

TESIS DOCTORAL 2024

DINÁMICA DE POBLACIONES HUMANAS DEL NEA **Aportes desde el estudio de linajes uniparentales**

LIC. DANIELA C. RODRÍGUEZ GOLPE



Obra original de Sebastián Dons. Serie migraciones, relatos de sueños dorados.

DIRECTORES:
DR. CLAUDIO M. BRAVI
DRA. JOSEFINA M. B. MOTTI



Universidad Nacional de La Plata
Facultad de Ciencias Naturales y Museo

*A los pueblos originarios de América
por su lucha incansable.*

I AGRADECIMIENTOS

Este trabajo no hubiera sido posible sin la guía y el acompañamiento de mis directores, de quienes admiro su integridad, tenacidad y pasión por la investigación de los linajes maternos. Agradezco a Jose, por su paciencia, su escucha, su crítica constructiva... A Claudio, por confiar en mí, a pesar de haber confesado en esa entrevista inicial que nunca había hecho una PCR. Y también por confiar después, a la distancia, en las buenas y en las otras. A Graciela Bailliet que me guió en los análisis del cromosoma Y. Al IMBICE que me abrió sus puertas para poder realizar este trabajo y a toda su gente. A mis compañeras de laboratorio: Marisol, Eli, Juli, Laura, Ma. Laura, Belén, Rita, Mariela, Marina y Cami, por siempre estar dispuestas a ayudarme. Al grupo de Becarios 1 con los que nos acompañamos en este desafío de ser becarios doctorales. A mi familia platense: Pau, Santi, Nacho y Simón, por todo el tiempo compartido. A la Universidad Nacional de La Plata, por mantenerse pública, gratuita y brindando educación de calidad. Al CONICET y a la Agencia Nacional de Promoción Científica y Tecnológica, por las becas y los subsidios de investigación otorgados de donde obtuvimos financiamiento. A los donantes que participaron de esta investigación y a las instituciones en donde se realizaron los muestreos. A los jurados: Emma Alfaro, Gonzalo Figueiro y Rodrigo Nores; por su atención en cada detalle, por sus sugerencias y correcciones que enriquecieron la versión final de este trabajo.

A mi familia, por su apoyo incondicional. A mi mamá, por su escucha, su generosidad inagotable. Por siempre apoyarme en mis decisiones. Por estar ahí, en los momentos más difíciles. A mi hermana Estefi y a mi sobri Piero, por esas videollamadas eternas cuando más las necesitaba, por mantenerse cerca aun estando lejos. A Estefi, además, por el diseño de las tapas y por su paciencia para enseñarme a usar Illustrator. Todas las figuras de esta tesis pudieron lograrse gracias a eso. A mi hermano Agus, por organizar todos los festejos antes de que la tesis estuviera escrita, por bancarme y delirar sobre otros caminos posibles, aunque yo siempre elija el mismo. A mi papá y a Vivi por ayudarme a mudarme a La Plata y también por ayudarme a volver. A papá por esa frase en el momento justo: “animate y soltala”. A Nacho (el Tito), por hacerme un lugar en su casa cuando comenzó toda esta historia. A la tía Claudia, por esperarme con la comida en esa última fase de experimentos agotadora, por hablar conmigo si lo necesitaba y por verme comer en silencio si no. Al tío Julio y a los primos, por los asados y las risas. A Darwin por ser mi *co-worker* favorito.

A mis amigas Agos, Car y Miqui por ser de fierro, por estar conmigo en todas. Por conocerme más que yo a mí misma y quererme así. A Flor, por recordarme vivir un día a la vez. A mis biólogos preferidos: Lau, Pato, Eva, Gon, Juan, Gonza, Aldi, Ro y Marcos por apoyarme

siempre. A Martín, por esa fe irracional en mí. Por alentarme a seguir cuando parecía que esta tesis no se terminaba nunca. A Seba, por cederme una de sus obras como portada de este trabajo sin dudarlo.

A mis compañeros de la Escuela Superior de Medicina, a los del comité de ética y a los del INIDEP; pero en especial a mi jefa Inés, por el tiempo para poder terminar de escribir estas líneas.

A Agustina por devolverme al eje cada vez que lo necesité. A cada uno que -sin entender demasiado lo que estaba haciendo- se dio cuenta de que este trabajo era importante para mí.

A ese 2 de Julio y a todo el camino recorrido antes y después.

¡Gracias!

II ÍNDICE

I AGRADECIMIENTOS.....	I
II ÍNDICE.....	III
III ÍNDICE DE TABLAS	VI
IV ÍNDICE DE FIGURAS	VIII
V RESUMEN	X
VI <i>ABSTRACT</i>	XIV
1 CAPÍTULO 1. INTRODUCCIÓN	1
Objetivo general.....	2
Objetivos específicos.....	2
1.1 Reconstrucción de la filogenia humana a partir de marcadores de herencia uniparental .	3
1.1.1 Marcadores de herencia uniparental.....	3
1.1.2 Filogeografía y distribución de las poblaciones americanas	8
1.2 Procesos de mestizaje y configuración de las poblaciones humanas americanas actuales	13
1.3 Definición de la zona de estudio	17
1.3.1 Geográfica	17
1.3.2 Arqueológica	18
1.3.3 Etnográfica y lingüística.....	22
1.3.4 Histórica	25
2 CAPÍTULO 2: MATERIALES Y MÉTODOS	28
2.1 Esquema de trabajo	28
2.2 Muestreos poblacionales	29
2.3 Encuestas genealógicas.....	30
2.4 Consideraciones éticas.....	30
2.4.1 Proceso de consentimiento informado, muestras biológicas y confidencialidad.....	30
2.4.2 Aprobaciones de los comités de ética.....	31
2.4.3 Devolución de resultados a la población participante	31
2.5 Procesamiento de muestras biológicas.....	32
2.5.1 Extracción de ADN.....	32
2.5.2 Tratamiento de muestras coaguladas.....	33
2.6 Análisis de Censos Nacionales.....	33
2.7 Primer nivel de análisis: migraciones intercontinentales	34
2.7.1 Tipificación de haplogrupos uniparentales	34
2.7.2 Linajes maternos	34

2.7.3 Linajes paternos	38
2.7.4 Comparación de datos moleculares y genealógicos	43
2.8 Segundo nivel de análisis: vínculos interregionales	43
2.8.1 Definición de sub-clados mitocondriales	43
2.9 Tercer nivel de análisis: distribución sudamericana de clados mitocondriales	58
2.9.1 Caracterización de alta resolución del haplogrupo B2e.....	59
3 CAPÍTULO 3: RESULTADOS	60
3.1 Primer nivel de análisis: migraciones intercontinentales	61
3.1.1 Linajes maternos	61
3.1.2 Linajes paternos	64
3.1.3 Linajes nativos: línea materna vs línea paterna	67
3.2 Segundo nivel de análisis: vínculos interregionales	68
3.2.1 Análisis de linajes maternos	68
3.3 Tercer nivel de análisis: distribución sudamericana de clados mitocondriales	98
3.3.1 Linaje A2bf.....	98
3.3.2 Linaje B2e	103
3.3.3 Linaje B2an	105
3.3.4 Linaje D1e.....	109
3.3.5 Linaje D1f1.....	113
4 CAPÍTULO 4. DISCUSIÓN.....	117
4.1 Migraciones intercontinentales	117
4.1.1 Linajes genéticos maternos y paternos.....	117
4.1.2 Historia oral.....	118
4.2 Relevancia de la Identificación de polimorfismos del ADNmt en región codificante. Una discusión metodológica.....	120
4.2.1 Asociación entre haplotipos de región control, PNS en regiones codificantes y clados dentro del haplogrupo A2.	120
4.3 Vínculos interregionales y distribución sudamericana de clados mitocondriales	122
4.3.1 Análisis de diversidad mitocondrial	122
4.3.2 Vínculos interregionales.....	124
5 CAPÍTULO 5. CONCLUSIONES	133
6 REFERENCIAS	135
7 ANEXOS	157
Anexo 2.1. Encuesta genealógica	157
Anexo 2.2 Hoja de información.....	158
Anexo 2.3 Formulario de consentimiento informado.....	160
Anexo 2.4 Aprobaciones de los comités de ética.....	161

Anexo 2.5 Informe para el participante	166
Anexo 2.6: Posters para las instituciones participantes.....	169
Anexo 2.7 Soluciones de trabajo.....	172
Anexo 2.8 Procesamiento de muestras coaguladas.....	176
Anexo 2.9. Árboles filogenéticos del ADNmt humano en donde figuran los clados nuevos definidos en este trabajo.	179
Anexo 2.10. Lista de abreviaturas	208
Anexo 3.1. Lista de muestras de la ciudad de Formosa y Corrientes en las que se secuenció la región control extendida de ADNmt. Se incluyen el rango de lectura, sitios polimórficos y asignación al clado más derivado.....	209
Anexo 3.2. Linajes de origen monofilético en 20 poblaciones de Argentina (número de individuos identificados).	220
Anexo 3.3. Asociación entre haplotipos de región control, PNS en regiones codificantes, y clados dentro de A2.	230
Anexo 3.4. Frecuencia (en porcentaje) de subhaplogrupos identificados en la ciudad de Formosa y su ocurrencia en localidades del NEA, Tartagal, NOA y Cuyo, obtenidas de la BBDD-RI.....	235
Anexo 3.5. Frecuencia de subhaplogrupos mayoritarios en regiones de Argentina y de países limítrofes (en porcentaje)	239
Anexo 3.6. Secuencias de región control de ADNmt recopiladas de la bibliografía e identificadas como A2bf por punto de muestreo, correspondientes al tercer nivel de análisis.	240
Anexo 3.7. Mitogenomas recopilados para la construcción de árboles filogenéticos del tercer nivel de análisis	242
Anexo 3.8. Secuencias de región control de ADNmt recopiladas de la bibliografía e identificadas como B2an por punto de muestreo, correspondientes al tercer nivel de análisis.	249
Anexo 3.9. Secuencias de región control de ADNmt recopiladas de la bibliografía e identificadas como D1e por punto de muestreo, correspondientes al tercer nivel de análisis.	251
Anexo 3.10. Secuencias de región control de ADNmt recopiladas de la bibliografía e identificadas como D1f1 por punto de muestreo, correspondientes al tercer nivel de análisis.	252

III ÍNDICE DE TABLAS

Tabla 2.1 Cebadores del Múltiplex I utilizados para la caracterización por APLP de linajes maternos.	36
Tabla 2.2 Cebadores y condiciones empleados para analizar linajes pertenecientes a N(xA,B).38	
Tabla 2.3. Cebadores utilizados para la caracterización por APLP de linajes paternos.	40
Tabla 2.4. Secuencias de cebadores para la asignación al clado I2	42
Tabla 2.5. Secuencias de cebadores para PN1.....	43
Tabla 2.6. Cebadores empleados para la amplificación y secuenciación de la región control... 44	
Tabla 2.7. Clados asignados por PNS de la región control extendida y de la región codificante 45	
Tabla 2.8 Cebadores y condiciones empleados para identificar clados regionales de A2.	46
Tabla 2.9 Cebadores y condiciones empleados para identificar clados regionales de B2.....	47
Tabla 2.10 Cebadores y condiciones empleados para identificar clados regionales de C1d.	47
Tabla 2.11 Cebadores y condiciones para identificar clados regionales de D1.	49
Tabla 2.12 Cebadores con resultados fallidos.	49
Tabla 2.13. Temperatura de hibridación para los ensayos PCR–APLP de ADNmt.....	50
Tabla 2.14. Identificación de PNS informativos en mitogenomas A2.....	51
Tabla 2.15. Secuencias de región control completa extendida y región codificante de ADNmt de poblaciones argentinas que conforman la base de datos del Laboratorio de Genética Molecular Poblacional, IMBICE (BBDD-RI) y que fueron comparadas con las secuencias obtenidas de la localidad de Formosa.	55
Tabla 2.16. Mitogenomas recopilados de la bibliografía que conforman la Base de Datos de Referencia de la Bibliografía (BBDD-RB) y que fueron comparados con las secuencias obtenidas de la localidad de Formosa.	56
Tabla 3.1. Origen continental de los haplogrupos mitocondriales, en porcentaje por localidad.	62
Tabla 3.2. Procedencia de los datos genealógicos para la línea materna, en porcentaje por localidad.	62
Tabla 3.3. Lugar de origen declarado de último ancestro conocido por vía materna, en porcentaje por localidad. NEA= Nordeste Argentino.....	63
Tabla 3.4. Frecuencia de los haplogrupos del cromosoma Y por localidad de muestreo, en porcentaje	65
Tabla 3.5. Procedencia de los datos genealógicos para la línea paterna, en porcentaje por localidad.	65
Tabla 3.6. Lugar de origen declarado de último ancestro conocido por vía paterna, en porcentaje por localidad. NEA= Nordeste Argentino.	66
Tabla 3.7. Diversidad intrapoblacional para 14 localidades del centro-oeste y noroeste argentino	69
Tabla 3.8. Linajes monofileticos identificados por haplogrupo en las 20 localidades de la BBDD-RI.....	70

Tabla 3.9. Distancias de Nei calculadas entre pares de poblaciones a partir de la frecuencia de clados nativos.....	94
Tabla 3.10. Definición de grupos en los AMOVA ensayados.	97
Tabla 3.11. Tabla AMOVA para los ensayos I, II y III.	98
Tabla 3.12 Frecuencia absoluta de linajes asignables a diferentes clados de B2e en poblaciones del NEA y Litoral	105
Tabla 3.13. Frecuencia absoluta de linajes asignables a diferentes clados de B2an en poblaciones del NEA y Litoral.	106

IV ÍNDICE DE FIGURAS

Figura 1.1. Genoma mitocondrial humano.....	5
Figura 1.2 El Cromosoma Y y su región no recombinante.....	7
Figura 1.3 Árbol filogenético del ADNmt.....	11
Figura 1.4. Árbol del cromosoma Y.....	13
Figura 2.1 Patrón de bandas obtenidas para los haplogrupos de ADNmt de la reacción Múltiple- L.....	35
Figura 2.2. Localización de los puntos de muestreo.....	57
Figura 3.1. Frecuencia de individuos que declaran un origen ultramarino (A) y un origen local (B) sus ancestros maternos más remotos y su correspondencia con los resultados moleculares, en porcentaje.....	64
Figura 3.2. Frecuencia relativa de linajes nativos americanos por línea materna (haplogrupos A- D) y por línea paterna (clado Q-M3) en Formosa y Corrientes y en las localidades muestreadas hasta el momento en la colección del IMBICE, en porcentaje.....	67
Figura 3.3. Red de haplotipos A2 ₍₆₄₎ . Red principal.....	73
Figura 3.4. Ampliación de red de haplotipos A2 ₍₆₄₎ . Linaje A2be.....	74
Figura 3.5. Red de haplotipos A2.....	75
Figura 3.6 Red de haplotipos B2. Red principal.....	77
Figura 3.7. Ampliaciones de red de haplotipos B2: B2+14470, B2+16178 y B2e.....	79
Figura 3.8 Red de haplotipos C1b. Red principal.....	81
Figura 3.9. Ampliación de red de haplotipos C1b. Linaje C1b16.....	82
Figura 3.10. Red de haplotipos C1c.....	84
Figura 3.11. Red de haplotipos C1d.....	86
Figura 3.12 Red de haplotipos D1.....	88
Figura 3.13 Histograma de frecuencias de subhaplogrupos encontrados en la localidad de Formosa.....	89
Figura 3.14. Diversidad de clados mitocondriales de la ciudad de Formosa.....	90
Figura 3.15. Distribución de clados mayoritarios de Formosa por localidad A) y por región B), incluyendo países limítrofes.....	92
Figura 3.16. Análisis de escalamiento multidimensional calculado a partir de las frecuencias de los 201 linajes de ADNmt en 20 poblaciones del NEA, NOA y COA.....	95
Figura 3.17 Barreras genéticas identificadas con el programa Barrier 2.2.....	96
Figura 3.18. Mapa de frecuencia de A2bf.....	99
Figura 3.19. Frecuencia de secuencias recopiladas A2bf según su población de pertenencia y detalle de la fracción indígena.....	100
Figura 3.20. Red de haplotipos de A2bf.....	101
Figura 3.21. Árbol filogenético de A2bf.....	102
Figura 3.22. Árbol filogenético de B2e.....	104

Figura 3.23. Mapa de frecuencia de B2an.	107
Figura 3.24. Red de haplotipos B2an.	108
Figura 3.25. Árbol filogenético de B2an.....	109
Figura 3.26. Mapa de frecuencia de D1e.	110
Figura 3.27. Red de haplotipos D1e.	111
Figura 3.28. Árbol filogenético de D1e.	112
Figura 3.29. Mapa de frecuencia de D1f1.....	114
Figura 3.30. Red de haplotipos D1f1.....	115
Figura 3.31. Árbol filogenético de D1f1.	116
Figura 4.1 Rutas de migración de B2an inferidas en base a los datos filogeográficos	126
Figura 4.2 Rutas de migración de B2e inferidas en base a los datos filogeográficos	129

V RESUMEN

Las poblaciones argentinas contemporáneas, al igual que en el resto de América Latina, se configuraron como resultado de un proceso de mestizaje entre los pueblos originarios y aquellos que arribaron desde tiempos coloniales. En Argentina, se registra un pico de flujo migratorio europeo entre 1860 y 1930, con poco aporte de población de origen africano y que se asentaron en zonas específicas del país. En particular, el Nordeste Argentino (NEA) se caracteriza por una gran diversidad étnica, cultural y lingüística en tiempos prehispánicos. En tiempos coloniales y de conformación del Estado Nacional, los movimientos poblacionales estuvieron influenciados por variaciones geopolíticas y por una íntima relación entre esta región con el Paraguay, Brasil y Uruguay.

Este trabajo de tesis se focaliza en las historias de poblamiento de las ciudades de Corrientes y Formosa y de sus áreas de influencia, en el marco de la región del NEA, a partir del estudio genético de linajes haploides de transmisión exclusivamente materna o paterna que permiten identificar un origen continental remoto. El objetivo es contribuir al conocimiento de la estructuración étnico/geográfica de la diversidad genética humana en el NEA mediante el análisis de ADN mitocondrial (ADNmt) y del cromosoma Y. Los datos moleculares se analizan en el marco de las evidencias bioarqueológicas, lingüísticas, históricas y demográficas, de manera de aportar a la discusión de patrones de distribución de las poblaciones que ya se encuentran estudiados o de plantear otros nuevos.

Se realizaron campañas de investigación en hospitales y centros de hemoterapia y se colectaron 425 muestras biológicas en las ciudades de Corrientes (CRR: N=151) y Formosa (FOR: N=274), las cuales, se analizaron para marcadores de linajes haploides maternos y paternos, previa extracción de ADN. Además, se realizaron encuestas genealógicas en donde se interrogó a los donantes acerca de sus lugares de nacimiento y de sus ancestros.

A fin de abordar las múltiples dimensiones que conforman la historia poblacional del NEA se desarrolló un diseño metodológico que estructura el análisis de la diversidad genética en tres ejes con distintos niveles de resolución. En el primer nivel, se propuso identificar el origen continental de los linajes maternos y paternos, y sus frecuencias relativas, en las localidades de Formosa y Corrientes, a partir de la tipificación de mutaciones diagnósticas del ADNmt y del Cromosoma Y. Los resultados obtenidos arrojan frecuencias mayoritarias de linajes originarios de América por vía materna (FOR=85,04%, CRR=89,40%). Los linajes del oeste de Eurasia (FOR=12,41%, CRR=9,93%) y de África Subsahariana (FOR=2,55%, CRR=0,66%) también están presentes. Por otra parte, más del 80% de los linajes paternos resultaron de origen euroasiático,

de los cuales el 55% corresponde al haplogrupo R1 (FOR=55,2%, CRR=54,7%). La fracción nativa de los linajes paternos resultó ser una de las frecuencias más bajas documentadas para Argentina (FOR= 4,9%; CRR=3,2%). Además, la comparación de los datos moleculares y los genealógicos permitió reconocer que los donantes declaran mayoritariamente un origen regional de sus ancestros paternos y maternos más remotos de los cuales se conserva memoria: para la línea materna, se corroboró una coincidencia entre el origen regional declarado y el resultado molecular; sin embargo, para la línea paterna hubo discrepancias entre el origen declarado (regional) y el resultado molecular (euroasiático) en la mayoría de los casos. Estos resultados podrían ser el reflejo de las sociedades criollas en tiempos coloniales y de la asimetría del mestizaje en la población criolla.

En un segundo nivel de análisis, se hizo foco en la fracción nativa en base a secuencias de la región control del ADNmt (RC) para realizar comparaciones a escala regional. En este sentido, se seleccionaron 169 muestras, teniendo en cuenta un origen local declarado para la línea materna y su membresía a un haplogrupo nativo y se secuenció la RC completa extendida en sentido 5'. El análisis preliminar de las secuencias permitió asignar una fracción a clados reconocibles por polimorfismos de nucleótido simple (PNS) de la RC extendida como A2ah, B2h, B2i1, B2b3a, C1b, C1b2, C1b3, C1d*, C1d1b, C1c*, D1f1, y D1j. Una segunda fracción de linajes fue portadora de haplotipos descritos en alta frecuencia para poblaciones indígenas chaqueñas o amazónicas, o bien en poblaciones cosmopolitas de la región, para algunos de los cuales se han publicado mitogenomas, pero no existe aún definición formal de clado. Una última fracción no es asignable a clados conocidos por carecer de PNS diagnóstico en RC. Esta última, así como los linajes C1b2 y C1d1b, fueron analizados con una batería de ensayos alelo-específicos de PNS de región codificante, algunos de ellos ya probados en el IMBICE y otros diseñados específicamente para este trabajo de tesis. Así se asignó o confirmó membresía en clados de distribución étnico/geográfica restringida y cuya presencia en la región es conocida o esperada, como A2ab, B2b, B2h, B2i, B2₊₆₀₅₆, C1b2, C1d1b, C1d1b1, D1a y D1e. Se construyó una colección de secuencias inéditas del IMBICE incluyendo 20 poblaciones de Argentina, a partir de la cual se realizaron análisis filogenéticos y filogeográficos. En particular, se construyeron redes medianas para los haplogrupos A2, B2, C1b, C1c, C1d y D1; se analizaron los agrupamientos de las diferentes poblaciones en estudio mediante el escalamiento multidimensional de distancias genéticas (Distancia de Nei y FST) y se identificaron barreras genéticas entre localidades. Los resultados demuestran que la muestra de Formosa está caracterizada por una gran diversidad de clados mitocondriales (N = 61). Los clados obtenidos en Formosa representan cerca del 80% de la diversidad encontrada para la región del NEA, 48% de la muestra del NOA y el 38% de la muestra del COA, según comparaciones con la base de datos construida, a partir de la cual se

lograron identificar un total de 201 linajes con presunto origen monofilético. Además, de los 61 clados encontrados, 49 de éstos están representados en la muestra de Formosa en frecuencias menores al 2%, agrupando el 44% de la diversidad total. Mientras tanto, el 48% de la diversidad restante está constituida por 12 subhaplogrupos mayoritarios A2zj: A2₍₆₄₎+(338-3438-16468), A2zk3: A2+(16266-10237-12858-14364), A2be, B2an, B2+16178, B2e, C1b16, C1b18, C1d1b, C1d1d, D1e y D1f1. Por otra parte, los análisis en relación a las distancias genéticas (MDS y Barrier) permitieron agrupar a la localidad de Formosa con el resto de las localidades del NEA y Litoral: Reconquista, Curuzú Cuatiá, La Paz y Gualeguaychú y separarlas del resto de las poblaciones del NOA y COA.

Finalmente, en el tercer y último nivel, se seleccionaron de las secuencias analizadas ciertos clados de interés (B2e, B2an, A2bf, D1e y D1f1) en base a información de la región codificante del ADNmt, conjuntamente con información obtenida de mitogenomas de la bibliografía. Se construyeron manualmente árboles filogenéticos por máxima parsimonia y se identificaron clados monofiléticos de diferente jerarquía, lo cual permitió explorar su distribución y estructuración a escala sudamericana o continental.

Los haplotipos compartidos entre las localidades del NEA y Litoral demuestran una gran afinidad entre estas poblaciones. Ciertos clados identificados presentan alta frecuencia y están representados por un gran número de haplotipos, por los que se los consideró de posible origen local y propios de la región litoraleña y mesopotámica como A2zh (A2+210), A2zd (A2+15262), B2zd (B2+16178), C1b18, C1b19b, C1b30a1 (C1b30+16126), C1b101 (C1b+16384), C1b102 (C1b+152) y C1b2a1a (C1b+263@). Por otra parte, ciertos subhaplogrupos permitieron establecer vínculos con otras regiones del Cono Sur. En particular, algunos de ellos permiten vincular a las poblaciones del NEA y Litoral con la amazonia brasilera (B2an, D1e, B2i1), Paraguay y Uruguay (A2be y B2zd). Además, otros linajes permitieron establecer afinidades con Venezuela y el Caribe (D1f1, A2ah, B2b3a, C1b2). Se pudieron establecer también vínculos con el Centro de Argentina (C1d1b, C1b16 y D1j), con los Andes Centrales (B2aj) y se destaca que no se encontraron vínculos con Patagonia. Es interesante resaltar que, pese a que Formosa conforma una provincia chaqueña, se registró una baja frecuencia de clados compartidos con pueblos originarios del Gran Chaco (sólo una secuencia fue asignada a A2bf).

Por último, los linajes maternos que predominan en Formosa y en el resto de las poblaciones del NEA y Litoral, son propios de esta región, demostrando que comparten una historia biológica común, y que, a su vez, se diferencian de los clados dominantes de otras regiones como el centro, centro-oeste, noroeste y sur de Argentina. Una fracción importante de los linajes que caracterizan a las poblaciones del NEA y el Litoral permiten establecer vínculos con poblaciones de Brasil, Uruguay y Paraguay, más que con poblaciones del resto de Argentina.

Las afinidades con Amazonia y en particular, con el mundo guaraní dominaron el paisaje formoseño, en comparación con los vínculos con el mundo chaqueño. Pese a que los análisis realizados se basan mayormente en la variabilidad encontrada en la RC de poblaciones argentinas, el abordaje metodológico implementado en este trabajo permitió aumentar la resolución de las secuencias obtenidas, a partir de secuenciación parcial de la Región Codificante. De esta manera, se lograron identificar ciertos grupos monofiléticos de linajes que no podrían haber sido reconocidos únicamente por variación en la RC, posibilitando, de esta forma, el subhaplogrupado de alta resolución. Además, la recopilación de secuencias en la bibliografía y en base de datos genéticas de los clados B2e, B2an, A2bf, D1e y D1f1, permitió reconstruir su distribución a nivel sudamericano y en base a esto proponer un origen y una posible ruta de expansión para algunos de ellos. El hecho de poder profundizar en la filogenia de los haplogrupos basales y de los sub-haplogrupos con distribución local diferencial, permitió revisar los modelos de poblamiento planteados hasta el momento y realizar un aporte para la generación de modelos de poblamiento regional.

VI ABSTRACT

Contemporary Argentine populations, as in the rest of Latin America, are the result of a process of miscegenation between natives and those who arrived from colonial times. In Argentina, there was a peak in the European migratory flow between 1860 and 1930, which settled in specific areas of the country, and with a scarce contribution from population of African origin. Northeast Argentina (NEA) is characterized by the presence of ethnic, cultural and linguistic diversity in pre-Hispanic times. In colonial times and during National State formation, population movements were influenced by geopolitical variations, and a close relationship between this region and Paraguay, Brazil and Uruguay.

This thesis focuses on the settlement histories of the cities of Corrientes and Formosa and their areas of influence, in the context of the NEA region, based on the genetic study of haploid lineages of exclusively maternal or paternal transmission that allow the identification of a remote continental origin. The aim is to contribute to the knowledge of the ethnic/geographic structuring of human genetic diversity in the NEA through the analysis of mitochondrial DNA (mtDNA) and the Y chromosome. The molecular data were analysed within the framework of bioarchaeological, linguistic, historical and demographic evidence, to contribute to the discussion of population distribution patterns that have already been studied and/or to propose new ones.

Research campaigns were carried out in hospitals and hemotherapy centers and 425 biological samples were collected in the cities of Corrientes (CRR: N=151) and Formosa (FOR: N=274), which were analysed for maternal and paternal haploid lineage markers, after DNA extraction. In addition, genealogical surveys were conducted in which donors were questioned about their places of birth and their ancestors.

To address the multiple dimensions that comprise the population history of the NEA, a methodological design was developed that structures the analysis of genetic diversity in three axes with different levels of resolution. At the first level, we proposed to identify the continental origin of the maternal and paternal lineages, and their relative frequencies, in the localities of Formosa and Corrientes, based on the typing of diagnostic mutations of the mtDNA and the Y chromosome. The results obtained show majority frequencies of Native American lineages through the maternal lineage (FOR=85.04%, CRR=89.40%). Western Eurasian (FOR=12.41%, CRR=9.93 %) and sub-Saharan African (FOR=2.55%, CRR=0.66%) lineages are also present. On the other hand, more than 80% of the paternal lineages were of Eurasian origin, of which 55% corresponded to haplogroup R1 (FOR=55.2%, CRR=54.7%). The Native American fraction of the paternal lineages was one of the lowest frequencies documented for Argentina (FOR= 4.9%;

CRR=3.2%). In addition, the comparison of molecular and genealogical data allowed us to recognize that donors mostly declare a local origin of their most remote paternal and maternal ancestors for whom there is memory. For the maternal line, a match between the declared local origin and the molecular result was corroborated; however, for the paternal line, there were discrepancies between the declared origin (local) and the molecular result (Eurasian) in most cases. These results could reflect “criollas” societies in colonial times and the asymmetry of miscegenation in the “criolla” population.

A second level of analysis focused on the Native American fraction based on mtDNA control region (CR) sequences for regional comparisons. In this regard, 169 samples were selected, considering a declared local origin for the maternal line and its membership to a native haplogroup, and the full 5' extended CR was sequenced. Preliminary analysis of the sequences made it possible to assign a fraction to clades recognizable by single nucleotide polymorphisms (SNP) of the extended CR as A2ah, B2h, B2i1, B2b3a, C1b, C1b2, C1b3, C1d*, C1d1b, C1c*, D1f1, y D1j. A second fraction of lineages carried haplotypes described in high frequency for indigenous populations of Chaco or the Amazon, or in cosmopolitan populations of the region, for some of which mitogenomes have been published, but there is still no formal clade definition. A final fraction is not assignable to known clades because it lacks diagnostic SNP in CR. This last, as well as the C1b2 and C1d1b lineages, were analysed with a battery of allele-specific assays for coding region SNP, some of them already tested at IMBICE and others designed specifically for this thesis work. Thus, membership in clades of restricted ethnic/geographic distribution and whose presence in the region is known or expected were assigned or confirmed as A2ab, B2b, B2h, B2i, B2₊₆₀₅₆, C1b2, C1d1b, C1d1b1, D1a and D1e. A collection of unpublished sequences of IMBICE including 20 populations from Argentina was built, from which phylogenetic and phylogeographic analyses were performed. Median networks were constructed for haplogroups A2, B2, C1b, C1c, C1d and D. The clustering of the different populations under study was analysed by multidimensional scaling of genetic distances (Nei's distance and FST) and genetic barriers between localities were identified. The results show that the Formosa sample is characterized by a high diversity of mitochondrial clades (N=61). The clades obtained in Formosa represent about 80% of the diversity found for the NEA region, 48% of the NOA sample and 38% of the COA sample, according to comparisons with the constructed database, from which a total of 201 lineages with presumed monophyletic origin were identified. In addition, of the 61 clades found, 49 of these are represented in the Formosa sample in frequencies of less than 2%, grouping 44% of the total diversity. Meanwhile, 48% of the remaining diversity is constituted by 12 major subhaplogroups (A2zj: A2₍₆₄₎+338-3438-16468, A2zk3: A2+16266-10237-12858-1436, A2be, B2an, B2+16178, B2e, C1b16, C1b18, C1d1b,

C1d1d, D1e and D1f1). On the other hand, the analysis of genetic distances (MDS and Barrier) allowed grouping the Formosa locality with the rest of the NEA and Litoral localities: Reconquista, Curuzú Cuatiá, La Paz and Gualeguaychú and separating them from the rest of the NOA and COA populations.

Finally, in the third and last level, certain clades of interest (B2e, B2an, A2bf, D1e and D1f1) were selected from the analysed sequences based on mtDNA coding region information, together with information obtained from mitogenomes in the literature. Phylogenetic trees were manually constructed by maximum parsimony and monophyletic clades of different hierarchy were identified, which allowed us to explore their distribution and structuring on a South American or continental scale.

The haplotypes shared between the NEA and Litoral localities demonstrate a great affinity between these populations. Certain clades identified have a high frequency and are represented by many haplotypes, so they were considered of possible local origin and typical of the littoral and mesopotamian region as A2zh (A2+210), A2zd (A2+15262), B2zd (B2+16178), C1b18, C1b19b, C1b30a1 (C1b30+16126), C1b101 (C1b+16384), C1b102 (C1b+152) and C1b2a1a (C1b+263@). On the other hand, certain subhaplogroups have made it possible to establish links with other regions of the Southern Cone. Some of them allow linking the NEA and Litoral populations with the Brazilian Amazon (B2an, D1e, B2i1), Paraguay and Uruguay (A2be and B2zd). In addition, other lineages allowed establishing affinities with Venezuela and the Caribbean (D1f1, A2ah, B2b3a, C1b2). Links could also be established with Central Argentina (C1d1b, C1b16 and D1j), with the Central Andes (B2aj) and it is noteworthy that no links were found with Patagonia. It is interesting to note that, although Formosa forms a Chaco province, a low frequency of shared clades with native peoples of the Gran Chaco was recorded (only one sequence was assigned to A2bf).

Finally, the maternal lineages that predominate in Formosa and in the rest of the populations of the NEA and Litoral, are typical of this region, showing that they share a common biological history, and that, at the same time, they differ from the dominant clades of other regions such as central, central-western, northwestern and southern Argentina. An important fraction of the lineages that characterize the NEA and Litoral populations allow us to establish links with populations from Brazil, Uruguay and Paraguay, rather than with populations from the rest of Argentina. The affinities with Amazonia and with the Guaraní world dominated the landscape of Formosa, in comparison with the links with the Chaco world. Although the analyses performed are based mainly on the variability found in the CR of Argentine populations, the methodological approach implemented in this work allowed increasing the resolution of the sequences obtained, based on partial sequencing of the coding region. In this way, it was

possible to identify certain monophyletic groups of lineages that could not have been recognized by CR variation alone, thus enabling high-resolution subhaplogrouping. In addition, the compilation of sequences in the literature and in genetic databases of the clades B2e, B2an, A2bf, D1e and D1f1, allowed us to reconstruct their distribution at the South American level and, based on this, to propose an origin and a possible expansion route for some of them. The fact of being able to deepen in the phylogeny of the basal haplogroups and of the sub-haplogroups with differential local distribution, allowed us to review the models of settlement proposed so far and to contribute to the generation of regional settlement models.

1 CAPÍTULO 1: INTRODUCCIÓN

La conformación actual de las poblaciones humanas, tanto en la Argentina como en el resto de América Latina, ha sido el resultado de procesos que operaron desde tiempos prehispánicos y coloniales, delineados por un poblamiento inicial de pueblos originarios y seguido por un proceso de mestizaje entre nativos y migrantes. Este último comenzó con el arribo de los europeos a América y continuó, posteriormente, con distintos pulsos migratorios de población extranjera.

Este trabajo de tesis se focaliza en las historias de poblamiento de las ciudades de Corrientes y Formosa y de sus áreas de influencia, en el marco de la región del Nordeste Argentino (NEA). Se abordan desde el estudio de la configuración genética de sus poblaciones humanas actuales, a fin de determinar el origen continental remoto tanto de los linajes maternos (ADN mitocondrial, ADNmt) como paternos (cromosoma Y). Se parte de la hipótesis general de que en las poblaciones cosmopolitas actuales es posible rastrear genes y linajes de poblaciones ya desaparecidas como aislados socioculturales, a pesar de los procesos de mestizaje que han ocurrido durante los últimos 500 años. Esto ha sido corroborado para otras regiones de Argentina (Motti, 2012; García et al., 2018; Schwab, 2018; entre otros). En este caso, se pondrá a prueba la hipótesis para la región del NEA. A su vez, teniendo en cuenta la estructura geográfica observable para el ADNmt, las hipótesis particulares planteadas son que existan similitudes en la distribución de linajes y subhaplogrupos dentro de la región del NEA y Litoral con los países colindantes en íntima relación con la región (Paraguay, sur de Brasil y Uruguay), y diferencias con otras regiones de la Argentina.

Objetivo general

Contribuir al conocimiento de la estructuración étnico/geográfica de la diversidad genética humana en el Nordeste de la República Argentina mediante el análisis de ADN mitocondrial y del cromosoma Y.

Objetivos específicos

1. Determinar los aportes relativos continente-específicos de origen euroasiático occidental (de Europa y Medio Oriente), africano subsahariano y nativo al acervo génico de linajes maternos y paternos presentes en poblaciones humanas de Formosa y Corrientes mediante tipificación de mutaciones haplogrupo-específicas.
2. Estimar el grado de concordancia entre los linajes genéticos identificados y el origen declarado por los individuos de estas localidades.
3. Obtener secuencias de alta calidad de la región control completa extendida hacia 5' incluyendo el ARNtThr (1300 pb) en linajes maternos de origen amerindio y africano provenientes de las poblaciones del NEA.
4. Analizar polimorfismos de nucleótido simple (PNS) identificadores de sub-haplogrupos/clados y definir diferentes jerarquías en base a la presencia de estas mutaciones diagnóstico.
5. Identificar los procesos de poblamiento que afectaron a la región del NEA; determinar las afinidades biológicas de los grupos humanos de la región, respecto de otras regiones e indagar en qué medida, los procesos post-hispánicos afectaron la conformación genética de las poblaciones.

1.1 Reconstrucción de la filogenia humana a partir de marcadores de herencia uniparental

En este trabajo el objeto de estudio serán las poblaciones humanas y se abordará desde el campo de la antropología biológica y la genética de poblaciones. Una población está conformada por un grupo de individuos que comparten un determinado tiempo y espacio y que pueden presentar diferencias en las frecuencias génicas respecto a otros grupos humanos.

A su vez, gran parte de esta tesis se encuentra atravesada por la filogeografía, campo de estudio que se enfoca en los principios y procesos en torno a la distribución geográfica de los linajes genealógicos y que se nutre de diversas disciplinas como son la genética molecular, la genética de poblaciones, la etología, la demografía, la sistemática filogenética, entre otras (Avice, 2000). La filogeografía hace uso de marcadores genéticos o secuencias de ADN con alta tasa mutacional, que no estén sujetas a recombinación (por ejemplo, el ADNmt) para poder realizar genealogías génicas; por otro lado, se basa en la teoría de la coalescencia (Kingman, 1982 a, b) y en una evolución neutral (Kimura, 1983) como marco teórico y en un análisis cladístico como método de análisis de datos (Lanteri & Confalonieri, 2003).

Para las poblaciones humanas, la estimación de edades de coalescencia y la asignación de una temporalidad a cada evento poblacional ha sido aplicado por Gomez-Carballa et al. (2018) al inferir los límites superior e inferior al poblamiento de Sudamérica en base a las edades de los clados mitocondriales con distribución panamericana y sudamericana, respectivamente.

1.1.1 Marcadores de herencia uniparental

Tanto el ADN mitocondrial como la región no recombinante del cromosoma Y presentan características que los hacen buenos marcadores para estudios poblacionales, en su capacidad para evidenciar la historia evolutiva humana y su correlato en la distribución geográfica. Gran parte de esta tesis se basa en el estudio de estos marcadores de herencia uniparental, por lo que las siguientes secciones se dedicarán a su descripción y caracterización.

1.1.1.1 El ADNmt humano

El ADNmt se encuentra dentro de las mitocondrias, orgánulos subcelulares que están involucrados en la homeostasis y el metabolismo celular, y representa una pequeña porción del genoma humano, en comparación con los tres mil millones de pares de bases (pb) del genoma nuclear. Fue secuenciado por primera vez por Anderson et al. (1981), quienes determinaron que

está compuesto por 16.569 pb formando una doble cadena circular. Estos autores, demostraron que existe cierta asimetría entre las dos cadenas debido a que una de ellas es rica en residuos de guanina (cadena pesada o *heavy strand*) mientras que la otra, es rica en residuos de citosina (cadena liviana o *light strand*). Esta última, fue utilizada como referencia de la secuencia de ADNmt humano, denominada secuencia de referencia de Cambridge (CRS por su acrónimo en inglés). Posteriormente, la secuencia fue corregida y actualmente se utiliza esta última versión: secuencia de referencia de Cambridge revisada o rCRS (Andrews et al., 1999). El ADNmt codifica para 37 genes, 13 polipéptidos –componentes esenciales de la cadena respiratoria-, 2 ARN ribosomales y 22 ARN de transferencia (Figura 1.1). Por otra parte, presenta una región no codificante de 1100 pb denominada región control, desde la posición 16024 a 576, que regula la replicación y la transcripción de la molécula (Anderson et. al., 1981). Esta región presenta a su vez, tres regiones de menor longitud que se caracterizan por su secuencia altamente variable en comparación con la del resto del genoma, denominadas regiones hipervariables I (RHV-I, 16024–16400), II (RHV-II, 44–340) y III (RHV-III, 438–576), respectivamente (Brandstätter et al., 2004).

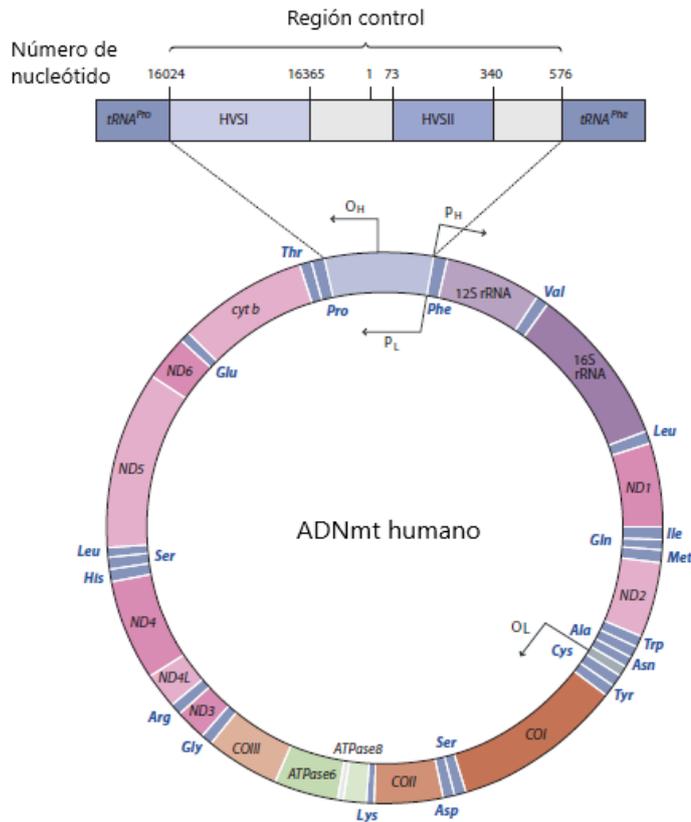


Figura 1.1. Genoma mitocondrial humano. ND1-ND6 y ND4L son genes del complejo I (NADH-ubiquinona oxidorreductasa). Cyt b es un gen de la subunidad del complejo III (ubiquinol-citocromo c oxidasa reductasa). COI – COIII son complejos IV (citocromo c oxidasa, o COX), y ATPase6 y ATPase8 son genes complejos de la subunidad V (ATP sintasa). Se intercalan dos genes de ARN ribosomal (12S RNA y 16S RNA) y 22 genes de ARN de transferencia entre los genes que codifican proteínas, designados por el nombre de los aminoácidos). Se amplía la región no codificante de 1,1 kb, que participa en la regulación de la transcripción y replicación de la molécula. OH y OL son los orígenes de la replicación del ADNmt de cadena pesada y cadena liviana, respectivamente. Figura tomada de Jobling et al. (2014), página 38.

Como se mencionó anteriormente, el ADNmt es uno de los marcadores más utilizados para los estudios filogenéticos, debido a que presenta ciertas propiedades únicas (Pakendorf & Stoneking, 2005):

- 1) *Alto número de copias por célula:* Esta característica hace posible analizar muestras contemporáneas degradadas para estudios forenses o poblacionales y la posibilidad de recuperar secuencias de ADNmt de restos antiguos.

- 2) *Herencia por vía materna*: Permite reflejar la ancestría materna de una población, a partir de rastrear estos linajes a través del tiempo.
- 3) *No recombinación*: Implica que el ADNmt se transmite en bloque de una generación a la siguiente.
- 4) *Alta tasa mutacional*: Posibilita el estudio de la variación intraespecífica en las poblaciones.

Particularmente, la tasa de sustitución en el ADNmt se ha calculado que excede de cinco a diez veces las del ADN nuclear (Brown et al., 1980). Incluso es mayor en las RHV de la región control debido a la existencia de gran cantidad de *hot spots* (puntos calientes), que son posiciones inestables a lo largo de la secuencia (éstos también pueden encontrarse en la región codificante). Particularmente se atribuye esta alta tasa mutacional a:

- ✓ la existencia de una alta concentración de radicales libres en las mitocondrias, producto de la energía generada en la fosforilación oxidativa;
- ✓ la ausencia de histonas como las que empaquetan el ADN nuclear, dejando al ADNmt expuesto a mutaciones;
- ✓ el ADNmt pasa por mayor cantidad de ciclos de replicación por unidad de tiempo ya que tiene una mayor tasa de recambio que el ADN nuclear (Jobling et al., 2014).

Cuando ocurre una mutación, ésta se da en una única molécula de ADNmt que se encuentra en una mitocondria específica. Sin embargo, se observa que la misma, puede presentarse en todas las copias de ADNmt de un individuo en la siguiente generación (mutación homoplásmica), mientras que en otros casos pueden presentarse en sólo algunas copias del genoma (mutación heteroplásmica). Este proceso ocurre debido a que en las poblaciones celulares mitocondriales puede coexistir más de un tipo de ADNmt en la célula (heteroplasmia) y éstas pasan por un cuello de botella intergeneracional (Jobling et al., 2014). Según esta hipótesis se produce una reducción en el número efectivo del número de genomas mitocondriales durante la embriogénesis temprana en la línea germinal femenina, limitando la variabilidad producida por mutación en el oocito primario; esta reducción podría deberse a una compartimentalización de los genomas en unidades segregantes homoplásmicas o por una amplificación preferencial de genotipos específicos (Chinnery, 2006).

1.1.1.2 El cromosoma Y y su región no recombinante

El cromosoma Y es el único de los cromosomas sexuales humanos que se hereda de padres a hijos varones. Conjuntamente con el cromosoma X, habría coevolucionado de un par de autosomas de un mamífero ancestral, hace 300 millones de años (Lahn & Page, 1999).

Con una longitud de aproximadamente 60 megabases (Mb), se conforma de una región de heterocromatina ubicada en la parte distal del brazo largo (Yq), con secuencias de ADN altamente repetitivas y otra de eucromatina con unos pocos genes, situada en el brazo corto (Yp) y en parte del brazo largo (Figura 1.2). A su vez, se pueden diferenciar tres regiones: dos pseudoautosómicas (aproximadamente 3 Mb de longitud) que recombinan con el cromosoma X y una región no recombinante o NRY (del inglés *Non-recombining Region of the Y-chromosome*) que constituye el 95% del cromosoma (Jobling & Tyler-Smith, 2003). Esto último permite que la mayor parte del cromosoma Y sea transmitido entero de padres a hijos, pudiendo recopilar la historia del linaje paterno a través del análisis de polimorfismos, que son de gran utilidad a la hora de realizar reconstrucciones genealógicas, estudios forenses, evolutivos y poblacionales (Y Chromosome Consortium, 2012).

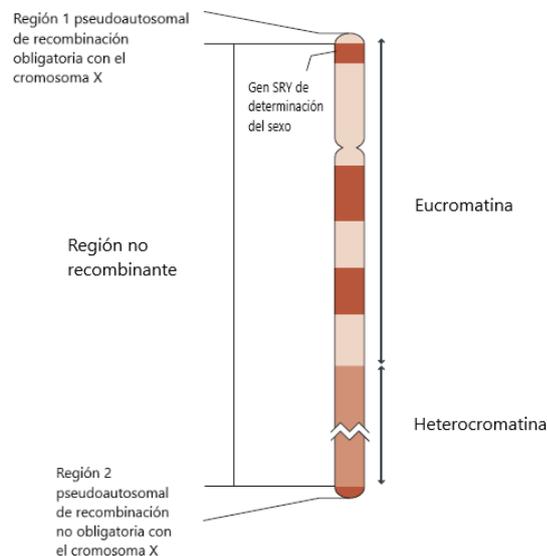


Figura 1.2. El cromosoma Y y su región no recombinante. Tomado de Jobling et al. (2014), página 37.

1.1.2 Filogeografía y distribución de las poblaciones americanas

Tanto el ADNmt como la NRY tienen en común que no están sujetos a recombinación. Esta característica posibilita que la única fuente de variabilidad sea la incorporación de nuevas variantes por mutación. La acumulación de mutaciones en forma secuencial provee información que se encuentra conservada en la secuencia a través del tiempo (Bravi, 2013). En este contexto se define un marcador polimórfico como aquel que presenta al menos dos formas distinguibles, llamadas alelos, con una frecuencia $\geq 1\%$ para el menos común de ellos. La combinación de dos o más estados alélicos de distintos marcadores en una misma molécula de ADN se denomina haplotipo (Jobling et al., 2004). Un conjunto de haplotipos de origen monofilético conforma un linaje o haplogrupo y se define como el conjunto de mutaciones características, heredadas todas en bloque, para las que puede suponerse un origen común (Bisso-Machado et al., 2016). Éstos últimos pueden ser utilizados para construir una filogenia robusta utilizando el principio de máxima parsimonia. A través de análisis filogenéticos se puede definir la jerarquía de las variantes genéticas; estas relaciones jerárquicas reflejan los efectos de la migración y la deriva génica, dos fuerzas que moldean la variabilidad poblacional (Underhill & Kivisild, 2007).

Utilizando polimorfismos binarios con bajas tasas de mutación como son los polimorfismos de nucleótido simple (PNS), una única filogenia puede ser fácilmente construida (Jobling & Tyler-Smith, 2003). La confiabilidad de los PNS bialélicos para la reconstrucción filogenética está dada por la baja probabilidad de que dos de estos eventos mutacionales ocurran en el mismo locus. Tal es así, que actualmente las versiones más recientes del árbol filogenético humano, tanto para el ADNmt como para el cromosoma Y, se basan en este tipo de marcadores.

1.1.2.1 Filogenia del ADN mitocondrial humano

El ADNmt es una fuente de información fundamental para el estudio de la diversidad genética en las poblaciones humanas. El árbol filogenético del ADNmt se construye sobre la base de un ancestro común por línea materna, originario de África subsahariana que dataría de hace aproximadamente 200.000 años (Behar et al., 2008) y se habría dispersado hacia fuera del continente hace aproximadamente 45.000 años (Mishmar et al., 2003; Macaulay et al., 2005). Tal como se muestra en la Figura 1.3, la estructura del árbol filogenético del ADNmt se conforma de una serie de haplogrupos nombrados de la A a la Z y de subhaplogrupos con menor o mayor variabilidad hacia el interior de cada linaje. El ancestro común más reciente de los humanos modernos se identifica con la letra "L". La rama L0 es una de las más antiguas y se encuentra en

poblaciones de África del sur y del este, principalmente en los pueblos khoisan; L1'5 (haplogrupos L1-L6) se encuentran en África subsahariana y, particularmente los haplogrupos de la rama L3, en el resto del mundo, conformando el haplogrupo del que derivan todos los demás linajes no africanos (Behar et al., 2008). Por ejemplo, el macrohaplogrupo M -que agrupa, entre otros, los haplogrupos C, D, G, E Q, Z- y ciertos subhaplogrupos de M se encuentran principalmente en Asia (Kivisild et al., 2002). Por otra parte, el macrohaplogrupo N -que incluye los haplogrupos A, I, R, S, W, X, Y- y varios subhaplogrupos de N son muy frecuentes en Eurasia. El haplogrupo R -que comprende los haplogrupos R0, B, F, H, J, P, T, U, V- habría divergido de N muy tempranamente, conformando uno de los tres linajes fundadores de Eurasia, conjuntamente con los macrohaplogrupos M y N (Macaulay et al., 2005).

El continente americano fue el último en ser poblado por la especie humana. Estudios realizados desde los campos de la arqueología y la genética concuerdan en delimitar la llegada de las poblaciones humanas al continente luego del Último Máximo Glacial, hace entre 15.000 y 18.000 años aproximadamente (Dillehay et al., 2008; Erlandson et al., 2015; Dawe & Kornfeld, 2017; Potter et al., 2017). Los haplogrupos A, B, C, D y X se consideran los linajes fundadores de América; mientras que X se encuentra restringido a Norteamérica, los haplogrupos A-D se distribuyen desde Canadá a Tierra del Fuego (Torroni et al., 1992; Bailliet et al., 1994; Bandelt et al., 1995; entre otros). El avance en técnicas moleculares y de secuenciación ha permitido una mayor resolución en el conocimiento de la variabilidad genética mitocondrial a partir del análisis de secuencias completas de la región control y de mitogenomas. De esta forma, se pudieron identificar un mayor número de linajes nativos fundadores, de los cuales 8 se conoce que ingresaron a Sudamérica: A2, B2, C1b, C1d, C1c, C4c, D1, D4h3a (Tamm et al., 2007; Achilli et al., 2008; Perego et al., 2009; Perego et al., 2010; Malhi et al., 2010). La distribución diferencial de los subclados de estos haplogrupos nativos, algunos de los cuales se encuentran dispersos por todo Sudamérica y otros restringidos, por ejemplo, al Cono Sur (D1j, D1g, B2i y C1b13) o a la Patagonia Fueguina (D4h3a5) (Bodner et al., 2012; de Saint Pierre et al., 2012), ha permitido refinar las filogenias de los 5 haplogrupos mayores basales (A-D y X). Tal como plantea Motti et al. (2017), esto último resulta necesario para poder confirmar los patrones de distribución de las poblaciones planteados hasta el momento o sugerir otros nuevos. Más aún, poder definir subhaplogrupos que se diferencian localmente aumentaría nuestra capacidad para interpretar la diferenciación poblacional hacia el interior del continente, y de esta forma, discutir modelos de poblamiento a escala regional (Motti et al., 2017).

La información acumulada acerca de la variabilidad genética de los pueblos originarios ha contribuido a mejorar la comprensión de la historia evolutiva y los procesos de migración de estos grupos humanos (Ribeiro dos Santos et al., 2013). En este sentido, se ha planteado que la

colonización del continente podría implicar desde una (Tamm et al., 2007; Kitchen et al., 2008; Fagundes et al., 2008) a múltiples oleadas migratorias desde Asia (Greenberg et al., 1986; Lell et al., 2002; Bortolini et al., 2003; Volodko et al., 2008; Ray et al., 2010; Perego et al., 2010). Reich et al. (2012), basándose en información genómica, señalan un patrón similar al modelo planteado inicialmente por Greenberg et al. (1986), postulando tres oleadas migratorias pero que no fueron equivalentes. Según estos autores, la primera de ellas habría poblado todo el continente e incluiría a los indígenas pertenecientes a la familia lingüística Amerindia (todos los nativos excepto los na-dene, los inuit y los aleutianos), que derivarían de una única población ancestral proveniente de Asia; seguida por al menos dos migraciones más con impacto restringido.

Por otra parte, se han planteado dos rutas para el ingreso a Sudamérica, una por la costa pacífica y otra por la vertiente atlántica (Rothhammer & Dillehay, 2009). Evidencias que provienen de estudios genéticos, craneométricos y lingüísticos apoyan este modelo, que implicó la separación temprana de poblaciones andinas (al oeste) y de poblaciones continentales al este (Cavalli-Sforza et al., 1994; Luiselli et al., 2000; Rothhammer et al., 2001; Tarazona-Santos et al., 2001; Keyeux et al., 2002; Pucciarelli et al., 2006; Wang et al., 2007; Rothhammer & Dillehay, 2009; Yang et al., 2010; Luisi et al., 2020; Campelo dos Santos et al., 2022). Las poblaciones pertenecientes a la región andina se caracterizan por un tamaño efectivo mayor y mayores niveles de flujo génico, lo que resulta en una relativa homogeneización de su acervo génico; por el contrario, en las poblaciones del este, en general pequeñas y aisladas, predominó la acción de la deriva (Tarazona-Santos et al., 2001). Estos factores se han reflejado en patrones genéticos distintivos y variabilidad morfológica y cultural entre las poblaciones del este y del oeste (Rothhammer & Dillehay, 2009).



Figura 1.3. Árbol filogenético del ADNmt. Tomado de Jobling et al. (2014), página 604

1.1.2.2 Filogenia del cromosoma Y humano

Las investigaciones en torno al cromosoma Y han contribuido a develar gran parte de la historia de las poblaciones humanas. Los estudios que describieron la variabilidad de la NRY, combinaron un número limitado de PNS y de STR (del inglés, *Short Tandem Repeats*, también conocidos como microsatélites) a partir de los cuales se construyeron los primeros árboles filogenéticos (Y Chromosome Consortium, 2002; Jobling & Tyler-Smith, 2003; Karafet et al., 2008). El descubrimiento de nuevos marcadores hizo posible que se complejizara cada vez más

la filogenia del cromosoma Y. En este sentido, van Oven et al. (2014), compilaron 417 marcadores y los resumieron en un árbol filogenético teniendo en cuenta los PNS más informativos. Un esquema reducido se presenta en la Figura 1.4. En particular, la raíz del árbol delimita al ancestro común más reciente por línea paterna de todos los humanos modernos, que, si bien aún se discute su antigüedad, el hecho de que los haplogrupos A y B se hayan encontrado en población africana, se encuentra en concordancia con el origen africano de los humanos modernos (Hammer et al., 1998).

Los avances en las tecnologías de secuenciación, en conjunto con el abaratamiento de los costos por parte de las empresas que brindan este servicio, han permitido secuenciar al cromosoma Y en forma completa. Más aún, el surgimiento de la secuenciación de nueva generación (*Next Generation Sequencing*, NGS), ha posibilitado la acumulación de masivos sets de datos genéticos, que fueron utilizados para redefinir árboles filogenéticos robustos (Jobling & Tyler-Smith, 2017), por ejemplo: Y Chromosome Consortium (2010); Wei et al. (2013); Larmuseau et al. (2015); 1000 Genomes Project Consortium (2016); entre otros. Recientemente, distintos autores han realizado reconstrucciones filogenéticas calibradas del cromosoma Y basados en datos genómicos, con aportes importantes al conocimiento filogenético del haplogrupo Q (Pinotti et al., 2019; Grugni et al., 2019; Paz Sepúlveda et al., 2022). El haplogrupo Q, definido por la mutación M242 (Bortolini et al., 2003), es el único haplogrupo con distribución panamericana y representa a prácticamente todos los linajes paternos de los nativos americanos en Mesoamérica y Sudamérica (Grugni et al., 2019). El subhaplogrupo Q-M3, propio de amerindios, es considerado como el linaje fundador de América (Pena et al., 1995; Bortolini et al., 2003). Otros linajes de Q como Q-M346 (xM3), se han encontrado en Chile, Bolivia, Argentina y el norte de Sudamérica, aunque en frecuencias más bajas y no tan ampliamente distribuidos como Q-M3 (Bailliet et al., 2011; Bailliet et al., 2009; Ramallo et al., 2009; Roewer, 2013).

Los aportes realizados por Pinotti et al. (2019) en base a los resultados filogenéticos y a las dataciones estimadas con 222 secuencias completas de todo el mundo pertenecientes a los haplogrupos Q y C, permiten estimar una ocupación inicial de Norteamérica hace 19.500 años. Además, estos autores identifican tres de cuatro linajes independientes como autóctonos y posiblemente como linajes fundadores: el linaje mayor Q-M3 y el linaje Q-CTS1780, que se encuentran en América; el linaje C3-MPB373 que se encuentra en Sudamérica con muy baja frecuencia y C3-P39/Z30536 encontrado en Norteamérica. A su vez, reafirman para el subhaplogrupo Q-M848 (perteneciente a Q-M3) lo que ya se había propuesto en trabajos anteriores: una rápida expansión de aproximadamente 15.000 años asociada a la colonización inicial de Sudamérica (Poznik et al., 2016), seguido por el establecimiento de una estructura espacial sustancial que surgió hace aproximadamente 12.300 años. Paz Sepúlveda et al. (2022),

basados en 102 secuencias genómicas del cromosoma Y, de las cuales 13 fueron secuencias nuevas obtenidas de poblaciones de Argentina y 89 fueron secuencias publicadas del resto del mundo (76 nativos y 13 euroasiáticos), determinaron que al interior de Q-M3, Q-M848 es el subhaplogrupo autóctono más frecuente en América. Además, describieron los tiempos de divergencia, la estructura espacial y nuevos PNS para Q-Z780, un subhaplogrupo autóctono encontrado en baja frecuencia en América y que provee evidencias genéticas del poblamiento de Sudamérica hace más de 18.000 años.

Cabe destacar que entre los haplogrupos extra-americanos, el haplogrupo R caracterizado por la mutación M207, es el más frecuente en Europa y se encuentra además en el centro, oeste y sur de Asia (Karafet et al., 2008, Myres et al., 2011, van Oven et al., 2014).

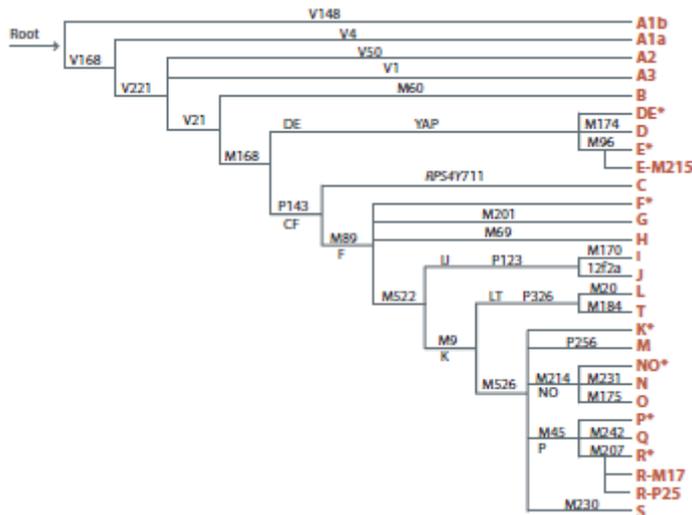


Figura 1.4. Árbol del cromosoma Y. Tomado de Jobling et al. (2014), página 607

1.2 Procesos de mestizaje y configuración de las poblaciones humanas americanas actuales

Los patrones de diversidad genética de las poblaciones americanas precolombinas han sido moldeados por el influjo de europeos desde la conquista y por un importante aporte de africanos durante los tiempos coloniales (Homburger et al., 2015; Ongaro et al., 2019). La constitución actual de las poblaciones regionales argentinas ha implicado procesos de mestizaje entre nativos y migrantes. En particular, en Argentina se registra un pico de flujo migratorio europeo entre 1860 y 1930 y, a diferencia de otros países aledaños, se caracteriza por el menor

aporte de población de origen africano (Studer, 1958). Datos del Censo Nacional de 1914 demuestran que, para ese entonces, alrededor de un 30% de la población total era extranjera y que el 60% de éstos eran varones. Los asentamientos de esta población inmigrante se orientaron a zonas geográficas específicas del país, como la región pampeana-litoraleña y Mendoza (Censo Nacional, 1914). Por otra parte, asentamientos diferenciales provenientes de países limítrofes conformaron otra fuente importante de heterogeneidad en la constitución de las poblaciones regionales argentinas. Si bien los grandes centros urbanos como Córdoba, Mendoza, Rosario, Tucumán y el Área Metropolitana de Buenos Aires (AMBA) han funcionado como atractores de migrantes de todos los países vecinos, los datos censales permiten reconocer un claro patrón de distribución sesgada en el que Cuyo y Patagonia atraen diferencialmente a chilenos, Salta, Jujuy y Tucumán –aunque también Mendoza, Córdoba y AMBA- a aquellos provenientes de Bolivia y las provincias del Nordeste y el Litoral a población paraguaya (Censo Nacional, 2010).

Estudios basados en el análisis de grupos sanguíneos y polimorfismos del ADN han reivindicado la importancia del aporte indígena en la población actual de Argentina (Dipierri et al., 1998; Avena et al., 1999; Martínez-Marignac et al., 1999; Morales et al., 2000; Alfaro et al., 2004). Los estudios genéticos de linajes haploides de transmisión exclusivamente materna o paterna, que permiten definir un origen continental remoto, han revelado un proceso de mestizaje entre las poblaciones nativas preexistentes y los migrantes -en su mayoría europeos y varones- de tipo sexo-asimétrico (para Argentina: Avena et al., 2007; Motti, 2012; García et al., 2018; Schwab, 2018; entre otros). La presencia de linajes maternos nativos puede incluso ser muy significativa en poblaciones urbanas de Argentina, alcanzando en algunos casos frecuencias mayores al 80% (v.g. en el NOA, Motti, 2012). Como contrapartida, la frecuencia de linajes paternos nativos en poblaciones cosmopolitas argentinas oscila alrededor del 10%, aumentando drásticamente hasta aproximadamente 40% en el NOA (ver, entre otros, Ramallo et al., 2009). De esta forma, se observa que los linajes maternos son en su mayoría autóctonos y los linajes paternos alóctonos.

Particularmente, para la región del NEA la mayor parte de los trabajos se ha enfocado en el estudio de poblaciones aborígenes. Respecto a los estudios realizados por línea materna, Cabana et al. (2006) analizaron secuencias RHV-I en 209 qom, pilagás y wichís de Chaco y Formosa, reevaluando la hipótesis formulada en el trabajo de Demarchi et al. (2001) en donde, a partir del análisis de las frecuencias de haplogrupos de las mismas poblaciones, se planteaba que la población del Gran Chaco presenta un patrón único de diversidad genética debido a una historia de poblamiento regional distintiva. Cabana et al. (2006) concluyen lo contrario, que ni la

estructura genética regional del Gran Chaco ni la subregional del Gran Chaco Central es única dentro de Sudamérica.

En otro trabajo realizado por Demarchi (2014), en donde se analizó la estructura genética de poblaciones nativas del Gran Chaco a partir de marcadores moleculares de herencia uni y biparental, se observó que las distancias genéticas son bajas y muestran, cada una de ellas, historias diferentes. Por ejemplo, los resultados sugieren la proximidad genética entre las familias lingüísticas matakó y guaykurú, lo que implicaría un flujo génico sostenido entre ellas. De esta forma, se puede afirmar que el lenguaje no parece ser una barrera importante para el flujo de genes en esta región. Estos resultados, podrían estar reflejando patrones culturales anteriores a la colonización europea o la ruptura de unidades poblacionales luego de ésta (Demarchi, 2014).

Por otra parte, Sala et al., (2010) analizaron STR autosómicos, marcadores del cromosoma Y así como secuencias RHV-I y RHV-II de 121 mbyá-guaraníes en 6 comunidades de la provincia de Misiones, indicando que los guaraníes argentinos son más cercanos genéticamente a los guaraníes de Brasil que a otros grupos nativos que se encuentran geográficamente más cercanos.

Sevini et al. (2013) publicaron secuencias parciales de la región control para 342 wichís y 210 criollos que conviven en una misma localidad rural de la provincia del Chaco, informando que gran parte de los haplotipos encontrados (106) son exclusivos de los wichís (36%) o de los criollos (51%), mientras que un 13,2% son compartidos entre ambas poblaciones. Asimismo, analizando la frecuencia de haplogrupos de la población criolla, estos autores reportaron un aporte mayoritario de haplogrupos nativos (98,1%) frente a un aporte europeo de sólo 1,9% por línea materna. Finalmente, analizando las distancias genéticas entre la población estudiada con otras de la bibliografía, informan que pese a un origen nativo americano común, un componente andino y otro amazónico emergen en los criollos y wichís, respectivamente.

Además del trabajo de Sevini et al. (2013), sólo tres trabajos publicaron secuencias mitocondriales derivadas de poblaciones “cosmopolitas” o mestizas asignadas a las provincias de Corrientes, Formosa y Misiones, dos de los cuales presentan tamaños muestrales reducidos: por un lado, Bobillo et al. (2010) informan secuencias de la región control completa para 26, 19 y 47 individuos, respectivamente, en donde se observa un aporte mayoritario de un componente nativo en Corrientes y Formosa (mayor al 80%), mientras que para Misiones el aporte autóctono es cercano al 47%, con un gran aporte europeo (49%) y un pequeño componente africano (4,3%). Por otra parte, Catelli et al. (2011) analizan las RHV-I y II en dos, uno y dos individuos de las provincias mencionadas, respectivamente. El tercer trabajo realizado en Misiones, analiza secuencias de RHV-I de 261 individuos (Badano et al., 2018), en donde se

observa un componente nativo americano mayoritario (68,6%) en comparación al componente alóctono (26,1% europeo y 5,3% africano).

En estudios de linajes paternos, que recopilan secuencias de poblaciones indígenas del NEA realizados por Ramallo et al. (2009), se analizan haplotipos de 93 individuos wichí de Formosa; mientras que otro estudio realizado por Jurado Medina et al. (2014) recopila haplotipos de 118 individuos pertenecientes a las etnias ayoreo, lengua, chorote, qom, wichí y mocoví del Gran Chaco, a partir de muestras de otros trabajos previos (Bianchi et al, 1998; Bailliet et al., 2009; Ramallo et al., 2009). Estos últimos autores, en concordancia con las conclusiones propuestas por Demarchi (2014) para la línea materna, plantean que las poblaciones chaqueñas muestran gran diversificación por línea paterna sin evidencia de aislamiento reproductivo entre ellas, ya que linajes idénticos fueron compartidos entre distintas poblaciones. Tal como ocurre para los estudios de ADNmt, los estudios previos que analizan los linajes paternos en poblaciones cosmopolitas del NEA son escasos. En particular, aquellos realizados por Ramallo (2009) y Jurado Medina (2015) han arrojado frecuencias variables de haplogrupos nativos (entre aproximadamente el 10% y el 40%) para diferentes localidades de Argentina. Por último, en el trabajo de Sevini et al. (2013) previamente mencionado, se estudiaron los linajes paternos de 100 wichí y 84 criollos, obteniéndose una frecuencia del haplogrupo nativo de 80% y 20%, respectivamente, en comparación con un componente europeo de 20% y 80%.

Se considera, tal como plantea Luisi et al. (2020), que centrarse sólo en las comunidades indígenas resulta insuficiente, al menos en Argentina, para poder caracterizar la diversidad genética originaria de América. Pese a la gran cantidad de información que se encuentra disponible sobre el poblamiento humano de América, la diversificación y la interacción local y regional entre poblaciones suele estar pobremente documentada (García et al., 2018). La finalidad de este trabajo consiste en contribuir al conocimiento de la frecuencia de linajes maternos y paternos en la región del NEA y su distribución a nivel local, regional y sudamericano. Un análisis pormenorizado de la fracción nativa, a partir de secuencias de la Región Control del ADNmt, permitirá hacer inferencias a escala regional, mientras que la elección de clados específicos mitocondriales en base a secuencias de fragmentos de la región codificante y de mitogenomas de la bibliografía, permitirá adentrarnos en un nivel de resolución aún mayor, de forma tal de dilucidar la distribución de estos linajes a escala continental. Los datos moleculares se analizarán en el marco de las evidencias bioarqueológicas, lingüísticas, históricas y demográficas, de manera de contribuir a la confirmación de patrones de distribución de las poblaciones que ya se encuentran estudiados y/o de plantear otros nuevos.

1.3 Definición de la zona de estudio

1.3.1 Geográfica

La región del Nordeste argentino (NEA) comprende las actuales provincias de Corrientes¹, Chaco, Misiones y Formosa; comienza a cobrar entidad como región en la década de 1950, cuando se provincializan las tres últimas jurisdicciones² (Valenzuela, 2014). Desde una perspectiva geohistórica, Ernesto Maeder (2015), en conjunto con otros autores, han intentado rastrear la existencia de la región antes de que se defina formalmente como tal. Particularmente, desde el punto de vista geográfico, Bruniard y Bolsi (1992) han definido al NEA como un “variado mosaico de paisajes geográficos” debido a la asociación de diversas formas naturales y culturales, por lo que resulta difícil de asimilar a este territorio como una región homogénea o como una región nodal y funcional teniendo en cuenta su génesis y su funcionamiento (Leoni, 2016).

El NEA, conjuntamente con las provincias de Santa Fe y Entre Ríos, integran la región del Litoral; la cual, debido a sus ambientes heterogéneos, se encuentra diferenciada en dos subregiones: la chaqueña y la mesopotámica. En particular, la subregión chaqueña, que agrupa a Chaco y Formosa, se trata de una extensa llanura boscosa, plana y uniforme, con clima predominantemente subtropical mientras que la subregión mesopotámica se encuentra delimitada por los cursos de los ríos Uruguay y Paraná, agrupando a las provincias de Misiones, Corrientes y Entre Ríos (Ministerio del Interior de la República Argentina, 2006). El relieve en esta subregión, varía de norte a sur, desde la meseta misionera hasta los esteros correntinos, las lomadas entrerrianas (de 100 a 200 metros de altura) y los bajos del delta (Hernández, 1995). A su vez, el clima es cálido y húmedo hacia el norte, y templado y seco al oeste, con precipitaciones que disminuyen de este a oeste.

A la hora de definir la zona de estudio de este trabajo, un criterio simple sería basarse en los límites actuales del NEA en términos geopolíticos. Sin embargo, se observa que históricamente, los movimientos poblacionales y el impacto cultural no necesariamente han respondido a factores exclusivamente geográficos y mucho menos a límites políticos (Hernández, 1995). Además, se debe tener en cuenta que los procesos poblacionales aquí estudiados anteceden a la conformación del Estado Argentino. En este sentido, estudios previos

¹ Se provincializó en 1814.

² Inicialmente se provincializó el Territorio Nacional del Chaco el 20 de julio de 1951 (ley 14037), luego Misiones el 10 de diciembre de 1953 (ley 14.294) y Formosa, el 28 de junio de 1955 (ley 14.408).

han interpretado el desenvolvimiento de las culturas indígenas prehispánicas en la Argentina basándose en la información etnoarqueológica disponible y el impacto de la llegada de los españoles en estos grupos. Por ejemplo, Isabel Hernández (1995) ha desarrollado el concepto de “área cultural” como criterio sistematizador que integra las clasificaciones regionales, culturales, lingüísticas y adaptativas en relación al espacio-hábitat. A su vez, Serrano (1947) estudió a los indígenas argentinos por núcleos étnicos o naciones, las cuales pueden o no integrar otras mayores dentro del continente o presentar, como los guaraníes, una dispersión geográfica discontinua. En base a los trabajos de estos dos autores se hará mención a los grupos indígenas en la región. Se describirá en esta sección las áreas culturales Litoral y Mesopotamia y la zona del Gran Chaco que comprende el Chaco Austral o argentino, complementando y ampliando de esta forma, la definición de la zona de estudio.

El área cultural Litoral y Mesopotamia abarca las provincias actuales de Misiones, Corrientes y Entre Ríos, los márgenes de los ríos Paraná y Uruguay y las provincias de Formosa, Chaco y Santa Fe. En la franja boreal se encuentran los ríos Paraguay y el Iguazú, mientras que en la franja austral la conforman el Delta del Paraná y el Río de La Plata. Esta región se caracteriza por un variado abanico de características climáticas y geográficas que deben tomarse en consideración a la hora de analizar la dinámica poblacional de la zona, así como la interrelación de ciertos grupos con el planalto brasileño y la zona colindante del Paraguay (Hernández, 1995).

Por otra parte, el Gran Chaco se ubica en la zona de Tierras Bajas de Sudamérica, comprendiendo un área de un millón de kilómetros cuadrados, que abarca en su extensión territorios de Argentina, Bolivia y Paraguay. En particular, el Chaco Austral o argentino se trata de una llanura extensa que limita al norte con el planalto central brasilero, al sur con el Río Salado y las Sierras de Córdoba y Guasayán, al este con los ríos Paraguay y Paraná y al oeste con las Sierras Subandinas. Incluye a las actuales provincias de Chaco y Formosa, el norte de Santa Fe y el oriente de Salta y Santiago del Estero. Se encuentra en íntima relación con el Chaco Boreal por su continuación en territorio de escasa pendiente y el curso de los ríos que es inconstante, con la presencia de ciertos cauces abandonados, incluyendo parte de los actuales territorios de Paraguay y Bolivia (Hernández, 1995).

1.3.2 Arqueológica

Las primeras investigaciones arqueológicas en la región del Nordeste argentino y Litoral datan de finales del siglo XIX y fueron realizadas bajo enfoques histórico-culturales que permitieron establecer distintas categorías de clasificación y análisis de las secuencias culturales, las cuales abarcaron dos grandes períodos: uno precerámico y otro cerámico (Serrano, 1972;

Caggiano, 1984; Rodríguez, 2001). Escudero et al. (2010) presentan una síntesis de los modelos y esquemas culturales de la región que encuentran puntos en común bajo clasificaciones como "cultura", "entidad", "fase" y/o "tradicición" (Lafón, 1971; Serrano, 1972; Caggiano, 1984; Ceruti, 1993; Ceruti & González, 2007; Rodríguez, 2001).

Las ocupaciones humanas iniciales en la región, fueron asociadas a la tradición "Paleoindígena" (Rodríguez, 2006) en la cuenca media del Río Uruguay durante la transición Pleistoceno-Holoceno y el Holoceno temprano (ca. 14.000-8200 años antes del presente, AP). Estas áreas comprenden el sudoeste de Brasil y el norte de Uruguay, sin embargo, para el territorio argentino en cercanías al río Uruguay, sólo se han identificado algunos hallazgos superficiales de puntas de proyectil cola de pescado y Pay Paso, además de un lito discoidal que se estima de entre ca. 10.725 y 10.140 años ¹⁴C AP de antigüedad, en función de hallazgos semejantes en contextos datados (Castro & Terranova, 2015; Castro, 2017; 2019).

En un período temprano desde la introducción de la cerámica, se identifica una cultura "Entrerriana o Básica del Litoral" que se trataba de cazadores y pescadores que dominaron en la Cuenca del Paraná y del río Uruguay medio e inferior, "caracterizados por una alfarería de formas simples, lisa o decorada, sin aditamentos zoomorfos"; esta cultura habría antecedido a la de los "Ribereños Plásticos" en el Paraná medio e inferior, que, conjuntamente con la cultura "Guaraní", se presentan más tardíamente en el registro arqueológico de la región (Serrano, 1972). Otras sistematizaciones han planteado las tradiciones de "Cazadores-recolectores", "Neolítica", "Tupí Guaraní generalizada" (Lafón, 1971), "Sabanas bajas", "Platense", "Planáltica", "Ribereña Paranaense" (que incluye a los complejos Goya-Malabrigo y Lechiguanas) y "Tupí guaraní" (Rodríguez, 2001). Por otra parte, Ceruti (1986; 2003) propuso un modelo de "entidades culturales" integrando los modelos de desarrollo cultural postulados previamente, pero enfocado en los cambios paleoclimáticos de la región. Entre las entidades culturales definidas por este autor se encuentran "La Esperanza", "Cancha de Luisa" y "Villa Cañás" para la última etapa árida (3000 a 1500 o 1000 AP) y "Goya-Malabrigo" para el cambio a la etapa húmeda (1500 a 1000 AP hasta el presente). Esta última entidad, sería una reformulación de la de los "Ribereños Plásticos" de Serrano (1972).

Cuando se analiza la arqueología de la región del Nordeste y Litoral, es necesario contextualizarla en la arqueología de las Tierras Bajas Sudamericanas y de sus ambientes fluviales. Los cursos hídricos que atraviesan la zona sur son los ríos Pilcomayo, Bermejo, Paraguay, Paraná, Uruguay y Río de la Plata, que han sido considerados "escenarios naturales privilegiados" para el desarrollo de la historia indígena, como resultado de la implementación de estrategias adaptativas para el manejo de ambientes lacustres, planicies húmedas y bosques ribereños (Bonomo et al., 2019). Las sociedades indígenas relacionadas con estos cursos de agua

basaron su economía en la pesca, la caza y la recolección, con asentamientos en los bordes de lagunas y ríos, albardones y elevaciones para superar las inundaciones (Lamenza et al., 2019). Una característica distintiva de la región es la presencia de montículos de tierra o cerritos de origen antrópico. Estas construcciones han sido depositarias de importantes hallazgos arqueológicos y protagonistas de una evolución constructiva que da cuenta de que su mantenimiento y reclamación cultural se mantuvo a lo largo de las generaciones (Bonomo et al., 2019). En particular, una parte de los montículos del Paraná Inferior se asignaron a la entidad arqueológica Goya-Malabrigo, y se ha determinado que, en principio, tenían funciones residenciales y funerarias; además, se han asociado con sociedades con cierta complejidad o jerarquías sociales y con prácticas hortícolas a pequeña escala (Politis & Bonomo, 2012; 2018a y b).

