

Universidade do Estado do Rio de Janeiro

Centro Biomédico Instituto de Biologia Roberto Alcântara Gomes

Julyana da Silva Varela Ribeiro

Padrões de herança paterna em populações da América do Sul

Julyana da Silva Varela Ribeiro

Padrões de herança paterna em populações da América do Sul

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Biociências, Universidade do Estado do Rio de Janeiro.

Orientadora: Profa. Dra. Leonor Gusmão

Coorientador: Prof. Dr. Elizeu Fagundes de Carvalho

CATALOGAÇÃO NA FONTE UERJ/REDE SIRIUS/BIBLIOTECA CB-A

R484	Ribeiro, Julyana da Silva Varela. Padrões de herança paterna em populações da América do Sul / Julyana da Silva Varela Ribeiro 2023. 167 f.
	Orientadora: Prof. ^a Dra. Leonor Gusmão Coorientador: Prof. Dr. Elizeu Fagundes de Carvalho
	Doutorado (Tese) - Universidade do Estado do Rio de Janeiro, Instituto de Biologia Roberto Alcântara Gomes. Pós-graduação em Biociências.
	1. Cromossomo Y - Teses. 2. Herança paterna - Genética. 3. Padrões de herança. 4. América do Sul – Etiologia. I. Carvalho, Elizeu Fagundes de. II. Gusmão, Leonor. III. Universidade do Estado do Rio de Janeiro. Instituto de Biologia Roberto Alcântara Gomes. IV. Título.
	CDU 575.113(892)

Bibliotecária: Ana Rachel Fonseca de Oliveira CRB7/6382

Autorizo apenas para fins acadêmicos e científicos, a repre	odução total ou parcial desta tese
desde que citada a fonte.	
Assinatura	Data

Julyana da Silva Varela Ribeiro

Padrões de herança paterna em populações da América do Sul

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Biociências, da Universidade do Estado do Rio de Janeiro.

Aprovada em 16 dezembro de 2023

Coorientador: Prof. Dr. Elizeu Fagundes de Carvalho Instituto de Biologia Roberto Alcântara Gomes – UERJ

Banca Examinadora:

Prof^a. Dra. Leonor Gusmão (Orientadora)

Instituto de Biologia Roberto Alcântara Gomes – UERJ

Prof^a. Dra. Cíntia Barros Santos Rebouças

Instituto de Biologia Roberto Alcântara Gomes – UERJ

Prof. Dr. Rodrigo Moura-Neto

Universidade Federal do Rio de Janeiro

Dr. Mário Campos Júnior

Fundação Oswaldo Cruz

Rio de Janeiro

AGRADECIMENTOS

À minha família pelo apoio incondicional e segurança que tornam mais aprazível minha jornada acadêmica.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão de bolsa de estudos durante o curso de Doutorado.

À minha orientadora Leonor Gusmão pela orientação atenciosa e dedicada, pelos conselhos no âmbito acadêmico e pessoal, e pela amizade generosa construída.

Ao meu coorientador Elizeu Carvalho por sua confiança e pelo apoio prestado durante a realização deste trabalho nas instalações do Laboratório de Diagnóstico por DNA.

Ao Instituto IPATIMUP, em Porto, Portugal, e em especial a todos que me receberam com carinho e permitiram a aquisição de conhecimentos durante a *Fellowship* concedida pela ISFG.

Aos colaboradores deste projeto, os outros profissionais e companheiros do Laboratório de Diagnóstico por DNA por sua recepção e apoio diário na realização da tese.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

RESUMO

RIBEIRO, Julyana da Silva Varela. **Padrões de herança paterna em populações da América do Sul.** 2023. 167f. Tese (Doutorado em Biociências) — Instituto de Biologia Roberto Alcântara Gomes, Universidade do Estado do Rio de Janeiro, 2023.

