the Central region, so lineage distributions in that area are likely under-resolved. Because most sequences were generated in the setting of virological failure, the lineage frequencies and temporal patterns inferred here probably reflect infections acquired several years earlier; the actual contribution of specific variants at the time of transmission may thus differ from our estimates, with subtype B likely overrepresented and BF1 recombinants underrepresented. At the genomic level, analyses were limited to the PR/RT region, precluding the detection of recombination events elsewhere and almost certainly underestimating the burden of complex recombinants and overall diversity. Finally, behavioural and demographic variables were derived from routine clinical records, and some, including gender and exposure, were coarse or inconsistently defined (e.g., sex assigned at birth and gender identity were not captured as separate variables), limiting the resolution of lineage-level associations with population subgroups. Despite these caveats, this analysis constitutes the most comprehensive molecular characterization of the HIV-1 epidemic in Uruguay to date and provides a robust foundation for future whole-genome sequencing and genomic surveillance efforts.

In summary, these findings highlight the broader potential of routine HIV genotyping data to strengthen integrated molecular surveillance in Latin America. The integration of clinical genotyping with epidemiological metadata, as implemented in the GENOUY-HIV framework, shows that datasets originally generated for drug resistance testing can be repurposed as cost-effective resources for real-time monitoring of viral diversity and transmission dynamics, an approach successfully adopted in several countries to inform prevention policies and guide public health responses [44–47]. Extending this strategy to whole-genome sequencing will enable higher-resolution detection of emerging recombinant forms and more precise reconstruction of transmission networks. In the longer term, linking these data with phylodynamic frameworks could support inference of key epidemiological parameters such as incidence, reproductive number (R_0), and patterns of geographic dissemination [20,48,49], thereby expanding their value for public health planning. Strengthening regional genomic cooperation among Argentina, Brazil, and Uruguay will be essential to monitor cross-border transmission and to foster coordinated prevention strategies. Collectively, these perspectives highlight the importance of developing an integrated molecular surveillance system that bridges clinical, epidemiological, and genomic data to guide evidence-based responses to HIV in Latin America.

REFERENCES

1. UNAIDS. (2025). AIDS, crisis and the power to transform: UNAIDS Global AIDS Update 2025. UNAIDS, Geneva.
2. Turk, G., Devisich, G., Valiente-Echeverria, F., Figueroa, M. I., Cassetti, I., Cahn, P., Cortes, C. P., Crabtree-Ramírez, B., Sued, O., & Laufer, N. (2024). Unity is strength: creation and future actions for a consortium for HIV cure research in Latin America and the Caribbean 'Lac-Cura.' In HIV Research and Clinical Practice (Vol. 25, Issue 1). Taylor and Francis Ltd.
3. UNAIDS. (2019). Mobility and stability – Advancing the health and rights of migrants in Latin America and the Caribbean. UNAIDS.
https://www.unaids.org/sites/default/files/media_asset/advancing-the-health-and-rights-of-migrants-in-latin-america-and-the-caribbean_en.pdf
4. Instituto Nacional de Estadística (Uruguay). (2025). Censo 2023: Cifras de Uruguay. Total de población y viviendas. INE, Montevideo.
<https://www.gub.uy/instituto-nacional-estadistica/censos2023pvh>
5. UNAIDS. AIDSinfo. <https://aidsinfo.unaids.org/> (accessed April 16, 2025).
6. Dirección General de la Salud – Departamento de Vigilancia en Salud. (2025). *Semana de la prevención de las infecciones de transmisión sexual. Boletín*

- epidemiológico*. Ministerio de Salud Pública. Montevideo, Uruguay.
<https://www.gub.uy/ministerio-salud-publica/sites/ministerio-salud-publica/files/documentos/publicaciones/Bolet%C3%ADn%20epidemiol%C3%B3gico%20de%20infecciones%20de%20transmisi%C3%B3n%20sexual%202025.pdf>
7. Dirección General de la Salud – Departamento de Vigilancia en Salud. (2015). *Situación epidemiológica del VIH/SIDA en Uruguay*. Ministerio de Salud Pública. Montevideo, Uruguay.
<https://www.gub.uy/ministerio-salud-publica/comunicacion/publicaciones/boletin-epidemiologico-vih-2015-uruguay>
 8. Carr, J. K., Avila, M., Gomez Carrillo, M., Salomon, H., Hierholzer, J., Watanaveeradej, V., Pando, M. A., Negrete, M., Russell, K. L., Sanchez, J., Bix, D. L., Andrade, R., Vinales, J., & McCutchan, F. E. (2001). Diverse BF recombinants have spread widely since the introduction of HIV-1 into South America.
 9. de Sá Filho, D. J., Sucupira, M. C., Caseiro, M. M., Sabino, E. C., Sobhie Diaz, R., & Mario Janini, L. (2006). Identification of Two HIV Type 1 Circulating Recombinant Forms in Brazil. *AIDS Research and Human Retroviruses*, 22(1), 1–13.
 10. Santos, A.F., Sousa, T.M., Soares, E.A., Sanabani, S., Martinez, A.M., Sprinz, E., Silveira, J., Sabino, E.C., Tanuri, A., & Soares, M.A. (2006) Characterization of a new circulating recombinant form comprising HIV-1 subtypes C and B in southern Brazil. *AIDS*, 20(16):2011-9.
 11. Ruchansky, D., Casado, C., Russi, J. C., Arbiza, J. R., & Lopez-Galindez, C. (2009). Identification of a new HIV Type 1 Circulating Recombinant Form (CRF38_BF1) in Uruguay. *AIDS Research and Human Retroviruses*, 23(3).
 12. Ríos, M., Delgado, E., Pérez-Álvarez, L., Fernández, J., Gálvez, P., de Parga, E. V., Yung, V., Thomson, M. M., & Nájera, R. (2007). Antiretroviral drug resistance and phylogenetic diversity of HIV-1 in Chile. *Journal of Medical Virology*, 79(6), 647–656.
 13. Delgado, E., Fernández-García, A., Pérez-Losada, M., Moreno-Lorenzo, M., Fernández-Miranda, I., Benito, S., Montero, V., Gil, H., Hernáez, S., Muñoz, J., Zubero-Sulibarria, M.Z., García-Bodas, E., Sánchez, M., Del Romero, J., Rodríguez, C., Elorduy, L., Bereciartua, E., Culebras, E., Rodríguez-Avil, I., Giménez-Alarcón, M.L., Martín-Salas, C., Gómez-González, C., García-Irure, J.J., Cenxual, G., Martínez-Sapiña, A., Maiques-Camarero, M., Pérez-Álvarez, L. & Thomson, M.M. (2021) Identification of CRF89_BF, a new member of an HIV-1 circulating BF intersubtype recombinant form family widely spread in South America. *Sci Rep*. 1;11(1):11442.
 14. Gräf, T., Bello, G., Andrade, P., Arantes, I., Pereira, J. M., da Silva, A. B. P., Veiga, R. v., Mariani, D., Boullosa, L. T., Arruda, M. B., Fernandez, J. C. C., Dennis, A. M.,

- Rasmussen, D. A., & Tanuri, A. (2021). HIV-1 molecular diversity in Brazil unveiled by 10 years of sampling by the national genotyping network. *Scientific Reports*, 11(1).
15. Thomson, M. M., Villahermosa, M. L., Vázquez-de-Parga, E., Cuevas, M. T., Delgado, E., Manjón, N., Medrano, L., Pérez-Alvarez, L., Contreras, G., Carrillo, M. G., Salomón, H., & Nájera, R. (2000). Widespread circulation of a B/F intersubtype recombinant form among HIV-1-infected individuals in Buenos Aires, Argentina. *AIDS*, 14(7), 897–899.
 16. Thomson, M. M., Delgado, E., Herrero, I., Villahermosa, M. L., Vázquez-de Parga, E., Cuevas, M. T., Carmona, R., Medrano, L., Pérez-Álvarez, L., Cuevas, L., & Nájera, R. (2002). Diversity of mosaic structures and common ancestry of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from Argentina revealed by analysis of near full-length genome sequences. *Journal of General Virology*, 83(1), 107–119.
 17. Hierholzer, A. J., Montano, S., Hoelscher, M., Negrete, M., Hierholzer, M., Avila, M. M., Carrillo, M. G., Russi, J. C., Vinales, J., Alava, A., Acosta, M. E., Gianella, A., Andrade, R., Sanchez, J. L., Carrion, G., Sanchez, J. L., Russell, K., Robb, M., Bix, D., ... Carr, J. K. (2002). Molecular Epidemiology of HIV Type 1 in Ecuador, Peru, Bolivia, Uruguay, and Argentina. *AIDS RESEARCH AND HUMAN RETROVIRUSES*, 18, 1339–1350
 18. Benedetti, A. (2014). The Argentina-Uruguay Border Space: A Geographical Description. *FRONTERA NORTE*, 27, 35–61. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0187-73722015000100002
 19. Dirección General de la Salud – División Epidemiología (Programa ITS/Sida-DEVISA-DLSP). (2011). *Informe de la situación nacional de VIH/Sida*. Ministerio de Salud Pública. Montevideo, Uruguay. <https://www.gub.uy/ministerio-salud-publica/comunicacion/publicaciones/informe-epidemiologico-vihsida-diciembre-2011>
 20. Bello, G., Aulicino, P. C., Ruchansky, D., Guimarães, M. L., Lopez-Galindez, C., Casado, C., Chiparelli, H., Rocco, C., Mangano, A., Sen, L., & Morgado, M. G. (2010). Phylodynamics of HIV-1 Circulating Recombinant Forms 12_BF and 38_BF in Argentina and Uruguay. *Retrovirology*, 7(22).
 21. Véras, N. M. C., Gray, R. R., de Macedo Brígido, L. F., Rodrigues, R., & Salemi, M. (2011). High-resolution phylogenetics and phylogeography of human immunodeficiency virus type 1 subtype C epidemic in South America. *Journal of General Virology*, 92(7), 1698–1709.

22. Alves, B. M., Siqueira, J. D., Prellwitz, I. M., Botelho, O. M., da Hora, V. P., Sanabani, S., Recordon-Pinson, P., Fleury, H., Soares, E. A., & Soares, M. A. (2019). Estimating HIV-1 genetic diversity in Brazil through next-generation sequencing. *Frontiers in Microbiology*, 10.
23. Alemán, Y., Vinken, L., Kourí, V., Pérez, L., Álvarez, A., Abrahantes, Y., Fonseca, C., Pérez, J., Correa, C., Soto, Y., Schrooten, Y., Vandamme, A. M., & van Laethem, K. (2015). Performance of an in-house human immunodeficiency virus type 1 genotyping system for assessment of drug resistance in Cuba. *PLoS ONE*, 10(2).
24. Vilsker, M., Moosa, Y., Nooij, S., Fonseca, V., Ghysens, Y., Dumon, K., Pauwels, R., Alcantara, L. C., vanden Eynden, E., Vandamme, A. M., Deforche, K., & de Oliveira, T. (2019). Genome Detective: An automated system for virus identification from high-throughput sequencing data. *Bioinformatics*, 35(5), 871–873.
25. Struck, D., Lawyer, G., Ternes, A. M., Schmit, J. C., & Bercoff, D. P. (2014). COMET: Adaptive context-based modeling for ultrafast HIV-1 subtype identification. *Nucleic Acids Research*, 42(18).
26. Siepel A, Halpern A, Macken C, & Korber B. (1995). A Computer Program Designed to Screen Rapidly for HIV Type 1 Intersubtype Recombinant Sequences. *AIDS Research and Human Retroviruses*, 11(11).
27. Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W., & Ray, S. C. (1999). Full-Length Human Immunodeficiency Virus Type 1 Genomes from Subtype C-Infected Seroconverters in India, with Evidence of Intersubtype Recombination. In *JOURNAL OF VIROLOGY* (Vol. 73, Issue 1).
28. Katoh, K., Misawa, K., Kuma, K.-I., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform.
29. Larsson, A. (2014). AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, 30(22), 3276–3278.
30. Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32(1), 268–274.
31. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermini, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587–589.
32. Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307–321.