La entidad Goya-Malabrigo se caracteriza por un estilo cerámico con apéndices zoomorfos y antropomorfos, campanas, cucharas, vasijas globulares cerradas, entre otros; habría ocupado la planicie de inundación del Paraná Medio e Inferior y los ambientes fluviales vecinos (Politis et al., 2018a; Bonomo et al., 2019) desde *ca.* 2400 años AP hasta los primeros siglos de la colonia (Castro et al., 2020). Politis y Bonomo (2012) redefinieron esta entidad cultural a “entidad arqueológica Goya-Malabrigo” incorporando ciertos rasgos distintivos como la construcción de montículos, un uso intensivo de la arcilla en las diferentes dimensiones de la cultura material y una economía mixta de caza, pesca y recolección y una horticultura de maíz, poroto y zapallo. Además, estos autores volvieron a poner en foco de discusión la hipótesis de filiación arawak de Goya-Malabrigo. Los arawak, al igual que los guaraníes, descienden de poblaciones que probablemente tuvieron su origen en la Amazonía. La recopilación de evidencias realizada por los autores, aplicando métodos actuales (p. ej., estudios de almidones y fitolitos, de macrorrestos vegetales obtenidos por flotación y microestratigrafía) bajo un nuevo enfoque “neo-difusionista”, no esencialista, y basados en la teoría social, revela que la expansión/dispersión de los grupos arawak parlantes puede ser mejor comprendida como un “complejo proceso de etnogénesis basado en viajes y desplazamientos de segmentos específicos de la sociedad e intercambios regionales y pan-regionales más que como la migración masiva de poblaciones arawak” (Politis & Bonomo, 2012; 2023).

Por otra parte, los guaraníes se han asociado a una alfarería policroma roja y/o negra sobre engobe blanco, corrugada, unguiculada y cepillada y a otros artefactos como hachas de piedra pulida, distintos tipos de vasijas y entierros en urna (Bonomo et al., 2019). Como se mencionó anteriormente, tendrían un origen amazónico y se habrían expandido a través del sistema fluvial de la cuenca del Plata alcanzando tardíamente la región del NEA y Litoral. Corrêa (2023) propuso el territorio del actual estado de Rondônia como posible centro de origen, desde

donde una parte de los hablantes de las lenguas de la familia lingüística Tupí se desplazaron al este del río Tapajós. Allí, la autora plantea que ocurrió la formación de los pueblos Tupí-Guaraní a partir de diversos procesos históricos; y su expansión ocurrió gradualmente desde el sector que se ubica entre los ríos Tapajós y Trocantis, hace aproximadamente 2500 años AP. De esta forma, los hablantes de lenguas Tupí-Guaraní llevaron su tradición cerámica con características propias a zonas más al sur e incluso a la región del NEA (Corrêa, 2023). En este sentido, Bonomo et al. (2015) recopilaron información de 1600 sitios arqueológicos atribuidos a los guaraníes y alrededor de 250 dataciones absolutas (^{14}C y termoluminiscencia) y construyeron un modelo de distribución geográfica y temporal, a partir del cual pudieron establecer dos pulsos principales de expansión de los guaraníes por vías fluviales: el primero entre el 0 al 300 AD que es cuando se alcanza el Río Paraná, al norte de la provincia de Corrientes; el segundo pulso, entre el 1000 AD y la conquista, que utiliza como eje el Río Uruguay y continúa hacia el sur hasta el Río de la Plata (entre los ca. 1340 y 1530 años AD). A pesar de la separación temporal de casi dos mil años y de la distancia geográfica entre los sitios arqueológicos (que abarcan una parte importante de la cuenca del Plata y el litoral atlántico brasileño), los guaraníes llevaron consigo un modo particular de “ver y relacionarse con el mundo socialmente compartido”, representado en la cultura material (Torino, 2021). La cerámica guaraní refleja “un modo de hacer común”, que se evidencia a partir de similitudes encontradas en la manufactura, forma, tamaño, tratamientos de superficie y motivos pintados de las vasijas (Bonomo, 2013; Loponte & Acosta, 2007, 2008; Noelli, 2004; Torino, 2021; 2024; entre otros). A su vez, el análisis de artefactos cerámicos realizado por Torino (2024) en el Delta del Paraná y el Río de La Plata, considerando los materiales utilizados, permite concluir que éstos estaban sujetos a la disponibilidad de recursos del ambiente y en el caso de las pastas, también al contacto con otros grupos. El modelo de expansión guaraní planteado por Bonomo et al. (2015) se comparó con otro realizado para la distribución de la entidad Goya-Malabrigo; concluyendo que los territorios ocupados por estas dos entidades culturales no se superponían (Politis et al., 2018a). Sin embargo, Torino (2024) revela una interacción entre estos grupos debido a que tanto el estilo Guaraní como el Goya-Malabrigo tenían recetas de pasta cerámica similares. En este sentido, la autora plantea que los guaraníes lograron preservar un modo de ser y de hacer que perduró a lo largo del tiempo y el espacio con una singularidad única, gracias al mantenimiento de relaciones de reciprocidad entre los diferentes tipos de asentamientos guaraní y pese a los condicionantes externos (poblaciones locales e invasión europea) en el límite meridional de su expansión territorial (desde los 390 ± 20 años AP hasta momentos posteriores a la conquista).

Recientemente, distintos autores han planteado que para comprender el origen y desarrollo de las ocupaciones humanas es necesario contextualizar la zona de estudio a una

escala más amplia (Bonomo et al., 2011; 2019). En este sentido, Politis y Tissera (2023) recopilan evidencias sobre las interacciones humanas producidas durante el Holoceno tardío en el centro-este de Argentina desde una perspectiva macrorregional haciendo foco tanto en el intercambio de ítems no locales (como metales y conchillas, etc.) como en el movimiento de personas, recursos (plantas y animales), representaciones e información entre distintas sociedades y ambientes. Por ejemplo, la conexión entre la región sur de los Andes y el bajo Paraná se puede establecer a partir de la recuperación de piezas de metal en contextos funerarios, entendiendo que tanto las entidades Goya-Malabrigo como Guaraní no habrían desarrollado una tecnología metalúrgica (Bonomo et al., 2017; entre otros); además, la presencia de piezas cerámicas con apéndices de animales no locales como cóndores y llamas en alfarería perteneciente a Goya-Malabrigo, también darían cuenta de estos intercambios macrorregionales (Bonomo et al., 2021; entre otros). Sin embargo, estas conexiones no habrían sido directas: los grupos de las Sierras Centrales, la Planicie Santiagueña y los pastizales de la Pampa habrían mediado estas interacciones a través de las vías fluviales de la cuenca del Río Salado (Politis & Tissera, 2023).

1.3.3 Etnográfica y lingüística

Las condiciones climáticas cambiantes entre los períodos húmedos del Pleistoceno final y la sequía que caracterizó al Holoceno, afectaron particularmente las zonas de Tierras Bajas que conforman el área de estudio, alcanzando tardíamente su configuración actual. Es de suma importancia considerar estos factores -climáticos, paisajísticos y ecológicos- a la hora de interpretar los datos arqueológicos y correlacionarlos con fuentes históricas. En particular, en esta sección se intenta describir a los grupos humanos que habitaban la actual región del NEA y Litoral antes de la llegada de los conquistadores. Se hace referencia a la caracterización de los grupos humanos realizada por Hernández (1995), quien propone una subdivisión de la región de estudio en dos categorías: una Región del Litoral y Mesopotamia y una Región Chaqueña.

Región del Litoral y Mesopotamia

Se hará referencia a los pueblos que residieron en la región del Litoral y Mesopotamia a partir de dos modalidades culturales diferentes: los grupos que se encontraban en el interior del territorio, que eran cazadores-recolectores: kaingang o guayaná, querandí y charrúa; y los pueblos agricultores que se desplazaban por ríos y esteros: guaraní y chaná-timbú.

- *Grupos cazadores-recolectores*

Los guayanás, pertenecientes a la familia lingüística *ge*, eran provenientes del Brasil meridional y ocupaban parte del territorio de Misiones y el nordeste de Corrientes (Serrano,

2012), que compartieron posteriormente con los guaraní-abá. Los querandíes, según las crónicas del siglo XVI, ocupaban una amplia franja cuyo centro sería la ciudad de Buenos Aires e incluían el sur de Santa Fe. Se conoce muy poco de esta etnia más allá de su carácter nómada ya que este etnónimo no registra continuidad en los siglos posteriores. Se han postulado diversas hipótesis que incluyen vínculos con los guaraníes (Conlazo, 1990), los tehuelches (Casamiquela, 1965) o los guaycurúes (Censabella, 1999). El último de los grupos no agricultores, los charrúas, conforman una familia lingüística de cinco entidades étnicas: los bohane, los guenoa, los yaro, los minuane y los charrúa propiamente dichos (Serrano, 1947; 1950). Eran característicos del Uruguay, ocupando prácticamente el actual territorio del mismo (bohane y guenoa), mientras que en la Argentina vivieron en los márgenes del río Uruguay, desde Yapeyú (yaro) hasta el sur de Entre Ríos (minuane) en los márgenes del río Gualeguay.

- *Grupos horticultores*

Los guaraníes, pertenecientes a la familia lingüística Tupi-Guaraní, son uno de los grupos étnicos de mayor dispersión en América meridional. Los hablantes de guaraní antiguo ocupaban parte de la costa atlántica al sur de Paranaguá, Brasil meridional, Uruguay y los alrededores del Río de la Plata y se encontraron asociados a los cursos de agua de las cuencas de los ríos Paraguay, Paraná y Uruguay (Bonomo et al., 2019). Al momento de la llegada de los conquistadores, ocupaban zonas continuas desde el litoral atlántico brasileño y parte de la cuenca del Amazonas hasta el Chaco oriental y del pie de los Andes, así como grandes superficies de la cuenca del Plata; estas áreas de expansión se alteraron drásticamente partir del siglo XVI, producto de cambios por la conquista y colonización europea (Bonomo et al., 2015).

Chaná-timbú se llamó en forma genérica al conjunto de “parcialidades” entre las que se puede mencionar a los timbú, chaná, mbeguá, caracaraes, corondás, colastinés, quiloazas, entre otros, que se encontraban en la región insular de Entre Ríos y Buenos Aires, además de la cuenca del Paraná hasta aproximadamente la latitud del río Corrientes (Serrano, 1972). Politis y Bonomo (2012) plantean elementos clave en favor de una continuidad histórica entre los indígenas prehispánicos y las sociedades indígenas coloniales, en donde Chaná-timbú sería la expresión etnográfica-histórica de una continuación de la entidad arqueológica Goya-Malabrigo (o su homóloga Ribereños Plásticos) en el bajo Paraná, al menos hasta el siglo XVII. En este sentido, estos autores se basan en fechados radiocarbónicos del siglo XVI de varios sitios atribuidos a Goya-Malabrigo, en la presencia de remanentes coloniales en sitios tardíos atribuidos a Goya-Malabrigo y en la presencia de cerámica característica de Goya-Malabrigo en tiempos coloniales tempranos (como en *Sancti Spiritu*, 1527–1529 y Santa Fe la Vieja, 1573–1660/1670; Ceruti 2009; Cocco, 2016) y en el solapamiento entre Goya-Malabrigo y la distribución geográfica y el territorio histórico conocido de Chaná-timbú (Politis et al. 2018; Serrano 1972).

Región chaqueña

Por otra parte, para la región chaqueña, el poblamiento inicial suele documentarse como tardío en comparación con otras zonas de Argentina y de Sudamérica (Miller, 1979). Particularmente en el Chaco austral, las características ecológicas han limitado la ocupación humana de esta zona del continente, debido a que los hábitats se encuentran emplazados en estrecha relación con los cursos de agua, a la dinámica estacional que existe de los Ríos Bermejo y Pilcomayo, sumado a las diferencias entre la estación seca y húmeda (Calandra & Salceda 2004). Hernández (1995) plantea que esta zona habría permanecido como una región “pantanosas, lacustre e inhabitable” hasta el séptimo o sexto milenio AC y que los primeros aborígenes habrían llegado del planalto brasilero o anteriormente, de las zonas andinas o periandinas que lo rodean. En este sentido, la autora describe tres tipos culturalmente diferenciados de población aborigen prehispánica en la región chaqueña: los chaqueños típicos representados por las familias lingüísticas mbayá-guaycurú y matabo-mataguayo, las culturas andinizadas, conformados por la familia lingüística lule-vilela y las culturas amazónicas, de las familias lingüísticas tupí-guaraní y arawak.

La familia lingüística mbayá-guaycurú se conforma por los pilagá que habitaron el Chaco central, en las costas del Pilcomayo; los guaycurú, que se ubicaron en la ribera del Río Paraná al norte de la provincia de Formosa, diezmados luego de la conquista; los qom que se encontraban en el curso medio e inferior del río Bermejo; los mocoví al noroeste de la provincia de Santiago del Estero y norte de Santa Fe; y los abipones en la costa septentrional del río Bermejo inferior, de los que no quedan representantes actualmente (Hernández, 1995).

Los matabo-mataguayo, por otra parte, estaban representados por diez familias que habitaron el Chaco argentino, de las que sólo tres persisten en nuestro territorio: wichí, chorote y chulupí. Se dedicaban a la recolección y secundariamente a la caza y la pesca; el estilo de vida de estos grupos se asemeja al de los guaycurú; sin embargo, a diferencia de éstos mantuvieron su tradicional forma de vida sin la incorporación del caballo.

En el Chaco argentino occidental y central, así como en ciertas zonas de Salta, Santiago del Estero y oriente de Tucumán, habitaban los grupos lule y vilela (grupos nómades cazadores, recolectores y pescadores) y también los tonocoté (sedentarios y agricultores) pertenecientes a la familia lingüística matabo-maká; si bien se incorporaron a las misiones religiosas no prosperaron hasta la actualidad (Martínez Sarasola, 2013).

En relación a los tupí-guaraní se mencionará a los chiriguano, los cuales inicialmente vivían en Paraguay o en Santa Catalina (Brasil) y migraron antes de la conquista hacia Bolivia y luego al Chaco nor-occidental y al noreste de Salta, donde se encuentran actualmente. Al

penetrar en el territorio dominaron tanto a los chané como a sus vecinos andinos, de modo que su alfarería si bien es policroma como la del resto de los guaraníes, se vio influenciada por los contactos con otros grupos (Hernández, 1995).

Por último, dentro de la familia lingüística arawak se encontraba a los chané, una cultura diversa que comprendía desde grupos semisedentarios pequeños hasta aldeas con gran número de habitantes y que, en su punto máximo de dispersión hacia el sur, alcanzó, en territorio argentino, el Chaco centro occidental. Compartieron el territorio con los chiriguano, quienes ejercieron una fuerte dominación sobre estos grupos. Sin embargo, se considera que el “alma arawak” debe haber tenido una gran influencia como para haberse mantenido como una entidad propia a pesar de su escaso número de representantes, siendo la comunidad de Tuyuntí en la provincia de Salta, la descendiente actual de ese núcleo (Martínez Sarasola, 2013).

1.3.4 Histórica

La región del NEA puede abordarse como región histórica teniendo en cuenta dos etapas: una primera etapa desde tiempos coloniales hasta el segundo tercio del siglo XIX, que involucra la formación del estado correntino y de las misiones jesuíticas de guaraníes, la marginalidad del Chaco y la presencia de Paraguay en la cuenca del Paraná; y una segunda etapa a partir del segundo tercio del siglo XIX, en donde la organización nacional del país impulsó la creación de nuevos territorios en el marco del Nordeste Argentino y en donde ya podía vislumbrarse un centro jerárquico en el sentido regional (Maeder, 2015).

Respecto a las historias de asentamiento colonial, las ciudades de Corrientes y de Asunción del Paraguay, presentan algunas características comunes: ambas fueron ciudades de fundación temprana (año 1588 y 1537, respectivamente) y su poblamiento fue gradual (criollos en su mayoría y núcleos indígenas sometidos, agrupados en reducciones bajo tutela franciscana o del clero secular). En contraposición, los territorios de Chaco y Formosa fueron ocupados militarmente más recientemente, entre 1872-1884, luego de la guerra de la Triple Alianza. En el caso de Misiones la ocupación se realizó a partir de programas de colonización sin intervención de las fuerzas militares. Por lo tanto, a diferencia de Corrientes, el poblamiento de Chaco, Formosa y Misiones fue reciente, poblándose de inmigración tanto interna como externa, mientras que los grupos indígenas de la zona se ubicaron en espacios marginales del territorio (Maeder, 2015).

Los guaraníes habrían llegado tardíamente al territorio argentino donde finalmente se contactaron con los conquistadores en el siglo XVI (Brochado, 1984; Noelli, 1993, 2004; Bonomo et al., 2015). El arribo de los europeos al Río de La Plata afectó drásticamente la historia indígena

regional y alteró el proceso de expansión poblacional que habían desarrollado los guaraníes (Bonomo et al., 2019). Las dataciones absolutas recopiladas por Bonomo et al. (2015) permiten entrever al menos 17 siglos de historia de movimientos poblacionales y una ocupación permanente de varias regiones, alcanzándose el límite meridional de su expansión territorial alrededor del año 650 AP. A diferencia de otros grupos indígenas, los guaraníes se integraron a la situación colonial, a través de las misiones jesuíticas (Bonomo et al., 2019). Los padres jesuitas pertenecían a una orden religiosa fundada por Ignacio de Loyola; comenzaron la enseñanza de la fe católica en Europa y, en América, se centró en la evangelización de los pueblos guaraníes, en especial en Paraguay y el NEA hacia fines del siglo XVI y principios del siglo XVII. Para mediados de 1630, había cuarenta reducciones fundadas bajo la protección de los sacerdotes de la Compañía de Jesús; sin embargo, este éxito inicial se vio opacado por la llegada de *bandeirantes* paulistas, formaciones militares conformadas por portugueses, mestizos e indios tupíes que tenían como objetivo la búsqueda de indios guaraníes para ser utilizados como mano de obra esclava en la costa brasileña y que habría impactado en el número de misiones efectivas, limitando el espacio ocupado por ellas al sur de Paraguay, nordeste de Misiones y Corrientes (Maeder y Poenitz, 2006). Se conoce que particularmente tres de estas reducciones en la cuenca de los ríos Ibicuy y Jacuy: Nuestra Señora de Anunciación (1630), San Carlos (1631) y Santo Tomé (1632), debieron trasladarse a un sitio seguro al resguardo de las *bandeiras* paulistas y se asentaron tras la frontera del río Uruguay; junto a Yapeyú, integraron las misiones del Uruguay. Otro ámbito de conflicto se abría para estos pueblos debido a los movimientos inestables de las parcialidades charrúas con los que se alternaron períodos de intercambios y acercamientos con otros de tensión y enfrentamientos bélicos.

Pese a estas dificultades, los jesuitas lograron consolidar las misiones en esta región alcanzando un número de unos treinta pueblos, desarrollando gran cantidad de actividades económicas y culturales, y dando lugar al proceso de urbanización y poblamiento más significativo del Río de La Plata en el siglo XVII y que se mantuvo hasta 1768; momento en el cual la monarquía española expulsó a los padres jesuitas y secularizó las misiones, creando en ese distrito la Gobernación de Misiones a cargo de las intendencias de Buenos Aires y Paraguay (Maeder, 2004). De esta forma, las misiones perdieron importancia gradualmente y los pueblos fueron incluidos en distritos vecinos como el Paraguay o Corrientes.

En resumen, se observa que en la construcción del Estado-Nación los indígenas tuvieron un lugar paradójico: en principio, fueron integrados y sometidos a la producción colonial, dispersos en las campañas y periferias urbanas luego de las guerras por la independencia; posteriormente, debieron “alterar su autorrepresentación identitaria en una Nación en donde no se podía seguir siendo ‘indio’, so pena de ser excluido de la ciudadanía” (Grosso, 2008). Como

menciona Segato (2007), “civilización se definió como ‘neutralidad étnica’, vaciada de toda particularidad y ‘barbarie’ como su antagónico otro interior en constante retirada y pugna por retorno”.

2 CAPÍTULO 2: MATERIALES Y MÉTODOS

2.1 Esquema de trabajo

Para abordar las múltiples dimensiones que conforman la historia poblacional del NEA, se implementó un diseño metodológico que estructura el análisis de la diversidad genética en tres ejes con distintos niveles de resolución. El primer nivel se centra en la historia local de las migraciones intercontinentales, basándose en el análisis de las frecuencias de haplogrupos de distintos orígenes continentales (América, Europa y Oriente Medio y África), tanto de los linajes maternos como paternos, en las ciudades de Formosa y Corrientes. Se parte de la **hipótesis general** de que en las poblaciones cosmopolitas actuales es posible rastrear genes y linajes de poblaciones ya desaparecidas como aislados socioculturales, debido a los procesos de mestizaje que han ocurrido durante los últimos 500 años. Esto último ha sido corroborado para otras regiones de Argentina (Motti, 2012; García et al., 2018; Schwab, 2018; entre otros). A su vez, los resultados acerca del origen continental de los linajes paternos y maternos son contrastados con la información genealógica proporcionada por los individuos a través de una encuesta.

Un segundo nivel de análisis hace foco en la fracción nativa de los haplogrupos; se parte del análisis de haplotipos del ADNmt, que combinan PNS y secuencias. Esta técnica de análisis otorga mayor poder de resolución, identificándose subhaplogrupos diferenciados regionalmente que permiten una aproximación a la historia regional. Para ello, se obtuvieron secuencias de la localidad de Formosa, se analizaron secuencias previamente recolectadas por nuestro grupo de investigación y también secuencias tomadas de la bibliografía. El objetivo de este segundo nivel de análisis consiste en establecer las afinidades de las poblaciones estudiadas en el contexto del Cono Sur. El abordaje de la historia evolutiva de las poblaciones desde un análisis molecular permitirá definir subhaplogrupos que se diferencian localmente, aumentando la capacidad para interpretar la diferenciación poblacional hacia el interior del continente, y de esta forma, discutir modelos de poblamiento a escala regional (Motti et al., 2017). Teniendo en cuenta la estructura geográfica observable para el ADNmt, se espera que existan similitudes en la distribución de clados y haplotipos dentro de la región del NEA y con los países colindantes en íntima relación con la región (Paraguay, sur de Brasil y Uruguay); y diferencias con otras regiones de la Argentina.

Finalmente, en el tercer y último nivel, se definen y analizan detalladamente nuevos clados mitocondriales caracterizados a partir de la combinación de polimorfismos de las regiones control y codificante. En base a una compilación de mitogenomas publicados e inéditos,

se seleccionaron PNS informativos para los que se diseñaron ensayos PCR³-APLP⁴ con el objetivo de testear su presencia en las poblaciones que forman parte de la colección del Instituto Multidisciplinario de Biología Celular (IMBICE), y también se secuenciaron regiones codificantes seleccionadas potencialmente informativas. La información molecular colectada se utilizó para proponer presuntos clados monofiléticos de diferente jerarquía, explorando su distribución y estructuración a escala sudamericana o continental, con el objetivo de identificar las principales rutas de poblamiento seguidas por las poblaciones que se asentaron en el NEA.

2.2 Muestreos poblacionales

Se realizaron viajes de campaña durante agosto de 2015 a las ciudades capitales de Formosa y Corrientes. Estas ciudades fueron seleccionadas debido a que atraen migrantes de otros lugares de dichas provincias. Los muestreos se realizaron en los Bancos de Sangre Central de ambas localidades y en un laboratorio privado en la ciudad de Formosa (Laboratorio de Análisis Clínicos Motter). Se invitó a participar en este estudio a todas aquellas personas mayores de 18 años (inclusive), que concurrieran a la institución para extracción de sangre, ya sea para donación o análisis clínicos. Se otorgó a cada voluntario un tubo de 15 mL codificado con anticoagulante (conteniendo de 50-100 μ L de EDTA 0,5M) para entregar al extraccionista. Las muestras de sangre fueron tomadas por personal de la institución capacitado para dicha tarea.

De las 506 muestras de sangre colectadas, se seleccionaron 425 pertenecientes a las 3 estaciones de muestreo: 274 de Formosa y 151 de Corrientes, de las cuales 142 y 95 fueron varones, respectivamente. No se incluyeron participantes en donde existiera una relación de parentesco que implicara compartir el mismo linaje materno o paterno. En caso de constatar una relación de parentesco entre los distintos individuos, se seleccionó a uno de ellos, de forma tal que esté representado un solo individuo de ese linaje materno o paterno según correspondiera. No fueron incluidos aquellos sujetos que conformaban parte del personal de la institución o familiares de éstos, o que hayan concurrido por otro medio que no sea espontáneamente (difusión por parte de cualquier medio de comunicación o un tercero). Las muestras de los dos puntos de colecta de Formosa se unificaron a los efectos de este trabajo, al no existir diferencias significativas entre ellas en cuanto al origen continental de sus haplogrupos

³ PCR: Por sus siglas en inglés *Polymerase Chain Reaction* o reacción en cadena de la polimerasa.

⁴APLP: Por sus siglas en inglés *Amplified Product Length Polymorphism* o polimorfismos de longitud de los productos de amplificación.

mitocondriales, ni de las frecuencias de haplogrupos del cromosoma Y (Test Exacto de Diferenciación Poblacional: $P > 0,05$ para todas las comparaciones. Raymond & Rousset, 1995. Datos no mostrados).

2.3 Encuestas genealógicas

Una vez terminado el proceso de consentimiento informado (ver siguiente sección), se realizó una encuesta genealógica para completar la información sobre el lugar de nacimiento del donante y de sus padres y abuelos por línea materna y paterna (Anexo 2.1). Con esta información se confeccionó una base de datos utilizando un código jerárquico de tres letras para cada localidad informada. Se designó primero al país, luego a la provincia y finalmente al departamento o partido. Por ejemplo, para un donante nacido en Corrientes capital se colocó el siguiente código: "ARG CRR CAP, Corrientes". De esta forma, se agilizó la búsqueda del último ancestro conocido (UAC) por vía materna y/o paterna, siendo, en la mayoría de los casos, de la tercera generación. Para aquellos encuestados en los que este dato no fuera conocido o certero, se consideró el lugar de nacimiento de la generación parental (madre o padre) o del mismo donante. Se trabajó con dos grupos de categorías: "regional" y "extrarregional". Se consideró que el lugar de origen era regional cuando el UAC tuvo como lugar de nacimiento a las provincias que conforman la región del NEA (Corrientes, Formosa, Misiones y Chaco) o los países limítrofes al NEA (Paraguay, Uruguay y Brasil). A las localidades informadas de otras provincias de la Argentina, de otros países de Sudamérica o de otros países europeos, se las consideró dentro de la categoría extrarregional.

A partir del origen declarado en la encuesta genealógica se detectó que ciertos linajes no eran locales del sitio de muestreo. Sin embargo, se tomó la precaución de no eliminar ni redistribuir las muestras en concordancia al origen declarado, de forma tal de no interferir en la distribución actual de los linajes maternos y paternos, siguiendo el mismo criterio que en Motti et al. (2013).

2.4 Consideraciones éticas

2.4.1 Proceso de consentimiento informado, muestras biológicas y confidencialidad

Los donantes fueron previamente informados y expresaron su voluntad de participar a través de la firma de un formulario de consentimiento informado. Se entregó al potencial

donante una hoja de información en donde se detallaban los objetivos del estudio, las características genéticas a investigar, la metodología a utilizar, los riesgos y beneficios que se derivan de la investigación, la forma en que se resguardará la confidencialidad de los datos y la información de contacto de los investigadores (Anexo 2.2). Se explicó toda esta información en un lenguaje comprensible y se dio la oportunidad de hacer preguntas. A su vez, se aclaró que la participación era voluntaria, pudiendo retirar su muestra del estudio en cualquier momento solicitándolo al investigador. En los casos en los que el sujeto aceptara participar, firmó una copia del formulario de consentimiento informado (Anexo 2.3) en presencia de un testigo.

Las muestras biológicas fueron codificadas con un sistema de tres letras y tres números para designar la estación de muestreo de donde provenían y el número de muestra. De esta forma, se resguardó la confidencialidad de los participantes durante su procesamiento en el laboratorio.

Los datos personales y la información genealógica de los participantes fueron volcados en una base de datos en los que se consignó su código de muestra, de modo tal que las mismas fueran identificables a la hora de la devolución de los resultados a la población. El acceso a esta base de datos se encontró restringido a los directores de este proyecto y a la tesista, asumiendo el compromiso de guardar la máxima confidencialidad, en concordancia con lo que establece la Ley 25.326 de Protección de Datos Personales y los consensos internacionales en materia de confidencialidad.

2.4.2 Aprobaciones de los comités de ética

Este trabajo se encuentra dentro de un proyecto mayor de investigación que fue aprobado por el Comité Provincial de Bioética de la Provincia de Jujuy, el Comité Institucional de Ética e Investigación en Salud del Ministerio de Salud y Desarrollo de la Provincia de Santiago del Estero y por el Comité de Ética en Investigaciones Biomédicas del IMBICE. Se desarrolla conforme a los lineamientos de la Declaración de Helsinki (2013) y las normas éticas vigentes, así como los Marcos Éticos de Referencia del CONICET para Ciencias Sociales y Humanidades y para las Investigaciones Biomédicas (Anexo 2.4).

2.4.3 Devolución de resultados a la población participante

Se informó a cada población participante los resultados obtenidos en esta investigación. Los mismos, fueron enviados individualmente a cada donante en sobre cerrado. Se informó

sobre el origen continental remoto por línea materna o paterna (en el caso de los varones) y el haplogrupo o linaje de pertenencia (Anexo 2.5). Por otra parte, se informaron los resultados para la población en general de la localidad muestreada a través de un afiche ilustrativo colocado en la institución en donde se realizó la recolección de las muestras (Anexo 2.6). En ningún caso se obtuvieron o informaron resultados con posibles implicancias médicas para los sujetos participantes.

2.5 Procesamiento de muestras biológicas

2.5.1 Extracción de ADN

2.5.1.1 Procesamiento en el sitio de muestreo

Una vez obtenidas las muestras de sangre se realizó una primera etapa de procesamiento en el sitio de muestreo que consistió en la separación de los linfocitos del resto de los componentes sanguíneos. Para ello, de 5 a 10 mL de las muestras de sangre se centrifugaron por 20-40 minutos a 3000 revoluciones por minuto (rpm), obteniendo una capa intermedia que se separó mecánicamente con una pipeta Pasteur. Los linfocitos se suspendieron en 10 mL de *buffer* de lisis de glóbulos rojos (ver Anexo 2.7 con la composición de todas las soluciones de trabajo mencionadas en esta sección y en la siguiente) de forma tal de lavar todos aquellos eritrocitos que pudieran haber sido arrastrados con ellos durante la separación. Se precipitó la fracción de interés centrifugando nuevamente por 20-40 minutos a 3000 rpm (este paso conjuntamente con el de lavado se repitió en caso de ser necesario). Se descartó el sobrenadante (excepto 1 mL) y los linfocitos fueron suspendidos en 3 mL de *buffer* de lisis de glóbulos blancos. De esta forma, se conservaron las muestras en el campo a temperatura ambiente y luego fueron trasladadas al IMBICE para continuar con su procesamiento.

2.5.1.2 Procesamiento en el laboratorio del IMBICE

Una vez en el laboratorio del IMBICE se colocaron 20 μ L de detergente SDS 20% en cada uno de los tubos que contenían las muestras de glóbulos blancos para lisar las células, es decir, romper las estructuras que confinan el citoplasma y liberar al medio su contenido. Se completaron con *buffer* de lisis de glóbulos blancos hasta un volumen final de 7 mL y se colocaron en rotador por 24 horas hasta obtener un lisado homogéneo. De cada muestra se

extrajeron 2 mL del lisado de linfocitos para resguardo y se colocaron en tubos con tapa a rosca en freezer a -18 °C para conformar parte de la colección del IMBICE, en caso de que se requiriera una nueva extracción de ADN. Se agregaron 2,5 mL de solución saturada de cloruro de sodio (NaCl 5,4 M) para lograr la precipitación de proteínas, se mezcló por inversión y se incubó en hielo por 10 minutos. Se centrifugó a 21°C por 10-20 minutos a 4500 rpm obteniendo un precipitado de proteínas que fue descartado y el sobrenadante se trasvasó a otro tubo de 15 mL, en el que se precipitaron los ácidos nucleicos completando el volumen del tubo con alcohol isopropílico. Se centrifugó a 21 °C por 10-20 minutos a 4500 rpm, obteniendo un *pellet* de ADN. Se descartó el sobrenadante y se lavó el *pellet* con etanol 70%. El etanol fue descartado y se rehidrató el ADN en 0,5-1 mL de Tris-EDTA (TE). Los tubos se cubrieron con parafilm con un orificio en el centro y se colocaron en baño húmedo a 37 °C por 24 hs, para lograr la evaporación del etanol. Pasado este tiempo los tubos se taparon nuevamente y fueron centrifugados por 2 minutos a 2500 rpm. Su contenido fue trasvasado con pipeta automática a tubos con tapa a rosca, que se conservaron en freezer a -18 °C en la colección de ADN del IMBICE.

2.5.2 Tratamiento de muestras coaguladas

Las muestras de sangre coaguladas se procesaron siguiendo el protocolo *Clotspin*[®] *Baskets* y *Gentra*[®] *Puregene*[®] *Blood Kit* QIAGEN que se detalla en el Anexo 2.8.

2.6 Análisis de Censos Nacionales

Se analizaron los Censos Nacionales de 1914 y 2010 a fin de determinar el origen de los migrantes en las provincias estudiadas (Tercer Censo Nacional, 1914; INDEC, 2010). Pese a la existencia de Censos nacionales anteriores (de los años 1869 y 1895), la elección se realizó teniendo en cuenta que el territorio de Chaco y Formosa fue ocupado militarmente entre 1872-1884, luego de la guerra de la Triple Alianza. Por otra parte, no se incluyeron los resultados del Censo 2022 por tratarse de un relevamiento posterior al muestreo de este trabajo.

2.7 Primer nivel de análisis: migraciones intercontinentales

2.7.1 Tipificación de haplogrupos uniparentales

Tanto para el diagnóstico de los haplogrupos mitocondriales como del cromosoma Y se utilizó la técnica PCR-APLP múltiple, que consiste en amplificar segmentos de ADN que contienen los polimorfismos de interés en una única reacción de PCR múltiple, siguiendo la metodología utilizada por Motti et al. (2009) y modificada por Beltramo (datos no publicados) para el ADN mitocondrial y de Jurado Medina (2015) para el cromosoma Y. Estos autores se basaron en los trabajos de Umetsu y colaboradores, quienes desarrollaron una técnica de tipificación alelo-específica de PNS para ADN mitocondrial (Umetsu et al., 2001; Umetsu et al., 2005). Para investigar cada polimorfismo se utilizaron tríos de cebadores: dos de ellos son alelo-específicos y difieren en la base que tienen en su extremo 3', de manera que reconocen la secuencia para el carácter derivado y ancestral; otro de los cebadores varía en su extensión de nucleótidos (de 1 a 7) en su extremo 5', es decir, lleva una cola de nucleótidos no complementarios. De esta forma, el producto amplificado tiene una longitud variable en función de cuál de los dos cebadores estuvo involucrado en la reacción.

2.7.2 Linajes maternos

2.7.2.1 Asignación de origen continental

Una exploración inicial permitió clasificar las muestras estudiadas según su origen continental. En este sentido, se consideró un origen nativo para aquellas que fueron asignables a los haplogrupos A, B, C y D; un origen de Europa y/o Medio Oriente para aquellas que resultaron N(xA,B); y un origen africano para las muestras asignables al paragrupo L(xM,N). Estas últimas se confirmaron mediante secuenciación de la región control. A su vez, las muestras pertenecientes al linaje M que no fueran C o D también fueron secuenciadas, a fin de poder discriminar un origen europeo, norafricano o de Medio Oriente (haplogrupo M1) de otro asiático oriental (por ejemplo, haplogrupos M7, G, Z, etc.).

Para el tamizaje inicial, se utilizaron cinco tríos de cebadores de los que fueron publicados por Umetsu et al. (2005) y adaptados a nuestras poblaciones por Motti et al. (2009), los cuales permiten determinar a los haplogrupos A, D, M8 (dentro del que se encuentran los haplogrupos C y Z) y los macrohaplogrupos M y N. El sexto trío de cebadores para diagnosticar

al haplogrupo B -que se analizaba en forma independiente (Motti et al., 2009, 2012)- fue agregado posteriormente al múltiplex por Beltramo (datos no publicados). Se decidió modificar el protocolo empleado por Motti et al. (2009) para estos análisis en nuestro laboratorio, debido a que la delección 8281-8289 -que se utilizaba como diagnóstica del haplogrupo B- al ser homoplásica puede presentarse asociada a diferentes haplogrupos (véase, por ejemplo Sokolova et al., 2002; Tarskaia et al., 2006; Achilli et al., 2008). El nuevo protocolo realizado analiza la posición 827, diagnóstica del clado B4b'd'e'j al que pertenecen los linajes B2 americanos (Beltramo, manuscrito de tesis doctoral en preparación), confiriendo una mayor especificidad a la asignación de este haplogrupo. A su vez, el hecho de que todas las mutaciones diagnósticas se resuelvan realizando una única reacción de PCR múltiplex tiene ciertas ventajas. Se optimizan tiempo y recursos de experimentación y, por otra parte, disminuyen las probabilidades de cometer errores en la asignación de haplogrupos. Esto último, se debe a que para poder asignar una muestra al haplogrupo A o B debe también ser asignable al clado N, de forma similar, para que una muestra pertenezca al haplogrupo C o D debe poder asignarse al clado M (Figura 2.1). Además, esta técnica permite detectar clados M(xC,D) y aquellos que pertenecen al linaje L(xM,N).

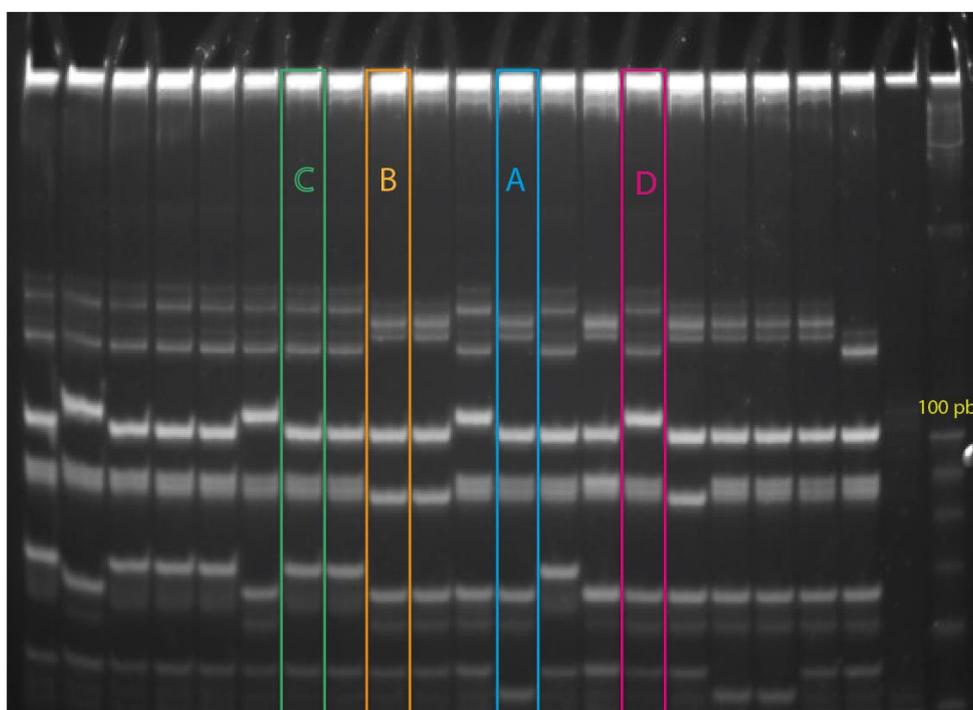


Figura 2.1 Patrón de bandas obtenidas para los haplogrupos de ADNmt de la reacción Múltiplex I. Las bandas derivadas para cada haplogrupo son de 149 pb (M), 136 pb (N), 107 pb (D), 85 pb (B), 70 pb (M8) y 52 pb (A). Las calles coloreadas muestran, en ese orden, linajes

asignables a los clados C, B, A y D, respectivamente. La última calle corresponde a un marcador de peso molecular y se indica la banda correspondiente a 100 pb.

Se detallan a continuación las secuencias de los cebadores utilizados pertenecientes al Múltiplex I (Tabla 2.1), indicando con letra minúscula las bases que corresponden a la región de no apareamiento.

Tabla 2.1 Cebadores del Múltiplex I utilizados para la caracterización por APLP de linajes maternos.

La letra F indica el cebador *forward* y R el *reverse*. Las minúsculas en las secuencias indican *mismatches* con la rCRS, a excepción de la posición 3' más extrema de los alelos non-rCRS. Conc. Indica la concentración. Los amplicones derivados se señalan con el nombre del clado entre paréntesis.

Haplogrupo	Cebador	Secuencia (5'-3')	Conc. (pmol/10 µL)	Tamaño amplicón (pb)
M	F-10400*T	taattaTACAAAAAGGATTAGACTGtgCT	4,7	149 (M)
	F-10400*C	TACAAAAAGGATTAGACaGAACC	4,7	142
	R-10400	GAAGTGAGATGGTAAATGCTAG	4,7	
N	F-10873*T	aatataCCTAATTATTAGCATCAaCCCT	5,1	136 (N)
	F-10873*C	CCTAATTATTAGCATCATtCCC	6,4	131
	R-10873	tGGTAtGAGTCAGGTAGTTAG	6,4	
D	F-5178*A	tgatcaaCGCACCTGAAACAAGA	4,4	107 (D)
	F-5178*C	gTCGCACCTGAAgCAAGC	3,2	102
	R-5178	CCCATTGgGCAAAAAGCC	3,2	
M8 (C)	R-15487*T	gtccttGGTCGCCTAGGAaGTCA	2,5	70 (M8)
	R-15487*A	tGGTCGCCTAGGAGtTCT	4,5	66
	F-15487	CTTACTTCTCTCCTTCTCTCC	2,5	
A	F-1736*A	taaAGCCAAACCATTTtCCCAA	2,8	55
	F-1736*G	AGCCAAACCATTTAtCCAG	2,8	52 (A)
	R-1736	aCCAGGTTTCAATTTCTATCG	2,8	
B	F-827*G	CGGGAAACAGCAGTGATTAG	1,75	85 (B)
	F-827*A	ttaACGGGAAACAGCAGTGAATA	1,75	89
	R-827	TGGCACGAAATTGACCAAC	1,75	

2.7.2.1.1 Condiciones de amplificación

La reacción de amplificación se llevó a cabo en un volumen final de 10 µL, con las siguientes condiciones: 10 ng de ADN genómico, 0,25 U/µL de Go[®] Taq DNA Polymerase (Promega), 5X *buffer* de reacción, cebadores en concentración óptima (Tabla 2.1), 100µM dNTPs

y agua ultrapura, para completar el volumen final. El ciclado se realizó en termociclador Biometra T3000, en las siguientes condiciones de amplificación: 3 minutos de desnaturalización inicial a 94°C, 33 ciclos de desnaturalización a 94°C por 1 minuto, hibridación a 55°C por 1 minuto y extensión a 72°C durante 3 minutos, extensión final de 10 minutos. Estas condiciones se obtuvieron a partir del método de prueba y error, ajustando las concentraciones de los cebadores para lograr que se amplifiquen todas las regiones de interés en forma simultánea en una única reacción de PCR. Estos resultados se resolvieron en geles de poliacrilamida al 10 % mediante electroforesis. Se sembraron 4 µL del producto de amplificación, que se corrieron durante 3,5 horas a 160 voltios conjuntamente con el/los marcador/es de peso molecular correspondiente/s. Posteriormente, los geles fueron teñidos en solución de TBE 1x con Gel Red®.

2.7.2.2 Asignación de haplogrupos del macro-haplogrupo N

Las muestras que resultaron pertenecientes al clado N(xA,B) fueron analizadas con ocho tríos de cebadores capaces de reconocer variantes específicas que permitieron asignar estas muestras a los haplogrupos H, I/X, J, K, T, U, V y W (Umetsu et al., 2005) (Tabla 2.2). De esta forma, se discriminó con mayor resolución dentro de haplogrupos de origen europeo y de Medio Oriente. Cabe aclarar que el trío utilizado para la asignación de I/X analizan la posición 1719, que conforma tanto la definición del clado N1'5 (al que pertenece el haplogrupo I) como del clado X2 (Umetsu et al., 2005), por lo que no se puede considerar a I/X como un grupo natural. Los ensayos de PCR-APLP fueron realizados en forma secuencial en el siguiente orden: primero se analizaron las muestras con el Múltiplex II (condiciones puestas a punto por Schwab, 2018, Tabla 2.2), luego, aquellas que resultaron no asignables a ningún clado de este experimento se ensayaron mediante los mónoplex I, II y III (Tabla 2.2). Los productos de amplificación se visualizaron en geles de poliacrilamida al 10 % mediante electroforesis.

Tabla 2.2 Cebadores y condiciones empleados para analizar linajes pertenecientes a N(xA,B).

Abreviaciones se detallan en la leyenda de la Tabla 2.1

Reacción	Haplogrupo	Cebador	Secuencia (5'-3')	Conc. pmol/10 µL	Tamaño amplificón (pb)
Múltiplex II	J	F-12612*A	taaCTCCATAATATTCAaCCCTGTA	2,5	53
		F-12612*G	CTCCATAATATTCAATtCCTGTG	2,5	50 (J)
		R-12612	CTATGATGGACCATGTAACG	2,5	
	H	F-7028*T	ttCGACACGTACTACGTaGTAGCT	1,25	59
		F-7028*C	GACACGTACTACGaGTAGCC	1,25	56 (H)
		R-7028	AATACAGCTCCTATTGATAGGAC	1,25	
	T	F-15607*A	tagCGATCCGTCCCTtACAAA	3,5	66
		F-15607*G	CGATCCGTCCCTAtCAAG	3,5	63 (T)
		R-15607	aGCTAGGATGAGGATGGATAG	3,5	
	U	F-12308*A	atatCATTGGTCTTAGaCCCCAAA	5,5	84
		F-12308*G	CATTGGTCTTAGGaCCCAAG	5,5	80 (U)
		R-12308	aTTAGGGTGGTTATAGTAGTGTGCA	5,5	
	K	F-10550*A	ttataTATATCGCTCACACCaCATA	1,25	95
		F-10550*G	TATATCGCTCACTCCTCATG	1,25	90 (K)
		R-10550	TGTTGAGGGTTATGAGAGTAGCT	1,25	
Mónoplex I	I/X	R-1719A	aatTTTGGGTAAATGGTTaGGT	2,5	74 (I/X)
		R-1719G	TTGGGTAAATGGaTTGGC	2,5	70
		F-1719	CTGAGCTAAACCTAGCCCC	2,5	
Mónoplex II	V	F-4580	TACCATCTTTCAGGCAC	2,5	
		R-4580*A	aaaGGTTAGAACTGGAATAAAAAGC aAGT	2,5	106 (V)
		R-4580*G	TTAGAACTGGAATAAAAtGCTAGC	2,5	101
Mónoplex III	W	F-8994*A	ttataTCATTCAACCAATtGCCCTA	2,5	121 (W)
		F-8994*G	TCATTCAACCAATAGtCCTG	2,5	116
		R-8994	AGAGGGAAGGTTAATGGTTG	2,5	

2.7.3 Linajes paternos

2.7.3.1 Tipificación de haplogrupos del cromosoma Y por PCR-APLP

Los cebadores para diagnosticar haplogrupos paternos se diseñaron para el reconocimiento preferencial de una variante alélica en función de las mutaciones características de los haplogrupos alóctonos de mayor frecuencia encontrados en Argentina y para los linajes propios de nativos americanos (Jurado Medina, 2015) siguiendo la técnica de Umetsu et al. (2005) y basados en la filogenia del cromosoma Y (Karafet et al., 2008; ISOGG, 2017). Se

utilizaron 18 PNS diagnósticos de los clados de interés, organizados en 4 múltiplex, según el protocolo desarrollado por Jurado Medina (2015). El múltiplex I, compuesto por 6 tríos de cebadores, comprende los marcadores M173, M242, M207, M172, M89 y M9, los cuales nos permiten discernir en qué rama basal dentro de la filogenia del cromosoma Y nos encontramos y, a partir de allí, determinar qué otro múltiplex debemos aplicar en nuestro análisis. En este sentido, el múltiplex II permite diagnosticar entre tres mutaciones dentro del clado Q (M242, M346 y M3) y se aplicó cuando el individuo presentó un alelo derivado para M242. Por otra parte, el múltiplex III diagnostica dentro del súperhaplogrupo F los distintos haplogrupos (P287, M253, M267 y M285) y se analizó en las muestras ancestrales para M9 (FxKLT). El múltiplex IV resuelve los subhaplogrupos del clado R1 (M343, P297, SRY10831.2 y M335) y se aplicó en las muestras derivadas para M173. En contraposición, el haplogrupo E (P152) se realizó en forma directa por PCR alelo-específico (Mónoplex I), sólo en aquellos individuos que resultaron de carácter ancestral para la mutación de M89. En la Tabla 2.3 se presentan las secuencias de los cebadores utilizados.

Las reacciones de PCR se realizaron utilizando un volumen final de 10 μ L, con las siguientes condiciones: 10ng de ADN genómico, 0,025 U/ μ L de Platinum® Taq DNA Polymerase (Invitrogen), 0,75 mM MgCl₂, 1x *buffer* de reacción, cebadores en concentración óptima (Tabla 2.3), 100 μ M dNTPs y agua ultrapura, para completar el volumen final. La reacción se llevó a cabo en un termociclador Biometra T3000. Después de 4 minutos de desnaturalización a 94°C, se realizaron 33 ciclos de amplificación, cada uno compuesto por: 1 min 94°C, 1 min 55°C y 1,5 min 72°C, seguido por un último paso de extensión final de 72°C por 10 min. Los resultados se resolvieron en geles de poliacrilamida al 10 % mediante electroforesis. Se sembraron 3 μ L del producto de amplificación con 3 μ L de *loading buffer*, que se corrieron durante 4,5 horas a 160 voltios conjuntamente con el/los marcador/es de peso molecular correspondiente/s. Posteriormente, los geles fueron teñidos en solución de TBE 1x con Gel Red®.

Tabla 2.3. Cebadores utilizados para la caracterización por APLP de linajes paternos.

La letra F indica el cebador *forward* y R el *reverse*. Las minúsculas en las secuencias indican *mismatches* con la rCRS. Conc. Indica la concentración. Los amplicones derivados se señalan en **negrita**.

Reacción	Haplogrupo	Cebador	Secuencia (5'-3')	Conc. pmol/ 10 µL	Tamaño amplicón (pb)
Múltiplex I	R1	F-M173*C	cCAAGGGCATTgAGAACC	3,75	150
		F-M173*A	taatgtatTCAAGGGCATaTAGAACA	3,75	158
		R-M173	CAGCCTTCAAAGCTTCTCCT	3,75	
	Q	F-M242*C	AAAGGTGACCAAGGTGCTC	2,81	131
		F-M242*T	AacaatAAAGGTtACCAAGGTGCTT	3,75	137
		R-M242	TCAGCATAATACCTTACCTAGAACAAC	2,81	
	R	F-M207*G	CgAGTCAAGCAAGcAATTTAG	5	124
		F-M207*A	gttcacTAAGTCAAGCAAGAAATcTAA	10	118
		R-M207	AAAAGCTGAAGGAAAAGTGGA	5	
	J2	F-M172*G	AAACCCATTTgGATGCTTG	6,25	96
		F-M172*T	ctaatAAACCCATTgTGATGCTTT	7,5	101
		R-M172	TTAAATATCAGCCAGGTACAGAGA	6,25	
	M89	R-M89*A	CgCAGGCAAAGTGAGAGATA	2	85
		R-M89*G	ttaaTCAGGCAAAGTGAGAGATG	2	89
		F-M89	CACAGAAGGATGCTGCTCA	2	
	M9	F-M9*G	GGCCTAAGATGGTTGgATG	3,75	70
		F-M9*CC	tttgGacCTAAGATGGTTGAATC	3,75	74
		R-M9	TGTAAGACATTGAACGTTTGgA	3,75	
Múltiplex II	Q*	F-M242*C	AAAGGTGACCAAGGTGCTC	3,75	131
		F-M242*T	aacaatAAAGGTtACCAAGGTGCTT	3,75	137
		R-M242	TCAGCATAATACCTTACCTAGAACAAC	3,75	
	Q1b1a1a	R-M3*T	cGTACCAGCTCTCCcAATTA	2,5	115
		R-M3*C	attttaaGTACCAGCTCTCCTAATTG	2,5	121
		F-M3	AAGGTACATTCGCGGGATA	2,5	
	Q1b	R-M346*G	cGCAATTTACAACgACAAGC	5	160
		R-M346*C	aataatgtGCAATTTACAAAGACAAGG	5	167
		F-M346	AAGGCCTGAAAATGTGGAAA	5	

Reacción	Haplogrupo	Cebador	Secuencia (5'-3')	Conc. pmol/10 µL	Tamaño amplicón (pb)
Múltiplex III	G2	F-P287*G	TGACTGACATGGCgTTGG	10	149
		F-P287*T	aaaatatatTGACTGACATtGCCTTGT	2,5	158
		R-P287	AAGGGTAATGAAGCCTATAGTCCA	3,75	
	I1	R-M253*A	TCAATGAAGAACCTGGAGAAA	6,25	132
		R-M253*G	ttattaTCAATGAAGAACCTGGAGAAAG	2,5	138
		F-M253	GCTGATCTGTTTTCTTTTTGGTG	2,5	
	J1	F-M267*G	ggTGAAGCATTtTgTAAATAG	3,75	117
		F-M267*T	taagcaTGAAGCATTtTGTAAATAT	3,75	121
		R-M267	CGTAGCTcCAAGCAATCCAC	3,75	
	G1	F-M285*G	CcCATCATCTACATTTCTCCTG	2,5	80
		F-M285*C	gtacaCTCATCTACATTTCTCCTC	2,5	85
		R-M285	ATCGAATCCGCTATCCAGACT	2,5	
Múltiplex IV	R1b	F-M343*G	GAGTGCCaTCGTGTTCCAC	2,5	140
		F-M343*A	attaaattAGTgaCCTCGTGTCCAA	2,5	147
		R-M343	ATATGCAAATGCCAGCGTTA	2,5	
	R1b1a1a	R-P297*G	GCCAAAGATGTATAAGcGGAAG	2,5	104
		R-P297*C	tctaaGCCAAAGATGTATAAGAGGAAC	2,5	109
		F-P297	TGGGAGAGAGAGACAGGTG	2,5	
	R1a1	F-SRY10831.2*C	AggTgAACCTTgAAAATgCTAC	2,5	76
		F-SRY10831.2*T	TCgATggTgAACCTTgTAAATgTTAT	2,5	72
		R-SRY10831.2	CATgggATCATTcAgTATCTgg	2,5	
	R1b1b1	R-M335*A	ATGCCCTTGAATTGTAAGcAAA	4.5	66
		R-M335*T	gattATGCCCTTGAATTGTAAGAAAT	4.5	62
		F-M335	cCTGTAACCTCCTAGAAAATTGGAAA	4.5	
Mónoplex I	E	F-P152*G	GCcCATCCGTTTTCTGAAG	3,75	141
		F-P152*C	ataaatGCTCATCCGTTTTCTGAAC	3,75	147
		R-152	TCTCaGGGTTTGTGCAGATT	3,75	

Como se mencionó anteriormente, las reacciones de PCR en múltiplex permiten realizar un control en la asignación de haplogrupos con el resultado de cada muestra. Por ejemplo, si una muestra presentara el carácter derivado para el marcador M207 (haplogrupo R) –luego de realizar el múltiplex I-, deberá presentar el carácter ancestral para M242 (haplogrupo Q) y derivado para M9 y M89. Esto resulta de gran utilidad a la hora de reducir los errores de asignación.

2.7.3.2 Tipificación del clado I2 por HRM

Aquellas muestras que resultaron ancestrales para los marcadores del múltiplex III se consideraron como posibles candidatos del clado I2. La tipificación de I2 se resolvió aplicando un protocolo desarrollado por Sala (datos no publicados) por *High Resolution Melting* (HRM), que consistió en realizar reacciones de PCR en tiempo real de 20 μ L de volumen final, conteniendo 10 μ L de “Mezcla Real 2X” marca Biodynamics (*Taq DNA Polimerasa*, dNTPs, Mg⁺⁺, conservantes, *buffer* y fluorocromo verde), 40 μ M de cebadores, 2 μ L de templado y agua libre de RNAsas y DNAsas hasta completar el volumen final (Tabla 2.4). Se utilizó un termociclador Rotor-Gene Q 5-plex HRM QIAGEN. El ciclado comprende un paso inicial de desnaturalización de 2 minutos de 95 °C, 40 ciclos de 20 segundos a 95 °C y de 20 segundos a 58 °C, 30 segundos a 72 °C y un ciclo de *melting* de 72 a 95 °C. Los datos obtenidos se analizaron a través del *software* integrado al equipo, a partir del cual se realizó la asignación de las muestras al genotipo ancestral o derivado. Esto último se llevó a cabo a partir de la comparación de los resultados con dos muestras -previamente secuenciadas- que fueron utilizadas como control.

Tabla 2.4. Secuencias de cebadores para la asignación al clado I2

Clado	Cebador	Secuencia (5´-3´)	Concentración (pmol/20 μ L)
I2	P215F-RT	GCCTGGAATGTAGACTAATGGT	40
	P215-R	AGCTTCCACAAATTACTACACACA	40

2.7.3.3 Tipificación por PCR-RFLP de muestras con posible origen africano

Se consideró a las muestras que resultaron derivadas para el haplogrupo E como posibles candidatas a pertenecer al clado africano E1b1a1. Se verificó la asignación a dicho clado mediante PCR-RFLP, utilizando el par de cebadores PN1 desarrollados por Bravi et al. (2000) (Tabla 2.5), que generan un fragmento de 218 pb. Se realizó una digestión enzimática con *AluI* (*New England Biolabs*) en un volumen final de 16 μ L por muestra, compuesta por 4 μ L de producto de PCR, 10 U de enzima, 1,6 μ L de Buffer 2 (*New England Biolabs*) y agua ultrapura hasta completar el volumen final. Se incubó durante 12 horas a 37 °C. El producto de amplificación (5 μ L) se sembró en un gel de agarosa al 2 % para comprobar el éxito de la reacción. Se consideró a las muestras derivadas para E1b1a1 a aquellas en las que la digestión enzimática resultó en dos fragmentos de 145 y 73 pb, mientras que en las ancestrales se observa una única banda de 218 pb.

Tabla 2.5. Secuencias de cebadores para PN1

Clado	Cebador	Secuencia (5´-3´)	Concentración (pmol/10 µL)	Tamaño amplicón
E1b1a1	PN1-F	CATAATTCCATTTTCCCTATTGC	2,5	218 pb
	PN1-B	GTCCTCTCCTTATTAACGTAAC	2,5	

2.7.4 Comparación de datos moleculares y genealógicos

Una vez obtenidos los datos moleculares y genealógicos del UAC de cada donante, se compararon los resultados. Sólo se pudo comparar la proporción de ancestros por vía materna que tuvieron un origen ultramarino con aquellos que pertenecen al superhaplogrupo N. Para la vía paterna, sólo se compararon los ancestros con un origen regional declarado (Argentina + países sudamericanos), con aquellos que pertenecen al clado Q-M3, mientras que el resto de las categorías utilizadas se agrupó como “ultramar”. Las demás categorías no pudieron ser homologadas con datos moleculares, y en ningún caso hubo ancestros declarados de origen africano.

2.8 Segundo nivel de análisis: vínculos interregionales

2.8.1 Definición de subclados mitocondriales

2.8.1.1 Nota sobre nomenclatura de clados mitocondriales

A lo largo de esta tesis se empleó la nomenclatura de clados mitocondriales según está establecida en la última versión de Phylotree (*mtDNA tree Build 17*, 18 Feb 2016, <http://phylotree.org/>, van Oven & Kayser, 2009), así como las propuestas posteriores realizadas en base a clados locales por Gómez-Carballa et al. (2018) para poblaciones andinas, Scheib et al. (2018) y Lindo et al. (2017) para poblaciones de Norteamérica, Brandini et al. (2018) para ecuatorianos, Schötig et al. (2015) para guatemaltecos, García et al. (2021), Motti et al. (2015), Bravi et al. (2022), Aquilano et al. (2023) y Figueiro et al. (2022) para sudamericanos.

Un análisis filogenético preliminar de los mitogenomas disponibles en la base de datos de Claudio M. Bravi (datos no mostrados) llevó a la definición de numerosos nuevos clados de presencia local en poblaciones del NEA y/o el este de Sudamérica. Para evitar confusiones con cualquier nomenclatura que fuera publicada mientras se escribe esta tesis, se decidió designar los nuevos subhaplogrupos con una nomenclatura que discontinúa el orden alfanumérico

establecido. Así, los nuevos clados de A2 comprenden desde A2za hasta A2zn y de A2ya hasta A2yc; las variantes de B2b van desde B2b50 a B2b53, y las de B2 de B2za a B2zg (Anexo 2.9). Las nuevas propuestas de C1 incluyen C1b100-C1b103, C1c100 y C1c101, y C1d1aa-C1d1ac; las de D1 son D1za-D1zh mientras que las de D4h3a abarcan de D4h3a90 a D4h3a92 (Anexo 2.9).

2.8.1.2 Secuenciación de la región control completa extendida

Se obtuvieron secuencias de la región control (RC) completa extendida en sentido 5' de 169 muestras de Formosa, seleccionadas en base a un origen local declarado para la línea materna. La RC y la porción 3' del ARNt-Thr fueron amplificadas en un fragmento único de 1340 pb con los cebadores F15792 y R698 (Tabla 2.6), siguiendo las condiciones descritas por Brandstätter et al. (2005). Todas las muestras fueron leídas en sentido *forward* con al menos los *primers* F15792 y F16475. Cuando hubo pérdida de calidad, como por ejemplo 3' de las posiciones 16193 y 315 por heteroplasmías de longitud, se emplearon los *primers* R186, R285, F314 y R484 (Tabla 2.6). Los amplicones fueron secuenciados en Macrogen (Corea del Sur). La obtención, interpretación, análisis y nomenclatura se realizó siguiendo las recomendaciones para tipificación del ADNmt de la *International Society for Forensic Genetics* (Carracedo et al., 2000; Parson et al., 2014). El análisis preliminar de las secuencias permitió asignar una fracción a clados reconocibles por PNS de la región control extendida como A2aa, A2ah, B2h, B2i1, B2b3a, C1b, C1b2, C1d*, C1d1b, C1c*, D1f1, y D1j (Tabla 2.7).

Tabla 2.6. Cebadores empleados para la amplificación y secuenciación de la región control

F15792	TCATTGGACAAGTAGCATCC
F16475	TAG CTA AAG TGA ACT GTA TCC
R186	GCCTGTAATATTGAACGTAGGTG
R285	GTTATGATGTCTGTGTGGAA
F314	CCGCTTCTGGCCACAGCACT
R484	TGAGATTAGTAGTATGGGAG
R698	GCATGTGTAATCTTACTAAGAG

2.8.1.3 Asignación de subhaplogrupos a partir de PNS de región codificante

En aquellos casos en los que no fue posible asignar las muestras a clados conocidos por carecer de PNS diagnósticos en la región control, se utilizó una batería de ensayos alelo-específicos de PNS de región codificante siguiendo la metodología de Umetsu et al. (2001),

Umetsu et al. (2005) y de Jurado Medina (2015). Estos ensayos permitieron asignar o confirmar membresía en clados de distribución étnico/geográfica restringida y cuya presencia en la región es conocida o esperada, como A2aa, B2b, B2h, B2i, B2+6056, C1b2, C1d1b, C1d1b1, D1a y D1e (Tabla 2.7).

Ciertos cebadores utilizados para diferenciar algunos de estos clados de interés ya habían sido diseñados por la Dra. Jurado Medina y puestos a punto por Schwab (2018) en poblaciones de Santiago de Estero y Tucumán en comparación con otras poblaciones del NOA. Además, dada la particularidad de clados encontrados en las muestras de Formosa y en otras poblaciones del NEA y Centro de Argentina, se diseñaron cebadores específicos para el desarrollo de esta tesis en colaboración con la Lic. Eliana Aquilano, a fin de poder interrogar sobre estas variantes y afinar aún más el conocimiento sobre la filogenia de estas poblaciones.

Tabla 2.7. Clados asignados por PNS de la región control extendida y de la región codificante (en negrita). Las variantes son transiciones respecto a rCRS salvo que se especifique lo contrario.

Haplogrupo	Polimorfismos diagnósticos
A2aa	A2 + 16239A + 16266
A2ah	A2 + 16097 + 16098
B2b	B2 + 6755
B2b3a	B2 + 152 + 271+ 16249 + 16312 + 6755
B2h	B2 + 16468 + 11821
B2i1	B2 + 430 + 485 +16311 + 6272
C1b	C1 + 493
C1b2	C1 + @263 + 493 + 12454
C1d	C1 + 194 + 16051
C1d1b	C1+ 194 + 195 + 16051 + 928
C1d1b1	C1+ 194 + 195 + 16051 + 928 + 7337
C1c	C1 + 15930
D1a	D1 + 5821
D1e	D1 + 146 + 3316
D1f1	D1 + 16142 + 16179 + 16295 + 16497
D1j	D1 + 152 + 16242 + 16311

2.8.1.3.1 Ensayos con cebadores previamente explorados

En relación a los ensayos con cebadores previamente explorados (Schwab, 2018), se realizaron ensayos con los múltiplex III-VI y dos ensayos monoplex designados IV y V (Tabla 2.8), en muestras candidatas a presentar estos polimorfismos tanto de Formosa como de otras

poblaciones del NEA. En particular, el múltiplex III indaga sobre cuatro polimorfismos del haplogrupo A2: 9039, 9899, 5147 y 3504 (Tabla 2.8), los múltiplex IV y el V interrogan acerca de tres polimorfismos cada uno del haplogrupo B2: 14470, 15671 y 6755; 11590, 8853 y 7055, respectivamente (Tabla 2.9), mientras que el VI interroga sobre dos variantes que permiten distinguir entre los clados C1d1b y C1d1b1 (Tabla 2.10). Por otra parte, los mónoplex IV y V indagan sobre otras variantes de A2 (7202) y B2 (6056), respectivamente.

Tabla 2.8 Cebadores y condiciones empleados para identificar clados regionales de A2.

Abreviaciones se detallan en la leyenda de la Tabla 2.3.

Reacción	Haplogrupo	Cebador	Secuencia (5´-3´)	Conc. pmol/10 µL	Tamaño amplicón (pb)
Múltiplex III	A2bc (A2+150- 9039)	F-9039*G	TACTGCAGGCCACCTACTaATG	2,5	74
		F-9039*A	cgtaTACTGCAGtCCACCTACTCATA	2,5	78
		R-9039	TAGAGGGAAGGTTAATGGTTGATA	2,5	
	A2az (A2+9899)	F-9899*C	ccTTCACCTTTACATCCcAACAC	3	85
		F-9899*T	gcttataTTCACCTTTACATCCcAACAT	3	90
		R-9899	CAGAAATAGTCAAACCACATCTACAAA	3	
	A2ax (A2+5147)	F-5147*G	CAACTTAAACTCCAGCAcCACG	2,5	130
		F-5147*A	aataaCAACTTAAACTCCAGCAaCACA	2,5	135
		R-5147	GGGC AAAAAGCCGGTTAG	2,5	
	A2bb (A2+150- 3504)	F-3504	CCGAACGAAAAATTCTAGGC	2	150
		R-3504*G	cGATGTAGAGGGTGATGGTG	2	158
		R-3504*A	gtaataatGATGTAGAGGGTGATGGTA	2	
Mónoplex IV	A2- 3834/7202	F-7202*G	CCACAACACTTTCTCGaCCTG	2,5	66
		F-7202*A	ataaCCACAACACTTTCTCGaCCTA	2,5	70
		R-7202	GGTGTATGCATCGGGGTAGT	2,5	

Tabla 2.9 Cebadores y condiciones empleados para identificar clados regionales de B2. Abreviaciones se detallan en la leyenda de la Tabla 2.3.

Reacción	Haplogrupo	Cebador	Secuencia (5'-3')	Conc. pmol/10 μ L	Tamaño amplicón (pb)
Múltiplex IV	B2-14470	R-14470*G	GGAATGATGGTTGTCTTaGGG	2,5	83
		R-14470*A	attaaGAATGATGGTTGTaTTTGGA	2,5	88
		F-14470	AAGACCTCAACCCCTGACC	2,5	
	B2ag1 (B2+15671)	F-15671*G	CAATAATCCCCATCCTCaATG	2,5	96
		F-15671*A	aatcgAATAATCCCCATCCTaCATA	2,5	100
		R-15671	TGAGGAGGTCTGCGGCTA	2,5	
	B2b	F-6755*G	cATCAATTGGCTTCgTAGGG	2,5	152
		F-6755*A	aaatactaATCAATTGGCTTCCTAGGA	2,5	159
		R-6755	GTGGAGTGTGGCGAGTCA	2,5	
Múltiplex V	B2ak (B2+11590)	F-11590*A	cATAACAAGCTCCATCTGCCTA	2,5	88
		F-11590*G	aatttTATAACAAGCTCCATCTGtCTG	2,5	93
		R-11590	TGTTACTACGAGGGCTATGTGG	2,5	
	B2aj (B2+8853)	F-8853*G	CATGGCCATCCaCTTATGG	2,5	110
		F-8853*A	aattaaCATGGCCATaCCCTTATGA	2,5	116
		R-8853	TGGGGATAAGGGGTGTAGGT	2,5	
	B2-7055	R-7055*C	CCTATGATGGCcAATACAGCC	2,5	127
		R-7055*T	ataaatgCTATGATGGCAAATACTgCT	2,5	133
		F-7055	CCGTAGGTGGCCTGACTG	2,5	
Mónoplex V	B2z (B2+6056)	F-6056*C	tGGCAACCTTCTAGGTAACGAC	2,5	74
		F-6056*T	aaatAGGCAACCTTCTAGGTAACGAT	2,5	78
		R-6056	TGAAGAAGATTATTACAAATGCATGG	2,5	

Tabla 2.10 Cebadores y condiciones empleados para identificar clados regionales de C1d. Las abreviaciones se detallan en la leyenda de la Tabla 2.3.

Reacción	Hg.	Cebador	Secuencia (5'-3')	Conc. pmol/10 μ L	Tamaño amplicón (pb)
Múltiplex VI	C1d1b	F-928*G	ACGATTAACCCAAGTCAtTAGg	2,5	113
		F-928*A	ataagACGATTAACCCAAGTCAtTAGa	2,5	118
		R-928	TCTATTTTGTGTCAACTGGAGTTTTT	2,5	
	C1d1b1	F-7337*G	TTGAGAAGCCTTCgTTCg	2,5	123
		F-7337*A	gtactaTTGAGAAGCCTTCGaTTCa	2,5	129
		R-7337	GATTTTATGTATACGGTTCTTCG	2,5	

2.8.1.3.2 Ensayos con cebadores diseñados para este trabajo de tesis

En el contexto de la información disponible hasta el momento en bases de datos y en publicaciones (aproximadamente 32.000 secuencias parciales o completas de la región control y 7.000 mitogenomas) y en la colección inédita de 200 mitogenomas de poblaciones argentinas del Laboratorio Genética Molecular Poblacional (IMBICE), se seleccionaron 9 PNS de la región codificante del ADN mitocondrial que definen 8 subhaplogrupos D1, de presencia confirmada o supuesta en poblaciones del NEA y de países limítrofes. Dos PNS definen D1a y D1e, presentes en poblaciones amazónicas, guaraníes y paraguayas. Otro PNS (998) define D1v, con distribución en centro, centro-oeste y noroeste de Argentina. Por último, la presencia de 11590 define el recientemente descrito D1x (Figueiro et al., 2022). Por otra parte, tres PNS definen dos clados aún innominados de presencia confirmada en Litoral/Tartagal/Brasil (6305T-13989), y en Chaco/Brasil/población wichí (9424). Se diseñaron 8 tríos de cebadores que permiten la amplificación alelo-específica, siguiendo la metodología de Umetsu et al. (2001), Umetsu et al. (2005) y de Jurado Medina (2015). Se realizaron ensayos de PCR en gradiente de temperatura (54-64 °C) a fin de obtener las condiciones óptimas de amplificación y la confirmación del correcto funcionamiento de los cebadores. De los 8 tríos diseñados, 6 pudieron ponerse a punto (Tabla 2.11) y dos fallaron en la especificidad de amplificación y fueron descartados (Tabla 2.12). El diseño de cebadores se realizó con la expectativa de poder combinar varias reacciones en múltiplex, sin embargo, en los ensayos de puesta a punto se obtuvieron mejores resultados de amplificación en ensayos de tipo mónoplex.

Tabla 2.11 Cebadores y condiciones para identificar clados regionales de D1.

Abreviaciones se detallan en la leyenda de la Tabla 2.3.

Reacción	Hg.	Cebador	Secuencia (5'-3')	Conc. pmol/10 μL	Tamaño amplicón (pb)
Monoplex VI	D1v	R-998*T	AGTCTATgTTGTGTCAGCTGGAGT	2,5	76
		R-998*C	atattaAGTCTATTTTGTGTCAACTGGAGC	2,5	82
		F-998	TTTTAGATCACCCCTCCCA	2,5	
Monoplex VII	D1a	F-5821*G	CAATcCAATcTGAAAATCACCTCG	2,5	122
		F-5821*A	taataaCAATaCAATATGgAAATCACCTCA	2,5	128
		R-5821	AACaGTCaGCGAACATCAGT	2,5	
Monoplex VIII	D1x	F-11150*G	CACACTTATCCGCACCTTGG	2,5	78
		F-11150*A	aatataCACACTTATCCGCACCTTGA	2,5	84
		R-11150	AGGAAGTATGTGCCTGCGTT	2,5	
Monoplex IX	D1za	R-6305*C	CTCCAGcGTGcGAGTAGTTC	2,5	107
		R-6305*A	aatatCTCCAGcGTGcGAGTAGTTA	2,5	112
		F-6305	ACCTCCCTCTCTCCTACTCCT	2,5	
Monoplex X	D1za	F-13989*C	GCCAtAACCTGCCgCTACTC	2,5	93
		F-13989*T	atatacGCCAtAACCTGCctCTACTT	2,5	99
		R-13989	TGATGGAGGTGGAGATTTGGTG	2,5	
Monoplex XI	D1	R-3591*C	AGGCaTAGGTTGAGGTTGAgC	2,5	118
		R-3591*T	tatttAGGCaTAGGTTGAGGTTGACT	2,5	123
		F-3591	CCGCCACATCTACCATCACC	2,5	

Tabla 2.12 Cebadores con resultados fallidos.

Abreviaciones se detallan en la leyenda de la Tabla 2.3.

Haplogrupo	Cebador	Secuencia (5'-3')	Conc. pmol/10 μl	Tamaño amplicón (pb)
D1e-3316	F-3316*G	TCTTCTTAACAgCATACCgATGG	2,5	126
	F-3316*A	ataataTCTTCTTAACAgCATACCgATGA	2,5	132
	R-3316	TTGGGGCCTTTGCGTAGTT	2,5	
D1-9424	F-9424*C	TACCAAGGCCAtCACACAgC	2,5	104
	F-9424*T	tattatTACCAAGGCCACCACAtAgT	2,5	110
	R-9424	AAGGCTCAGAAAATCCTGCGA	2,5	

2.8.1.3.3 Condiciones de amplificación

El resultado de amplificación se corrió en gel de poliacrilamida al 10% por electroforesis, los patrones de bandas obtenidos se distinguieron en base al tamaño diferencial de los amplicones. Las condiciones de amplificación para ADNmt se detallaron en la sección 2.7.2.1.1,

la temperatura en la fase de hibridación fue variable en función del múltiplex/mónoplex ensayado (ver Tabla 2.13).

Tabla 2.13. Temperatura de hibridación para los ensayos PCR–APLP de ADNmt.

Ensayo	Temperatura de hibridación
Mónoplex VII	54 °C
Múltiplex I, V; Mónoplex X	55 °C
Múltiplex II, IV y VI; Mónoplex I-III	57 °C
Múltiplex III	58 °C
Mónoplex V, VIII, IX	61 °C
Mónoplex IV, VI, XI	63 °C

2.8.1.4 Búsqueda y tipificación de PNS en región codificante de linajes A2 mediante secuenciación

Algunos subhaplogrupos previamente descriptos de A2 son portadores de PNS diagnóstico en la región control. Así, por ejemplo, los linajes A2aa se distinguen por la presencia de 16239A-16266 mientras que la combinación 16097-16098 es diagnóstica de A2ah, y la presencia del muy derivado conjunto 59-152-297-16111@-16126-16278 define membresía en A2be, clado ampliamente distribuido en Litoral+NEA y países vecinos. Asimismo, otros haplotipos comparten la presencia de dos o más PNS en región control que permiten asumirlos como parte de grupos monofiléticos, tales como aquellos portadores de 16084-16215 (A2zm), o de 16126-16259 (A2yc). Por último, las restantes secuencias de Formosa y del resto del Litoral y NEA son o bien idénticas a las secuencias nodales de A2 y A2₍₆₄₎, o bien portadoras de uno o unos pocos PNS derivados. Con el objetivo de asignar el máximo número de secuencias a grupos monofiléticos, se realizó una exhaustiva revisión de mitogenomas A2 disponibles en la base de datos de Claudio M. Bravi, conformada tanto por secuencias públicamente disponibles como otras inéditas obtenidas en el IMBICE o donadas por clientes de Family Tree DNA. Este ejercicio puso en evidencia la existencia de cuatro regiones enriquecidas en PNS potencialmente informativos (Tabla 2.14):

1. entre posiciones 2900 y 3960;
2. en el segmento 4216-4706;
3. en el fragmento 10200-10400;
4. entre posiciones 14100 y 14900.

Tabla 2.14. Identificación de PNS informativos en mitogenomas A2.

PNS región codificante	PNS región control	Clado	GenBank ID	País	Población
3435 / 14178 14755 14861		A2ya = A2bg	EU095198 EU095204 OL638617	Brasil Brasil Brasil	<i>Suruí</i> <i>Katuena</i> Porto Alegre
2905 3579 / 14587		A2za1a	KJ923817 OL638679	Brasil Brasil	<i>Guarani</i> Porto Alegre
3834	279	A2zc1	KM656459 MN894801 OL638649	Venezuela Brasil Brasil	<i>Makiritare</i> Porto Alegre Porto Alegre
3083	16422	A2ze	EU095194	Paraguay	<i>Aché</i>
3397		A2zm1	MH981874	Paraguay	Alto Paraná
3693		A2zb	EU095202 KY645725 KY645788	Brasil Colombia Colombia	<i>Yanomama</i> <i>Tuyuca</i> <i>Piapoco</i>
3438	16468/16092/185	A2zj	OL638909 KP172420	Brasil Argentina	Porto Alegre Criollo
4216	143	A2bh	OL638628 EU095203 MH981883	Brasil Brasil Paraguay	Porto Alegre <i>Kayapó</i> Alto Paraná
4158	59	A2zz	MN894745	Brasil	Distrito Federal
4550 / 10398			EU095199	Brasil	<i>Wayampi</i>
4706	nodal		MH981853	Paraguay	Alto Paraná
10203		A2zi	MN894753 OL638658 MW057672	Brasil Brasil Uruguay	Distrito Federal Porto Alegre Montevideo
10237 12858	16266 / nodal	A2zk	KM656458 MG571124 MH981872 MN894720	Venezuela Peru Paraguay Brasil	<i>Makiritare</i> Amazonas Alto Paraná Distrito Federal
14040 14502	16104	A2zg	MH981892	Paraguay	Alto Paraná
15262	nodal		MH981906		

Teniendo en cuenta esta información, se secuenciaron tres regiones en un total de 53 individuos, a saber:

- ✓ 31 individuos para el fragmento FR2 (F2480-R4860), con rango de lectura 2550-3565, o 2550-3610, o 2550-4372, en Macrogen (Corea del Sur).
- ✓ 19 individuos para el fragmento FR7 completo (rango 13845-16031), en Celemics (Corea del Sur).
- ✓ 3 individuos para una porción del fragmento FR7 (rango 14030-14841), en Macrogen.

Los fragmentos FR2 y FR7 fueron amplificados empleando los *primers* y condiciones descritas en Lyons et al. (2013).