Os padrões de mistura genética que se observam atualmente nas populações sulamericanas é o resultado do encontro de grupos nativos com europeus, e com populações africanas, como também de movimentos migratórios que caracterizaram os séculos XX e XXI. Dessa maneira, é importante a construção de bancos de dados genéticos que permitam capturar a alta diversidade e heterogeneidade genética presente nos diferentes grupos populacionais sul-americanos. Devido à sua natureza não recombinante e transmissão uniparental, os marcadores do cromossomo Y são amplamente utilizados em genética populacional para inferir genealogias e migrações de populações humanas. O objetivo deste estudo foi contribuir para o estudo das linhagens paternas de populações da América do Sul, caracterizando marcadores do cromossomo Y. Portanto, 2.202 amostras de homens não aparentados, residentes na Argentina, Brasil, Colômbia, Equador e Paraguai, foram genotipadas para o PowerPlex Y23. Além disso, parte dessas amostras foi analisada para 67 Y-SNPs. Para além do objetivo principal, foram feitas análises de segregação em pares paifilho, no âmbito de um estudo colaborativo, com o objetivo de aumentar a precisão das estimativas de taxa de mutação em Y-STRs. Uma alta diversidade haplotípica foi encontrada nas populações estudadas (>0,9997), semelhante ao observado para outras 14 populações latino-americanas, para os 23 marcadores Y-STR. A análise de distâncias genéticas (F_{ST}) entre as populações estudadas neste trabalho e outras populações da América do Sul, assim como com populações europeias, africanas e americanas nativas, revelou uma maior proximidade das populações sul-americanas com as populações ibéricas, exceto Bolívia e Peru, muito provavelmente devido ao maior componente nativo destas populações. Na análise de haplogrupos definidos por Y-SNPs, foi observada uma predominância de linhagens paternas europeias nas populações estudadas (>73,5%), sendo o haplogrupo R1b-S116, originário da Península Ibérica, o mais frequente. A proporção relativa de linhagens de origem nativo-americana e africana variou entre países, sendo que a maior proporção de linhagens nativas foi observada no Equador e as populações brasileira e colombiana mostraram um predomínio das linhagens africanas sobre as nativas, confirmando o esperado a partir de dados históricos. Finalmente, as taxas de mutação calculadas para 201 duos pai-filho, em conjunto com as geradas por outros grupos, num total de 84.715 duos, mostraram que: (i) mutações de uma etapa são mais comuns do que as de várias etapas; (ii) os alelos longos são mais propensos à mutação do que os curtos e tendem a perder repetições enquanto os curtos a ganhar; e (iii) a taxa de mutação e a idade do pai estão positivamente correlacionadas. Com o conjunto dos resultados obtidos contribuiu-se para o conhecimento sobre os processos evolutivos e migratórios que deram origem as populações sul-americanas contemporâneas.

Palavras-chave: América do Sul; marcadores STR; marcadores SNP; taxa de mutação; cromossomo Y.

ABSTRACT

RIBEIRO, Julyana da Silva Varela. **Patterns of paternal inheritance in South American populations.** 2023. 167f. Tese (Doutorado em Biociências) – Instituto de Biologia Roberto Alcântara Gomes, Universidade do Estado do Rio de Janeiro, 2023.

The paternal admixture patterns currently observed in South American populations result from the encounter of native groups with Europeans and African populations. In addition, recent migrations from Europe and other countries in South America have also contributed to their genetic structure. Therefore, it is important to build genetic databases that capture the high genetic diversity and heterogeneity in different South American population groups. Due to their non-recombining nature and uniparental transmission, Y chromosome markers are widely used in population genetics to infer genealogical history and migrations of human populations. This study aimed to contribute to the study of the paternal lineages of South American populations, characterizing Y chromosome markers. Therefore, 2,202 samples from unrelated men, residing in Argentina, Brazil, Colombia, Ecuador and Paraguay, were genotyped for PowerPlex Y23. Furthermore, part of these samples was analyzed for 67 Y-SNPs. In addition to the main objective, marker-specific mutation rates were estimated in parent-child pairs, as part of a collaborative study, to increase the precision of mutation rate estimates in Y-STRs. A high haplotypic diversity was found in the studied populations (>0.9997), similar to what was observed for 14 other Latin American populations, for the 23 Y-STR markers. The analysis of genetic distances (F_{ST}) between the populations studied in this work and other populations in South America, as well as with European, African and Native American populations, revealed a greater proximity between the South American and the Iberian populations, except Bolivia and Peru, most probably due to the greater native component of these populations. In the analysis of haplogroups defined by Y-SNPs, a predominance of European paternal lineages was observed in the populations studied (>73.5%), with haplogroup R1b-S116, originating from the Iberian Peninsula, being the most frequent. The relative proportion of lineages of Native American and African origin varied between countries, with the highest proportion of native lineages being observed in Ecuador, and Brazilian populations and Colombia showing a predominance of African lineages over native ones, confirming what was expected from historical data. Finally, mutation rates calculated for 201 parent-child duos, together with those generated by other groups, totaling 84,715 allelic transfers, showed that: (i) single-step mutations showed to be more common than multi-step ones; (ii) longer alleles showed to be nearly twice more mutable than the shorter ones and the loss of repeats showed to be nearly twice more likely than the gain while short ones gain them; and (iii) the mutation rate and the father's age are positively correlated. The results obtained contributed to knowledge about the evolutionary and migratory processes that gave rise to contemporary South American populations.