33. Hemelaar, J., Elangovan, R., Yun, J., Dickson-Tetteh, L., Fleminger, I., Kirtley, S., Williams, B., Gouws-Williams, E., Ghys, P. D., Abimiku, A. G., Agwale, S., Archibald, C., Avidor, B., Barbás, M. G., Barre-Sinoussi, F., Barugahare, B., Belabbes, E. H., Bertagnolio, S., Bix, D., & Zhang, R. (2018). Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *The Lancet Infectious Diseases*, 19(2), 143–155.
34. Roda, E. M., & Luiz, S. S. (2020). The tourism from Argentina in Uruguay: A quantitative behavioral analysis. *Cuadernos de Turismo*, 45, 239–264.
35. Espinosa, A., Vignoles, M., Carrillo, M. G., Sheppard, H., Donovan, R., Peralta, L. M., Rossi, D., Radulich, G., Salomón, H., & Weissenbacher, M. (2004). Intersubtype BF Recombinants of HIV-1 in a Population of Injecting Drug Users in Argentina. In *EPIDEMIOLOGY AND SOCIAL SCIENCE*.
36. Aulicino, P. C., Bello, G., Guimaraes, M. L., Ruchansky, D., Rocco, C., Mangano, A., Morgado, M. G., & Sen, L. (2011). Longitudinal analysis of HIV-1 BF1 recombinant strains in vertically infected children from Argentina reveals a decrease in CRF12_BF pol gene mosaic patterns and high diversity of BF unique recombinant forms. *Infection, Genetics and Evolution*, 11(2), 349–357.
37. Clemente, I. (2010). The Uruguayan Brazilian Border Region and Binational Relations: Past and Perspective. *Revista Uruguay de Ciencia Política*. 19(1).
38. Braga, A. & Riggati, D. (2012). Crossing the borderline takes you to heaven: working classes consumption tourism and frontier conurbations in South America. *International Forum on Urbanism*. "6th Conference of the International Forum on Urbanism (IFoU): TOURBANISM, Barcelona, 25-27 gener". Barcelona: IFoU, 2012, p. 1-13.
39. Patinõ-Galindo, J. A., Torres-Puente, M., Bracho, M. A., Alastrué, I., Juan, A., Navarro, D., Galindo, M. J., Ocete, D., Ortega, E., Gimeno, C., Belda, J., Domínguez, V., Moreno, R., & González-Candelas, F. (2017). The molecular epidemiology of HIV-1 in the Comunidad Valenciana (Spain): Analysis of transmission clusters. *Scientific Reports*, 7(1).
40. Le Vu, S., Ratmann, O., Delpech, V., Brown, A. E., Gill, O. N., Tostevin, A., Dunn, D., Fraser, C., & Volz, E. M. (2019). HIV-1 Transmission Patterns in Men Who Have Sex with Men: Insights from Genetic Source Attribution Analysis. *AIDS Research and Human Retroviruses*.
41. Shiino, T., Hachiya, A., Hattori, J., Sugiura, W., & Yoshimura, K. (2020). Nation-Wide Viral Sequence Analysis of HIV-1 Subtype B Epidemic in 2003–2012 Revealed a

- Contribution of Men Who Have Sex With Men to the Transmission Cluster Formation and Growth in Japan. *Frontiers in Reproductive Health*, 2.
42. Ávila, M. M., Pando, M. A., Carrion, G., Martinez Peralta, L., Salomon, H., Gomez Carrillo, M., Sanchez, J., Maulen, S., Hierholzer, J., Marinello, M., Negrete, M., Rusull, K. L., & Carr, J. K. (2002). Two HIV-1 Epidemics in Argentina: Different Genetic Subtypes Associated With Different Risk Groups. *Journal of Acquired Immune Deficiency Syndromes*, 29, 422–426.
 43. Gomez-Carrillo, M., Pampuro, S., Duran, A., Losso, M., Harris, D. R., Read, J. S., Duarte, G., de Souza, R., Soto-Ramirez, L., & Salomón, H. (2006). Analysis of HIV Type 1 Diversity in Pregnant Women from Four Latin American and Caribbean Countries. *AIDS Research and Human Retroviruses*, 22(11), 1186–1191.
 44. Dunn, D. (2014). The increasing genetic diversity of HIV-1 in the UK, 2002-2010. *AIDS*, 28(5), 773–780.
 45. Poon, A. F. Y., Gustafson, R., Daly, P., Zerr, L., Demlow, S. E., Wong, J., Woods, C. K., Hogg, R. S., Kraiden, M., Moore, D., Kendall, P., Montaner, J. S. G., & Harrigan, P. R. (2016). Near real-time monitoring of HIV transmission hotspots from routine HIV genotyping: An implementation case study. *The Lancet HIV*, 3(5), e231–e238.
 46. Kouyos, R.D., von Wyl, V., Yerly, S., Böni, J., Taffé, P., Shah, C., Bürgisser, P., Klimkait, T., Weber, R., Hirschel, B., Cavassini, M., Furrer, H., Battegay, M., Vernazza, P.L., Bernasconi, E., Rickenbach, M., Ledergerber, B., Bonhoeffer, S., Günthard, H.F. (2010) Molecular epidemiology reveals long-term changes in HIV type 1 subtype B transmission in Switzerland. *J Infect Dis.* 15;201(10):1488-97.
 47. Andrade, P., Arantes, I., Tanuri, A., Bello, G. & Gräf, T. Characterization of HIV-1 Transmission Clusters Inferred from the Brazilian Nationwide Genotyping Service Database. (2022) *Viruses*. 12;14(12):2768.
 48. Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L. N., Daly, J. M., Mumford, J. A., & Holmes, E. C. (2004). Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. *Science*, 327(303).
 49. Bello, G., Eyer-Silva, W. A., Couto-Fernandez, J. C., Guimarães, M. L., Chequer-Fernandez, S. L., Teixeira, S. L. M., & Morgado, M. G. (2007). Demographic history of HIV-1 subtypes B and F in Brazil. *Infection, Genetics and Evolution*, 7(2), 263–270.

FIGURES AND TABLES

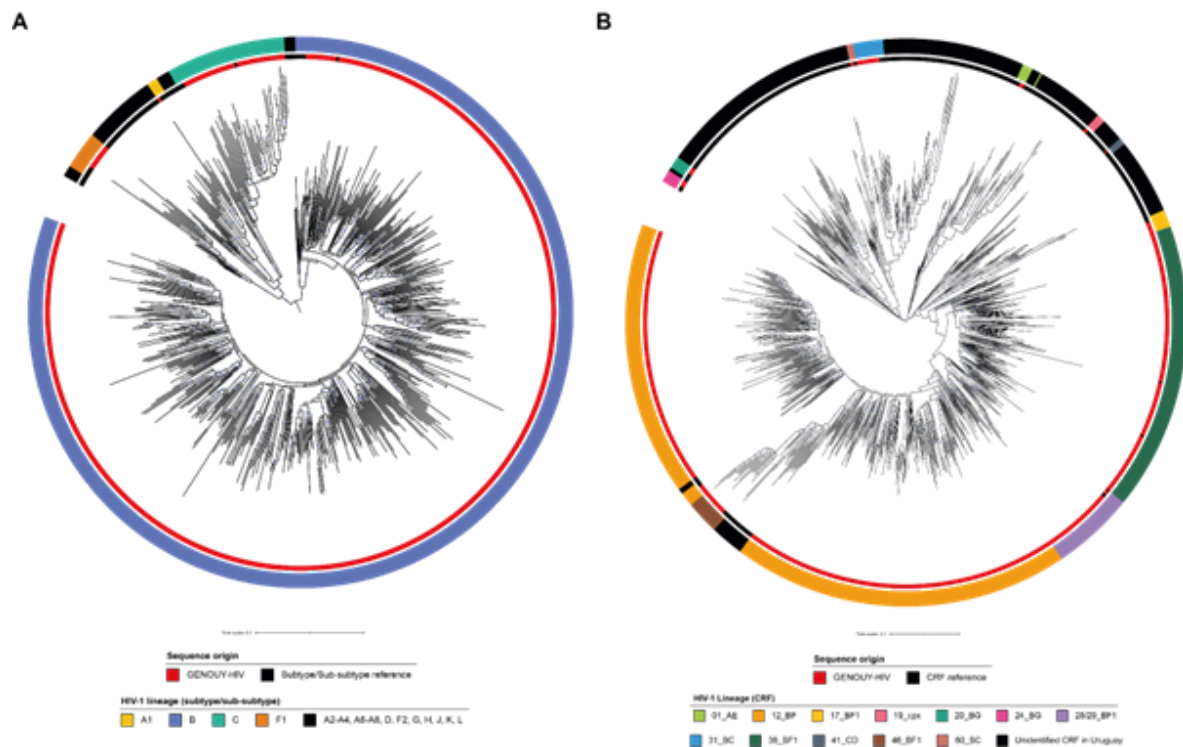


Figure 1. Maximum Likelihood phylogenetic inference of HIV-1 lineages circulating in Uruguay. (A) Subtypes and sub-subtypes tree. Based on 605 sequences, 549 from the GENOUY-HIV database and 56 reference sequences representing all pure HIV-1 subtypes and sub-subtypes. **(B) CRFs tree.** Based on 967 sequences, 611 from the GENOUY-HIV database and 356 reference sequences representing all described CRFs. In both trees, the inner ring indicates sequences from the GENOUY-HIV database (red) and reference sequences (black). The outer ring shows the different lineages identified (color legend). The tree was midpoint-rooted. All nodes marked with circles have SH-aLRT support ≥ 0.80 . Branch lengths are scaled to nucleotide substitutions per site. Reference sequences were obtained from the Los Alamos HIV Database.

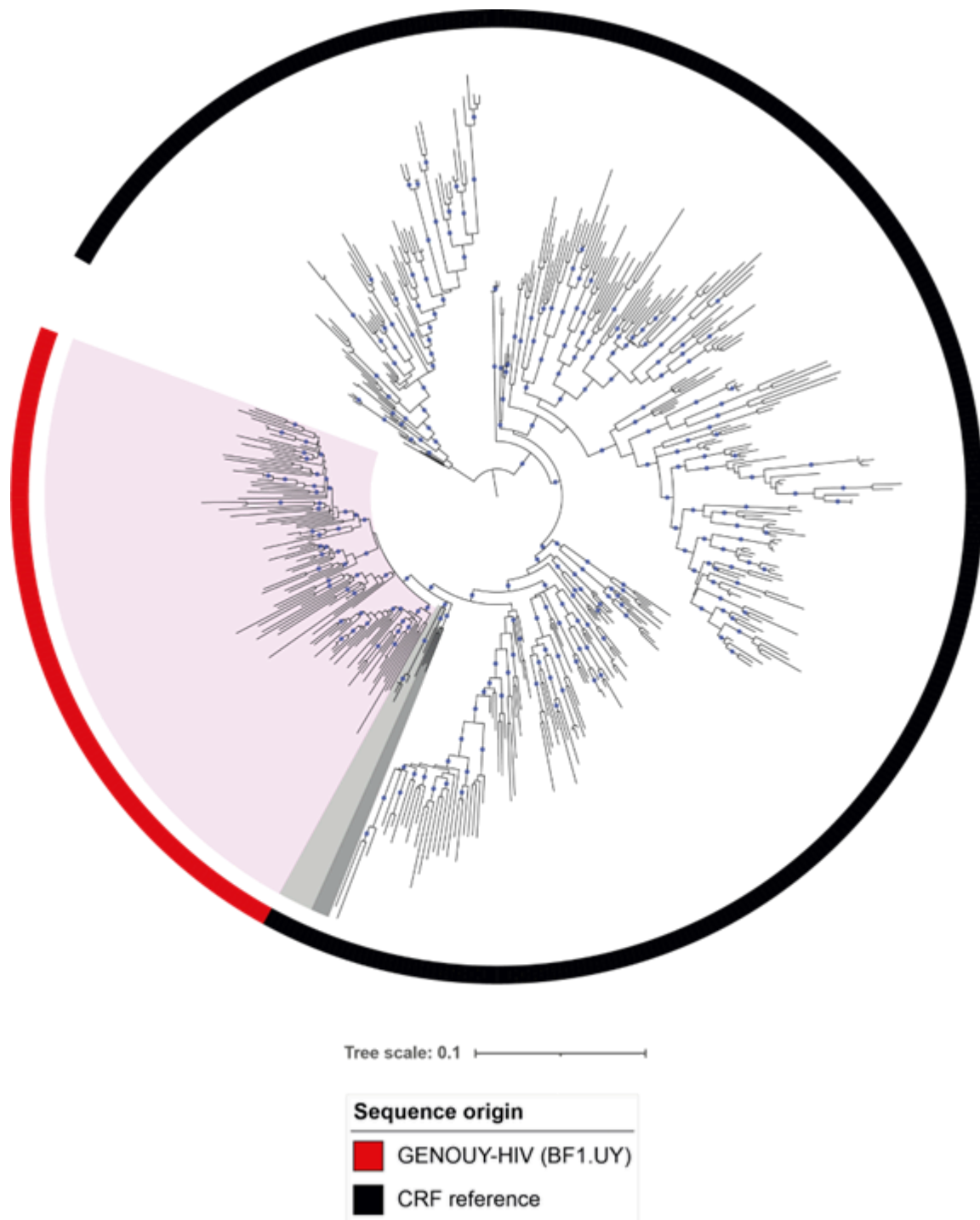


Figure 2. Maximum-likelihood phylogeny of unassigned B/F1 recombinant *pol* sequences. The analysis was based on 464 sequences, including 108 B/F1 recombinants from GENOUY-HIV and 356 reference sequences representing all described CRFs. The outer ring indicates GENOUY-HIV sequences (red) and reference sequences (black). The tree was midpoint-rooted. All branches marked with a dot have SH-aLRT support ≥ 0.80 . Branch lengths are shown on a scale of nucleotide substitutions per site. The BF1.UY clade, defined here as a strongly supported monophyletic group of B/F1 recombinants that does not include any CRF reference sequences, is highlighted in pink (SH-aLRT > 0.90). The light gray clade includes reference sequences corresponding to CRFs 38_BF1 and 42_BF1, while the darker gray clade indicates references to 17_BF1. Reference sequences were obtained from the Los Alamos HIV Database.