Todas las muestras del fragmento FR2 fueron leídas en Macrogen con al menos el *primer* F2625, algunas también con F3316 y F4142. Las 3 muestras del fragmento FR7 fueron leídas en Macrogen con F13976 para el rango 14030-14841, las restantes 19 fueron leídas en forma completa en Celemics.

2.8.1.5 Análisis de variabilidad intrapoblacional

A partir de las secuencias de región control obtenidas de Formosa se calcularon los siguientes índices de diversidad:

1. Índice de diversidad haplotípica: Indica la probabilidad de que dos haplotipos tomados de la muestra al azar sean diferentes entre sí. Se calculó a partir del algoritmo de Nei (1987).

$$\hat{H} = \frac{n}{n-1} \left(1 - \sum_{i=1}^k p_i^2\right)$$

2. Índice de diversidad nucleotídica: Indica la probabilidad de que dos nucleótidos homólogos tomados de la muestra al azar sean diferentes (Tamura & Nei, 1993).
3. Número medio de diferencias nucleotídicas entre pares de haplotipos: Indica el promedio de diferencias nucleotídicas entre todos los pares de haplotipos posibles de una muestra (Tamura & Nei, 1993).

$$\hat{\pi} = \frac{n}{n-1} \sum_{i=1}^k \sum_{j=1}^k p_i p_j \hat{d}_{ij}$$

Los análisis se realizaron utilizando el programa Arlequín 3.1 (Excoffier et al., 2005) y se construyó una tabla con los resultados que fueron comparados con los obtenidos para otras poblaciones de Argentina (Motti, 2012; Schwab, 2018).

2.8.1.6 Análisis comparativos de los linajes maternos

Una vez definidos los subhaplogrupos, se estimaron las frecuencias de los mismos tanto para Formosa como para una selección de poblaciones de la región publicadas en la bibliografía. Se construyeron bases de datos compilando esta información, teniendo en cuenta los subhaplogrupos con un nivel de jerarquía por debajo de los linajes fundadores, definidos según Perego et al. (2010). Además, se realizaron comparaciones específicas de ciertos subhaplogrupos de interés (algunos de ellos descritos en la Tabla 2.5) teniendo en cuenta un nivel mayor de resolución. La abundancia relativa y distribución geográfica de los clados aquí definidos fueron explorados visualmente. Por otra parte, las secuencias obtenidas se utilizaron en la construcción de redes medianas (Bandelt et al., 1999) que se utilizaron para la interpretación de la diversidad dentro de cada clado.

2.8.1.6.1 Construcción de redes de haplotipos

Las redes se realizaron utilizando el programa Network 10.2 y el algoritmo *median joining*; sin embargo, ciertas ramas fueron modificadas en forma manual teniendo en cuenta, además del criterio de parsimonia, criterios de coherencia geográfica e información de región codificante que se obtuvo para este trabajo de tesis. En cada una de las redes se asignaron pesos diferenciales a los distintos caracteres tomando en consideración la escala de posiciones variables publicada por Soares et al. (2009) y otros criterios que se relacionan con la estabilidad de ciertos caracteres en haplogrupos particulares. Por ejemplo, para la posición 16519 existe una gran estabilidad en el haplogrupo B2, mientras que en el resto de los haplogrupos se presenta de forma muy inestable. Para la edición de las imágenes se utilizó el programa Adobe Illustrator 2023.

2.8.1.6.2 Construcción de bases de datos

A fin de ampliar el número de secuencias disponibles comparables con las obtenidas de la localidad de Formosa, se construyeron dos bases de datos. Una de ellas se conformó a partir de una colección de 1952 secuencias del Laboratorio de Genética Molecular Poblacional del

Instituto Multidisciplinario de Biología Celular (IMBICE), agrupando información de un total de 20 localidades de distintas regiones de Argentina (incluida Formosa) (Tabla 2.15). Se mencionará esta última con la sigla BBDD-RI (Base de Datos de Referencia IMBICE).

La segunda base de datos se construyó a partir de 546 mitogenomas obtenidos de la bibliografía o de *GenBank*, y se identificará en esta tesis con la sigla BBDD-RB (Base de Datos de Referencia de la Bibliografía) (Tabla 2.16). El criterio de elección de las muestras en la BBDD-RB fue representar ciertas localidades o regiones de Argentina que no estaban presentes en la BBDD-RI y que se consideró eran relevantes para las comparaciones. Por ejemplo, se incluyeron muestras de San Luis y Córdoba en representación de la región del Centro del país y del Chaco como parte del Gran Chaco Argentino. Además, se incluyeron mitogenomas de países limítrofes que agruparon localidades del sur de Brasil, de Uruguay y el Alto Paraná en Paraguay. Se asignó a cada punto de muestreo un número y un identificador (ID). En algunos casos el identificador coincide con una única localidad; por ejemplo, para Formosa el ID es FOR y coincide con el punto de muestreo 1 y en otros casos se agrupó a más de una localidad en un único punto de muestreo. A modo de ejemplo, las muestras correspondientes a las localidades de Santiago del Estero les corresponde el ID SDE y un único punto de muestreo 11. El mismo criterio se usó para agrupar las localidades de países limítrofes. La localización geográfica de los puntos de muestreo se puede visualizar en la Figura 2.2. Los casos en los que se agruparon más de una localidad en un punto de muestreo representan en el mapa puntos intermedios a la localización real de las localidades representadas por ese punto.

La construcción de estas bases de datos permitió realizar análisis comparativos a nivel local y regional, además de análisis filogenéticos y filogeográficos que se describirán más adelante.

Tabla 2.15. Secuencias de región control completa extendida y región codificante de ADNmt de poblaciones argentinas que conforman la base de datos del Laboratorio de Genética Molecular Poblacional, IMBICE (BBDD-RI) y que fueron comparadas con las secuencias obtenidas de la localidad de Formosa.

Fuente	Población	Provincia/Localidad	Región	N	ID	Punto de muestreo
Este trabajo	Urbana	Formosa, Formosa	NEA	169	FOR	1
IMBICE, inédito	Urbana	Santa Fe, Reconquista	NEA	79	REQ	2
IMBICE, inédito	Urbana	Corrientes, Curuzú Cuatiá	NEA	140	CC	3
IMBICE, inédito	Urbana	Entre Ríos, La Paz	NEA	56	LPZ	4
IMBICE, inédito	Urbana	Entre Ríos, Gualeguaychú	NEA	115	GUA	5
Motti, 2012	Urbana	Salta, Tartagal	Tierras Bajas	169	TG	6
Motti, 2012	Urbana	Jujuy, La Quiaca	NOA	169	LQ	7
Motti, 2012	Urbana	Jujuy, Maimará	NOA	162	MM	8
Motti, 2012	Urbana	Jujuy, San Salvador de Jujuy	NOA	34	JUJ	9
Schwab, 2018	Urbana	Tucumán, Tucumán	NOA	83	TMN	10
Schwab, 2018	Urbana	Santiago del Estero, Santiago del Estero	NOA	83	SDE	11
Motti, 2012	Urbana	Catamarca, Santa María	NOA	111	SMA	12
Motti, 2012	Urbana	Catamarca, Belén	NOA	61	BEL	13
Motti, 2012	Urbana	La Rioja, La Rioja	NOA	78	LR	14
Motti, 2012	Urbana	La Rioja, Chepes	NOA	44	CHE	15
Motti, 2012	Urbana	San Juan, Calingasta	Cuyo	65	CAL	16
Motti, 2012	Urbana	San Juan, San Juan	Cuyo	99	SNJ	17
Motti, 2012	Urbana	Mendoza, Villa Tulumaya	Cuyo	63	TUL	18
Motti, 2012	Urbana	Mendoza, Mendoza	Cuyo	55	MZA	19
IMBICE, inédito	Urbana	Mendoza, Malargüe	Cuyo	117	MLG	20

Tabla 2.16. Mitogenomas recopilados de la bibliografía que conforman la Base de Datos de Referencia de la Bibliografía (BBDD-RB) y que fueron comparados con las secuencias obtenidas de la localidad de Formosa.

Fuente	Población	Localidad/País	Región	N	N total	ID	Punto de muestreo
García et al., 2021	Urbana	Villa Atamisqui, Santiago del Estero, Argentina	NOA, Argentina	12	16	SDE	11
		Otras localidades, Santiago del Estero, Argentina		4			
García et al., 2021	Urbana	La Toma, San Luis, Argentina	Centro, Argentina	14	24	SL	21
		Otras localidades, San Luis, Argentina		10			
García et al., 2021	Urbana	Córdoba, Córdoba, Argentina	Centro, Argentina	22	52	CBA	22
		Chancaní, Córdoba, Argentina		1			
		Cruz Alta, Córdoba, Argentina		5			
		Otras localidades, Córdoba, Argentina		24			
Sevini et al., 2014 <i>GenBank direct submission</i>	Criollo	Argentina, Chaco, Chaco	Gran Chaco argentino	55	124	CHA_ARG	23
	Wichí			69			
Simão et al., 2019	Urbana	Alto Paraná, Paraguay	Paraguay	90	107	PAR	24
Simão et al., 2021	Urbana	Paraguay		17			
Figueiro et al., 2022	Urbana	Montevideo, Uruguay	Uruguay	13	32	URU	25
		Tacuarembó, Uruguay		13			
		Trinidad, Uruguay		6			
Fagundes et al., 2008	Indígena	Brasil	Brasil	47	191	BRA	26
Rieux et al., 2014	Indígena			8			
Ávila et al., 2019	Urbana			39			
Ávila et al., 2022	Urbana			97			



Figura 2.2. Localización de los puntos de muestreo. Se denotan con distintos colores las localidades de Argentina de la región del NEA (lila), NOA (amarillo), Cuyo (verde) y Centro (celeste) y de localidades de países limítrofes (Uruguay, Paraguay y Brasil, en gris). La localidad de Formosa se distingue por el número 1. Los números hacen referencia al punto de muestreo de cada localidad asignado en la Tabla 2.15 y Tabla 2.16.

2.8.1.6.3 Subhaplogrupos mayoritarios y comparaciones a nivel local y regional

En base a las frecuencias obtenidas se definieron ciertos subhaplogrupos como mayoritarios en la muestra de Formosa y se indagó su presencia en otras poblaciones de la región, a partir de muestras analizadas en las bases de datos construidas para este trabajo. En este sentido, se calcularon las frecuencias de estos clados para las regiones del 1-Nordeste argentino (NEA), 2-Noroeste argentino (NOA), 3-Tartagal, 4-Cuyo, 5-Uruguay (URU), 6- Alto Paraná, Paraguay (PAR), 7-Gran Chaco argentino (CHA_ARG), 8-Brasil (BRA) y 9-Centro de Argentina (CENTRO) a partir de la sumatoria de frecuencias obtenidas para 1- Formosa, Curuzú Cuatiá, Reconquista, La Paz, Gualaguaychú; 2- La Quiaca, Maimará, San Salvador de Jujuy, Santiago del Estero, Tucumán, La Rioja, Chepes, Catamarca, Belén y Santa María; 3-Tartagal; 4-Calingasta, San Juan, Mendoza, Malargüe y Villa Tulumaya; 5-Uruguay; 6- Paraguay este; 7- Chaco argentino, 8-Brasil y 9- Córdoba y San Luis, respectivamente, a fin de realizar comparaciones regionales que incluyeron países limítrofes.

2.8.1.7 Análisis de la estructuración geográfica de la variabilidad

Para las secuencias del ADNmt, se exploró el agrupamiento de las diferentes poblaciones en estudio, mediante el escalamiento multidimensional de distancias genéticas (distancia de Nei; Nei, 1972). La estructura genética de las poblaciones se analizó a través del análisis molecular de la varianza (Excoffier et al., 1992). Además, se identificaron barreras genéticas entre localidades con el programa Barrier 2.2 (Manni et al., 2004) a partir de la estimación de las distancias de Nei entre ellas (NTSYSpc Versión 2.11, coeff=NEI72). En suma, estos análisis permitirán poner a prueba distintas hipótesis de afinidades interregionales entre las poblaciones estudiadas, como se mencionó anteriormente.

2.9 Tercer nivel de análisis: distribución sudamericana de clados mitocondriales

Se seleccionaron ciertos clados de interés (A2bf, B2an, B2e, D1e y D1f1) para analizar su distribución y estructuración geográfica a nivel sudamericano, a fin de identificar rutas de poblamiento. La selección de linajes se realizó teniendo en cuenta su abundancia relativa, importancia filogeográfica o escasa representación en estudios publicados, en base al análisis de las 169 secuencias de Formosa en conjunto con las aproximadamente 29.000 secuencias de la base de datos del IMBICE.

Se compilaron y analizaron secuencias de poblaciones sudamericanas disponibles en la bibliografía, en las bases de datos genéticas (BBDD-RI, BBDD-RB, Genbank) y obtenidas por donación de clientes de Family Tree DNA, pertenecientes al haplogrupo A2 (para A2bf), B2 (para B2an y B2e) y D1 (para D1e y D1f1). Estas recopilaciones se realizaron a fin de poder asignar nuevas secuencias a los subhaplogrupos de interés y de establecer un rango más amplio de su distribución. Del total de secuencias mitocondriales analizadas, se seleccionaron aquellas de las que se disponía de información tanto de la RHV-I como de la RHV-II y se unificaron los rangos de lectura, con el objetivo de contar con secuencias de la misma longitud para realizar una red de haplotipos de cada linaje de interés y maximizar el número de secuencias disponibles. Además, se realizaron mapas que permitieron visualizar la frecuencia de cada linaje por punto de muestreo y la geolocalización de los mitogenomas recopilados. A su vez, se construyeron manualmente árboles filogenéticos en base a los mitogenomas recopilados.

2.9.1 Caracterización de alta resolución del haplogrupo B2e

Un análisis filogenético preliminar de mitogenomas disponibles dejó en evidencia que en las poblaciones del NEA y Litoral existían dos ramas del clado B2e: B2e1, definido por la presencia de 15672-15781 (García et al., 2021); y B2e2, diagnosticado por 986 y con una rama interna definida por la presencia de 9899.

Inicialmente, las 21 muestras asignadas a B2e de la colección del NEA y Litoral fueron interrogadas para la posición 9899 mediante el ensayo PCR-APLP descrito en Tabla 2.9. Las nueve muestras portadoras del alelo ancestral para 9899 fueron analizadas para 15672-15781 mediante secuenciación del fragmento FR7 en el rango 15236-16024 con el *primer* F15190 (5'-CTTACTATCCGCCATCCCATA-3). Cuatro de estas últimas muestras, no asignables a B2e1, fueron analizadas para presencia de 986 mediante secuenciación del fragmento FR1 en el rango 460-1287 con el *primer* F1-402 (5'-ATCTTTTGGCGGTATGCACTTT-3').

3 CAPÍTULO 3: RESULTADOS

Los resultados de este trabajo se articulan en los tres ejes planteados en el esquema de trabajo, que comprenden tres niveles de resolución de los datos y a su vez involucran distintas escalas de análisis (local, regional y continental).

En el primer nivel se trabajó con frecuencias de haplogrupos de linajes maternos y paternos en las localidades de Formosa y Corrientes. A partir de la tipificación de ciertas mutaciones diagnósticas del ADNmt y del cromosoma Y, se determinaron los orígenes continentales de las poblaciones (procedencia de América, Europa o Medio Oriente o África Subsahariana). Además, se contrastaron los resultados moleculares con información genealógica, que se obtuvo interrogando a los donantes acerca de cuál recordaban que era el lugar de nacimiento del último ancestro conocido por vía materna y paterna.

En un segundo nivel de análisis, se hizo foco en la fracción nativa en base a secuencias de la región control del ADNmt, para realizar comparaciones a escala regional. En este sentido, se obtuvieron secuencias de la localidad de Formosa y se trabajó con secuencias previamente obtenidas de una colección de nuestro laboratorio y otras recopiladas de la bibliografía.

Finalmente, en el tercer y último nivel, se seleccionaron de las secuencias analizadas ciertos clados de interés en base a información de la región codificante del ADNmt, conjuntamente con información obtenida de mitogenomas de la bibliografía y de esta forma, se utilizaron para postular presuntos clados monofiléticos de diferente jerarquía, explorando su distribución y estructuración a escala sudamericana o continental.

3.1 Primer nivel de análisis: migraciones intercontinentales

3.1.1 Linajes maternos

3.1.1.1 Frecuencias de haplogrupos mitocondriales

Los resultados demostraron una presencia mayoritaria de linajes nativos por vía materna en cada una de las localidades de muestreo (FOR= 85,04%, CRR= 89,40%) (Tabla 3.1). Por otra parte, el aporte de linajes europeos y de Medio Oriente fue variable entre ellas (FOR=12,41%, CRR=9,93%). Es interesante destacar la relativamente alta proporción de linajes africanos (en relación a la fracción autóctona) en la muestra FOR (2,55%), siendo menor en CRR (0,66%) (Tabla 3.1). Sin embargo, un Test Exacto de Diferenciación Poblacional (TEDP) (Raymond & Rousset, 1995) muestra que no existen diferencias significativas en cuanto a los aportes relativos continentales entre ambas poblaciones ($P > 0,05$ para todas las comparaciones).

Entre los linajes nativos el haplogrupo C es el que se encontró en mayor frecuencia en las dos localidades muestreadas (FOR= 26,64%, CRR= 28,48%); en Corrientes, además, predominó el haplogrupo B (CRR= 28,48%) (Tabla 3.1). Por otra parte, la caracterización de linajes al interior del macro-haplogrupo N de origen europeo o de Medio Oriente, refleja una frecuencia predominante del haplogrupo H en las dos poblaciones estudiadas (FOR= 3,65%, CRR= 4,64%), seguida por el haplogrupo U (FOR= 2,55%), y por el haplogrupo I/X (FOR= 2,19%, CRR=1,99%) (Tabla 3.1). El resto de los haplogrupos presentaron frecuencias menores al 2% en ambas localidades.

Tabla 3.1. Origen continental de los haplogrupos mitocondriales, en porcentaje por localidad.

Origen continental	Hg	Formosa		Corrientes	
		N	Frec %	N	Frec %
Nativo	A	57	20,80	34	22,52
	B	63	22,99	43	28,48
	C	73	26,64	43	28,48
	D	40	14,60	15	9,93
	Total	233	85,04	135	89,40
Europeos y de Medio Oriente	H	10	3,65	7	4,64
	J	3	1,09	0	0
	K	2	0,73	0	0
	T	5	1,82	2	1,32
	U	7*	2,55	1	0,66
	V	0	0	0	0
	W	0	0	1	0,66
	I/X	6	2,19	3	1,99
N (x H, J, K, T, U, V, W, I/X)	3	1,09	1	0,66	
Total	34	12,41	15	9,93	
Africanos	L (x M, N)	7	2,55	1	0,66
Total		274	100	151	100

*Dos individuos resultaron K-U

3.1.1.2 Información genealógica para la línea materna

Respecto a la información genealógica brindada por los donantes, se recopiló el dato sobre el último ancestro conocido (UAC) por vía materna, siendo en un 74% de los casos de la tercera generación (Tabla 3.2).

Tabla 3.2. Procedencia de los datos genealógicos para la línea materna, en porcentaje por localidad.

	N	Datos genealógicos		
		Abuela	Madre	Donante
Formosa	274	73,4	23,7	2,9
Corrientes	151	74,8	21,2	4,0
TOTAL	425	73,9	22,8	3,3

Los resultados demostraron un mayor componente “regional” (85,6%) -en el corto plazo- del origen declarado por los donantes por línea materna (Tabla 3.3). El 62,4% declara que

sus ancestros nacieron en la región del NEA y un 22,4% en Paraguay. De aquellos que declararon un origen extrarregional, 10,4% fue del resto de Argentina, 0,7% del resto de Sudamérica y 3,3% de origen europeo. Se destaca que el 74,8% de los donantes de Corrientes declara que sus ancestros nacieron dentro de la misma provincia y que se registra un aporte de 32,1% de ancestros de origen paraguayo en Formosa. Esta frecuencia nos indica la influencia de Paraguay en la región del NEA respecto a los procesos poblacionales, tendencia que ya podía observarse en los datos reportados en el Censo Nacional de 1914. Este fenómeno es especialmente notable en Formosa donde, sobre una población de 8.774 habitantes, el 45% era extranjero y, a su vez, de esta fracción el 84% provenía de Paraguay y 13,6% era de origen ultramarino; mientras que para Corrientes, sobre una población de 24.462 habitantes, el 7% era extranjero, y entre ellos, el 16,9% era de origen paraguayo y el 39% era de origen ultramarino.

Tabla 3.3. Lugar de origen declarado de último ancestro conocido por vía materna, en porcentaje por localidad. NEA= Nordeste Argentino.

Línea materna					
			Formosa	Corrientes	Total
Origen declarado	Regional	Formosa	107	1	364
		Corrientes	24	113	
		Chaco	9	5	
		Misiones	3	3	
		NEA	143	122	
		Paraguay	88	7	
		Uruguay	1	2	
	Brasil	0	1		
	Extrarregional	Resto de Argentina	31	13	61
		Resto de Sudamérica	2	1	
Europa		9	5		
N			274	151	425

3.1.1.3 Comparación entre resultados moleculares y genealógicos para la línea materna

Cuando se realizaron las comparaciones de los orígenes identificados genéticamente con aquellos declarados por los donantes, se observó que 14 de los 425 donantes analizados para la línea materna declaran un origen ultramarino, de los cuales 11 (79%) fueron corroborados por el resultado molecular, mientras que los tres restantes resultaron portadores

de linajes nativos (Figura 3.1.A). Estas discrepancias en las comparaciones por vía materna podrían explicarse por pérdida de este dato en la historia familiar, casos de adopción, apropiación ilegítima, entre otros. En aquellos que declaran un origen regional de sus ancestros (N=411), aproximadamente 89% (N=365) presenta ancestría nativa. Sin embargo, un 11% (N=46) presenta un origen alóctono, ya que los resultados moleculares arrojan que el 9% es de origen europeo y el 2% de origen africano (Figura 3.1.B).

Los resultados de esta sección se analizaron en conjunto, debido a que no hubo diferencias significativas entre las localidades estudiadas.

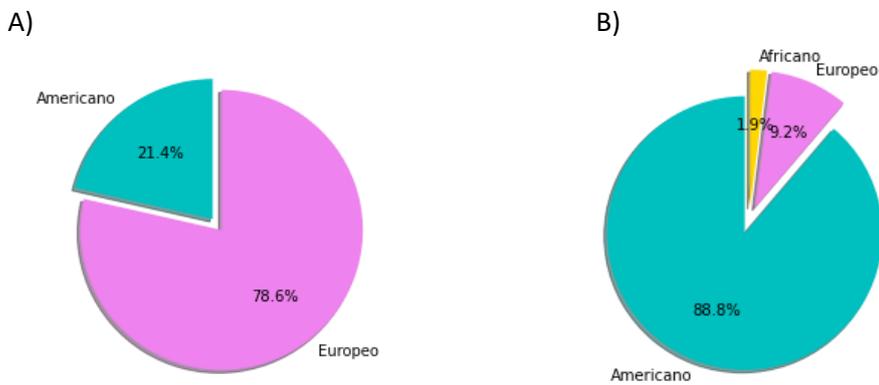


Figura 3.1. Frecuencia de individuos que declaran un origen ultramarino (A) y un origen regional (B) para ancestros maternos más remotos y su correspondencia con los resultados moleculares, en porcentaje.

3.1.2 Linajes paternos

3.1.2.1 Frecuencias de haplogrupos del cromosoma Y

Más del 80% de los linajes paternos en las poblaciones muestreadas correspondieron a un origen ancestral europeo o de Medio Oriente, con valores de alrededor de 55% para el haplogrupo R1 (FOR= 55,2%, CRR= 54,7%) (Tabla 3.4). Cabe destacar la baja proporción de linajes nativos Q-M3 (FOR: 4,2%; CRR: 3,2%) en relación a la fracción alóctona.

No se encontraron diferencias significativas entre ambas poblaciones en cuanto a los aportes relativos de los clados (TEDP: $P > 0,05$).

Tabla 3.4. Frecuencia de los haplogrupos del cromosoma Y por localidad de muestreo, en porcentaje

	Haplogrupos	Formosa		Corrientes	
		N	Frec. (%)	N	Frec. (%)
Alóctono	E	16	11,2	8	8,4
	K (xQ,R)	8	5,6	16	16,8
	F (xG,I,J)	2	1,4	0	0,0
	G1	0	0,0	0	0,0
	G2	6	4,2	3	3,2
	J1	5	3,5	2	2,1
	J2	12	8,4	7	7,4
	I1	0	0,0	2	2,1
	I2	8	5,6	2	2,1
	R1*	1	0,7	3	3,2
	R1a1	4	2,8	4	4,2
	R1b1a1a	74	51,7	45	47,4
	Q (xM3)	1	0,7	0	0,0
Nativo	Q1a2a1a1 (Q-M3)	6	4,2	3	3,2
Total		143	100	95	100

3.1.2.2 Información genealógica para la línea paterna

Los datos de los donantes que informaron acerca de su UAC por línea paterna fueron en un 55,3% de los casos de la tercera generación a nivel global (FOR=59,2%; CRR=49,5%), registrándose, además, una frecuencia mayor al 30% de datos que provenían de la segunda generación (FOR=34,5%; CRR=42,1%). El resto de la información provino de los donantes (FOR=6,3%; CRR=8,4%) (Tabla 3.5).

Tabla 3.5. Procedencia de los datos genealógicos para la línea paterna, en porcentaje por localidad.

	Datos genealógicos			
	N	Abuelo	Padre	Donante
Formosa	142	59,2	34,5	6,3
Corrientes	95	49,5	42,1	8,4
TOTAL	237	55,3	37,6	7,2

Se observa un mayor componente “regional” (aproximadamente 80%) en el corto plazo, exhibido a través de la historia oral (Tabla 3.6), preponderantemente en aquellos cuyo lugar de nacimiento corresponde al NEA y especialmente para la localidad de Corrientes. Se registra un aporte del 31% de ancestros de origen paraguayo en Formosa.

Tabla 3.6. Lugar de origen declarado de último ancestro conocido por vía paterna, en porcentaje por localidad. NEA= Nordeste Argentino.

Línea paterna						
			Formosa	Corrientes	Total	
Origen declarado	Regional	Formosa	47	0	190	
		Corrientes	9	71		
		Chaco	7	3		
		Misiones	0	0		
		NEA	63	74		
		Paraguay	44	6		
		Uruguay	2	0		
		Brasil	0	1		
	Extrarregional	Resto de Argentina	14	7	47	
		Resto de Sudamérica	0	0		
		Europa	19	7		
	N			142	95	237

3.1.2.3 Comparación entre resultados moleculares y genealógicos para la línea paterna

Se obtuvieron los datos genealógicos para la línea paterna y se analizaron en contraste con los resultados moleculares, en los casos que eran comparables.

Los resultados demuestran una mayor frecuencia de donantes varones que declaran un origen regional de sus ancestros paternos más remotos (210/236), en comparación con aquellos que declaran origen ultramarino (26/236). En ese sentido, los resultados moleculares indican que 25 de los 26 que declaran un origen ultramarino efectivamente son portadores de linajes euroasiáticos, mientras que el restante lleva uno de origen nativo. Por el contrario, de aquellos que mencionan un origen regional de sus ancestros, sólo 4% de los resultados moleculares

coinciden con la información declarada, mientras que aproximadamente 96% (N= 201) tiene un origen euroasiático.

3.1.3 Linajes nativos: línea materna vs línea paterna

Por último, se compilaron los resultados de frecuencias de linajes nativos para la línea materna y paterna de este trabajo con los de 15 localidades de Argentina del NOA y COA (Motti, 2012; Jurado Medina, 2015; Schwab, 2018; Ramallo et al., 2009) analizados previamente (Figura 3.2).

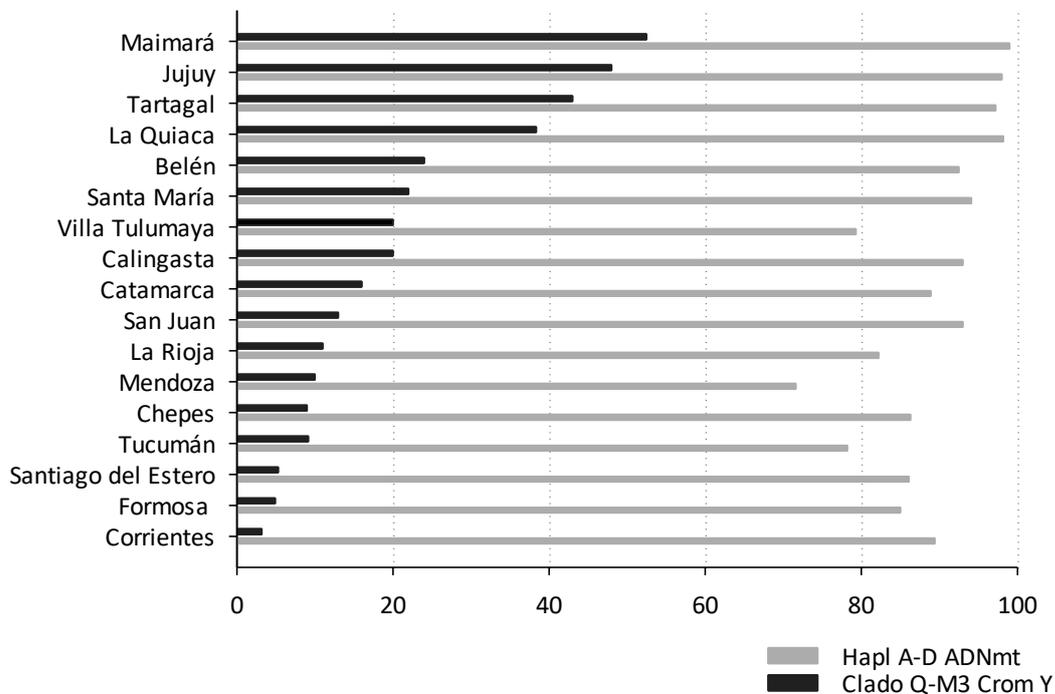


Figura 3.2. Frecuencia relativa de linajes nativos por línea materna (haplogrupos A-D) y por línea paterna (clado Q-M3) en Formosa y Corrientes y en las localidades muestreadas hasta el momento en la colección del IMBICE, en porcentaje.

3.2 Segundo nivel de análisis: vínculos interregionales

3.2.1 Análisis de linajes maternos

3.2.1.1 Obtención de secuencias de región control del ADNmt

Del 85% de las muestras de la ciudad de Formosa que fueron asignadas a los haplogrupos nativos A-D, se seleccionaron 169 (teniendo en cuenta un origen regional declarado para la línea materna), y se obtuvieron secuencias de la región control (RC) completa extendida en sentido 5' tal como se muestran en el Anexo 3.1.

3.2.1.2 Análisis de variabilidad intrapoblacional

Se estimó la variabilidad intrapoblacional de las secuencias nativas entre las posiciones 16024-16569 y 1-576 de la ciudad de Formosa. No se consideraron las variaciones en el número de residuos citosina en los homopolímeros en torno a 16184-16193, 303-309, 568-573 ni las variaciones en el número de repetidos CA en el microsatélite 514-523. Los resultados se analizaron en comparación con los obtenidos para 14 poblaciones por Motti (2012) y Schwab (2018). Si se toma en consideración la diversidad haplotípica, Formosa presenta junto con otras cinco poblaciones (Maimará, Mendoza, Tartagal, La Quiaca y Jujuy) los mayores valores (alrededor de 0,99) (Tabla 3.7). Sin embargo, en relación al porcentaje de haplotipos únicos, los valores se encuentran cercanos a los obtenidos en Belén y Santiago del Estero, y se ubican en el tercio más bajo en relación al conjunto de las 14 localidades comparadas. Esto se atribuye a que, en la mayoría de los casos, los haplotipos encontrados se encuentran en al menos dos individuos de la muestra de Formosa.

Tabla 3.7. Diversidad intrapoblacional para 14 localidades del centro-oeste y noroeste argentino

Población	N	h	s	k	Diversidad Nucleotídica	Diversidad haplotípica	Haplotipos únicos	
							N	%
Jujuy	34	32	81	15,791 +/- 7,223	0,0141 +/- 0,0072	0,9964 +/- 0,0082	30	88,24
La Quiaca	169	126	200	12,294 +/- 5,580	0,0108 +/- 0,0054	0,9951 +/- 0,0015	98	57,99
Tartagal	167	127	192	14,748 +/- 6,633	0,0132 +/- 0,0066	0,9945 +/- 0,0017	106	63,47
Mendoza	55	48	85	13,200 +/- 6,031	0,0118 +/- 0,0060	0,9939 +/- 0,0051	43	78,18
Formosa	169	123	154	12,881 +/- 4,345	0,0116 +/- 0,0002	0,9934 +/- 0,0019	68	40,23
Maimará	162	107	171	11,906 +/- 5,415	0,0106 +/- 0,0053	0,9920 +/- 0,0020	75	46,3
V. Tulumaya	63	50	96	15,968 +/- 7,216	0,0143 +/- 0,0072	0,9898 +/- 0,0053	43	68,25
San Juan	100	67	105	14,684 +/- 6,628	0,0131 +/- 0,0066	0,9822 +/- 0,0059	51	51
Santa María	111	62	108	14,939 +/- 6,732	0,0133 +/- 0,0067	0,9810 +/- 0,0046	39	35,14
Chepes	44	31	58	12,564 +/- 5,777	0,0113 +/- 0,0057	0,9799 +/- 0,0099	22	50
S. del Estero	83	50	88	11,796 +/- 5,394	0,0105 +/- 0,0053	0,9774 +/- 0,0065	38	45,78
Belén	61	37	67	13,221 +/- 6,031	0,0118 +/- 0,0060	0,9727 +/- 0,0094	26	42,62
S.M. de Tucumán	83	57	108	12,331 +/- 5,624	0,0110 +/- 0,0056	0,9680 +/- 0,0116	48	57,83
Calingasta	65	35	74	14,006 +/- 6,365	0,0126 +/- 0,0063	0,9553 +/- 0,0145	23	35,38
La Rioja	78	40	69	11,896 +/- 5,440	0,0107 +/- 0,0054	0,9257 +/- 0,0234	28	35,9

h: número de haplotipos.

s: número de sitios polimórficos.

k: número promedio de las diferencias entre 2 secuencias, T-N.

El orden de las filas desde Jujuy a La Rioja es descendente en función de los valores de la diversidad haplotípica.

En color lila se marcan las localidades con menor % de haplotipos únicos

3.2.1.3 Definición de subhaplogrupos mitocondriales

Para la asignación de subhaplogrupos se analizaron polimorfismos encontrados tanto en la región control extendida como en la región codificante en las secuencias de Formosa y se los clasificó teniendo en cuenta su aparición en Phylotree_{mt} y en la bibliografía. Por otra parte, se definieron nuevos clados teniendo en cuenta las combinaciones de polimorfismos que aún no se encontraban descriptas. Se construyó una tabla que describe las frecuencias de subhaplogrupos definidos en las 20 localidades de la Base de Datos de Referencia IMBICE (BBDD-RI) que incluye los de la localidad de Formosa (Anexo 3.2). Se lograron identificar un total de 201 linajes con presunto origen monofilético. En la Tabla 3.8 se describen los totales de los linajes definidos por haplogrupo.

Tabla 3.8. Linajes monofiléticos identificados por haplogrupo en las 20 localidades de la BBDD-RI

	Haplogrupo								Total
	A2	B2	C1b	C1c	C1d	C4c	D1	D4h3a	
Linajes monofiléticos identificados	49	58	29	7	12	1	35	10	201

3.2.1.3.1 Relevancia de la Identificación de polimorfismos del ADN_{mt} en región codificante

Si bien en las últimas décadas ha aumentado considerablemente el número de mitogenomas disponibles en la bibliografía y en las bases de datos genéticas, el volumen de información de secuencias de región control es mucho mayor. El hecho de tener un mayor número otorga a las secuencias de región control una ventaja en relación a los mitogenomas, ya que permitirían realizar una gran cantidad de comparaciones con las secuencias de otras zonas geográficas y asignarlas a ciertos clados a partir de las variantes compartidas. Sin embargo, ciertos polimorfismos de la región control son hipervariables, con lo cual, la interpretación de un origen común es difícil de establecer si sólo se dispone de este tipo de información. En este sentido, es importante destacar que disponer de ciertos PNS de región codificante aportaría en la definición de clados monofiléticos y en la reconstrucción filogenética. Tal es así, que en este trabajo de tesis la interpretación de las secuencias de región control se realizó en conjunto con ciertos PNS que se consideraron relevantes para explorar en la región codificante, algunos de los cuales fueron encontrados por secuenciación de fragmentos específicos y otros por APLP, tanto en muestras de Formosa como en otras muestras de la BBDD-RI. Esto último posibilitó la

definición de nuevos clados, la asignación de clados ya sea que se trate de linajes sin variantes compartidas o con variantes homoplásicas en región control y la posibilidad de reconocer linajes y asignarlos a un subhaplogrupo determinado, cuando éstos presentaban patrones diversos en la región control. En la tabla presentada en el Anexo 3.3, se puede observar la asociación entre haplotipos de región control, PNS en regiones codificantes y clados dentro del haplogrupo A2.

3.2.1.3.2. Construcción de redes de haplotipos

Una vez identificados los presuntos linajes monofiléticos de la muestra de Formosa y de haber establecido su ocurrencia en otras 19 localidades de la BBDD-RI (Anexo 3.2 y Anexo 3.4), se construyeron redes de haplotipos para los haplogrupos nativos A2, B2, C1b, C1d, C1c y D1, a modo de representar las respectivas hipótesis filogenéticas.

3.2.1.3.2.1. Haplogrupo A2

El haplogrupo A2 presenta frecuencias que oscilan entre el 8% al 32% del total de los linajes nativos entre las muestras analizadas en la BBDD-RI. Las frecuencias más bajas corresponden a Tucumán, Santa María y La Rioja, mientras que las más altas se encuentran en Curuzú Cuatiá, Reconquista y Villa Tulumaya.

Se construyeron dos redes medianas del haplogrupo A2 partiendo de dos secuencias nodales A2₍₆₄₎ (Figura 3.3) y A2 (Figura 3.5). La red A2₍₆₄₎ principal se realizó a partir de 72 secuencias de poblaciones del NEA y Litoral y agrupa 34 haplotipos (Figura 3.3). Tiene forma estrellada y presenta amplia diversidad de linajes. Desde su secuencia nodal se separan 17 ramas principales, una de las cuales corresponde al linaje A2be, que se describe por separado más adelante. El haplotipo con mayor representación es el que corresponde al clado A2zh, agrupando doce individuos de las localidades de Gualeguaychú, Curuzú-Cuatiá, Reconquista y Formosa en un único haplotipo. El resto de los haplotipos representados agrupan de uno a cinco individuos y caracterizan clados como A2zj, A2zl, A2aa, A2zk3, A2ab, A2ah1, A2zb y A2zd.

Además, la búsqueda de una batería de PNS por APLP en la región codificante (2905-3579, 3083, 14502-14040, 3834, 4158, 3693) permitió diferenciar 6 ramas de la red que se asignaron a los clados A2za1a, A2ze, A2zg1, A2zc1, A2zz y A2zb1, respectivamente. Es interesante resaltar que el haplotipo que caracteriza a A2za1a se encontró además en mitogenomas de Paraguay (2300544, Simão et al., 2021) y de Brasil, tanto en población urbana (PUC462, Ávila et al., 2022) como guaraní (KJ923817, Rieux et al., 2014). La presencia de 3693 vincula a Formosa con un clado de amplia distribución en la Amazonía (yanomama y mestizos

de Brasil, Fagundes et al., 2008, Ávila et al., 2022; Tuyuca, Piapoco y Achagua de Colombia, Arias et al., 2017).

La rama correspondiente a A2be se amplió en otra red que agrupa el 27% de las secuencias analizadas del haplogrupo A2₍₆₄₎. Este linaje fue originalmente definido por Figueiro et al. (2022), quien reportó su presencia en mitogenomas de Uruguay y de Brasil. En este trabajo, las 27 secuencias recopiladas de localidades del NEA y Litoral agrupan un total de nueve haplotipos, cinco de ellos encontrados en individuos de la localidad de Formosa (Figura 3.4). Se pudieron identificar subhaplogrupos de menor jerarquía como A2be1a4a, también encontrado en mitogenomas de Brasil (Z1759 y PUC643, Ávila et al., 2022); A2be1a1, también registrado en mitogenomas de Uruguay (MW057673, Figueiro et al., 2022) y de Brasil (PUC732, Ávila et al., 2022); y A2be1a1d, encontrado en un mitogenoma de Uruguay (MW057675, Figueiro et al., 2022).

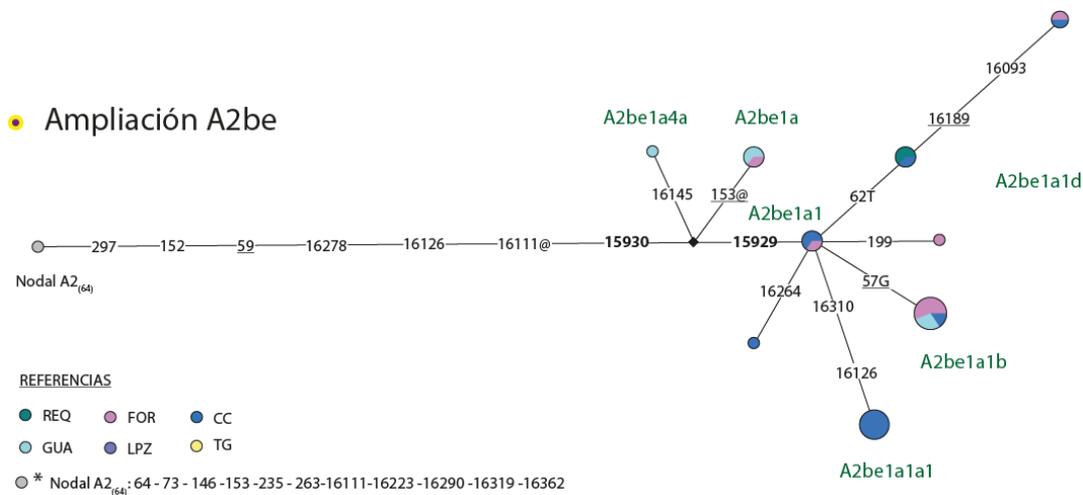


Figura 3.4. Ampliación de red de haplotipos A2₍₆₄₎. Linaje A2be. La figura se desprende de la red principal de A2₍₆₄₎ (Figura 3.3). Detalles en la leyenda de la Figura 3.3

La red A2 se construyó en base a 26 secuencias de poblaciones del NEA y NOA y agrupa 13 haplotipos (Figura 3.5). Desde la secuencia nodal parten cinco ramas principales, la primera de ellas porta la variante diagnóstica 16051 que concentra 17 secuencias en 8 haplotipos. Esta rama caracteriza al linaje A2bf, compartido tanto por las secuencias presentadas aquí como por otras uruguayas y chaqueñas (Figueiro et al., 2022). En esta red, las secuencias pertenecientes a A2bf representan el 65% de las muestras de la red del haplogrupo A2.

Dos de las ramas principales se caracterizan por estar enraizadas en sitios diagnósticos de la región codificante que fueron encontrados por APLP (Anexo 3.3). Se trata de la variante 3866, que agrupa un haplotipo que se encuentra representado por un individuo de Formosa y la variante 3435, que concentra dos haplotipos. Este último linaje fue identificado como A2ya y

se ha encontrado en un mitogenoma de Brasil (PUC487, Ávila et al., 2022). Un haplotipo encontrado en tres individuos de Formosa que fue asignado al clado A2yc1a, es compartido por dos mitogenomas de Paraguay (MH981840, Simão et al., 2019; 2300550, Simão et al., 2021).

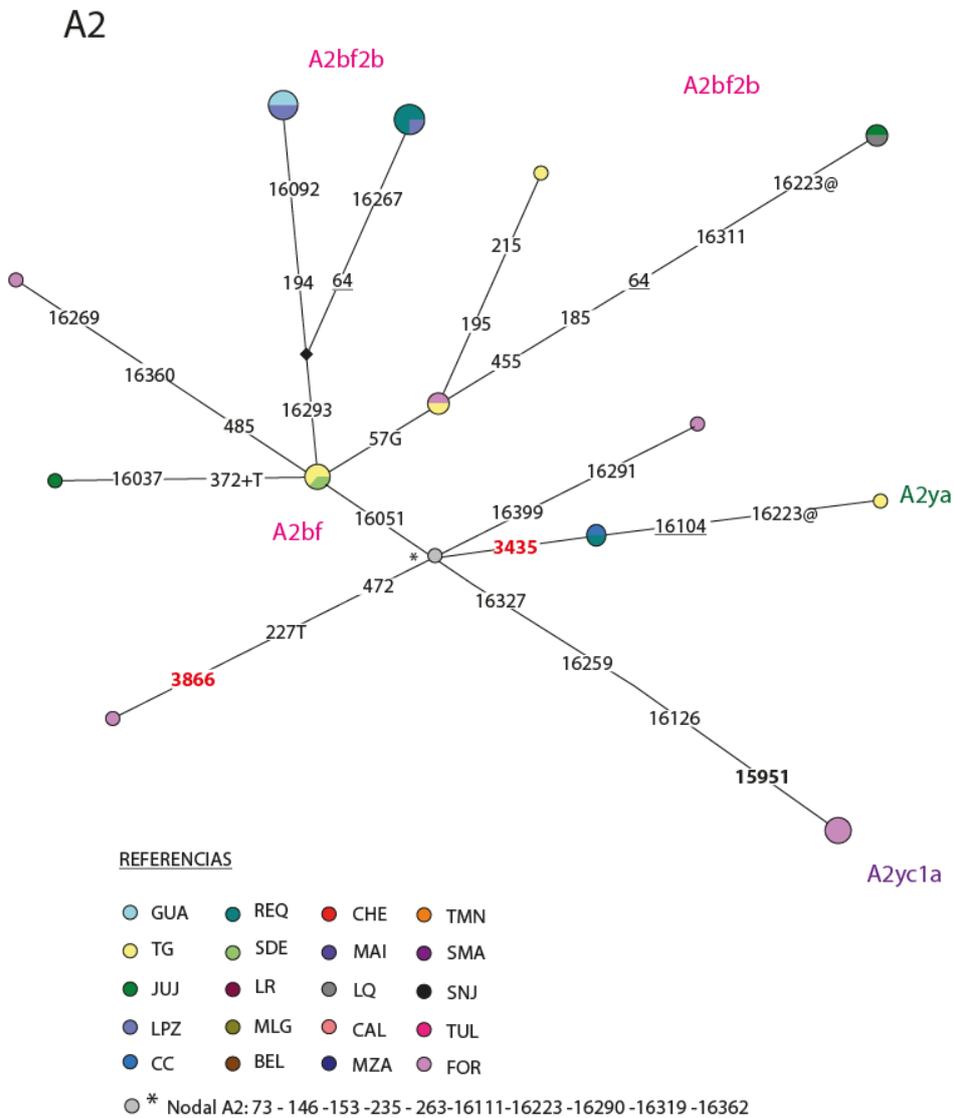


Figura 3.5. Red de haplotipos A2. Detalles en la leyenda de la Figura 3.3

3.2.1.3.2.2 Haplogrupo B2

El haplogrupo B2 resultó ser el más frecuente de la localidad de Formosa, alcanzando el 31% de las secuencias analizadas. En otras localidades de Argentina de la BBDD-RI, se encuentra representado con frecuencias que varían desde el 5% en Chepes y cercanas al 70% en La Quiaca y Maimará.

Las redes del haplogrupo B2 se realizaron a partir de 177 secuencias de poblaciones del NEA. Ochenta y siete secuencias conforman la red principal agrupando 46 haplotipos (Figura 3.6). Tiene forma de estrella y una gran diversidad de linajes que conforman 14 ramas principales, tres de las cuales se amplían en redes por separado (Figura 3.7) y conforman los linajes de más alta frecuencia (B2+14470, B2+16178 y B2e).

La búsqueda de PNS por APLP permitió caracterizar tres ramas de la red: la rama B2+14470, que se detallará más adelante; B2+6755 a partir de la cual se pudieron identificar los linajes B2b15, B2b* y B2b3a; y B2+7055 que agrupó una muestra de Formosa con tres haplotipos de Tartagal. Otros linajes encontrados permitieron establecer afinidades entre las muestras de Formosa y NEA y Litoral con otras regiones sudamericanas. Por ejemplo: B2o1a, también se ha encontrado en mitogenomas de población cosmopolita de Bolivia (KC503932, Taboada-Echalar et al., 2013) y de Paraguay (MH981857, Simão et al., 2019); B2zf1, representado en tres secuencias de Formosa, es compartido por mitogenomas de comunidades indígenas de Brasil: en un individuo Kaingang (KJ923825, Rieux et al., 2014) y otro Guaraní (EU095219, Fagundes et al., 2008); B2i1 se ha asociado a secuencias de Brasil (Fagundes et al., 2008; Avila et al., 2022) y en particular B2i1a1 (B2i1a+161479A+34) al Alto Paraná en Paraguay (MH981885, Simão et al., 2019); B2h se ha encontrado en Aché de Paraguay (EU095206, Fagundes et al., 2008; KJ923805, Rieux et al., 2014) y en población cosmopolita de Brasil (Ávila et al., 2022) y Alto Paraná (Simão et al., 2021; 2019), B2za y en particular B2za1a se ha asociado a secuencias de Perú (Barbieri et al., 2017), Paraguay (Simão et al., 2019) y Chaco (Sevini et al., 2014).

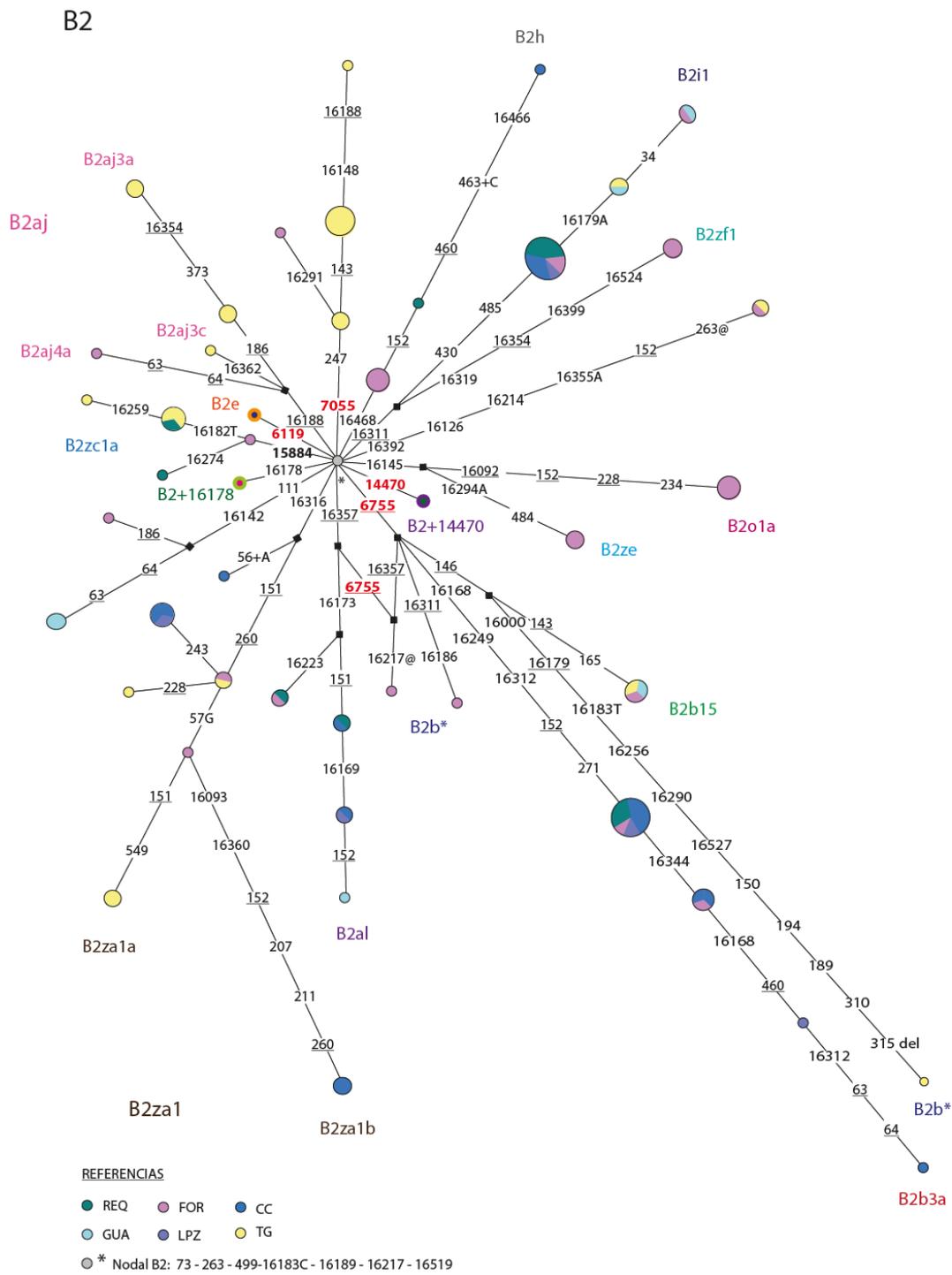


Figura 3.6 Red de haplotipos B2. Red principal. Detalles en la leyenda de la Figura 3.3

Como se mencionó anteriormente, tres de las ramas de la red principal de B2 se ampliaron en otras redes que se muestran en la Figura 3.7:

B2+14470

La primera red corresponde al linaje B2+14470, se realizó a partir de 30 secuencias y agrupa 14 haplotipos que representa el 17% de las secuencias de la red del haplogrupo B2. En particular, la posición diagnóstica 14470 es recurrente en clados asignables a B2. A partir de los análisis realizados se pudo diferenciar un clado monofilético con la combinación B2+14470₁₀₃₋₁₅₂₋₁₆₂₄₁, que caracteriza al linaje B2an definido por Figueiro et al. (2022) e identificado en mitogenomas de Uruguay y en un individuo Xavante de Brasil (Fagundes et al., 2008). Este clado se encontró representado por 24 secuencias pertenecientes a poblaciones exclusivamente del NEA y Litoral y agrupó ocho haplotipos, de los cuales cinco se encontraron en la localidad de Formosa. Además, se pudieron identificar clados de menor jerarquía que portan las mutaciones diagnósticas 146-455+T, como B2an2 y B2an2a (B2an2+16241@).

B2+16178

La segunda red amplía el linaje B2+16178 (B2zd), se construyó a partir de 35 secuencias y conforma 9 haplotipos. Es el linaje con mayor representación del haplogrupo B2, alcanzando frecuencias cercanas al 20% considerando el total de secuencias analizadas en las redes B2. El haplotipo nodal es el que se encontró en mayor frecuencia, agrupando 20 individuos pertenecientes a las localidades de Formosa y del NEA y Litoral presenta afinidades con secuencias de Brasil (Ávila et al., 2022). Se identificaron otros haplotipos que se definieron como pertenecientes a los clados menores B2zd2 (B2zd+152) y B2zd4 (B2zd+16354A), el primero asociado a una muestra de Tartagal y el segundo compartido por un mitogenoma de Paraguay (MH981921).

B2e

Por último, la tercera red que comprende el linaje B2e, se realizó a partir de 25 secuencias y agrupa 8 haplotipos. Este linaje se caracteriza por presentar las variantes diagnósticas 6119 y 14049. 6119 fue confirmada en todas las secuencias agrupadas en esta red mediante secuenciación. Además, se pudieron identificar otros clados de menor jerarquía como B2e1, definido por la presencia de 15672-15781 (García et al., 2021); y B2e2, diagnosticado por 986 y con una rama interna (B2e2a) definida por la variante 9899 mediante PCR-APLP.

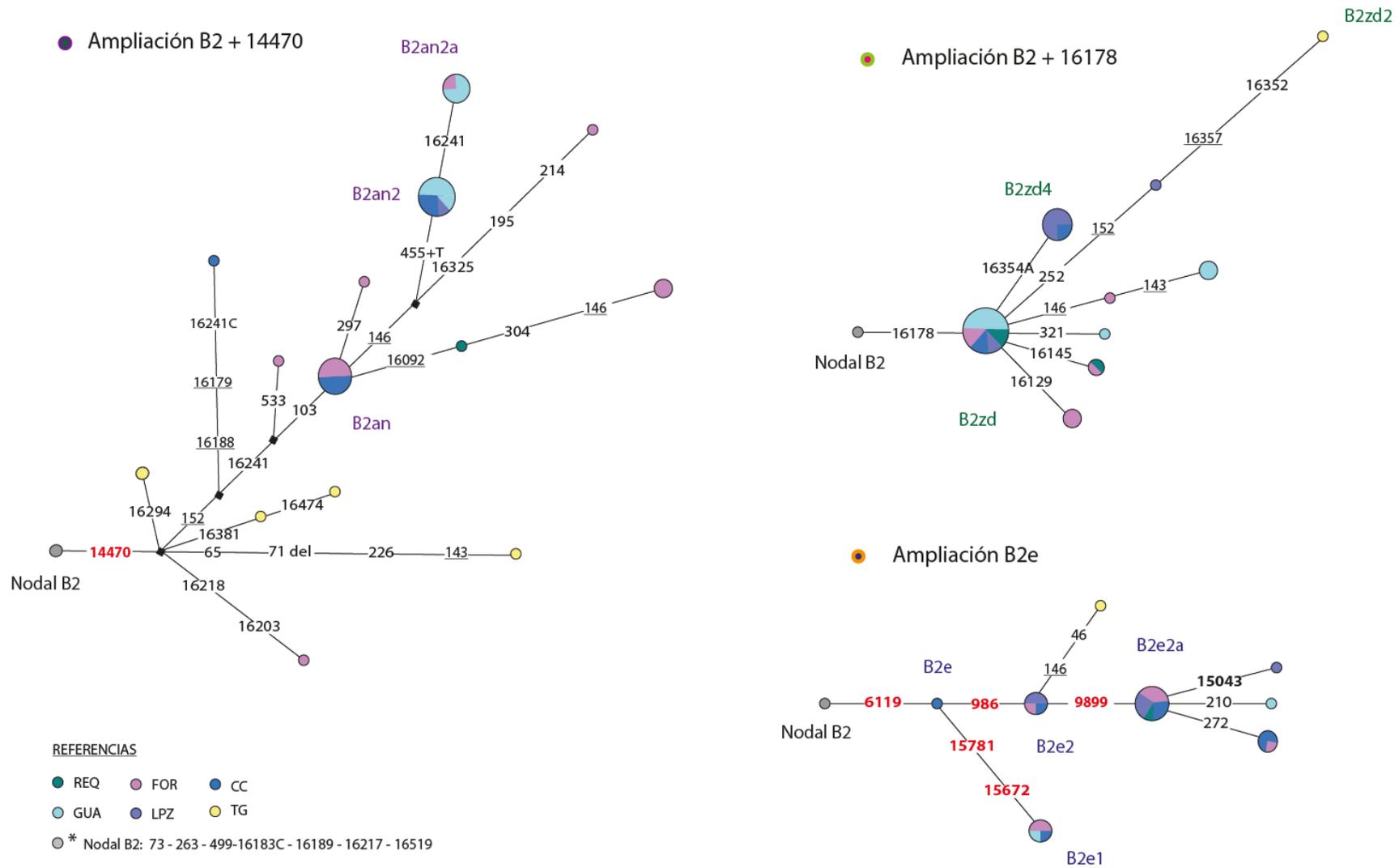


Figura 3.7. Ampliaciones de red de haplotipos B2: B2+14470, B2+16178 y B2e (Figura 3.6). Detalles en la leyenda de la Figura 3.3

3.2.1.3.2.3 Haplogrupo C1

El haplogrupo C1 se encuentra representado en el 27% de las secuencias de Formosa. En el resto de las localidades analizadas, disminuye sus frecuencias gradualmente desde el centro oeste del país (Mendoza, San Juan y La Rioja) hacia las provincias de Salta y Jujuy. Las redes correspondientes a C1 se realizaron agrupando las secuencias de toda la BBDD-RI en cada uno de los tres subhaplogrupos que la conforman.

Sub-haplogrupo C1b

Al subhaplogrupo C1b pertenecen el 66% de las muestras asignadas a C1 de la BBDD-RI. La red principal se realizó a partir de 120 secuencias que comprenden 44 haplotipos (Figura 3.8). Tiene forma estrellada y desde su secuencia nodal se separan 12 ramas principales, una de las cuales representa al clado C1b16 que se muestra en una red por separado (Figura 3.9).

Ciertos linajes identificados sólo estuvieron representados en secuencias pertenecientes a las localidades del NEA y Litoral, como C1b+16384 (C1b101), C1b+152 (C1b102), C1b+263@ (C1b2a1a) y C1b+16126 que agrupan 3, 6, 5 y 2 haplotipos, respectivamente. Otros, pudieron asociarse a regiones tanto de Argentina como de países aledaños. Por ejemplo, se encontraron dos haplotipos del clado C1b30 definido por Figueiro et al. (2022) como compartido por muestras uruguayas y chaqueñas. En particular, en esta red C1b30+16126 caracteriza al linaje C1b30a1, que agrupa 17 secuencias de Curuzú-Cuatiá, La Paz, Gualeguaychú y Formosa. Este mismo haplotipo se ha identificado en un mitogenoma del Chaco argentino (KP172342, Sevini et al., 2014) y en otro de Paraguay (MH981830, Simão et al., 2019).

Cuatro de las ramas principales de la red de C1b están definidas a través de polimorfismos diagnósticos de la región codificante que fueron encontrados por APLP: 4736, 5054C, 10029, 4676A y 12813. La rama que presenta los polimorfismos C1b+207-150-10029 caracteriza al linaje C1b17 (García et al., 2021), que se encuentra representada por dos haplotipos que agrupan secuencias de Formosa y Gualeguaychú con otras de Chepes, Santiago del Estero y Tucumán. A su vez, la rama que presenta la variante diagnóstica 4736 caracteriza al linaje C1b18 (García et al., 2021) agrupando el 24% de las secuencias analizadas en esta red y conformando un total de 11 haplotipos, ocho de los cuales presentan la variante 5054C que caracteriza al clado C1b18a y que fue compartido principalmente por secuencias tanto de Formosa como de otras localidades del NEA y Litoral. Por otra parte, este clado permite establecer afinidades con el sur de Brasil (Ávila et al., 2019; Ávila et al., 2022) y Paraguay (MH981924, MH981908; Simão et al., 2019). Otro haplotipo presente en una secuencia de

Formosa fue identificado como C1b18b (C1b18+16187) (García et al., 2021), reforzando el vínculo con Paraguay (MH981889, MH981891, MH981918, Simão et al., 2019). Finalmente, la rama que se define por el polimorfismo 4676A, agrupa 8 haplotipos correspondientes al linaje C1b19 (García et al., 2021). La muestra de Formosa identificada como C1b19b (C1b+152-143-16266) se agrupó con secuencias de Tucumán y Santiago del Estero.

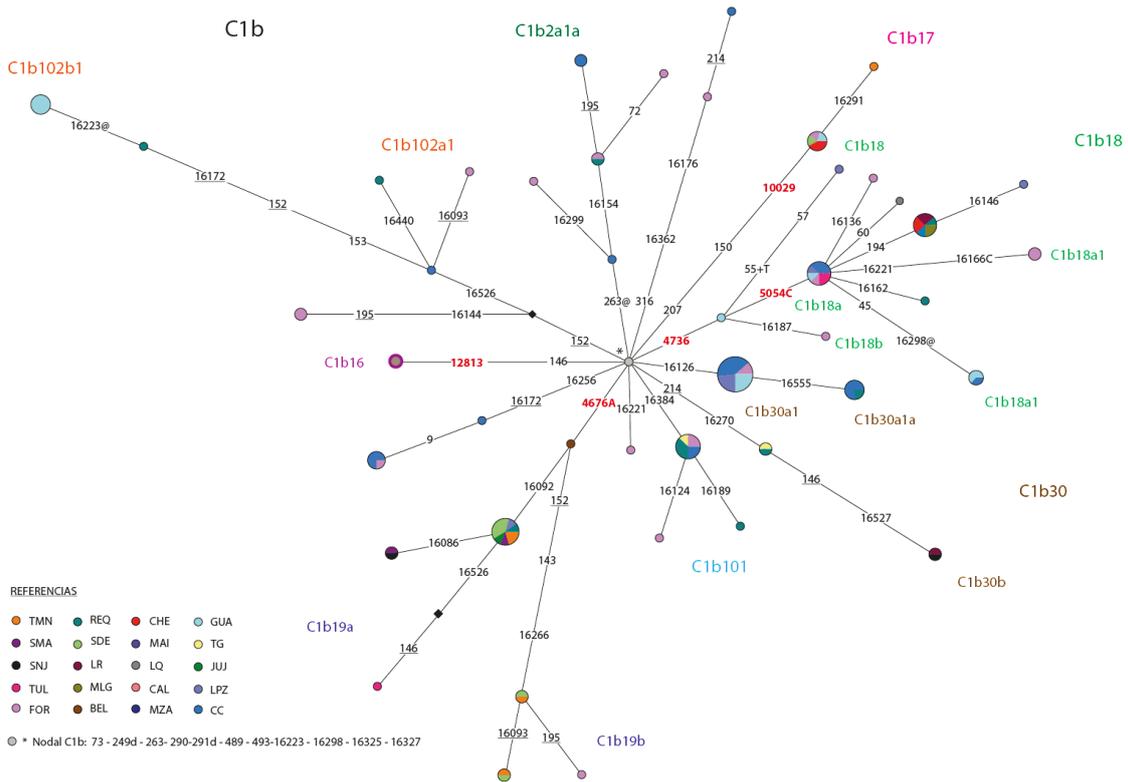


Figura 3.8 Red de haplotipos C1b. Red principal. Detalles en la leyenda de la Figura 3.3

Linaje C1b16

La red del linaje C1b16 se realizó a partir del conjunto de 82 secuencias y agrupa 20 haplotipos (Figura 3.9). Este linaje se caracteriza por las variantes C1b + 12813-146 y fue definido por García et al. (2021) como un linaje propio del Centro y Centro-Oeste Argentino pero que ocasionalmente puede encontrarse en otras regiones del país e incluso en Bolivia, en Gran Chaco (Sevini et al., 2014) y en Uruguay (MW057664, MW057665, MW057663; Figueiro et al., 2022). En esta red se observa un haplotipo central de C1b16 de gran representación en las secuencias analizadas y entre las que se encuentran las cuatro muestras asignadas a este linaje de Formosa entre otras de Curuzú-Cuatiá, Tartagal y de localidades del NOA. Sin embargo, el haplotipo de más alta frecuencia es el que porta la variante 14587 diagnóstica de C1b16a que está representado en más de un 50% por muestras de Tucumán, Santiago del Estero y Tartagal.

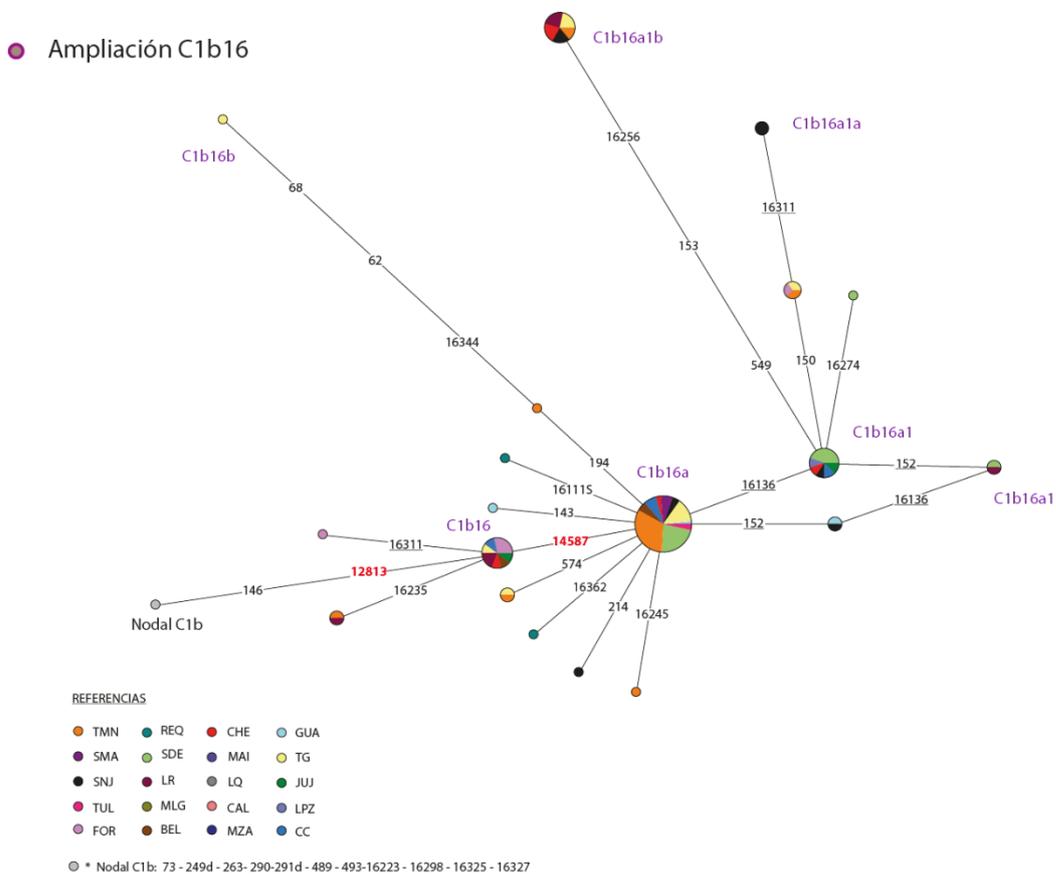


Figura 3.9. Ampliación de red de haplotipos C1b. Linaje C1b16. La figura se desprende de la red principal de C1b (Figura 3.14). Detalles en la leyenda de la Figura 3.3

✚ Sub-haplogrupo C1c

El subhaplogrupo C1c, definido por la presencia de 1888-15930, agrupa el 12% de las muestras asignadas a C1 de la BBDD-RI. La red se realizó a partir de 45 secuencias que comprenden 20 haplotipos (Figura 3.10). Una única secuencia de Formosa pudo ser asignada a este subhaplogrupo y se encuentra en la rama que porta las variantes C1c+16311-152-151-16390-16368. Este clado fue denominado como C1c100 y agrupa 6 haplotipos. Además, se pudieron definir otros clados menores como C1c100a1 (C1c100+214) en una única secuencia de Santiago del Estero y C1c100a2 (C1c100+16071) representado por cuatro haplotipos en dos secuencias de Tartagal y en una secuencia de Formosa, Santiago del Estero y Tucumán.

Por otra parte, la rama principal que porta las variantes 16189 y 195 conduce a un haplotipo que caracteriza al clado C1c101 presente principalmente en secuencias de Santiago del Estero, pero también de Tartagal, Tucumán y Reconquista y que se ha encontrado en un mitogenoma del Chaco argentino (KP172360, Sevini et al., 2014).

Finalmente, otra rama principal de la red C1c que se caracteriza por la variante 16169 pudo ser identificada como una rama interna del linaje C1c9 (C1c+14581) denominada C1c9a (C1c+14581-1019-16169-16362) (García et al., 2021) agrupando 8 haplotipos. C1c9a+16145 presenta afinidades con el Centro de Argentina, mientras que C1c9a+16192A-16209 (C1c9a1) ha sido definida por García et al. (2021) como exclusiva del Gran Chaco ya que se encontró en población Wichí del Chaco argentino (KP172354, KP172346, KP172347, Sevini et al., 2014). En esta red se identificaron cuatro haplotipos con esas variantes, tres de ellos presentes en secuencias de Tartagal y uno en Maimará.

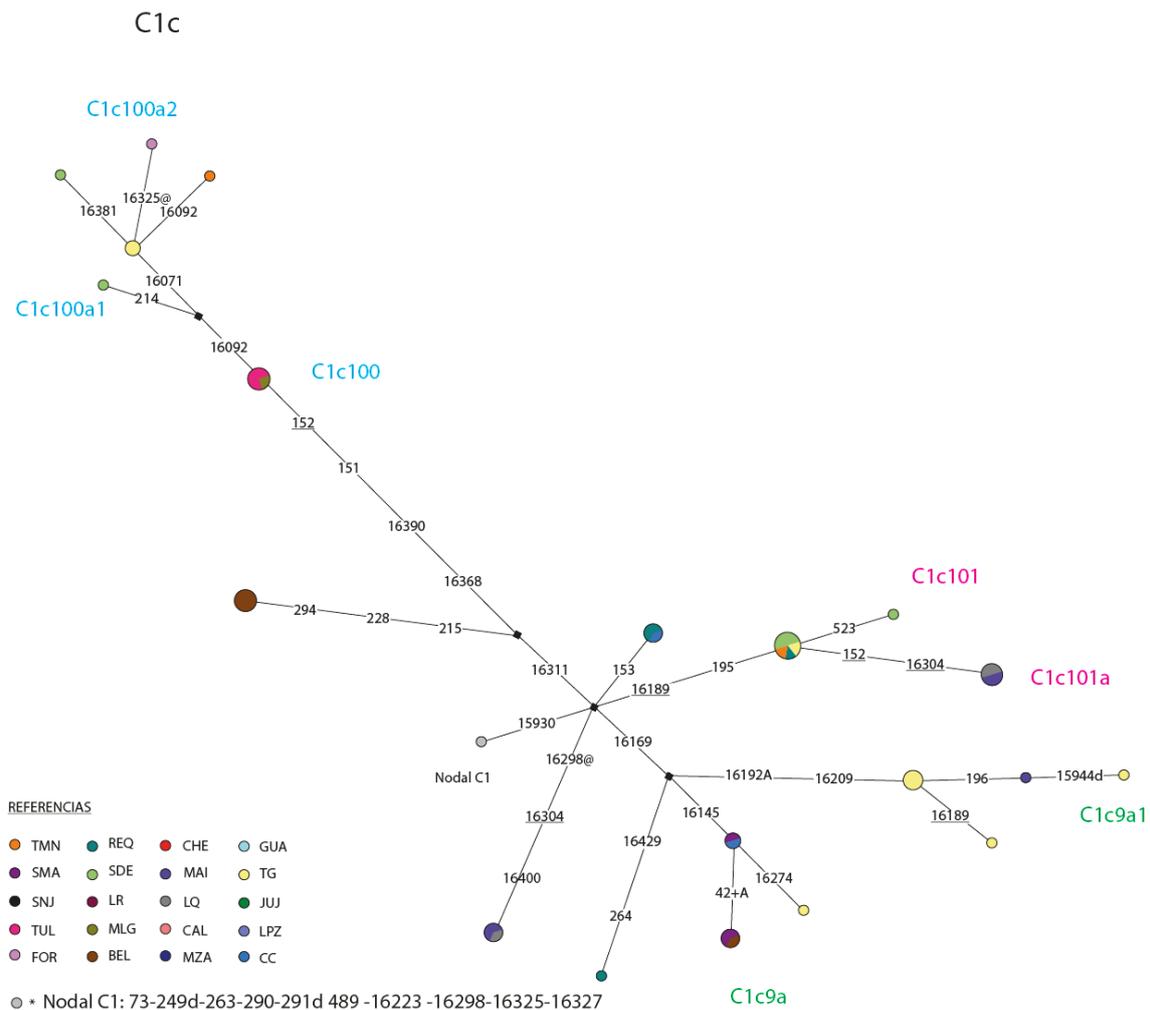


Figura 3.10. Red de haplotipos C1c. Detalles en la leyenda de la Figura 3.3

✚ Sub-haplogrupo C1d

El sub-haplogrupo C1d representa el 22% de las muestras asignadas a C1 de la BBDD-RI y se caracteriza por la presencia de las variantes 194 y 16051. La red se realizó a partir de 84 secuencias que comprenden 22 haplotipos (Figura 3.11). El clado C1d+16172, que caracteriza al linaje C1d1d, se encuentra representado en alta frecuencia, agrupando 13 individuos en un haplotipo central presente en secuencias de Formosa, Curuzú-Cuatiá, Reconquista y San Juan. Otra rama que se desprende de este haplotipo presenta la variante 310 y conforma otro haplotipo presente en dos individuos de Formosa.

A partir de la búsqueda de PNS en región codificante se lograron identificar 11 haplotipos pertenecientes al clado C1d1+928 (C1d1b), a su vez, la presencia de la variante 7337 hizo posible asignar a 7 de estos haplotipos como pertenecientes al clado C1d1b1. En particular, C1d1b1+204 (C1d1b1a, García et al., 2021) fue encontrado en alta frecuencia agrupando secuencias de Santa

María, Santiago del Estero, Formosa, Tucumán, Tartagal y San Juan; mientras que C1d1b1+16129 agrupó secuencias correspondientes Formosa, Tucumán y Santiago del Estero. Por otra parte, C1d1b2 (C1d1b+16311-16271-16259; García et al., 2021) se encontró representado por dos haplotipos, agrupando secuencias pertenecientes a localidades de Santiago del Estero, Tucumán, Formosa, Curuzú-Cuatiá, La Quiaca, Maimará y Tartagal. Este clado se ha encontrado en dos mitogenomas de Santiago del Estero del trabajo de García et al. (2021), en otro de Salta (HM107323, Perego et al., 2010) y en otro de población wichí del Chaco (KP172348, Sevini et al., 2014). Otros clados identificados en menor frecuencia que también se encontraron en secuencias de Formosa fueron C1d1aa2a2, C1d1e y C1d1ab1.

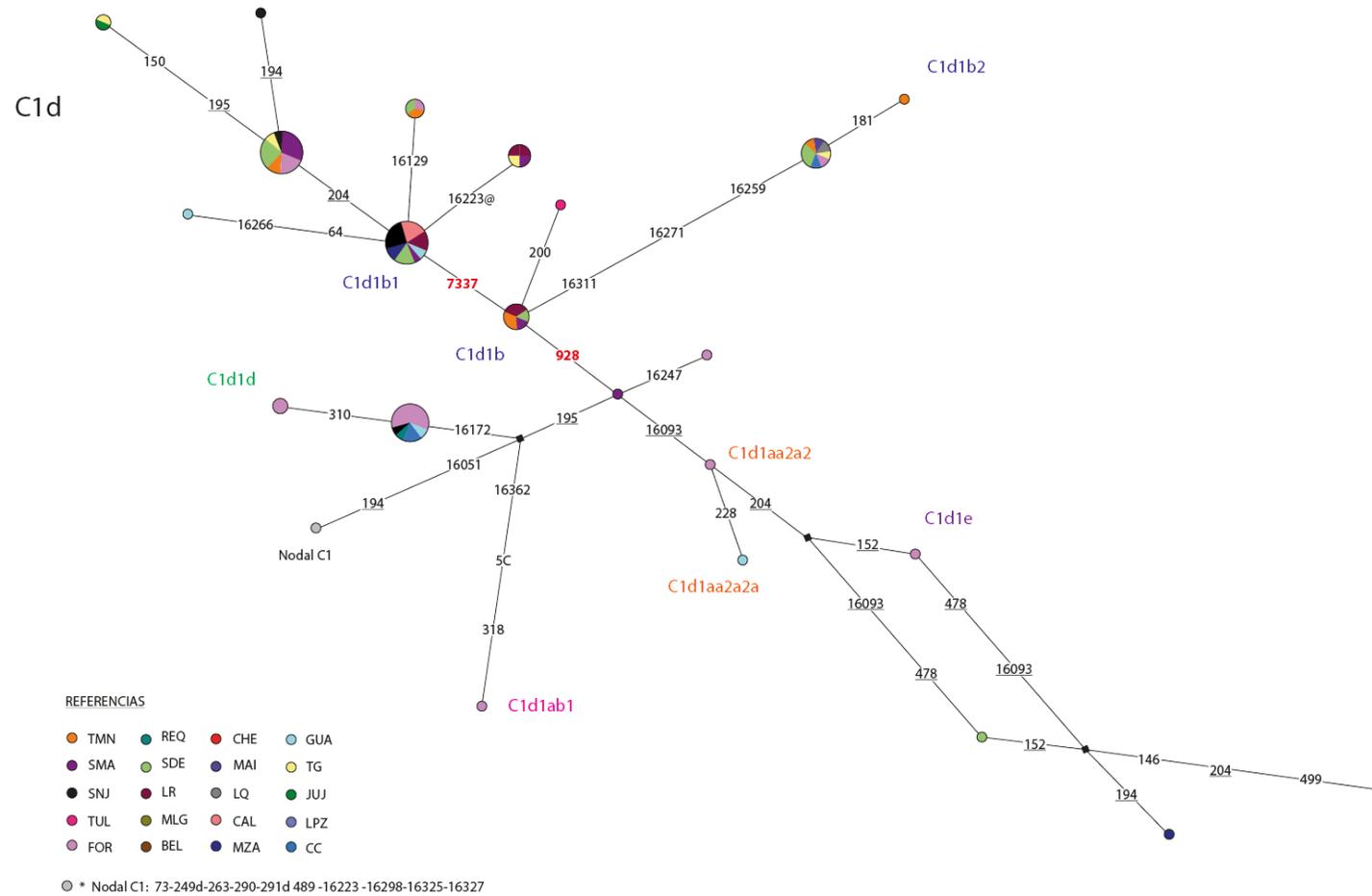


Figura 3.11. Red de haplotipos C1d. Detalles en la leyenda de la Figura 3.3

3.2.1.3.2.4 Haplogrupo D1

El haplogrupo D1 se encuentra representado por el 17% de las secuencias asignadas a haplogrupos nativos de la BBDD-RI. La red D1 se realizó a partir de 69 secuencias y agrupa 27 haplotipos (Figura 3.12). Tiene forma estrellada y de la secuencia nodal se separan 9 ramas principales.