Keywords: South America; STR markers; SNP markers; mutation rate; Y chromosome.

LISTA DE FIGURAS

Figura 1 - l	Entrada e dispersão do homem moderno nas Américas17
Figura 2 -	Mapa da Argentina dividida em regiões geográficas: Noroeste, Nodeste, Cuyo,
	Pampas e Patagônia. A capital Buenos Aires está localizada na região dos
	Pampas19
Figura 3 -	Distribuição dos povos indígenas na Terra do Fogo Yámana, Selk'nam, Haush e
	marginalmente os Kawésqar. 20
Figura 4 -	Mapa do Brasil dividido em regiões geopolíticas: Norte, Nordeste, Centro-Oeste,
	Sudeste e Sul. A capital Brasília está localizada no Centro-Oeste
Figura 5 -	Mapa da Colômbia dividida em regiões naturais: Caribe, Andes, Orinoquía,
	Amazônica e Pacífico. A capital Bogotá está localizada na região dos
	Andes
Figura 6 -	Mapa do Equador dividido em regiões naturais: Andes, Pacífico, Amazônia, e a
	região insular (Ilhas Galápagos). A capital Quito está localizada na região
	andina
Figura 7 -	Mapa do Paraguai: sua subdivisão nas regiões (Ocidental e Oriental) e suas
	fronteiras com outros países
Figura 8 - 0	O cromossomo Y
Figura 9 -	Marcadores Y-STR mais utilizados em estudos de genética populacional e forense,
	bem como sua localização no cromossomo Y
Figura 10 -	- Árvore filogenética de haplogrupos do cromossomo Y
Figura 11	- Mapa geográfico da América do Sul, com destaque para os países de origem das
	2.202 amostras deste estudo
Figura 12 -	- Árvore filogenética de marcadores do cromossomo Y
Figura 13	3 - Gel de poliacrilamida da amplificação em multiplex Q de Y-
	SNPs71
Figura 14 -	- Eletroferograma da reação de SNaPshot do Multiplex R
Figura 15 -	- Representação esquemática dos <i>loci</i> DYS389 e DYS385
Figura 16	- Eletroferograma do perfil genético da amostra TDF_113 obtido através da
	tinagem com o kit PowerPlex® Y23

Figura 17 - Exemplos de microvariantes em três amostras, detectados usando o kit
PowerPlex® Y23
Figura 18 - Exemplos de alelos nulos em duas amostras, detectados usando o kit PowerPlex®
Y2385
Figura 19 - Regiões de azoospermia AZFa, AZFb e AZFc em Yq, e Y-STR 86
Figura 20 - Amostra MARA83
Figura 21 - Exemplos de alelos extra em 3 amostras, detectados usando o kit PowerPlex®
Y2389
Figura 22 - MDS $plot$ das distâncias genéticas $F_{\rm ST}$ pareadas, calculadas entre a Terra do Fogo,
a Colômbia, o Equador, o Rio de Janeiro, o Maranhão e outras populações da
América do Sul, bem como populações de referência europeia, africana e nativo-
americana96
Figura 23 - Classificação das amostras miscigenadas sul-americanas, segundo a ancestralidade
dos haplogrupos
Figura 24 - Panorama da ancestralidade paterna em populações miscigenadas da América do
Sul116