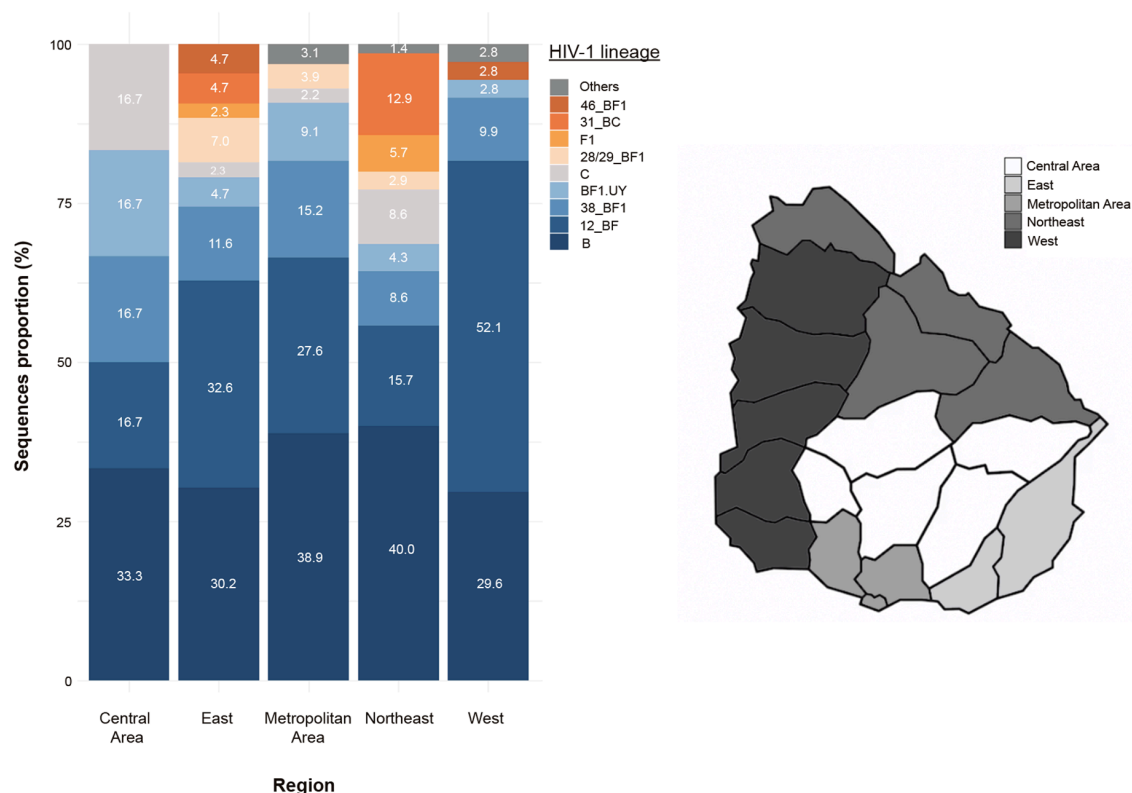
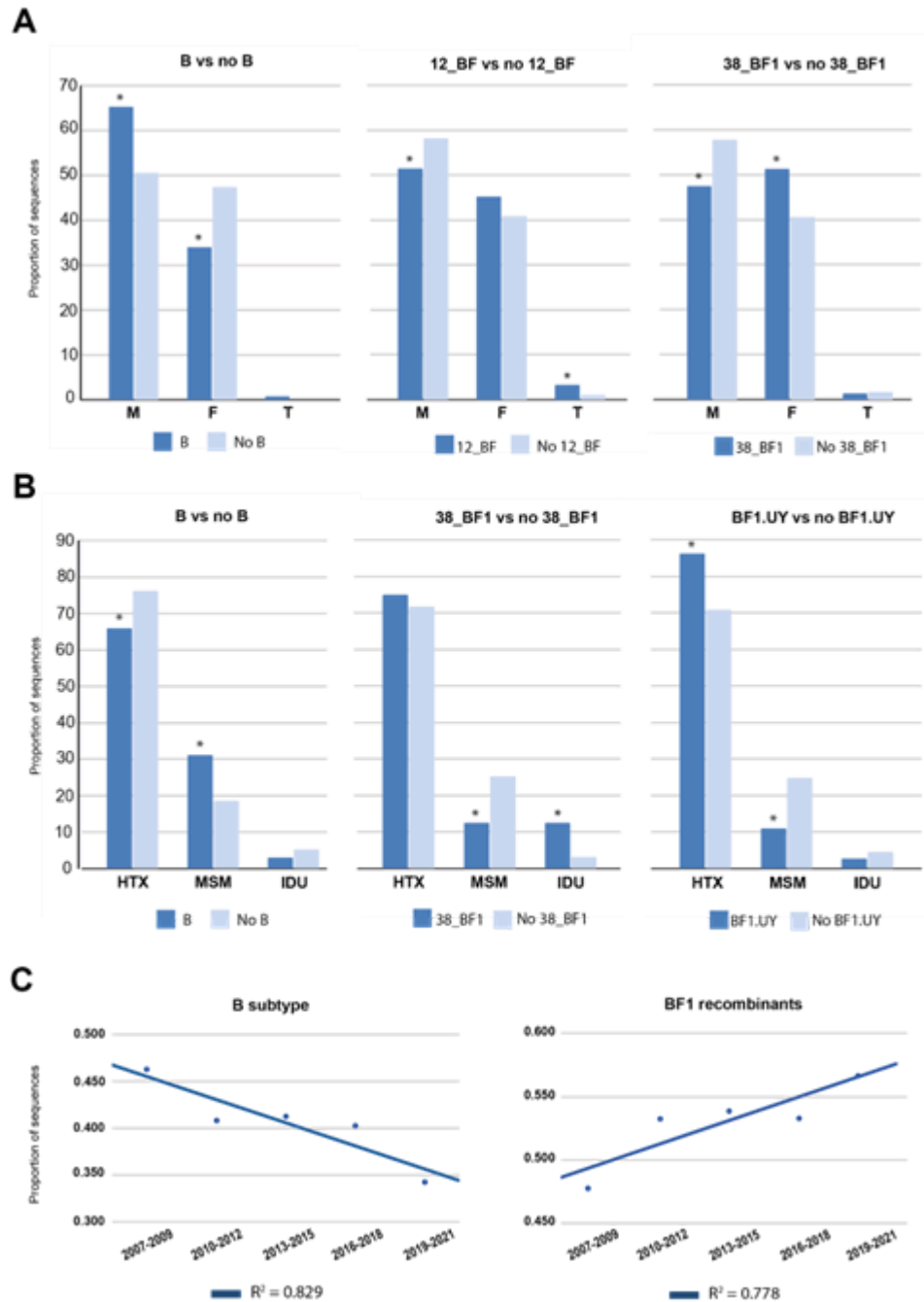


Figure 3. Geographic distribution of circulating HIV-1 lineages in Uruguay. Stacked bar charts show the proportion of HIV-1 lineages among sequences with available information on place of acquisition, stratified by macro-region (Northeast, West, Central, East and Metropolitan Area). The most frequent lineages at the national level are shown in blue, while lineages that are not common nationwide but reach $\geq 2.0\%$ within a given region are shown in orange. The “Others” category (grey) groups all remaining lineages with $< 2.0\%$ frequency in every region. On the right, a map of Uruguay indicates the boundaries and names of the macro-regions used in the analysis.



SUPPLEMENTARY FIGURES AND TABLES

Supplementary Table 1.

Lineage	Sequences (%)	Time range
B	498 (39.3)	2007-2021
12_BF	353 (27.8)	2007-2021
38_BF1	163 (12.9)	2007-2021
BF1.UY	108 (8.5)	2007-2021
28/29_BF1	43 (3.4)	2008-2021
C	40 (3.2)	2009-2021
46_BF1	15 (1.2)	2010-2021
31_BC	14 (1.1)	2011-2021
F1	10 (0.8)	2008-2021
17_BF1	7 (0.6)	2008-2019
20_BG	4 (0.3)	2014-2020
24_BG	4 (0.3)	2011-2021
01_AE	3 (0.2)	2012-2017
19_cpx	2 (0.2)	2021
60_BC	2 (0.2)	2017-2019
41_CD	1 (0.1)	2010
A1	1 (0.1)	2009

Supplementary Table 2.

Lineage	Gender			Total	p-value
	M (%)	F (%)	T (%)		
B	322 (65,2)	168 (34,0)	4 (0,8)	494	0.000
Other lineages combined (non B)	372 (50,5)	348 (47,3)	16 (2,2)	736	
12_BF	171 (51,5)	150 (45,2)	11 (3,3)	332	0.005
Other lineages combined (non 12_BF)	523 (58,2)	366 (40,8)	9 (1,0)	898	
38_BF1	75 (47,5)	81 (51,3)	2 (1,3)	158	0.036
Other lineages combined (non 38_BF1)	619 (57,7)	435 (40,6)	18 (1,7)	1072	
28/29_BF1	21 (48,8)	21 (48,8)	1 (2,3)	43	0.378
Other lineages combined (non 28/29_BF1)	673 (56,7)	495 (41,7)	19 (1,6)	1187	
C	19 (47,5)	21 (52,5)	0	40	0.380
Other lineages combined (non C)	675 (56,7)	495 (41,6)	20 (1,7)	1190	
BF1.UY	46 (46,5)	51 (51,5)	2 (2,0)	99	0.083
Other lineages combined (non BF1.UY)	648 (57,3)	465 (41,1)	18 (1,6)	1131	

Supplementary Table 3.

Lineage	Transmission route			Total	p-value
	HTX (%)	MSM (%)	IDU (%)		
B	223 (66,0)	105 (31,1)	10 (3,0)	338	0.000
Other lineages combined (non B)	393 (76,2)	96 (18,6)	27 (5,2)	516	
12_BF	169 (73,8)	50 (21,8)	10 (4,4)	229	0.777
Other lineages combined (non 12_BF)	447 (71,5)	151 (24,2)	27 (4,3)	625	
38_BF1	84 (75,0)	14 (12,5)	14 (12,5)	112	0.000
Other lineages combined (non 28/29_BF1)	532 (71,7)	187 (25,2)	23 (3,1)	742	
28/29_BF1	19 (70,4)	7 (25,9)	1 (3,7)	27	0.932
Other lineages combined (non 28/29_BF1)	597 (72,2)	194 (23,5)	36 (4,3)	827	
C	21 (80,8)	5 (19,2)	0	26	0.693
Other lineages combined (non C)	595 (71,9)	196 (23,7)	37 (4,5)	828	
BF1.UY	63 (86,3)	8 (11,0)	2 (2,7)	73	0.013
Other lineages combined (non BF1.UY)	553 (70,8)	193 (24,7)	35 (4,5)	781	

ANEXO 2. Manuscrito derivado de la presente tesis

Revista: -

Estado: Próximo a submitir

Spatiotemporal analysis of HIV-1 subtype B transmission networks in Uruguay

Authors: Podchibiakin, A. A.¹, Arantes, I.², Flieller, R.³, Ruchansky, D.³, Brasesco, M.³, Chiparelli, H.³, Bello, G.², Mir, D¹.

Affiliation:

¹Universidad de la República - CENUR, Department of Biological Sciences, Genomics and Bioinformatics Unit, Salto, Uruguay

²Fiocruz, Oswaldo Cruz Institute, Arbovirus and Hemorrhagic Virus Laboratory, Rio de Janeiro, Brazil

³Ministry of Public Health, Department of Public Health Laboratories, Virology Unit, Montevideo, Uruguay

***Corresponding authors: G. Bello, D. Mir**

Conflicts of Interest and Source of Funding: The authors have no conflict of interest. This work was supported by the Comisión Sectorial de Investigación Científica (CSIC–Uruguay) (CSIC_VUSP_Modalidad-2-2021), by the Fondo Carlos Vaz Ferreira (FVF) from the Dirección Nacional de Innovación, Ciencia y Tecnología (DICYT), Ministry of Education and Culture (MEC), Uruguay (Project FVF_2023_482), and by the L'Oréal Uruguay For Women in Science Fellowship Program. Agustina A. Podchibiakin is the recipient of a predoctoral scholarship (POS_NAC_2023_1_178397) from ANII (Uruguay).

ABSTRACT

Objective: Reconstruct the evolutionary and transmission dynamics of the HIV-1 subtype B epidemic in Uruguay and assess the relative contribution of cross-border introductions and local transmission networks.

Methods: A total of 491 HIV-1 subtype B pol sequences sampled in Uruguay between 2007 and 2021 were analyzed. Phylogenetic inference and clade assignment were performed using maximum likelihood approaches, integrating regional and global reference datasets. Transmission clusters were identified based on phylogenetic support and ancestral state reconstruction. Time-scaled phylogenies and demographic histories were inferred using Bayesian coalescent and birth–death models to estimate the timing of introductions, effective population size (N_e), and effective reproductive number (R_e).

Results: The Uruguayan HIV-1 subtype B epidemic is overwhelmingly dominated by the pandemic lineage (B_{PAN} ; 96.3%), with minimal contribution from Caribbean non-pandemic lineages (3.7%). A total of 229 independent viral introductions were identified, of which only 22.7% resulted in detectable local transmission clusters, indicating that most introductions failed to establish sustained transmission. The epidemic is largely driven by two major Uruguayan-specific clades ($B_{PAN-UY-I}$ and $B_{PAN-UY-II}$), both emerging in the mid-to-late 1980s and accounting for the largest and most persistent transmission networks. Phylodynamic analyses revealed sustained exponential growth from the 1980s until approximately 2010, followed by stabilization and recent declines in transmission ($R_e < 1$). Nearly half of transmission events occurred within the first two years post-infection. Phylogeographic reconstruction identified Brazil and Argentina as the primary sources of viral introductions into Uruguay.