La rama D1+16142-16179-16497 que caracteriza al clado D1f1 agrupa 5 haplotipos de los cuales tres se encuentran representados en secuencias de Formosa, además de otro haplotipo que se consideró pre-D1f1 por no presentar la variante 16497. Este linaje está ampliamente distribuido en mitogenomas de poblaciones indígenas de Brasil (waiwai: EU095235, katuena: EU095238 y tyrió: EU095240, EU095241; Fagundes et al., 2008) y se lo ha reportado en secuencias de Paraguay (Simão et al., 2019) y de Venezuela (Brandini et al., 2018). En este trabajo se lo encontró principalmente en secuencias del NEA y Litoral.

D1zd (D1+16325@-16291) se ha encontrado en cuatro haplotipos compartidos por secuencias de Formosa y Tartagal, y en un criollo del Chaco (KP172395; Sevini et al., 2014).

D1+16187A agrupa a un haplotipo que se encontró representado en secuencias de Formosa, Curuzú-Cuatiá, Reconquista y Gualaguaychú y fue denominado como D1zc. Además, se ha encontrado en mitogenomas de Uruguay (MW057681, Figueiro et al., 2022) y de Brasil (PUC633, Ávila et al., 2022).

En la rama D1+146 se pudieron identificar dos clados: D1e1 (D1+146+273) presente exclusivamente en secuencias de Formosa, y D1za (D1+146-13989-6305T-143) asociado a secuencias de Formosa, Curuzú-Cuatiá y Tartagal. Estos linajes, permitieron establecer afinidades con poblaciones de Paraguay (D1e1: MH981846 y D1za: MH981835, MH981886) (Simão et al., 2019).

Algunas de las ramas fueron caracterizadas por polimorfismos de la región codificante. Una de ellas es D1j1+5004; esta variante es diagnóstica del clado D1j1a y agrupa a un haplotipo representado por muestras de Gualaguaychú y Formosa. Este clado se lo ha encontrado en un mitogenoma de Uruguay (MW057676, Figueiro et al., 2022), en criollos del Chaco (KP172370, KP172400; Sevini et al., 2014) y en secuencias de Santiago del Estero y Tucumán (Schwab, 2018). A su vez el haplotipo D1j1a+146 representado en individuos de Gualaguaychú y de La Paz también fue compartido por un criollo del Chaco (KP172382, Sevini et al., 2014). Por otra parte, D1j1a1 (D1j1a+5325T) agrupa dos haplotipos, uno presente en secuencias de Formosa y Tartagal y otro con la variante 16172 en una muestra de Formosa.

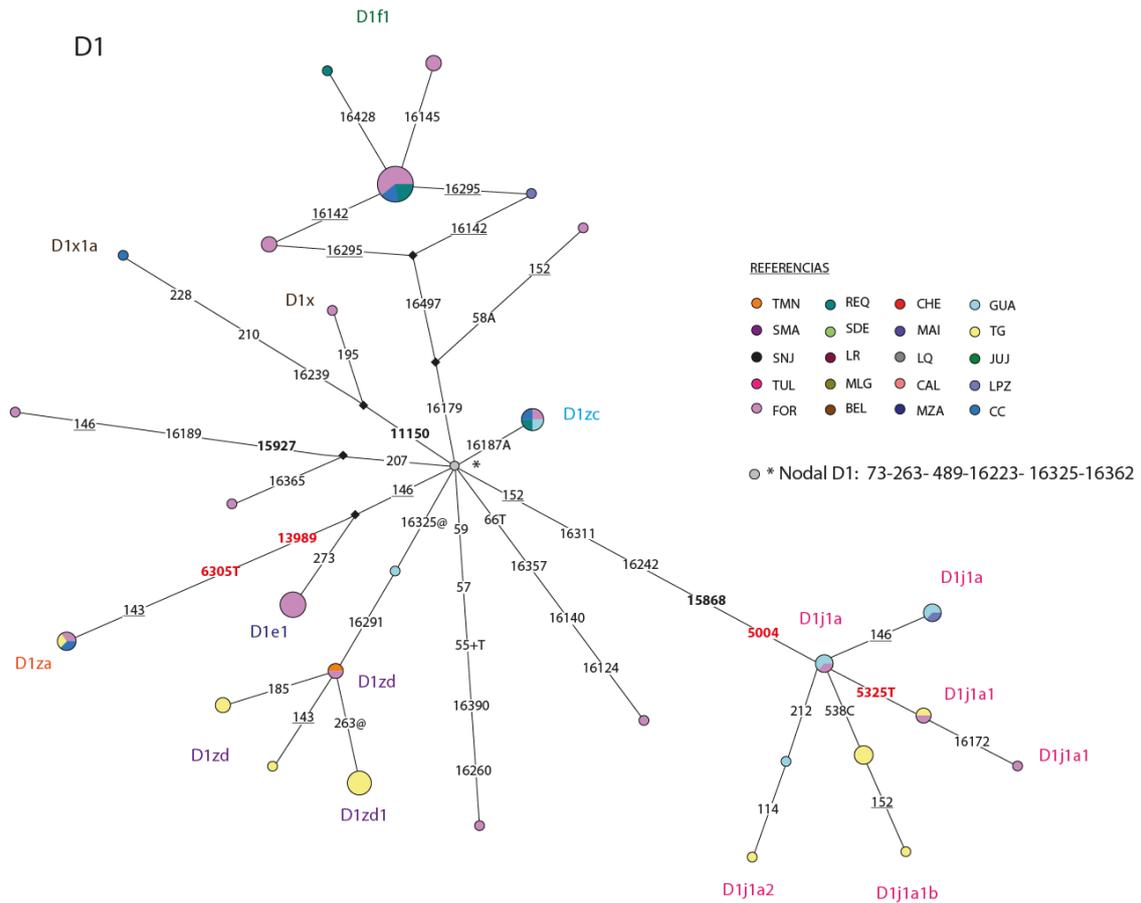


Figura 3.12 Red de haplotipos D1. Detalles en la leyenda de la Figura 3.3

3.2.1.3.2.5 Haplogrupo D4h3a

En este trabajo se han definido tres nuevos subhaplogrupos de D4h3a: D4h3a90 (D4h3a+16296) presente en una muestra de Reconquista que comparte haplotipo con un mitogenoma de Paraguay (MH981882, Simão et al., 2019); D4h3a91 (D4h3a+106-111d) hallado en una muestra de Formosa compartido por una secuencia de Paraguay (LS186) informada por Strobl et al. (2019), y D4h3a92 (D4h3a+16114A-16288) presente en una muestra de Formosa y compartido por un mitogenoma de Brasil (PUC601, Ávila et al., 2022) (ver Anexo 2.9).

3.2.1.4 Identificación de clados mitocondriales en la ciudad de Formosa y su presencia en otras localidades y a escala regional

3.2.1.4.1 Clados propios de la localidad de Formosa

A modo de visualizar de forma más resumida la información contenida en el Anexo 3.3 se construyó una tabla en la que se calculó la frecuencia de linajes identificados en la muestra de la ciudad de Formosa (en %) y su ocurrencia en localidades del NEA, Tartagal, NOA y Cuyo. Particularmente, se observa que la localidad de Formosa se caracteriza por una gran diversidad de clados mitocondriales (N = 61) (Anexo 3.4). Además, a través de los totales parciales de la tabla mencionada, se puede vislumbrar que los linajes identificados representan cerca del 80% de la diversidad encontrada para las secuencias recopiladas de la región del NEA y un 57% de la muestra de Tartagal, mientras que para las regiones del NOA y Cuyo representan frecuencias menores al 50% (45% y 31%, respectivamente) (Anexo 3.4).

3.2.1.4.2 Definición de subhaplogrupos mayoritarios en Formosa

A fin de determinar la presencia de linajes monofiléticos propios de Formosa en otras localidades de Argentina y a escala regional, se exploraron las frecuencias de los subhaplogrupos encontrados en Formosa, observando un punto de quiebre entre ellas en un valor de 1,7%. Este punto permitió diferenciar entre dos categorías de subhaplogrupos: los subhaplogrupos minoritarios (en morado), que se definieron como aquellos que presentaron frecuencias menores al 2%, y los subhaplogrupos mayoritarios (en gris) con frecuencias de aparición mayores al 2% (Figura 3.13).

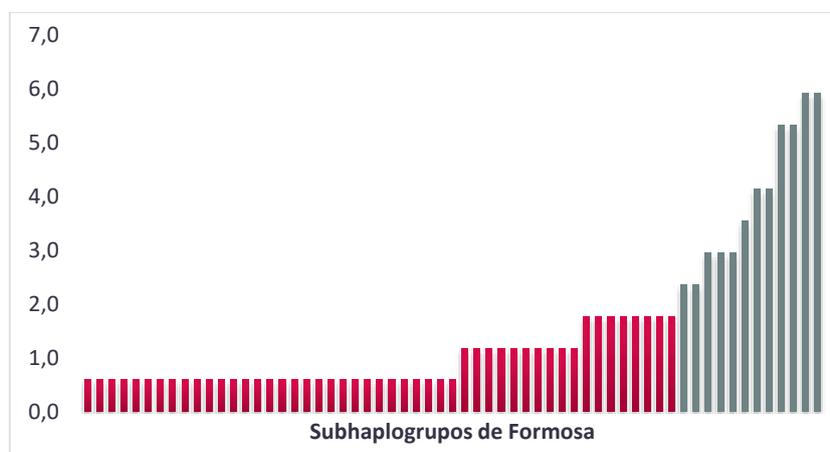


Figura 3.13 Histograma de frecuencias de subhaplogrupos encontrados en la localidad de Formosa

Es interesante resaltar que de los 61 clados encontrados, 49 están representados en la muestra de Formosa en frecuencias menores al 2%, agrupando el 44% de la diversidad total. Por otra parte, 48% de la diversidad restante está constituida por 12 subhaplogrupos mayoritarios (Figura 3.14). Por último, el 8% de los haplotipos encontrados no pudieron asignarse a ningún clado específico (haplotipos no asignados).

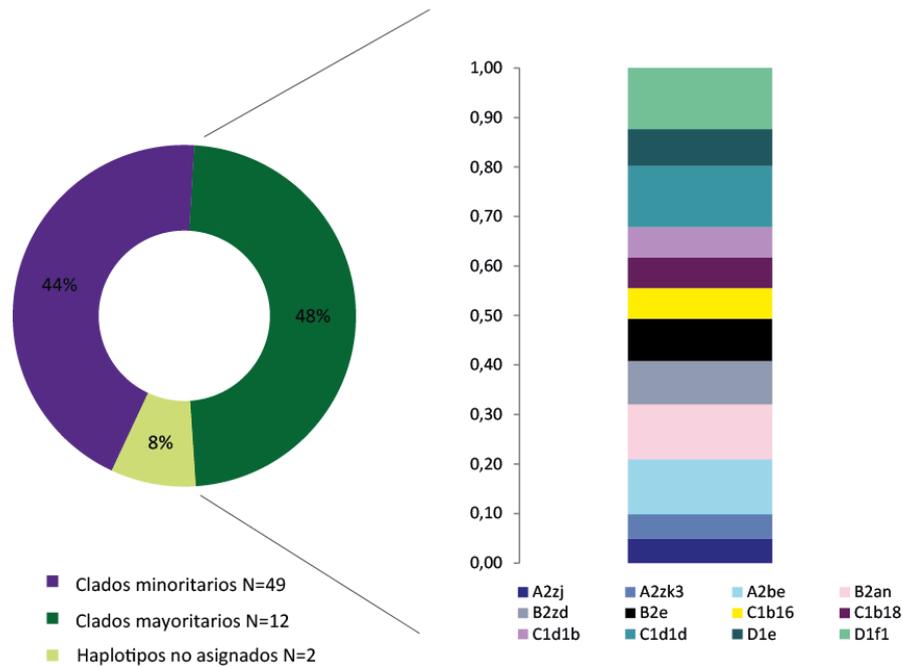


Figura 3.14. Diversidad de clados mitocondriales de la ciudad de Formosa

3.2.1.4.3 Presencia de subhaplogrupos mayoritarios en otras localidades y a escala regional

Mediante comparaciones con las bases de datos construidas (BBDD-RI: Tabla 2.15 y BBDD-RB: Tabla 2.16), se pudo identificar que los 12 subclados mayoritarios encontrados en Formosa (A2zj: A2₍₆₄₎+338-3438-16468, A2zk3: A2+16266-10237-12858-14364, A2be, B2an, B2zd: B2+16178, B2e, C1b16, C1b18, C1d1b, C1d1d, D1e y D1f1) se distribuyen en frecuencias y combinaciones variables en las localidades estudiadas (Anexo 3.5, Figura 3.15.A) y a escala regional (Figura 3.15.B). En particular, combinaciones de 6 a 10 subhaplogrupos mayoritarios constituyen entre el 25% y el 40% de la variabilidad encontrada en las muestras de localidades del NEA y Litoral como Reconquista, Curuzú Cuatiá, La Paz y Gualedguaychú y de Brasil, Uruguay y Paraguay; contrariamente a lo que ocurre en localidades del NOA (La Quiaca, Maimará y Jujuy), Cuyo y CENTRO en donde se observan combinaciones de 1 a 4 subhaplogrupos (como C1d1b,

C1b16, C1d1d y C1b18), representando menos del 20% de la variabilidad encontrada (figura 3.16.A). Un caso excepcional se da en Santiago del Estero y Tucumán en donde se encontraron 4 (C1d1b, C1b16, C1d1d y B2an) y 2 (C1b16 y C1d1b) de estos clados respectivamente, agrupando alrededor del 30% de la diversidad en estas localidades (Figura 3.15.A). A su vez, un dato significativo es que 10 de los subhaplogrupos descritos están presentes en muestras de Paraguay, alcanzando el 25% de la diversidad, mientras que, en el Chaco argentino, con menor diversidad de linajes, no alcanza el 15% (Figura 3.15.A). En este sentido, se observa que la columna correspondiente al Chaco argentino se asemeja más a Tartagal en cuanto al número de los subhaplogrupos mayoritarios presentes (cinco de ellos son compartidos por ambas localidades: A2zj, B2zd, B2e, C1b16 y C1d1b) y en relación a la frecuencia alcanzada en comparación con Formosa, el resto de las localidades del NEA y Litoral y los países limítrofes incluidos.

Si se toma en consideración una escala regional (que incluye países limítrofes) para visualizar la distribución de los subhaplogrupos mayoritarios presentes en la muestra de Formosa, se observa que estos últimos representan el 37% de la diversidad total del NEA y Litoral y entre el 10 al 12% de la diversidad de Cuyo, NOA y CENTRO (Figura 3.15.B). Es interesante destacar que los linajes presentes en Uruguay, alcanzan cerca del 35% de la diversidad, mientras que los de Paraguay y Brasil acumulan frecuencias del 25%; además se presentan en combinaciones de 6, 10 y 10 subhaplogrupos respectivamente (Figura 3.15.B).

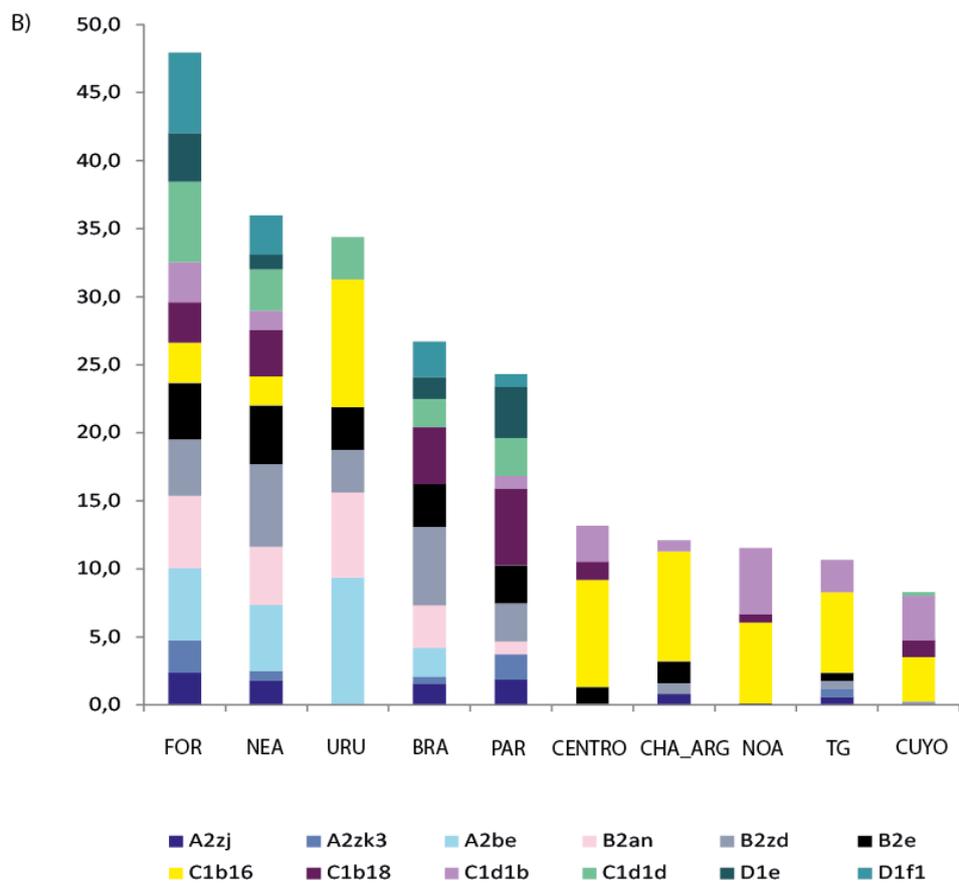
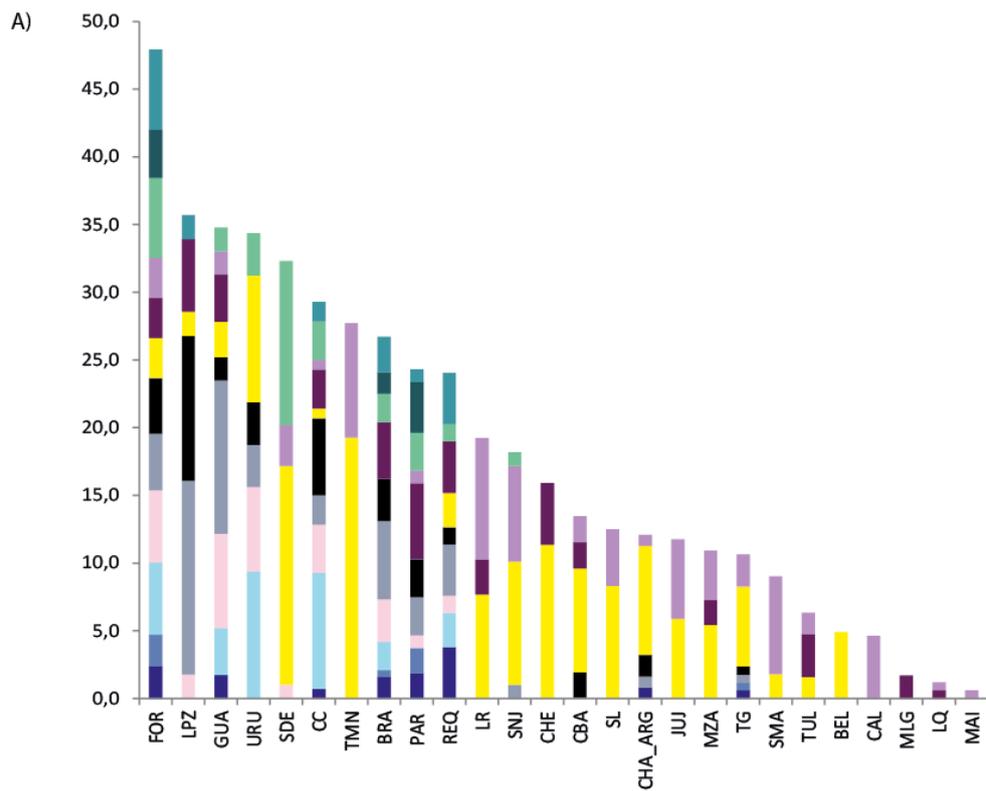


Figura 3.15. Distribución de clados mayoritarios de Formosa por localidad A) y por región B), incluyendo países limítrofes.

3.2.1.5 Análisis de las distancias genéticas entre poblaciones

La identificación de linajes monofiléticos y el cálculo de su frecuencia relativa en las localidades analizadas de la BBDD-RI (Anexo 3.2), fueron empleados para el cálculo de las distancias de Nei entre pares de poblaciones y la construcción de una matriz de distancias (Tabla 3.9), a partir de la cual, se realizaron un análisis de escalamiento multidimensional (MDS), y la identificación de barreras geográficas entre las poblaciones analizadas.

Tabla 3.9. Distancias de Nei calculadas entre pares de poblaciones a partir de la frecuencia de clados nativos

	LQ	MM	JUJ	TG	BEL	SMA	TMN	SDE	LRI	CHE	SNJ	CAL	TUL	MZA	MLG	GUA	LPZ	CC	REQ	FOR	
LQ	0																				
MM	0,152	0																			
JUJ	0,561	0,638	0																		
TG	1,023	1,132	0,766	0																	
BEL	0,849	0,599	0,580	1,388	0																
SMA	0,777	0,624	0,451	0,985	0,266	0															
TMN	3,024	2,588	0,609	0,703	1,037	0,604	0														
SDE	2,462	2,252	0,566	0,742	1,111	0,613	0,079	0													
LRI	3,109	2,737	0,693	1,089	0,994	0,515	0,201	0,333	0												
CHE	3,498	3,096	0,857	1,125	1,011	0,695	0,220	0,397	0,137	0											
SNJ	2,488	2,273	0,856	0,884	1,130	0,480	0,294	0,350	0,436	0,424	0										
CAL	3,390	2,682	1,748	2,024	1,976	0,928	1,244	1,380	1,173	0,970	0,450	0									
TUL	2,112	1,891	1,597	1,514	1,915	1,346	1,102	1,116	1,100	0,889	0,712	1,019	0								
MZA	1,491	1,388	0,764	0,991	0,999	0,675	0,443	0,502	0,371	0,382	0,460	1,020	0,506	0							
MLG	2,481	2,214	2,419	2,283	2,760	2,130	2,369	2,584	1,886	1,665	1,391	0,696	1,069	0,800	0						
GUA	3,654	3,705	1,662	1,411	2,209	1,678	1,207	1,235	1,241	1,153	1,152	1,906	1,763	1,453	3,061	0					
LPZ	3,136	3,758	1,672	1,502	3,139	2,846	2,130	1,961	2,163	2,084	1,996	3,802	2,693	2,238	3,437	0,443	0				
CC	3,020	3,159	3,801	1,698	4,065	3,572	2,987	2,563	3,086	2,751	2,147	4,261	2,505	2,410	3,705	0,663	0,599	0			
REQ	3,492	3,335	2,222	1,352	3,480	3,409	2,147	1,903	2,970	2,443	2,216	4,680	2,476	2,293	3,493	0,706	0,691	0,511	0		
FOR	2,805	2,911	1,726	1,390	2,679	1,956	1,377	1,262	1,690	1,685	1,427	2,643	2,092	1,715	3,326	0,559	0,759	0,491	0,677	0	

3.2.1.5.1 Análisis de escalamiento multidimensional a partir de frecuencias haplotípicas

Se realizó un análisis de escalamiento multidimensional que permitió representar en un espacio geométrico de dos dimensiones las proximidades existentes entre las poblaciones estudiadas (Figura 3.16). Se obtuvo un valor de stress final = 0,15 que sugiere un ajuste aceptable según el rango propuesto por Kruskal (1964a,b). En el ordenamiento de las poblaciones se puede observar un cline este-oeste y cierto cline norte-sur. Se destaca principalmente el hecho de que Formosa se encuentre agrupado al resto de las poblaciones del NEA (Curuzú Cuatiá, Reconquista, La Paz y Gualeguaychú) y se separe de las poblaciones que corresponden al NOA y a Cuyo.

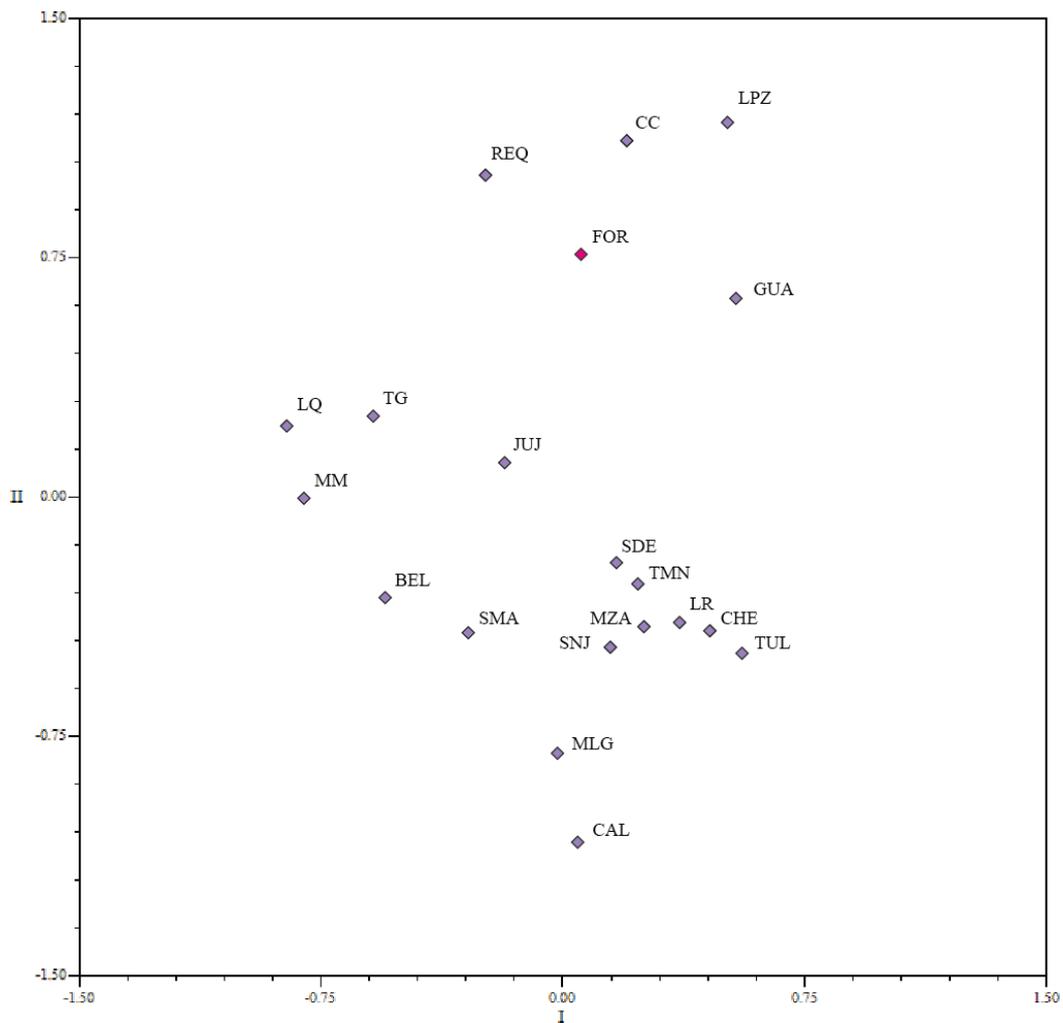


Figura 3.16. Análisis de escalamiento multidimensional calculado a partir de las frecuencias de los 201 linajes de ADNmt en 20 poblaciones del NEA, NOA y COA.

3.2.1.5.2 Identificación de barreras genéticas

Utilizando el programa Barrier, se diagramaron 4 barreras genéticas (Figura 3.17) en base a las distancias de Nei generadas a partir de las frecuencias haplotípicas (Anexo 3.3). La primera barrera (barrera a) divide las localidades del NEA del resto de las localidades del NOA y COA. La segunda barrera (b) aísla a la localidad de Calingasta, mientras que lo mismo ocurre con la barrera c y Tartagal. Por último, la cuarta barrera (d) separa las localidades del NOA del COA.

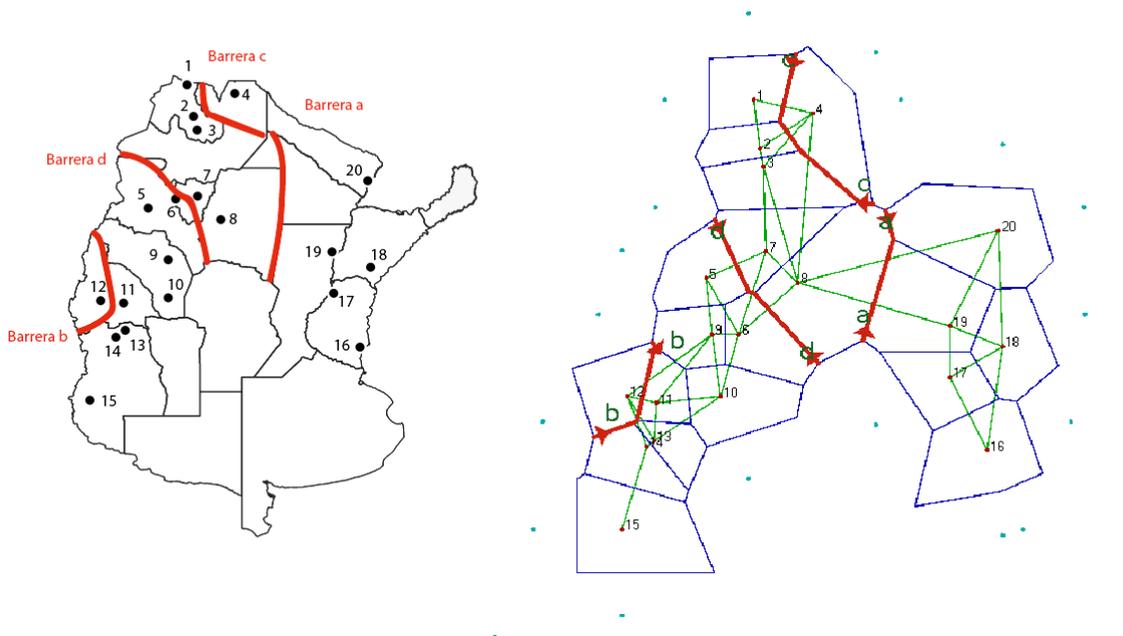


Figura 3.17 Barreras genéticas identificadas con el programa Barrier 2.2. De derecha a izquierda: 1. Mapa geométrico con las barreras propuestas por el programa Barrier 2.2. 2. Barreras superpuestas en un mapa de Argentina con división política. Localidades: 1 = LQ, 2 = MM, 3 = JUJ, 4 = TG, 5 = BEL, 6 = SMA, 7 = TMN, 8 = SDE, 9 = LR, 10 = CHE, 11 = SNJ, 12 = CAL, 13 = TUL, 14 = MZA, 15 = MLG, 16 = GUA, 17 = LPZ, 18 = CC, 19 = REQ, 20 = FOR. Las letras a-d indican el orden de prioridad de las barreras.

3.2.1.5.3 Análisis Molecular de la Varianza (AMOVA)

Con el objetivo de analizar la estructuración de las poblaciones, se ensayaron tres AMOVAs (Excoffier et al., 1992) con el programa Arlequín. Se estableció el nivel de significación $p < 0,01$ para 10.000 permutaciones. Se consideraron distintos agrupamientos de las poblaciones, a fin de analizar la distribución de la variabilidad genética entre diferentes niveles jerárquicos (localidad y región)

En los cuatro ensayos el grupo 1 incluyó a las localidades del NEA y Litoral, el grupo 2 a La Quiaca y Maimará como grupo diferenciado del NOA, en concordancia con lo informado por Motti et al. (2013), el grupo 3 contuvo al resto de las localidades del NOA y el grupo 4 a las localidades del COA (Tabla 3.10). En particular, en el ensayo I se colocó a Formosa en conjunto con las localidades del NEA y Litoral, en el ensayo II se consideró a Formosa con las localidades del NOA, en el ensayo III se evaluó a Formosa como grupo separado (grupo 5) y en el ensayo IV se consideró a Formosa como grupo separado con Tartagal (Tabla 3.10).

Tabla 3.10. Definición de grupos en los AMOVA ensayados.

Ensayos	Grupos				
	1	2	3	4	5
I	LPZ, CC, GUA, REQ, TG, FOR	LQ, MAI	JUJ, TMN, SDE, SMA, BEL	LRI, CHE, SNJ, CAL, MZA, MLG, TUL	
II	LPZ, CC, GUA, REQ, TG	LQ, MAI	JUJ, TMN, SDE, SMA, BEL, FOR	LRI, CHE, SNJ, CAL, MZA, MLG, TUL	
III	LPZ, CC, GUA, REQ, TG	LQ, MAI	JUJ, TMN, SDE, SMA, BEL	LRI, CHE, SNJ, CAL, MZA, MLG, TUL	FOR
IV	LPZ, CC, GUA, REQ	LQ, MAI	JUJ, TMN, SDE, SMA, BEL	LRI, CHE, SNJ, CAL, MZA, MLG, TUL	TG, FOR

Considerando los componentes de la varianza y los porcentajes de variación obtenidos en los cuatro ensayos, se puede concluir que el agrupamiento que maximiza la variabilidad entre los grupos y minimiza la variabilidad entre poblaciones dentro del grupo es el propuesto para el ensayo I, ubicando a Formosa con el resto de las poblaciones del NEA y Litoral (Tabla 3.11). En este ensayo 2,19% de la variabilidad observada se explica por diferencias entre grupos y 2,58% corresponde a la variabilidad hallada entre poblaciones pertenecientes al mismo grupo. El 95,23% restante se explica por variabilidad intrapoblacional y se considera un porcentaje esperado en cualquier población humana. Los tres valores resultaron significativos ($p < 0,01$).

Tabla 3.11. Tabla AMOVA para los ensayos I, II y III.

Porcentaje de variación (%)			
	Entre grupos	Entre poblaciones dentro del grupo	Dentro de las poblaciones
I	2,19	2,58	95,23
II	1,55	3,05	95,39
III	1,80	2,79	95,40
IV	1,90	2,68	95,42

3.3 Tercer nivel de análisis: distribución sudamericana de clados mitocondriales

Se indagó sobre la presencia de ciertos clados mitocondriales identificados en el segundo nivel de análisis a fin de establecer su distribución a nivel regional y sudamericano. Los linajes que se desarrollarán en profundidad en esta sección son los siguientes: A2bf, B2e, B2an, D1e y D1f1.

3.3.1 Linaje A2bf

A2bf es un subhaplogrupo mitocondrial descrito por primera vez recientemente, cuya distribución fue asociada a secuencias mitogenómicas uruguayas y chaqueñas (Figueiro et al., 2022). Su presencia puede ser detectada tanto en mitogenomas como en secuencias de la región control a partir de la variante diagnóstica 16051.

En este trabajo, se pudieron identificar 224 secuencias como pertenecientes al subhaplogrupo A2bf, de las cuales 209 fueron secuencias de región control (Anexo 3.6) y 15 fueron mitogenomas (Anexo 3.7: mitogenomas 1-15). En la figura 3.18 se puede observar la distribución geográfica del linaje A2bf. Es interesante resaltar que cuando se tienen en cuenta secuencias de la región control, es posible detectar la presencia de A2bf en un rango mucho más amplio de distribución, que involucra -excepto Misiones- todas las localidades de la región del NEA y Litoral, ciertas localidades del NOA (Santiago del Estero, Tucumán, Salta y Jujuy), Paraguay, Bolivia, Chile, Venezuela y la Guayana francesa en su límite norte; Buenos Aires y Río Negro en su límite sur (Figura 3.19). Resulta necesario aclarar que la presencia de 16051 sobre linajes A2 de Centro y Norteamérica no debería ser interpretada como A2bf, ya que la existencia de al menos cuatro mitogenomas permite inferir la ocurrencia de al menos tres eventos de homoplasia, a saber:

1. en linajes de un subhaplogrupo A2 no descrito caracterizado por la presencia de 3394-4113, como en el zapoteca Z13 (Mizuno et al., 2014) y el mexicano-americano HQ012094 (Kumar et al., 2011);
2. en la secuencia A2w salvadoreña #190 (Aizpurua-Iraola et al., 2023) y
3. en el zapoteca A2d #Z95 (Mizuno et al., 2014).

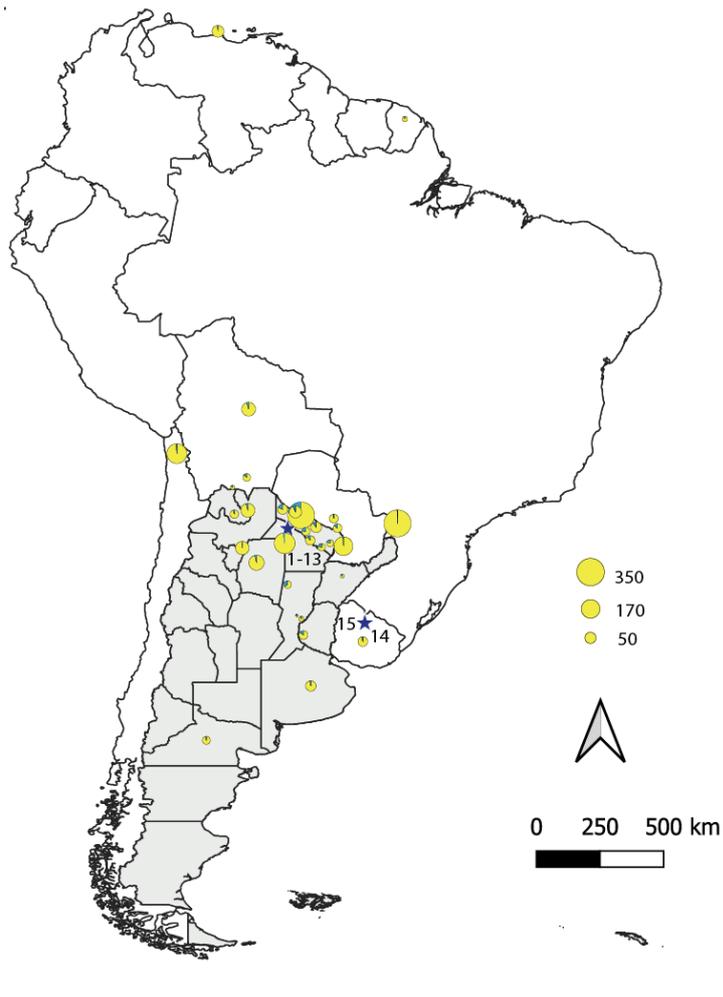


Figura 3.18. Mapa de frecuencia de A2bf (celeste) representado a escala en función del número de individuos nativos N(A,B,C,D) por punto de muestreo (amarillo). Las estrellas azules indican la localización geográfica de los mitogenomas representados en el árbol filogenético.

Por otra parte, el 84% de las secuencias A2bf recopiladas pertenecen a poblaciones indígenas, mayoritariamente de población wichi, pero también se han encontrado representadas en población qom, mocoví y pilagá (Anexo 3.6). Un 95% de éstas pertenece a localidades del NEA, las 9 secuencias restantes corresponden a un kaliña de la Guayana Francesa y a cuatro quechua y cuatro wichi de Bolivia (Figura 3.19).

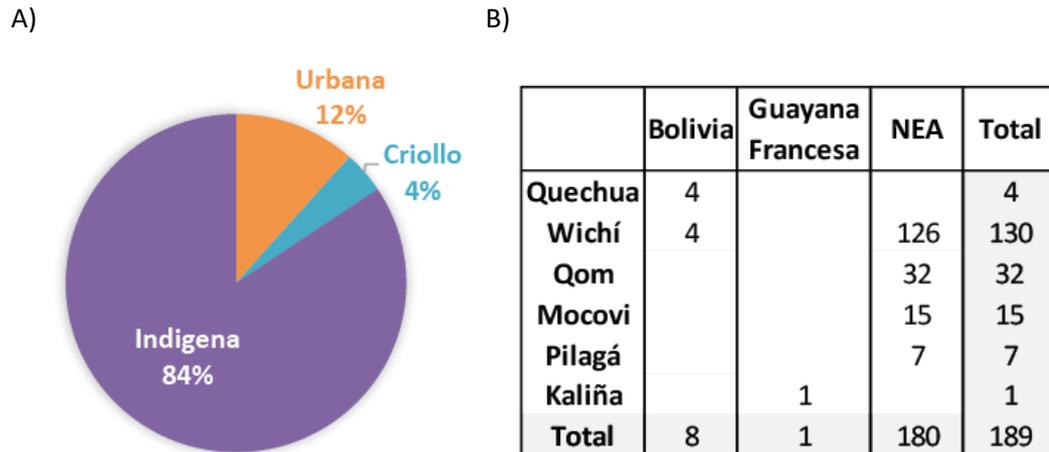


Figura 3.19. Frecuencia de secuencias recopiladas A2bf según su población de pertenencia y detalle de la fracción indígena. A) Frecuencia (en %) de A2bf según su población de pertenencia (urbana, criolla o indígena). B) Tabla que muestra la composición de la fracción indígena según número de secuencias por comunidad y región geográfica.

Se construyó una red de haplotipos en base a las secuencias recopiladas (Figura 3.20), resultando en una red con forma estrellada que agrupa 28 haplotipos en trece ramas principales. El haplotipo central recopila más del 70% de las secuencias de la red, de las cuales más de un 70% pertenece a población wichi y un 16% a población qom provenientes del NEA y Litoral.

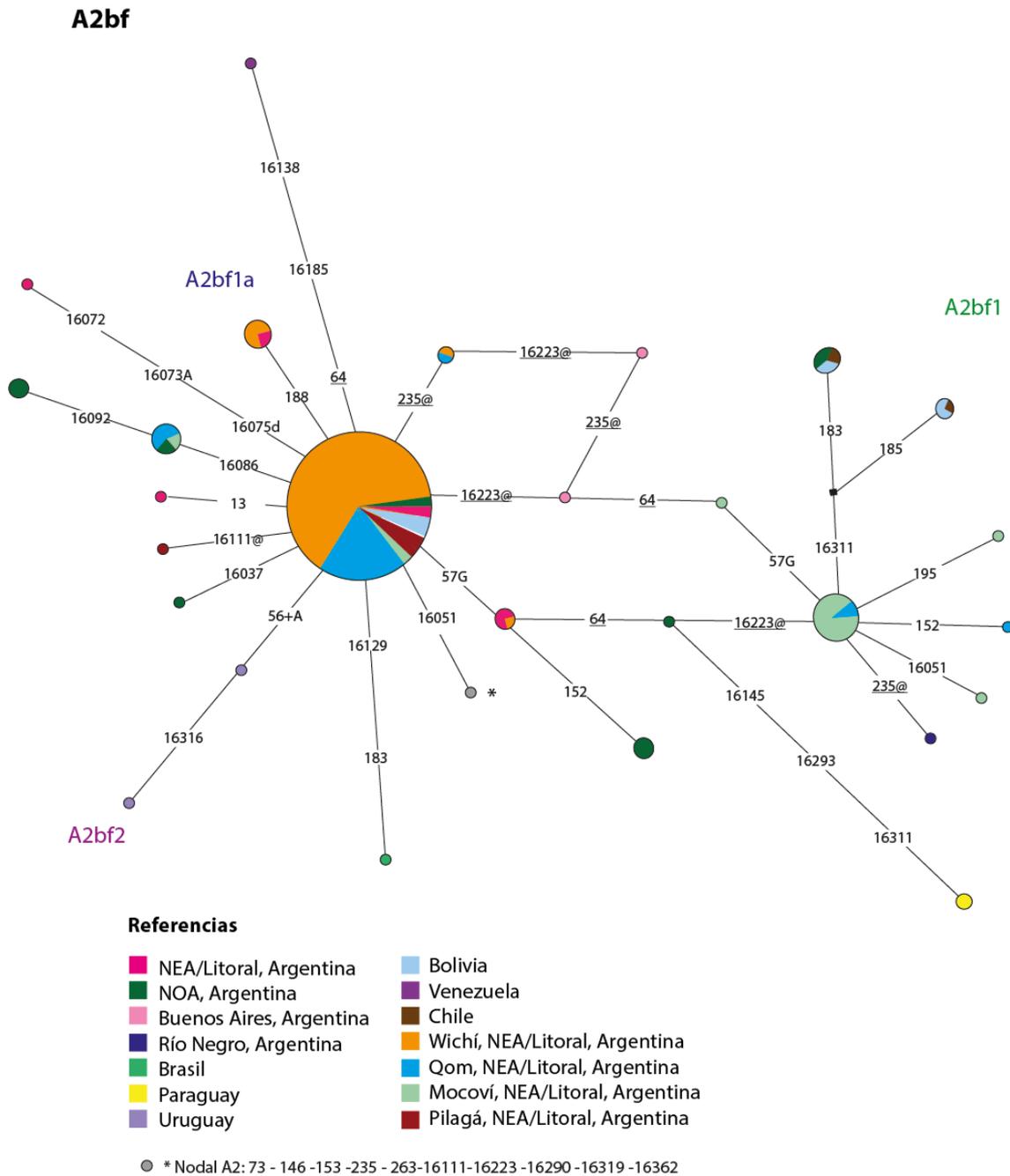


Figura 3.20. Red de haplotipos de A2bf. Los haplotipos se encuentran representados por círculos de tamaño proporcional a la frecuencia, cuya combinación de colores refiere a las poblaciones de origen (ver Referencias). Los rombos negros corresponden a haplotipos hipotéticos. Todos los cambios son transiciones, otros cambios se indican con un sufijo. Las posiciones subrayadas son homoplásicas, el símbolo “@” indica reversión respecto de la nodal. El haplotipo nodal está señalado mediante asterisco. *Indels* en el tracto 303-309, 311-315 no fueron tenidos en cuenta. Se indican en la figura los sub-linajes identificados. Rango de lectura: 16024-16362 y 1-250.

Por último, se construyó un árbol filogenético en base a un total de 15 mitogenomas compilados de la BBDD-RB (Anexo 3.7: mitogenomas 1-15, Figura 3.21). El análisis filogenético permitió reconocer tres clados con distribución geográfica particular: A2bf1* [= A2bf1(x A2bf1a)] que está presente en localidades del NEA y especialmente en población wichi, mocoví y pilagá, en Chile, Bolivia y muestras del NOA; A2bf1a, que se encontró en población wichi y criolla del NEA; mientras que A2bf2 está restringido a Uruguay (Figura 3.20 y Figura 3.21).

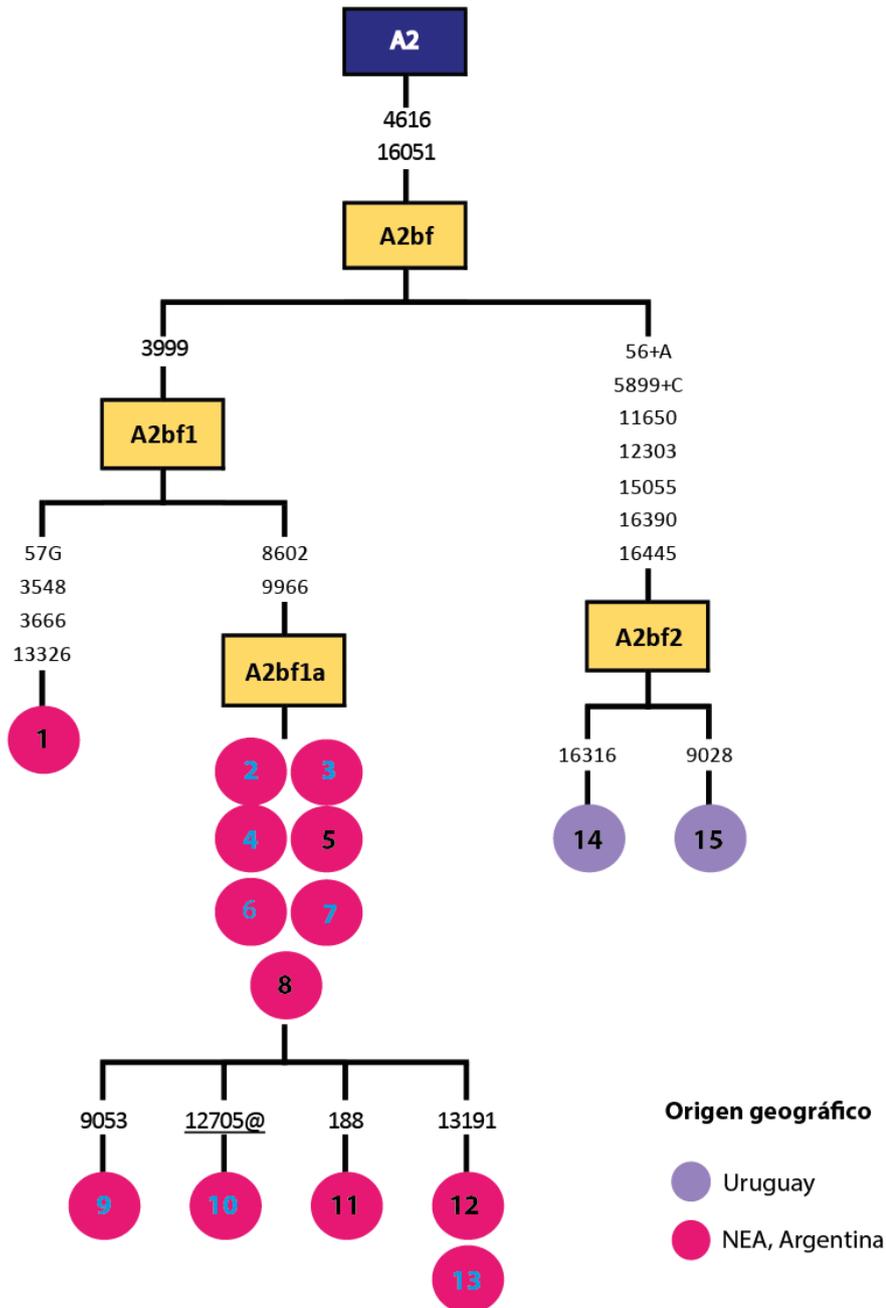


Figura 3.21. Árbol filogenético de A2bf. Se muestran las posiciones variables respecto de la rCRS, “@” indica reversión. En celeste se indican los mitogenomas pertenecientes a población wichi.

3.3.2 Linaje B2e

B2e es un sub-haplogrupo mitocondrial cuya distribución a nivel continental es todavía poco conocida ya que sólo puede ponerse en evidencia cuando se obtienen secuencias mitogenómicas, o cuando se hace una búsqueda *ad-hoc* de sus variantes diagnósticas 6119-14049. La información disponible apunta a su presencia en poblaciones amazónicas de Colombia (Lippold et al., 2014; Arias et al., 2017), así como en indígenas y población cosmopolita de Brasil (Fagundes et al., 2008), Paraguay (Simão et al., 2019), Argentina (Sevini et al., 2014) y Uruguay. Una anomalía filogeográfica está representada por el caso de una secuencia derivada de un afroamericano de Alabama que comparte un polimorfismo privado con un nativo amazónico. Hasta el momento, B2e parecía estar ausente en poblaciones andinas de Colombia, Ecuador, Perú y Bolivia (Tito et al., 2013; Zheng et al., 2012; Gómez-Carballeda et al., 2018; Brandini et al., 2018).

Se construyó un árbol filogenético en base a un total de 39 mitogenomas, de los cuales 27 derivaron de la literatura, 9 fueron obtenidos por donación de clientes de Family Tree DNA (cuyas mutaciones privadas no son mostradas), y tres son parte de la BBDD-RI (Anexo 3.7: mitogenomas 16-54, Figura 3.22). A partir de este árbol se diseñaron estrategias de tipificación de posiciones diagnósticas de clados derivados de B2e. Dentro de las 53 secuencias de Formosa que fueron asignadas al haplogrupo nativo B2, un tamizaje mediante secuenciación de la posición 6119, uno de los PNS diagnósticos de B2e, resultó en el hallazgo de 7 muestras asignables a B2e. Al momento de realizar estos análisis existían dos ramas del clado B2e: B2e1, definido por la presencia de 15672-15781 (García et al., 2021); y B2e2, diagnosticado por 986 y con una rama interna definida por la variante 9899. Las 7 muestras candidatas y otras 17 muestras de otras poblaciones del NEA y Litoral (Reconquista, Curuzú-Cuatiá, La Paz y Gualaguaychú, previamente asignadas a B2e) más una muestra de Tartagal, fueron analizadas para 9899, que define B2e2a, mediante PCR-APLP. Las muestras no asignadas a B2e2a fueron secuenciadas para la búsqueda de los PNS 15672-15781, que definen B2e1. Además, 5 de las muestras fueron analizadas para 986 mediante secuenciación.

El análisis filogenético permitió reconocer 6 clados, B2e1 a B2e6, que reunieron 37 de los 39 mitogenomas y que presentaron una distribución geográfica particular (Figura 3.22). Mientras que B2e1 y B2e2 están presentes en Brasil, Argentina, Paraguay y Uruguay, B2e5 y B2e6 están restringidos a la Amazonía colombiana, B2e3 tiene una distribución disyunta en el nordeste brasileño y la Guainía colombiana, y B2e4 reúne a un afroamericano de Estados Unidos con un waiwai amazónico (Figura 3.22).

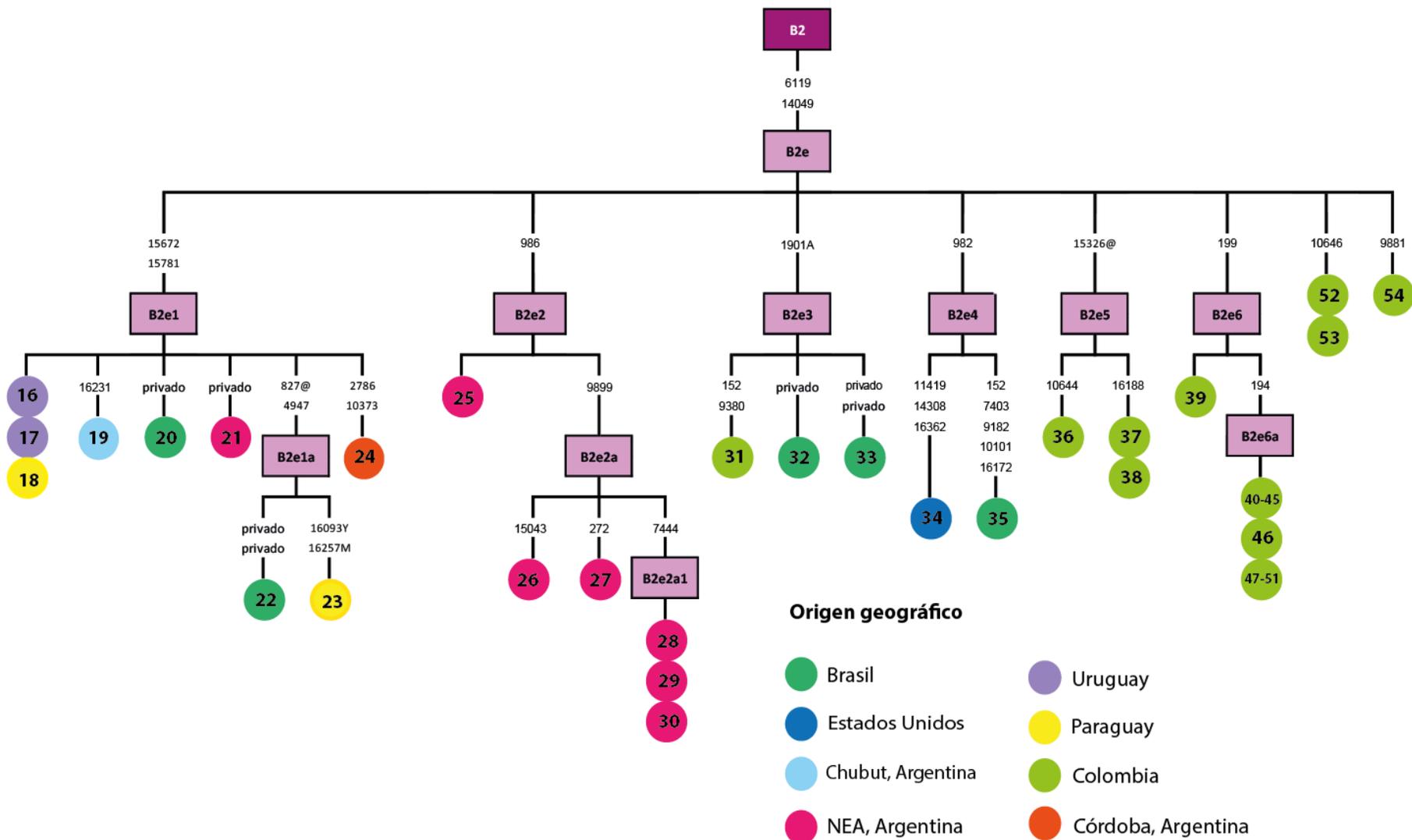


Figura 3.22. Árbol filogenético de B2e. Se muestran las posiciones variables respecto de la rCRS, “@” indica reversión.

Quince de las 25 muestras B2e analizadas en este trabajo (60%) resultaron asignables a B2e2a mientras que cinco tuvieron membresía en B2e2 y cuatro en B2e1. La muestra restante fue asignada al paragrupo B2e* (Tabla 3.12). La tabla 3.12 muestra los resultados obtenidos. Un total de 8 haplotipos para B2e fueron identificados y se construyó una red de haplotipos que se presentó en la Figura 3.7 en el segundo nivel de análisis (sección 3.2.1.3.2.2).

Tabla 3.12 Frecuencia absoluta de linajes asignables a diferentes clados de B2e en poblaciones del NEA y Litoral

Poblaciones	N total nativos	B2e*	B2e1	B2e2	B2e2a	Total
Formosa	169		2	1	4	7
Reconquista	79				1	1
Curuzú Cuatiá	140	1	1	1	5	8
La Paz	56			2	4	6
Guauguaychú	115		1		1	2
Tartagal	169			1		1
Total	728	1	4	5	15	25

3.3.3 Linaje B2an

La distribución del linaje B2an no está muy difundida en la bibliografía, ya que fue descrito recientemente por Figueiro et al., (2022). Estos autores denotaron la presencia de cuatro mutaciones (103, 152, 14470 y 16241) compartidas por dos mitogenomas oriundos de Uruguay y otro perteneciente a un individuo Xavante de Brasil (Fagundes et al., 2008). Debido a que tres de estas mutaciones pertenecen a la región control, este linaje se puede rastrear tanto en mitogenomas como en secuencias de región control, lo cual amplía el abanico de búsqueda de secuencias candidatas.

Tal como fue descrito en la sección 3.2.1.3.2.2, B2an fue identificado por sus PNS característicos en las secuencias de la región control en la BBDD-RI, y luego confirmado por APLP de la mutación diagnóstica 14470. De esta manera, se lograron asignar 24 secuencias a B2an de poblaciones del NEA y Litoral, 15 de las cuales se asignaron a clados de menor jerarquía como B2an2 (B2an+146-455+T) y B2an2a (B2an+146-16241@) (Tabla 3.13).

Tabla 3.13. Frecuencia absoluta de linajes asignables a diferentes clados de B2an en poblaciones del NEA y Litoral.

	B2an	B2an2	B2an2a	Total
Formosa	5	3	1	9
Gauleguaychú		5	3	8
Curuzú Cuatiá	3	2		5
La Paz		1		1
Reconquista	1			1
Total	9	11	4	24

Se identificaron 118 secuencias como pertenecientes al subhaplogrupo B2an, de las cuales 104 fueron secuencias de región control (Anexo 3.8) y 14 fueron mitogenomas (Anexo 3.7: mitogenomas 55-68). En la Figura 3.23 se observa la distribución geográfica del linaje B2an. En particular, las mayores frecuencias alcanzadas de B2an en población cosmopolita son del 14,29% en el Distrito Federal de Brasil (Nogueira, 2019) y del 12,62% en el sudeste brasilero (Fridman et al., 2014). En población parakaná del Amazonas brasilero se observa que alcanza frecuencias del 36,8% (Ramallo et al., 2013). Además, en localidades del NEA las frecuencias de B2an varían entre 1,3 al 7% en población cosmopolita (Beltramo, tesis en preparación; este trabajo) y 6,4% en población Mbyá Guaraní de Misiones (Sala et al., 2010).

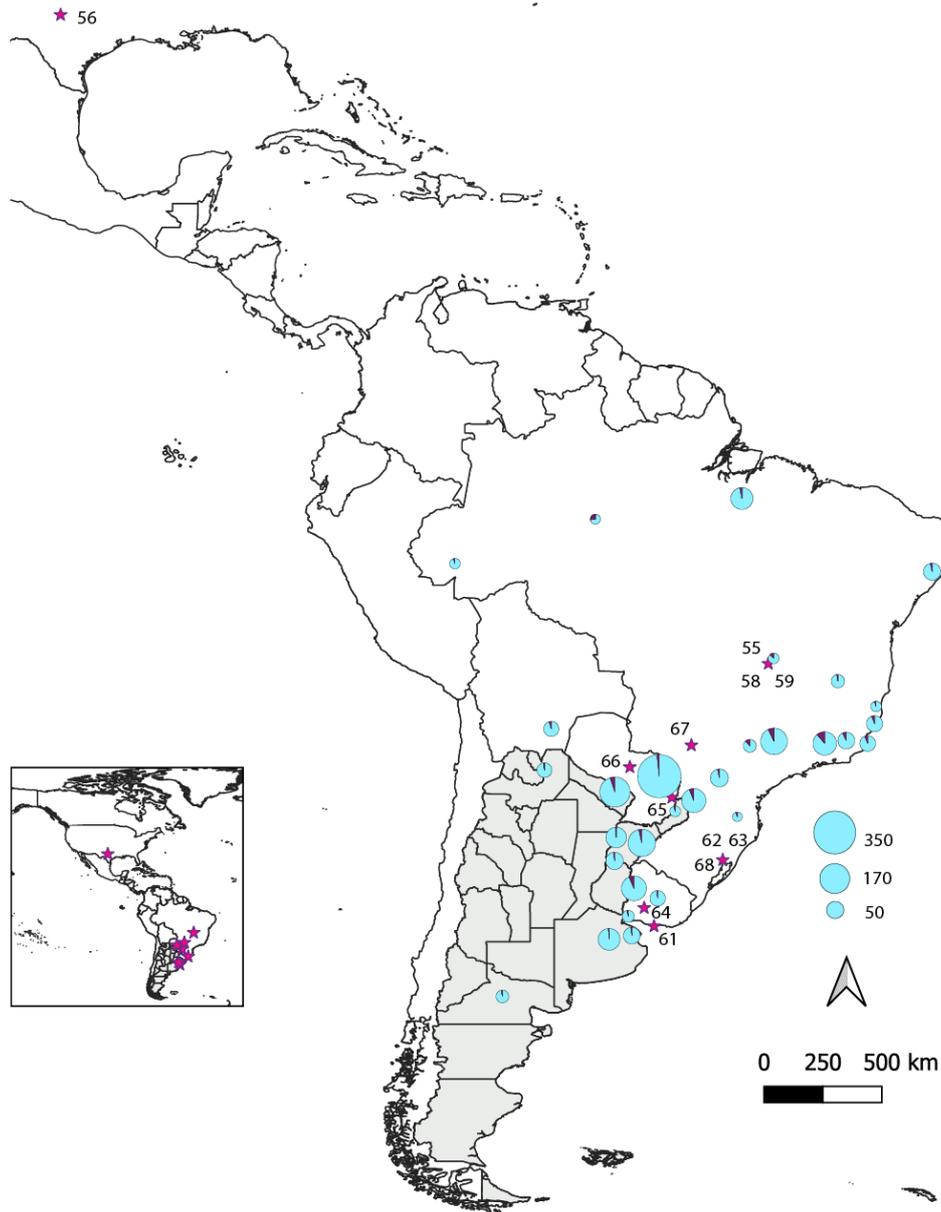


Figura 3.23. Mapa de frecuencia de B2an. Se representa la frecuencia de B2an (morado) a escala en función del número de individuos nativos N(A,B,C,D) (turquesa) por punto de muestreo. Las estrellas indican los puntos de muestreo de los mitogenomas representados en el árbol filogenético. Los mitogenomas 57 y 60 no se encuentran representados porque su localización es desconocida.

Se construyó una red de haplotipos, que resultó en una red con forma estrellada que agrupa 28 haplotipos y presenta 18 ramas principales (Anexo 3.8, Figura 3.24). El haplotipo central concentra la mayor cantidad de secuencias que corresponden en más de un 73% a secuencias de Brasil, pero también incluye secuencias del NEA, de Paraguay, del Ramal de Jujuy, de Bolivia y de Buenos Aires. La mayor parte de los haplotipos descritos en la red corresponden

al subhaplogrupo B2an, sin embargo, pudieron caracterizarse otros clados de menor jerarquía como B2an1 (B2an+16240C) y B2an3 (B2an+16093) en secuencias de Brasil, B2an2 (B2an+146-455+T) principalmente en secuencias del NEA y Litoral y en secuencias de Brasil y de Uruguay y B2an2a (B2an2+16241@) compartido por secuencias del NEA y Litoral, Paraguay y una secuencia de Río Negro. Es interesante resaltar, que los haplotipos que agrupan las secuencias asignadas a B2an2, incluyen las secuencias recopiladas de Mbya Guaraní de Misiones (Sala et al., 2010).

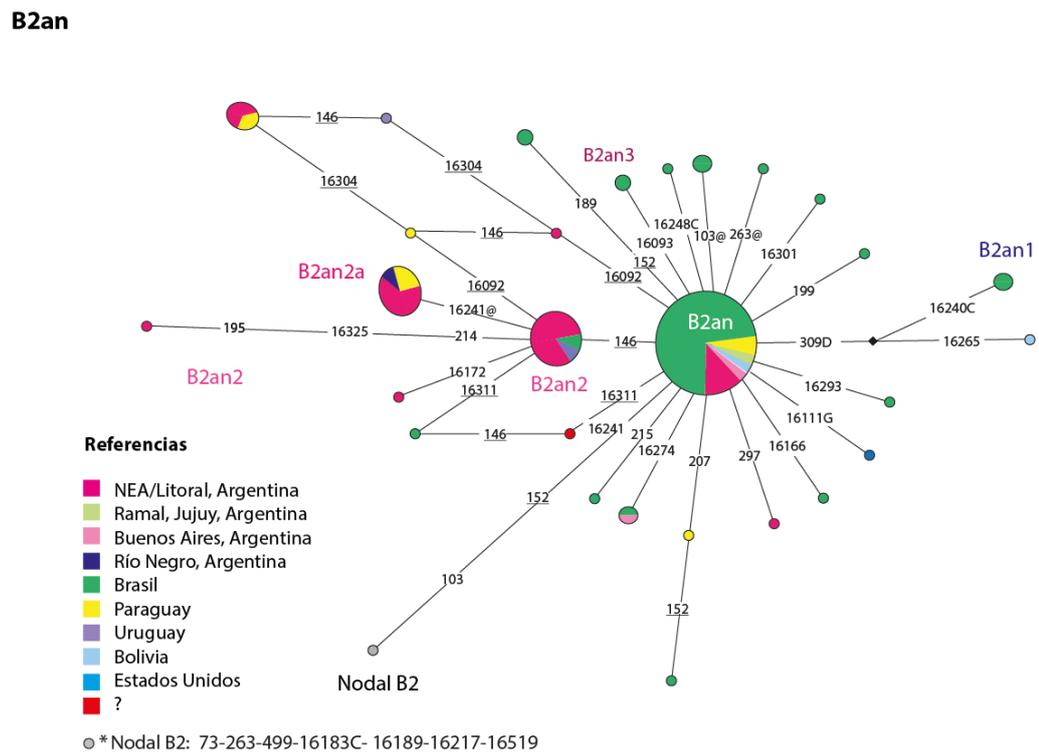


Figura 3.24. Red de haplotipos B2an. Ver detalles en leyenda de la Figura 3.20.

Rango de lectura: 16024-16365 y 73-340.

Por último, se construyó un árbol filogenético en base a los 14 mitogenomas recopilados (Anexo 3.7: mitogenomas 55-68, Figura 3.25). El análisis filogeográfico permite plantear un origen de B2an en la Amazonía brasilera, dada su alta frecuencia con una posterior diferenciación y expansión hacia el sur de Brasil, Paraguay, Uruguay, NEA y Litoral.

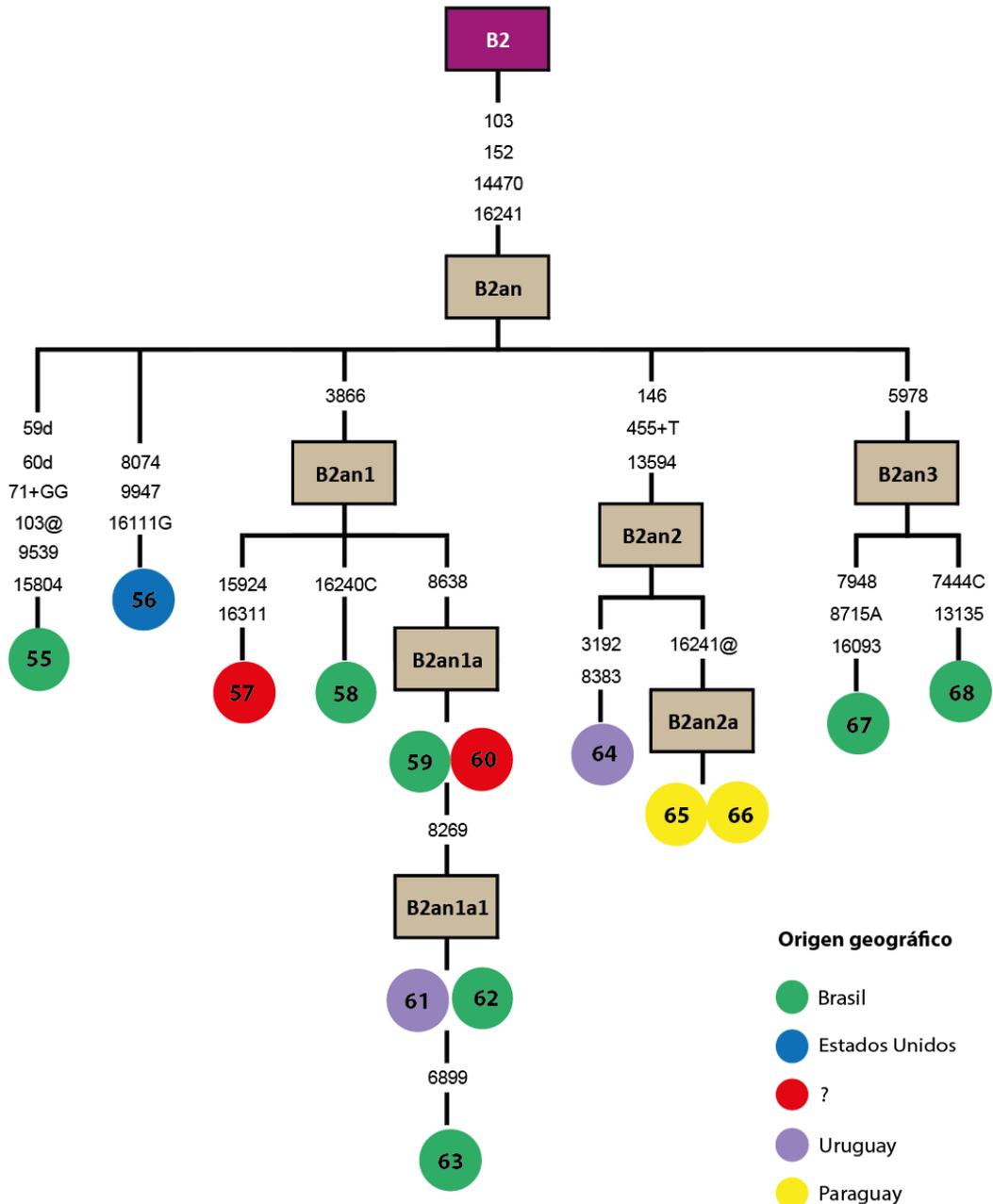


Figura 3.25. Árbol filogenético de B2an. Se muestran las posiciones variables respecto de la rCRS, “@” indica reversión.

3.3.4 Linaje D1e

D1e es un subhaplogrupo mitocondrial que se ha descrito hasta el momento en mitogenomas de Brasil (Lippold et al., 2014) y Paraguay (Simão et al., 2019) y que se caracteriza por presentar la combinación de mutaciones 146-3316-11248.

Para la asignación al subhaplogrupo D1e, se identificó por secuenciación de la región control extendida a las secuencias asignables al haplogrupo D1+146 de la BBDD-RI; de esta

manera se pudieron asignar seis secuencias de Formosa a este linaje (ver sección 3.2.1.3 de este trabajo). Además, se identificaron 32 secuencias asignables a D1e, 12 de ellas de región control (Anexo 3.9) y 20 mitogenomas (Anexo 3.7: mitogenomas 69-88). En la Figura 3.26 se puede observar la frecuencia de D1e de las secuencias recopiladas, lo que permitió ampliar la distribución de este linaje a la región del NEA, además de la ya conocida (Brasil y Paraguay). Los estudios poblacionales realizados indican que las mayores frecuencias de este linaje ocurren en ciertas localidades del NEA (Bobillo et al., 2010) (Formosa: 6,7% y Misiones: 4,6%). En este trabajo, se observa que la frecuencia de D1e no supera el 4% en la localidad de Formosa.



Figura 3.26. Mapa de frecuencia de D1e. Se representa la frecuencia de D1e (turquesa) a escala en función del número de individuos nativos N(A,B,C,D) por punto de muestreo (lila). Las estrellas amarillas indican los puntos de muestreo de los mitogenomas representados en el árbol filogenético.

Se construyó una red de haplotipos que agrupa cuatro haplotipos en dos ramas principales, cada una de las cuales corresponde a un clado de menor jerarquía dentro de D1e (Figura 3.27). La primera rama corresponde al subhaplogrupo D1e1 que agrupa secuencias del NEA y Paraguay; la segunda rama, porta la mutación diagnóstica 16261 que caracteriza al subhaplogrupo D1e2, propio de secuencias de Brasil. Esta misma diferenciación se refleja en el árbol filogenético que se realizó en base a los 20 mitogenomas recopilados (Anexo 3.7: mitogenomas 69-88, Figura 3.28). Es interesante resaltar que el 100% de los mitogenomas asignados a D1e2 (N=12) corresponden a población karitiana de Brasil.

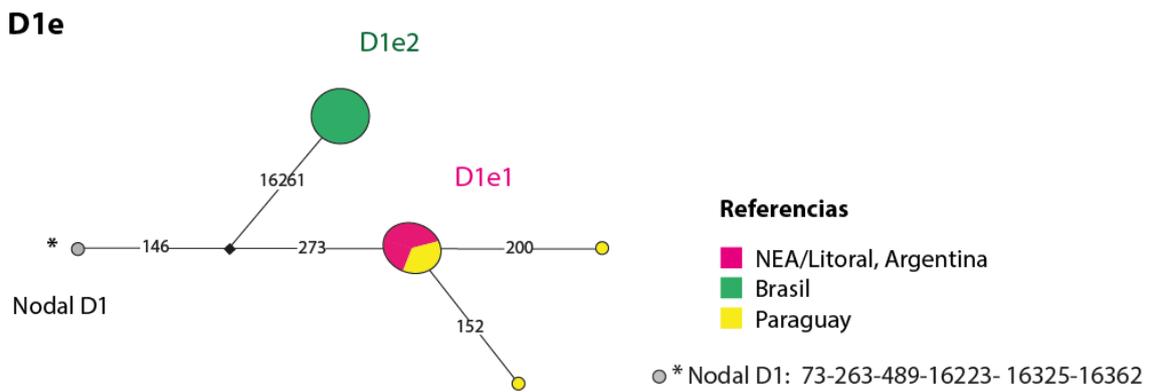


Figura 3.27. Red de haplotipos D1e. Ver detalles en leyenda de la Figura 3.20. Rango de lectura: 16024-16569 y 1-576.

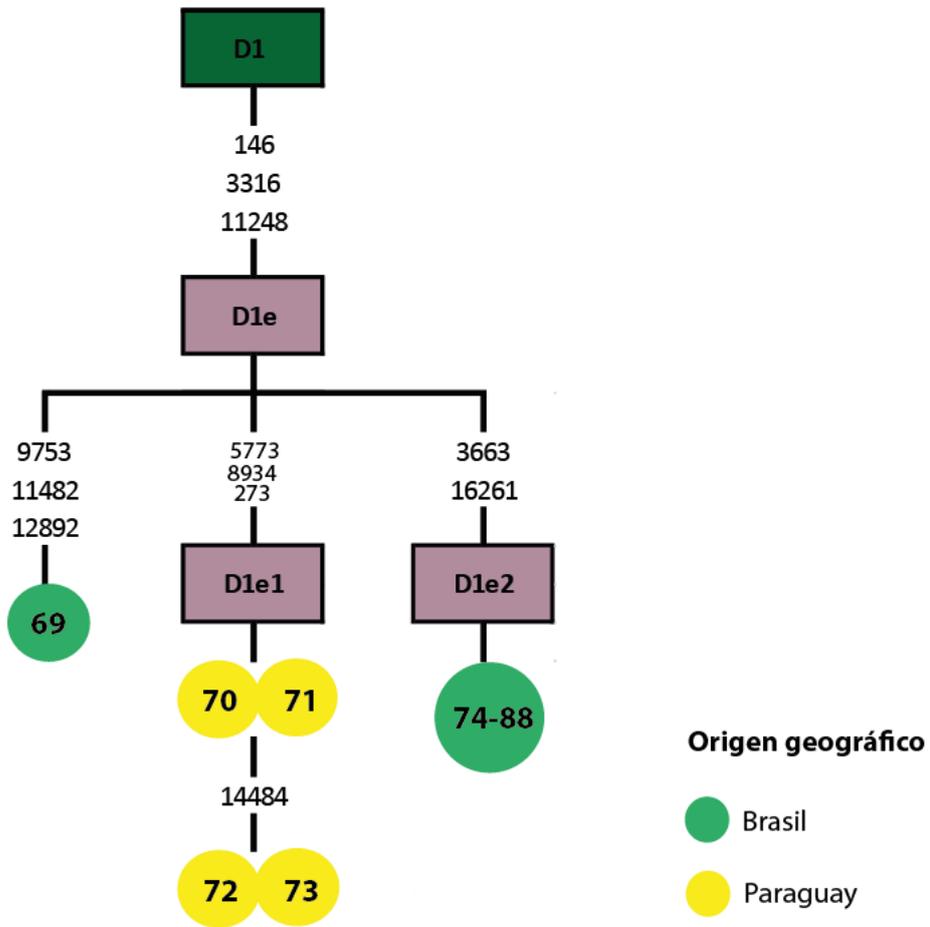


Figura 3.28. Árbol filogenético de D1e. Se muestran las posiciones variables respecto de la rCRS.

3.3.5 Linaje D1f1

D1f es un subhaplogrupo mitocondrial caracterizado por la presencia de 16142, que se ha documentado en mitogenomas de Norteamérica y de Sudamérica (Brandini et al., 2018). Por otra parte, D1f1 se caracteriza por presentar la combinación de mutaciones 10874-16179-16295-16497. A nivel sudamericano, se ha registrado su presencia en secuencias de región control de Panamá (Perego et al., 2012) y Uruguay (Pagano et al., 2005) y en mitogenomas de Venezuela (Brandini et al., 2018), Brasil (Fagundes et al., 2008; Lippold et al., 2014; Ávila et al., 2022) y Paraguay (Simão et al., 2019; Simão et al., 2021). Debido a que tres de las cuatro mutaciones diagnósticas de D1f1 se encuentran en la región control, su búsqueda en este tipo de secuencias es posible, lo que permite ampliar el conocimiento acerca su distribución.