LISTA DE TABELAS

Tabela 1.	Conjunto de marcadores Y-STR utilizados pela comunidade
	forense41
Tabela 2.	23 <i>loci</i> do kit PowerPlex® Y23
Tabela 3.	Reagentes e volumes utilizados por amostra na preparação da mistura de reação
	para amplificação com o kit PowerPlex® Y2360
Tabela 4.	Condições termocíclicas para a amplificação com o kit PowerPlex®
	Y2360
Tabela 5.	Marcadores Y-SNP e seus respectivos multiplexes (continua)63
Tabela 6.	Conjuntos de oligonucleotídeos das PCR multiplexes realizadas66
Tabela 7.	Condições termocíclicas utilizadas para amplificações dos multiplexes e
	monoplex deste estudo
Tabela 8.	Conjuntos de oligonucleotídeos SBE das multiplexes (continua)72
Tabela 9.	Microvariantes encontradas nas amostras deste estudo
Tabela 10.	Alelos nulos encontrados nas amostras deste estudo
Tabela 11.	Alelos extra encontrados nas amostras deste estudo
Tabela 12.	Diversidade haplotípica para 23 loci Y-STR em diferentes populações
	miscigenadas sul-americanas 93
Tabela 13.	Matriz das distâncias genéticas de $F_{\rm ST}$ entre todos os pares de populações,
	juntamente com os correspondentes valores de probabilidade de não diferenciação
	obtidos para 10.000 permutações
Tabela 14.	Diversidade de haplogrupos em populações miscigenadas sul-americanas, bem
	como suas referências. 98
Tabela 15.	Frequência dos haplogrupos no Equador, Paraguai e Terra do
	Fogo
Tabela 16.	Frequência de haplogrupos atribuídos pelo software
	NEVGEN
Tabela 17.	Estimativa das taxas de mutação para 33 Y-STRs e intervalos de confiança
	correspondentes

Tabela 18.	Número de transmissões alélicas pai-filho analisadas considerando intervalos de
:	idade paterna (no momento do nascimento do filho) e a taxa de mutação
	correspondente 122
Tabela 19.	Número de transferências de alelos (Transf) e mutações (Mut) observadas
1	compatíveis com alterações envolvendo um número inteiro de repetições, bem
1	como número de mutações compatível com ganhos ou perdas, considerando as
;	seguintes categorias para o alelo paterno: alelo modal, mais curto ou mais longo
	que o alelo modal124