Conclusions: The HIV-1 subtype B epidemic in Uruguay represents a mature and highly compartmentalized transmission system, characterized by frequent but largely unsuccessful viral introductions and long-term persistence driven by a limited number of early-established founder lineages. These findings highlight the importance of historical contingency and transmission network structure in shaping epidemic dynamics and underscore the continued role of regional connectivity in sustaining viral diversity.

Keywords: HIV-1; subtype B; molecular epidemiology; Uruguay.

INTRODUCTION

The Human Immunodeficiency Virus Type 1 Subtype B (HIV-1B) gave rise to the earliest large-scale epidemics established outside Africa, initiating the global phase of the HIV pandemic [1,2]. Although HIV-1B represents only ~10–12% of infections worldwide, it is the most geographically dispersed HIV genetic variant, predominating in several regional epidemics and serving as the reference model for molecular and translational research [3–5]. In particular, HIV-1B accounts for >90% of cases in the Caribbean, ~76% in Latin America, and >80% in Western and Central Europe and North America, while remaining

relatively rare in sub-Saharan Africa. Historical and phylogenetic evidence trace its origin to Central Africa (Kinshasa) in the mid-20th century, followed by a single introduction into Haiti in the mid-1960s. A subsequent founder event from Haiti to the United States in the late 1960s ultimately drove its expansion across the Americas and Europe [1,2,6].

In Latin America, the HIV-1B epidemic comprises two phylogenetically distinct components of viral diversity, reflecting alternative evolutionary trajectories following the early establishment of the virus in the Caribbean. One corresponds to the pandemic lineage (B_{PAN}), which established itself in North America and spread across the continent, now accounting for more than 90% of Subtype B infections in most mainland Latin American countries [6,7]. The other comprises a set of Caribbean-derived clades (B_{CAR}), originating from early diversification within the Caribbean basin [7,8]. While B_{CAR} clades dominate several Caribbean island epidemics, representing a persistent component of local transmission, they are detected only sporadically or at low prevalence (1–10%) in most of mainland Latin America [7,9].

Following its introduction into Latin America, B_{PAN} diversified into a limited number of large transmission clades. These B_{PAN} clades were established through multiple introductions between the late 1960s and the 1980s and subsequently expanded within specific national and sub-regional contexts, giving rise to stable and long-lasting epidemic lineages [10]. Prominent examples include several country-specific clades that have become central to national epidemics in Argentina, Brazil, El Salvador, Honduras, Mexico and Peru, as well as two regional clades spanning Argentina/Brazil and Honduras/El Salvador. This pattern reveals a regional epidemic characterized by strong geographic compartmentalization and predominantly within-country transmission of B_{PAN} founder clades [10,11].

Within this regional framework, the evolutionary history of HIV-1B in Uruguay remains largely unexplored. Nationwide molecular surveillance indicates that HIV-1B predominates in Uruguay (39.3%) and is strongly associated with transmission among men who have sex with men [12]. Yet, the relative contribution of the B_{PAN} and B_{CAR} clades to the national epidemic, their frequency of introduction, and the long-term persistence of local transmission networks have not been assessed. To address these gaps, this study conducts a comprehensive, nationwide phylodynamic analysis. The objective is to reconstruct the evolutionary and demographic history of HIV-1B in Uruguay, as well as to infer its cross-border and local transmission patterns, thereby placing national dynamics within their proper regional context to better understand the forces shaping this epidemic.

METHODS

Subtype B clade assignment A total of 491 HIV-1B pol sequences (HXB2 positions 2358–3251), generated from samples collected in Uruguay between 2007 and 2021, were analyzed. All sequences were provided by the Department of Public Health Laboratories (DLSP, Ministry of Public Health, Uruguay), stored in the GENOUY-HIV national database, and previously quality-checked and subtyped [12]. GenBank accession numbers for all sequences analyzed in this study are provided in Supplementary Table 3. For clade assignment, the dataset was aligned with curated reference sequences comprising B_{PAN} sequences from United States (USA) and France (n = 300), B_{CAR} lineage from the Caribbean (n = 200) [13,14] and major B_{PAN} Latin American clusters [10]. Clade placement was performed using the Evolutionary Placement Algorithm (EPA) implemented in RAXML v8.2.4 [15]. To further characterize the placement of Uruguayan sequences within B_{PAN} diversity and delineate Uruguayan B_{PAN} clades, B_{PAN}-assigned sequences were then stratified into regional datasets and re-analyzed together with their corresponding reference panels (Table S1). Sequences classified within South American B_{PAN} clusters were aligned with reference sequences from B_{PAN}-SAM, B_{PAN}-AR-II, B_{PAN}-AR-III, B_{PAN}-BR-II, B_{PAN}-BR-III, B_{PAN}-BR-IV, and B_{PAN}-PE; those falling within Central American diversity were analysed with B_{PAN}-CAM references; and those placed within Mexican diversity were analyzed with B_{PAN}-MX-I and B_{PAN}-MX-II. Uruguayan sequences assigned to B_{PAN} but not falling within any previously described Latin American cluster were aligned with B_{PAN} reference sequences from the USA and France. Independent ML phylogenetic analyses were then performed for each B_{PAN} Latin American cluster and for the sequences that branched outside the known clades, resulting in a total of 11 phylogenetic inferences. All phylogenetic trees were reconstructed with PhyML [16,17] using the Subtree Pruning and Regrafting (SPR) heuristic search strategy. Branch support was assessed using the approximate likelihood-ratio test (aLRT) based on the Shimodaira-Hasegawa-like procedure [18]. All trees were visualized using FigTree v1.4.4 program [19].

ML phylogeographic analyses ML phylogenetic trees were evaluated for temporal signal and best-fitting root using TempEst v1.5 [20] and subsequently time-scaled with TreeTime v.0.8.6 [21]. Ancestral geographic reconstruction was performed with PastML [22], under the maximum likelihood/maximum a posteriori (MAP) approach and an F81-like model. Uruguayan transmission clusters were defined as highly supported monophyletic clades (aLRT ≥ 0.85) whose ancestral node was inferred to be in Uruguay and that were composed predominantly (>90%) of Uruguayan sequences. Clades were classified according to size as large (n > 30), medium (n = 10-30), or small (n = 2-9).

Bayesian time-scale evolutionary analyses The time of the most recent common ancestor (T_{MRCA}) of all B_{PAN} Uruguayan clades was estimated using a Bayesian Markov Chain Monte

Carlo (MCMC) approach, as implemented in BEAST v1.10.4 [23,24], with BEAGLE to improve run-time [25]. Dated phylogenies were inferred with the flexible Bayesian Skyline coalescent model [26], the GTR + I + Γ 4 nucleotide substitution model, and a relaxed uncorrelated lognormal molecular clock [27]. A uniform prior was applied to the clock rate ($1.7\text{--}2.1 \times 10^{-3}$ substitutions/site/year), based on estimates from previous studies [14,28]. MCMC chains were run for $10\text{--}30 \times 10^7$ generations to ensure Effective Sample Size (ESS) values above 200. Posterior distributions and 95% Highest Posterior Density (HPD) intervals were inspected using Tracer v1.7.1 [29]. Maximum clade credibility trees (MCCT) were summarized with TreeAnnotator v1.10.4 and visualized in FigTree v1.4.4.

Bayesian demographic reconstructions The demographic history of major B_{PAN} Uruguayan clades was next reconstructed using Bayesian MCMC approaches. Changes across time in their effective number of infections (N_e) were estimated using the coalescent Bayesian Skygrid (BSKG) model [30], as implemented in BEAST v1.10.4 [23,24]. Changes across time in their effective reproductive number (R_e) were estimated using the Birth-death Skyline (BDSKY) model [31], as implemented in BEAST 2.1 [32]. For BDSKY, the sampling rate (δ) was set to zero for the period before the oldest sample and estimated afterward. The BDSKY prior settings were as follows: becoming non-infectious rate (log-normal, mean = 0.25; standard deviation = 0.5); reproductive number (log-normal, mean = 0.0, s.d. = 1.0); sampling proportion (beta, alpha = 1.0, beta = 25.0). The origin parameter was conditioned to root height, and R_e was estimated in a piecewise manner over eight time intervals defined from the date of the most recent sample up to the root of the tree. The MCMC chain was run until all relevant parameters reached ESS >200. The inferred temporal evolution of the R_e was plotted with the bdskytool R package (<https://github.com/laduplessis/bdskytools>). The timing of secondary transmission events was inferred using TreeStat software [24] from internode distances in MCCTs and grouped into discrete post-infection intervals (<0.5, 0.5–2, 2–5, and >5 years).

RESULTS

HIV-1B introductions and local transmission clusters. ML phylogenetic analyses showed that the vast majority of Uruguayan HIV-1B sequences ($n = 473$, 96.3%) clustered within B_{PAN} , whereas only 3.7% ($n = 18$) were assigned to B_{CAR} (Supplementary Figure 1). Among B_{PAN} sequences, 37.0% ($n = 175$) clustered within previously described country- or regional-specific B_{PAN} clades (as defined by Mir et al., 2015), while a substantial proportion (63.0%, $n = 298$) did not associate with any B_{PAN} clade described to date (Table 1). These sequences were analyzed in an independent phylogenetic inference ($B_{\text{PAN-UY}}$), which resolved them into a mixture of Uruguayan-specific transmission subclades and phylogenetically isolated introductions. Across the 11 ML phylogenies inferred for the major

B_{PAN} clades and for the B_{PAN-UY} dataset, 229 independent introductions of B_{PAN} viruses into Uruguay were identified; 52 of these introductions gave rise to detectable local transmission clusters, whereas the remaining 177 corresponded to sporadic introductions without evidence of onward local spread (singletons). Among the 52 transmission clusters, 2 were large (3.8%; $n \geq 30$), 9 were medium-sized (17.3%; $10 \leq n < 30$), and 41 were small (78.9%; $2 \leq n < 10$) (Figure 1A). Notably, the two largest clusters were detected exclusively within B_{PAN-UY} subclades and were designated $B_{PAN-UY-I}$ ($n = 37$) and $B_{PAN-UY-II}$ ($n = 30$). Medium and small transmission clusters were identified within seven B_{PAN} clades previously described by Mir et al. (2015) ($B_{PAN-SAM}$, $B_{PAN-AR-II}$, $B_{PAN-AR-III}$, $B_{PAN-BR-II}$, $B_{PAN-BR-III}$, $B_{PAN-MX-I}$, $B_{PAN-MX-II}$), as well as across multiple B_{PAN-UY} subclades. In contrast, $B_{PAN-BR-IV}$, $B_{PAN-CAM}$, and B_{PAN-PE} were restricted to sporadic introductions only (Supplementary Figures 2–3).

Bayesian time-scale evolutionary analyses and demographic reconstruction. Bayesian time-scaled phylogenies were inferred for South American B_{PAN} clusters involving Uruguayan transmission networks (B_{PAN-UY} , $B_{PAN-SAM}$, $B_{PAN-AR-II}$, $B_{PAN-AR-III}$, $B_{PAN-BR-II}$, and $B_{PAN-BR-III}$). The overall mean substitution rate was 1.8×10^{-3} substitutions/site/year (95% HPD: 1.7×10^{-3} – 2.1×10^{-3}) and the coefficient of variation of branch-specific rates was 0.36 (95% HPD: 0.26–0.47), indicating substantial among-lineage rate heterogeneity and supporting the use of a relaxed uncorrelated lognormal clock model. Within these clade-specific phylogenies, introductions of B_{PAN} lineages into Uruguay were dated between 1978 (95% HPD: 1972–1985) and 2013 (95% HPD: 2008–2016), and most events between 1995 and 2015 were not associated with the establishment of sustained local transmission networks (Figure 1B). The oldest inferred clade (mean T_{MRCA} : 1978; 95% HPD: 1972–1985) corresponded to an intermediate-sized transmission cluster within the B_{PAN-UY} phylogeny (Supplementary Table 1–2). The mean T_{MRCA} estimates for the two major Uruguayan B_{PAN} subclades, $B_{PAN-UY-I}$ and $B_{PAN-UY-II}$, were 1985 (95% HPD: 1978–1989) and 1987 (95% HPD: 1978–1988), respectively, consistent with their independent emergence in the mid-to-late 1980s (Figure 2A).

To further characterize the evolutionary and demographic history of the Uruguayan B_{PAN} epidemic, a joint phylodynamic analysis was performed for $B_{PAN-UY-I}$ and $B_{PAN-UY-II}$, the two major and most densely sampled Uruguayan subclades, providing robust signal for parameter estimation. A BSKG model was employed to reconstruct their combined N_e through time, and a BDSKY model was applied to infer temporal variation in the R_e (Figure 2B). The BSKG analysis indicated that the combined B_{PAN-UY} population experienced a sustained phase of exponential growth beginning in the mid-1980s and extending until approximately 2010, reflected by a marked and continuous increase in N_e . Under the BDSKY model, R_e was estimated as piecewise constant across eight equidistant intervals

spanning from 1981 to 2021. Posterior median R_e values were >1 during the earliest intervals, corresponding to the late 1980s and early 2000s, and subsequently declined toward values close to 1 during the mid-2000s to approximately 2015. In the most recent intervals, posterior median R_e values remained <1 , consistent with a transition toward reduced transmission intensity (Figure 2B). Overall, the combined BSKG–BDSKY analyses support a temporal pattern characterized by early epidemic expansion, followed by partial stabilization and a more recent decline in transmission within the major Uruguayan B_{PAN-UY} subclades.