Se recopilaron un total de 71 secuencias asignables a D1f1: 62 de región control (Anexo 3.10) y 9 mitogenomas (Anexo 3.7, mitogenomas 89-97). Inicialmente, se analizaron las secuencias asignadas al haplogrupo D1 de la BBDD-RI (ver sección 3.2.1.3 de este trabajo) y se pudieron asignar 15 secuencias a D1f1 correspondientes a localidades del NEA y Litoral, 10 de las cuales pertenecen a individuos de Formosa, tres a Reconquista y dos a Curuzú-Cuatiá. Una secuencia de La Paz se asignó a un linaje pre-D1f1 ante la ausencia de 16295. Por otra parte, se recopilaron dos secuencias de Corrientes y una secuencia de Misiones, pudiendo notar una marcada presencia de D1f1 en el NEA y Litoral. En la figura 3.29, se documenta un representante en una muestra poblacional de Panamá y frecuencias que varían entre el 4,5 (Castro de Guerra et al., 2011) y el 12% (Gómez Carballa et al., 2012) en Venezuela. Una marcada representación de este linaje se puede determinar en base a la bibliografía analizada en el sur de Brasil y en el NEA y Litoral, mediante los resultados de este trabajo (Anexo 3.10).

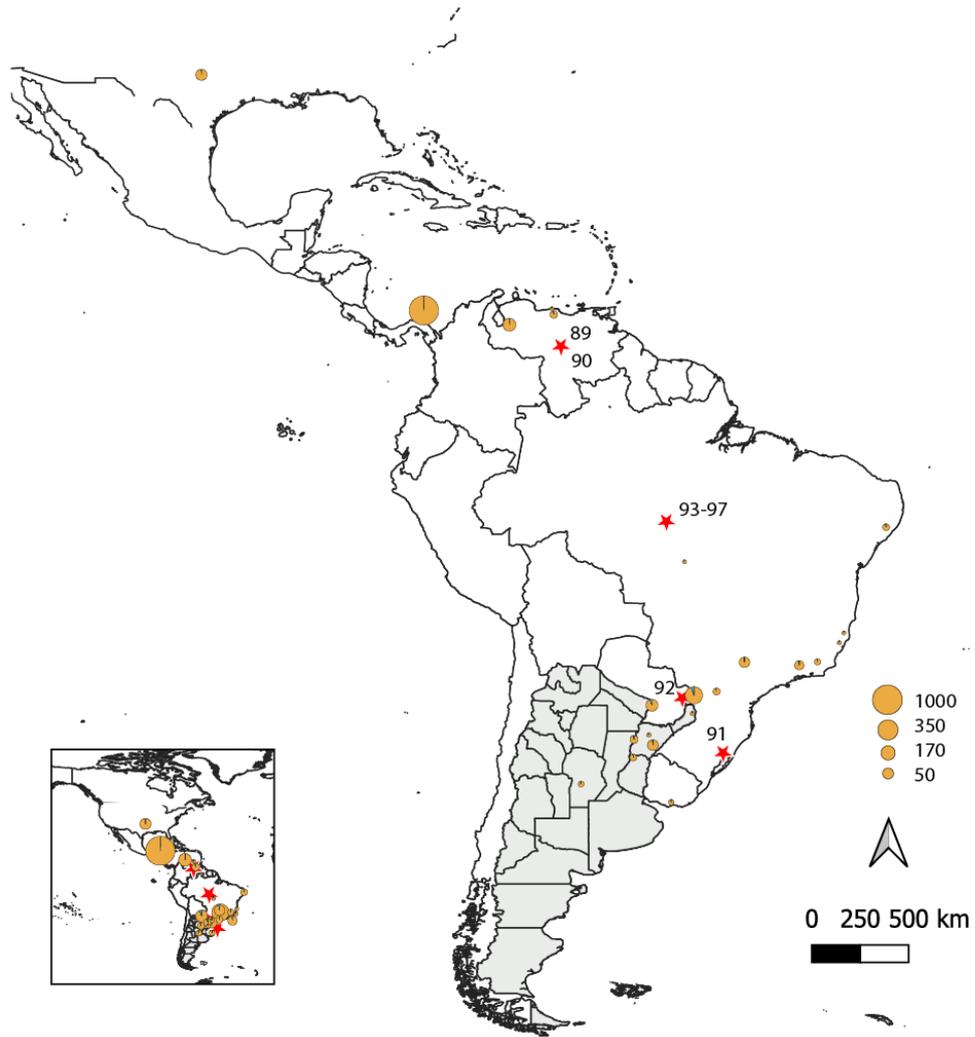


Figura 3.29. Mapa de frecuencia de D1f1 (turquesa) representado a escala en función del número de individuos nativos N(A,B,C,D) por punto de muestreo (naranja). Las estrellas (rojas) indican los puntos de muestreo de los mitogenomas representados en el árbol filogenético.

Se construyó una red que presentó 9 haplotipos (Figura 3.30). El haplotipo central fue el que concentró mayor número, agrupando secuencias del NEA y Litoral + Paraguay en un 50%, otro 25% de este haplotipo lo integran secuencias de Brasil y el resto, secuencias de Venezuela, y un individuo de Panamá, Estados Unidos, Uruguay y Córdoba, respectivamente.

D1f1

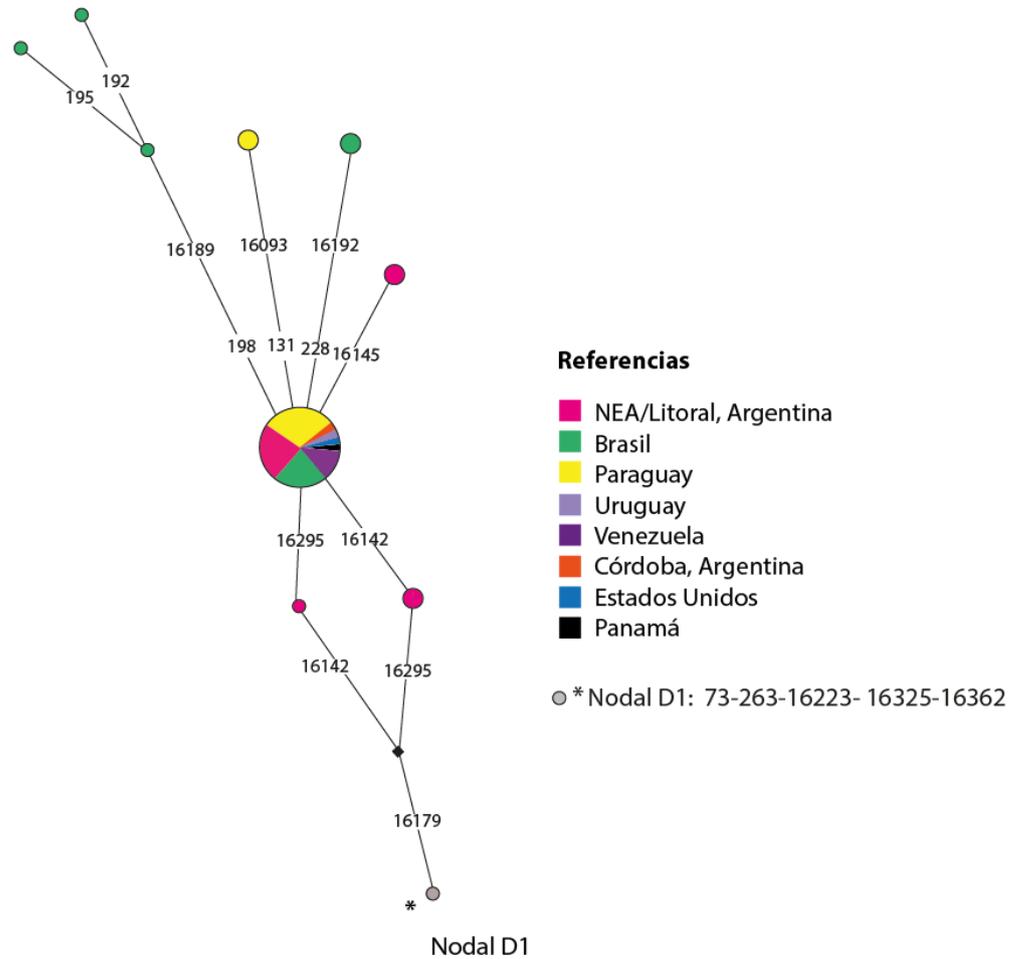


Figura 3.30. Rede de haplotipos D1f1. Ver detalhes em legenda de la Figura 3.20. Rango de lectura: 16024-16365 y 73-340.

Por último, se construyó un árbol filogenético en base a los 9 mitogenomas recopilados (Anexo 3.7, mitogenomas 89-97, Figura 3.31). Se destaca la presencia de dos mitogenomas de origen venezolano, a partir de los cuales parten el resto de las ramas del árbol.

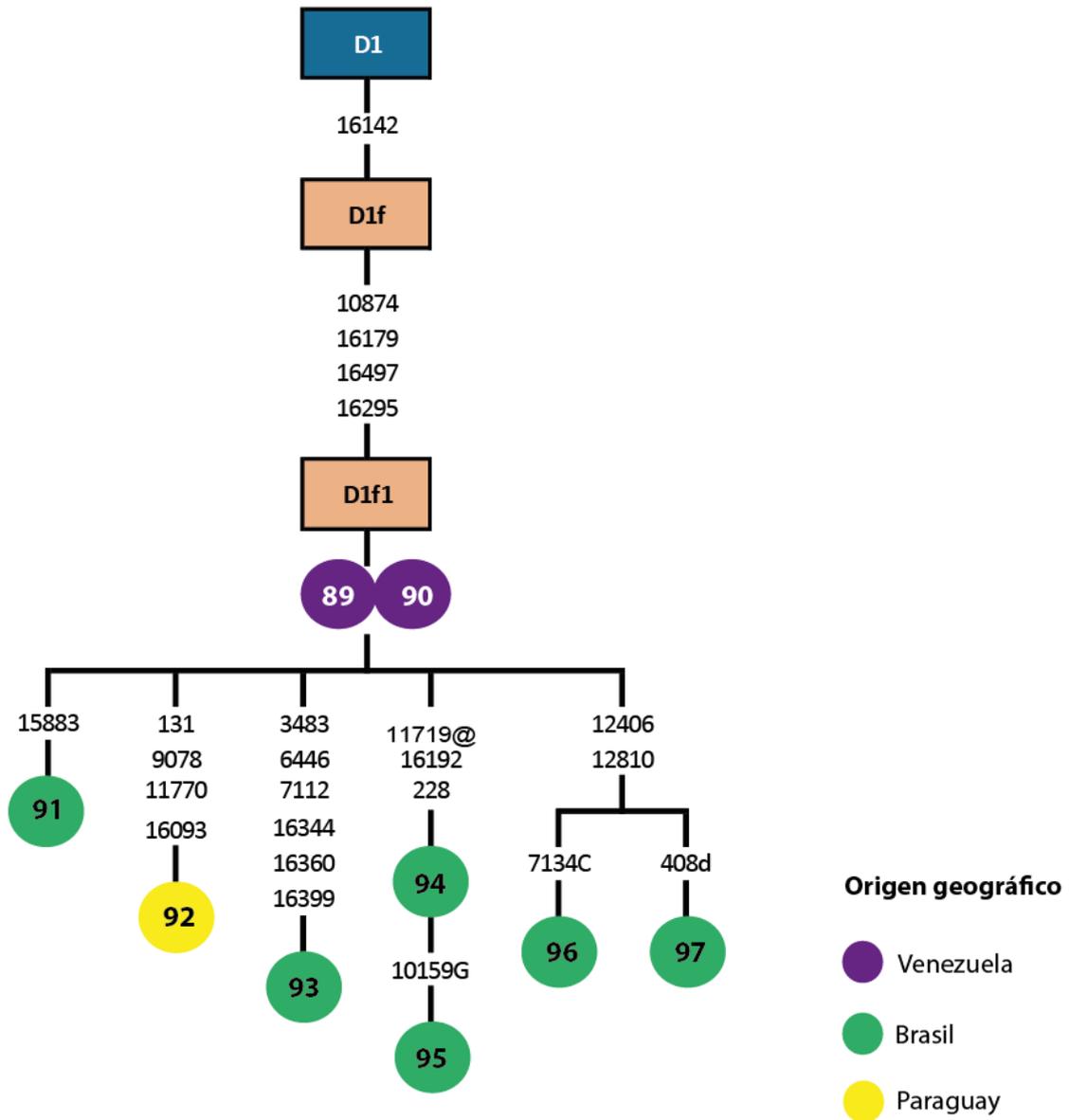


Figura 3.31. Árbol filogenético de D1f1. Se muestran las posiciones variables respecto de la rCRS, “@” indica reversión

4 CAPÍTULO 4. DISCUSIÓN

4.1 Migraciones intercontinentales

4.1.1 Linajes genéticos maternos y paternos

Una fracción importante de las personas que contribuyeron al acervo génico actual de Argentina arribó durante las migraciones ultramarinas entre 1860-1930. El trabajo de Dipierri et al. (1998) fue pionero en demostrar un proceso de mestizaje entre las poblaciones preexistentes y los migrantes -en su mayoría europeos y varones-, con aportes diferenciales según el sexo, utilizando tanto marcadores mitocondriales como del cromosoma Y. De esta forma, se observa que, mayoritariamente, los linajes maternos presentan ancestría nativa y los linajes paternos realizan un aporte alóctono. En este sentido, nuestros resultados se asemejan a los resultados de estudios previos y a los realizados en nuestro laboratorio en 15 localidades de la Argentina (Motti, 2012; Jurado Medina, 2015; Schwab, 2018). Particularmente, se observa que la frecuencia de linajes nativos maternos en Corrientes y Formosa es mayor al 85%, al igual que en la mayoría de las localidades del NOA y centro-oeste de Argentina analizadas previamente por nuestro grupo (Figura 3.2). Este predominio de linajes de origen nativo se relaciona con un menor ingreso de inmigrantes ultramarinos en la zona norte del país, tanto para el NEA como para el NOA, a diferencia de un mayor flujo de migrantes de ultramar en regiones como la pampeana y la patagónica, en donde se reportan frecuencias del componente nativo del 50% para la provincia de Buenos Aires y de alrededor del 65% para las provincias patagónicas (Avena et al., 2009; Bobillo et al., 2010; Parolín et al., 2013).

Contrariamente, para los linajes paternos se destaca la muy baja frecuencia del haplogrupo nativo, particularmente en las muestras de Corrientes (3,2%) y Formosa (4,9%). Además, la representación del aporte que realiza el clado Q-M3 al acervo génico de distintas localidades de Argentina permite vislumbrar que las frecuencias obtenidas para las provincias del NEA son, junto a las de Santiago del Estero (5,3%), las más bajas informadas hasta el momento para Argentina (Figura 3.2). Para contextualizar estos resultados, es necesario tener en cuenta el escenario histórico de la Guerra de la Triple Alianza, reconocida como una de las contiendas más largas y sangrientas de la historia latinoamericana, en donde se enfrentaron Argentina, Uruguay y Brasil contra el Paraguay desde 1864 a 1870. Baratta (2013) realiza un análisis de la producción historiográfica sobre la guerra en donde destaca que, pese a la activa participación de la Argentina en la contienda, se ha desdibujado su registro de la historia, a pesar

del impacto demográfico de la misma debido a la gran cantidad de bajas y su significación clave en la consolidación del Estado Argentino. Esto se ha puesto de manifiesto en trabajos realizados en las últimas décadas con nuevas perspectivas. En particular, se ha dado relevancia al papel de Corrientes durante la guerra del Paraguay en la historiografía argentina en forma reciente. Por ejemplo, Ramírez Braschi (2004), realizando un análisis de la prensa en esta provincia, coloca a Corrientes como escenario geográfico del campo de batalla durante los primeros años de la guerra, y luego como retaguardia y campo de aprovisionamiento de los ejércitos aliados, con lo cual la dinámica de la ciudad se vio alterada radicalmente hasta 1869. En este contexto, es altamente probable que haya ocurrido una disminución drástica de varones, con la consecuente pérdida de linajes nativos, tal como se observa a través de los análisis moleculares. Por otra parte, teniendo en cuenta que Formosa es una ciudad atractora de migrantes oriundos del Paraguay, lo cual pudo vislumbrarse cuando se analizó la información genealógica de los donantes, así como también a través de los censos nacionales de 1914 y de 2010, y teniendo en cuenta el “Holocausto masculino” vivido durante la Guerra de la Triple Alianza por el pueblo paraguayo (Baratta, 2013), consideramos que este mismo escenario histórico permite explicar la baja frecuencia de linajes paternos nativos en ambas localidades.

4.1.2 Historia oral

Respecto a los resultados obtenidos a partir del análisis de las encuestas genealógicas realizadas a los donantes, se observa una preservación en mayor frecuencia de la información genealógica proveniente de la abuela materna que del abuelo paterno (74% y 56% respectivamente), indicando una mayor profundidad temporal en la historia oral registrada para la línea materna. A su vez, se destaca que predomina un origen regional del ancestro más remoto declarado (materno: 85,6%; paterno: 80,2%) y se verifica una fuerte influencia de migrantes provenientes de Paraguay en la ciudad de Formosa, tendencia que ya podía observarse en los datos reportados en el Censo Nacional de 1914. Este fenómeno es especialmente notable en Formosa donde, sobre una población de 8.774 habitantes, el 45% era extranjero y, a su vez, de esta fracción el 84% provenía de Paraguay y 13,6% era de origen ultramarino; mientras que, para Corrientes, sobre una población de 24.462 habitantes, el 7% era extranjero, y entre ellos el 16,9% era de origen paraguayo y el 39% era de origen ultramarino. Por lo tanto, la frecuencia obtenida nos indica la influencia de Paraguay en la región del NEA respecto a los procesos poblacionales y, contextualizada con los datos censales, permite inferir una continuidad histórica de esta influencia hasta la actualidad.

Además, la comparación de datos moleculares y genealógicos permite dar cuenta de que una mayor parte de los donantes declara un origen regional de sus ancestros paternos y

maternos más remotos conocidos en las localidades estudiadas. En particular, para los resultados obtenidos de la línea materna de aquellos que declararon un origen regional del ancestro más remoto conocido (en 89% se verificó ancestría nativa), se considera que esta imagen podría ser el reflejo de las poblaciones criollas de la región en tiempos coloniales. Resultados semejantes se han obtenido por Motti (2012) en 1734 muestras de la Región del NOA y Centro Oeste argentino.

Por otra parte, para la línea paterna se destaca un componente mayoritario que declara un origen regional (89%) donde, sin embargo, en aproximadamente el 96% de los casos se identificó un origen molecular euroasiático. Estos resultados reflejan la antigüedad de los eventos de mestizaje, que probablemente se remontan a las primeras etapas de colonización del territorio, razón por la cual no ha quedado registrado en la historia oral. Por otra parte, dan cuenta de la asimetría del mestizaje en la población criolla. Además, se observa una mayor pérdida de la historia paterna por paternidades no conocidas. En línea con estos resultados, en un trabajo realizado por Muzzio *et al.* (2012), en el que se evalúa la proporción de transmisión de apellidos por línea paterna en diferentes regiones de Argentina, se demostró una alta proporción de paternidades no reconocidas entre la población general, una muy rápida pérdida de asociación entre los linajes paternos y los apellidos, y significativas diferencias regionales.

La interpretación de las discrepancias entre los resultados de la historia oral y de los análisis moleculares puede derivar en explicaciones complejas. Tal como plantea Sans (2009), una razón es que reflejen los procesos de construcción de la identidad nacional, en donde los varones indígenas habrían sido integrados a la sociedad luego de las guerras y las mujeres indígenas por medio de uniones con varones europeos o criollos; en cualquier caso, habrían ocultado o perdido la identificación de su origen. La idea de nación construida a partir de la negación, invisibilización y borramiento de otras alteridades —en especial, indígena y afrodescendiente—, oculta cualquier evidencia que no se corresponda con dicho canon (García *et al.*, 2016).

Por último, se debe tener en cuenta que la determinación de la ancestría a través de la historia oral —particularmente en este trabajo, del origen del ancestro más remoto conocido— está limitada por el conocimiento de las personas acerca de sus ancestros, ya que en la mayoría de los casos sólo se remonta a dos o tres generaciones ascendientes. Así como también, por el desconocimiento de los ancestros más recientes, dado por múltiples factores, incluida la adopción (Race, Ethnicity, and Genetics Working Group, 2005).

4.2 Relevancia de la Identificación de polimorfismos del ADNmt en región codificante. Una discusión metodológica.

La ventaja de contar con secuencias de la región control del ADNmt permite realizar un número mucho mayor de comparaciones para la asignación de subhaplogrupos, en comparación con el menor número de mitogenomas disponibles. Sin embargo, tal como ya habían advertido Simão et al. (2019), cuando se considera sólo la información de la región control, la asignación a un determinado subhaplogrupo es diferente en un gran porcentaje de muestras de la que se realizaría si se contara además, con información de la región codificante; esto ocurre en tres escenarios:

1. cuando la asignación a un determinado subhaplogrupo sólo es posible a partir de polimorfismos de la región codificante;
2. cuando la clasificación se realiza en otra rama debido a que los polimorfismos obtenidos de la región codificante son diferentes a aquellos que se requerían para su clasificación basada en región control;
3. cuando aun faltando ciertas mutaciones en la región control, otras mutaciones de la región codificante permiten una clasificación con mayor resolución.

En este trabajo de tesis, la obtención de PNS de la región codificante y su análisis en conjunto con los polimorfismos de la región control, permitió definir nuevos grupos monofiléticos, la asignación a clados ya sea que se trate de linajes sin variantes compartidas o con variantes homoplásicas en región control y reconocer linajes y asignarlos a un subhaplogrupo determinado, cuando éstos presentaban polimorfismos diversos en la región control.

4.2.1 Asociación entre haplotipos de región control, PNS en regiones codificantes y clados dentro del haplogrupo A2.

Una fracción importante de los linajes maternos pertenecientes al haplogrupo A2, cuando se analizan a nivel de las secuencias de región control, son o bien idénticos a las secuencias nodales de A2 y A2₍₆₄₎, o bien portadores de uno o unos pocos polimorfismos derivados, en ocasiones homoplásicos. La exploración de los fragmentos FR2 (2900-3960), FR7 (14100-14900) y FR6 (12494-13557) en secuencias asignadas al haplogrupo A2 (Anexo 3.3), puso en evidencia la existencia de tres regiones enriquecidas en polimorfismos informativos de diferentes subhaplogrupos frecuentes en Sudamérica extra-andina. En particular, en 16 de las 31 muestras secuenciadas para parte del fragmento FR2 se detectó al menos un PNS. Seis muestras con el haplotipo nodal para A2₍₆₄₎ fueron portadoras de 2905-3579, que determinan

membresía en el subhaplogrupo A2za1a y ha sido detectado en guaraníes (Rieux et al., 2014) y población cosmopolita de Porto Alegre (Ávila et al. 2022) y Paraguay (Simão et al., 2021). En dos muestras con la secuencia nodal A2 se halló 3435, diagnóstico del subhaplogrupo A2bg1 encontrado de suruí, waiwai y población cosmopolita de Porto Alegre. La presencia de 3693 en FOR056 la vincula con un clado de amplia distribución en la Amazonía -yanomama de Brasil- (Ávila et al., 2022); -tuyuca, piapoco y achagua de Colombia- (Arias et al., 2017). Con 485-16269-16360 en región control, FOR056 es asignable a A2zb1, también hallado en población cosmopolita brasilera. La presencia de 3484 asociada a 279-16526 permite asignar FLP156 al clado A2zc1, hallado en muestras actuales y antiguas del Caribe (Puerto Rico, República Dominicana, Bahamas) (Vilar et al., 2014; Fernandes et al., 2021), así como en makiritare de la amazonía venezolana (Nieves-Colón et al., 2019), y en población cosmopolita de Brasil (Ávila et al., 2022) y Paraguay (Simão et al., 2019; 2022).

Con 4158 y 153@, CC068 y FLP178 son parte de A2zz. Por último, la coocurrencia de 3083 y 16422 permite emparentar FOR071 con población brasilera y con los aché de Paraguay, asignables a A2ze (Fagundes et al., 2008).

La secuenciación del fragmento FR7 también resultó en numerosas asignaciones. Cinco individuos portadores del haplotipo nodal para A2(64), así como un sexto con 16092, compartieron la presencia de 15262 y asignación a A2zd. Ocho muestras compartieron 143 y 13708, siendo asignadas a A2bh, descrito tanto en nativos kayapó, como en paraguayos y brasileros cosmopolitas. Mientras que FLP208 comparte presencia de 16104 y 14040-14052 con brasileros y paraguayos, TG060 también lleva 16104 pero acompañada de 14178-14755-14861 y es por lo tanto asignable a A2ya. Dos muestras de Formosa compartieron 14491-15884, PNS aún no informados en mitogenomas A2.

Cuatro individuos de Formosa y uno de Tartagal fueron portadores de 16266. La secuenciación parcial del fragmento FR6 en tres de ellos permitió detectar la presencia de 12858 y por lo tanto asignarlos a un clado que aquí denominamos A2zk3 y que ha sido bautizado como A2ar por García-Olivares 2023. A2ar está ampliamente distribuido en la Amazonía de Venezuela (makiritare) y de Perú (chachapoyas) (Nieves-Colón et al., 2019), en el Caribe (Cuba, Puerto Rico, Rep. Dominicana) (Fernandes et al., 2021), Brasil (Ávila et al., 2019), Paraguay (Simão et al., 2019) y Argentina (este trabajo).

En resumen, las asignaciones resultantes mediante tamizaje de posiciones informativas de FR2, FR6 y FR7 demuestran que la secuenciación parcial de región codificante es una herramienta que permite el subhaplogrupado de alta resolución.

4.3 Vínculos interregionales y distribución sudamericana de clados mitocondriales

4.3.1 Análisis de diversidad mitocondrial

El análisis de la fracción nativa de la localidad de Formosa, a partir de la obtención de secuencias de la región control extendida y de la búsqueda de ciertos PNS de la región codificante, permitió asignar las secuencias a diferentes subclados nativos de ADN mitocondrial, en contexto con las secuencias de otras 19 localidades de Argentina. De esta manera, se pudo interpretar la variabilidad genética local y regional a partir de linajes monofiléticos de menor jerarquía, que constituyeron las variables operacionales para analizar su frecuencia y su distribución geográfica. El hecho de que los linajes fueran definidos en relación a PNS de la región codificante además de los polimorfismos propios de la región control, permitió asegurar su monofilia y que las relaciones de afinidad entre las poblaciones analizadas den cuenta de procesos que ocurrieron con posterioridad al poblamiento americano y que tuvieron lugar en Sudamérica. Es decir, cada uno de los 201 linajes definidos en este trabajo constituye un grupo monofilético de haplotipos cuyo ancestro común habitó el continente en momentos posteriores al poblamiento americano.

La diversidad mitocondrial observada en las poblaciones actuales es el resultado de procesos y eventos evolutivos que la originaron en el pasado; así, una alta diversidad puede atribuirse a eventos de migración o crecimiento demográfico, mientras que una diversidad baja podría resultar de cuellos de botella o descensos del tamaño poblacional. Las estimaciones de variabilidad intrapoblacional realizadas en base a las secuencias obtenidas de Formosa, indican que esta localidad se agrupa entre las poblaciones que presenta mayores valores de diversidad haplotípica (alrededor de 0,99) conjuntamente con Maimará, Mendoza, Tartagal, La Quiaca y Jujuy (Motti, 2012). Los haplotipos encontrados se hallaron en al menos dos individuos de la muestra lo que resultó en valores bajos en el porcentaje de haplotipos únicos (40%). Además, la localidad de Formosa se caracterizó por una gran diversidad de linajes mitocondriales, pudiéndose identificar 61 linajes monofiléticos. Desde un punto de vista histórico más reciente, Formosa constituye una ciudad capital atractora de migrantes, especialmente de Paraguay. Este flujo génico habría sido sostenido en el tiempo según se ha corroborado en los Censos nacionales de 1914 y 2010 (ver sección anterior), lo cual podría explicar, al menos en parte, su gran diversidad. Sin embargo, el origen de esta diversidad también podría explicarse por procesos migratorios más antiguos, relacionados con las dinámicas de los pueblos originarios que habitaban la región en tiempos prehispánicos o del poblamiento inicial de Sudamérica. Para analizar las evidencias al respecto, es necesario considerar a Formosa como parte integral de la

Cuenca del Plata, emplazada entre los ríos Pilcomayo y Bermejo en dirección norte-sur, y por el río Paraguay en su borde oriental. En este sentido, Rodríguez (2006) menciona que esta cuenca tuvo mayor diversidad cultural que áreas vecinas (el Chaco, La Pampa y el Centro) y lo relacionó a su gran diversidad ambiental y climática. Según este autor, miembros de una misma *entidad cultural* (grupos cazadores-recolectores, mayoritariamente) se dispersaron ocupando zonas ribereñas de distintos ambientes, motivando a que se produjeran cambios adaptativos en la tecnología, subsistencia, estrategias de explotación de recursos y patrones de asentamiento, originando a su vez, distintos *tipos culturales*. Estudios que relevan la arqueología de los ríos Pilcomayo, Bermejo y Paraguay señalan a estos ríos como “enclaves estratégicos” para poder comprender la dinámica poblacional de la región y como corredores importantes que interconectaron el mundo andino, amazónico y pampeano y que generaron un escenario singular que se materializó en un mundo chaqueño (Lamenza et al., 2019). Investigaciones realizadas en el sudeste de la provincia de Formosa evidencian la heterogeneidad de la vegetación, siendo las selvas y los pastizales hidrófilos los que ocupan gran parte de esta área de estudio, y en donde, además, convergen especies de linaje amazónico (Contreras et al., 2015). Algunas de las especies selváticas “no chaqueñas” encontradas ocurren en las Yungas de las Sierras Subandinas, por lo que se sugiere que el río Bermejo y sus tributarios constituirían la principal vía de transporte de dichas especies (Biani et al., 2006; Contreras et al., 2015), mientras que las especies paranaenses encontradas en los bosques de albardón y en los bosques mesófilos del sudeste de Formosa habrían migrado a través de los ríos Paraná y Paraguay (Oakley et al., 2005). Estas investigaciones refuerzan las conexiones a nivel macrorregional a través de las vías fluviales, que habrían permitido el movimiento tanto de poblaciones humanas como de otras formas de vida.

A escala continental, el ingreso a Sudamérica de las poblaciones humanas se ha planteado a través de dos rutas migratorias, una por la costa pacífica y otra por la vertiente atlántica; este modelo implicó la separación temprana de poblaciones andinas (al oeste) y de poblaciones continentales al este, resultando en patrones genéticos distintivos y variabilidad morfológica y cultural entre estas poblaciones (Rothhammer & Dillehay, 2009). En particular, las poblaciones del este, fueron más pequeñas y aisladas, por lo que se favoreció la acción de la deriva génica (Tarazona-Santos et al., 2001). Bortolini et al. (2003) sugieren que la tribalización temprana, y su posterior aislamiento, hizo que fenómenos de naturaleza estocástica tuvieran gran impacto en la historia evolutiva de las poblaciones de Tierras Bajas (Amazonia, Macizo Central Brasileño y Chaco). De esta manera, la gran diferenciación de las poblaciones del este podría estar reflejada en los patrones genéticos de nuestros resultados.

4.3.2 Vínculos interregionales

Los análisis filogenéticos y filogeográficos permitieron detectar un gran número de clados mitocondriales compartidos entre Formosa y las demás localidades del NEA y Litoral. Esta semejanza quedó en evidencia, además, con los análisis de distancias genéticas. En particular, el análisis de escalamiento multidimensional realizado a partir de frecuencias haplotípicas de 201 linajes de ADNmt, en 20 poblaciones de localidades de Argentina, permitió diferenciar a las poblaciones del NEA y Litoral del resto de las poblaciones del NOA y COA. Además, la primera barrera identificada por el programa Barrier, estableció esta misma separación entre las poblaciones del NEA y Litoral del resto de las poblaciones estudiadas. Por otra parte, a partir de una serie de ensayos AMOVA, se obtuvo que el mejor agrupamiento de las poblaciones se logra cuando se ubica a Formosa con el resto de las localidades del NEA y Litoral.

Ciertos clados encontrados en las secuencias analizadas presentaron alta frecuencia y estaban representados por un gran número de haplotipos, por los que se los consideró de posible origen local y propios de la región litoraleña y mesopotámica como A2zh (A2+210), A2zd (A2+15262), B2zd (B2+16178), C1b18, C1b19b, C1b30a1 (C1b30+16126), C1b101 (C1b+16384), C1b102 (C1b+152) y C1b2a1a (C1b+263@).

Por otra parte, ciertos subhaplogrupos encontrados en la muestra de Formosa permitieron establecer vínculos con la Región Chaqueña y con otras regiones del Cono Sur. A su vez, pudieron inferirse clados de llegada más reciente a la región en comparación a otros de origen más antiguo. Se consideró como clados de introducción reciente a aquellos que presentaron baja frecuencia en las localidades del NEA y Litoral y mayor frecuencia y diversidad en otras regiones de Sudamérica, siguiendo el criterio planteado en García et al. (2021). Se definieron seis categorías para describir los vínculos regionales que se describen a continuación:

Vínculos con la región Chaqueña

Contrariamente a lo esperado, dado que Formosa conforma una provincia chaqueña, la frecuencia de clados compartidos con pueblos originarios del Gran Chaco fue muy baja. La vinculación de Formosa con el mundo chaqueño se pudo establecer a partir de los análisis filogenéticos realizados de A2bf (A2+16051). En la muestra analizada en este trabajo se encontró sólo una secuencia correspondiente a este subhaplogrupo, mientras que, en otros trabajos realizados en la región, se lo encontró mayoritariamente en población indígena del NEA, particularmente en muestras del Chaco (Cabana et al., 2006; Sevini et al., 2014), por lo que se propone que éste podría ser su lugar de origen. La diversificación de ciertos clados de menor jerarquía al interior de A2bf podrían haber ocurrido por diferenciación local. En este sentido, el análisis filogeográfico realizado permitió reconocer una distribución geográfica particular de

estos sub-clados: A2bf1 se encontró en secuencias de localidades del NEA y especialmente en población wichi, mocoví y pilagá, en Chile, Bolivia y muestras del NOA; A2bf1a, se identificó en población wichi y criolla del NEA; mientras que A2bf2 está restringido a Uruguay. La baja representatividad de linajes chaqueños en Formosa da cuenta de una conformación de las sociedades de esta localidad a partir de población mestiza que habría arribado al territorio en migraciones postconquista. Tal como plantea Maeder (2015), el poblamiento de Chaco por parte de la sociedad criolla fue reciente, como resultado de inmigración tanto interna como externa; desplazando a los grupos indígenas que ya se encontraban en el territorio, los cuales tuvieron que ocupar espacios marginales.

Vínculos con Amazonas, el Alto Paraná y Uruguay

Una fracción importante de los linajes encontrados en las muestras de Formosa y compartidos por otras muestras del NEA y Litoral permiten vincular esta región con Brasil, Uruguay y Paraguay. Algunos de ellos son: A2aa, A2ah, A2be, A2zj, A2yc, A2ze, A2zz, A2za1a, B2b3a, B2e, B2h, C1b2, C1d1d, B2an, B2i1, D1e y D1f1.

▪ *Vínculos con la Amazonía brasilera*

Particularmente, los clados B2an y D1e que fueron analizados a nivel sudamericano, en conjunto con la presencia de B2i1 en secuencias de Formosa y del NEA y Litoral, permiten establecer afinidades con la Amazonia brasilera.

La distribución sudamericana de B2an que se pudo reconstruir a partir de los datos de este trabajo conjuntamente con los aportados por Figueiro et al. (2022), permite proponer un origen de este linaje en la Amazonia brasilera, una diferenciación temprana de tres clados B2an1, B2an2 y B2an3 y una expansión relativamente rápida por el este, alcanzando el sur de Brasil, Paraguay y el NEA y Litoral (Figura 4.1). En este sentido, se observa una distribución particular de cada clado: para B2an1 se propone una expansión siguiendo una ruta atlántica hasta el sur de Brasil en donde se diferencian posteriormente otros clados locales B2an1a y B2an1a1, llegando este linaje hasta Uruguay. Para B2an3 se observa una expansión más acotada que alcanza el sur de Brasil, no habiéndose encontrado secuencias de este linaje en otras regiones hasta el momento. Por último, B2an2 se expande por el centro-este de Brasil hasta alcanzar el este de Paraguay, desde donde se propone ingresaría a la región del NEA y Litoral y a Uruguay. Este subhaplogrupo se encontró en representantes de población Mbya Guaraní de Misiones (Sala et al., 2010) por lo que se plantea que podría estar asociado con la expansión guaraní, de llegada tardía a nuestro territorio. Al respecto, Bonomo et al. (2015) realizaron un estudio detallado de la distribución geográfica y temporal de los sitios arqueológicos atribuidos a los guaraníes y detallan la expansión de los mismos por vías fluviales, a través del relevamiento

de las coordenadas geográficas de los sitios y dataciones disponibles y de análisis espaciales con herramientas SIG. Como evidencia en apoyo a esta hipótesis, el área de expansión guaraní sugerida por estos autores coincide con la propuesta para este linaje. Por último, se considera que B2an2a se habría diferenciado localmente en la región del NEA y Litoral. La llegada de este linaje a Río Negro, según se observa en una secuencia reportada por Bobillo et al. (2010), se propone que habría ocurrido por migraciones recientes.

Otros linajes que dan cuenta de la conexión entre Formosa y el mundo guaraní fueron A2za1a (KJ923817, Rieux et al., 2014) y B2zf1 (EU095219, Fagundes et al., 2008) ambos presentes en población guaraní.

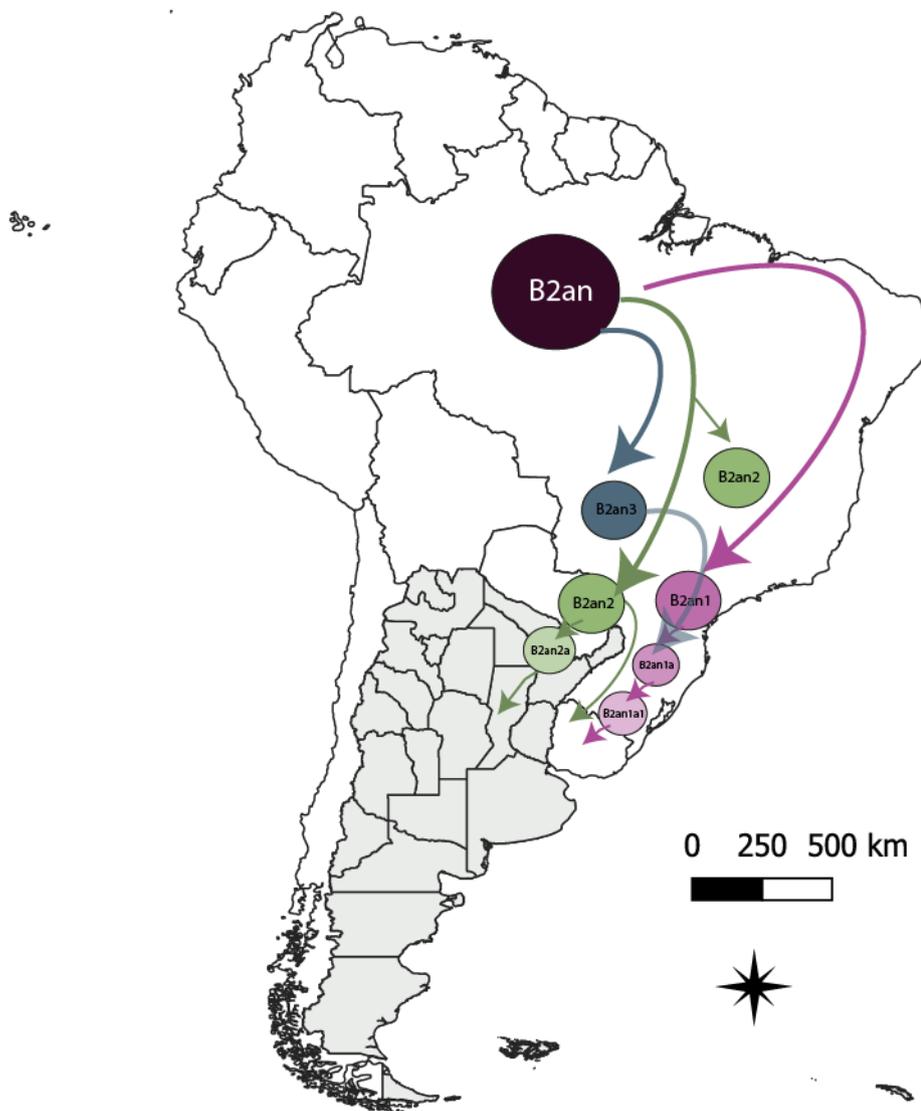


Figura 4.1 Rutas de migración de B2an inferidas en base a los datos filogeográficos

En base a los datos aportados en este trabajo se sugiere un origen amazónico de D1e. El clado de distribución brasilera encontrado en población karitiana (D1e2) no se encuentra anidado con el clado encontrado en Paraguay y NEA y Litoral (D1e1), por lo que se hipotetiza una separación temprana en dos grupos poblacionales que se diferenciaron localmente, uno que permaneció en Brasil y otro que alcanzó Paraguay, en conexión permanente con el NEA y Litoral.

La presencia de B2i1 pone en evidencia el vínculo entre Formosa y el resto de las localidades del NEA y Litoral con el Amazonas brasilero. Este clado fue identificado por Fagundes et al. (2008) en población kayapó del Amazonas central de Brasil. Tanto B2i1 como B2i2 son grupos hermanos que pertenecen a B2i, dando cuenta de una separación temprana de dos grupos poblacionales: uno de ellos, según la hipótesis propuesta por de Saint Pierre et al. (2012b), se habría expandido siguiendo la ruta costera del pacífico hasta la Patagonia argentina (B2i2) y el otro siguiendo una ruta continental al este de Sudamérica (B2i1).

Por otra parte, A2aa y en particular A2aa+16153+16555 permitió vincular las secuencias de Formosa y del NEA y Litoral con secuencias del sur de Brasil (Ávila et al., 2022).

- *Vínculos con Alto Paraná*

Ciertos clados de menor jerarquía al interior de linajes como B2i1 y A2yc encontrados en Formosa, permitieron establecer afinidades con el Alto Paraná en Paraguay. Si bien estos dos linajes se han reportado en secuencias de Brasil (Fagundes et al., 2008; Ávila, et al., 2019; Ávila et al., 2022), los haplotipos encontrados en este trabajo identificados como A2yc1a y B2i1a1 (B2i1a+161479A+34) son compartidos con mitogenomas de Paraguay (MH981840 y MH981885; Simão et al., 2019). Al interior del clado C1b18, otro haplotipo de Formosa fue identificado como C1b18b (C1b18+16187) (García et al., 2021), reforzando el vínculo con Paraguay (MH981889, MH981891, MH981918, Simão et al., 2019). Linajes como B2h y A2ze (A2+3083-16422) vinculan a las secuencias analizadas aquí con la población brasilera y con los aché de Paraguay (Fagundes et al., 2008; Rieux et al., 2014). A2zj, de gran representación en las secuencias analizadas aquí, también permitió establecer un vínculo con secuencias del este de Paraguay (Simão et al., 2021).

- *Vínculos con Uruguay*

Ciertos linajes identificados como subhaplogrupos mayoritarios en las secuencias de Formosa, permitieron establecer vínculos con Uruguay: A2be, B2an, B2zd, B2e y C1d1d. En particular, A2be representó alrededor de un 5% de los linajes nativos de Formosa y del NEA y Litoral y un 30% de las secuencias analizadas pertenecientes al haplogrupo A2₍₆₄₎ de estas poblaciones. A su vez, la identificación de ciertos clados de menor jerarquía permitieron vincular

a las poblaciones analizadas aquí con otras de Uruguay (Figueiro et al., 2022) y de Brasil (Ávila et al., 2022).

- *Clados pan-sudamericanos*

Dos de los clados mayoritarios encontrados en las secuencias analizadas de Formosa fueron considerados de diferenciación temprana teniendo en cuenta su distribución a nivel sudamericano. Uno de ellos es D1f1 que se describe en la siguiente sección, ya que no sólo permite establecer vínculos entre las secuencias aquí analizadas con Brasil, sino también con Venezuela y el Caribe. El otro linaje es B2e.

La distribución continental de B2e que se conoce hasta la fecha, conjuntamente con los datos aportados en este trabajo, nos permiten proponer un origen de este linaje en el norte de Sudamérica, y una expansión relativamente rápida por el este (Figura 4.2). Los clados de distribución sureña, B2e1 y B2e2, no están anidados dentro de los clados de distribución norteña (B2e3-B2e6), por lo que en principio B2e se habría diferenciado tempranamente en el norte de Sudamérica y su expansión sería anterior a la etnogénesis del pueblo guaraní. Las estimaciones arqueológicas y lingüísticas sitúan el origen de la familia Tupí-guaraní hace aproximadamente 2500 años (Corrêa, 2023). Se observa que la distribución geográfica de B2e1 y B2e2 coincide con el área de expansión guaraní propuesta por Bonomo et al. (2015). Sin embargo, la llegada al territorio argentino puede ser tanto anterior a la expansión guaraní (nótese la presencia de B2e1 en una muestra de la BBDD-RI perteneciente a población tehuelche de Chubut, aunque sin lugar de nacimiento conocido) como posterior, debido al flujo migratorio desde Brasil, Paraguay y Uruguay, como consecuencia de fenómenos de migración postconquista.

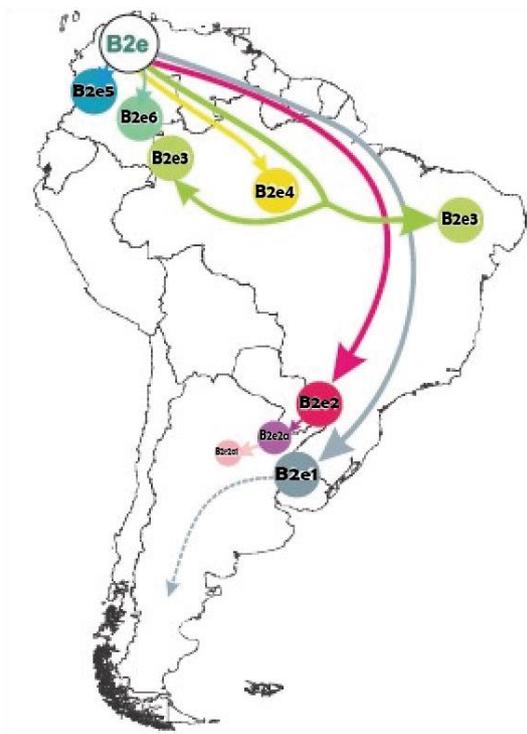


Figura 4.2 Rutas de migración de B2e inferidas en base a los datos filogeográficos

 *Vínculos con Venezuela y el Caribe*

Gran parte de los subhaplogrupos compartidos por poblaciones del NEA y Litoral con Paraguay, Uruguay y Brasil también fueron encontrados en el Caribe o Venezuela, como por ejemplo los casos de A2ah, A2zk3, B2b3a, C1b2 y D1f (Marrero et al., 2007; Vilar et al., 2014; Sans et al., 2015; Gómez-Carballa et al., 2018; Brandini et al., 2018; Nieves-Colón et al., 2020; Fernandes et al., 2021; Figueiro et al., 2022; García-Olivares 2023).

En particular, D1f1 fue uno de los haplogrupos mayoritarios encontrados en la localidad de Formosa. Del análisis filogeográfico realizado y teniendo en cuenta el origen propuesto de D1f en Norte o Centro América por Brandini et al. (2018), se puede inferir un origen de D1f1 en Venezuela y una posterior migración siguiendo una ruta atlántica alcanzando el norte de Brasil, sur de Brasil, Paraguay y el NEA y litoral. Dada la escasa divergencia entre las secuencias encontradas en Venezuela y en el resto de las poblaciones analizadas, se plantea que se trata de una expansión muy tardía de estos grupos poblacionales. Una expansión más reciente podría relacionarse al mundo guaraní, sin embargo, la ausencia de guaraníes en Venezuela descartaría este escenario.

Por otra parte, la búsqueda de posiciones informativas por tamizaje y secuenciación parcial de ciertos fragmentos de la región control realizadas para el haplogrupo A2 (para más detalle ver Anexo 4.1), permitió identificar dos clados que refuerzan las evidencias en favor de los vínculos del NEA y Litoral con Venezuela y el Caribe. El primero es el clado A2zk3 (A2+16266+14364) (o A2ar, García-Olivares, 2023) y el segundo es A2zc1, ambos se encuentran

ampliamente distribuidos en la Amazonía de Venezuela (Makiritare), en el Caribe (Puerto Rico, Rep. Dominicana) (Fernandes et al., 2021) y en población cosmopolita de Brasil (Ávila et al., 2019; Ávila et al., 2022) y Paraguay (Simão et al., 2019; Simão et al., 2022). Además, A2zc1 se encontró en muestras antiguas del Caribe (Vilar et al., 2014); en Argentina, particularmente en este trabajo, A2zk3 se encontró en cuatro individuos de Formosa y uno de Tartagal, mientras que A2zc1 se encontró en una muestra de Formosa.

A2ah, B2b3a, C1b2 también están presentes en poblaciones modernas y antiguas del Caribe, adonde habrían arribado presumiblemente desde lo que hoy es Venezuela (Vilar et al., 2014, Nieves-Colón et al., 2020; Fernandes et al., 2021)

Vínculos con la Región Central de Argentina

Otros clados de interés encontrados en las muestras de Formosa y compartidos por las secuencias del NEA y Litoral fueron C1b+146 (C1b16, García et al., 2021), D1j y C1d1b. Motti (2012) y Motti et al. (2017) han descripto a estos tres linajes como indicadores de un vínculo estrecho entre las poblaciones del centro y el centro oeste argentino, y de la permanencia de flujo génico entre ellas.

En particular, García (2012) y Motti (2012) plantearon un origen local para D1j en el área central del país, ya sea en las sierras centrales o en los valles riojanos, basándose en las altas frecuencias encontradas de este sub-haplogrupo. En este trabajo, los linajes D1j1a (D1j1+5004) y D1j1a1 (D1j1a+5325T y D1j1a+5325T+16172) se identificaron en muestras de Formosa, y fueron compartidos por muestras de Tartagal y de otras secuencias del NEA y Litoral.

El clado C1d1b se ha encontrado únicamente en muestras del sur de Sudamérica (Perego et al., 2010) y ha sido reportado por distintos autores en poblaciones de Argentina (García et al., 2021; Perego et al., 2010; Schwab, 2018), con una distribución continua que alcanza su frecuencia máxima en la región central y que disminuye gradualmente hacia el norte y hacia el sur (García et al., 2018; Pauro, 2015). Además, recientemente se han informado representantes de este linaje en Paraguay (Simão et al., 2019). En este trabajo, los haplotipos encontrados en muestras de Formosa, por ejemplo: C1d1b2 (C1d1b₊₁₆₂₅₉₋₁₆₂₇₁₋₁₆₃₁₁) han sido identificados en poblaciones contemporáneas del NOA (Motti, 2012; Perego et al., 2010), en Montevideo (Figueiro et al., 2011), en el Gran Chaco (Cabana et al., 2006; Sevini et al., 2014) y en el centro de Argentina (García et al., 2018). Además, este haplotipo se ha encontrado tanto en muestras actuales como arqueológicas de las provincias de Santiago del Estero y Córdoba, lo que sugiere un flujo génico entre la región central de Argentina y el Gran Chaco (Schwab, 2018; Tavella et al., 2020; García et al., 2021), que queda en evidencia por la presencia de este haplotipo en las secuencias de Formosa analizadas en este trabajo.

Por último, C1b16 es un clado propio del Centro y Centro-Oeste Argentino y presenta vínculos con otras regiones del cono sur como Bolivia y el Gran Chaco (García et al., 2021). Los haplotipos encontrados en las muestras de Formosa pudieron ser identificados como C1b16 y C1b16a, siendo este clado el que se encontró en mayor frecuencia en las secuencias de la BBDD-RI analizadas y asociado a secuencias de Tucumán, Santiago del Estero y Tartagal y en menor frecuencia a secuencias de Cuyo.

Vínculos con los Andes Centrales

La presencia de un único haplotipo asignado a B2aj en Formosa junto a su ausencia en otras localidades del NEA y Litoral, a excepción de Tartagal, permiten dar cuenta de la escasez de vínculos de la región estudiada con la región andina. La presencia de B2aj en Tartagal se explica debido a la ubicación de esta ciudad en el punto de contacto entre el mundo andino y el mundo chaqueño. Este linaje está mayormente distribuido en Perú, Bolivia y la porción argentina de los Andes; su alta frecuencia en Bolivia en conjunto con ciertos subhaplogrupos compartidos entre Bolivia y Argentina, sugieren un patrón de dispersión y flujo génico desde Bolivia a la región andina norte de Argentina (Gómez-Carballa et al., 2018). Los haplotipos compartidos entre la región andina de Bolivia y Argentina señalados por estos autores, coinciden con los encontrados en este trabajo: B2aj4a, B2aj3a y B2aj3c, poniendo en evidencia un flujo génico favorecido de norte a sur y limitado con las poblaciones del este.

Vínculos con Patagonia

En este trabajo no se encontraron evidencias de vínculos entre Formosa ni el resto de las localidades del NEA y Litoral con Patagonia. En una única muestra de Reconquista se encontró un individuo asignado a D1g5. Sin embargo, tratándose del subhaplogrupo de más amplia distribución de D1g, que no se restringe a Patagonia y que se ha encontrado además en poblaciones actuales de las provincias de Catamarca, La Rioja, San Juan, Mendoza, Córdoba y San Luis (Motti, 2012; García et al., 2018; García et al., 2021) y en la región central de Chile (Gómez-Carballa et al., 2016), se concluye que su presencia en el Litoral, no implica vínculos con esa región. Además, es interesante resaltar la total ausencia de clados patagónicos C1b13 y B2i2 y D4h3a5 tanto en Formosa como en el resto de las localidades del NEA y Litoral.

En resumen, los análisis realizados en este trabajo, en conjunto con los resultados preliminares de Beltramo (tesis doctoral en curso), indican que los linajes maternos que predominan en Formosa y en el resto de las poblaciones del NEA y Litoral, son particulares de esta región, demostrando que comparten una historia biológica común, y que, a su vez, se diferencian de los clados dominantes de otras regiones como el centro, centro-oeste, noroeste

y sur de Argentina. Una fracción importante de los linajes que caracterizan a las poblaciones del NEA y el Litoral permiten establecer vínculos con poblaciones de Brasil, Uruguay y Paraguay, más que con poblaciones del resto de Argentina. Pese a que los análisis realizados se basan mayormente en la variabilidad encontrada en la región control (RC) de poblaciones argentinas, el abordaje metodológico implementado en este trabajo permitió aumentar la resolución de las secuencias obtenidas, a partir de secuenciación parcial de la región codificante. De esta manera, se lograron identificar ciertos grupos monofiléticos de linajes que no podrían haber sido reconocidos únicamente por variación en la RC. Además, la recopilación de secuencias en la bibliografía y en base de datos genéticas de ciertos clados de interés, permitió reconstruir su distribución a nivel sudamericano. Las afinidades con Amazonia y en particular, con el mundo guaraní dominaron el paisaje formoseño, en comparación con los vínculos con el mundo chaqueño.

5 CAPÍTULO 5: CONCLUSIONES

- Los linajes nativos de Formosa y Corrientes reflejan un aporte diferencial y de tipo sexo asimétrico al acervo génico, que fue dominante para los linajes mitocondriales (frecuencias mayores al 85% en ambas localidades) y minoritario para los linajes paternos (frecuencias menores al 5% en ambas localidades).
- La contribución que realiza el clado Q-M3 de las localidades de Formosa y Corrientes al acervo génico, corresponde a las frecuencias más bajas informadas hasta el momento (3,2% para Corrientes y 4,9% para Formosa) junto con las de Santiago del Estero (5,3%), en comparación a otras 14 localidades de Argentina.
- A partir de las encuestas genealógicas, se registra una mayor preservación y transmisión oral acerca del lugar de nacimiento de las abuelas maternas, respecto del de los abuelos paternos.
- La comparación de los datos moleculares y los genealógicos permite reconocer que los donantes declaran mayoritariamente un origen regional de sus ancestros paternos y maternos más remotos:

- ✓ para la línea materna, se corroboró un origen regional declarado con el resultado molecular;
- ✓ sin embargo, para la línea paterna hubo discrepancias entre el origen declarado (regional) y el resultado molecular (euroasiático) en la mayoría de los casos.

Estos resultados podrían ser el reflejo de las sociedades criollas en tiempos coloniales y de la asimetría del mestizaje en la población criolla.

- El análisis de secuencias de alta calidad de la región control completa en combinación con la secuenciación parcial de región codificante permitió:
 - ✓ definir con certeza nuevos grupos monofiléticos de linajes maternos;
 - ✓ demostrar la asociación de varios haplotipos de la región control con linajes descritos previamente;
 - ✓ mapear la distribución y frecuencia relativa de todos estos clados en Formosa en otras localidades del NEA y Litoral y en distintas regiones tanto de Argentina como de países limítrofes.

Se considera, por lo tanto, una herramienta que permite subhaplogrupado de alta resolución.

- Los análisis en relación a las distancias genéticas permiten agrupar a la localidad de Formosa con el resto de las localidades del NEA y Litoral: Reconquista, Cruzú-Cuatiá, La Paz y Gualguaychú y separarlas del resto de las poblaciones del NOA y COA.

- Los haplotipos compartidos entre las localidades del NEA y Litoral corroboran una gran afinidad entre estas poblaciones y permiten visualizar vínculos principalmente con poblaciones de la amazonia brasilera (B2an, D1e, B2i1), Paraguay y Uruguay (A2be, B2zd). Además, otros linajes permitieron establecer afinidades con Venezuela y el Caribe (D1f1, A2ah, B2b3a, C1b2).
- Otros linajes encontrados permitieron vincular a Formosa y al NEA y Litoral con el Centro de Argentina (C1d1b, C1b16 y D1j) y con los Andes Centrales (B2aj). No se encontraron vínculos con Patagonia.
- Pese a que Formosa conforma una provincia chaqueña, se registró una baja frecuencia de clados compartidos con pueblos originarios del Gran Chaco (sólo una secuencia fue asignada a A2bf).
- Los análisis de la distribución sudamericana de los clados mitocondriales realizados en este trabajo, en particular de B2e, B2an, A2bf, D1e y D1f1 permitieron proponer un origen y una posible ruta de expansión para ciertos clados de interés.
- Profundizar en la filogenia de los haplogrupos basales y de los sub-haplogrupos con distribución local diferencial, permitió revisar los modelos de poblamiento planteados hasta el momento y realizar un aporte para la generación de modelos de poblamiento regional.

Corolario

En los últimos años, las investigaciones que analizan la variabilidad genética de las poblaciones argentinas han aumentado considerablemente, generando una gran cantidad de datos moleculares. Sin embargo, son las interpretaciones bioantropológica, histórica y demográfica las que colaboran a dar un marco de análisis a los datos genéticos, a fin de poder reconstruir las historias fracturadas y suprimidas de las poblaciones humanas.

A partir del diseño metodológico planteado en este trabajo y de la investigación que se deriva del mismo, se contribuyó a incrementar el conocimiento de la historia local, a través de la aplicación de técnicas que aportan una mirada independiente e innovadora y cuyos resultados permiten sumar un nuevo abordaje a las interpretaciones sobre los procesos que intervinieron en la conformación de las poblaciones del Nordeste argentino.

Futuras investigaciones podrán valerse de los datos generados en este trabajo para seguir reconstruyendo las memorias e identidades regionales de nuestro país y dar cuenta del complejo entramado cultural que habita cada región.

6 REFERENCIAS

1000 Genomes Project Consortium. (2010). A map of human genome variation from population scale sequencing. *Nature*, 467(7319), 1061. <https://doi.org/10.1038/nature09534>.

Achilli A., Perego U. A., Bravi C. M., Coble M. D., Kong Q. & Scott R. (2008). The phylogeny of the four pan-American mtDNA haplogroups: Implications for evolutionary and disease studies. *PLoS One*, 3(3), e1764. <https://doi.org/10.1371/journal.pone.0001764>.

Aizpurua-Iraola J., Abdeli A., Benhassine T., Calafell F. & Comas, D. (2023). Whole mitogenomes reveal that NW Africa has acted both as a source and a destination for multiple human movements. *Scientific Reports*, 13(1), 10395. <https://doi.org/10.1038/s41598-023-37549-4>.

Alves-Silva J., da Silva Santos M., Guimarães P. E., Ferreira A. C., Bandelt H. J., Pen, S. D., & Prado V. F. (2000). The ancestry of Brazilian mtDNA lineages. *The American Journal of Human Genetics*, 67(2), 444-461. <https://doi.org/10.1086/303004>.

Anderson S., Bankier A. T., Barrell B. G. G., Bruijn M. D., De Bruijn M. H., Coulson A. R., ... Young I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature*, 290(April), 457-465. <https://doi.org/10.1038/290457a0>.

Andrews R. M., Kubacka I., Chinnery P. F., Lightowlers R. N., Turnbull D. M. & Howell N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genetics*, 23(2), 147. <https://doi.org/10.1038/13779>.

Aquilano E., de la Fuente C., Motti J.M.B., Rodríguez Golpe D., Bravi C.M. (2023). Sequencing errors in Native American mitogenomes: impact on clade definitions, haplogroup assignation, and beyond. *Mitochondrion*. 70:54-58. <https://doi.org/10.1016/j.mito.2023.03.005>

Arias, L., Barbieri, C., Barreto, G., Stoneking, M. & Pakendorf, B. (2017) High-resolution mitochondrial DNA analysis sheds light on human diversity, cultural interactions and population mobility in Northwestern Amazonia. *American Journal of Physical Anthropology*, 165, 238-255. <https://doi.org/10.1002/ajpa.23345>.

Avena S., Goicoechea A., Bartomioli M., Fernández V., Cabrera A., Dugoujon J.M., Dejean C., Fabrykant G., Carnese F.R. (2007) Mestizaje en el sur de la región pampeana (Argentina). Su estimación mediante el análisis de marcadores proteicos y moleculares uniparentales. *Revista Argentina de Antropología Biológica* 9: 59-76.

Avena S.A., Goicoechea A., Rey J. et al. (1999). Análisis de la Participación del Componente Indígena en una Muestra Hospitalaria de la Ciudad de Buenos Aires. *Revista Argentina de Antropología Biológica* 2: 211-225.

Avena S.A., Parolin M.L., Dejean C.B., Ríos Part M.C., Fabrikant G., Goicoechea A.S., Dugoujon J.M. y Carnese F.R. (2009) Mezcla génica de linajes uniparentales en Comodoro Rivadavia

(Provincia de Chubut, Argentina). *Revista Argentina de Antropología Biológica*, 11(1), 25-41. Recuperado de: <https://revistas.unlp.edu.ar/raab/article/view/261>.

Ávila E., Graebin P., Chemale G., Freitas J., Kahmann A., Alho C.S. (2019). Full mtDNA genome sequencing of Brazilian admixed populations: A forensic-focused evaluation of a MPS application as an alternative to Sanger sequencing methods. *Forensic Science International: Genetics* 42: 154-164. <https://doi.org/10.1016/j.fsigen.2019.07.004>.

Ávila E., Speransa P.A., Lindholz C.G., Kahmann A., Alho C.S. (2022). Haplotype distribution in a forensic full mtDNA genome database of admixed Southern Brazilians and its association with self-declared ancestry and pigmentation traits. *Forensic Science International: Genetics* 57:102650. <https://doi.org/10.1016/j.fsigen.2021.102650>.

Avice J.C. (2000). *Phylogeography. The history and formation of species* (Harvard Un). Cambridge, Massachusetts.

Badano I., Sanabria, D. J., Totaro, M. E., Rubinstein, S., Gili, J. A., Liotta, D. J., et al. (2018). Mitochondrial DNA ancestry, HPV infection and the risk of cervical cancer in a multiethnic population of northeastern Argentina. *PloS One*, 13(1), e0190966. <https://doi.org/10.1371/journal.pone.0190966>.

Baeta M., Núñez C., Cardoso S., Palencia-Madrid L., Piñeiro-Hermida S., Arriba-Barredo M., Villanueva-Millán M.J., M de Pancorbo M. (2015). Different Evolutionary History for Basque Diaspora Populations in USA and Argentina Unveiled by Mitochondrial DNA Analysis. *PLoS One*;10(12):e0144919. <https://doi.org/10.1371/journal.pone.0144919>

Bailliet G., Ramallo V., Muzzio M., García A., Santos M. R., Alfaro E. L., et al. (2009). Brief communication: Restricted geographic distribution for Y-Q* paragroup in South America. *American Journal of Physical Anthropology*, 140(3), 578–582. <https://doi.org/10.1002/ajpa.21133>.

Bailliet G., Ramallo V., Muzzio M., Santos M. R., Motti J. M. B., Bianchi N. O., & Bravi C. M. (2011). Antecedentes y nuevos aportes en el estudio del cromosoma Y en poblaciones humanas sudamericanas. *Journal of Basic and Applied Genetics*, 22(1), 1–9.

Bailliet G., Rothhammer F., Carnese F., Bravi C. M. & Bianchi N. (1994). Founder mitochondrial haplotypes in amerindian populations. *American Journal of Human Genetics*, 55(1), 27–33.

Bandelt H. J., Forster P. & Röhl A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular biology and evolution*, 16(1), 37-48.

Bandelt H.J., Forster P., Sykes B., Richards M. (1995). Mitochondrial portraits of human populations using median networks. *Genetics*; 147:743-753.

Baratta M. V. (2013). La Guerra del Paraguay y la historiografía argentina. *História da Historiografia: International Journal of Theory and History of Historiography*, Ouro Preto, 7(14),

98–115. Disponible en línea:
<https://www.historiadahistoriografia.com.br/revista/article/view/614>.

Barbieri C., Sandoval J. R., Valqui J., Shimelman A., et al. (2017). Enclaves of genetic diversity resisted Inca impacts on population history. *Scientific Reports*, 7(1), 1–12. <https://doi.org/10.1038/s41598-017-17728-w>

Barbosa A. B., da Silva L. A., Azevedo D. A., Balbino V. Q. & Mauricio-da-Silva L. (2008). Mitochondrial DNA control region polymorphism in the population of Alagoas state, north-eastern Brazil. *Journal of Forensic Sciences*; 53(1):142–146. <https://doi.org/10.1111/j.1556-4029.2007.00619.x>

Behar D.M., Villems R., Soodyall H., Blue-Smith J., Pereira L., Metspalu E., Scozzari R., Makkan H., Tzur S., Comas D., Bertranpetit J., Quintana-Murci L., Tyler-Smith C., Wells R.S., Rosset S.; Genographic Consortium. (2008). The dawn of human matrilineal diversity. *American Journal of Human Genetics*, May; 82(5): 1130-40.

Behar D.M., van Oven M., Rosset S., Metspalu M., Loogväli E.L., Silva N.M., Kivisild T., Torroni A., Villems R. (2012) A "Copernican" reassessment of the human mitochondrial DNA tree from its root. *American Journal of Human Genetics*, Apr 6; 90(4): 675-84.

Beltramo J (en curso). Origen continental, afinidades filogeográficas y estructuración geográfica de los linajes maternos y paternos en poblaciones humanas del Nordeste y Litoral de Argentina. Manuscrito de tesis doctoral en preparación, Universidad Nacional de La Plata.

Bernardo S., Hermida R., Desidério M., Silva D. A., & de Carvalho E. F. (2014). MtDNA ancestry of Rio de Janeiro population, Brazil. *Molecular biology reports*, 41, 1945-1950.

Bianchi N.O., Catanesi C.I., Bailliet G., Martinez-Marignac V.L., Bravi C.M., Vidal-Rioja L.B., Herrera R.J., López-Camelo J.S. (1998). Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. *American Journal of Human Genetics*. Dec; 63 (6): 1862-71.

Biani N., Vesprini J. & Prado D. (2006). Conocimiento sobre el gran Chaco Argentino en el siglo XX. En Arturi, M. F., Frangi J. L. & Goya, J. F. (eds.). *Ecología y Manejo de los bosques de Argentina*, pp. 1-19. La Plata.

Bisso-Machado R., Ramallo V., Salzano F.M., Bortolini M.C. (2016). La dispersión de Homo sapiens y el poblamiento temprano de América. En: Madrigal L & González-José, R. *Introducción a la Antropología Biológica*. Asociación Latinoamericana de Antropología Biológica. ISBN: 978-987-33-9562-8. 467-494

Bobillo M.C., Zimmermann B., Sala A., Huber G., Röck A., Bandelt H.J., Corach D., Parson W. (2010) Amerindian mitochondrial DNA haplogroups predominate in the population of Argentina: towards a first nationwide forensic mitochondrial DNA sequence database. *International Journal of Legal Medicine* 124(4), 263-8.

Bodner M., Perego U. A., Huber G., Fendt L., Röck A. W., Zimmermann B. ... & Bobillo M. C. (2012). Rapid coastal spread of First Americans: Novel insights from South America's Southern Cone mitochondrial genomes. *Genome Research*, 22(5), 811–820. <https://doi.org/10.1101/gr.131722.111>

Bonomo M. (2013). Reanálisis de la colección de Samuel Lothrop procedente del delta del Paraná. *Relaciones de la Sociedad Argentina de Antropología*, XXXVIII (1), 169-198.

Bonomo M., Cabanillas E.D & Montero R. (2017). Archaeometallurgy in the Paraná Delta (Argentina): Composition, manufacture, and Indigenous routes. *Journal of Anthropological Archaeology* 47:1–11.

Bonomo M., Costa Angrizani R., Apolinaire E. & Noelli F. S. (2015) A model for the Guaraní expansion in the La Plata Basin and littoral zone of southern Brazil, *Quaternary International*, Volume 356, 54-73 <https://doi.org/10.1016/j.quaint.2014.10.050>.

Bonomo M., Di Prado V., Silva C., Scabuzzo C., Ramos van Raap M.A., Castiñeira C., Colobig M.M y Politis G. (2019) Las poblaciones indígenas prehispánicas del río Paraná Inferior y Medio. En: Dossier Arqueología y ríos de las Tierras Bajas de América del Sur. *Revista del Museo de La Plata*. Vol 4, (2): 585-620.

Bonomo M., G. Politis, L. Bastourre & Moreira G. (2021). Humanized nature: Symbolic representation of fauna in pottery from the Paraná River of South America. In *South American contributions to world archaeology*. Edited by M. Bonomo and S. Archila, 411–46. Springer-Nature.

Bonomo M., Politis G.G. & Gianotti G. (2011). Montículos, jerarquía social y horticultura en las sociedades indígenas del Delta del río Paraná (Argentina). *Latin American Antiquity* 22(3):297–333.

Bortolini M. C., Salzano F. M., Thomas M. G., Stuart S., Nasanen S. P. K., Bau C. H. D., ... Ruiz-Linares A. (2003). Y-chromosome evidence for differing ancient demographic histories in the Americas. *American Journal of Human Genetics*, 73(3), 524–539. <https://doi.org/10.1086/377588>.

Brandini S., Bergamaschi P, Cerna MF, Gandini F, et al. (2018) The Paleo-Indian Entry into South America According to Mitogenomes. *Molecular Biology and Evolution* 35(2):299-311

Brandstätter A., Sanger T, Lutz-Bonengel S, Parson W, Beraud-Colomb E, et al. (2005) Phantom mutation hotspots in human mitochondrial DNA. *Electrophoresis* 26(18): 3414-29. <https://doi.org/10.1002/elps.200500307>.

Brandstätter A., Peterson C.T., Irwin J.A., Mpoke S., Koech D.K., Parson W. & Parsons T.J. (2004). Mitochondrial DNA control region sequences from Nairobi (Kenya): Inferring phylogenetic parameters for the establishment of a forensic database. *International Journal of Legal Medicine*, 118(5), 294–306. <https://doi.org/10.1007/s00414-004-0466-z>.

Bravi C. M. (2013). La evidencia molecular del poblamiento humano de América; Universidad Nacional de La Plata, Facultad de Ciencias Naturales y Museo; Revista del Museo de La Plata; 20; 177; 8-2013; 40-57.

Bravi C.M., Bailliet G., Martinez-Marignac V.L., Bianchi N.O. (2000). Origin of YAP+ lineages of the human Y-chromosome. *American Journal of Physical Anthropology*. Jun; 112(2): 149-58.

Bravi C. M., Motti J.M.B., Beltram, J., Schwab M., Rodríguez Golpe D., Aquilano E. (2022). Letter to the Editor: The Devil is in the dental. *American Journal of Biological Anthropology*, 1-3. <https://doi.org/10.1002/ajpa.24627>

Brochado P.J. (1984) An Ecological Model of the Spread of Pottery and Agriculture into Eastern South America (Tesis doctoral). Department of Anthropology. University of Illinois at Urbana-Champaign.

Brown W. M. (1980). Polymorphism in mitochondrial DNA of humans as revealed by restricción endonucleasa análisis. *PNAS*, 77(6), 3605–3609. <https://doi.org/10.1073/pnas.77.6.3605>.

Bruniard, E. & Bolsi, A. (1992). Región agro-silvo-ganadera con frentes pioneros de ocupación del Nordeste. En: Roccatagliata, J (coord.) La Argentina. Geografía general y los marcos regionales. Buenos Aires, Planeta.

Cabana G., Merriwether D.A., Hunley K., Demarchi D.A (2006). Is the Genetic Structure of Gran Chaco Populations Unique? Interregional Perspectives on Native South American Mitochondrial DNA Variation. *American Journal of Physical Anthropology*, 131:108-119.

Caggiano M.A. (1984). Prehistoria del noreste argentino, sus vinculaciones con la República Oriental del Uruguay y sur de Brasil. *Pesquisas, Antropología* 38, pp. 1-109.

Calandra H.A. & Salceda, S.A. (2004). El territorio y sus ocupantes: ¿Qué, quienes cómo y cuándo? *Folia Histórica del Nordeste* 15, pp.107-128.

Campelo dos Santos A.L., Owings A., Sullasi H.S.L., Gokcumen O., DeGiorgio M., Lindo J. (2022) Genomic evidence for ancient human migration routes along South America's Atlantic coast. *Proceedings of the Royal Society B* 289: 20221078. <https://doi.org/10.1098/rspb.2022.1078>

Capodiferro M. R., Aram B., Raveane A., Migliore N. R., Colombo G., Ongaro L., ... & Achilli A. (2021). Archaeogenomic distinctiveness of the Isthmo-Colombian area. *Cell*, 184(7), 1706-1723.

Cardena M. M., Ribeiro-dos-Santos A., Santos S., Mansur A. J., Pereira A. C., & Fridman C. (2013). Assessment of the relationship between self-declared ethnicity, mitochondrial haplogroups and genomic ancestry in Brazilian individuals. *PloS One*, 8(4), e62005. <https://doi.org/10.1371/journal.pone.0062005>.

Cardoso S., Palencia-Madrid L., Valverde L., Alfonso-Sanchez M. A., Gomez-Perez L., Alfaro, E. ... & de Pancorbo M. M. (2013). Mitochondrial DNA control region data reveal high prevalence

of Native American lineages in Jujuy province, NW Argentina. *Forensic Science International: Genetics*, 7(3), e52-e55. <https://doi.org/10.1016/j.fsigen.2013.01.007>.

Carracedo A., Bär W., Lincoln P. J., Mayr W., Morling N., Olaisen B., ... & Wilson M. (2000). DNA commission of the international society for forensic genetics: guidelines for mitochondrial DNA typing. *Forensic Science International: Genetics*, 110(2), 79-85.

Casamiquela R. (1965). Rectificaciones y ratificaciones. Hacia una interpretación definitiva del panorama etnológico de la Patagonia y área septentrional adyacente, Cuadernos del Sur. Bahía Blanca: Universidad Nacional del Sur.

Castro de Guerra D., Figuera Pérez C., Bravi C.M., Saunier J. et al. (2012) Sequence variation of mitochondrial DNA control region in North Central Venezuela. *Forensic Science International: Genetics* 6(5): e131-3. <https://doi.org/10.1016/j.fsigen.2011.11.004>.

Castro J.C., Bonomo M., Venanzi L.G & Cornero S. (2020). Perros indígenas en el Noreste argentino. *Latin American Antiquity* 31(4):853–70.

Castro J.C. & Terranova E. (2015). Indicadores paleoindios en la provincia de Entre Ríos (Argentina). *Comechingonia. Revista de Arqueología* 19(1), pp. 183-190.

Castro J.C. (2017). Investigaciones arqueológicas en la cuenca media e inferior del río Uruguay (provincia de Entre Ríos). Tesis Doctoral. Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata, La Plata, 536 pp.

Castro J.C. (2019) Río Uruguay. Una síntesis arqueológica. *Revista del Museo de La Plata* 4(2), pp. 541-584. <https://doi.org/10.24215/25456377e088>

Catelli M.L., Alvarez-Iglesias V., Gómez-Carballa A., Mosquera-Miguel A., Romanini C., Borosky A., Amigo J., Carracedo A., Vullo C., Salas A. (2011). The impact of modern migrations on present-day multi-ethnic Argentina as recorded on the mitochondrial DNA genome. *BMC Genetics*, 12(1), 77. <https://doi.org/10.1186/1471-2156-12-77>

Cavalli-Sforza L. L., Menozzi, P. & Piazza A. (1994). The history and geography of human genes. Princeton university press.

Censabella M. (1999). Las lenguas indígenas de la Argentina: una mirada actual. Eudeba. ISBN: 950-23-0956-1.

Ceruti C. (2009). La cerámica de Santa Fe la Vieja: hacia una revaloración del componente indígena: la macroetnia Chaná-Timbú. En: Chiavazza H; Ceruti C.N. (Eds.). Arqueología de ciudades americanas del siglo XVI. Mendoza (ARG): Universidad Nacional de Cuyo, p. 369-400.

Ceruti C. N. y González M. I. (2007). Modos de vida vinculados con ambientes acuáticos del Nordeste y Pampa Bonaerense de Argentina. Relaciones de la Sociedad Argentina de Antropología XXXII, Buenos Aires. ISSN 0325-2221

Ceruti C. N. (1986). Algo Sobre Crítica y Autocrítica en Arqueología. *Revista de Antropología*. 1(1):19-ss. Buenos Aires.

Ceruti C.N. (1993). Arqueología. En: Renna, A. (ed.) Nueva Enciclopedia de la Provincia de Santa Fe IV, Santa Fe, Sudamérica, pp. 557-580.

Ceruti C. N. (2003). Entidades culturales presentes en la Cuenca del Paraná Medio (Margen entrerriana). Mundo de Antes. Instituto de Arqueología y Museo. 3 pp 110-135. Tucumán.

Cihlar J. C., Strobl C., Lagacé R., Muenzler M., Parson W. & Budowle B. (2020). Distinguishing mitochondrial DNA and NUMT sequences amplified with the precision ID mtDNA whole genome panel. *Mitochondrion*, 55, 122-133.

Chinnery P. F. (2006). Mitochondrial DNA in Homo sapiens. In Human Mitochondrial DNA and the Evolution of Homo sapiens (pp. 3-15). Berlin, Heidelberg: Springer Berlin Heidelberg.

Cocco G. et al. (2016). Sancti Spiritus, 1527-1529: aportes al estudio de los primeros asentamientos europeos en Sudamérica. In: Calvo L. M. & Cocco, G. (Eds.). Primeros asentamientos españoles y portugueses en la América central y meridional: siglos XVI y XVII. Santa Fe (ARG): Universidad Nacional del Litoral, p. 201-220.

Conlazo D. (1990). Los indios de Buenos Aires: (siglos XVI y XVII). Fondo De Cultura Económica.

Contreras S., Contreras F., Lutz A. & Zucol A. (2015). Contribución al conocimiento florístico del Chaco Oriental, sudeste de la provincia de Formosa, Argentina. *Boletín de la Sociedad Argentina de Botánica* 50(4), pp. 531-574.

Corrêa Â. A. (2023). Pindorama de Mboîa e îakaré: continuidade e mudança na trajetória das populações Tupi. *Revista de Arqueologia*, 36(2), pp. 194–221.

Dawe R. & Kornfeld M. (2017). Nunataks and valley glaciers: Over the mountains and through the ice. *Quaternary International*, 444:56-71

de Araujo L. F., Fonseca A. S., Muys B. R., Praça J. R., Bueno R. B., Lorenzi J. C., ... & Silva W. A. (2015). Mitochondrial genome instability in colorectal adenoma and adenocarcinoma. *Tumor Biology*, 36, 8869-8879.

de Saint Pierre M., Bravi C.M., Motti J.M.B., Fuku N., Tanaka M., Llop E., Bonatto S.L., Moraga M. (2012a) An alternative model for the early peopling of southern South America revealed by analyses of three mitochondrial DNA haplogroups *PLoS One* 7(9): e43486. <https://doi.org/10.1371/journal.pone.0043486>.

de Saint Pierre M., Gandini F., Perego U. A., Bodner M., et al. (2012b). Arrival of paleo-Indians to the southern cone of South America: New clues from mitogenomes. *PLoS One*, 7(12), e51311. <https://doi.org/10.1371/journal.pone.0051311>.