LISTA DE ABREVIATURAS E SIGLAS

A Adenina
C Citosina

CEISH-UDLA Comitê de Ética em Human Research da Universidad de

las Américas

DNA Deoxyribonucleic acid

DTT Ditiotreitol

dNTPs Desoxirribonucleotídeos Fosfatados

EAAF Equipo Argentino de Antropología Forense

FTA Flinders Technology Associates

FTDNA Family Tree DNA

G Guanina

GHEP-ISFG Grupo de Trabalho de Língua Espanhola e Portuguesa

da Sociedade Internacional de Genética Forense

HD Haplotype diversity

IDEAus Instituto de Diversidad y Evolución Austral

LDD Laboratório de Diagnósticos por DNA

MDS MultiDimensional Scaling

MilliQ Água desionizada

MPS Sequenciamento massivo paralelo

MSY Male-specific Y chromosome region

mtDNA DNA mitocondrial

NRY Non-recombining Y chromosome region

PAR Pseudoautosomal Region

PCR Polymerase Chain Reaction

pH Potencial hidrogeniônico

RFU Relative Fluorescence Units

RNA Ribonucleic acid

SBE Single Base Extension

SNP Single Nucleotide Polymorphism

SMM Single Step mutation model

SSC Solução salina de citrato de sódio

STR Short Tandem Repeat

SWGDAM Scientific Working Groupon DNA Analysis Methods

T Timina

UERJ Universidade do Estado do Rio de Janeiro

YBP Anos antes do presente

YCC Y Chromosome Consortium

YHRD *Y-STR Haplotype Reference Database*

Y-SNP SNP específico do cromossomo Y

Yp Braço curto do cromossomo Y

Yq Braço longo do cromossomo Y

Y-STR STR específico do cromossomo Y

LISTA DE SIMBOLOS

°C Grau Celsius

% Porcentagem

± Mais ou menos

cm Centímetro

 $F_{\rm ST}$ Fixation index

Km² Kilometro quadrado

M Concentração molar

Mb Megabase

min Minuto

MgCl₂ Cloreto de magnésio

mL Mililitro

mm² Milímetro quadrado

mM Milimolar

N Número

ng Nanograma

Nm Nanômetro

p-valor Valor de probabilidade

pb Par de bases

pg Picograma

rpm Rotações por minuto

S Segundo

Ml Microlitro

Mm Micromolar

SUMÁRIO

INTRODUÇÃO	15
1 JUSTIFICATIVA	50
2 OBJETIVOS	51
2.1 Objetivo geral	51
2.2 Objetivos específicos	51
3 MATERIAL E MÉTODOS	53
3.1 Amostras populacionais	53
3.2 Métodos de extração de DNA	55
3.2.1 Extração orgânica de DNA.	55
3.2.2 Extração de DNA por resina Chelex	56
3.2.3 Extração de DNA por PureLink TM Genomic DNA Mini Kit (<i>ThermoFisher Scienti</i>)	<u>fic)</u> 57
3.3 Análise de marcadores do cromossomo Y	57
3.3.1 Genotipagem dos <i>loci</i> STR do cromossomo Y utilizando o kit PowerPlex® Y23	58
3.3.2 Genotipagem dos marcadores Y-SNP utilizando o kit SNaPshot TM Multiplex (Appl	<u>ied</u>
Biosystems TM)	60
3.4 Análises estatísticas	77
4 RESULTADOS E DISCUSSÃO	80
4.1 Análise populacional de haplótipos definidos por Y-STRs	80
4.1.1 Alelos microvariantes observados nas populações amostradas	81
4.1.2 Alelos nulos observados nas populações amostradas	84
4.1.3 Alelos extra (duplicações) observados nas populações amostradas	89
4.1.4 <u>Diversidade haplotípica</u>	92
4.1.5 <u>Análise de distâncias genéticas (F_{ST})</u>	94
4.2 Análise populacional de haplogrupos definidos por Y-SNPs	97
4.2.1 <u>Diversidade de haplogrupos</u> .	98
4.2.2 Frequência e origem dos haplogrupos encontrados.	99
4.3 Ancestralidade paterna das populações da América do Sul	112
4.3.1 Ancestralidade das populações estudadas.	112
4.3.2 Panorama da América do Sul	114
4.4 Determinação de taxas de mutação em Y-STRs em pares pai-filhos	118
4.4.1 Modelo mutacional passo a passo (Single Step mutation model)	119
4.4.2 <u>Variação da taxa de mutação com a idade do pai</u>	121

4.4.3 <u>Variação da taxa de mutação com o número de repetições</u>
CONCLUSÕES125
REFERÊNCIAS
ANEXO A - Documento da aprovação pelo Comitê de Ética (Paraguai)157
ANEXO B - Documento da aprovação pelo Comitê de Ética em Human Research da
Universidad de las Américas - CEISH-UDLA 2017-0301158
ANEXO C - Documento da aprovação pelo Comitê de Ética em Investigações
Biomédicas (IMBICE)
ANEXO D - Documento da aprovação pelo Comitê de Ética - Pontifica Universidade
Javeriana160
ANEXO E - Parecer do Comitê de Ética em Pesquisa do Hospital Pedro Ernesto162
ANEXO F - Documento da aprovação pelo Comitê de Ética em Pesquisa - Hospital
Pedro Ernesto (Rio de Janeiro, Brasil)
ANEXO G - The Ancestry of Eastern Paraguay: A Typical South American Profile with
a Unique Pattern of Admixture (artigo publicado)164
ANEXO H Tierra Del Fuego: What Is Left from the Precolonial Male Lineages?
(artigo publicado)165
ANEXO I - Male lineages in Brazilian populations and performance of haplogroup
prediction tools (artigo publicado)
ANEXO J - Microsatellites Mutation Modeling Through the Analysis of the Y-
Chromosome Transmission: results of a GHEP-ISFG Collaborative Study
(artigo publicado)