Finally, analysis of internode distances in MCCTs indicated that nearly half of all inferred secondary transmission events (49.3%) occurred within the first two years following infection. An additional 29.2% of transmission events occurred between 2 and 5 years post-infection, while the remaining 21.5% took place more than 5 years after infection (Figure 2C).

South American sources of the Uruguayan B_{PAN} epidemic. Ancestral state reconstruction of geographic origin was restricted to South American B_{PAN} clusters contributing to the Uruguayan epidemic and embedded within previously described regional transmission networks ($B_{PAN-SAM}$, $B_{PAN-AR-II}$, $B_{PAN-AR-III}$, $B_{PAN-BR-II}$, and $B_{PAN-BR-III}$). PastML analyses consistently identified Brazil as the primary source of Uruguayan B_{PAN} diversity across these lineages. Among the 20 Uruguayan transmission clades identified within them, 16 (80.0%) were inferred to have originated in Brazil, whereas 4 (20.0%) were assigned an Argentine origin. A similar pattern was observed for independent introductions that did not result in detectable local transmission clusters: among 61 such Uruguayan sequences, ancestral state reconstructions assigned 45 (73.8%) to a Brazilian and 16 (26.2%) to an Argentine origin (Supplementary Table 2).

DISCUSSION

Across Latin America, HIV-1B epidemics are structured by repeated viral introductions into partially isolated transmission systems, generating strong geographic and network-level compartmentalization and pronounced heterogeneity in lineage-level transmission success [10,33]. The Uruguayan epidemic conforms closely to this phylodynamic regime and is characterized by a high frequency of introductions that fail to establish sustained local transmission. Despite continuous viral influx from a wide range of B_{PAN} lineages circulating regionally, most introductions remained epidemiologically unsuccessful, either persisting as isolated sequences or generating only limited local clustering. This pattern was particularly evident for several lineages that have achieved substantial epidemic expansion elsewhere in Latin America, such as $B_{PAN-BR-IV}$, $B_{PAN-CAM}$ and B_{PAN-PE} , which in Uruguay were detected

exclusively as sporadic introductions without evidence of onward transmission. Even among frequently introduced South American clades ($B_{\text{PAN-SAM}}$, $B_{\text{PAN-AR-II}}$, $B_{\text{PAN-AR-III}}$, $B_{\text{PAN-BR-II}}$ and $B_{\text{PAN-BR-III}}$) and Mexican lineages ($B_{\text{PAN-MX-I}}$ and $B_{\text{PAN-MX-II}}$), local amplification was generally modest, indicating strong constraints on transmission establishment following introduction. Non-pandemic Caribbean lineages (B_{CAR}) contributed only marginally to the Uruguayan epidemic and similarly failed to establish sustained transmission, consistent with their limited epidemiological impact in the Southern Cone [7].

Against this background of frequent but largely transmission-limited introductions, only two Uruguayan-specific B_{PAN} subclades, $B_{\text{PAN-UY-I}}$ and $B_{\text{PAN-UY-II}}$, account for the majority of long-term epidemic persistence in the country. These subclades comprise the largest and most durable transmission clusters identified in Uruguay and can be interpreted as national-scale counterparts of the major B_{PAN} founder clades that sustain long-term epidemics elsewhere in Latin America. Time-scaled phylogenetic analyses indicate that both lineages emerged independently during the mid-to-late 1980s, coinciding with the early expansion phase of the subtype B epidemic in the Southern Cone and the diversification of B_{PAN} lineages across the region [10,11,14,34]. In contrast, viral introductions occurring after the mid-1990s were markedly less likely to establish sustained transmission, predominantly resulting in small clusters or unclustered infections. This temporal asymmetry mirrors patterns reported for other South American subtype B epidemics, in which older clades tend to exhibit longer growth phases and larger effective population sizes than more recent lineages [10–14], consistent with a phylodynamic regime in which early lineage establishment and historical contingency play a central role in shaping long-term epidemic structure.

Phylodynamic analyses of the two dominant Uruguayan B_{PAN} lineages indicate that the HIV-1B epidemic has reached a mature phase. The inferred trajectories are characterized by sustained expansion from the mid-1980s until approximately 2010, followed by stabilization and more recent signals of declining transmission, in line with patterns reported for long-established subtype B epidemics in Latin America [10,14,35]. Within this mature epidemic context, R_e estimates remained persistently above the epidemic threshold for several decades, approached criticality during the mid-2000s to mid-2010s, and declined to subcritical values thereafter. This temporal transition in R_e coincides with the period following the early universalization of antiretroviral therapy in Uruguay (initiated in 1996–1997) [36], suggesting a gradual population-level impact on transmission consistent with treatment-as-prevention frameworks [37,38]. The partial divergence between coalescent- and birth–death-based reconstructions in the most recent periods is expected in

advanced epidemic settings and reflects the reduced sensitivity of coalescent-based approaches to near-present transmission dynamics [10,35].

Temporal reconstruction of transmission events based on internode distances indicates that a substantial fraction of HIV-1B spread in Uruguay occurs relatively early after infection, with nearly half of inferred transmission events concentrated within the first two years (49.3%) and more than three quarters occurring within the first five years. This distribution is consistent with established models of HIV transmission heterogeneity, in which elevated infectiousness during early and recent infection, together with delays in diagnosis and treatment initiation, generates a disproportionate contribution to onward spread [39,40]. Notably, a considerable fraction of inferred transmission events (21.5%) occurs more than five years after infection, indicating that epidemiologically relevant transmission is not restricted to acute or recent stages but persists among individuals who remain untreated, disengaged from care, or are diagnosed late. Similar contributions of long-standing infections to subtype B transmission have been reported in regional phylodynamic analyses, highlighting the role of delayed diagnosis and incomplete retention in sustaining transmission beyond early infection windows [41]. Taken together, these results indicate that while early post-infection stages concentrate much of the transmission potential, failures along the diagnosis–treatment continuum allow a non-negligible fraction of transmission to accumulate over extended periods, contributing to epidemic persistence in a mature transmission landscape.

Within this constrained epidemic context, phylogeographic reconstructions indicate that the Uruguayan HIV-1 subtype B epidemic is largely shaped by repeated cross-border introductions, predominantly from Brazil and, to a lesser extent, from Argentina. This pattern is consistent with a Latin American dissemination regime dominated by short-range cross-border spread, strong geographic compartmentalization, and limited long-distance mixing [10,33]. In contrast, subtype B epidemics in Europe are driven by dense transnational networks and frequent long-range viral exchanges, resulting in weaker spatial structure and highly intermixed phylogenies [42].

These findings should be interpreted in light of several limitations. The dataset is derived from routine antiretroviral resistance genotyping within the national HIV surveillance framework and therefore represents a non-random subset of the infected population. It is enriched for individuals engaged in clinical care and with detectable viremia, as resistance testing is primarily performed under specific clinical indications [43]. In addition, because sampling often occurs years after infection, temporal uncertainty may affect estimates of lineage emergence and recent transmission dynamics. Analyses were restricted to a partial

pol region, limiting phylogenetic resolution and potentially leading to underestimation of cluster size, misclassification of introductions, and incomplete detection of recombination events. Moreover, incomplete sampling of transmission chains may lead to phylogenetic fragmentation, causing multiple sequences to be interpreted as independent introductions and thus biasing estimates of the number of introduction events.

Overall, the HIV-1 subtype B epidemic in Uruguay exemplifies a mature, highly compartmentalized phylodynamic system in which long-term epidemic persistence is shaped by the interplay between early lineage establishment, selective amplification within highly connected transmission networks, and persistent structural gaps along the diagnosis–treatment continuum. Rather than being sustained by continuous viral influx, persistence is driven by the long-term success of a limited number of early-established founder lineages. This pattern situates Uruguay within a broader Latin American phylodynamic regime, in which contemporary subtype B epidemic structure reflects not only ongoing transmission processes but also the enduring imprint of early historical contingencies and network formation, thereby linking regional emergence to present-day persistence.

BIBLIOGRAFÍA

1. Gilbert, M. T., Rambaut, A., Wlasiuk, G., Spira, T. J., Pitchenik, A. E., & Worobey, M. (2007). The emergence of HIV/AIDS in the Americas and beyond. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47), 18566–18570. <https://doi.org/10.1073/pnas.0705329104>
2. Faria, N. R., Rambaut, A., Suchard, M. A., Baele, G., Bedford, T., Ward, M. J., Tatem, A. J., Sousa, J. D., Arinaminpathy, N., Pépin, J., Posada, D., Peeters, M., Pybus, O. G., & Lemey, P. (2014). The early spread and epidemic ignition of HIV-1 in human populations. *Science*, 346(6205), 56–61. <https://doi.org/10.1126/science.1256739>
3. Nair, M., Gettins, L., Fuller, M., Kirtley, S., & Hemelaar, J. (2024). Global and regional genetic diversity of HIV-1 in 2010–21: systematic review and analysis of prevalence. *The Lancet Microbe*, 5(11):100912. [https://doi.org/10.1016/S2666-5247\(24\)00151-4](https://doi.org/10.1016/S2666-5247(24)00151-4)
4. Hemelaar, J., Elangovan, R., Yun, J., Dickson-Tetteh, L., Fleminger, I., Kirtley, S., Williams, B., Gouws-Williams, E., Ghys, P. D., Abimiku, A. G., Agwale, S., Archibald, C., Avidor, B., Barbás, M. G., Barre-Sinoussi, F., Barugahare, B., Belabbes, E. H., Bertagnolio, S., Bix, D., & Zhang, R. (2019). Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *The Lancet Infectious Diseases*, 19(2), 143–155. [https://doi.org/10.1016/S1473-3099\(18\)30647-9](https://doi.org/10.1016/S1473-3099(18)30647-9)

5. Junqueira, D. M., & de Matos Almeida, S. E. (2016). HIV-1 subtype B: Traces of a pandemic. *Virology* 495, 173–184. <https://doi.org/10.1016/j.virol.2016.05.003>
6. Worobey, M., Watts, T. D., McKay, R. A., Suchard, M. A., Granade, T., Teuwen, D. E., Koblin, B. A., Heneine, W., Lemey, P., & Jaffe, H. W. (2016). 1970s and 'Patient 0' HIV-1 genomes illuminate early HIV/AIDS history in North America. *Nature*, 539(7627), 98–101. <https://doi.org/10.1038/nature19827>
7. Cabello, M., Junqueira, D. M., & Bello, G. (2015). Dissemination of nonpandemic Caribbean HIV-1 subtype B clades in Latin America. *AIDS*, 29(4), 483–492. <https://doi.org/10.1097/QAD.0000000000000552>
8. Arantes, I., Esashika Crispim, M. A., Nogueira da Guarda Reis, M., Martins Araújo Stefani, M., & Bello, G. (2019). Reconstructing the Dissemination Dynamics of the Major HIV-1 Subtype B Non-Pandemic Lineage Circulating in Brazil. *Viruses*, 11(10), 909. <https://doi.org/10.3390/v11100909>
9. Junqueira, D. M., de Medeiros, R. M., Matte, M. C., Araújo, L. A., Chies, J. A., Ashton-Prolla, P., & Almeida, S. E. (2011). Reviewing the history of HIV-1: spread of subtype B in the Americas. *PloS One*, 6(11), e27489. <https://doi.org/10.1371/journal.pone.0027489>
10. Mir, D., Cabello, M., Romero, H., & Bello, G. (2015). Phylodynamics of major HIV-1 subtype B pandemic clades circulating in Latin America. *AIDS*, 29(14), 1863–1869. <https://doi.org/10.1097/QAD.0000000000000770>
11. Junqueira, D. M., de Medeiros, R. M., Gräf, T., & Almeida, S. E. (2016). Short-Term Dynamic and Local Epidemiological Trends in the South American HIV-1B Epidemic. *PloS One*, 11(6), e0156712. <https://doi.org/10.1371/journal.pone.0156712>
12. Podchibiakin, A. A., Flieller, R., Ruchansky, D., Brasesco, M., Cortinas, M. N., Lizasoain, A., Chiparelli, H., Cabrera, S., Bello, G., Soler, A. M., & Mir, D. Nationwide molecular epidemiology of HIV-1 in Uruguay (2007–2021): lineage diversity, BF1 recombinant complexity, and epidemiological patterns. Submitted for publication.
13. Cabello, M., Mendoza, Y., & Bello, G. (2014). Spatiotemporal dynamics of dissemination of non-pandemic HIV-1 subtype B clades in the Caribbean region. *PloS One*, 9(8), e106045. <https://doi.org/10.1371/journal.pone.0106045>
14. Mendoza, Y., Martínez, A. A., Castillo Mewa, J., González, C., García-Morales, C., Avila-Ríos, S., Reyes-Terán, G., Armien, B., Pascale, J. M., & Bello, G. (2014). Human immunodeficiency virus type 1 (HIV-1) subtype B epidemic in Panama is mainly driven by

dissemination of country-specific clades. *PloS One*, 9(4), e95360. <https://doi.org/10.1371/journal.pone.0095360>

15. Stamatakis A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>

16. Guindon, S., Lethiec, F., Duroux, P., & Gascuel, O. (2005). PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic acids research*, 33(Web Server issue), W557–W559. <https://doi.org/10.1093/nar/gki352>

17. Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology*, 59(3), 307–321. <https://doi.org/10.1093/sysbio/syq010>

18. Anisimova, M., & Gascuel, O. (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Systematic biology*, 55(4), 539–552. <https://doi.org/10.1080/10635150600755453>

19. Rambaut A. (2018), *FigTree*. In (Version 1.4). Figtree. <http://tree.bio.ed.ac.uk/software/figtree/>. accessed May 7 2025.