Delgado M., Rodríguez F., Kassadjikova L., & Fehren-Schmitz L. (2021). A paleogenetic perspective of the Sabana de Bogotá (northern South America) population history over the Holocene (9000–550 cal BP). *Quaternary International*, 578, 73–86. <https://doi.org/10.1016/j.quaint.2020.08.03>

Demarchi D.A., Panzetta-Dutarin G.M., Motran C.C., López De Basualdo M.A., Marcellino A.J. (2001) Mitochondrial DNA haplogroups in Amerindian populations from the Gran Chaco. *American Journal of Physical Anthropology*, 115:199–203. <https://doi.org/10.1002/ajpa.1074>

Demarchi D. A. (2014). Análisis de la estructura genética en poblaciones nativas del Gran Chaco. *Folia Histórica del Nordeste*, 22, 169-185. <http://dx.doi.org/10.30972/fhn.02255>

Dillehay T. D., Ramírez C., Pino M., Collins M. B. & Rossen J. (2008). Monte Verde: Seaweed, food, medicine, and the peopling of South America. *Science*, 320, 784–786.

Dipierri J., Alfaro E., Martínez Marignac V., Bailliet G., Bravi C., Cejas S., & Bianchi N. (1998). Paternal Directional Mating in Two Amerindian Subpopulations Located at Different Altitudes in Northwestern Argentina. *Human Biology*, 70(6), 1001-1010

Dos Reis R. S., Simão F., dos Santos Stange V., Garcia F. M., Dettogni R. S., Stur E., ... & Louro I. D. (2019). A view of the maternal inheritance of Espírito Santo populations: the contrast between the admixed and Pomeranian descent groups. *Forensic Science International: Genetics*, 40, 175-181.

Erlandson J.M., Braje T.D., et al. (2015). Ecology of the kelp highway: did marine resources facilitate human dispersal from Northeast Asia to the Americas? *The Journal of Island and Coastal Archaeology*, 10:392-411.

Escudero S., Feuillet Terzaghi M. R., & Núñez Camelino M. (2010). Modelos teóricos imperantes en la arqueología del Nordeste argentino: los casos de las Provincias de Corrientes y Santa Fe (Argentina). *Arqueología*, 17, 97-118. <https://doi.org/10.34096/arqueologia.t17.n0.1839>

Excoffier L., Laval L. G. & Schneider S. (2005). An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, 47–50. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>

Excoffier L., Smouse P., Quattro J. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA RFLP data. *Genetics* 131:479-491.

Fagundes N. J. R., Kanitz R., Eckert R., Valls A. C. S., Bogo M. R., Salzano F. M., et al. (2008). Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *American Journal of Human Genetics*, 82(3), 583–592. <https://doi.org/10.1016/j.ajhg.2007.11.013>.

Fernandes D.M., Sirak K.A., et al. (2021) A genetic history of the pre-contact Caribbean. *Nature* 590:103-110.

Figueiro G., Mut P., Ale L., Flores-Gutiérrez S., Greif G., Hidalgo P.C., Luna L., Ackermann E., Negro R. G. et al. (2022). Filogeografía de mitogenomas indígenas de Uruguay. *Revista Argentina de Antropología Biológica* 24(1): 042. <https://doi.org/10.24215/18536387e042>.

Fridman C., Gonzalez R. S., Pereira A. C. & Cardena M. M. S. G. (2014). Haplotype diversity in mitochondrial DNA hypervariable region in a population of southeastern Brazil. *International Journal of Legal Medicine*, 128, 589-593.

Galagedara G.R.S. (2018). Mitochondrial D-Loop characterization of the Amazonian Gorotire Kayapó population. Binghamton University, Master's Thesis.

García A., Nores R. et al. (2021) Ancient and modern mitogenomes from Central Argentina: new insights into population continuity, temporal depth and migration in South America. *Human Molecular Genetics* 30:1200-1217.

García A., Oliveira Rufino R., Bergese A.B., Agüero J.F., Cuevas A., et al. (2016). El cruce entre las antropologías: una mirada interdisciplinaria en torno a la genética de poblaciones, las memorias familiares y la construcción identitaria; Universidad Nacional de Córdoba. Facultad de Filosofía y Humanidades. Museo de Antropología; *Revista del Museo de Antropología*; 9; 2; 12; 105-112.

García A., Pauro M., Bailliet G., Bravi C.M., Demarchi D.A. (2018). Genetic variation in populations from central Argentina based on mitochondrial and Y chromosome DNA evidence. *Journal of human genetics*, 63(4), 493–507. <https://doi.org/10.1038/s10038-017-0406-7>.

García A., Pauro M., et al. (2012). Phylogeography of mitochondrial haplogroup D1: an early spread of subhaplogroup D1j from Central Argentina. *American Journal of Physical Anthropology* 149: 583-90.

García-Olivares V., Rubio-Rodríguez L. A., Muñoz-Barrera A., Díaz-de Usera A., Jáspez D., Lñigo-Campos A., ... & Flores C. (2023). Digging into the admixture strata of current-day Canary Islanders based on mitogenomes. *iScience*, 26(1). <https://doi.org/10.1016/j.isci.2022.105907>.

Garita-Onandía, Pauro M., Altuna M.E., Demarchi D., García A. (2021). Patrones de migración y niveles de diversidad genética de linajes maternos en la población de La Esperanza, Jujuy. *Revista Argentina de Antropología Biológica*, 23(2):1-24, <https://doi.org/10.24215/18536387e033>

Gayà-Vidal M., Moral P., Saenz-Ruales N., Gerbault P., Tonasso L., Villena M., et al. (2011) mtDNA and Y-chromosome diversity in Aymaras and Quechuas from Bolivia: different stories and special genetic traits of the Andean Altiplano populations. *American Journal of Physical Anthropology* 145:215-30.

Gómez-Carballa A., Ignacio-Veiga A., Álvarez-Iglesias V., Pastoriza-Mourelle A., Ruiz Y., Pineda L., ... & Salas A. (2012). A melting pot of multicontinental mtDNA lineages in admixed Venezuelans. *American journal of physical anthropology*, 147(1), 78-87.

Gómez-Carballa A., Moreno F., Álvarez-Iglesias V., Martín-Torres F., García-Magariños M., Pantoja-Astudillo J. A., ... & Salas A. (2016). Revealing latitudinal patterns of mitochondrial DNA diversity in Chileans. *Forensic Science International: Genetics*, 20, 81-88.

Gómez-Carballa A., Pardo-Seco J., Brandini S., Achilli A., Perego U.A., Coble M.D. et al. (2018). The peopling of South America and the trans-Andean gene flow of the first settlers. *Genome Research* 28(6):767-779.

Greenberg J. H., Turner C. G., & Zegura S. L. (1986). The settlement of the Americas: A 103 comparison of the linguistic, dental, and genetic evidence. *Current Anthropology*, 27(5), 477–497. <https://doi.org/10.1086/203472>.

Grosso J.L. (2008). Indios muertos, negros invisibles. Hegemonía, Identidad y Añoranza. Córdoba, Encuentro Grupo Editor. 252 p.

Grugni V., Raveane A., Ongaro L., Battaglia V., Trombetta B., Colombo G., et al. (2019). Analysis of the human Ychromosome haplogroup Q characterizes ancient population movements in Eurasia and the Americas. *BMC biology*, 17(1):3. <https://doi.org/10.1186/s12915-018-0622-4>

Guevara E. K., Palo J. U., King J. L., Buś M. M., Guillen S., Budowle B., & Sajantila A. (2021). Autosomal STR and SNP characterization of populations from the Northeastern Peruvian Andes with the ForenSeq™ DNA Signature Prep Kit. *Forensic Science International: Genetics*, 52, 102487.

Hammer M. F., Karafet T., Rasanayagam A., Wood E. T., Altheide T. K., Jenkins T., ... & Zegura S. L. (1998). Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. *Molecular biology and evolution*, 15(4), 427-441.

Hartmann A., Thieme M., Nanduri L. K., Stempfl T., Moehle C., Kivisild T., & Oefner P. J. (2009). Validation of microarray-based resequencing of 93 worldwide mitochondrial genomes. *Human mutation*, 30(1), 115-122.

Hernández, I. (1995). Los indios de Argentina. Colección pueblos y lenguas indígenas 4. Madrid, Ed MAPFRE, 335p.

Homburger J. R., Moreno-Estrada A., Gignoux C. R., Nelson D., Sanchez E., Ortiz-Tello P., et al. (2015). Genomic insights into the ancestry and demographic history of South America. *PLoS Genetics*, 11(12), e1005602. <https://doi.org/10.1371/journal.pgen.1005602>

INDEC. (2010). Censo Nacional de Población, Hogares y Viviendas.

ISOGG (2017) ISOGG Y-DNA haplogroup Q and its Subclades-2017.In: ISOGG. http://www.isogg.org/tree/ISOGG_HapgrpQ.html. Accessed 20 Apr 2017

Jobling M. A. & Tyler-Smith C. (2003). The human Y chromosome: An evolutionary marker comes of age. *Nature Reviews Genetics*, 4(8), 598–612. <https://doi.org/10.1038/nrg1124>.

Jobling M. A. & Tyler-Smith C. (2017). Human Y-chromosome variation in the genome-sequencing era. *Nature Reviews Genetics* 18, 485–497. <https://doi.org/10.1038/nrg.2017.36>

Jobling M., Hollox E., Kivisild T., Hurles M., & Tyler-Smith C. (2014). *Human evolutionary genetics* (Garland Sc). New York.

Jobling M., Hurles M. & Tyler-Smith C. (2004). Human Evolutionary Genetics: Origins, Peoples & Disease. Garland Science, ISBN: 9780203487211, 650p.

Jurado Medina L.S., Ramallo V., Calandra H., Lamenza G., Braunstein J., Salceda S., Bailliet G. (2014). Linajes paternos del Gran Chaco, un abordaje desde el ADN. *Folia Histórica del Nordeste* (22): 187-202. <http://dx.doi.org/10.30972/fhn.02256>.

Jurado Medina, L. S. (2015). Tipificación de marcadores uniparentales en poblaciones mestizas de Argentina, Tesis Doctoral Inédita, La Plata, Universidad Nacional de La Plata.

Just R. S., Irwin J. A., & Parson W. (2015). Mitochondrial DNA heteroplasmy in the emerging field of massively parallel sequencing. *Forensic Science International: Genetics*, 18, 131-139.

Karafet T. M., Mendez F. L., Meilerman M. B., Underhill P. A., Zegura S. L., & Hammer M. F. (2008). New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Research*, 18(5), 830–838. <https://doi.org/10.1101/gr.7172008>.

Keyeux G., Rodas C., Gelvez N. & Carter D. (2002). Possible migration routes into South America deduced from mitochondrial DNA studies in Colombian Amerindian populations. *Human Biology*, 211-233.

Kimura M. (1983). The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge.

Kingman J.F.C. (1982a). The coalescent. *Stochastic Processes and their Applications* 13:235-248.

Kingman J.F.C. (1982b). On the genealogy of large populations. *Journal of Applied Probability* 19:27-43.

Kitchen A., Miyamoto M. M. & Mulligan C. J. (2008). A three-stage colonization model for the peopling of the Americas. *PloS One*, 3(2), e1596. <https://doi.org/10.1371/journal.pone.0001596>.

Kivisild T., Tolk H. V., Parik J., Wang Y., Papiha S. S., Bandelt H. J., & Villems R. (2002). The emerging limbs and twigs of the East Asian mtDNA tree. *Molecular biology and evolution*, 19(10), 1737-1751.

Kruskal, J.B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29, 1–27. <https://doi.org/10.1007/BF02289565>.

Kruskal J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29(2), 115–129.

Kumar S., Bellis C., Zlojutro M., Melton P., Blangero J., Curran J.E. (2011) Large scale mitochondrial sequencing in Mexican Americans suggests a reappraisal of Native American origins. *BMC Evolution Biology* 11:293.

Lafon C.R. (1971). Introduccion a la Arqueologia del Nordeste argentino. *Relaciones de la Sociedad Argentina de Antropología* 5(2), pp. 119-152.

Lahn B. T. & Page D. C. (1999). Four evolutionary strata on the human X chromosome. *Science (New York, N.Y.)*, 286(5441), 964–967. <https://doi.org/10.1126/science.286.5441.964>.

Lamenza G.N., Calandra H.A., Salceda S.A. (2019). Arqueología de los ríos Pilcomayo, Bermejo y Paraguay En: Dossier: Arqueología de los ríos de Tierras Bajas de América del Sur. *Revista del Museo de La Plata*, Vol 4 (2), 481-510.

Lanteri A.A. & Confalonieri V.A. (2003). Filogeografía: Objetivos, metodos y ejemplos. Una Perspectiva Latinoamericana de La Biogeografía, 185–193.

Larmuseau M. H., Van Geystelen A., Kayser M., van Oven M. & Decorte R. Towards a consensus Y-chromosomal phylogeny and Y-SNP set in forensics in the next-generation sequencing era. *Forensic Science International: Genetics*, vol. 15, pp. 39-42, 2015.

Lee E. J. & Merriwether D. A. (2015). Identification of whole mitochondrial genomes from Venezuela and implications on regional phylogenies in South America. *Human Biology*, 87(1), 29-38.

Lell J. T., Sukernik R. I., Starikovskaya Y. B., Su B., Jin L., Schurr, T. G., ... & Wallace D. C. (2002). The dual origin and Siberian affinities of Native American Y chromosomes. *The American Journal of Human Genetics*, 70(1), 192-206.

Leoni M. S. (2016) La región NEA en la perspectiva de Ernesto Maeder. XXXVI Encuentro de Geohistoria regional. ISBN: 978-987-4450-00-5

Lindo J., Achilli A., Perego U. A., Archer D., et al. (2017). Ancient individuals from the North American Northwest Coast reveal 10,000 years of regional genetic continuity. *Proceedings of the National Academy of Sciences*, 114(16), 4093-4098.

Lippold S., Xu H., Ko A., Li M., Renaud G., Butthof A., ... & Stoneking M. (2014). Human paternal and maternal demographic histories: insights from high-resolution Y chromosome and mtDNA sequences. *Investigative genetics*, 5, 1-17.

Loponte D. & Acosta A. (2007). Horticultores amazónicos en el humedal del Paraná Inferior: Los primeros datos isotópicos de la dieta. En C. Bayón, A. Pupio, M. I González, N. Flegenheimer y M. Frére (Eds.), *Arqueología de las Pampas* (Tomo I, pp. 75-93). Sociedad Argentina de Antropología.

Loponte D. & Acosta A. (2008). Estado actual y perspectivas de la arqueología de la “Tradición Tupiguarani” en Argentina. En A. Prous y T. Andrade Lima (Eds.), *Os ceramistas Tupiguarani* (Vol. 2, pp. 197–215). Sigma, Belo Horizonte.

Luiselli D., Simoni L., Tarazona-Santos E., Pastor S., Pettener D. (2000). Genetic structure of Quechua-speakers of Central Andes and geographic patterns of gene frequencies in South Amerindian populations. *American Journal of Physical Anthropology* 113:5–17.

Luisi P., García A., Berros J.M., Motti J.M.B., Demarchi D.A., Alfaro E., et al. (2020) Fine-scale genomic analyses of admixed individuals reveal unrecognized genetic ancestry components in Argentina. *PLoS One* 15(7): e0233808. <https://doi.org/10.1371/journal.pone.0233808>

Lyons E.A., Scheible M.K., Sturk-Andreaggi K., Irwin J.A., Just R.S. (2013) A high-throughput Sanger strategy for human mitochondrial genome sequencing. *BMC Genomics* 14:881

Macaulay V. (2005). Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science*, 308(5724), 1034–1036. <https://doi.org/10.1126/science.1109792>.

Maeder E. & Poenitz A.J.E. (2006). Corrientes Jesuítica. Historia de las misiones de Yapeyú, La Cruz, Santo Tomé y San Carlos en la etapa jesuítica y en el período posterior, hasta su disolución. Ediciones Al Margen, La Plata. 129p.

Maeder E. (2004). Misiones: Historia de la tierra prometida. Buenos Aires, Eudeba, 204p.

Maeder E. (2015). Los grupos subalternos en el nordeste del Virreinato del Río de la Plata. Rosario, Prohistoria Ediciones.

Malhi R.S., Cybulski J.S., Tito R.Y., Johnson J., et al. (2010) Mitochondrial haplotype C4c confirmed as a founding genome in the Americas. *American Journal of Physical Anthropology* 141: 494-497.

Manni F., Guérard E., Heyer E. (2004). Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by —Monmonier’s algorithm. *Human Biology*; 76(2):173-190.

Mann P. C., Cooper M. E., Ryckman K. K., Comas B., Gili J., Crumley S., ... & Murray J. C. (2013). Polymorphisms in the fetal progesterone receptor and a calcium-activated potassium channel isoform are associated with preterm birth in an Argentinian population. *Journal of perinatology*, 33(5), 336-340.

Marrero A.R., Silva-Junior W.A., Bravi C.M., et al. (2007). Demographic and evolutionary trajectories of the Guarani and Kaingang natives of Brazil. *American Journal of Physical Anthropology* 132(2):301-10

Martínez Sarasola C. (2013). Nuestros paisanos los indios. Vida, historia y destino de las comunidades indígenas de Argentina. Buenos Aires, Editorial Del Nuevo Extremo, 816p.

Martínez-Marignac V.L., Bravi C.M., Lahitte H.B., Bianchi N.O. (1999). Estudio del ADN mitocondrial de una muestra de la ciudad de La Plata. *Revista Argentina de Antropología Biológica*; 2(1):281–300.

Mazières S., Guitard E., Crubézy E., Dugoujon J. M., Bortolini M. C., Bonatto S. L., ... & Salzano F. M. (2008). Uniparental (mtDNA, Y-chromosome) Polymorphisms in French Guiana and Two Related Populations—Implications for the Region's Colonization. *Annals of Human Genetics*, 72(1), 145-156.

Mishmar D., Ruiz-Pesini E., Golik P., Macaulay V., Clark A.G., Hosseini S., Brandon M., Easley K., Chen E., Brown M.D., Sukernik R.I., Olckers A., Wallace D.C. (2003). Natural selection shaped regional mtDNA variation in humans. *Proceedings of the National Academy of Sciences of the United States of America*; 100(1):171-6.

Mizuno F., Gojobori J., Wang L., Onishi K., Sugiyama S., Granado, J., ... & Ueda S. (2014). Complete mitogenome analysis of indigenous populations in Mexico: its relevance for the origin of Mesoamericans. *Journal of Human Genetics*, 59(7), 359-367.

Monson K. L. (2002). The mtDNA Population Database: An Integrated Software and Database Resource for Forensic Comparison by Monson, Miller, Wilson, DiZinno, and Budowle (Forensic Science Communications, April 2002). *Research and Technology*, 4(2).

Morales J., Dipierri J.E., Alfaro E., Bejarano I.F. (2000). Distribution of the ABO system in the Argentine northwest: miscegenation and genetic diversity. *Interciencia*, 25, 378-386.

Motti J.M.B. (2012). Caracterización de linajes maternos en la población actual del noroeste y centro-oeste argentinos. Facultad de Ciencias Naturales y Museo, UNLP. Tesis doctoral.

Motti J.M.B., Beltramo J., Schwab M.E., Jurado-Medina L., et al. (2017) Diferenciación regional en Sudamérica: aportes de la evidencia mitocondrial. *Intersecciones en Antropología*, vol. 18, núm. 3 pp. 271-282 Universidad Nacional del Centro de la Provincia de Buenos Aires, Buenos Aires, Argentina.

Motti J.M.B., Hagelberg E., Lindo J., Malhi R., Bravi C.M., Guichón R.A. (2015). Primer genoma mitocondrial en restos humanos de la costa de Santa Cruz, Argentina. *Magallania* 43:119-131

Motti J.M.B., Muzzio M., Ramallo V., et al. (2013) Origen y distribución espacial de linajes maternos nativos en el noroeste y centro oeste argentinos. *Revista Argentina de Antropología Biológica* 13:3-14

Motti J.M.B., Rodenak B., Bailliet G., Bravi C.M. (2009a) Tipificación de haplogrupos mitocondriales mediante multiplex PCR-AFLP. *Novenas Jornadas Nacionales de Antropología Biológica*, Puerto Madryn, Argentina.

Muzzio M., Motti J. M. B., Chiarullo S. M., Bravi C. M., & Bailliet G. (2012). Genealogical surveys show a high rate of non-paternal surname transmission with regional differences in Argentina. *HOMO*, 63(1), 43-49.

Myres N. M., Rootsi S., Lin A. A., Jarve M., King R. J., Kutuev I., ... Underhill P. A. (2011). A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *European Journal of Human Genetics*, 19, 95–101. <https://doi.org/10.1038/ejhg.2010.146>.

Nakatsuka N., Lazaridis I., Barbieri C., Skoglund P., Rohland N., Mallick S., et al. (2020) A Paleogenomic Reconstruction of the Deep Population History of the Andes. *Cell*. May 28;181(5):1131-1145.e21.

Nei M. (1972). Genetic distance between populations. *The American Naturalist*, 106(949), 283-292. doi:10.1086/282771

Nieves-Colón M.A., Pestle W.J., Reynolds A.W., et al. (2020) Ancient DNA Reconstructs the Genetic Legacies of Precontact Puerto Rico Communities. *Molecular Biology and Evolution* 37(3): 611-626.

Noelli F.S. (2004) La distribución geográfica de las evidencias arqueológicas guaraní. *Revista Indias* 64(230), 17-34.

Nogueira T.L. (2019) Ancestralidade genômica matrilinear de indivíduos das regiões Centro-Oeste, Norte e Nordeste do Brasil. Universidade do Estado do Rio de Janeiro, PhD Thesis. <http://www.bdt.d.uerj.br/handle/1/16124>

Novelletto A. (2007). Y chromosome variation in Europe: Continental and local processes in the formation of the extant gene pool. *Annals of Human Biology*, 34(2), 139–72. <https://doi.org/10.1080/03014460701206843>

Oakley L., Prado D. & Adámoli J. (2005) Aspectos biogeográficos del corredor fluvial Paraguay-Paraná. *Miscelánea INSUGEO* 14, pp. 245-258.

Ongaro, L., Scliar, M. O., Flores, R., Raveane, A., Marnetto, D., Sarno, S., ... Montinaro, F. (2019). The genomic impact of European colonization of the Americas. *Current Biology*, 29(23), 3974–3986.e4. <https://doi.org/10.1016/j.cub.2019.09.076>

Pagano S., Sans M., Pimenoff V., Cantera A. M., Alvarez J. C., Lorente J. A., ... & Sajantila A. (2005). Assessment of HV1 and HV2 mtDNA variation for forensic purposes in an Uruguayan population sample. *Journal of forensic sciences*, 50(5), JFS2004362-4.

Pakendorf B., & Stoneking M. (2005). Mitochondrial DNA and human evolution. *Annu. Rev. Genomics Hum. Genet.*, 6, 165-183. <https://doi.org/10.1146/annurev.genom.6.080604.162249>.

Palencia L., Valverde L., Álvarez A., Cainé L. M., Cardoso S., Alfonso-Sánchez M. A., ... & de Pancorbo M. M. (2010). Mitochondrial DNA diversity in a population from Santa Catarina (Brazil): predominance of the European input. *International journal of legal medicine*, 124, 331-336.

Parolín M. L., Avena, S. A., Fleischer, S., Pretell, M., Di Fabio Rocca, F., Rodríguez, D. A., ... & Carnese F. R. (2013). Análisis de la diversidad biológica y mestizaje en la ciudad de Puerto Madryn (Prov. de Chubut, Argentina). *Revista Argentina de Antropología Biológica*, 15(1), 61-75. Recuperado a partir de <https://revistas.unlp.edu.ar/raab/article/view/693>.

Parson W., Gusmão L., Hares D. R., Irwin J. A., Mayr W. R., Morling N., ... Parsons T. J. (2014). DNA Commission of the International Society for Forensic Genetics: Revised and extended guidelines for mitochondrial DNA typing. *Forensic Science International: Genetics*, 13, 134–142. <https://doi.org/10.1016/j.fsigen.2014.07.010>

Pauro M. (2015) Análisis molecular de linajes uniparentales en poblaciones humanas del centro de Argentina. Universidad Nacional de Córdoba, Córdoba. Tesis doctoral.

Paz Sepúlveda P., Jurado Medina L., Ramallo V., Muzzio M., Motti J., Santos M.R., Dipierri J., Gómez, E., Demarchi D. Bravi C., Salceda S., Bailliet, G. (2020). Linajes paternos autóctonos de Gran Chaco analizados con microsatélites. *Revista del Museo de La Plata*. 5. 553-562.

Pena S.D.J., Santos F.R., Bianchi N.O., Bravi C.M., Carnese F.R., Rothhammer F., Gerelsaikhan T. Munkhtuja B., Oyunsuren T. (1995). A major founder Y-chromosome haplotype in Amerindians. *Nature Genetics*, 11:15.

Perego U. A., Achilli A., Angerhofer N., Accetturo M., Pala M., Olivieri A., ... Torroni A. (2009). Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. *Current Biology*, 19(1), 1–8. <https://doi.org/10.1016/j.cub.2008.11.058>

Perego U. A., Angerhofer N., Pala M., Olivieri A., Lancioni H., Kashani B. H., ... Torroni A. (2010). The initial peopling of the Americas: A growing number of founding mitochondrial genomes from Beringia. *Genome Research*, 20(9), 1174–1179. <https://doi.org/10.1101/gr.109231.110>.

Perego U. A., Lancioni H., Tribaldos M., Angerhofer N., Ekin J. E., Olivieri, A., ... & Achilli A. (2012). Decrypting the mitochondrial gene pool of modern Panamanians. *PLoS One*, 7(6), e38337. <https://doi.org/10.1371/journal.pone.0038337>.

Pinotti T., Bergström A., Geppert M., Bawn M., Ohasi D., Shi W., Lacerda D.R., et al. (2019). Y Chromosome Sequences Reveal a Short Beringian Standstill, Rapid Expansion, and early Population structure of Native American Founders. *Current Biology*. 2019 Jan 7;29(1):149-157.e3. <https://doi.org/10.1016/j.cub.2018.11.029>.

Poletto M.M., Malaghini M., Silva J.S., Bicalho M.G., Braun-Prado K. (2019). Mitochondrial DNA control region diversity in a population from Parana state-increasing the Brazilian forensic database. *International Journal of Legal Medicine*, 133(2):347-351.

Politis G.G & Bonomo M. (2023) Goya-Malabrigo y la expansión arawak 10 años después. *Revista de Arqueología* Vol 36 (1), 87-102.

Politis G.G. & Bonomo M. (2018b). Estado actual y perspectivas de Goya-Malabrigo, una sociedad indígena del Noreste Argentino. En: Goya-Malabrigo. Arqueología de una sociedad indígena del Noreste Argentino. Edited by G. Politis and M. Bonomo, 9–44. Tandil: Editorial UNICEN.

Politis G.G., Domic A.I, Bonomo M. & Capriles J. (2018a). Modelamiento de los parámetros ambientales de la distribución de Goya-Malabrigo mediante Maxent. En: Goya-Malabrigo: Arqueología de una sociedad indígena del noreste argentino. Editado por Politis G. & Bonomo M., 289–305. Tandil: Editorial UNICEN.

Politis, G.G. & Bonomo M. (2012). La entidad arqueológica Goya-Malabrigo (Ríos Paraná y Uruguay) y su filiación Arawak. *Sociedade de Arqueologia Brasileira, Revista de Arqueologia* 25(1):10–46.

Potter B.A., Reuther J.D., Holliday V.T. et al. (2017). Early colonization of Beringia and Northern North America: Chronology, routes, and adaptive strategies. *Quaternary Int* 444:36-55.

Poznik G. D., Xue Y., Mendez F. L., Willems T. F., Massaia A., Sayres M. A. W., et al. (2016). Punctuated bursts in human male demography inferred from 1,244 worldwide Y chromosome sequences. *Nature genetics*, vol. 48, p. 593.

Prieto L., Zimmermann B., Goios A., Rodriguez-Monge A., Paneto G. G., Alves C., ... & Parson W. (2011). The GHEP–EMPOP collaboration on mtDNA population data—A new resource for forensic casework. *Forensic Science International: Genetics*, 5(2), 146-151.

Pucciarelli H.M., Neves W.A., González-José R., Sardi M.L., Rozzi F.R., Struck A., Bonilla M.Y. (2006). East-West cranial differentiation in pre-Columbian human populations of South America. *Homo*. 57(2):133-50.

Race, Ethnicity and Genetics Working Group. (2005). The use of racial, ethnic, and ancestral categories in human genetics research. *American Journal of Human Genetics*, 77, 519-532. <https://doi.org/10.1086/491747>.

Ramallo V., Bisso-Machado R., Bravi C., Coble M. D., Salzano F. M., Hünemeier T. & Bortolini M. C. (2013). Demographic expansions in South America: enlightening a complex scenario with genetic and linguistic data. *American Journal of Physical Anthropology* 150 (3): 453–463.

Ramallo V., Mucci J. M., García A., Muzzio M., Motti J. M. B., Santos M. R., ... Bailliet G. (2009). Comparison of Y-chromosome haplogroup frequencies in eight provinces of Argentina. *Forensic Science International: Genetics Supplement Series*, 2(1), 431–432. <https://doi.org/10.1016/j.fsigss.2009.08.047>.

Ramallo V., Santos M. R., Muzzio M., Motti J. M. B., Salceda S. A. & Bailliet, G. (2009). Linajes masculinos y su diversidad en comunidades Wichí de Formosa. Universidad Nacional de Córdoba. Facultad de Filosofía y Humanidades; *Revista del Museo de Antropología*; 2; 1; 21-12-2009; 67-74.

Ramírez Braschi, D. (2004) La guerra de la Triple Alianza a través de los periódicos correntinos. Corrientes: Moglia ediciones.

Ray N., Wegmann D., Fagundes N. J., Wang S., Ruiz-Linares A., & Excoffier L. (2010). A statistical evaluation of models for the initial settlement of the American continent emphasizes the importance of gene flow with Asia. *Molecular biology and evolution*, 27(2), 337-345.

Raymond M. & Rousset F. (1995). An exact test for population differentiation. *Evolution*; 49:1280-1283. <https://doi.org/10.2307/2410454>.

- Rebolledo-Jaramillo** B., Obregon M. G., Huckstadt V., Gomez A., & Repetto G. M. (2021). Contribution of Mitochondrial DNA Heteroplasmy to the Congenital Cardiac and Palatal Phenotypic Variability in Maternally Transmitted 22q11. 2 Deletion Syndrome. *Genes*, 12(1), 92.
- Reich** D., Patterson N., Campbell D., Tandon A., Mazieres S., Ray N., ... Ruiz-Linares A. (2012). Reconstructing Native American population history. *Nature*, 488(7411), 370–374. <https://doi.org/10.1038/nature11258>.
- Rieux** A., Eriksson A., Li M., Sobkowiak B., Weinert L.A., Warmuth V., et al. (2014). Improved calibration of the human mitochondrial clock using ancient genomes. *Molecular Biology and Evolution* 31:2780-2792
- Rodríguez Golpe** D.C., Aquilano E., Beltramo J., Schwab M.E. et al. (2019). Linajes B2e en el NEA y Litoral de Argentina. XIV Jornadas Nacionales de Antropología Biológica, San Salvador de Jujuy, Argentina.
- Rodríguez** J. A. (2006). El poblamiento y el proceso de la secuencia cultural prehistórica de la cuenca del Plata. *Folia Histórica del Nordeste*, Nº 16 (Resistencia, 2006) IIGHI, CONICET - IH, UNNE.
- Rodríguez** J. A. (2001). Poblamiento prehistórico de la Mesopotamia Argentina. *Folia Histórica del Nordeste* 15. IIGHI-Instituto de Historia-UNNE, 129-146.
- Roewer** L., Nothnagel M., Gusmão L., Gomes V., González M., Corach D., ... Krawczak M. (2013). Continent-wide decoupling of Y-chromosomal genetic variation from language and geography in native South Americans. *PLoS Genetics*, 9(4), 1–16. <https://doi.org/10.1371/journal.pgen.1003460>.
- Rothhammer** F. & Dillehay T. D. (2009). The late Pleistocene colonization of South America: an interdisciplinary perspective. *Annals of human genetics*, 73(5), 540-549.
- Rothhammer** F., Llop E., Carvallo P. & Moraga M. (2001). Origin and evolutionary relationships of native Andean populations. *High Altitude Medicine & Biology*, 2(2), 227-233.
- Sala** A., Argüelles C.F., Marino M.E., Bobillo C., Fenocchio A., Corach D. (2010) Genetic analysis of six communities of Mbyá-Guaraní inhabiting northeastern Argentina by means of nuclear and mitochondrial polymorphic markers. *Human Biology* 82(4):433-56. doi: 10.3378/027.082.0406.
- Sala** A., Caputo M. & Corach D. (2019). Genetic structure of Mataco-Guaycurú speakers from Argentina and the extent of their genetic admixture with neighbouring urban populations. *Scientific Reports*, 9(1), 1-13. <https://doi.org/10.1038/s41598-019-54146-6>.
- Sanches** N. M., Paneto G. G., Figueiredo R. F., de Mello A. O. & Cicarelli R. M. (2014). Mitochondrial DNA control region diversity in a population from Espírito Santo state, Brazil. *Molecular biology reports*, 41, 6645-6648.

- Sandoval** J. R., Revollo S., Cuella, C., Lacerda D. R., Jota M. S., Fujita R., & Santos F. R. (2023). Genetic portrait of the Amazonian communities of Peru and Bolivia: The legacy of the Takanan-speaking people. *Annals of Human Genetics*, 87(5), 210-221.
- Sans** M. (2009). "Raza", adscripción étnica y genética en Uruguay. RUNA, Archivo Para Las Ciencias Del Hombre, 30(2), 163-174. <https://doi.org/10.34096/runa.v30i2.771>.
- Sans** M., Mones P., Figueiro G., Barreto I., Motti J.M.B., Coble M.D., et al. (2015). The mitochondrial DNA history of a former native American village in northern Uruguay. *American Journal of Human Biology* 27:407–416
- Scheib** C.L., Li H., Desai T., Link V., Kendall C., Dewar G., Griffith P.W., Mörseburg A. et al. (2018) Ancient human parallel lineages within North America contributed to a coastal expansion. *Science* 360(6392):1024-1027.
- Söchtig** J., Álvarez-Iglesias V., Mosquera-Miguel A., Gelabert-Besada M., et al. (2015) Genomic insights on the ethno-history of the Maya and the 'Ladinos' from Guatemala. *BMC Genomics* 16 (1), 131. <https://doi.org/10.1186/s12864-015-1339-1>.
- Schwab** M. (2018). Afinidades filogeográficas y estructuración geográfica de los linajes maternos y paternos presentes en poblaciones humanas del Noroeste argentino. Facultad de Ciencias Naturales y Museo, UNLP. Tesis doctoral.
- Segato** R.L. (2007). La nación y sus otros: raza, etnicidad y diversidad religiosa en tiempos de políticas de la identidad. Prometeo libros, Buenos Aires, 350 p.
- Serrano** A. (1972). Líneas fundamentales de la arqueología del litoral (una tentativa de periodización). Córdoba, Instituto de Antropología XXXII.
- Serrano** A. (2012). Los aborígenes argentinos. Síntesis etnográfica. Buenos Aires, CS Ediciones, 282p.
- Sevini** F., Yao D.Y., Lomartire L., Barbieri A., Vianello D., Ferri G., et al. (2013). Analysis of population substructure in two sympatric populations of Gran Chaco, Argentina. *PLoS One* 8(5): e64054.
- Sevini F.**, Vianello D., Barbieri C., Iaquilano N., De Fanti S., Luiselli D., Franceschi C. & Franceschi, Z. (2014). Human mitochondrial genomes reveal population structure and different phylogenies in Gran Chaco (Argentina). Unpublished. *GenBank direct submission*.
- Shoubridge** E.A. (2001). Cytochrome c oxidase deficiency. *American journal of medical genetics*, 106(1), 46-52.
- Simão** F., Ribeiro J., Vullo C., Catelli L., Gomes V., Xavier C., ... & Gusmão L. (2021). The ancestry of eastern Paraguay: a typical south American profile with a unique pattern of admixture. *Genes*, 12(11), 1788.

Simão F., Strobl C., Vullo C., Catelli L., Machado P., Huber N., Schnaller L., Huber G., Xavier C. et al. (2019) The maternal inheritance of Alto Paraná revealed by full mitogenome sequences. *Forensic Science International: Genetics* 39:66-72.

Soares P., Ermini L., Thomson N., Mormina M., Rito T., Röhl A., et al. (2009) Correcting for purifying selection: an improved human mitochondrial molecular clock. *American Journal of Human Biology* 84: 740-759.

Söchtig J., Álvarez-Iglesias V., Mosquera-Miguel A., Gelabert-Besada M., et al. (2015) Genomic insights on the ethno-history of the Maya and the 'Ladinos' from Guatemala. *BMC Genomics* 16 (1), 131. <https://doi.org/10.1186/s12864-015-1339-1>.

Sokolova V. A., Vasilyev V. B., Delefosse T., Hänni C., Rochet D. & Godinot C. (2002). A Russian family of Slavic origin carrying mitochondrial DNA with a 9-bp deletion in region V and a long C-stretch in D-loop. *Mitochondrion*, 1(6), 479-483.

Strobl C., Cihlar J. C., Lagacé R., Wootton S., Roth C., Hube N., ... & Parson W. (2019). Evaluation of mitogenome sequence concordance, heteroplasmy detection, and haplogrouping in a worldwide lineage study using the Precision ID mtDNA Whole Genome Panel. *Forensic Science International: Genetics*, 42, 244-251.

Studer E. F.S. (1958). La Trata de Negros en el Río de la Plata Durante el Siglo XVIII. Publicaciones del Instituto de Historia Argentina Doctor Emilio Ravignani. Buenos Aires, Universidad de Buenos Aires. Number 101. Pp. 378.

Taboada-Echalar P., Álvarez-Iglesias V., Heinz T., Vidal-Bralo et al. (2013). The genetic legacy of the pre-colonial period in contemporary Bolivians. *PLoS One*, 8(3), e58980. <https://doi.org/10.1371/journal.pone.0058980>.

Tamm E., Kivisild T., Reidla M., Metspalu M., Smith D.G., Mulligan C.J., Bravi C.M. et al. (2007) Beringian standstill and spread of Native American founders. *PLoS One* 2(9): e829. <https://doi.org/10.1371/journal.pone.0000829>

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):512-26.

Tarazona-Santos E., Carvalho-Silva D. R., Pettener D., Luiselli D., De Stefano G. F., Labarga C. M., Rickards O., Tyler Smith C., Pena S. D. & Santos F. R. (2001) Genetic differentiation in South America is related to environmental and cultural diversity: evidence from the Y chromosome. *American Journal of Human Genetics* 68,1485–1496.

Tarskaia L. A., & Melton P. (2006). Comparative analysis of mitochondrial DNA of Yakuts and other Asian populations. *Russian Journal of Genetics*, 42, 1439-1446.

Tavella M.P., Demarchi D.A. & Nores R. (2020) Diversidad genética en restos humanos de contextos arqueológicos tardíos de la provincia de Santiago del Estero a partir del estudio de ADN mitocondrial. *Revista Argentina de Antropología Biológica*, 22, 013.

Templeton A.R. (2006). Populations Genetics and Microevolutionary Theory. Department of Biology. Washington University St. Louis, Missouri. Wiley. 705 pp.

Tercer Censo Nacional. (1914). República Argentina. Población Tomo II.

Tito R.Y., Polo S.I. & Lewis, C.M. Jr. (2013). Mitochondrial genomes and founder effects within the Americas. *Unpublished*.

Torino R. (2021). El rol de la cerámica en la expansión guaraní: ¿Cómo se relacionan la cultura material, los nuevos espacios y las personas? *Revista del Museo de Antropología*, 14(2), 81-92.

Torino R. (2024). Modos de hacer la cerámica guaraní en el límite meridional de su expansión poblacional (Delta del Paraná y Río de la Plata). Universidad Nacional de La Plata, Tesis doctoral inédita. <https://doi.org/10.35537/10915/167575>

Torroni A., Schurr T.G., Cabell M.F., Brown M.D., Neel J.V., Larsen M., Smith D.G., Vullo C.M., Wallace D.C. (1993). Asian affinities and continental radiation of the four founding Native American mtDNAs. *American Journal of Human Genetics*; 53(3):563-90.

Umetsu K., Tanaka M., Yuasa I., Saitou N., Takeyasu I., Fuku N., ... Osawa M. (2001). Multiplex amplified product-length polymorphism analysis for rapid detection of human mitochondrial DNA variations. *Electrophoresis*, 22(16), 3533–8.

Umetsu K., Tanaka M., Yuasa, I., Adachi N., Miyoshi A., Kashimura S., ... Osawa M. (2005). Multiplex amplified product-length polymorphism analysis of 36 mitochondrial single-nucleotide polymorphisms for haplogrouping of East Asian populations. *Electrophoresis*, 26(1), 91–98. <https://doi.org/10.1002/elps.200406129>.

Underhill P. A. & Kivisild T. (2007). Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annual Review of Genetics*, 41, 539-564.

Valenzuela C. (2014). Principales Problemáticas y Potencialidades de La Región Nordeste. *Geograficando* 10(2). Disponible en línea: http://www.geograficando.fahce.unlp.edu.ar/article/view/Geov10n02a07/html_24.

van Oven M. & Kayser M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human Mutation* 30: E386-E394. <http://www.phylotree.org>.

van Oven M., Van Geystelen A., Kayser M., Decorte R. & Larmuseau M. H. (2014). Seeing the wood for the trees: A minimal reference phylogeny for the human Y chromosome. *Human Mutation*, 35(2), 187–191. <https://doi.org/10.1002/humu.22468>.

Vilar M.G., Melendez C., Sanders A.B., Walia A., Gaijeski J.B., et al. (2014) Genetic diversity in Puerto Rico and its implications for the peopling of the Island and the West Indies. *American Journal of Physical Anthropology* 155(3):352-68.

Volodko N. V., Starikovskaya E. B., Mazunin I. O., Eltsov N. P., Naidenko P. V., Wallace D. C., & Sukernik R. I. (2008). Mitochondrial genome diversity in arctic Siberians, with particular reference to the evolutionary history of Beringia and Pleistocenic peopling of the Americas. *The American Journal of Human Genetics*, 82(5), 1084-1100.

Wang S., Lewis C., Jakobsson M., Ramachandran S., Ray N. et al. (2007). Genetic variation and population structure in Native Americans. *PLoS Genet* 3(11): e185. <https://doi.org/10.1371/journal.pgen.0030185>.

Wei W., Ayub Q., Chen Y., McCarthy S., Hou Y., Carbone I., et al. (2013). A calibrated human Y chromosomal phylogeny based on resequencing. *Genome research*, vol. 23, pp. 388-395. <https://doi.org/10.1101/gr.143198.112>.

Y Chromosome Consortium. (2012). A nomenclature system for the tree of human Y-Chromosomal binary haplogroups. *Genome Research*, 12(2), 339–348. <https://doi.org/10.1101/gr.217602>.

Yang N.N., Mazieres S., Bravi C.M., Ray N., Wang S., Burley M-W et al. (2010) Contrasting patterns of nuclear and mtDNA diversity in Native American populations. *Annals of Human Genetics* 74(6):525–38.

Zheng HX,... & Jin L. (2012). MtDNA analysis of global populations support that major population expansions began before Neolithic Time. *Scientific Reports* 2, 745. <https://doi.org/10.1038/srep00745>.

7 ANEXOS

Anexo 2.1. Encuesta genealógica

<input type="checkbox"/>						
	VARÓN	Muestra N°	C Reg Hemo	Lab privado	Otro	Código final
<input type="checkbox"/>	<input type="checkbox"/>					
	MUJER					

DONANTE **Apellido y nombres (edad)**

DONANTE **Lugar de nacimiento** (localidad o paraje / distrito/ provincia / región / país)

¿Desea recibir sus resultados por correo electrónico? Por favor indíquenos su dirección:

1- PADRE **Apellido y Nombres** **Lugar de nacimiento** (localidad / distrito / provincia / región / país)

Pertenencia socio-étnica*¹: _____

2- ABUELO PATERNO **Apellido y Nombres** **Lugar de nacimiento** (localidad / distrito / provincia / región / país)

Pertenencia socio-étnica*¹: _____

3- ABUELA PATERNA **Apellido y Nombres** **Lugar de nacimiento** (localidad / distrito / provincia / región / país)

Pertenencia socio-étnica*¹: _____

4- MADRE **Apellido y Nombres** **Lugar de nacimiento** (localidad / distrito / provincia / región / país)

Pertenencia socio-étnica*¹: _____

5- ABUELO MATERNO **Apellido y Nombres** **Lugar de nacimiento** (localidad / distrito / provincia / región / país)

Pertenencia socio-étnica*¹: _____

6- ABUELA MATERNA **Apellido y Nombres** **Lugar de nacimiento** (localidad / distrito / provincia / región / país)

Pertenencia socio-étnica*¹: _____

*¹ Favor de preguntar al donante si sabe que alguno de sus padres/abuelos es/era descendiente de inmigrantes o perteneciente a alguna parcialidad socio-étnica (por ej. **judío sefaradí, árabe, afrodescendiente, criollo, mestizo, nativo, aborigen**, etc.). Cuando el donante desconoce el lugar de nacimiento de algún ancestro, favor de consignar el lugar de residencia habitual (e indicar en la encuesta con la leyenda “*vive o vivía en*”).

Anexo 2.2 Hoja de información

INFORMACION PARA EL POTENCIAL DONANTE

Proyecto de Investigación

“Aportes continentales diferenciales en la conformación de las poblaciones humanas de América Latina”

Datos generales del proyecto

Investigadores Responsables

Dr. Claudio M. Bravi (Investigador Adjunto CONICET, IMBICE, La Plata)

Dra. Graciela Bailliet (Investigador Independiente CONICET, IMBICE, La Plata)

Otros Integrantes del Grupo de Trabajo

Lic. Julieta Beltramo (Becario ANPCyT, IMBICE, La Plata)

Lic. Laura Jurado Medina (Becario CONICET, IMBICE, La Plata)

Lic. Marina Muzzio (Becario CONICET, IMBICE, La Plata)

Lic. María Rita Santos (Becario CONICET, IMBICE, La Plata)

Lic. Virginia Ramallo (Becario CONICET, IMBICE, La Plata)

Lic. Josefina Motti (Becario CONICET, IMBICE, La Plata)

Mag. José Edgardo Dipierri (Docente-Investigador, Universidad Nacional de Jujuy)

Lic. Emma Laura Alfaro (Docente-Investigador, Universidad Nacional de Jujuy)

Unidad Ejecutora

Laboratorio de Genética Molecular Poblacional

Instituto Multidisciplinario de Biología Celular (IMBICE)

Calle 526 entre 10 y 11, (1900) La Plata, República Argentina

TeleFax: (0221) 421-0112

¿Cómo contactarse con nosotros?

Envíe un e-mail a linajes2007@yahoo.com.ar o bien escriba o llame por teléfono al IMBICE.

Financiación

CONICET: Proyecto de Investigación Plurianual 2010 N° 1

Agencia Nacional de Promoción Científica y Tecnológica: PICT 2008 N° 715

INFORMACIÓN PARA EL PARTICIPANTE VOLUNTARIO

Estamos invitando a usted y a otras personas de su comunidad a participar de un proyecto de investigación científica que permitirá entender mejor el origen y la evolución de las poblaciones humanas que habitan diferentes regiones de América Latina. El proyecto implica la obtención de una muestra de sangre o saliva y de información sobre padres y abuelos de cada donante.

1- Objetivos del Proyecto de Investigación

Este proyecto tiene entre sus objetivos conocer que proporción de los ancestros de los habitantes actuales de diferentes regiones de América era originaria de este continente o bien llegó aquí desde Europa, Asia o África.

2- Un poco más de información...

Todas las personas recibimos de nuestros padres la información genética que determina nuestro aspecto físico y otras características individuales como el color de pelo y de ojos, el grupo y factor sanguíneo, etc. El estudio de una parte de la información genética nos permite conocer dónde se originaron nuestros antepasados.

La población americana es el resultado de una larga historia de mezcla y de migraciones de personas nacidas en diferentes partes del mundo. Muchos habitantes actuales del continente son descendientes directos de los primeros ocupantes de estas tierras, los pueblos originarios o indígenas. Otros son hijos, nietos, bisnietos, etc. de inmigrantes llegados de Europa (gallegos, vascos, napolitanos, rusos, alemanes, etc.), Medio Oriente (sirios, libaneses, etc.) o Lejano Oriente (japoneses, coreanos, chinos), mientras que otros tienen abuelos o bisabuelos "criollos", de los que no se conoce con certeza su origen. También hay personas que tienen algún lejano ancestro nacido en África.

3- Metodología a emplear

Las muestras de sangre o saliva donadas para este proyecto serán procesadas en el Laboratorio de Genética Molecular Poblacional del Instituto Multidisciplinario de Biología Celular (IMBICE), ubicado en la ciudad de La Plata, República Argentina.

En cada muestra donada para este estudio se analizarán diferentes porciones del ADN. Algunos de los análisis permiten conocer el origen de algunos de los antepasados del donante. El estudio del **ADN mitocondrial**, por ejemplo, permitirá saber si la abuela materna o la bisabuela materna, o la tatarabuela materna, o la madre de tatarabuela materna etc. hacia atrás varias generaciones...era de origen europeo, africano o americano. De modo parecido, el estudio del **Cromosoma Y** permitirá conocer el origen del ancestro remoto por vía paterna de los donantes varones.

La información genética obtenida en su comunidad será comparada con la obtenida por nosotros y otros colegas en estudios similares realizados en otras regiones de Argentina y en otros países del continente.

4- Riesgos del estudio,

La participación como voluntario en este proyecto no implica ningún riesgo para el donante. La extracción de sangre puede ocasionar una leve molestia, así como una pequeña mancha roja en el lugar del pinchazo, y más raramente que le baje la tensión arterial.

5- Confidencialidad de los datos

La información obtenida es confidencial y anónima: en ningún lugar se hará público el nombre de las personas participantes ni sus características genéticas. Sólo serán publicados datos generales para cada población estudiada.

6- Calidad del Participante

Su participación es voluntaria, y usted puede elegir retirarse en cualquier momento del estudio, poniéndose en contacto con los investigadores responsables en la dirección, el teléfono o el e-mail indicados arriba. Usted tiene derecho a recibir los resultados de los análisis que se hagan con su muestra, así como solicitar que se le devuelva o destruya el sobrante de su muestra.

Su participación será muy agradecida y contribuirá a la reconstrucción de la historia de las poblaciones humanas de América Latina.

Si usted acepta participar de este estudio, le agradeceremos que preste su conformidad por escrito completando y firmando el Formulario de Consentimiento que se le entrega por separado.

Quedamos a su disposición para aclarar cualquier duda sobre este proyecto.

Dr. Claudio Bravi - Dra. Graciela Bailliet: Investigadores responsables

Anexo 2.3 Formulario de consentimiento informado

Consentimiento Informado del Proyecto de Investigación “Aportes continentales diferenciales en la conformación de las poblaciones humanas de América Latina”

Yo,, DNI N°.....con domicilio legal en calle..... N°..... de la localidad de de la provincia de declaro que he sido informado sobre la realización de un estudio genético para conocer el origen o procedencia de los antepasados de los habitantes actuales de América y averiguar qué proporción de ellos es originaria de este continente o bien llegó aquí desde Europa, Asia o África.

Manifiesto acceder voluntariamente a participar en este proyecto, del cual he sido informado a través de una entrevista con miembros del grupo de investigación (abajo firmantes), quienes me comunicaron ampliamente sobre las características y alcances del estudio y me entregaron una hoja de “Información para el Participante” donde se detallan los objetivos, características genéticas a investigar, metodología, beneficios, riesgos y confidencialidad de los datos. Luego de leer esta hoja de información he podido realizar libremente cualquier pregunta relacionada con el proyecto.

Comprendo las características del trabajo y acepto libre y voluntariamente que me sea tomada una muestra de sangre o de saliva para ser utilizada en el presente estudio. Acepto además aportar datos sobre el origen o procedencia de mis padres, abuelos y bisabuelos para los fines indicados en la hoja de información. He comprendido que si lo deseo puedo retirarme de la investigación sin tener que dar explicaciones, como así también solicitar los datos obtenidos con mi muestra y reclamar la muestra sobrante del estudio.

La extracción de sangre o saliva (tachar lo que no corresponda) fue realizada por..... Cargo..... DNI:.....

La muestra es recibida por el investigador responsable quien la codifica y archiva el consentimiento informado en el Laboratorio de Bioantropología de la Universidad Nacional de Jujuy. La toma de la muestra y su codificación se realizó ante la presencia del/los testigos abajo firmantes.

Dado a losdías del mes de.....de 200....

_____	_____
Firma del donante	Aclaración y DNI
_____	_____
Firma y cargo del entrevistador	Aclaración y DNI
_____	_____
Firma y cargo del extraccionista	Aclaración y DNI
_____	_____
Firma del testigo	Aclaración y DNI

Investigadores Responsables: Dr. Claudio Bravi - Dra. Graciela Bailliet

Instituto Multidisciplinario de Biología Celular (IMBICE), Calle 526 e/ 10 y 11, La Plata.
Tel/Fax: (0221) 421-0112

Anexo 2.4 Aprobaciones de los comités de ética

COMITÉ PROVINCIAL DE BIOÉTICA PROVINCIA DE JUJUY

Ref. Evaluación de Protocolo

San Salvador de Jujuy, 22 de diciembre del 2008.

Sr. Ministro de Salud
Dr. Víctor Urbani

PRESENTE

Por la presente me dirijo a Ud., a fin de elevar para su conocimiento, el informe de APROBACION realizado por el Comité Provincial de Bioética de Jujuy, el día 22 de Diciembre del 2008, sobre el protocolo de Investigación, presentado por el Dr. Dipierri, sobre "Aportes continentales diferenciales en la conformación de las poblaciones humanas de América Latina"

Sin más, saluda a Ud. muy atte.


Dra. María de La Paz Bassio
Vice Pres


Lic. María Giardino
Presidente


Ministerio de Salud
Provincia de Jujuy



Comité Institucional de Ética de Investigación en Salud (CIEIS).
Santiago del Estero-República Argentina.
E-mail: pcieissgo@gmail.com

Santiago del Estero, 28 de Mayo de 2012.

Al Sr. Investigador Principal del Protocolo N° 017/2009 "Aportes Continentales diferenciales en la conformación de las poblaciones humanas de América Latina"

Dr. Claudio M. Bravi
Investigador Adjunto CONICET
Laboratorio de Genética Molecular de

Ref. Actualización Registro Protocolos CIEIS

De mi mayor consideración:

Me dirijo a Ud., con el fin de comunicarle que se registró la Renovación de la Autorización del protocolo de referencia de acuerdo a vuestra solicitud, para los años 2011 y 2012, venciendo dicha registración el 11/12/2012.

Debe destacarse que dicha renovación del Registro es un acto **ADMINISTRATIVO**, a los fines registrales, y que **NO MODIFICA** el plazo de la autorización emitida por el CIES para la ejecución del protocolo en cuestión, **vigente por un año a partir del reclutamiento del primer paciente en la Provincia de Santiago del Estero**, hecho que debe ser informado por los canales habituales al Comité.

Sin otro particular, hago propicia la oportunidad para saludarlo atentamente.-




Dr. Daniel Wozniak
PRESIDENTE
C. I. E. I. S.
Min. de Salud y Des. Social



Banco Central de
Sangre
Dr. César Guerra
SI.PRO.SA
TUCUMAN

San Miguel de Tucumán, 21 de enero de 2013.-

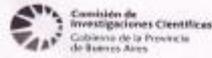
**Al Investigador Adjuato
CONICET, IMBICE, La Plata
Dr. Claudio Bravi**

En carácter de Directora General de Programa Integrado de Salud (PRIS) SI.PRO.SA, me dirijo a Usted a fin de comunicarle que el Banco Central de Sangre dependiente del PRIS, cuenta con la autorización correspondiente para participar del estudio multicéntrico "Aporte Continentales Diferenciales en la conformación de las Poblaciones Humanas de América Latina", organizado por el laboratorio de Genética Molecular Poblacional IMBICE, CONICET, a realizarse durante el corriente año.

Sin más, aprovecho la oportunidad para saludarle atte, encontrándome a su disposición para lo que considere necesario.

DR. SANDRA TIRADO
DIRECTORA GENERAL
DEL PROG. INT. DE SALUD
SI.PRO.SA.

Nota.
Se solicita devolución de informe final
del proyecto.



La Plata, 13 de Agosto de 2009.-

Dra. Graciela Bailliet
Dr. Claudio M. Bravi
IMBICE
S/D

Estimados Dres. Bailliet / Bravi:

Tenemos el agrado de dirigirnos a uds. con el objeto de informarles que el Comité de Ética en Investigaciones Biomédicas del IMBICE evaluó y aprobó las modificaciones incorporadas al texto de la Cartilla de Información para el Donante y a la organización del formulario de Consentimiento Informado, ambos elaborados en el marco del Proyecto "Aportes continentales diferenciales en la conformación de las poblaciones humanas de América Latina", que planean desarrollar en el laboratorio de Genética Molecular Poblacional del IMBICE.

Sin otro particular saludamos a uds. atentamente.

Dra. Lidia A. Vidal Rioja

Comité de Ética en Investigaciones Biomédicas
IMBICE
Dra. Lidia A Vidal Rioja - Coordinador
Dr. Ricardo S. Calandra
Prof. Jorge Asuaje
Dr. Luis Julio Couyet
Dr. Eduardo Luis Tinant

La Plata, 12 de Julio de 2017.-

Ante quien corresponda

Por la presente informo que el Comité de Ética en Investigaciones Biomédicas del IMBICE evaluó los antecedentes y grado de avance del proyecto "**Aportes continentales diferenciales en la conformación de las poblaciones humanas de América Latina**", que realizan los Dres Graciela Bailliet y Claudio M Bravi en el laboratorio de Genética Molecular Poblacional del IMBICE, y comprobó la continuación del mismo bajo las normas del "Marco Ético de Referencia para las Investigaciones Biomédicas en Humanos" (CONICET Res. D No. 1047/05). En consecuencia, el Comité de Ética en Investigaciones Biomédicas del IMBICE declara vigente la aprobación del texto de la Cartilla de Información para el donante y la Organización del Formulario de Consentimiento Informado, emitida el 13 de Agosto de 2009.

Sin otro particular, saludo con atenta consideración.

Dra Lidia A de Vidal Rioja

Comité de Ética en Investigaciones Biomédicas IMBICE

- RENIS CE000023

Dra Lidia A Vidal Rioja –Coordinadora
Prof. Jorge Asuaje - Miembro
Dr. Luis Julio Çouyet - Miembro
Dr. Eduardo Luis Tinanti - Miembro

Anexo 2.5 Informe para el participante



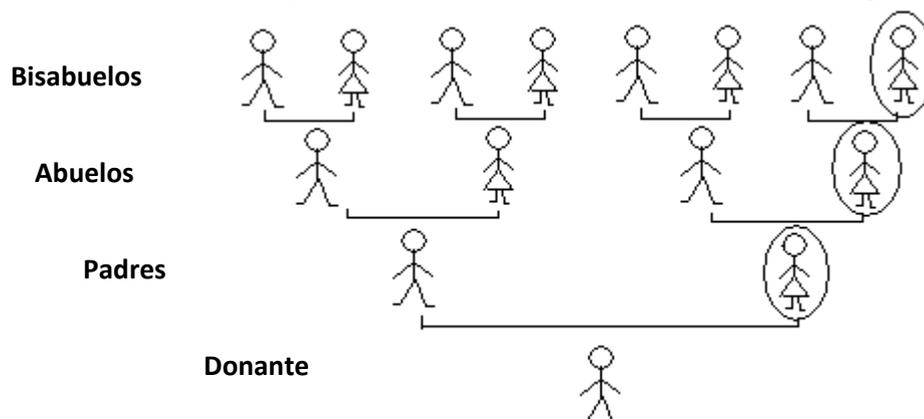
Calle 526 y Camino General Belgrano
B1906APO La Plata, Buenos Aires Argentina
Tel. +54 221 421 0112
www.imbice.gov.ar

Estimado participante:

En primer lugar queremos agradecerle por su participación en el proyecto **Aportes continentales diferenciales en la conformación de las poblaciones humanas de América Latina**, para el cual usted donó una muestra biológica y proporcionó información genealógica en el año 2015. Hoy podemos informarle el resultado de nuestros estudios con su muestra.

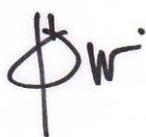
Ante todo, es necesario comprender que cada ser humano es un ser único: la forma de la nariz, el color de pelo, el grupo sanguíneo, etc. son rasgos que se configuran a partir de la información genética que hemos heredado de nuestros padres y que ellos a su vez heredaron de los suyos. Es así que en cada persona se reúne la herencia de decenas de individuos a los que denominamos ancestros. Esta herencia se encuentra almacenada en cada una de las células del cuerpo, en el material genético conocido como ADN.

Mediante este estudio no pretendemos conocer acerca de todos sus ancestros, sino que analizamos sólo una pequeña porción del ADN, llamado ADN mitocondrial, que se hereda exclusivamente a través de las madres. Con el estudio del ADN mitocondrial no podemos obtener información sobre su papá o sobre sus abuelos paternos, pero sabremos algo acerca de su línea materna (su mamá, su abuela materna, la madre de esta abuela y así sucesivamente hacia atrás), como lo muestra el siguiente esquema:



Tampoco estudiamos rasgos visibles como el color de ojos o la forma de la nariz, ni aquellas características que puedan producir enfermedades. Este no es un estudio médico sino que lo que “leemos” en el ADN nos cuenta sobre las migraciones de las personas.

Así, podemos informarle que por **vía materna**, usted tiene un ancestro (cercano o remoto) que era **nativo de América**.



Dr. Claudio Bravi



Dra. Graciela Bailliet

Laboratorio de Genética Molecular Poblacional
Instituto Multidisciplinario de Biología Celular
La Plata, julio de 2018 E-mail: linajes2007@yahoo.com.ar

Certificado de análisis de ADN mitocondrial

Por medio de la presente certifico que el linaje mitocondrial de la muestra obtenida de

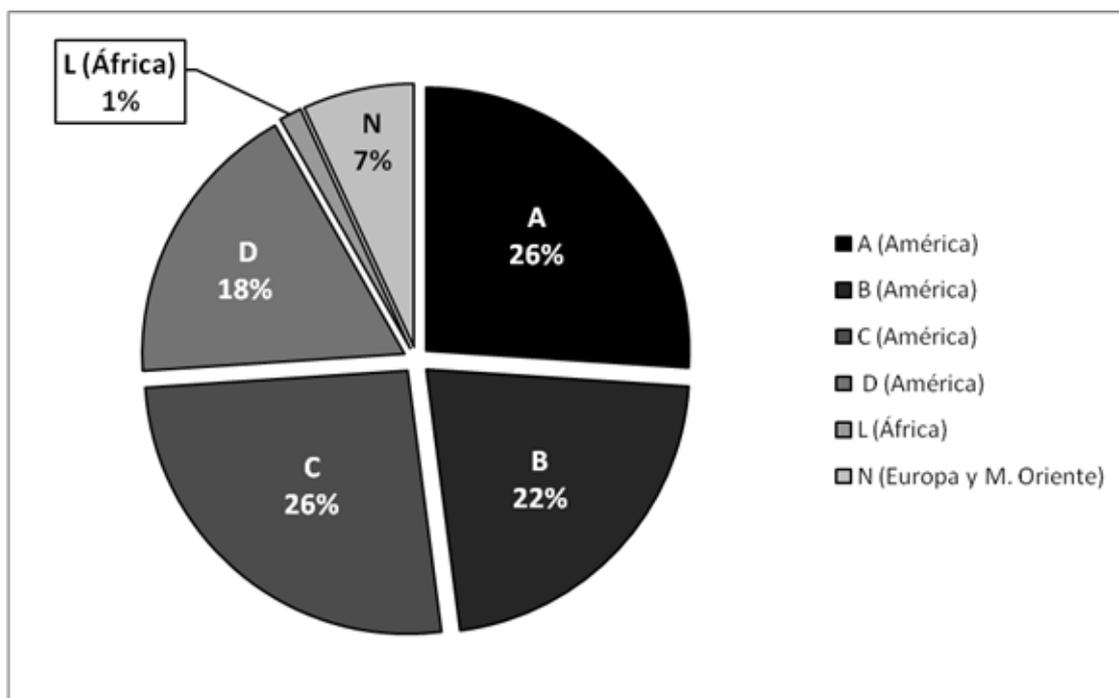
es asignable al siguiente clado del árbol mitocondrial humano:

Haplogrupo: A

Origen continental del linaje: América

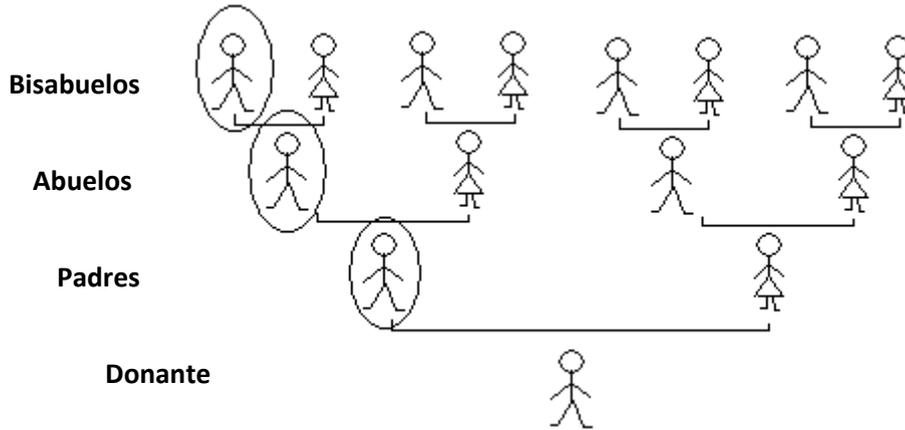
Origen continental de los linajes maternos analizados en una muestra de 75 voluntarios obtenida en el Centro Regional de Hemoterapia, Formosa, septiembre 2015

El 92% de los 75 individuos analizados resultaron portadores de linajes maternos de origen nativo americano (haplogrupos A-D). Los restantes se repartieron entre aquellos de origen africano (haplogrupo L, 1%) y los originados en Europa y/o Medio Oriente (haplogrupo N, 7%).



Certificado de análisis de haplogrupos del cromosoma Y

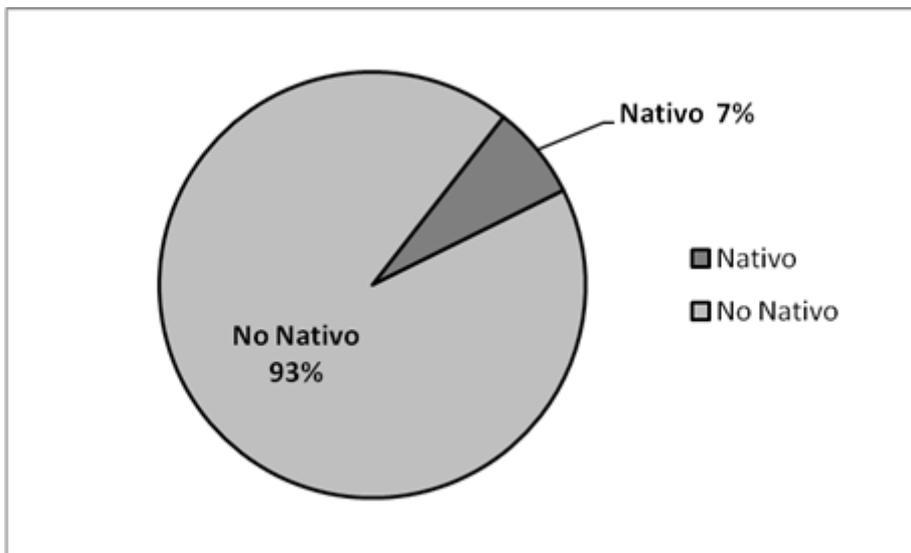
En los varones nos es posible estudiar un segmento del ADN que se encuentra en el cromosoma Y, que por lo tanto es exclusivo de los varones, y que sigue la línea paterna de modo semejante a la herencia de los apellidos. De esta manera podemos obtener información sobre su padre, su abuelo paterno, el padre de este abuelo y así sucesivamente hacia atrás, como lo muestra el siguiente esquema:



Así, podemos informarle que su linaje paterno es atribuible al **haplogrupo Q-M3** y por lo tanto tiene un origen remoto en **América**.

Origen continental de los linajes paternos analizados en una muestra de 28 voluntarios obtenida en el Centro Regional de Hemoterapia, Formosa.

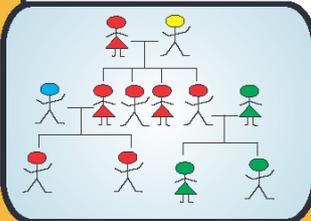
Sólo dos de los 28 varones analizados (7%) fueron portadores de linajes masculinos de origen nativo americano.



Anexo 2.6: Posters para las instituciones participantes.

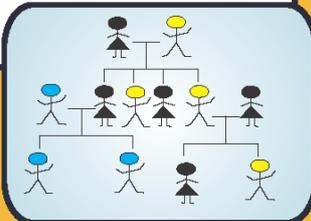
Origen continental de los linajes maternos y paternos de Formosa

¿Qué es un linaje materno?
Es el conjunto de ancestros y descendientes conectados a través de la línea femenina, es decir, a través de las madres.



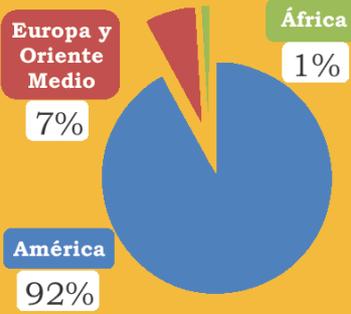
Origen de los linajes maternos

¿Qué es un linaje paterno?
Es el conjunto de ancestros y descendientes conectados a través de la línea masculina, es decir, a través de los padres.

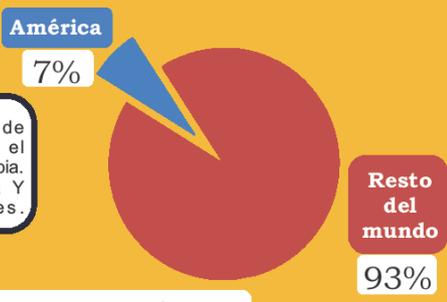


Origen de los linajes paternos

¿Cómo se estudian?
Analizando dos tipos de ADN: para los linajes maternos, el ADN mitocondrial que se hereda sólo a través de las madres; y para los linajes paternos, el cromosoma "Y", que sólo poseen los hombres y por tanto sólo se trasmite de padres a hijos varones.



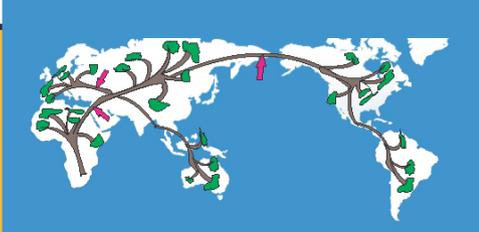
Origen de los linajes maternos



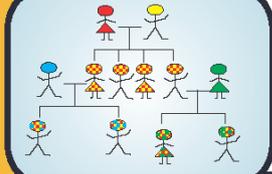
Origen de los linajes paternos

Datos obtenidos a partir de 103 muestras tomadas en el Centro Regional de Hemoterapia. El estudio del cromosoma Y se realizó en 28 varones.

¿Cómo podemos conocer el origen continental de un linaje?
Analizando las mutaciones que han ocurrido en el "árbol genealógico" de los seres humanos.



¿Qué pasa con el resto de los ancestros?
Por medio del ADN nuclear, recibimos partes iguales de información genética, tanto de lado paterno como materno. Pero éste no fue analizado en nuestro estudio.



¿Qué son las mutaciones?
Son cambios que ocurren en el código de la herencia llamado ADN



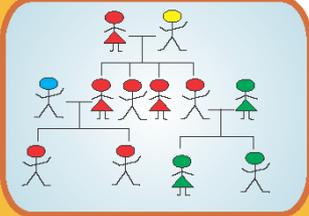
Contacto
Investigadores Responsables: Claudio Bravi y Graciela Bailliet
Laboratorio de Genética Molecular Poblacional.
Instituto Multidisciplinario de Biología Celular (IMBICE)
Calle 526 s/n entre 10 y 11, 1900 La Plata, República Argentina.
Tel. (0221) 421-0112 / e-mail: linajes2007@yahoo.com.ar



Origen continental de los linajes maternos y paternos de Formosa

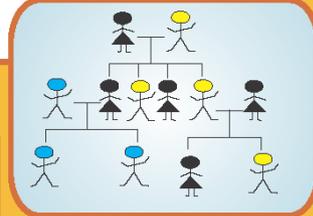
¿Qué es un linaje materno?

Es el conjunto de ancestros y descendientes conectados a través de la línea femenina, es decir, a través de las madres.



¿Qué es un linaje paterno?

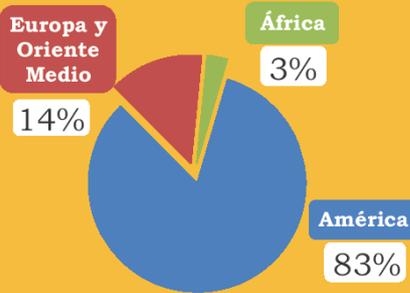
Es el conjunto de ancestros y descendientes conectados a través de la línea masculina, es decir, a través de los padres.



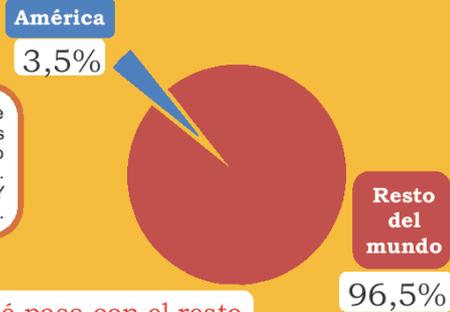
¿Cómo se estudian?

Analizando dos tipos de ADN: para los linajes maternos, el ADN mitocondrial que se hereda sólo a través de las madres; y para los linajes paternos, el cromosoma "Y", que sólo poseen los hombres y por tanto sólo se trasmite de padres a hijos varones.

Origen de los linajes maternos



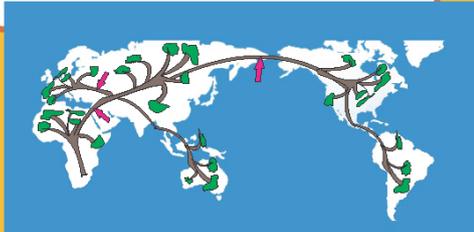
Origen de los linajes paternos



Datos obtenidos a partir de 228 muestras tomadas en el Laboratorio Dr. Raymundo Motter. El estudio del cromosoma Y se realizó en 116 varones.

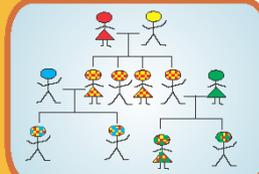
¿Cómo podemos conocer el origen continental de un linaje?

Analizando las mutaciones que han ocurrido en el "árbol genealógico" de los seres humanos.



¿Qué pasa con el resto de los ancestros?

Por medio del ADN nuclear, recibimos partes iguales de información genética, tanto de lado paterno como materno. Pero éste no fue analizado en nuestro estudio.



¿Qué son las mutaciones?

Son cambios que ocurren en el código de la herencia llamado ADN



Contacto

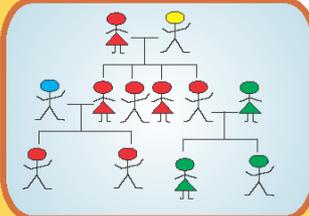
Investigadores Responsables: Claudio Bravi y Graciela Bailliet
Laboratorio de Genética Molecular Poblacional.
Instituto Multidisciplinario de Biología Celular (IMBICE)
Calle 526 s/n entre 10 y 11, 1900 La Plata, República Argentina.
Tel. (0221) 421-0112 / e-mail: linajes2007@yahoo.com.ar



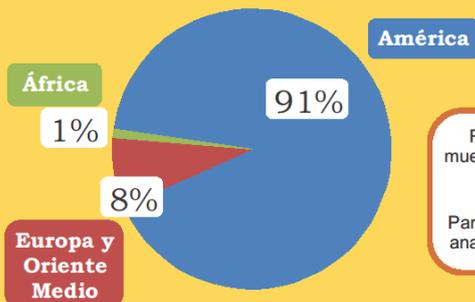
Origen continental de los linajes maternos y paternos de Corrientes

¿Qué es un linaje materno?

Es el conjunto de ancestros y descendientes conectados a través de la línea femenina, es decir, a través de las madres.

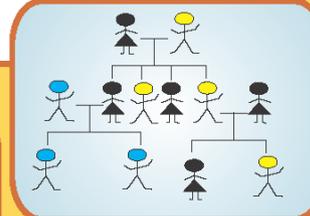


Origen de los linajes maternos



¿Qué es un linaje paterno?

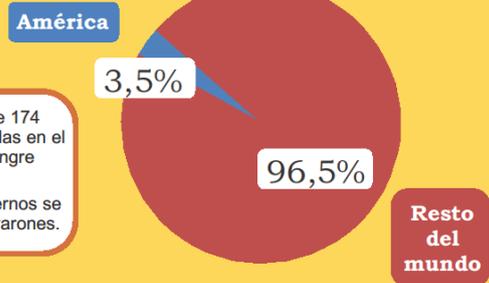
Es el conjunto de ancestros y descendientes conectados a través de la línea masculina, es decir, a través de los padres.



Origen de los linajes paternos

¿Cómo se estudian?

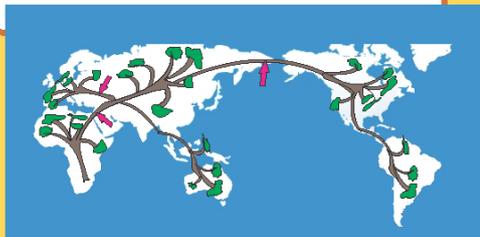
Analizando dos tipos de ADN: para los linajes maternos, el ADN mitocondrial que se hereda sólo a través de las madres; y para los linajes paternos, el cromosoma "Y", que sólo poseen los hombres y por tanto sólo se trasmite de padres a hijos varones.



Resultados de 174 muestras obtenidas en el Banco de Sangre Central. Para linajes paternos se analizaron 114 varones.

¿Cómo podemos conocer el origen continental de un linaje?

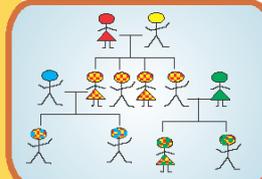
Analizando las mutaciones que han ocurrido en el "árbol genealógico" de los seres humanos.



Contacto

Investigadores Responsables:
 Claudio M. Bravi, Graciela Bailliet
 Lab. de Genética Molecular Poblacional
 Instituto Multidisciplinario de Biología Celular (IMBICE)
 (1900) La Plata, República Argentina
 Tel. (0221) 421-0112 / e-mail: linajes2007@yahoo.com.ar

¿Qué pasa con el resto de los ancestros?



Por medio del ADN nuclear, recibimos partes iguales de información genética, tanto de lado paterno como materno. Pero éste no fue analizado en nuestro estudio.

¿Qué son las mutaciones?

Son cambios que ocurren en el código de la herencia llamado ADN



Buenos Aires
Provincia



Anexo 2.7 Soluciones de trabajo

Solución de lisis de glóbulos rojos (SLGR) 20x

Para 1000 mL:

- ✓ 1,58 gr de Bicarbonato de Amonio (NH_4HCO_3)
- ✓ 123,05 gr de Cloruro de Amonio (NH_4Cl)
- ✓ Disolver en 800 mL de H_2O destilada
- ✓ Completar hasta 1000 mL con H_2O destilada
- ✓ Autoclavar
- ✓ En el campo, se lleva la solución a 1x mezclando 50 mL de SLGR 20x en 950 mL de H_2O destilada

Solución de lisis de glóbulos blancos (SLGB) 1x

Para 1000 mL:

- ✓ 100 mL de Tris-Cl 1M
- ✓ 50 mL de NaCl 1M
- ✓ 80 mL de EDTA 0,5 M
- ✓ 0,5 gr de Azida Sódica
- ✓ Completar con H_2O destilada hasta 1000 mL
- ✓ Autoclavar en dos botellas de 500 mL en cada una marcando el volumen final.
- ✓ Autoclavar en dos botellas de 500 mL en cada una marcando el volumen final.
- ✓ Agregar 10 mL de SDS al 20%

NaCl 1 M

Para 1000 mL:

- ✓ Disolver 58,44 gr de NaCl (1 mol) en 800 mL de H_2O destilada calentando la solución.
- ✓ Agregar H_2O destilada hasta completar el volumen final de 1000 mL
- ✓ Autoclavar en dos botellas de 500 mL en cada una marcando el volumen final.
- ✓ Autoclavar en dos botellas de 500 mL en cada una marcando el volumen final.

NaCl 5,4 M

Para 1000 mL:

- ✓ Disolver 315,576 gr de NaCl (5,4 moles) en 700 mL de H_2O destilada calentando la solución.

- ✓ Agregar H₂O destilada hasta completar el volumen final de 1000 mL (tener en cuenta que la solución se encuentra cercana al punto de saturación).
- ✓ Autoclavar en dos botellas de 500 mL en cada una marcando el volumen final.
- ✓ Completar con H₂O destilada en caso de que haya disminuido el volumen final.

SDS 20%

Para 100 mL:

- ✓ Disolver 20 gr de SDS⁺ (por sus siglas en inglés *Sodium Dodecyl Sulfate*) en 80 mL de H₂O destilada.
- ✓ Colocar un imán y agitar (suavemente para evitar que se genere espuma) hasta disolver.
- ✓ Enrasar a 100 mL con H₂O destilada.

Precaución al manipular el SDS, producto nocivo para las vías respiratorias.

Etanol 70%

Para 1000 mL:

En una probeta de 1000 mL

- ✓ Agregar 700 mL de etanol 96%.
- ✓ Enrasar con H₂O destilada a 1000 mL.

Etil-diamino-tetracético sal disódica (EDTA) 0,5 M

Para 500 mL:

- ✓ 93,05 gr de EDTA
- ✓ 480 mL de H₂O destilada
- ✓ Llevar a pH 8 con NaOH (10 gr aproximadamente)
- ✓ Completar con H₂O destilada hasta 500 mL

Tris-Cloruro 1M pH 8 (Tris-Cl)

Para 500 mL:

- ✓ 60,57gr de Tris-Base
- ✓ 400 mL de H₂O destilada
- ✓ Agregar HCl hasta alcanzar pH>7,8
- ✓ Completar con H₂O destilada hasta 500 mL

Buffer Tris-EDTA (TE)

Para 100 mL:

- ✓ Tris-Cloruro 1 M pH 8 1 mL
- ✓ EDTA 0,5 M 0,2 mL
- ✓ Enrasar con H₂O destilada a 100 mL.
- ✓ Ajustar el pH a 8 con NaOH
- ✓ Autoclavar y preservar en heladera

Buffer Tris-Borato-EDTA (TBE) 10x

Para 1000 mL:

- ✓ Tris base 108 gr
- ✓ Ácido Bórico 55 gr
- ✓ EDTA 0,5 M 40 mL
- ✓ Aforar con H₂O a 1000 mL

TBE 1x

Para 1000 mL:

En una probeta de 1000 mL

- ✓ Agregar 100 mL de TBE 10x
- ✓ Enrasar con H₂O destilada a 1000 mL.

Poliacrilamida 40%

Para 500 mL:

- ✓ Acrilamida 190 gr*
- ✓ Bisacrilamida 10 gr
- ✓ Aforar a 500 mL con H₂O destilada

***Precauciones en el manejo de acrilamida:** Usar guantes y barbijo durante su manipulación al tratarse de una sustancia con efectos carcinogénicos y teratogénicos, con riesgo de efectos graves para la salud en caso de exposición prolongada por inhalación o ingestión.

Persulfato de amonio 40%

- ✓ 2 gr de Persulfato de amonio
- ✓ 5 mL de H₂O destilada
- ✓ Agitar hasta disolver
- ✓ Alicuotar en tubos de 1,5 mL y envolver en papel aluminio.

- ✓ Conservar en freezer.

Gel de poliacrilamida 10%

Para 1 gel:

- ✓ 5 mL de Poliacrilamida 40%
- ✓ 20 mL de TBE
- ✓ 20 μ L de TEMED (tetrametiletilendiamina)
- ✓ 40 μ L de Persulfato de amonio 40%

Marcadores de peso molecular (*Ladders*)

De 25 pb:

- ✓ 2 μ L de ladder
- ✓ 15 μ L de buffer
- ✓ 23 μ L de TE

De 100 pb:

- ✓ 5 μ L de ladder
- ✓ 5 μ L de buffer

De 10 pb:

- ✓ 3 μ L de ladder
- ✓ 7 μ L de buffer

De 1000 pb:

- ✓ 0,2 μ L de ladder
- ✓ 5,2 μ L de buffer

Anexo 2.8 Procesamiento de muestras coaguladas

Materiales necesarios:

- Kit para extracción de sangre Gentra Puregene Blood Kit (1000 mL), cat no. 158389
- Proteínasa K Puregene (650 µL) o (5 mL), cat nos. 158918 y 158920
- Solución de glucógeno (500 µL), cat. no. 158930
- Canastas Clotspin (50), cat. no. 158932
- Isopropanol 100%
- Etanol 70% (usar alcohol desnaturalizado sin otras sustancias como metanol o metiletilcetona)
- Pipetas y tips
- Tubos para centrífuga de 50 mL
- Centrífuga capaz de alcanzar 2000 x *g* que pueda contener tubos de 50 mL
- Baños de agua de 37°C, 55°C y 65°C
- Vortex
- Tubo rotador
- Hielo en escamas

Protocolo QIAGEN:

Antes de comenzar, se prepararon los baños de agua a temperaturas de 37°C, 55°C y 65°C para realizar los pasos 1, 18 y 33 del protocolo. Se partió de muestras de sangre coaguladas que habían permanecido en freezer al menos 30 minutos (de -70°C a -80°C).

1. Se descongeló rápidamente la muestra de sangre coagulada y congelada a 37 °C y luego se colocó en hielo.
2. Se invirtió el tubo de muestra para aflojar el coágulo.
3. Se colocó 5–7 mL de coágulo de sangre en la cesta Clotspin contenida en el tubo de 50 mL.
4. Se centrifugó a 2000 x *g* durante 5 min para dispersar el coágulo.
5. Se agregó 15 mL de solución de lisis de glóbulos rojos a la canasta Clotspin para enjuagar la mayor cantidad posible de coágulo residual. Levantar con cuidado la cesta Clotspin en el tubo para permitir que la solución de lisis de glóbulos rojos fluya libremente.
6. Se retiró el inserto y, utilizando una punta de pipeta, se transfirió cualquier material de coágulo retenido en la cesta Clotspin al filtrado en el tubo de 50 mL.
7. Vortear vigorosamente durante 3 s a alta velocidad para dispersar aún más el material coagulado.

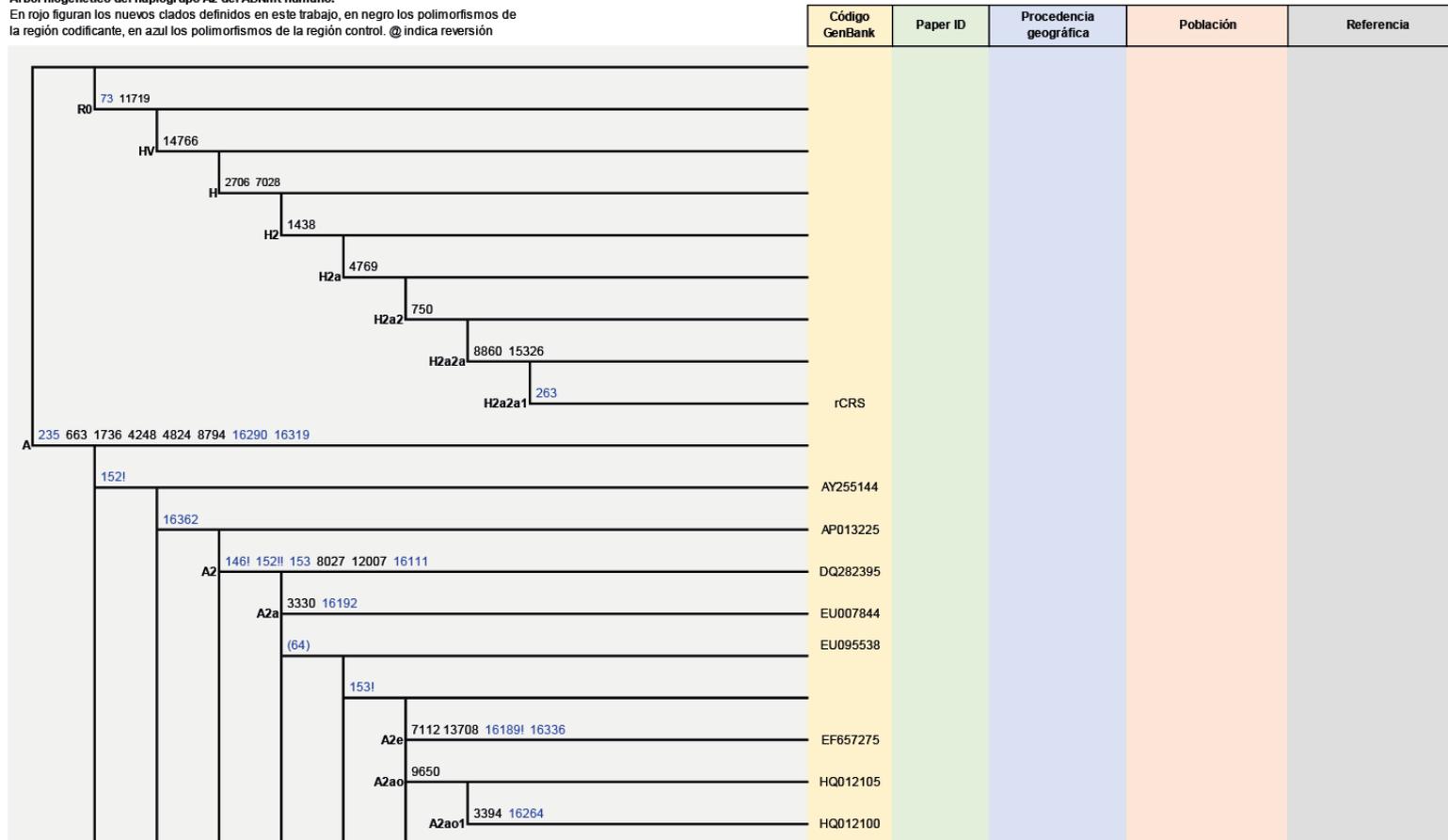
8. Se colocó la muestra en un rotador de tubos durante 5 min a temperatura ambiente (15–25 °C).
9. Vortear el tubo vigorosamente durante 3 s a alta velocidad.
10. Centrifugar a 2000 x g durante 5 min para sedimentar los glóbulos blancos y las partículas del coágulo. Nota: El pellet será de color rojizo.
11. Se desechó con cuidado el sobrenadante, teniendo cuidado de que el sedimento permanezca en el tubo.
12. Agregar 5 mL de solución de lisis de rojos al precipitado y agitar vigorosamente durante 3 s a alta velocidad.
13. Incubar en un rotador de tubos durante 5 min a temperatura ambiente.
14. Centrifugar a 2000 x g durante 5 min para sedimentar los glóbulos blancos y las partículas.
15. Se desechó con cuidado el sobrenadante, dejando 200 µL de líquido residual.
16. Vortear el tubo vigorosamente a alta velocidad para resuspender el sedimento en el líquido residual, esto facilita enormemente la lisis celular en el paso 17.
17. Se agregaron 5 mL de solución de lisis celular y 25 µL de Proteinasa K Puregene (20 mg/mL). Vortear vigorosamente durante 10 s a alta velocidad para iniciar la lisis celular y la digestión de proteínas.
18. Se completó la lisis celular incubando a 55 °C durante 2 h a toda la noche, hasta que se eliminen todas las partículas. Para facilitar la digestión, agite vigorosamente durante 10 s a alta velocidad al menos 3 veces durante la incubación.
19. Incubar en hielo durante 5 min.
20. Agregar 1,67 mL de solución de precipitación de proteínas al lisado celular y agitar vigorosamente durante 20 s a alta velocidad.
21. Centrifugar a 2000 x g durante 10 min.
22. Incubar en hielo durante 2 min.
23. Se colocaron 5 mL de isopropanol y 10 µL de solución de glucógeno (20 mg/ml) en un recipiente limpio de 50 mL (tubo para centrifuga).
24. Se agregó el sobrenadante del paso 21 vertiéndolo con cuidado. Asegurarse de no desalojar el sedimento de proteína al transferir el sobrenadante.
25. Se mezcló invirtiendo el tubo suavemente 50 veces.
26. Centrifugar a 2000 x g durante 3 min. El ADN debe ser visible como una pequeña bolita blanca.
27. Se desechó con cuidado el sobrenadante. Escurrir el tubo en un trozo limpio de papel absorbente, cuidando que el pellet quede en el tubo.

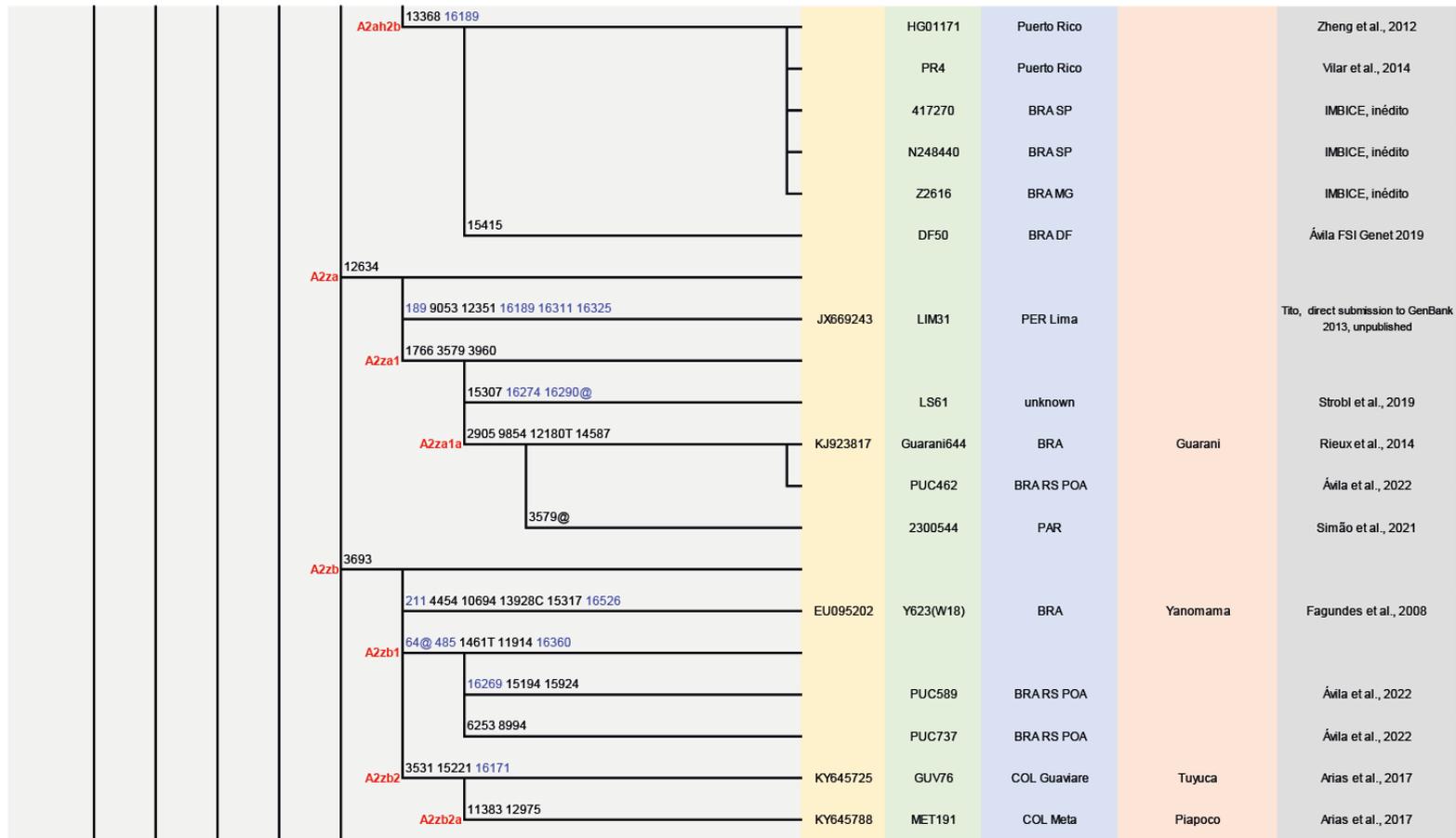
28. Se agregaron 5 mL de etanol al 70 % y se invirtió el tubo varias veces para lavar el sedimento de ADN.
29. Centrifugar a 2000 x g durante 1 min.
30. Se desechó con cuidado el sobrenadante. Escurrir el tubo en un trozo limpio de papel absorbente, cuidando que el pellet quede en el tubo. Es posible que el sedimento esté suelto y se desprenda fácilmente.
31. Se secó el ADN al aire a temperatura ambiente durante 10 min.
32. Se agregaron 500 μ L de solución de hidratación de ADN al tubo que contiene el sedimento.
33. Incubar a 65 °C durante 1 h para disolver el ADN.
34. Incubar a temperatura ambiente durante la noche con agitación suave. Asegúrese de que la tapa del tubo esté bien apretada cerrado para evitar fugas. A continuación, las muestras se centrifugaron brevemente y se transfirieron a un tubo de almacenamiento.

Anexo 2.9. Árboles filogenéticos del ADNmt humano en donde figuran los clados nuevos definidos en este trabajo.
 Detalle de abreviaturas utilizadas en el Anexo 2.10

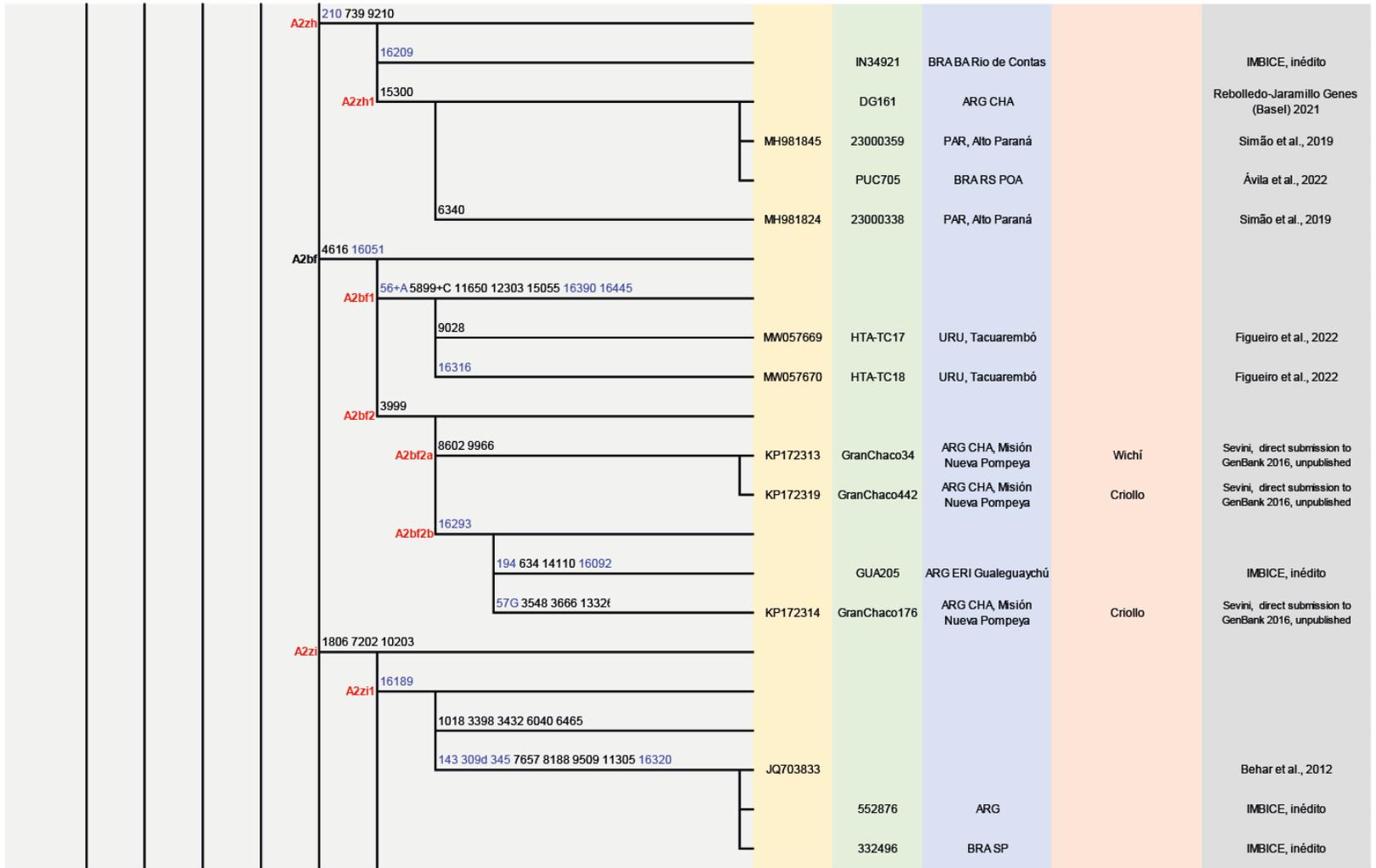
Árbol filogenético del haplogrupo A2 del ADNmt humano.

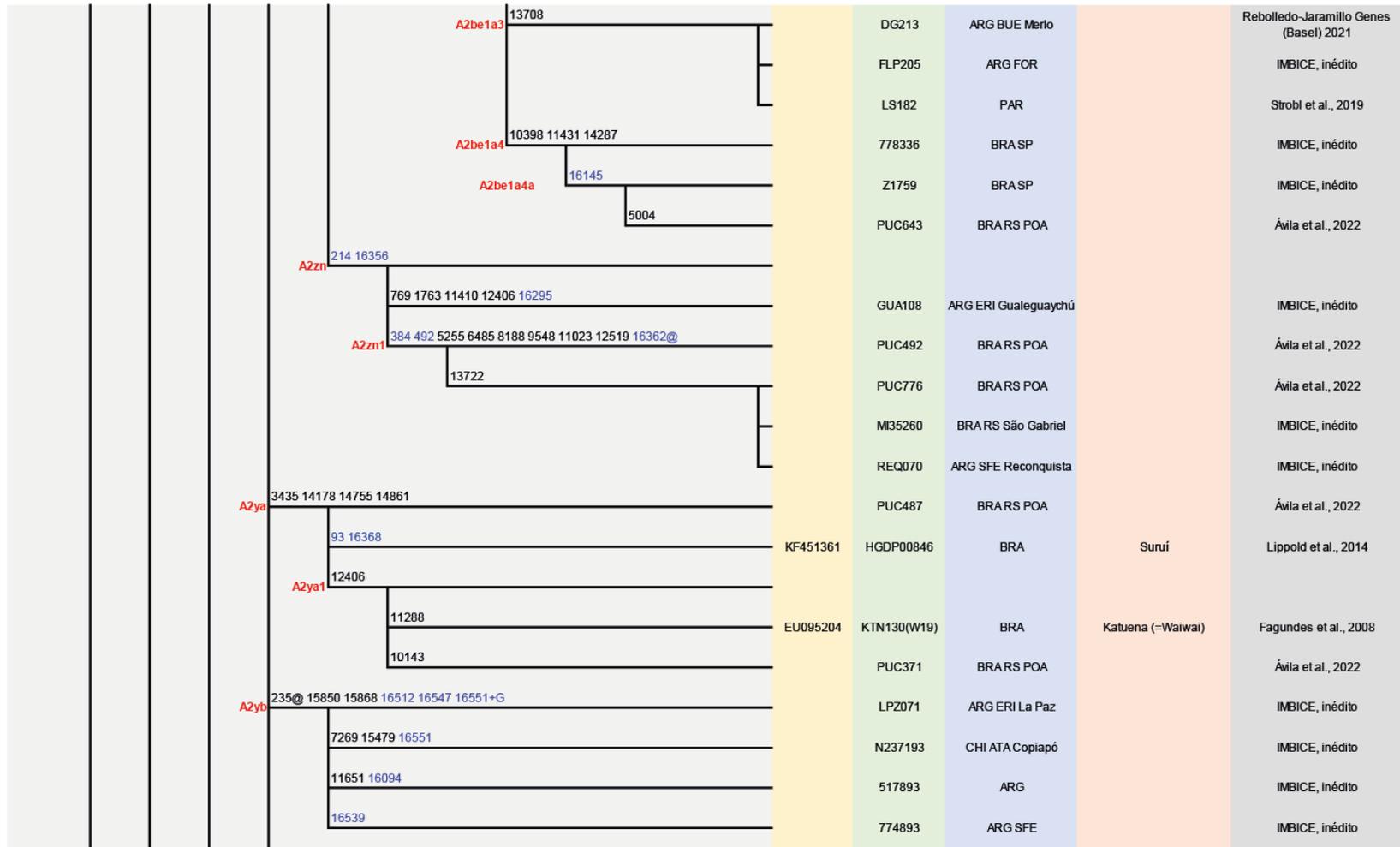
En rojo figuran los nuevos clados definidos en este trabajo, en negro los polimorfismos de la región codificante, en azul los polimorfismos de la región control. @ indica reversión

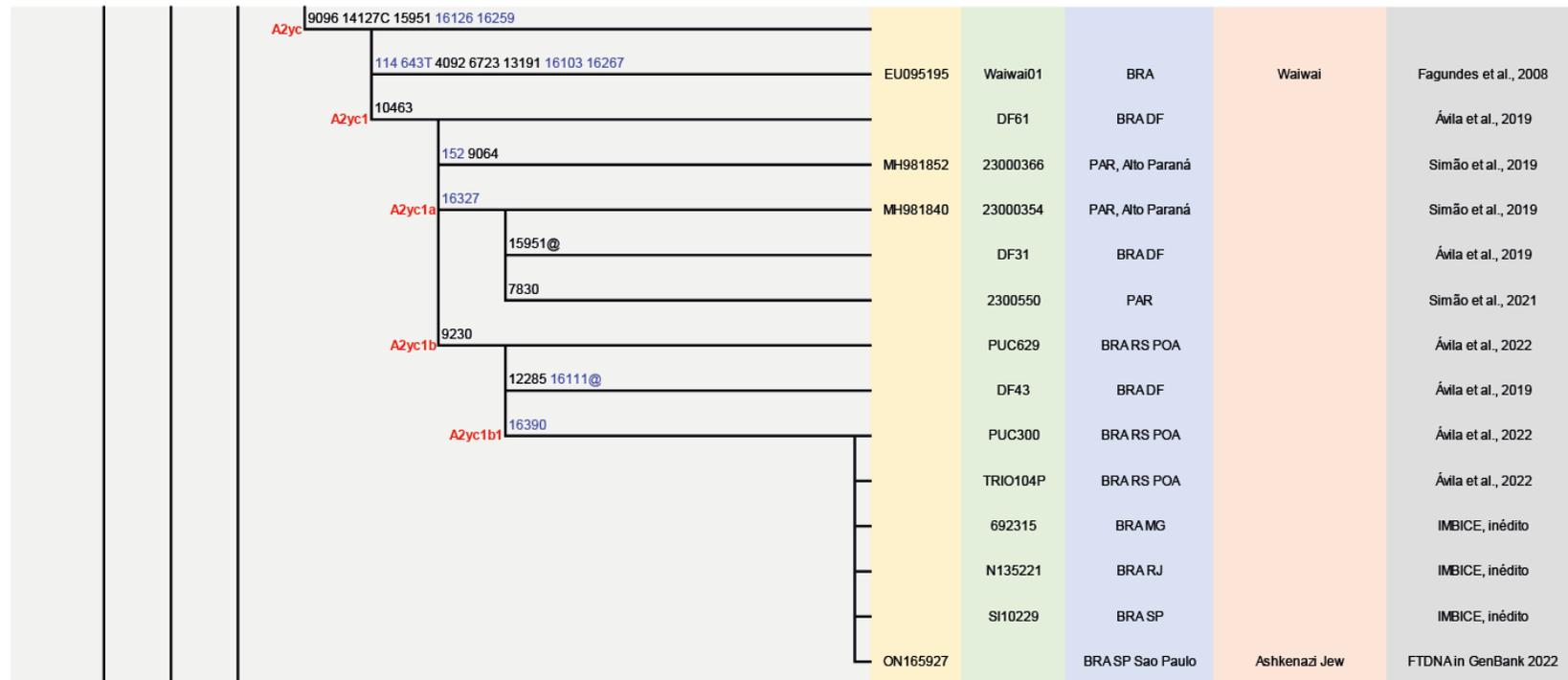




A2zc	3834								
	5030 10511 12372	JQ704350		unknown		unknown		Behar et al., 2012	
A2zc1	279 15244 16526		PR1	Puerto Rico				Vilar et al., 2014	
A2zc1a	4824@		I13541	Puerto Rico		aDNA		Fernandes et al., 2021	
	2267G 6842 12509C 15218 16111@	KM656459	F530	VEN, Amazonas		Yekuana (=Makiritare)		Lee et al., 2015	
	15784		2300545	PAR				Simão et al., 2021	
	153@		RS06	BRA RS POA				Ávila FSI Genet 2022	
A2zd	7981 13788A 15262		LPZ107	ARG ERI La Paz				IMBICE, inédito	
	12354	MH981906	23000511	PAR, Alto Paraná				Simão et al., 2019	
A2ze	3083 8276d 15734 16422								
	8870	EU095194	Aché30	PAR		Aché		Fagundes et al., 2008	
A2zf	150 14290 16311	JX669184	HY017	PER Junin, Huancayo				Tito, direct submission to GenBank 2013, unpublished	
	14443	JX669347	PE821C	PER Pasco, Cerro de Pasco				Tito, direct submission to GenBank 2013, unpublished	
	16093	JX669381	PE922	PER Pasco, Cerro de Pasco				Tito, direct submission to GenBank 2013, unpublished	
	9591 16422		GUA089	ARG ERI Gualeguaychú				IMBICE, inédito	
A2zg	573+C(n) 8843 9482 14040 14502		MI19040	BRA SP				IMBICE, inédito	
A2zg1	16104								
	5147		LPZ115	ARG ERI La Paz				IMBICE, inédito	
	408A 11172	MH981892	23000406	PAR, Alto Paraná				Simão et al., 2019	
A2zg1a	3531		528453	BRA MG				IMBICE, inédito	
	8853		LS297	unknown		unknown		Strobl et al., 2019	

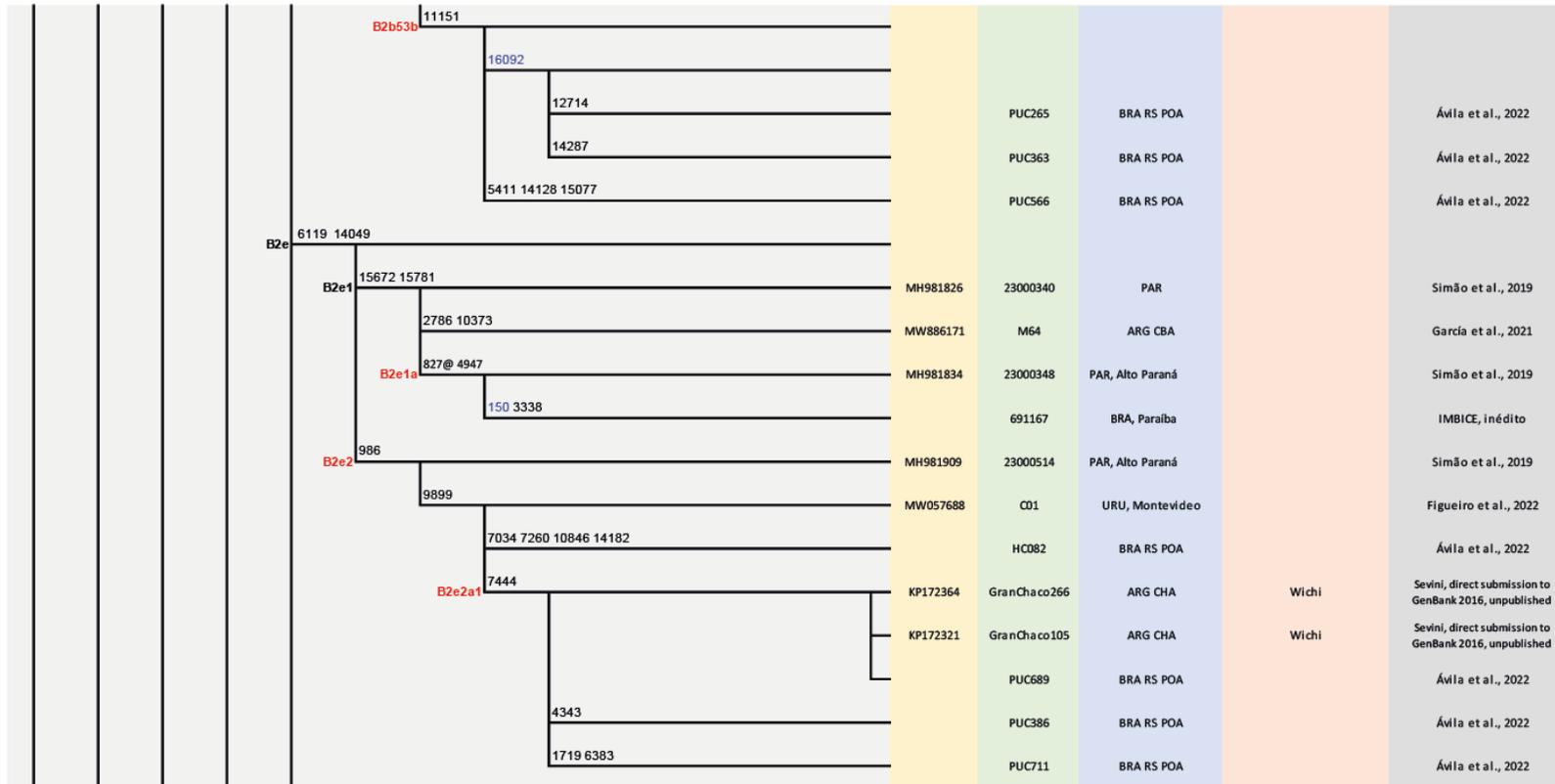


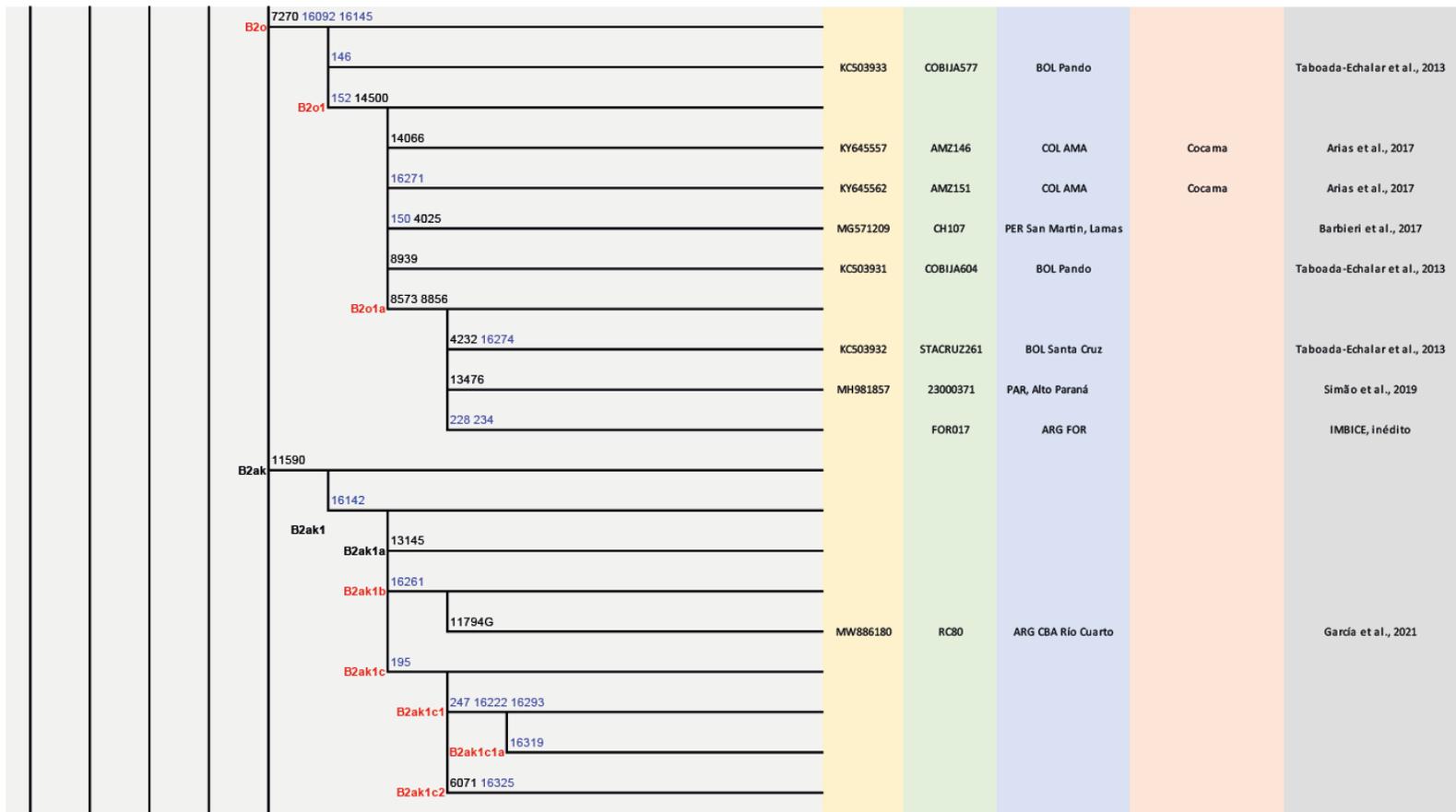


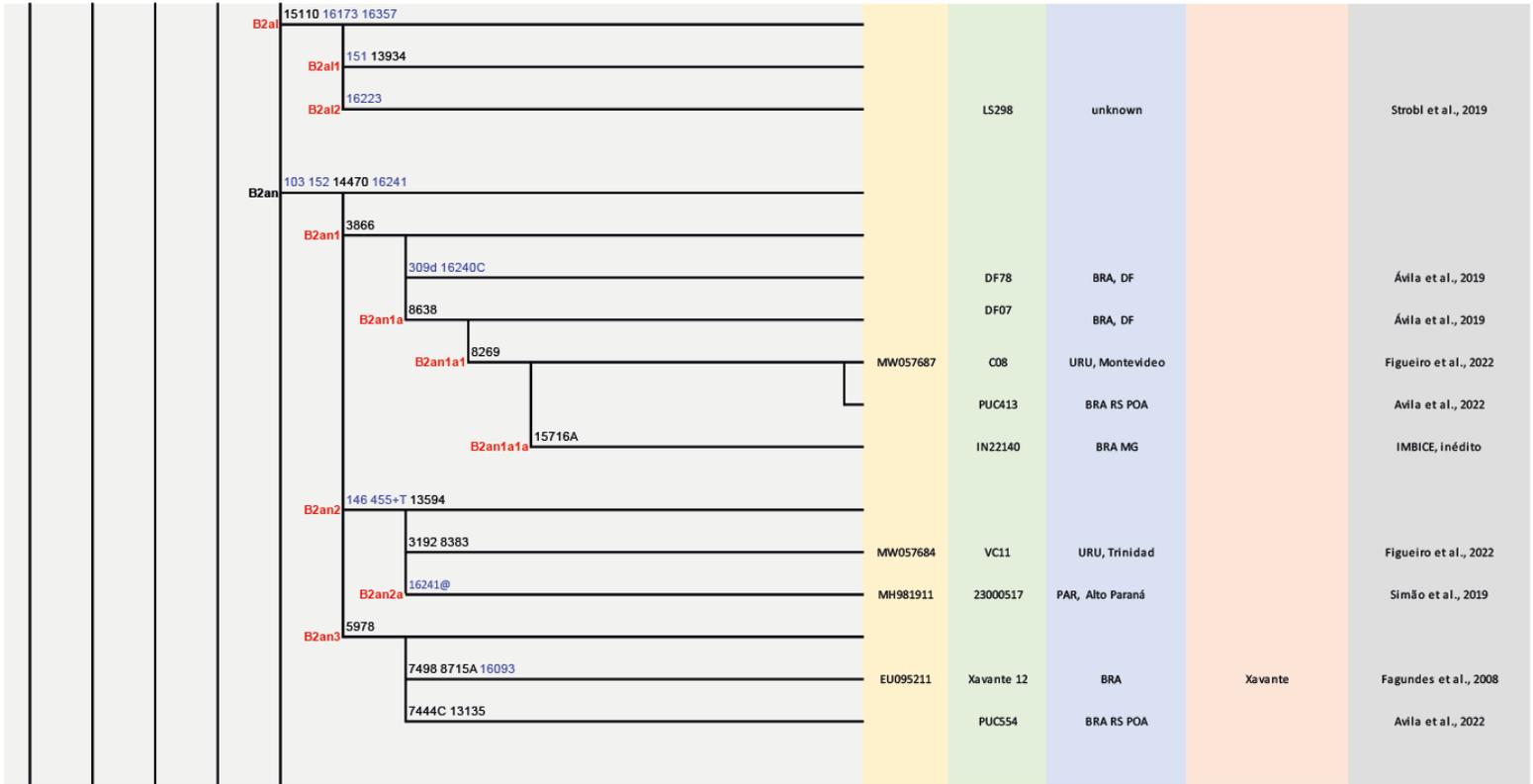


Árbol filogenético del haplogrupo B2 del ADNmt humano.
 En rojo figuran los nuevos clados definidos en este trabajo, en negro los polimorfismos de la región codificante, en azul los polimorfismos de la región control. @ indica reversión

827 15535		Código GenBank	Paper ID	Procedencia geográfica	Población	Referencia
B4b'd'e'	499 4820 13590					
B4b	3547 4977 6473 9950 11177	EU095548				
B2	16111 16483	KC711022				
B2a	6755	EU095532				
B2b	143 8296 16309					
B2b50	8946		B395930	CHI		IMBICE, inédito
	199 8701 10793	KC503928	Mco-10	ARG	Mataco	Taboada-Echalar et al., 2013
B2b51	309d 8185 11383 15968	JX669298	PE308	Perú, Lima		Tito, direct submission to GenBank 2013, unpublished
B2b51a	58 16356 16362					
B2b51a1	56+C 153 14162 16086		553714	CHI		IMBICE, inédito
B2b52	9804					
	4086 5563 6701 8584 10352 12456 15758 16153	MW886184	SRC13	ARG SI Santa Rosa de Conlara		García et al., 2021
	1709 3397 4012 4113 5153 7247 10295 15097 15115 16390		A98	unknown		Chiar et al., 2020
B2b53	146 152 2056 2875 5291 13928C					
	1664 3394 13500	MW057685	C05	URU, Montevideo		Figueiro et al., 2022
B2b53a	6286					
	13932 16168	EU095210	Xavante 04	BRA	Xavante	Fagundes et al., 2008
	1462		DF72	BRA DF		Ávila et al., 2019



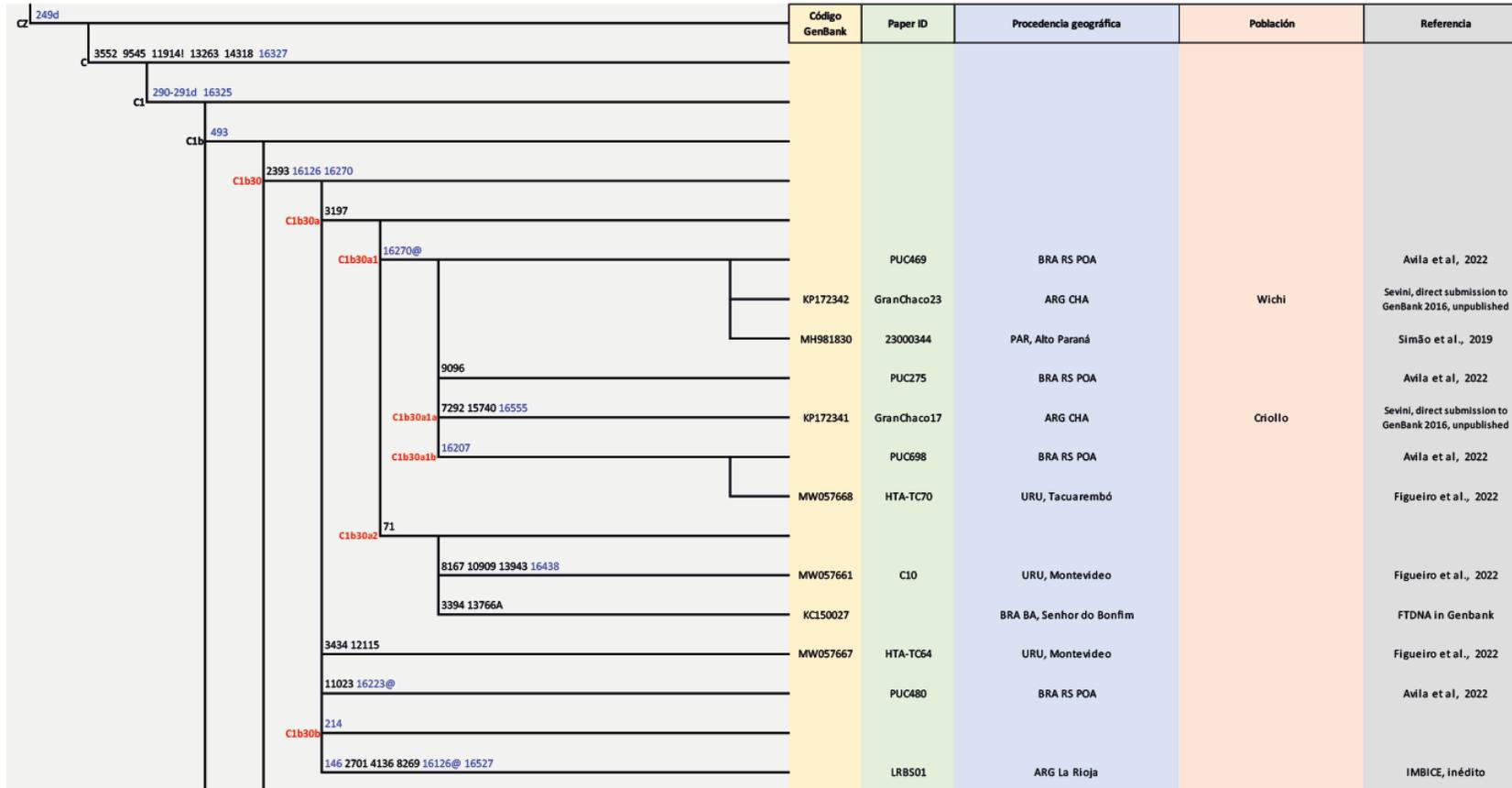




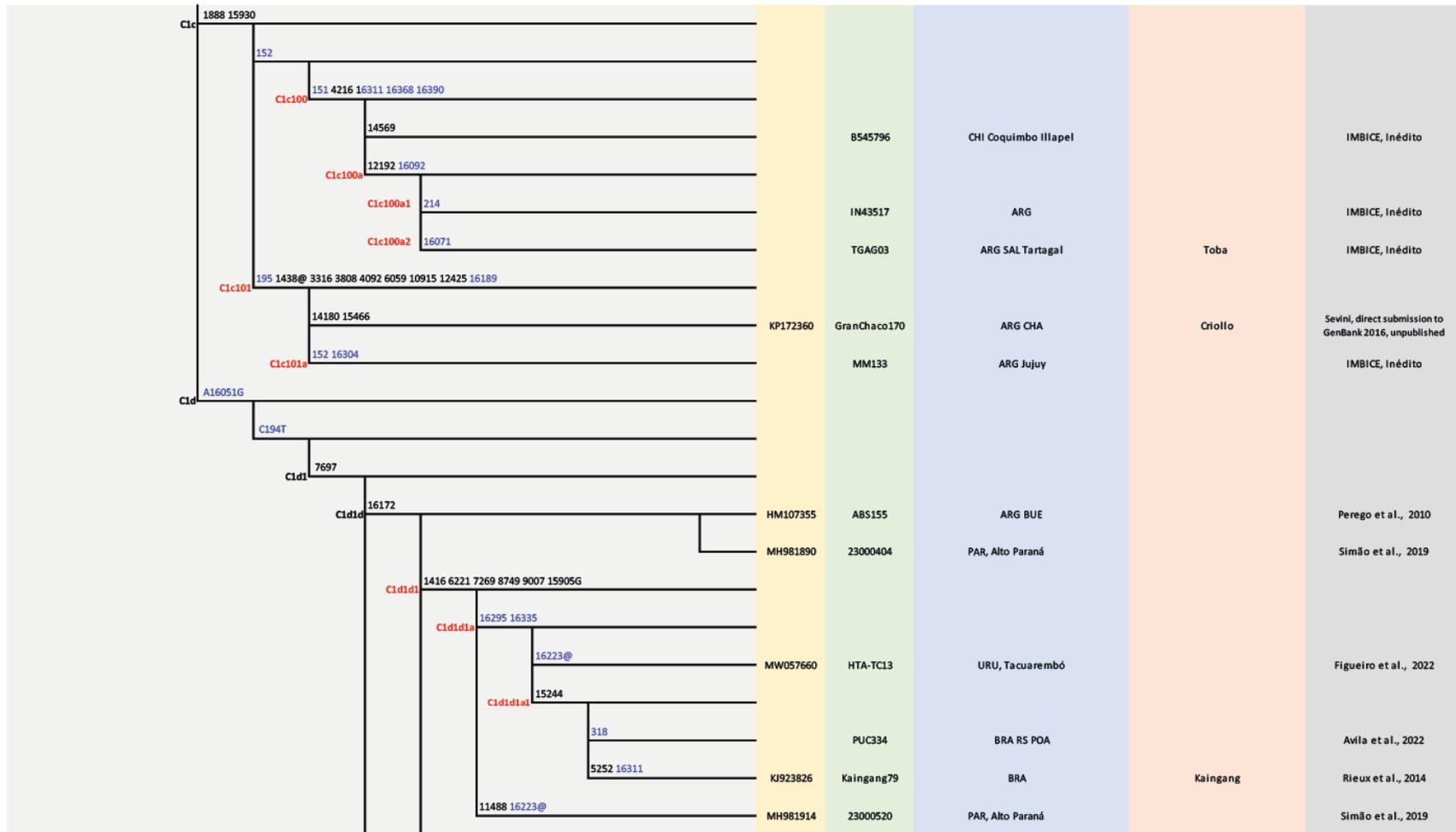
B2ze	12771	EU095207	Gaviao 23	BRA	Gavão	Fagundes et al., 2008	
	152 6471 10101 11125 12390 16261 16284		LS183			PAR	Strobl et al., 2019
	484 5743 6899 9935 11243 16145 16294A						
B2zf	1598 3615 4232 16311 16319 16524	KJ923825	Kaingang106	BRA	Kaingang	Rieux et al., 2014	
	16354		EU095219			KRC33	BRA
B2zf1	1598@ 7227T 7251	FLP212		ARG FOR		IMBICE, inédito	
	16399						
B2zf2	569 9664 14545 16173	PUC357		BRA RS POA		Avila et al., 2022	
	3710 12696 13602 13753 15316T 15895 16126 16372	621429		BRA SP		IMBICE, inédito	
B2zg	5557	359741		BRA RJ		IMBICE, inédito	

Árbol filogenético del haplogrupo C1 del ADNmt humano.

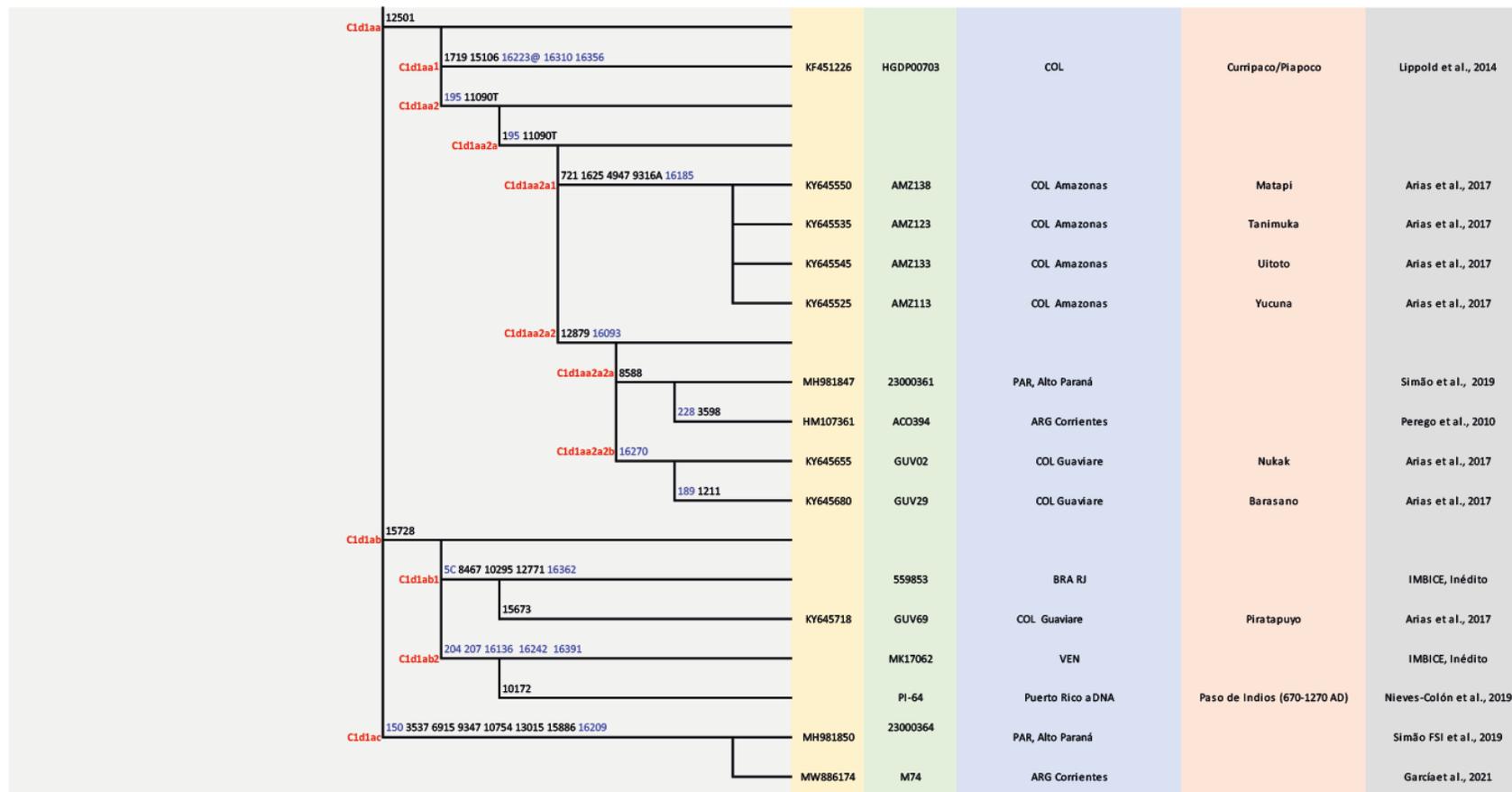
En rojo figuran los nuevos clados definidos en este trabajo, en negro los polimorfismos de la región codificante, en azul los polimorfismos de la región control. @ indica reversión



Cib100	10043 16192						
	204 4452 7340 10248 16173 16346		TG 099	ARG SAL Tartagal		IMBICE, inédito	
	207 679A 3369 6293 16356		SJ087	ARG San Juan		IMBICE, inédito	
Cib101	2404 8772 13827 16384		REQ 098	ARG SFE Reconquista		IMBICE, inédito	
	3338 12285		N3735	BOL Beni, Santa Ana del Yacuma		IMBICE, inédito	
Cib102	13212 16526						
	Cib102a	8764 8860@ 14233					
		Cib102a1	152 15734		DF58	BRA DF	Ávila et al., 2019
		Cib102a1a	5460 6956 16440		DF22	BRA DF	Ávila et al., 2019
	Cib102b	153 16172					
Cib102bi	7844 12223@	MW886206	VA47	ARG SDE Villa Atamisqui		García et al., 2021	
Cib103	8078 16292 16362		DF05	BRA DF		Ávila et al., 2019	
			N247427	BRA BA		IMBICE, inédito	
			628335	Suriname		IMBICE, inédito	
	3531		837057	BRA PE		IMBICE, inédito	
	4991		8197717	BRA AL		IMBICE, inédito	
	3764 4924	EU095227	ARL58(W02)	BRA	Arara Laranjal	Fagundes et al., 2008	
	Cib103a	8047		DF84	BRA DF		Ávila et al., 2019
		1719 8552		PUC523	BRA RS POA		Ávila et al., 2022

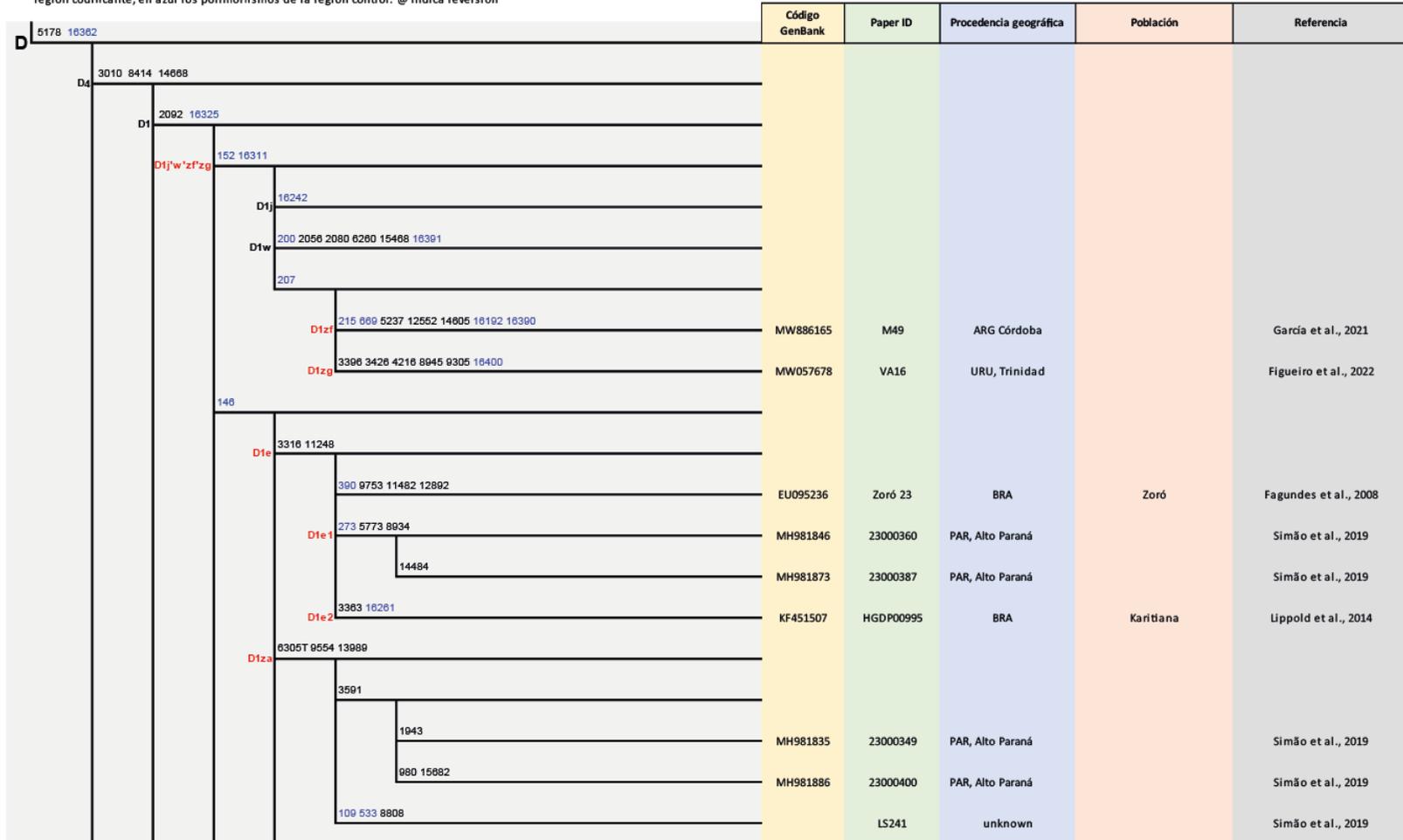


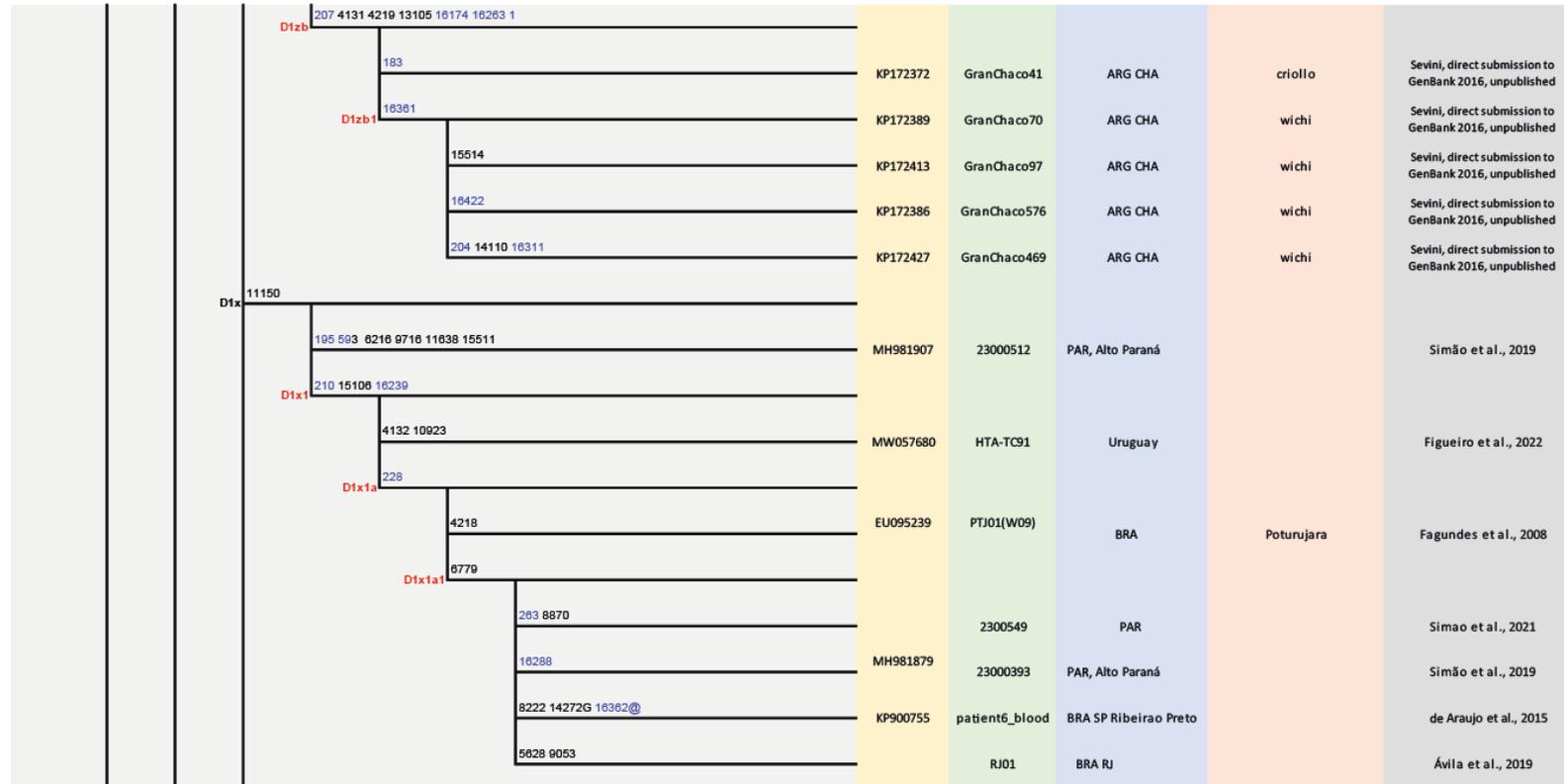
C1d1d2	1452 5836 8425 15852							
	6272 13770 15884	EU095537	Coreguaje 1-54	COL	Coreguaje	Tamm et al., 2007		
	235 13604 13934	KY645840	PUT284	COL, Putumayo	Kamentsa	Arias et al., 2017		
	709							
	C1d1d2a	16293	KY645709	GUV59	COL, Guaviare	Carijona	Arias et al., 2017	
		7397 9055	KY645895	PUT341	COL, Putumayo	Coreguaje	Arias et al., 2017	
		195 6347A 13215 14569	KY680980	NAT032	ECU PIC, Conocoto	Quitú-Cara	Brandiniet al., 2019	
		15628	HM107356	S-681199	BRA MS		Perego et al., 2010	
	13934		PUC725	BRA RS POA		Avila et al., 2022		
	146 195 7598 8440		HG01556	COL, Medellin		Capodiferro et al., 2021		
	197 1246 2262 6545 9148 9921 16209	HM107368	S-671578	Nicaragua, Puerto Cabezas		Perego et al., 2010		
	10837	MH981829	23000343	PAR, Alto Paraná		Simão et al., 2019		
	16311							
	C1dig	16287						
C1dig1		3338 4955 7354 11167 15034						
		C1dig1a	4048 4652 12930	HM107346	ABS326	ARG BUE	Perego et al., 2010	
				MH981900	23000415	PAR, Alto Paraná		Simão et al., 2019
					2300537	PAR		Siamo et al., 2021
		MW057658	HTA-TC100	URU, Tacuarembó		Figueiro et al., 2022		
C1dig1b		182 16093		PUC617	BRA RS POA		Avila et al., 2022	
		7534@		DF11	BRA DF		Avila et al., 2019	
C1dig2		146 13470 16184						
		14128		PUC572				
		4667 13350	HM107348	S-681241	URU		Perego et al., 2010	
		MW057659	C09	URU, Montevideo				



Árbol filogenético del haplogrupo D1 del ADNmt humano.

En rojo figuran los nuevos clados definidos en este trabajo, en negro los polimorfismos de la región codificante, en azul los polimorfismos de la región control. @ indica reversión

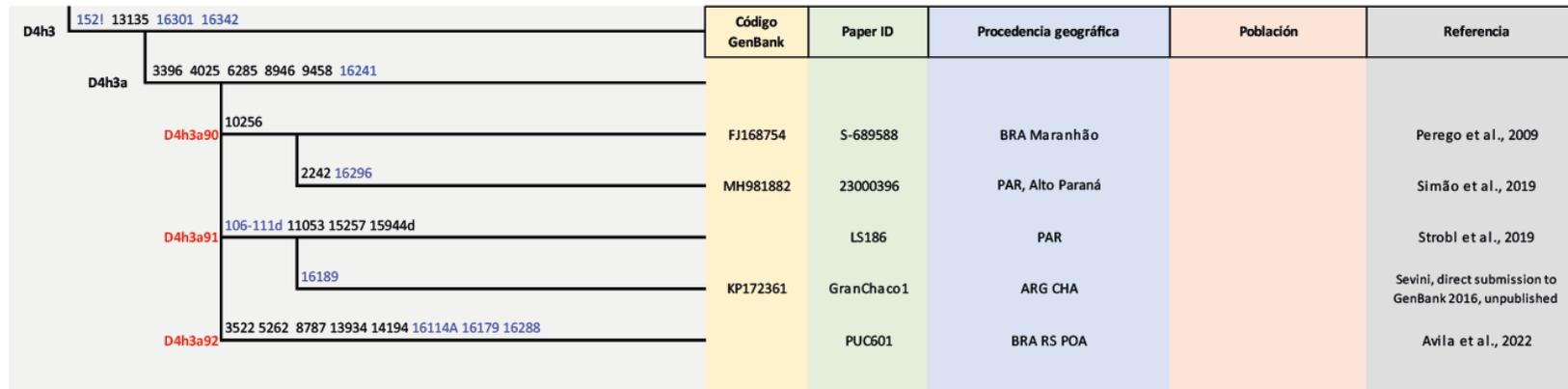




Dtzc	15519 16187A	JN253391	ABS311	ARG BUE	Criollo	Bodner et al., 2012
	5585	MW057681	A06	URU, Montevideo		Figueiro et al., 2022
	16148		PUC633	BRA RS POA		Avila et al., 2022
Dtzd	16291 16325@	KP172395	GranChaco14	ARG CHA	Criollo	Sevini, direct submission to GenBank 2016, unpublished
	2156+A 16093	JN253392	ARN127	ARG RNE		Bodner et al., 2012
	249 15709 16344		2300538	PAR		Simao et al., 2021
	Dtzd1 263@		B297987	ARG		IMBICE, inédito
Dtze	143 8705				Chinchawas, this 750-1100 CE	
	Dtze1 5085 5417 11305 16147 16291 16355		CWS2	PER ANC Huaraz, Pira		Nakatsuka et al., 2020
	Dtze2 6446 6719G 14605 16154 16193 16319	JX669287	PE029	PER Lima		Tito, direct submission to GenBank 2013, unpublished
Dtzh	195 297 4550 6485 15388 16189				Criollo	
	7080	KP172381	GranChaco507	ARG CHA		Sevini, direct submission to GenBank 2016, unpublished
	Dtzh1 4703 14048 16342		TMN 086	ARG TUC		IMBICE, inédito

Árbol filogenético del haplogrupo D4h3a del ADNmt humano.

En rojo figuran los nuevos clados definidos en este trabajo, en negro los polimorfismos de la región codificante, en azul los polimorfismos de la región control. @ indica reversión



Anexo 2.10. Lista de abreviaturas

IMBICE	Instituto Multidisciplinario de Biología Celular	SMA (ARG)	Santa María	MS (BRA)	Mato Grosso do Sul
ARG	Argentina	BEL (ARG)	Belén	ATA (CHI)	Atacama
CRR (ARG)	Corrientes	LR (ARG)	La Rioja	PAR	Paraguay
FOR (ARG)	Formosa	CHE (ARG)	Chepes	VEN	Venezuela
ERI (ARG)	Entre Ríos	CAL (ARG)	Calingasta	PER	Perú
FLP	Formosa, Laboratorio Privado Motter	SNJ (ARG)	San Juan	AMA (PER)	Amazonas
CBS	Corrientes, Banco de Sangre	TUL (ARG)	Villa Tulumaya	ECU	Ecuador
CHA (ARG)	Chaco	MZA (ARG)	Mendoza	PIC (ECU)	Pichincha
CBA (ARG)	Córdoba	MLG (ARG)	Malargüe	BOL	Bolivia
SAL (ARG)	Salta	URU	Uruguay	CHI	Chile
BUE (ARG)	Buenos Aires	COL	Colombia		
TUC (ARG)	Tucumán	BRA	Brasil		
SFE (ARG)	Santa Fe	DF (BRA)	Distrito Federal		
RNE (ARG)	Río Negro	POA (BRA)	Porto Alegre		
LPZ (ARG)	La Paz	SP (BRA)	São Paulo		
GUA (ARG)	Gualedguaychú	MT (BRA)	Mato Grosso		
CC (ARG)	Curuzú Cuatiá	RS (BRA)	Río Grande Do Sul		
REQ (ARG)	Reconquista	MG (BRA)	Minas Gerais		
TG (ARG)	Tartagal	BA (BRA)	Bahía		
LQ (ARG)	La Quiaca	CE (BRA)	Ceará		
MM (ARG)	Maimará	PE (BRA)	Pernambuco		
TMN (ARG)	Tucumán	RJ (BRA)	Río de Janeiro		
SDE (ARG)	Santiago del Estero	PI (BRA)	Piauí		

Anexo 3.1. Lista de muestras de la ciudad de Formosa y Corrientes en las que se secuenció la región control extendida de ADNmt. Se incluyen el rango de lectura, sitios polimórficos y asignación al clado más derivado.