20. 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 evolution*, 2(1), vew007. <https://doi.org/10.1093/ve/vew007>

21. Sagulenko, P., Puller, V., & Neher, R. A. (2018). TreeTime: Maximum-likelihood phylodynamic analysis. *Virus evolution*, 4(1), vex042. <https://doi.org/10.1093/ve/vex042>

22. Ishikawa, S. A., Zhukova, A., Iwasaki, W., & Gascuel, O. (2019). A Fast Likelihood Method to Reconstruct and Visualize Ancestral Scenarios. *Molecular biology and evolution*, 36(9), 2069–2085. <https://doi.org/10.1093/molbev/msz131>

23. Drummond, A. J., Nicholls, G. K., Rodrigo, A. G., & Solomon, W. (2002). Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*, 161(3), 1307–1320. <https://doi.org/10.1093/genetics/161.3.1307>

24. 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 evolution*, 4(1), vey016. <https://doi.org/10.1093/ve/vey016>

25. Suchard, M. A., & Rambaut, A. (2009). Many-core algorithms for statistical phylogenetics. *Bioinformatics*, 25(11), 1370–1376. <https://doi.org/10.1093/bioinformatics/btp244>
26. Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular biology and evolution*, 22(5), 1185–1192. <https://doi.org/10.1093/molbev/msi103>
27. Drummond, A. J., Ho, S. Y., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology*, 4(5), e88. <https://doi.org/10.1371/journal.pbio.0040088>
28. Divino, F., de Lima Guerra Corado, A., Gomes Naveca, F., Stefani, M. M., & Bello, G. (2016). High Prevalence and Onward Transmission of Non-Pandemic HIV-1 Subtype B Clades in Northern and Northeastern Brazilian Regions. *PloS One*, 11(9), e0162112. <https://doi.org/10.1371/journal.pone.0162112>
29. Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic biology*, 67(5), 901–904. <https://doi.org/10.1093/sysbio/syy032>
30. Gill, M. S., Lemey, P., Faria, N. R., Rambaut, A., Shapiro, B., & Suchard, M. A. (2013). Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Molecular Biology and Evolution*, 30(3), 713–724. <https://doi.org/10.1093/molbev/mss265>
31. Stadler, T., Kouyos, R., von Wyl, V., Yerly, S., Böni, J., Bürgisser, P., Klimkait, T., Joos, B., Rieder, P., Xie, D., Günthard, H. F., Drummond, A. J., Bonhoeffer, S., & Swiss HIV Cohort Study (2012). Estimating the basic reproductive number from viral sequence data. *Molecular Biology and Evolution*, 29(1), 347–357. <https://doi.org/10.1093/molbev/msr217>
32. Russel, P. M., Brewer, B. J., Klaere, S., & Bouckaert, R. R. (2019). Model Selection and Parameter Inference in Phylogenetics Using Nested Sampling. *Systematic Biology*, 68(2), 219–233. <https://doi.org/10.1093/sysbio/syy050>
33. Pagán, I., & Holguín, Á. (2013). Reconstructing the Timing and Dispersion Routes of HIV-1 Subtype B Epidemics in The Caribbean and Central America: A Phylogenetic Story. *PLoS One*, 8(7). <https://doi.org/10.1371/journal.pone.0069218>
34. Bello, G., Guimarães, M. L., & Morgado, M. G. (2006). Evolutionary history of HIV-1 subtype B and F infections in Brazil. *AIDS*, 20(5), 763–768. <https://doi.org/10.1097/01.aids.0000216377.84313.52>

35. Arantes, I., Gräf, T., Andrade, P., Oliveira Chaves, Y., Guimarães, M. L., & Bello, G. (2022). Dissemination Dynamics of HIV-1 Subtype B Pandemic and Non-pandemic Lineages Circulating in Amazonas, Brazil. *Frontiers in Microbiology*, 13, 835443. <https://doi.org/10.3389/fmicb.2022.835443>
36. Ministerio de Salud Pública. (2001, noviembre 26). *Crecimiento del SIDA en Uruguay*. <https://archivo.presidencia.gub.uy/noticias/archivo/2001/noviembre/2001112612.htm>
37. Granich, R. M., Gilks, C. F., Dye, C., De Cock, K. M., & Williams, B. G. (2009). Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. *The Lancet*, 373(9657), 48–57. [https://doi.org/10.1016/S0140-6736\(08\)61697-9](https://doi.org/10.1016/S0140-6736(08)61697-9)
38. Montaner, J. S., Lima, V. D., Barrios, R., Yip, B., Wood, E., Kerr, T., Shannon, K., Harrigan, P. R., Hogg, R. S., Daly, P., & Kendall, P. (2010). Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. *The Lancet*, 376(9740), 532–539. [https://doi.org/10.1016/S0140-6736\(10\)60936-1](https://doi.org/10.1016/S0140-6736(10)60936-1)
39. Pilcher, C. D., Tien, H. C., Eron, J. J., Vernazza, P. L., Leu, S. Y., Stewart, P. W., Goh, L. E., & Cohen, M. S. (2004). Brief but efficient: Acute HIV infection and the sexual transmission of HIV. *The Journal of Infectious Diseases*, 189(10), 1785–1792. <https://academic.oup.com/jid/article/189/10/1785/2191338>
40. Hollingsworth, T. D., Anderson, R. M., & Fraser, C. (2008). HIV-1 transmission, by stage of infection. *Journal of Infectious Diseases*, 198(5), 687–693. <https://doi.org/10.1086/590501>
41. Arantes, I., Bello, G., Darcissac, E., Lacoste, V., & Nacher, M. (2021). Using phylogenetic surveillance and epidemiological data to understand the HIV-1 transmission dynamics in French Guiana. *AIDS*, 35(6), 979–984. <https://doi.org/10.1097/QAD.0000000000002817>
42. Paraskevis, D., Pybus, O., Magiorkinis, G., Hatzakis, A., Wensing, A. M. J., van de Vijver, D. A., Albert, J., Angarano, G., Åsjö, B., Balotta, C., Boeri, E., Camacho, R., Chaix, M. L., Coughlan, S., Costagliola, D., de Luca, A., de Mendoza, C., Derdelinckx, I., Grossman, Z., & Vandamme, A. M. (2009). Tracing the HIV-1 subtype B mobility in Europe: A phylogeographic approach. *Retrovirology*, 6, 49 <https://doi.org/10.1186/1742-4690-6-49>
43. Flieller, R., & Cabrera, S. (2019). *Virus de inmunodeficiencia humana y resistencia a antirretrovirales*. Cátedra de Enfermedades Infecciosas, Facultad de Medicina, Universidad de la República.

TABLES AND FIGURES

Table 1. Phylogenetic classification of HIV-1 subtype B *pol* sequences from Uruguay across major pandemic and non-pandemic lineages.

Lineage	Sub-lineage	N (%)	Sampling range
B_{PAN}		473 (100%)	2007 – 2021
	B_{PAN-UY}	298 (63.0%)	2007 – 2021
	B_{PAN-SAM}	94 (19.8%)	2007 – 2021
	B_{PAN-AR-II}	21 (4.4%)	2008 – 2021
	B_{PAN-BR-III}	15 (3.2%)	2007 – 2021
	B_{PAN-MX-II}	13 (2.7%)	2008 – 2021
	B_{PAN-BR-II}	10 (2.1%)	2010 – 2019
	B_{PAN-MX-I}	8 (1.7%)	2013 – 2015
	B_{PAN-AR-III}	4 (0.8%)	2012 – 2015
	B_{PAN-CAM}	4 (0.8%)	2008 – 2015
	B_{PAN-BR-IV}	3 (0.6%)	2015 – 2019
	B_{PAN-PE}	3 (0.6%)	2009 – 2018
B_{CAR}		18 (100%)	2007–2021

The table summarizes the distribution of 491 HIV-1 subtype B *pol* sequences from Uruguay (HXB2 positions 2358–3251), sampled between 2007 and 2021, according to their phylogenetic assignment within pandemic (B_{PAN}) and non-pandemic Caribbean (B_{CAR}) lineages. For B_{PAN}, sequences are further stratified into previously described regional or country-specific clades and into Uruguayan-specific diversity (B_{PAN-UY}). Values indicate the number and proportion of sequences per lineage, together with their corresponding sampling time ranges.

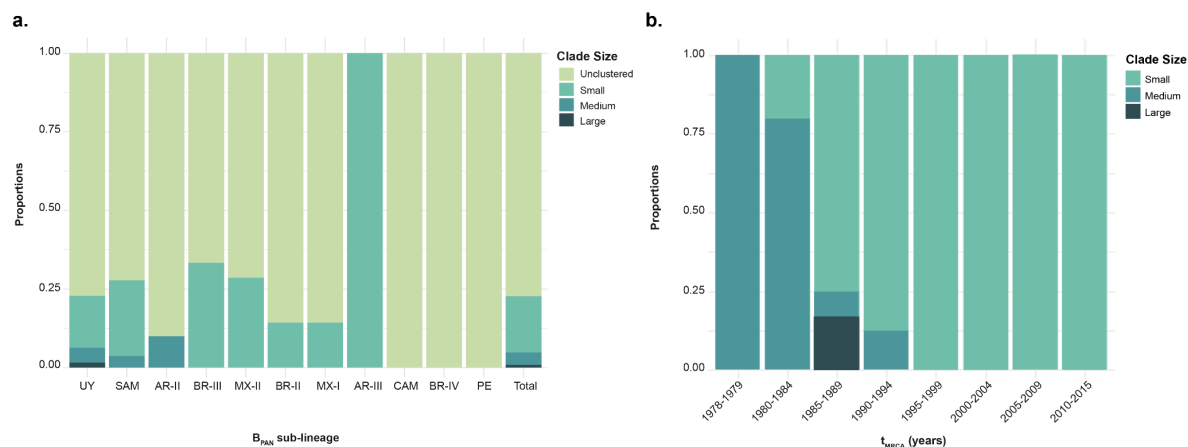
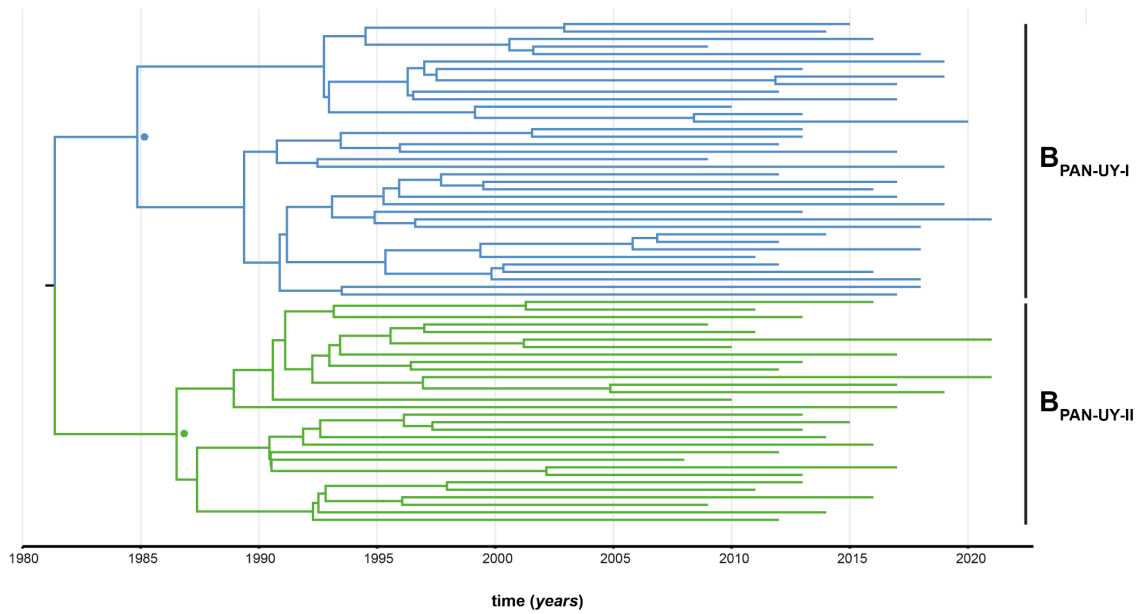


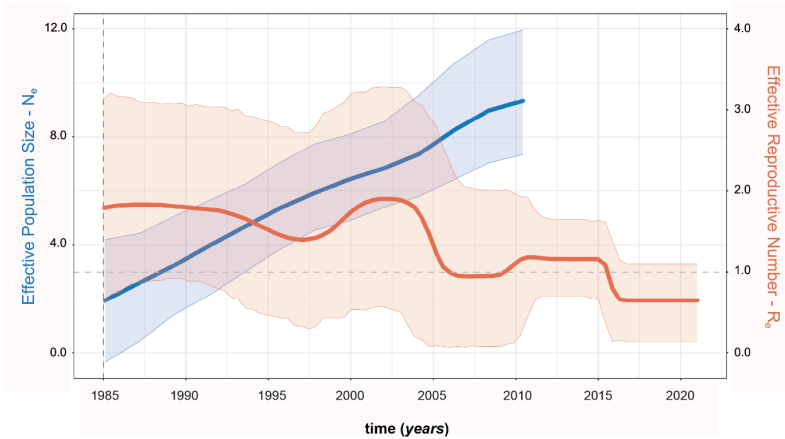
Figure 1. Distribution and temporal structure of BPAN transmission clusters in Uruguay.