ID	Haplogrupo ADNmt	Rango	Sitios polimórficos
FOR 020	A2aa	15837-16569; 1-676	16111 16223 16239A 16290 16319 16362 64 73 146 153 235 263 309+C 315+C 523-524d 663
FLP 151	A2ab	15843-16569; 1-676	16111 16223 16290 16291 16319 16362 64 73 146 153 207 235 263 315+C 489 523-524d 663
FLP 036	A2ah1	15843-16569; 1-676	16094 16097 16098 16111 16223 16290 16319 16362 64 73 146 153 235 263 309+C 315+C 512T 523-524d 663
FLP 215	A2ah1	15843-16569; 1-676	16094 16097 16098 16111 16223 16290 16319 16362 64 73 146 153 235 263 309+C 315+C 512T 523-524d 663
FOR 057	A2be1a1	15837-16569; 1-676	16126 16223 16278 16290 16319 16362 59 64 73 146 152 153 199 235 263 297 309+C 315+C 523-524d 573+C 663
FLP 130	A2be1a1	15843-16569; 1-676	16126 16223 16278 16290 16319 16362 59 64 73 146 152 153 235 263 297 309+C 315+C 523-524d 573+CC 663
FLP 197	A2be1a1b	15843-16569; 1-676	16126 16223 16278 16290 16319 16362 57G 59 64 73 146 152 153 235 263 297 315+C 523-524d 573+CCC 663
FLP 042	A2be1a1b	15843-16569; 1-575	16126 16223 16278 16290 16319 16362 57G 59 64 73 146 152 153 235 263 297 315+C 523-524d 573+CCCC 663
FLP 047	A2be1a1b	15843-16569; 1-676	16126 16223 16278 16290 16319 16362 57G 59 64 73 146 152 153 235 263 297 315+C 523-524d 573+CCC 663
FOR 030	A2be1a1b	15837-16569; 1-573	16126 16223 16278 16290 16319 16362 57G 59 64 73 146 152 153 235 263 297 315+C 523-524d 573+CCC 663
FLP 102	A2be1a	15843-16569; 1-676	16126 16223 16278 16290 16319 16362 59 64 73 146 152 235 263 297 309+CC 315+C 523-524d 663
FLP 205	A2be1a	15843-16569; 1-676	16126 16223 16278 16290 16319 16362 16519 59 64 73 146 152 235 263 297 309+C 315+C 523-524d 663
FLP 219	A2be1a1d	15843-16569; 1-676	16093 16126 16189 16193+C 16223 16278 16290 16319 16362 59 62T 64 73 146 152 153 235 263 297 309+C 315+C 523-524d 573+C 663
FOR 055	A2zh	15843-16569; 1-676	16111 16223 16290 16319 16362 16519 64 73 146 153 210 235 263 315+C 523-524d 663
FOR 071	A2ze	15837-16569; 1-676	16111 16223 16290 16319 16362 16422 16519 64 73 146 153 235 263 309+C 315+C 523-524d 663

ID	Haplogrupo ADNmt	Rango	Sitios polimórficos
FLP 178	A2zz	15843-16569; 1-676	16111 16223 16290 16319 16362 59 64 73 146 235 263 309+C 315+C 523-524d 663
FLP 159	A2yc	15843-16569; 1-676	16111 16126 16223 16259 16290 16319 16327 16362 16519 73 146 153 235 263 315+C 523-524d 663
FLP 160	A2yc	15843-16569; 1-676	16111 16126 16223 16259 16290 16319 16327 16362 16519 73 146 153 235 263 315+C 523-524d 663
FLP 082	A2yc	15843-16569; 1-676	16111 16126 16223 16259 16290 16319 16327 16362 16519 73 146 153 235 263 315+C 523-524d 663
FLP 154	A2bf	15843-16569; 1-676	16051 16111 16223 16290 16319 16362 57G 73 146 153 235 263 315+C 523-524d 663
FLP 072	A2zk3	15843-16569; 1-676	16111 16223 16266 16290 16319 16362 64 73 146 235 263 309+CC 315+C 523-524d 663
FLP 071	A2zk3	15843-16569; 1-676	16111 16223 16266 16290 16319 16362 16519 64 73 146 153 235 263 309+C 315+C 523-524d 663
FLP 037	A2zk3a	15843-16569; 1-676	16111 16223 16266 16290 16319 16325 16362 64 73 146 153 235 263 309+C 315+C 523-524d 663
FLP 181	A2zk3a	15843-16569; 1-676	16111 16223 16266 16290 16319 16325 16362 64 73 146 153 235 263 309+C 315+C 523-524d 663
FLP 192	A2zj	15843-16569; 1-676	16092 16111 16223 16290 16319 16362 16468 16519 64 73 146 153 235 263 309+C 315+C 338 523-524d 573+CC 663
FLP 203	A2zj	15843-16569; 1-676	16092 16111 16223 16290 16319 16362 16468 16519 64 73 146 153 235 263 309+C 315+C 338 523-524d 573+CC 663
FLP 066	A2zj	15843-16569; 1-676	16092 16111 16223 16290 16319 16362 16468 16519 64 73 146 153 185 235 263 309+C 315+C 338 523-524d 573+C 663
FLP 104	A2zj1a	15843-16569; 1-676	16092 16111 16223 16290 16319 16362 16468 16519 64 73 93 146 153 185 235 263 315+C 338 523-524d 573+C 663
FOR 003	A2zl	15837-16569; 1-676	16111 16223 16290 16319 16362 16519 73 143 146 153 235 263 309+CC 315+C 523-524d 663
FLP 231	A2zl	15843-16569; 1-588	16111 16223 16290 16319 16362 64 73 143 146 153 235 263 309+CC 315+C 523-524d 663
FOR 011	A2zl	15837-16569; 1-676	16111 16189 16193+C 16223 16290 16319 16362 64 73 143 146 153 204 235 263 315+C 523-524d 663
FOR 056	A2zb1	15837-16569; 1-676	16111 16223 16269 16290 16319 16360 16362 73 146 153 235 263 309+C 315+C 485 523-524d 663

ID	Haplogrupo ADNmt	Rango	Sitios polimórficos
FLP 156	A2zc1	15843-16569; 1-574	16111 16223 16290 16319 16362 16526 64 73 146 153 235 263 279 309+C 315+C 485 523-524d 573+CCCCC 663
FLP 012	A2*	15843-16569; 1-676	16038 16111 16129 16183C 16189 16193+C 16223 16290 16311 16319 16362 16381 16519 64 73 146 153 235 263 309+C 315+C 523-524d 593 663
FLP 164	A2*	15843-16569; 1-676	16111 16192 16223 16242 16290 16319 16362 16398 64 73 131 146 153 235 263 309+C 315+C 524+AC 663
FLP 119	A2*	15843-16569; 1-676	16111 16223 16290 16319 16362 16390 16519 64 73 146 153 235 263 309+CC 315+C 523-524d 663
FLP 202	A2*	15843-16569; 1-676	16111 16223 16290 16319 16362 16519 64 73 146 153 235 263 309+C 315+C 523-524d 663
FLP 031	A2*	15843-16569; 1-676	16111 16223 16290 16319 16362 73 146 153 227T 235 263 309+C 315+C 472 523-524d 663
FLP 086	A2*	15843-16569; 1-676	16111 16223 16290 16291 16319 16362 16399 73 146 153 235 263 309+C 315+C 523-524d 663

ID	Haplogrupo ADNmt	Rango	Sitios polimórficos
FLP 185	B2zc1a	15834-16569; 1-573	16182C 16183C 16189 16193+CC 16217 16519 73 263 309+CC 315+C 499 524+AC 573+CCCCC
FOR 017	B2o1a	15844-16569; 1-676	16092 16145 16182C 16183C 16189 16193+C 16217 16519 73 152 228 234 263 315+C 499
FOR 044	B2o1a	15837-16569; 1-676	16092 16145 16182C 16183C 16189 16193+C 16217 16519 73 152 228 234 263 315+C 499
FLP 087	B2o1a	15843-16569; 1-676	16092 16145 16182C 16183C 16189 16193+C 16217 16519 73 152 228 234 263 315+C 499
FLP 222	B2an	15837-16569; 1-676	16183C 16189 16193+C 16217 16241 16519 73 103 152 263 309+C 315+C 499
FLP 224	B2an	15837-16569; 1-676	16182C 16183C 16189 16193+C 16217 16241 16519 73 103 152 263 309+CC 315+C 499
FLP 230	B2an	15837-16569; 1-676	16183C 16189 16193+C 16217 16241 16519 73 103 152 263 297 309+CC 315+C 499
FLP 177	B2an	15837-16569; 1-676	16183C 16189 16193+C 16217 16241 16519 73 103 152 263 309+CC 315+C 499
FOR 038	B2an	15843-16569; 1-676	16182C 16183C 16189 16193+C 16217 16241 16519 73 152 263 309+CC 315+C 499 523-524d 533
FLP 085	B2an2	15837-16569; 1-676	16092 16182C 16183C 16189 16193+C 16217 16241 16304 16519 73 103 146 152 263 309+CC 315+C 499
FLP 076	B2an2	15838-16569; 1-676	16092 16182C 16183C 16189 16193+C 16217 16241 16304 16519 73 103 146 152 263 309+CC 315+C 499
FLP 003	B2an2	15834-16569; 1-676	16183C 16189 16193+C 16217 16241 16325 16519 73 103 146 152 195 214 263 309+CC 315+C 499
FLP 017	B2an2a	15844-16569; 1-676	16183C 16189 16193+C 16217 16519 73 103 146 152 263 309+CC 315+C 455+T 499
FOR 007	B2+16218	15837-16569; 1-676	16183C 16189 16193+CC 16203 16217 16218 16519 73 263 309+C 315+C 499
FLP 043	B2e1	15844-16569; 1-676	16183C 16189 16193+C 16217 16519 73 263 315+C 499 524+ACAC
FLP 049	B2e2a	15849-16569; 1-676	16183C 16189 16193+C 16217 16519 73 263 315+C 499
FLP 052	B2e2a	15843-16569; 1-676	16183C 16189 16193+C 16217 16519 73 263 309+C 315+C 499 524+ACAC
FLP 143	B2e2a	15834-16569; 1-676	16183C 16189 16193+C 16217 16519 73 263 272 309+CC 315+C 499 524+AC

ID	Haplogrupo ADNmt	Rango	Sitios polimórficos
FLP 077	B2e1	15837-16569; 1-676	16183C 16189 16193+C 16217 16519 73 263 309+C 315+C 499 524+ACAC
FLP 080	B2e2a	15843-16569; 1-676	16183C 16189 16193+C 16217 16519 73 263 309+C 315+C 499
FLP 035	B2e2	15834-16569; 1-676	16183C 16189 16193+C 16217 16519 73 263 309+C 315+C 499 524+AC
FLP 179	B2i1	15844-16569; 1-676	16179A 16183C 16189 16193+CC 16217 16311 16519 73 263 309+C 315+C 430 485 499
FLP 210	B2i1a1	15834-16569; 1-676	16179A 16183C 16189 16193+C 16217 16311 16519 34 73 263 309+CC 315+C 430 485 499 524+AC
FOR 018	B2h	15842-16569; 1-676	16183C 16189 16193+C 16217 16468 16519 73 263 309+CC 315+C 499 524+AC
FLP 125	B2h	15838-16569; 1-660	16183C 16189 16193+C 16217 16468 16519 73 263 309+CC 315+C 499 524+AC
FLP 137	B2h	15829-16569; 1-655	16183C 16189 16193+C 16217 16468 16519 73 263 309+CC 315+C 499 524+AC
FLP 053	B2b3a2	15844-16569; 1-676	16168 16182C 16183C 16189 16193+C 16217 16249 16312 16519 73 152 263 271 309+CC 315+C 499
FLP 120	B2b3a2	15830-16569; 1-676	16168 16182C 16183C 16189 16193+C 16217 16249 16312 16344 16519 73 152 263 271 315+C 499
FOR 062	B2*	15837-16569; 1-676	16092 16126 16182C 16183C 16189 16193+CC 16214 16355A 16519 73 152 309+TCC 315+CC 499
FOR 028	B2zf1	15837-16569; 1-676	16183C 16189 16193+CC 16217 16311 16319 16354 16399 16519 16524 73 263 309+CC 315+C 499
FLP 150	B2zf1	15838-16569; 1-676	16183C 16189 16193+C 16217 16311 16319 16354 16519 16524 73 263 315+C 499
FLP 212	B2zf1	15837-16569; 1-676	16183C 16189 16193+C 16217 16311 16319 16354 16519 16524 73 263 315+C 499 523-524d
FOR 012	B2za1a	15837-16569; 1-676	16182C 16183C 16189 16193+C 16217 16316 16519 73 151 260 263 309+CC 315+C 499 523-524d
FLP 127	B2za1a	15837-16569; 1-676	16182C 16183C 16189 16193+CCC 16217 16316 16519 57G 73 151 260 263 309+CC 315+C 499 523d
FLP 068	B2aj4a	15844-16569; 1-676	16183C 16188 16189 16217 16519 63 64 73 263 315+C 499
FLP 111	B2b*	15834-16569; 1-676	16182C 16183C 16186 16189 16217 16311 16519 73 263 315+C 499

ID	Haplogrupo ADNmt	Rango	Sitios polimórficos
FLP 226	B2b*	15834-16569; 1-676	16182C 16183C 16189 16193+C 16357 16519 73 263 309+C 315+C 499
FLP 214	B2b15	15837-16569; 1-676	16182C 16183C 16189 16193+C 16217 16519 73 143 146 165 263 315+C 499
FLP 207	B2+7055-247	15843-16569; 1-658	16182C 16183C 16189 16193+C 16217 16291 16519 73 247 263 309+CC 315+C 499
FLP 237	B2ze	15837-16569; 1-676	16145 16182C 16183C 16189 16193+C 16217 16294A 16326 16519 73 263 309+C 315+C 484 499
FOR 026	B2ze	15837-16569; 1-620	16145 16182C 16183C 16189 16193+C 16217 16294A 16519 73 263 309+C 315+C 484 499
FOR 035	B2zd	15837-16569; 1-676	16178 16183C 16189 16193+CCC 16217 16519 73 263 309+CC 499
FLP 116	B2zd	15837-16569; 1-676	16178 16183C 16189 16193+CC 16217 16519 73 263 309+CC 315+C 499
FLP 028	B2zd	15837-16569; 1-676	16178 16182C 16183C 16189 16193+CCC 16217 16519 73 263 309+CC 315+C 499
FOR 022	B2zd	15837-16569; 1-676	16178 16183C 16189 16193+CC 16217 16519 73 146 263 315+C 499
FLP 148	B2zd	15838-16569; 1-676	16129 16178 16183C 16189 16193+C 16217 16519 73 263 309+CC 315+C 499
FLP 041	B2zd	15834-16569; 1-676	16129 16178 16183C 16189 16193+CC 16217 16519 73 263 309+C 315+C 499
FLP 189	B2zd	15837-16569; 1-676	16145 16178 16183C 16189 16193+CC 16217 16519 73 263 309+CC 315+C 499
FLP 161	B2*	15836-16569; 1-676	16142 16183C 16189 16193+C 16217 16519 73 111 186 263 309+C 315+C 499 573+CC
FLP 225	B2a12	15837-16569; 1-676	16173 16182C 16183C 16189 16193+CC 16217 16223 16357 16519 73 263 309+CC 315+C 499
FOR 059	B2*	15848-16569; 1-676	16183C 16189 16193+C 16217 16519 73 196 263 309+C 315+C 499
FLP 118	B2*	15837-16569; 1-676	16092 16170 16183C 16189 16193+C 16217 16519 73 263 315+C 499
FOR 015	B2*	15837-16569; 1-676	16183C 16189 16193+C 16217 73 143 217 263 315+C 374 499

ID	Haplogrupo ADNmt	Rango	Sitios polimórficos
FOR 006	C1b18a1	15837-16569; 1-676	16166C 16221 16223 16298 16325 16327 73 249d 263 290-291d 309+C 315+C 489 493 523-524d
FOR 009	C1b18a1	15837-16569; 1-598	16166C 16221 16223 16298 16325 16327 73 249d 263 290-291d 309+C 315+C 489 493 523-524d
FOR 010	C1b*	15843-16569; 1-676	16176 16223 16298 16325 16327 16362 73 249d 263 290-291d 309+C 315+C 316 489 493 523-524d
FLP 153	C1b30a1	15837-16569; 1-676	16126 16223 16298 16325 16327 73 249d 263 290-291d 309+C 315+C 489 493 523-524d
FOR 047	C1b30a1	15843-16569; 1-676	16126 16223 16298 16325 16327 73 249d 263 290-291d 309+C 315+C 489 493 523-524d
FOR 040	C1b101	15843-16569; 1-676	16223 16298 16325 16327 16384 73 249d 263 290-291d 309+C 315+C 489 493 523-524d
FLP 054	C1b101	15837-16569; 1-676	16223 16298 16325 16327 16384 73 249d 263 290-291d 309+C 315+C 489 493 523-524d
FLP 121	C1b101	15837-16569; 1-676	16124 16223 16298 16325 16327 16384 73 249d 263 290-291d 309+C 315+C 489 493 523-524d
FOR 045	C1b16a1a	15837-16569; 1-676	16136 16223 16298 16325 16327 16519 73 92 98A 146 150 249d 263 290-291d 309+C 315+C 489 493 523-524d
FOR 060	C1b*	15837-16569; 1-676	16172 16223 16256 16298 16325 16327 16519 9 73 249d 263 290-291d 309+CC 315+C 489 493 523-524d
FLP 006	C1b102a1	15837-16569; 1-676	16093 16223 16298 16325 16327 16526 73 152 249d 263 290-291d 309+CC 315+C 489 493 523-524d
FLP 046	C1b18b	15837-16569; 1-676	16187 16223 16298 16325 16327 73 249d 263 290-291d 309+C 315+C 489 493 523-524d
FLP 078	C1b18a	15837-16569; 1-676	16223 16298 16325 16327 16519 73 249d 263 290-291d 309+C 315+C 489 493 523-524d
FLP 062	C1b18	15837-16569; 1-676	16136 16223 16298 16325 16327 73 249d 263 290-291d 315+C 489 493 523-524d
FOR 024	C1b17	15837-16569; 1-676	16223 16298 16325 16327 16519 73 150 207 249d 263 290-291d 309+C 315+C 489 493 523-524d
FLP 234	C1b16	15837-16569; 1-676	16223 16298 16325 16327 16519 73 146 249d 263 290-291d 309+C 315+C 489 493 523-524d
FLP 101	C1b16	15837-16569; 1-676	16223 16298 16325 16327 16519 73 146 249d 263 290-291d 309+C 315+C 489 493 523-524d
FLP 109	C1b16	15837-16569; 1-676	16223 16298 16325 16327 16519 73 146 249d 263 290-291d 309+C 315+C 489 493 523-524d

ID	Haplogrupo ADNmt	Rango	Sitios polimórficos
FLP 199	C1b16	15837-16569; 1-676	16223 16298 16311 16325 16327 73 146 249d 263 290-291d 315+C 489 493 523-524d
FLP 187	C1b19b	15837-16569; 1-676	16223 16266 16298 16325 16327 73 143 152 195 249d 263 290-291d 315+C 489 493 523-524d
FLP 141	C1b*	15837-16569; 1-676	16144 16223 16298 16325 16327 73 152 195 249d 263 290-291d 309+CC 315+C 489 493 523-524d
FLP 146	C1b*	15837-16569; 1-676	16144 16223 16298 16325 16327 73 152 195 249d 263 290-291d 309+CC 315+C 489 493 507A 523-524d
FLP 065	C1b2a1a	15837-16569; 1-676	16154 16223 16298 16325 16327 16519 73 249d 290-291d 309+CC 315+C 489 493 523-524d
FLP 108	C1b2a1a	15837-16569; 1-676	16154 16223 16298 16325 16327 16519 72 73 249d 290-291d 309+C 315+C 489 493 523-524d
FOR 054	C1b2a1	15829-16569; 1-676	16223 16298 16299 16325 16327 16519 73 249d 290-291d 315+C 489 493 523-524d
FLP 060	C1b*	15837-16569; 1-676	16221 16223 16298 16325 16327 73 249d 263 290-291d 309+CC 315+C 489 493 523-524d
FLP 026	C1c100a2	15837-16569; 1-676	16071 16092 16223 16298 16311 16325 16327 16368 16390 73 151 152 249d 263 290-291d 309+C 315+C 489

ID	Haplogrupo ADNmt	Rango	Sitios polimórficos
FOR 008	C1d1b1a	15837-16569; 1-676	16051 16223 16298 16325 16327 73 194 195 204 249d 263 290-291d 309+C 315+C 489 499 523-524d
FOR 025	C1d1b1 +16129	15837-16569; 1-598	16051 16129 16223 16298 16325 16327 73 194 195 249d 263 290-291d 309+C 315+C 489 523-524d
FLP 170	C1d1b1a	15837-16569; 1-598	16051 16223 16298 16325 16327 73 194 195 204 249d 263 290-291d 309+C 315+C 489 523-524d
FLP 027	C1d1b1a	15837-16569; 1-598	16051 16223 16298 16325 16327 73 194 195 204 249d 263 290-291d 309+C 315+C 489 523-524d
FLP 211	C1d1b2	15837-16569; 1-598	16051 16223 16259 16271 16298 16311 16325 16327 73 194 195 249d 263 290-291d 315+C 489 523-524d
FOR 016	C1d1aa2a2	15837-16569; 1-676	16051 16093 16223 16298 16325 16327 73 194 195 249d 263 290-291d 315+C 489 523-524d
FLP 112	C1d1aa2a2	15837-16569; 1-676	16051 16093 16223 16298 16325 16327 73 152 194 195 204 249d 263 290-291d 309+C 315+C 489 523-524d
FLP 129	C1d ⁺⁴⁶⁸⁸⁻ 5656-16247-195	15837-16569; 1-676	16051 16223 16247 16298 16325 16327 16519 73 194 195 249d 263 290-291d 309+C 315+C 489
FLP 022	C1d1d	15837-16569; 1-676	16051 16172 16223 16298 16325 16327 73 194 249d 263 290-291d 309+C 315+C 489 523-524d
FLP 048	C1d1d	15837-16569; 1-676	16051 16172 16223 16298 16325 16327 73 194 249d 263 290-291d 309+CC 315+C 489 523-524d
FLP 074	C1d1d	15837-16569; 1-676	16051 16172 16223 16298 16325 16327 73 194 249d 263 290-291d 309+C 315+C 489 523-524d
FLP 103	C1d1d	15837-16569; 1-573	16051 16172 16223 16298 16325 16327 73 194 249d 263 290-291d 309+C 315+C 489 523-524d 573+C
FLP 135	C1d1d	15837-16569; 1-676	16051 16172 16223 16298 16325 16327 73 194 249d 263 290-291d 309+C 315+C 489 523-524d
FLP 145	C1d1d	15837-16569; 1-676	16051 16172 16223 16298 16325 16327 73 194 249d 263 290-291d 295 297C 299A 308d 309d 310 315+C 424d 489 523-524d
FLP 157	C1d1d	15837-16569; 1-676	16051 16172 16223 16298 16325 16327 73 194 249d 263 290-291d 309+C 315+C 489 523-524d
FLP 168	C1d1d	15837-16569; 1-676	16051 16172 16223 16298 16325 16327 73 194 249d 263 290-291d 295 297C 308d 309d 310 315+C 424d 489 523-524d
FLP 175	C1d1d	15837-16569; 1-676	16051 16172 16223 16298 16325 16327 73 194 249d 263 290-291d 309+CC 315+C 489 523-524d
FLP 206	C1d1d	15837-16569; 1-676	16051 16172 16223 16298 16325 16327 73 194 249d 263 290-291d 309+C 315+C 489 523-524d
FLP 063	C1d1ab1	15837-16569; 1-676	16051 16223 16298 16325 16327 16362 5C 73 194 249d 263 290-291d 315+C 318 489 523-524d

ID	Haplogrupo ADNmt	Rango	Sitios polimórficos
FLP 201	D1e1	15837-16569; 1-676	16223 16325 16362 73 146 263 273 309+CC 315+C 489
FLP 227	D1e1	15837-16569; 1-676	16223 16325 16362 73 146 263 273 309+C 315+C 489 516
FLP 235	D1e1	15837-16569; 1-676	16223 16325 16362 73 146 263 273 309+C 315+C 489
FLP 019	D1e1	15837-16569; 1-676	16223 16325 16362 73 146 263 273 309+C 315+C 489
FOR 061	D1e1	15837-16569; 1-676	16223 16325 16362 73 146 263 273 309+CC 315+C 489
FLP 045	D1f1	15837-16569; 1-676	16142 16179 16223 16295 16325 16362 16497 16519 73 152Y 263 309+C 315+C 489 524+AC
FLP 136	D1f1	15837-16569; 1-676	16142 16179 16223 16295 16325 16362 16497 16519 519 73 263 309+C 315+C 489 524+AC
FLP 138	D1f1	15837-16569; 1-676	16142 16179 16223 16295 16325 16362 16497 16519 73 263 309+CC 315+C 489 524+AC
FLP 167	D1f1	15837-16569; 1-676	16142 16179 16223 16295 16325 16362 16497 16519 73 263 309+C 315+C 489 524+AC
FLP 182	D1f1	15837-16569; 1-676	16142 16179 16223 16295 16325 16362 16497 16519 73 263 309+CC 315+C 489 524+AC
FLP 133	D1f1	15837-16569; 1-676	16179 16223 16295 16325 16362 16497 16519 73 263 309+C 315+C 489 524+AC
FLP 174	D1f1	15837-16569; 1-676	16179 16223 16295 16325 16362 16497 16519 73 263 309+CC 315+C 489 524+AC
FLP 186	D1f1	15837-16569; 1-676	16142 16145 16179 16223 16295 16325 16362 16497 16519 73 263 309+C 315+C 489 524+AC
FLP 191	D1f1	15837-16569; 1-676	16142 16145 16179 16223 16295 16325 16362 16497 16519 73 263 309+CC 315+C 489 524+AC
FLP 059	D1f1	15837-16569; 1-676	16142 16145 16179 16223 16295 16325 16362 16497 16519 73 263 309+C 315+C 489
FLP 050	D1j1a	15837-16569; 1-676	16223 16242 16311 16325 16362 73 152 263 309+C 315+C 489
FLP 232	D1j1a1	15837-16569; 1-676	16172 16223 16242 16311 16325 16362 73 152 263 315+C 489
FLP 221	D1j1a1	15837-16569; 1-676	16223 16242 16311 16325 16362 73 152 263 309+C 315+C 489 524+AC

ID	Haplogrupo ADNmt	Rango	Sitios polimórficos
FOR 005	D1*	15837-16569; 1-676	16223 16260 16325 16362 16390 10A 55+T 57 59 73 263 309+C 315+C 489
FLP 013	D1*	15837-16569; 1-676	16124 16140 16223 16325 16357 16362 66T 73 263 309+C 315+C 489
FOR 053	D1*	15837-16569; 1-676	16223 16325 16362 16365 73 207 263 309+C 315+C 489
FLP 034	D1*	15837-16569; 1-676	16179 16223 16325 16362 58A 73 152 263 315+C 489
FOR 013	D1*	15843-16569; 1-676	16189 16193+C 16223 16325 16362 16519 73 146 207 263 315+C 489 524+ACAC
FLP 083	D4h3a92	15837-16569; 1-676	16114A 16223 16241 16288 16301 16320 16342 16362 73 152 263 309+C 315+C 489 524+AC
FLP 124	D4h3a91	15837-16569; 1-676	16223 16241 16301 16342 16362 16519 73 106-111d 152 263 309+C 315+C 489 523-524d
FLP055	L3b	15843-16569; 1-684	15944d 093 124 223 278 362 519 73 263 315+C 523-524d
FLP094	L1c2	15843-16569; 1-684	093 129 181 187 189 223 265C 278 286G 294 311 360 519 527 73 151 152 182 186A 189C 195 198 247 263 297 315+C 316 524+ACAC
FLP095	L2a1l3	15843-16569; 1-620	129 189 192 223 278 294 309 390 73 143 146 152 189 195 263 315+C
FLP183	L3e1d	15843-16569; 1-684	15942 15968 176 223 256 327 73 150 152 189 200 263 315+C
FLP204	L1b2'3	15843-16569; 1-673	126 187 189 223 264 270 278 311 519 73 152 182 185T 195 247 263 315+C 357 523-524d
FLP009	L3e1a1	15843-16569; 1-673	15942 185 223 311 327 73 150 185 189 200 263 309+C 315+C
FLP091	L2a1	15843-16569; 1-673	037 129 223 278 294 309 390 73 143 146 152 195 263 315+C
CBS078	L3e1a	15843-16569; 1-673	15924 15942 185 189 223 327 519 73 150 189 200 263 315+C

Anexo 3.2. Linajes de origen monofilético en 20 poblaciones de Argentina (número de individuos identificados).

Hg	Clado	LQ	MAI	JUJ	TG	BEL	SMA	TMN	SDE	LRI	CHE	SNJ	CAL	TUL	MZA	MLG	GUA	LPZ	CC	REQ	FOR	TOT	
A2	A2 ₍₆₄₎ +12634/1766-2905-3579-3960-9854-12180T-14587	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	0	6	
	A2zb1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	
	A2zc1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	
	A2ze	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	
	A2+3435-14178-14755-14861	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	4	
	A2 ₍₆₄₎ +153@-4158-9878-6497/13659/	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	0	1	5	
	A2 ₍₆₄₎ +7981-13788A-15262	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	0	6	
	A2aa	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	5	
	A2ah	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	4	
	A2as1 (<i>sensu</i> Brandini et al., 2018)	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
	A2 ₍₆₄₎ +3083-15734-16422-8276d	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	3	
	A2bc_A2 ₍₆₄₎ +150-9039	0	0	0	0	0	0	0	0	0	1	1	2	0	8	3	2	0	0	0	0	0	17
	A2bb_A2 ₍₆₄₎ +150-3504	0	0	0	0	1	0	2	2	1	2	3	3	3	0	0	1	0	0	0	0	0	18
	A2az_A2 ₍₆₄₎ +9899	2	3	0	1	1	2	3	4	1	1	4	1	3	2	0	1	0	3	1	0	33	
	A2 ₍₆₄₎ +7202	0	0	0	0	3	3	1	0	0	0	4	0	0	0	0	0	0	0	0	0	0	11
	A2ax_A2+5147	0	0	0	0	2	0	0	0	2	4	1	0	1	3	1	0	0	1	0	0	15	
	A2+8843-9482-14040-14502-573+C/16104	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	2
	A2 ₍₆₄₎ +739-9210-15300-210	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	6	0	4	1	1	15	
	A2 ₍₆₄₎ +200-251-@523-524d	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
	A2 ₍₆₄₎ +239	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
A2ba_A2 ₍₆₄₎ +16129-195	0	0	0	0	0	0	0	1	1	3	1	0	1	0	3	0	0	0	0	0	0	10	

Hg	Clado	LQ	MAI	JUJ	TG	BEL	SMA	TMN	SDE	LRI	CHE	SNJ	CAL	TUL	MZA	MLG	GUA	LPZ	CC	REQ	FOR	TOT
A2	A2 ₍₆₄₎ +@663-16209-@16362-238	1	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	4
	A2bf_A2+16051	1	0	2	4	0	0	0	1	0	0	0	0	0	0	0	2	3	0	3	1	17
	A2+16512-16547-16551iG+@235	1	2	0	2	0	0	0	0	0	0	1	0	0	1	0	1	2	4	2	0	16
	A2+@16111-16234G-16256	4	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
	A2 ₍₆₄₎ +152-194	2	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
	A2 ₍₆₄₎ +16189-16209-@153	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2
	A2 ₍₆₄₎ +16192-16209	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	3
	A2 ₍₆₄₎ +16156-16189-16263-16438-573A	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	2
	A2+16266-16437-356iC-499	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	A2 ₍₆₄₎ +16287	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	A2 ₍₆₄₎ +16111@-182-194-195-235@	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
	A2au_A2 ₍₆₄₎ +16291-16465-10-146@	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	A2ab	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	3
	A2 ₍₆₄₎ +16189-16356	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2
	A2 ₍₆₄₎ +16095	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	A2 ₍₆₄₎ +16294	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	2
	A2 ₍₆₄₎ +1806-7202-10203-16189	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2
	A2 ₍₆₄₎ +338-3438-16468	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0	1	3	4	11
	A2+16266/10237-12858/14364/7853-16325	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	5
	A2 ₍₆₄₎ +2850-4216-15968-143	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	2	3	8
A2 ₍₆₄₎ +15884-14491	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	

Hg	Clado	LQ	MAI	JUJ	TG	BEL	SMA	TMN	SDE	LRI	CHE	SNJ	CAL	TUL	MZA	MLG	GUA	LPZ	CC	REQ	FOR	TOT
A2	A2 ₍₆₄₎ +3397-13926	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	3
	A2be (<i>sensu</i> Figueiro et al., 2022)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	12	2	9	27
	A2+16209	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	3
	A2 ₍₆₄₎ +16356-214/16362@-384-492	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	1	0	4
	A2+(9096-14127C-15951-16126-16259)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3
	A2(64)+200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
	A2+(2222-636-10604-12630-14569-16092-45)	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
	A2 sin asignar	7	2	2	7	0	2	1	0	0	0	0	2	1	2	2	3	1	2	1	7	42
	Total A2	29	22	7	23	8	9	7	9	7	12	18	8	18	13	14	31	10	45	23	40	353
B2	B2b+16156	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	3
	B2b2b (<i>sensu</i> García et al., 2021)	2	3	1	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	9
	B2b3a2 (<i>sensu</i> García et al., 2018)	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	6	2	2	14
	B2b15 (<i>sensu</i> García et al., 2021)	0	0	0	1	0	1	1	0	1	1	2	4	4	1	0	1	0	0	0	1	18
	B2b50_B2b+16309-143	1	6	0	0	1	3	0	0	0	0	0	0	0	0	3	0	0	0	0	0	14
	B2b+8185-11383-15968-309d/16356-16362	0	1	0	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	4
	B2b12a (<i>sensu</i> Brandini et al., 2018)	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
	B2b9c (<i>sensu</i> Brandini et al., 2018)	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	B2b52_B2b+9804/16390	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	5	3	0	10
	B2b*	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	B2b*	3	2	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	9
	B2e	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	2	6	8	1	7	26
	B2h	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	3	5
B2i1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	1	2	3	2	11	

Hg	Clado	LQ	MAI	JUJ	TG	BEL	SMA	TMN	SDE	LRI	CHE	SNJ	CAL	TUL	MZA	MLG	GUA	LPZ	CC	REQ	FOR	TOT
B2	B2i2	0	0	0	2	0	0	0	0	0	0	4	7	1	1	16	0	0	0	0	0	31
	B2ab (<i>sensu</i> Brandini et al., 2018)	9	6	0	8	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25
	B2am (<i>sensu</i> García et al., 2021)	2	1	2	2	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	11
	B2ai (<i>sensu</i> Gómez-Carballa et al., 2018)	2	3	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	7
	B2y2_B2+709+16261+16319	10	0	1	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0	14
	B2ag (<i>sensu</i> Brandini et al., 2018)	17	19	3	2	20	16	0	1	0	0	0	0	0	1	0	0	0	0	0	0	79
	B2aj (<i>sensu</i> Gómez-Carballa et al., 2018)	30	21	3	5	1	6	1	1	1	0	1	0	1	2	3	0	0	0	0	1	77
	B2+3357/16437	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
	B2ah (<i>sensu</i> Brandini et al., 2018)	1	1	1	0	0	0	0	0	0	3	0	0	0	0	1	0	0	0	0	0	7
	B2+(12192-16390/207)	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
	B2aa1 (<i>sensu</i> Brandini et al., 2018)	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	B2ak1 (<i>sensu</i> García et al., 2021)	2	0	0	2	2	6	1	1	5	0	7	0	2	2	2	1	0	0	0	0	33
	B2+14470	5	9	0	4	0	0	0	0	0	0	0	0	2	1	0	0	0	1	0	1	23
	B2an (<i>sensu</i> Figueiro et al., 2022)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	1	5	1	9	24
	B2+(7055-247)	0	0	0	8	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	1	12
	B2za_B2+(3372-16316-151)	5	0	0	5	0	0	0	0	0	0	1	0	0	0	0	0	1	5	0	2	19
	B2+16258-16357-103-310	2	2	1	4	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	10
	B2+16357-195	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	2
	B2+16291-16311-16356-16362-103-217-228-310-514	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	B2+143-7202-7598-16289	1	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
B2+9452-13135/197-207-292	0	0	0	4	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	6	
B2+16092-16126-16214-16355A-152-263@	0	0	1	1	1	1	1	0	0	0	3	1	1	0	2	0	0	0	0	1	13	
B2+(12397 16390/16092 16342)	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	3	

Hg	Clado	LQ	MAI	JUJ	TG	BEL	SMA	TMN	SDE	LRI	CHE	SNJ	CAL	TUL	MZA	MLG	GUA	LPZ	CC	REQ	FOR	TOT
B2	B2 sin asignar	2	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
	B2+16256-16288-185	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
	B2+60iT-513	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
	B2+479-480	2	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
	B2+6956	1	1	1	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
	B2+16359-103-146-151	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
	B2zb_B2+14791-93-195	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	2	1	2	3	0	12
	B2zc_B2+15884	0	2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	8
	B2+16086-16220-16294-56d-58-59G-61iT-64-65G-103-200	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	B2+228T	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	B2+93-195	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2
	B2z2 (<i>sensu</i> García et al., 2021)	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
	B2zd_B2+16178	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	13	8	3	3	7	36
	B2+16142-111(non-11590)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	1	3
	B2al (<i>sensu</i> García et al., 2021)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	2	1	7
	B2+12771	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	3
	B2o1a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3
	B2zfi_B2+3615-4232-16311-16319-16354-16524	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3
	B2+16391-60+T-195-309d	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
	B2+54	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	B2+152	1	3	0	0	1	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	8
	B2 sin asignar	4	5	1	10	0	1	1	0	0	0	2	0	0	0	0	0	3	4	2	3	36
Total B2		115	113	16	78	29	42	12	9	13	2	23	14	16	9	30	36	27	45	23	53	705

Hg	Clado	LQ	MAI	JUJ	TG	BEL	SMA	TMN	SDE	LRI	CHE	SNJ	CAL	TUL	MZA	MLG	GUA	LPZ	CC	REQ	FOR	TOT
C1b	C1b3a (<i>sensu</i> García et al., 2021)	0	2	0	0	1	8	1	0	0	2	10	13	1	0	1	1	0	0	0	0	40
	C1b13	0	1	0	0	0	0	0	0	1	1	1	6	3	3	28	0	0	0	0	0	44
	C1b+16166C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	2
	C1b21a (<i>sensu</i> Brandini et al., 2018)	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	C1b30 (<i>sensu</i> Figueiro et al., 2022)	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	4	4	11	2	2	26
	C1b26a (<i>sensu</i> Brandini et al., 2018)	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	3
	C1b+11084	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
	C1b+214-5201A-16145-16192	2	1	0	0	0	0	0	0	1	2	0	0	0	0	0	2	0	0	0	0	8
	C1b19b (<i>sensu</i> García et al., 2021)	0	0	1	0	1	3	4	6	0	0	1	0	1	0	0	0	1	0	1	1	20
	C1b+204-207	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	3
	C1b16 (<i>sensu</i> García et al., 2021)	0	0	2	10	3	2	16	14	6	5	9	0	1	3	0	3	1	1	2	5	83
	C1b18 (<i>sensu</i> García et al., 2021)	1	0	0	0	0	0	0	0	2	2	0	0	2	1	2	4	3	4	3	5	29
	C1b+16093-16129-16300-16362	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	0	0	0	0	0	5
	C1b17 (<i>sensu</i> García et al., 2021)	0	0	0	0	0	0	1	1	0	2	0	0	0	0	0	1	0	0	0	1	6
	C1b+16189-16311	1	2	0	0	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8
	C1b+16390-150-194	0	0	0	1	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	4
	C1b+15904-16297-16362-228T	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	3
	C1b+16157-16209	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	C1b+16129	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3
C1b+16192-16144-16291-16362	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	2	

Hg	Clado	LQ	MAI	JUJ	TG	BEL	SMA	TMN	SDE	LRI	CHE	SNJ	CAL	TUL	MZA	MLG	GUA	LPZ	CC	REQ	FOR	TOT	
C1b	C1b100_C1b+16192-10043/16356-207-679A-3369-6293	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	2	0	0	5	
	C1b101_C1b+2404-8772-13827-16384	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	3	10	
	C1b2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	1	3	7	
	C1b102_C1b+13212-16526/152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	1	2	1	9	
	C1b+4973-5978-10143-16172-16256	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	1	5	
	C1b103_C1b+16292-16362-960+C-8078	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	0	0	4	
	C1b+16327@-16465-361+A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	
	C1b+16144-152-195	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	
	C1b+316-16176-16362	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	
	C1b sin asignar	3	4	2	4	0	1	3	2	1	0	0	2	3	2	1	2	2	2	2	0	1	35
	Total C1b	8	13	5	18	8	20	27	26	12	14	24	22	15	11	39	26	13	33	15	26	375	
C1c	C1c10 (<i>sensu</i> García et al., 2021)	0	0	0	1	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	4	
	C1c9a (<i>sensu</i> García et al., 2021)	0	1	0	6	1	3	0	0	0	0	0	0	0	0	0	0	0	1	1	0	13	
	C1c100_C1c+152-4216-16311-16368-16390-151	0	0	0	2	0	0	1	2	0	0	0	0	3	0	1	0	0	0	0	1	10	
	C1c+16311-215-228-294	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	
	C1c101_C1c+16189-195	2	2	0	1	0	0	1	3	0	0	0	0	0	0	0	0	0	0	1	0	10	
	C1c+16298@-16304-16400	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	
	C1c+5562-153	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	4	
	C1c nodal	1	0	0	2	0	0	0	0	1	0	0	0	0	0	2	1	1	2	2	0	12	
	Total C1c	4	5	0	12	5	3	2	5	1	0	2	0	4	0	3	1	1	4	7	1	60	

Hg	Clado	LQ	MAI	JUJ	TG	BEL	SMA	TMN	SDE	LRI	CHE	SNJ	CAL	TUL	MZA	MLG	GUA	LPZ	CC	REQ	FOR	TOT	
C4c	Total C4c	0	0	0	0	0	0	0	0	0	2	2	1	1	0	0	0	0	0	0	0	6	
C1d	C1d+(15924-16051-16320-16344-152)	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	
	C1d+(4688-5656-16247-195)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	
	C1d1+(16086-55iT-57-59)	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	C1d1ab1_C1d1+(15728/16362-5C)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	C1d1aa2a2_C1d1+12501/11090T-195/12879-16093/	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2	3
	C1d1d	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2	0	4	1	10	18	
	C1d1ac_C1d1+150-3537-6915-9347-10754-13015-15886-16209	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	2
	<i>C1d1g</i> (sensu Figueiro et al., 2022)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
	<i>C1d1f</i> (sensu García et al., 2021)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	3
	C1d1+(16249-199)	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	2
	C1d1e (sensu García et al., 2021)	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	4
	C1d1b	1	1	2	4	0	8	7	12	7	0	7	3	1	2	0	2	0	1	0	5	63	
Total C1d	2	2	2	4	0	9	8	13	7	0	9	4	2	3	3	6	1	6	1	19	101		
D1	D1f4 García 2021	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	4	
	D1f1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	10	16		
	D1g (D1g* o D1g3)	0	0	0	1	0	0	0	0	0	0	2	0	0	1	3	0	0	0	0	0	7	
	D1g+16189	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	4	
	D1+2885	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	3	
	D1g*	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
	D1g1b	0	0	0	0	0	0	0	0	0	0	1	0	0	0	3	0	0	0	0	0	4	
	D1g5	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	0	0	0	1	0	6	

Hg	Clado	LQ	MAI	JUJ	TG	BEL	SMA	TMN	SDE	LRI	CHE	SNJ	CAL	TUL	MZA	MLG	GUA	LPZ	CC	REQ	FOR	TOT	
D1	D1g5a1	0	0	0	0	0	0	1	0	0	0	2	0	3	1	1	0	0	0	0	0	8	
	D1g5a	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	0	0	0	0	4	
	D1g2a	0	0	0	0	0	0	0	0	0	0	0	2	0	2	3	0	0	0	0	0	7	
	D1g4	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	2	
	D1g4a	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	2	
	D1g6	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	3	
	D1u (<i>sensu</i> García et al., 2021)	6	3	0	2	0	3	0	0	0	0	0	1	0	0	2	0	0	0	0	0	17	
	D1zb1_D1+16174-16263-16361-16390-146-207	0	0	0	7	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8
	D1+16293-97-106-111d	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3	
	D1zd_D1+@16325	0	0	0	8	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	11	
	D1+16190	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	
	D1-16311-152	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	3	
	D1w García 2021	0	0	0	0	0	0	2	3	0	0	0	0	0	0	0	2	0	0	1	0	8	
	D1zg_ D1+16311-152/16400-207 (D1j'w <i>sensu</i> García et al., 2021)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	3	
	D1j	0	1	3	6	8	15	19	14	35	13	11	5	3	6	4	5	1	0	0	3	152	
	D1zh1_D1+16189-195-297-16342	0	0	0	0	0	0	2	1	0	1	0	0	0	0	0	0	0	0	0	0	4	
	D1v_D1+998/16126-10-55-56-64-279	0	0	0	0	0	0	0	0	0	0	1	3	0	0	1	2	0	0	0	0	7	
	D1v_D1+998/152-234	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
	D1v_D1+998/150	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	
	D1za_D1+6305T-13989-146/3591-4883-7444C-11440-143	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	3	

Hg	Clado	LQ	MAI	JUJ	TG	BEL	SMA	TMN	SDE	LRI	CHE	SNJ	CAL	TUL	MZA	MLG	GUA	LPZ	CC	REQ	FOR	TOT
D1	D1+6480-11830-12200-15924-16189-146	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
	D1a2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
	D1e	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	6
	D1zc_D1+16187A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	4
	D1x1a_D1+11150/15106-210/16239-228	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2
	D1ze2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	2
	D1a	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2
	D1 sin asignar	1	1	1	5	0	0	0	1	0	0	1	0	0	0	2	1	1	1	0	5	20
	Total D1	10	7	4	31	11	22	27	20	39	14	20	15	7	17	28	15	4	7	9	28	335
D4h3a		0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
		1	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
	D4h3a5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	2
	D4h3a11	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	D4h3a90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
	D4h3a91	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	D4h3a92	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
		0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
		0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
		0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Total D4h3a	1	0	0	3	0	6	0	0	0	0	1	1	0	0	2	0	0	0	1	2	17	
Total linajes nativos	169	162	34	169	61	111	83	82	79	44	99	65	63	53	119	115	56	140	79	169	1952	

Anexo 3.3. Asociación entre haplotipos de región control, PNS en regiones codificantes, y clados dentro de A2.

Las secuencias de región control en fondo gris son haplotipos nodales de A2 o A2(64). Tanto en las secuencias de región control como en los fragmentos de región codificante se destacan en negrita los PNS derivados. n.d. = no disponible

Muestra	Región Control extendida		FR2 (F2480-R4860)		Fragmento FR7	Fragmento FR7	FR7 (Celemics)	Fragmento FR6	Clado
	Región 5'	Región 3'	rango	PNS	rango 13700~14841	rango 14030-14841	rango 13845-16031	rango 12494-13557	
CC 051	111 223 290 319 362	64 73 146 153 235 263 309+CC 315+C 523dd	2550-3610	2706 2905 3579	n.d.	n.d.	n.d.	n.d.	A2za1a
CC 110	111 223 290 319 362	64 73 146 153 235 263 315+C 523dd	2550-3610	2706 2905 3579	n.d.	n.d.	n.d.	n.d.	A2za1a
CC 153	111 223 290 319 362	64 73 146 153 235 263 315+C 523dd	2550-3610	2706 2905 3579	n.d.	n.d.	n.d.	n.d.	A2za1a
REQ 076	111 223 290 319 362	64 73 146 153 235 263 315+C 523dd	2550-3610	2706 2905 3579	n.d.	n.d.	n.d.	n.d.	A2za1a
REQ 117	111 223 290 319 362	64 73 146 153 235 263 315+C 523dd	2550-3610	2706 2905 3579	n.d.	n.d.	n.d.	n.d.	A2za1a
REQ 136	111 223 290 319 362	64 73 146 153 235 263 315+C 523dd	2550-3610	2706 2905 3579	n.d.	n.d.	n.d.	n.d.	A2za1a
CC 090	111 223 290 319 362	73 146 153 235 263 309+C 315+C 523dd	2550-3565	2706 3435	n.d.	n.d.	n.d.	n.d.	A2ya
REQ 013	111 223 290 319 362	73 146 153 235 263 309+C 315+C 523dd	2550-3565	2706 3435	n.d.	n.d.	n.d.	n.d.	A2ya
TG 60	104 111 290 319 362	73 146 153 235 263 309+C 315+C 523dd		n.d.	n.d.	14178 14755 14766 14861	n.d.	n.d.	A2ya
CC 017	129 223 290 319 362	59 62 64 73 146 152 153 235 263 309+CC 315+C 523dd	2550-4372	2706 3612 4248	n.d.	n.d.	n.d.	n.d.	

Muestra	Región Control extendida		FR2 (F2480-R4860)		Fragmento FR7	Fragmento FR7	FR7 (Celeemics)	Fragmento FR6	Clado
	Región 5'	Región 3'	rango	PNS	rango 13700~14841	rango 14030-14841	rango 13845-16031	rango 12494-13557	
FOR056	111 223 269 290 319 360 362	73 146 153 235 263 309+C 315+C 485 523-524d	2550- 4372	2706 3693 4248	n.d.	n.d.	n.d.	n.d.	A2zb1
FLP156	111 223 290 319 362 526	64 73 146 153 235 263 279 309+C 315+C 523-524d 573+CCCC	2550- 4372	2706 3834 4248	n.d.	n.d.	n.d.	n.d.	A2zc1
FLP031	111 223 290 319 362	73 146 153 227T 235 263 309+C 315+C 472 523-524d 663	2550- 4372	2706 3866 4248	n.d.	n.d.	n.d.	n.d.	
LPZ SLV 03	111 223 290 319 362 519	73 146 235 263 309+CC 315+C 523dd	2550- 4372	2706 4002 4248	n.d.	n.d.	n.d.	n.d.	
CC 068	111 223 274 290 319 362	64 73 146 235 263 315+C 523dd	2550- 4372	2706 4158 4248	n.d.	n.d.	n.d.	n.d.	A2zz
FLP178	111 223 290 319 362	59 64 73 146 235 263 309+C 315+C 523-524d	2550- 4372	2706 4158 4248	n.d.	n.d.	n.d.	n.d.	A2zz
FLP012	15908 038 111 129 183C 189 193+C 223 290 311 319 362 381 519	64 73 146 153 235 263 309+C 315+C 523-524d [593]	2550- 4372	2706 4248	n.d.	n.d.	n.d.	n.d.	
REQ A04	111 129R 189 223 290 319 362	64 73 146 153 235 263 315+C 485 523dd	2550- 4372	2706 4248	n.d.	n.d.	14766 15326	n.d.	
FLP119	[15884] 111 223 290 319 362 390 519	64 73 146 153 235 263 309+CC 315+C 523-524d 663	2550- 4843	2706 4248 4769 4824	n.d.	n.d.	14491 14766 15326 15884	n.d.	
FLP202	[15884] 111 223 290 319 362 519	64 73 146 153 235 263 309+C 315+C 523-524d 663		n.d.	n.d.	n.d.	14491 14766 15326 15884	n.d.	
LPZ 107	111 223 290 319 362	64 73 146 153 235 263 309+C 315+C 523dd	2550- 4843	2706 4248 4769 4824	n.d.	n.d.	14766 15262 15326	n.d.	A2zd
CC OÑA 04	111 223 290 319 362	64 73 146 153 235 263 309+C 315+C 523dd	2550- 4843	2706 4248 4769 4824	n.d.	n.d.	14766 15262 15326	n.d.	A2zd

Muestra	Región Control extendida		FR2 (F2480-R4860)		Fragmento FR7	Fragmento FR7	FR7 (Celemics)	Fragmento FR6	Clado
	Región 5'	Región 3'	rango	PNS	rango 13700~14841	rango 14030-14841	rango 13845-16031	rango 12494-13557	
CC 021	111 223 290 319 362	64 73 146 153 235 263 309+C 315+C 523dd	2550-4843	2706 4248 4769 4824	n.d.	n.d.	14766 15262 15326	n.d.	A2zd
REQ 073	111 223 290 319 362 519	64 73 146 153 235 263 309+C 315+C 523dd	2550-4843	2706 4248 4769 4824	n.d.	n.d.	14766 15262 15326	n.d.	A2zd
REQ 074	092 111 223 290 319 362	64 73 146 153 235 263 309+C 315+C 523dd	2550-4372	2706 4248	n.d.	n.d.	14766 15262 15326	n.d.	A2zd
REQ 60	111 223 290 319 362 519	64 73 146 153 235 263 309+C 315+C 523dd	2550-4843	2706 4248 4769 4824	n.d.	n.d.	14766 15262 15326	n.d.	A2zd
J036	111 223 290 319 362	64 73 146 153 235 263 315+C 523dd		n.d.	n.d.	n.d.	14587 14766 15326	n.d.	
CAL 31	111 223 290 319 362	64 73 146 153 235 263 309+C 315+C 523dd		n.d.	n.d.	n.d.	14766 15326 15401 15574	n.d.	.
TG 066	111 223 290 319 362	64 73 146 153 235 263 309+C 315+C		n.d.	n.d.	n.d.	14766 15326	n.d.	.
MLG 071	111 223 290 319 362	64 73 146 153 235 263 309+CC 315+C 523dd		n.d.	n.d.	n.d.	14766 15289 15326	n.d.	.
LQ 061	111 223 290 319 362 519	64 73 146 153 235 263 315+C 523dd		n.d.	n.d.	n.d.	14569 14766 15326	n.d.	.
MM 132	111 223 290 319 362 519	64 73 146 153 235 263 315+C 523dd		n.d.	n.d.	n.d.	14569 14766 15326	n.d.	.
TG BS 27	092 111 223 290 319 362	45 64 73 146 153 235 263 309+CC 315+C 523dd	2550-4372	2706 4248	n.d.	n.d.	14569 14766 15326	n.d.	.
GUA 155	111 223 290 319 362 519	64 73 146 152 153 235 263 309+C 315+C 523dd		n.d.	n.d.	n.d.	14766 15326	n.d.	.
GUA 117	111 223 290 319 362 519	64 73 146 153 178R 235 263 309+C 315+C 523dd		n.d.	n.d.	n.d.	14364 14766 15326 15603 15951	n.d.	

Muestra	Región Control extendida		FR2 (F2480-R4860)		Fragmento FR7	Fragmento FR7	FR7 (Celemics)	Fragmento FR6	Clado
	Región 5'	Región 3'	rango	PNS	rango 13700~14841	rango 14030-14841	rango 13845-16031	rango 12494-13557	
TG BS 5 sal	092 111 189 223 290 319 362	64 73 146 153 235 263 309+CC 315+C 523dd	2550-4372	2706 4248	n.d.	n.d.	14766 15326	n.d.	.
CC PER 02	048 111 171 223 290 319 362 519	64 73 146 153 235 263 309+CC 315+C 523dd	2550-4372	2706 4248	n.d.	n.d.	n.d.	n.d.	
FLP164	111 192 223 242 290 319 362 398	64 73 131 146 153 235 263 309+C 315+C 524+AC 663	2550-4372	2706 4248	n.d.	n.d.	n.d.	n.d.	
FOR071	111 223 290 319 362 422 519	64 73 146 153 235 263 309+C 315+C 523-524d	2679-3527	2706 3083	n.d.	n.d.	n.d.	n.d.	A2ze
GUA 089	111 223 290 311 319 362 422	64 73 146 150 153 235 263 315+C 523dd	2550-4843	2706 4248 4769 4824	n.d.	n.d.	n.d.	n.d.	A2zf
CC 134	111 223 290 319 362 519	64 73 146 152 153 235 263 309+C 315+C 523dd		n.d.	n.d.	n.d.	14233 14766 15326 15607	n.d.	
FLP208	104 111 189 223 290 319 362 519	64 73 146 153 235 263 309+C 315+C 523-524d 573+CC		n.d.	n.d.	14040 14502 14766	n.d.	n.d.	A2zg1
TG AG 41	111 223 266 290 319 362	64 73 146 153 235 263 315+C 523dd		n.d.	n.d.	n.d.	n.d.	12705 12858	A2zk3
FLP072	111 223 266 290 319 362	64 73 146 235 263 309+CC 315+C 523-524d 663		n.d.	n.d.	n.d.	n.d.	12705 12858	A2zk3
FLP071	[15927] 111 223 266 290 319 362 519	64 73 146 153 235 263 309+C 315+C 523-524d 663		n.d.	n.d.	n.d.	n.d.	12705 12858	A2zk3
CC 088	[15968] 111 223 290 319 362	64 73 143 146 153 235 263 309+C 315+C 523dd		n.d.	13708 14766	n.d.	n.d.	n.d.	A2zl
CC 096	[15968] 111 223 290 319 362	64 73 143 146 153 235 263 309+C 315+C 523dd		n.d.	13708 //	n.d.	n.d.	n.d.	A2zl

Muestra	Región Control extendida		FR2 (F2480-R4860)		Fragmento FR7	Fragmento FR7	FR7 (Celemics)	Fragmento FR6	Clado
	Región 5'	Región 3'	rango	PNS	rango 13700~14841	rango 14030-14841	rango 13845-16031	rango 12494-13557	
REQ 018	[15968] 111 223 290 319 362	64 73 143 146 153 235 263 309+C 315+C 523dd		n.d.	13708 //	n.d.	n.d.	n.d.	A2zl
FOR011	[15968] 111 189 193+C 223 290 319 362	64 73 143 146 153 204 235 263 315+C 523-524d		n.d.	13708 14766	n.d.	n.d.	n.d.	A2zl
REQ 057	[15968] 111 223 290 319 362	73 143 146 153 235 263 309+C 315+C 523dd		n.d.	13708 14620 //	n.d.	n.d.	n.d.	A2zl
FOR003	[15968] 111 223 290 319 362 519	73 143 146 153 235 263 309+CC 315+C 523-524d [663]		n.d.	13708 13768 //	n.d.	n.d.	n.d.	A2zl
FLP231	[15968] 111 223 290 319 362 519	64 73 143 146 153 235 263 309+CC 315+C 523-524d		n.d.	13708 13768 14605 14766	n.d.	n.d.	n.d.	A2zl
MLG 077	111 223 224Y 290 319 362 519	73 143 146 153 235 263 309+CC 315+C 523dd		n.d.	13708 13768 14605 14766	n.d.	n.d.	n.d.	A2zl
GUA 007	111 209 223 290 319 362 519	73 146 153 235 263 315+C 523dd	2550-4843	2706 4248 4769 4824	n.d.	n.d.	n.d.	n.d.	

Anexo 3.4. Frecuencia (en porcentaje) de subhaplogrupos identificados en la ciudad de Formosa y su ocurrencia en localidades del NEA, Tartagal, NOA y Cuyo, obtenidas de la BBDD-RI.

En negrita se registran los polimorfismos de la región codificante.

Hg	Clado	Polimorfismos diagnósticos	Formosa	NEA	Tartagal	NOA	Cuyo
A2	A2zb1	A2 ₍₆₄₎ + (3693 /485-16269-16360- 1461T-11914 -/64@)	0,59	0,18	0	0	0
	A2zc1	A2 + (3834-15244 -16526-279)	0,59	0,18	0	0	0
	A2ze	A2 + (3083 -15734-16422- 8276d)	0,59	0,18	0	0	0
	A2zz	A2 ₍₆₄₎ + (153@- 4158-9878-6497/13659)	0,59	0,89	0	0	0
	A2aa	A2 + (16239A-16266)	0,59	0,72	0,59	0	0
	A2ah	A2 + (16097 -16098)	1,18	0,54	0	0,12	0
	A2zg1	A2 + (8843-9482-14040-14502 -573+C/16104)	0,59	0,36	0	0	0
	A2zh	A2 ₍₆₄₎ + (739-9210-15300 -210)	0,59	2,15	0,00	0,12	0,50
	A2bf	A2 + (16051- 4616)	0,59	1,61	2,37	0,48	0
	A2ab	A2 + 16291-207-489	0,59	0,54	0	0	0
	A2zj	A2 ₍₆₄₎ + 338- 3438 -16468	2,37	1,79	0,59	0	0
	A2zk3	A2 + (16266/ 10237-12858/14364)	2,37	0,72	0,59	0	0
	A2zl	A2 ₍₆₄₎ + (2850-4216 -15968-143)	1,78	1,25	0	0	0,25
		A2 ₍₆₄₎ + (15884 + 14491)	1,18	0,36	0	0	0
	A2be	A2 + (3535 -16111@ -59-64- 12618 -15929-15930-16126-16278-152-297- 13708)	5,33	4,83	0	0	0
	A2yc	A2 + (9096-14127C -15951-16126-16259)	1,78	0,54	0	0	0
		A2 sin asignar	2,37	1,97	4,14	1,33	1,75
	Total A2	23,67	18,25	8,28	2,06	2,51	

Hg	Clado	Polimorfismos diagnósticos	Formosa	NEA	Tartagal	NOA	Cuyo
B2	B2b3a	B2 + (152 -271-16249 -16312 - 6755)	1,18	2,15	0	0,12	0,25
	B2b15	B2+ (6755 -143 -146)	0,59	0,36	0,59	0,48	2,76
	B2b*	B2 + 6755	1,18	0,36	0,59	0,73	0
	B2e	B2 + (6119- 6473)	4,14	4,29	0,59	0,12	0
	B2h	B2 + (16468 - 11821)	1,78	0,89	0	0	0
	B2i1	B2 + (430 -485 -16311 - 6272)	1,18	1,79	0,59	0	0
	B2aj	B2 + (8853 -15883 -16188)	0,59	0,18	2,96	7,76	1,75
		B2+14470	0,59	0,36	2,37	1,70	0,75
	B2an	B2 + (103-152- 14470 -16241)	5,33	4,29	0	0	0
		B2+(7055 -247)	0,59	0,18	4,73	0,36	0
	B2za	B2+(3372 -16316-151)	1,18	1,43	2,96	0,61	0,25
		B2+(16092-16126-16214-16355A-152-263@)	0,59	0,18	0,59	0,48	1,75
	B2zc	B2+15884	0,59	0,54	1,78	0,24	0
	B2zd	B2+16178	4,14	6,08	0,59	0	0,25
		B2+(16142-111, non-11590)	0,59	0,54	0	0	0
	B2al	B2 +(151- 13934 -15110 -16173 -16357)	0,59	1,25	0	0	0
	B2ze	B2+ 12771	1,18	0,36	0	0	0,25
	B2o1a	B2 + (7270 -16092-16145/ 14500 -152/ 8573 - 8856 /228-234)	1,78	0,54	0	0	0
	B2zf1	B2+(3615-4232 -16311-16319-16354-16524)	1,78	0,54	0	0	0
		B2 sin asignar	1,78	2,15	5,92	1,45	0,50
	Total B2	31,36	28,44	24,26	14,06	8,52	

Hg	Clado	Polimorfismos diagnósticos	Formosa	NEA	Tartagal	NOA	Cuyo
C1b	C1b30a1	C1b + 16270 + 16126 + 2393	1,18	4,11	0,59	0,12	0,25
	C1b19	C1b+(4676A-15462 /143-152-16266/195)	0,59	0,54	0	1,82	0,50
	C1b16	C1b +146 + 12813	2,96	2,15	5,92	5,82	3,26
	C1b18	C1b + 4736	2,96	3,40	0	0,61	1,25
	C1b17	C1b + 207 + 150 + 10029 + 12361	0,59	0,36	0	0,48	0
	C1b101	C1b+2404-8772-13827-16384	1,78	1,61	0,59	0	0
	C1b2	C1 + (493-7013)	1,78	1,25	0	0	0
	C1b102	C1b+13212-16526/152	0,59	1,61	0	0	0
		C1b+(4973-5978-10143-16172-16256)	0,59	0,89	0	0	0
		C1b+(16144-152-195)	1,18	0,36	0	0	0
		C1b+(316-16176-16362)	0,59	0,36	0	0	0
		C1b sin asignar	0,59	1,25	2,37	1,94	2,01
	Total C1b		15,38	17,89	9,47	10,79	7,27
C1c	C1c100	C1c+(152-4216-16311-16368-16390-151)	0,59	0,18	1,18	0,36	1,00
C1d		C1d+(4688-5656-16247-195)	0,59	0,18	0	0	0
	C1d1ab1	C1d1+(15728/16362-5C)	0,59	0,18	0	0	0
	C1d1aa2a2	C1d1+(12501/11090T-195/12879-16093/)	1,18	0,54	0	0	0
	C1d1d	C1+(051-172 - 7697- 928-7337)	5,92	3,04	0	0	0,25
	C1d1b	C1+ (194- 195 -16051- 928)	2,96	1,43	2,37	4,61	3,26
		Total C1d		11,24	5,37	2,37	4,61

Hg	Clado	Polimorfismos diagnósticos	Formosa	NEA	Tartagal	NOA	Cuyo
D1	D1f1	D1 + 16142 + 16179 + 16295 + 16497	5,92	2,86	0	0	0
	D1zd	D1+@16325	0,59	0,36	4,73	0,12	0
	D1j	D1 + (152 -16242- 16311)	1,78	1,61	3,55	13,09	7,27
	D1za	D1+6305T- 13989 -146/ 3591-4883-7444C-11440 -143	0,59	0,36	0,59	0	0
	D1e	D1+(3316 -146)	3,55	1,07	0	0	0
	D1zc	D1+16187A	0,59	0,72	0	0	0
	D1x1a	D1+11150/15106-210/16239-228	0,59	0,36	0	0	0
		D1 sin asignar	2,96	1,43	2,96	0,48	0,75
Total D1			16,57	8,77	11,83	13,70	8,02
D4h3a	D4h3a91		0,59	0,18	0	0	0
	D4h3a92		0,59	0,18	0	0	0
	Total D4h3a			1,18	0,36	0	0
Totales parciales			100	79,25	57,40	45,58	30,83
Total linajes nativos			100	100	100	100	100

Anexo 3.5. Frecuencia de subhaplogrupos mayoritarios en regiones de Argentina y de países limítrofes (en porcentaje)

Clado	FOR	NEA	URU	BRA	PAR	CENTRO	CHA_ARG	NOA	TG	CUYO	SDE
A2zj	2,37	1,79	0	1,57	1,87	0	0,81	0	0,59	0	0
A2zk3	2,37	0,72	0	0,52	1,87	0	0	0	0,59	0	0
A2be	5,33	4,83	9,38	2,09	0	0	0	0	0	0	0
B2an	5,33	4,29	6,25	3,14	0,93	0	0	0	0	0	0
B2zd	4,14	6,08	3,13	5,76	2,80	0	0,81	0	0,59	0,25	0
B2e	4,14	4,29	3,13	3,14	2,80	1,32	1,61	0,12	0,59	0	1,01
C1b16	2,96	2,15	9,38	0	0	7,89	8,06	5,95	5,92	3,26	16,16
C1b18	2,96	3,40	0	4,19	5,61	1,32	0	0,59	0	1,25	0
C1d1b	2,96	1,43	0	0	0,93	2,63	0,81	4,88	2,37	3,26	15,15
C1d1d	5,92	3,04	3,13	2,09	2,80	0	0	0	0	0,25	0
D1e	3,55	1,07	0	1,57	3,74	0	0	0	0	0	0
D1f1	5,92	2,86	0	2,62	0,93	0	0	0	0	0	0
Frecuencia acumulada	47,93	35,96	34,38	26,70	24,30	13,16	12,10	11,53	10,65	8,27	32,32

Anexo 3.6. Secuencias de región control de ADNmt recopiladas de la bibliografía e identificadas como A2bf por punto de muestreo, correspondientes al tercer nivel de análisis.

Localidad	Latitud	Longitud	N(A2bf)	N(A,B,C,D)	Frec rel A2bf (%)	Población	Referencia
ARG_LaEsperanza	-24,1	-64,84	4	94	4,3	Urbana	Garita-Onandía, 2021
ARG_SanSalvadordeJujuy	-24,18	-65,3	2	35	5,7	Urbana	Motti, 2012
ARG_Sgo_del_Estero	-27,78	-64,26	4	109	3,7	Urbana	Schwab, 2018
ARG_Formosa	-26,18	-58,18	1	169	0,6	Urbana	Este trabajo
BOL	-16,81	-64,78	4	91	4,4	Quechua	Gayá -Vidal et al., 2011
ARG_LaQuiaca	-22,1	-65,59	1	9	11,1	Urbana	Motti et al., 2012
CHI_Iquique	-20,23	-70,13	2	190	1,1	Urbana	Gómez-Carballea et al., 2016
BOL_Tar	-21,51	-64,7	4	29	13,8	Wichí	Sandoval et al., 2023
ARG_SantaFe	-31,6	-60,7	1	13	7,7	Qom	Sala et al., 2019
ARG_Rosario	-32,9	-60,68	6	37	16,2	Qom	Sala et al., 2019
ARG_LagunaBlanca	-25,12	-58,25	6	37	16,2	Qom	Sala et al., 2019
ARG_SanMartin	-26,53	-59,34	8	27	29,6	Qom	Sala et al., 2019
ARG_Tostado	-29,22	-61,77	14	27	51,9	Mocoví	Sala et al., 2019
ARG_IngJuarez	-23,89	-61,85	11	48	22,9	Wichí	Sala et al., 2019
ARG_Ibarreta	-25,21	-59,85	6	56	10,7	Pilagá	Sala et al., 2019
PAR	-25,51	-54,61	2	350	0,6	Urbana	Simão et al., 2021
ARG_BuenosAires	-36,62	-60,21	2	52	3,8	Urbana	Catelli et al., 2011
ARG_MisionNP	-24,92	-61,49	4	206	1,9	Criollo	Sevini et al., 2014
ARG_MisionNP	-24,92	-61,49	92	342	26,9	Wichi	Sevini et al., 2014
ARG_Tucuman	-26,8	-65,22	1	87	1,1	Urbana	Mann et al., 2013
ARG_RioNegro	-40,45	-67,59	1	30	3,3	Urbana	Bobillo et al., 2010

Localidad	Latitud	Longitud	N(A2bf)	N(A,B,C,D)	Frec rel A2bf (%)	Población	Referencia
ARG_FOR_toba	-26,22	-58,2	6	24	25,0	Qom	Cabana et al., 2006, Demarchi comm pers
ARG_CHA_wichi	-25,42	-60,19	10	32	31,3	Wichi	Cabana et al., 2006, Demarchi comm pers
ARG_CHA_toba	-26,02	-60,22	5	43	11,6	Qom	Cabana et al., 2006, Demarchi comm pers
ARG_FOR_Pilagá	-24,51	-58,51	1	37	2,7	Pilagá	Cabana et al., 2006, Demarchi comm pers
ARG_FOR_Wichi	-24,15	-61,37	5	70	7,1	Wichi	Cabana et al., 2006, Demarchi comm pers
ARG_SantaFe_dos	-31,3	-60,95	1	5	20,0	Mocovi	Tamm et al., 2007
URU_Artigas	-33,4	-56,47	2	45	4,4	Urbana	Sans et al., 2015
Guayana_Francesa	4,39	-53,31	1	14	7,1	Kaliña	Mazieres et al., 2008
VEN_Caracas	10,47	-66,9	1	67	1,5	Caraqueño	Castro de Guerra et al., 2012
ARG_Corrientes	-28,47	-57,73	1	7	14,3	Urbana	Catelli et al., 2011

Anexo 3.7. Mitogenomas recopilados para la construcción de árboles filogenéticos del tercer nivel de análisis

#	Haplogrupo	ID	Código Genbank	Punto geográfico	Población	Referencia
1	A2bf1	GranChaco176	KP172314	Argentina, Chaco, Misión Nueva Pompeya	Criollo	Sevini et al., 2014
2	A2bf1a	GranChaco34	KP172313	Argentina, Chaco, Misión Nueva Pompeya	Wichí	Sevini et al., 2014
3	A2bf1a	GranChaco324	KP172317	Argentina, Chaco, Misión Nueva Pompeya	Wichí	Sevini et al., 2014
4	A2bf1a	GranChaco366	KP172318	Argentina, Chaco, Misión Nueva Pompeya	Wichí	Sevini et al., 2014
5	A2bf1a	GranChaco442	KP172319	Argentina, Chaco, Misión Nueva Pompeya	Criollo	Sevini et al., 2014
6	A2bf1a	GranChaco33	KP172369	Argentina, Chaco, Misión Nueva Pompeya	Wichí	Sevini et al., 2014
7	A2bf1a	GranChaco9	KP172391	Argentina, Chaco, Misión Nueva Pompeya	Wichí	Sevini et al., 2014
8	A2bf1a	GranChaco154	KP172416	Argentina, Chaco, Misión Nueva Pompeya	Criollo	Sevini et al., 2014
9	A2bf1a	GranChaco18	KP172435	Argentina, Chaco, Misión Nueva Pompeya	Wichí	Sevini et al., 2014
10	A2bf1a	GranChaco430	KP172375	Argentina, Chaco, Misión Nueva Pompeya	Wichí	Sevini et al., 2014
11	A2bf1a	GranChaco273	KP172316	Argentina, Chaco, Misión Nueva Pompeya	Criollo	Sevini et al., 2014

#	Haplogrupo	ID	Código Genbank	Punto geográfico	Población	Referencia
12	A2bf1a	GranChaco15	KP172312	Argentina, Chaco, Misión Nueva Pompeya	Criollo	Sevini et al., 2014
13	A2bf1a	GranChaco427	KP172374	Argentina, Chaco, Misión Nueva Pompeya	Wichí	Sevini et al., 2014
14	A2bf2	HTA-TC18	MW057670	Uruguay, Tacuarembó	Urbana	Figueiro et al., 2022
15	A2bf2	HTA-TC17	MW057669	Uruguay, Tacuarembó	Urbana	Figueiro et al., 2022
16	B2e1			Uruguay, Maldonado	Urbana	Family Tree DNA
17	B2e1			Uruguay, Durazno	Urbana	Family Tree DNA
18	B2e1		MH981826	Paraguay, Alto Paraná	Urbana	Simão et al., 2019
19	B2e1			Argentina, Chubut	Tehuelche	IMBICE, Inédito
20	B2e1			Brasil, Paraná	Guarapuavá	Family Tree DNA
21	B2e1			Argentina, Santa Fe	Urbana	Family Tree DNA
22	B2e1a			Brasil	Urbana	Family Tree DNA
23	B2e1a		MH981834	Paraguay, Alto Paraná	Urbana	Simão et al., 2019
24	B2e1	M64		Córdoba, Argentina	Urbana	García et al., 2021
25	B2e2			Argentina, Santa Fe	Urbana	Family Tree DNA
26	B2e2a	LPZ058		Argentina, La Paz	Urbana	IMBICE, Inédito
27	B2e2a	CC024		Argentina, Curuzú Cuatiá	Urbana	IMBICE, Inédito
28	B2e2a1		KP172321	Argentina, Chaco, Misión Nueva Pompeya	Wichí	Sevini et al., 2014

#	Haplogrupo	ID	Código Genbank	Punto geográfico	Población	Referencia
29	B2e2a1		KP172384	Argentina, Chaco, Misión Nueva Pompeya	Wichí	Sevini et al., 2014
30	B2e2a1			Argentina, Entre Ríos	Urbana	Family Tree DNA
31	B2e3		KY645637	Colombia, Guainía, Río Inírida	Puinave	Arias et al., 2017
32	B2e3			Brasil, Alagoas	Urbana	Family Tree DNA
33	B2e3			Brasil, Paraíba	Urbana	Family Tree DNA
34	B2e4		KM101574	Estados Unidos, Alabama	Afro-Americano	Just et al., 2015
35	B2e4		EU095209	Brasil	Waiwai	Fagundes et al., 2008
36	B2e5		KY645522	Colombia, Amazonas, Mirití-Paraná river	Tanimuka	Arias et al., 2017
37	B2e5		KY645712	Colombia, Guaviare, Upper Vaupés river	Barasano	Arias et al., 2017
38	B2e5		KY645523	Colombia, Guaviare, Upper Vaupés river	Barasano	Arias et al., 2017
39	B2e6		KY645627	Colombia, Guainia, Inirida river	Puinave	Arias et al., 2017
40	B2e6a			Colombia, Guainia, Inirida river	Puinave	
41	B2e6a			Colombia, Guainia, Inirida river	Puinave	
42	B2e6a			Colombia, Guainia, Inirida river	Puinave	

#	Haplogrupo	ID	Código Genbank	Punto geográfico	Población	Referencia
43	B2e6a			Colombia, Guainia, Inirida river	Puinave	
44	B2e6a			Colombia, Guainia, Inirida river	Puinave	
45	B2e6a			Colombia, Guainia, Inirida river	Puinave	
46	B2e6a		KY645620	Colombia, Guainia, Atabapo river	Curripaco	Arias et al., 2017
47	B2e6a			Colombia	Curripaco o Piapoco	
48	B2e6a			Colombia	Curripaco o Piapoco	
49	B2e6a			Colombia	Curripaco o Piapoco	
50	B2e6a			Colombia	Curripaco o Piapoco	
51	B2e6a			Colombia	Curripaco o Piapoco	
52	B2e+10646		KY645693	Colombia, Guaviare, Upper Vaupés river	Tukano	Arias et al., 2017
53	B2e+10646		KY645714	Colombia, Guaviare, Upper Vaupés river	Desano	Arias et al., 2017
54	B2e+9881		KY645707	Colombia, Guaviare, Upper Vaupés river	Carijona	Arias et al., 2017

#	Haplogrupo	ID	Código Genbank	Punto geográfico	Población	Referencia
55	B2an	DF78		Brasil, Distrito Federal	Urbana	Ávila et al., 2019
56	B2an	USA_TX_0102		Estados Unidos, Texas	Urbana	King et al., 2015
57	B2an1	LS321		?	Urbana	Strobl et al., 2019
58	B2an1	DF78	MN894781	Brasil, Distrito Federal	Urbana	Ávila et al., 2019
59	B2an1a	DF07	MN894715	Brasil, Distrito Federal	Urbana	Ávila et al., 2019
60	B2an1a	LS302		?	Urbana	Strobl et al., 2019
61	B2an1a1	C08	MW057687	Uruguay, Montevideo	Urbana	Figueiro et al 2022
62	B2an1a1	PUC413	OL638648	Brasil, Porto Alegre	Urbana	Ávila et al., 2022
63	B2an1a1	PUC543	OL638732	Brasil, Porto Alegre	Urbana	Ávila et al., 2022
64	B2an2	VC11	MW057684	Uruguay, Trinidad	Urbana	Figueiro et al., 2022
65	B2an2a	23000517	MH981911	Paraguay, Alto Paraná	Urbana	Simão et al., 2019
66	B2an2a	LS185		Paraguay	Urbana	Strobl et al., 2019
67	B2an3	Xavante 12	EU095211	Brasil	Xavante	Fagundes et al., 2008
68	B2an3	PUC554	OL638742	Brasil, Porto Alegre	Urbana	Ávila et al., 2022
69	D1e	Zoró 23	EU095236	Brasil	Zoró	Fagundes et al., 2008
70	D1e1	23000360	MH981846	Paraguay, Alto Paraná	Urbana	Simão et al., 2019
71	D1e1	23000376	MH981862	Paraguay, Alto Paraná	Urbana	Simão et al., 2019

#	Haplogrupo	ID	Código Genbank	Punto geográfico	Población	Referencia
72	D1e1	23000387	MH981873	Paraguay, Alto Paraná	Urbana	Simão et al., 2019
73	D1e1	23000401	MH981887	Paraguay, Alto Paraná	Urbana	Simão et al., 2019
74	D1e2	HGDP01000	EU597510	Brasil	Karitiana	Hartmann et al., 2009
75	D1e2	HGDP00995	KF451507	Brasil	Karitiana	Lippold et al., 2014
76	D1e2	HGDP01014	KF451525	Brasil	Karitiana	Lippold et al., 2014
77	D1e2	HGDP01015	KF451526	Brasil	Karitiana	Lippold et al., 2014
78	D1e2	HGDP01018	KF451529	Brasil	Karitiana	Lippold et al., 2014
79	D1e2	HGDP01009	KF451520	Brasil	Karitiana	Lippold et al., 2014
80	D1e2	HGDP01013	KF451524	Brasil	Karitiana	Lippold et al., 2014
81	D1e2	HGDP01004	KF451515	Brasil	Karitiana	Lippold et al., 2014
82	D1e2	HGDP01007	KF451518	Brasil	Karitiana	Lippold et al., 2014
83	D1e2	HGDP00999	KF451511	Brasil	Karitiana	Lippold et al., 2014
84	D1e2	HGDP01001	KF451513	Brasil	Karitiana	Lippold et al., 2014
85	D1e2	HGDP01016	KF451527	Brasil	Karitiana	Lippold et al., 2014
86	D1e2	HGDP01012	KF451523	Brasil	Karitiana	Lippold et al., 2014
87	D1e2	HGDP01017	KF451528	Brasil	Karitiana	Lippold et al., 2014
88	D1e2	HGDP00996	KF451508	Brasil	Karitiana	Lippold et al., 2014
89	D1f1	13-1995	KY681012	Venezuela	Urbana	Brandini et al., 2018

#	Haplogrupo	ID	Código Genbank	Punto geográfico	Población	Referencia
90	D1f1	13-2417	KY681014	Venezuela	Urbana	Brandini et al., 2018
91	D1f1	PUC304	OL638564	Brasil, Porto Alegre	Urbana	Ávila et al., 2022
92	D1f1	23000342	MH981828	Paraguay, Alto Paraná	Urbana	Simão et al., 2019
93	D1f1	HGDP01009	KF451520	Brasil	Karitiana HGDP	Lippold et al., 2014
94	D1f1	Waiwai 05	EU095235	Brasil	Waiwai	Fagundes et al., 2008
95	D1f1	KTN18 (W10)	EU095238	Brasil	Katuena (=Waiwai)	Fagundes et al., 2008
96	D1f1	TYR04(W07)	EU095240	Brasil	Tyrió	Fagundes et al., 2008
97	D1f1	TYR16(W08)	EU095241	Brasil	Tyrió	Fagundes et al., 2008

Anexo 3.8. Secuencias de región control de ADNmt recopiladas de la bibliografía e identificadas como B2an por punto de muestreo, correspondientes al tercer nivel de análisis.

Localidad	Latitud	Longitud	N(B2an)	N(A,B,C,D)	Frec rel B2an (%)	Población	Referencia
ARG_CuruzuCuatia	-29,79	-58,05	5	140	3,57	urbana	IMBICE, inédito
ARG_Formosa	-26,18	-58,18	9	169	5,33	urbana	Este trabajo
ARG_Gualeguaychu	-33,01	-58,52	8	115	6,96	urbana	IMBICE, inédito
ARG_LaPaz	-30,74	-59,65	1	56	1,79	urbana	IMBICE, inédito
ARG_Reconquista	-29,15	-59,65	1	79	1,27	urbana	IMBICE, inédito
ARG_Misiones	-26,59	-54,17	7	109	6,42	Mbyá Guaraní	Sala & Corach, pers. comm.
ARG_Misiones_dos	-26,89	-55,03	1	22	4,55	urbana	Bobillo et al., 2010
ARG_BuenosAires	-36,62	-60,21	1	89	1,12	urbana	Bobillo et al., 2010
ARG_BuenosAires_dos	-35,1	-58,38	1	52	1,92	urbana	Catelli et al., 2011
ARG_BuenosAires_tres	-34,58	-58,47	1	27	3,70	Basque Diaspora	Baeta et al., 2015
ARG_RioNegro	-40,45	-67,59	1	30	3,33	urbana	Bobillo et al., 2010
BRA_Amazonas	-5,57	-60,78	7	19	36,84	Parakanã	Ramallo et al., 2013
BRA_Para	-4,49	-50,61	3	90	3,33	Gorotire Kayapó (Gé)	Galagedara GRS, Tesis de maestría, 2018
BRA_Alagoas	-9,65	-36,68	2	55	3,64	Alagoans	Barbosa et al., 2008
BRA_DistritoFederal	-15,73	-47,9	3	21	14,29	urbana	Nogueira, tesis doctoral, 2019
BRA_EspirituSanto	-19,24	-40,5	1	21	4,76	urbana (ES)	Sanches et al., 2014
BRA_SE	-22,4	-44,68	13	103	12,62	urbana (SP, MG, ES, RJ)	Fridman et al., 2014
BRA_MinasGerais	-17,49	-43,36	1	33	3,03	"White" brazilians mainly MG	Alves-Silva et al.,2000

Localidad	Latitud	Longitud	N(B2an)	N(A,B,C,D)	Frec rel B2an (%)	Población	Referencia
BRA_EspirituSanto_dos	-20,11	-40,81	3	49	6,12	urbana (ES)	dos Reis et al., 2019
BRA_SaoPaulo	-22,19	-49,71	4	31	12,90	urbana (SP)	Prieto et al., 2011
BRA_RiodeJaneiro	-21,95	-42,88	4	53	7,55	urbana (RJ)	Simão et al., 2018
BRA_SaoPaulo_dos	-22,36	-48,46	10	132	7,58	urbana (SP)	Cardena et al., 2013
BRA_RiodeJaneiro_dos	-21,55	-41,29	3	46	6,52	Cariocas (RJ)	Bernardo et al., 2014
BRA_Acre	-8,83	-70,95	1	21	4,76	urbana	Nogueira, Tesis doctoral, 2019
BRA_SantaCatarina	-27,24	-50,49	1	17	5,88	urbana	Palencia et al., 2010
BRA_Parana	-24,71	-52,09	2	60	3,33	urbana	Poletto et al., 2018
ARG_JUJ_Ramal	-24,03	-64,63	1	41	2,44	urbana	Cardoso et al., 2013
BOL_	-21,04	-64,15	2	43	4,65	guarani	Sandoval et al., 2023
PAR_Este	-25,51	-54,61	6	350	1,71	urbana	Simão et al., 2021
URU_Artigas	-33,4	-56,47	1	45	2,22	urbana	Sans et al., 2015

Anexo 3.9. Secuencias de región control de ADNmt recopiladas de la bibliografía e identificadas como D1e por punto de muestreo, correspondientes al tercer nivel de análisis.

Localidad	Latitud	Longitud	N(D1e)	N(A,B,C,D)	Frec rel D1e (%)	Población	Referencia
ARG_Formosa	-26,18	-58,18	6	169	3,55	Urbana	Este trabajo
ARG_Misiones	-26,59	-54,17	1	22	4,55	Urbana	Bobillo et al., 2010
PAR_este	-25,51	-54,61	3	350	0,86	Urbana	Simão et al., 2021
ARG_Formosa_dos	-26,64	-57,98	1	15	6,67	Urbana	Bobillo et al., 2010
ARG_Rosario	-32,9	-60,68	1	37	2,70	Tobas	Sala et al., 2019

Anexo 3.10. Secuencias de región control de ADNmt recopiladas de la bibliografía e identificadas como D1f1 por punto de muestreo, correspondientes al tercer nivel de análisis.

Localidad	Latitud	Longitud	N(D1f1)	N(A,B,C,D)	Frec rel D1f1 (%)	Población	Referencia
USA_TX	31,56	-99,08	1	148	0,68	urbana	Monson et al., 2002
VEN_Caracas	10,47	-66,9	3	67	4,48	caraqueños	Castro de Guerra et al., 2011
URU	-34,78	-56,14	1	37	2,70	urbana	Pagano et al., 2005
ARG_Reconquista	-29,14	-59,64	3	79	3,80	urbana	IMBICE, inédito
ARG_CuruzuCuatia	-29,79	-58,05	2	140	1,43	urbana	IMBICE, inédito
ARG_LaPaz	-30,74	-59,65	1	56	1,79	urbana	IMBICE, inédito
ARG_Formosa	-26,18	-58,18	10	169	5,92	urbana	Este trabajo
ARG_Corrientes	-28,56	-58,12	2	23	8,70	urbana	Bobillo et al., 2010
ARG_Misiones	-26,59	-54,17	1	22	4,55	urbana	Bobillo et al., 2010
PAN	9,17	-79,69	1	1010	0,10	urbana	Perego et al., 2012
BRA_EspirituSanto	-19,2	-40,39	1	20	5,00	urbana	dos Reis et al., 2019
BRA_Rio_de_Janeiro	-21,95	-42,88	1	46	2,17	cariocas	Bernardo et al., 2014
BRA_SE	-22,4	-44,68	2	103	1,94	urbana (SP, MG, ES, RJ)	Fridman et al., 2014
BRA_SaoPaulo	-22,19	-49,71	4	132	3,03	urbana (brasileros blancos/pardos)	Cardena et al., 2013
BRA_Paraná	-24,71	-52,09	1	60	1,67	urbana	Poletto et al., 2018
BRA_Roraima	1,54	-61,58	1	8	12,50	urbana	Nogueira tesis doctoral, 2019

Localidad	Latitud	Longitud	N(D1f1)	N(A,B,C,D)	Frec rel D1f1 (%)	Población	Referencia
BRA_Mato Grosso	-12,69	-54,87	1	20	5,00	urbana	Nogueira tesis doctoral, 2019
ARG_Misiones_dos	-26,89	-55,03	1	2	50,00	urbana	Catelli et al., 2011
VEN_Caracas_dos	10,51	-66,9	2	17	11,76	urbana	Gómez-Carballa et al., 2012
VEN_Merida	8,61	-71,18	1	197	0,51	urbana	Gómez-Carballa et al., 2012
ARG_RioCuarto	-33,12	-64,34	1	40	2,50	urbana	García et al., 2018
PAR_este	-25,51	-54,61	19	350	5,43	urbana	Simão et al., 2021
BRA_Alagoas	-9,65	-36,68	1	55	1,82	urbana	Barbosa et al., 2008
BRA_EspirituSanto_dos	-20,11	-40,81	1	21	4,76	urbana	Sanches et al., 2014