(A) Proportional distribution of Uruguayan BPAN sequences across transmission cluster size categories—large ($n \geq 30$), medium ($10 \leq n < 30$), small ($2 \leq n < 10$)—and non-clustered introductions (singletons), shown separately for each BPAN sub-lineage analyzed in independent maximum-likelihood (ML) phylogenies. (B) Temporal distribution of BPAN transmission clusters according to their estimated time to the most recent common ancestor (TMRCA), grouped into discrete time intervals based on molecular clock analyses. Stacked bars represent the relative contribution of small, medium, and large clusters within each TMRCA interval.

a.



b.



c.

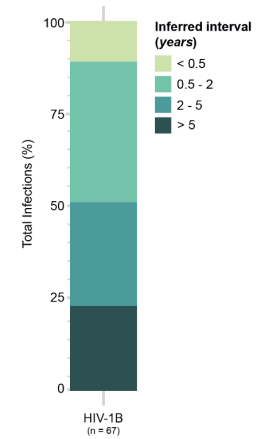


Figure 2. Time-scaled phylogeny and joint phylodynamic reconstruction of the major Uruguayan B_{PAN} sublineages. (A) Bayesian time-scaled phylogeny of the two dominant Uruguayan-specific BPAN sublineages, $B_{PAN-UY-I}$ and $B_{PAN-UY-II}$, inferred under a relaxed uncorrelated lognormal molecular clock in BEAST. Branch lengths are proportional to time (years), as indicated by the horizontal axis, and colors denote the two independently emerging sublineages. The asterisk (*) indicate the origin node of each clade (PP = 1). (B) Joint phylodynamic reconstruction of the combined $B_{PAN-UY-I}$ and $B_{PAN-UY-II}$ sublineages. Temporal changes in the effective number of infections (N_e) were inferred using the Bayesian SkyGrid (BSKG) model (posterior median shown as a solid blue line, with shaded 95% HPD interval; left y-axis), while temporal variation in the effective reproductive number (R_e) was estimated using the Birth–Death Skyline (BDSKY) model (posterior median shown as a solid orange line, with shaded 95% HPD interval; right y-axis). A horizontal dashed line indicates the epidemic threshold ($R_e = 1$). (C) Distribution of inferred secondary transmission events according to time since infection, estimated from internode distances in maximum clade credibility (MCC) trees of large and medium-sized Uruguayan B_{PAN} transmission clusters. Bars represent the proportion of transmission events

occurring within discrete post-infection time intervals (<0.5, 0.5–2, 2–5, and >5 years).

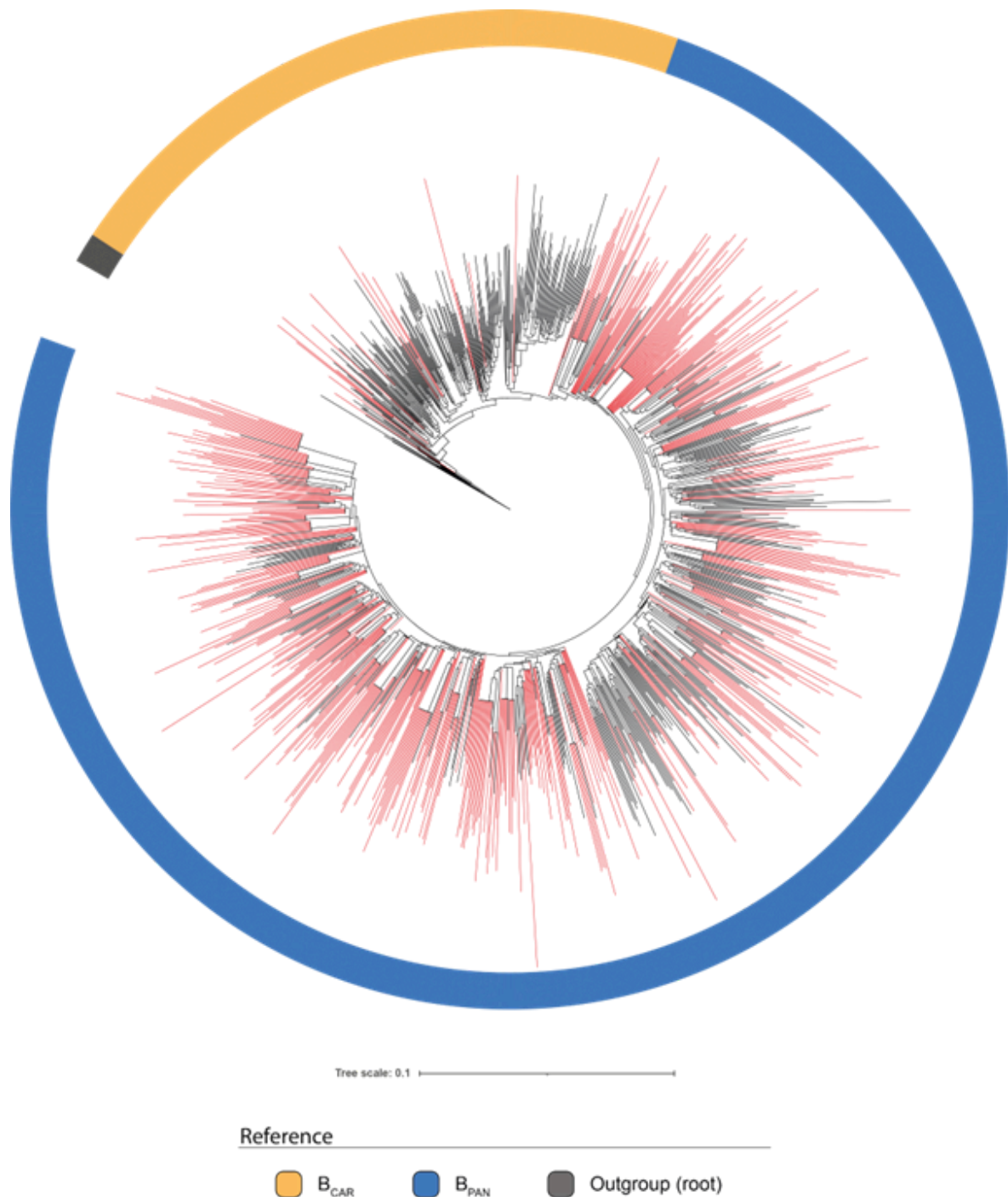
SUPPLEMENTARY TABLES AND FIGURES

Supplementary Table 1.

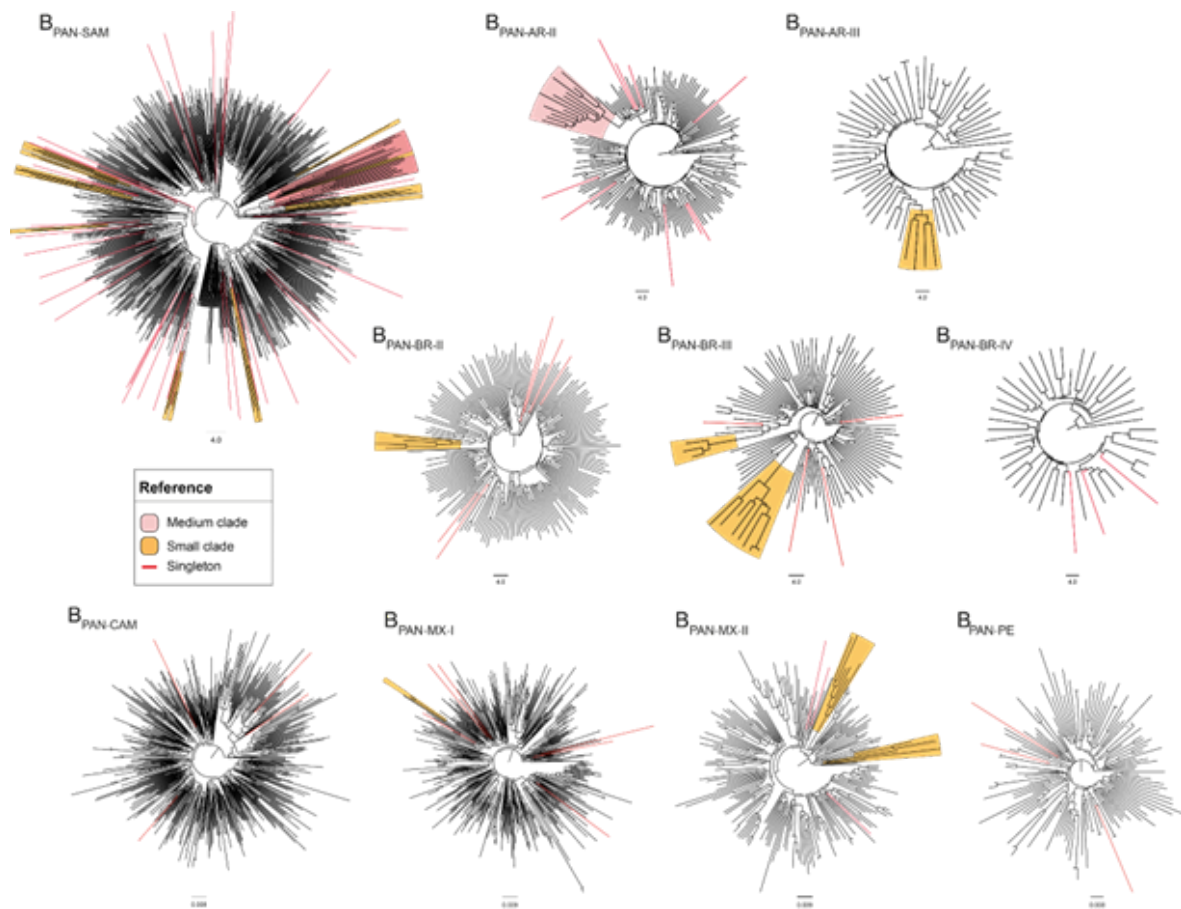
Sublineage	Cluster	<i>N</i>	<i>T</i> _{MRC1}
BPAN-UY			
B_{PAN-UY-I}		37	1985 (1978 – 1989)
B_{PAN-UY-II}		30	1987 (1978 – 1988)
B _{PAN-UY-III}		4	1993 (1986 – 1999)
B _{PAN-UY-IV}		2	2010 (2009 – 2011)
B _{PAN-UY-V}		3	1986 (1978 – 1993)
B _{PAN-UY-VI}		4	1988 (1982 – 1995)
B _{PAN-UY-VII}		4	1987 (1980 – 1994)
B _{PAN-UY-VIII}		10	1982 (1974 – 1988)
B _{PAN-UY-IX}		2	1998 (1991 – 2004)
B _{PAN-UY-X}		2	2002 (1998 – 2006)
B _{PAN-UY-XI}		6	1992 (1985 – 1998)
B _{PAN-UY-XII}		10	1989 (1983 – 1995)
B _{PAN-UY-XIII}		3	1999 (1992 – 2006)
B _{PAN-UY-XIV}		18	1980 (1972 – 1987)
B _{PAN-UY-XV}		2	2011 (2006 – 2016)
B _{PAN-UY-XVI}		12	1978 (1972 – 1984)
B _{PAN-UY-XVII}		5	1998 (1991 – 2006)
B _{PAN-UY-XVIII}		2	2001 (1995 – 2008)
B _{PAN-UY-XIX}		2	2005 (1993 – 2013)
B _{PAN-UY-XX}		3	2005 (1999 – 2010)
B _{PAN-UY-XXI}		2	1994 (1986 – 2002)
B _{PAN-UY-XXII}		2	2009 (2007 – 2011)
B _{PAN-UY-XXIII}		10	1979 (1973 – 1986)
B _{PAN-UY-XXIV}		10	1978 (1972 – 1985)
B _{PAN-UY-XXV}		2	2003 (1996 – 2009)
B _{PAN-UY-XXVI}		2	2002 (1996 – 2009)
B _{PAN-UY-XXVII}		4	1999 (1992 – 2006)
B _{PAN-UY-XXVIII}		3	1988 (1981 – 1993)
B _{PAN-UY-XXIX}		2	1992 (1985 – 2000)

Supplementary Table 2.

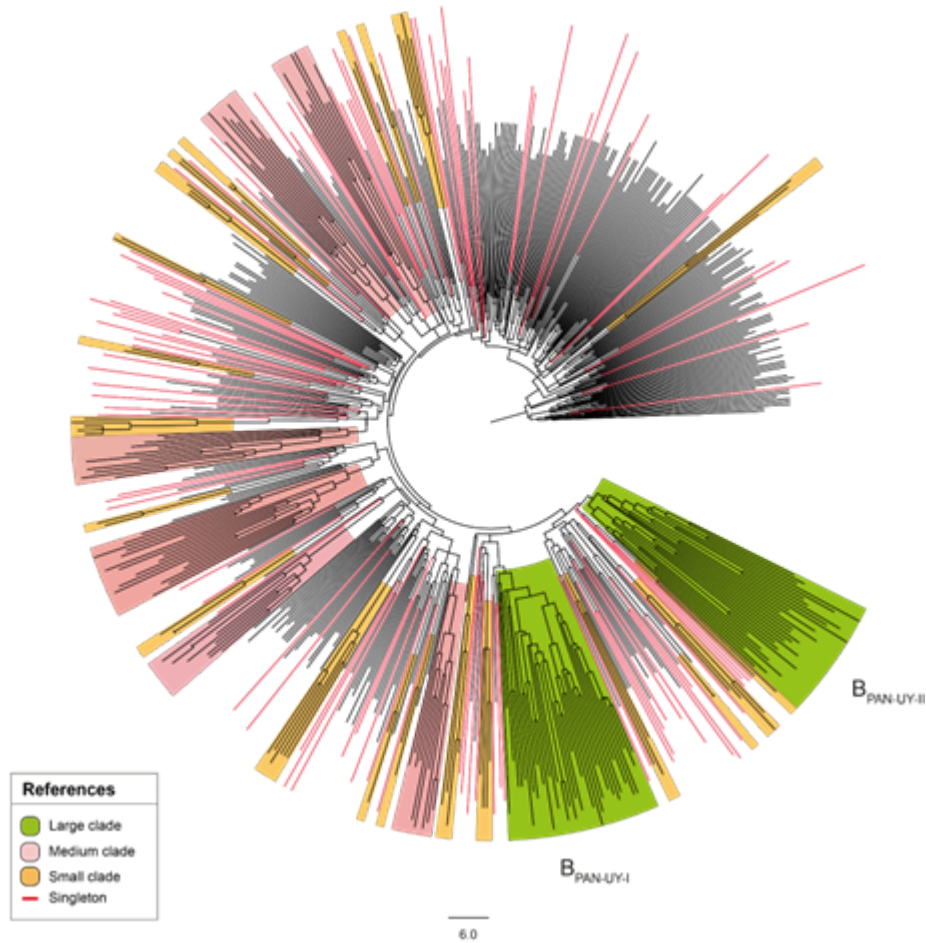
Sublineage	Cluster	N	$T_{MRC\Delta}$	Geographic origin
B_{PAN-SAM}				
	B _{PAN-SAM-I}	4	1987 (1981 – 1994)	Brasil
	B _{PAN-SAM-II}	2	2004 (1997 – 2011)	Brasil
	B _{PAN-SAM-III}	3	1990 (1981 – 1998)	Brasil
	B _{PAN-SAM-IV}	2	2002 (1996 – 2008)	Brasil
	B _{PAN-SAM-V}	2	2005 (1998 – 2012)	Brasil
	B _{PAN-SAM-VI}	2	1996 (1987 – 2003)	Brasil
	B _{PAN-SAM-VII}	10	1982 (1975 – 1988)	Brasil
	B _{PAN-SAM-VIII}	2	2013 (2008 – 2016)	Argentina
	B _{PAN-SAM-IX}	2	2005 (1999 – 2012)	Argentina
	B _{PAN-SAM-X}	2	1984 (1974 – 1993)	Brasil
	B _{PAN-SAM-XI}	2	1997 (1989 – 2005)	Brasil
	B _{PAN-SAM-XII}	13	1980 (1973 – 1987)	Brasil
	B _{PAN-SAM-XIII}	2	1990 (1982 – 1999)	Brasil
	B _{PAN-SAM-XIV}	5	1989 (1982 – 1994)	Brasil
	B _{PAN-SAM-XV}	2	1986 (1978 – 1993)	Brasil
B_{PAN-AR-II}				
	B _{PAN-AR-II}	12	1990 (1984 – 1997)	Argentina
B_{PAN-AR-III}				
	B _{PAN-AR-III}	4	1988 (1982 – 1995)	Argentina
B_{PAN-BR-II}				
	B _{PAN-BR-II}	4	1988 (1981 – 1995)	Brasil
B_{PAN-BR-III}				
	B _{PAN-BR-III-I}	8	1990 (1983 – 1996)	Brasil
	B _{PAN-BR-III-II}	3	2005 (1999 – 2010)	Brasil



Supplementary Figure 1. Maximum-likelihood phylogenetic tree used for clade assignment of Uruguayan HIV-1 subtype B sequences. Maximum-likelihood phylogeny of HIV-1 subtype B pol gene sequences inferred using RAxML. Uruguayan sequences from the GENOUY-HIV database are shown in red, whereas reference sequences corresponding to the pandemic (BPAN, n = 300) and non-pandemic Caribbean (BCAR, n = 200) subtype B lineages are shown in black. The outer ring shows clade assignment, with B_{CAR} in yellow and B_{PAN} in blue; the grey cluster corresponds to subtype D sequences used as an outgroup (n = 10).. Branch lengths are proportional to the number of nucleotide substitutions per site.



Supplementary Figure 2. Maximum-likelihood phylogenetic trees of major B_{PAN} lineages contributing to the Uruguayan HIV-1 subtype B epidemic. Independent maximum-likelihood phylogenetic trees were reconstructed for each B_{PAN} lineage detected in Uruguay using PhyML. Uruguayan transmission clusters are highlighted according to size, with medium-sized clades shown in pink and small clades in yellow, whereas red branches indicate sporadic introductions without evidence of onward local transmission (singletons). All highlighted clusters are supported by aLRT values ≥ 0.85 . Trees were midpoint-rooted, and branch lengths are proportional to the number of nucleotide substitutions per site.



Supplementary Figure 3. Maximum-likelihood phylogenetic tree of the B_{PAN-UY} dataset.

Maximum-likelihood phylogenetic tree inferred using PhyML for B_{PAN-UY} sequences, comprising Uruguayan HIV-1 subtype B sequences not assigned to any previously described Latin American B_{PAN} clade. Uruguayan transmission clusters are highlighted according to size, with large clades shown in green, medium-sized clades in pink, and small clades in yellow, whereas red branches indicate sporadic introductions without evidence of onward local transmission (singletons). All highlighted clusters are supported by aLRT values ≥ 0.85 . The tree was midpoint-rooted, and branch lengths are proportional to the number of nucleotide substitutions per site.

Supplementary Table 3. Genbank Accession Numbers of GENOUY-HIV HIV-1B sequences

ACCESSION NUMBERS (HIV-1B GENOUY-HIV)									
PZ351607	PZ351757	PZ351903	PZ352081	PZ352277	PZ352451	PZ352597	PZ352751		
PZ351611	PZ351758	PZ351911	PZ352086	PZ352281	PZ352453	PZ352598	PZ352754		
PZ351612	PZ351759	PZ351912	PZ352087	PZ352283	PZ352454	PZ352599	PZ352757		
PZ351615	PZ351760	PZ351913	PZ352089	PZ352284	PZ352457	PZ352600	PZ352764		
PZ351617	PZ351761	PZ351917	PZ352092	PZ352285	PZ352458	PZ352602	PZ352768		
PZ351620	PZ351764	PZ351920	PZ352093	PZ352288	PZ352460	PZ352609	PZ352770		
PZ351621	PZ351772	PZ351921	PZ352096	PZ352289	PZ352461	PZ352613	PZ352771		
PZ351622	PZ351779	PZ351926	PZ352101	PZ352292	PZ352463	PZ352614	PZ352773		
PZ351627	PZ351780	PZ351927	PZ352103	PZ352294	PZ352465	PZ352615	PZ352774		
PZ351632	PZ351782	PZ351931	PZ352113	PZ352300	PZ352467	PZ352617	PZ352776		
PZ351635	PZ351783	PZ351934	PZ352115	PZ352306	PZ352468	PZ352618	PZ352781		
PZ351636	PZ351784	PZ351943	PZ352116	PZ352312	PZ352469	PZ352623	PZ352783		
PZ351637	PZ351790	PZ351947	PZ352122	PZ352314	PZ352471	PZ352624	PZ352784		
PZ351640	PZ351794	PZ351948	PZ352124	PZ352317	PZ352474	PZ352628	PZ352788		
PZ351643	PZ351795	PZ351950	PZ352129	PZ352325	PZ352478	PZ352633	PZ352790		
PZ351648	PZ351799	PZ351957	PZ352131	PZ352327	PZ352480	PZ352634	PZ352791		
PZ351649	PZ351800	PZ351958	PZ352132	PZ352331	PZ352481	PZ352636	PZ352795		
PZ351652	PZ351802	PZ351960	PZ352137	PZ352334	PZ352482	PZ352642	PZ352796		
PZ351655	PZ351803	PZ351962	PZ352140	PZ352337	PZ352484	PZ352649	PZ352798		
PZ351656	PZ351804	PZ351967	PZ352149	PZ352341	PZ352485	PZ352654	PZ352804		
PZ351657	PZ351805	PZ351973	PZ352154	PZ352343	PZ352489	PZ352657	PZ352807		
PZ351658	PZ351806	PZ351974	PZ352162	PZ352345	PZ352492	PZ352658	PZ352809		
PZ351659	PZ351809	PZ351977	PZ352164	PZ352348	PZ352494	PZ352659	PZ352810		
PZ351661	PZ351810	PZ351979	PZ352165	PZ352355	PZ352495	PZ352663	PZ352825		
PZ351665	PZ351811	PZ351981	PZ352166	PZ352359	PZ352501	PZ352664	PZ352826		
PZ351666	PZ351814	PZ351982	PZ352169	PZ352360	PZ352503	PZ352668	PZ352827		
PZ351668	PZ351815	PZ351985	PZ352171	PZ352361	PZ352512	PZ352669	PZ352830		
PZ351672	PZ351818	PZ351991	PZ352172	PZ352363	PZ352514	PZ352671	PZ352833		
PZ351673	PZ351821	PZ351992	PZ352175	PZ352370	PZ352515	PZ352672	PZ352834		
PZ351674	PZ351822	PZ351993	PZ352177	PZ352371	PZ352518	PZ352673	PZ352835		
PZ351677	PZ351827	PZ351994	PZ352184	PZ352379	PZ352519	PZ352676	PZ352839		
PZ351678	PZ351828	PZ352001	PZ352185	PZ352380	PZ352522	PZ352677	PZ352840		
PZ351680	PZ351829	PZ352004	PZ352189	PZ352382	PZ352524	PZ352679	PZ352842		
PZ351682	PZ351830	PZ352005	PZ352191	PZ352385	PZ352527	PZ352680	PZ352845		
PZ351683	PZ351832	PZ352007	PZ352193	PZ352386	PZ352528	PZ352682	PZ352848		
PZ351685	PZ351833	PZ352010	PZ352195	PZ352387	PZ352530	PZ352687	PZ352850		
PZ351686	PZ351834	PZ352013	PZ352196	PZ352388	PZ352532	PZ352690	PZ352851		
PZ351687	PZ351836	PZ352015	PZ352198	PZ352393	PZ352535	PZ352694	PZ352853		
PZ351688	PZ351841	PZ352016	PZ352202	PZ352395	PZ352539	PZ352696	PZ352854		
PZ351691	PZ351846	PZ352018	PZ352210	PZ352400	PZ352542	PZ352702	PZ352856		
PZ351700	PZ351853	PZ352019	PZ352215	PZ352401	PZ352545	PZ352706	PZ352857		
PZ351702	PZ351854	PZ352022	PZ352217	PZ352403	PZ352549	PZ352708	PZ352860		
PZ351704	PZ351856	PZ352024	PZ352221	PZ352407	PZ352553	PZ352710	PZ352862		
PZ351706	PZ351857	PZ352025	PZ352223	PZ352408	PZ352554	PZ352712	PZ352866		
PZ351708	PZ351858	PZ352029	PZ352224	PZ352409	PZ352555	PZ352714	PZ352867		
PZ351709	PZ351864	PZ352031	PZ352227	PZ352411	PZ352559	PZ352718	PZ352868		
PZ351713	PZ351866	PZ352032	PZ352230	PZ352413	PZ352561	PZ352719	PZ352869		
PZ351720	PZ351867	PZ352036	PZ352231	PZ352415	PZ352563	PZ352722	PZ352871		
PZ351725	PZ351868	PZ352038	PZ352236	PZ352417	PZ352564	PZ352726	PZ352872		
PZ351728	PZ351870	PZ352046	PZ352240	PZ352419	PZ352565	PZ352729	PZ352873		
PZ351733	PZ351871	PZ352056	PZ352241	PZ352421	PZ352567	PZ352730			
PZ351735	PZ351875	PZ352058	PZ352242	PZ352423	PZ352568	PZ352732			
PZ351741	PZ351876	PZ352062	PZ352247	PZ352425	PZ352570	PZ352734			
PZ351743	PZ351878	PZ352065	PZ352249	PZ352426	PZ352571	PZ352736			
PZ351744	PZ351879	PZ352067	PZ352250	PZ352428	PZ352575	PZ352738			
PZ351746	PZ351884	PZ352070	PZ352251	PZ352429	PZ352578	PZ352739			
PZ351749	PZ351885	PZ352071	PZ352253	PZ352434	PZ352581	PZ352740			
PZ351751	PZ351886	PZ352073	PZ352258	PZ352435	PZ352582	PZ352741			
PZ351752	PZ351889	PZ352075	PZ352263	PZ352445	PZ352585	PZ352743			
PZ351753	PZ351897	PZ352076	PZ352268	PZ352446	PZ352588	PZ352744			
PZ351755	PZ351899	PZ352078	PZ352269	PZ352448	PZ352593	PZ352745			
PZ351756	PZ351900	PZ352079	PZ352270	PZ352449	PZ352594	PZ352746			
	PZ351901	PZ352080	PZ352274	PZ352450	PZ352595	PZ352747			