INTRODUÇÃO

A Genética de Populações é a área da ciência que estuda a diversidade genética populacional, ao nível de indivíduos que compõem uma população e entre populações distintas, bem como sua variação ao longo do tempo. Uma vez observado que a distribuição das frequências alélicas e genotípicas numa população ao longo das gerações é alterada é possível apontar a atuação de forças evolutivas. A mutação é a força evolutiva que produz a fonte primária da variabilidade genética entre as populações. Embora ocorram espontaneamente, as taxas de mutação são lentas e não são suficientes para produzir mudanças rápidas no *genepool* de uma população. Após a criação da variação pela mutação, a diversidade genética é moldada e distribuída por outros processos, tais como seleção, deriva gênica e fenômenos de migração. Logo, tais forças evolutivas em conjunto são responsáveis pela variabilidade genética que se observa nas populações atuais e sua distribuição ao nível mundial.

Populações mais distantes tendem a acumular mais diferenças do que populações vizinhas, o que permite a caracterização populacional através de estudos genéticos (CLARCK e HARTL, 2007; MARTINS, 2007). Assim, a diversidade populacional é reflexo das diferenças acumuladas ao longo do tempo entre as populações e, por conseguinte, caracteriza as relações evolutivas e de ancestralidade (BAMCHAD *et al.*, 2004; CLARCK e HARTL, 2007).

Em genética de populações, a migração consiste no deslocamento de indivíduos de uma área geográfica para outra que já é habitada por indivíduos de um grupo distinto, criando-se um fluxo gênico entre estas populações, com a formação de uma população com uma composição genética distinta (dita miscigenada), como por exemplo, o ocorrido nas populações contemporâneas americanas. A análise da variabilidade e estrutura genética das populações atuais possibilita compreender quais os eventos migratórios ocorridos ao longo do tempo e a sua importância na distribuição mundial contemporânea.

Portanto, o estudo dos padrões de variabilidade genética em populações humanas gera conhecimento sobre sua história, e fornece informações sobre a maneira como evoluíram até os dias de hoje (RAMACHANDRAN, 2010).

Populações da América do Sul

Nas últimas décadas, estudos com marcadores genéticos têm contribuído para corroborar ou complementar os fatos históricos acerca da origem do homem moderno. Diferentes estudos com marcadores autossômicos e de linhagem paterna (cromossomo Y) e materna (DNA mitocondrial) apontaram que as populações africanas possuem os níveis de diversidade mais elevados, e que essa diversidade diminui à medida que nos distanciamos deste continente (CAVALLI-SFORZA e FELDMAN, 2003; CAMPBELL e TISHKOFF, 2008; LÓPEZ, VAN DORP e HELLENTHAL, 2015). Estudos com dados craniométricos também apontaram uma origem africana do homem moderno (ou Homo sapiens) (MATSUMURA et al., 2019). Logo, a teoria mais aceita para a formação das populações modernas aponta uma origem dos primeiros grupos populacionais humanos há, aproximadamente, 200 mil anos atrás, na África, uma vez que as linhagens mais antigas se encontram neste continente e possivelmente são as ancestrais de todas as outras até agora encontradas (UNDERHILL et al., 2001; BELLWOOD e NESS, 2014). A partir deste continente, um pequeno grupo migrou e originou os primeiros europeus e asiáticos (FORSTER, 2004; PENA et al., 2009). Entretanto, ainda estão em debate detalhes sobre alguns dos padrões de dispersão humana (REYES-CENTENO et al., 2014).

A América foi o último continente a receber as populações humanas migrantes, entre 14.000 e 20.000 YBP (anos antes do presente), vindas da Ásia (MARTIN, 1973; FIEDEL, 2000; FOSTER, 2004; PENA et al., 2009; REGUEIRO et al., 2013; RAGHAVAN et al., 2015; LLAMAS et al., 2016). A travessia de grupos nômades asiáticos muito provavelmente ocorreu através de uma grande ponte de terra conhecida como Beríngia, entre a leste da Sibéria e o Alaska, na América do Norte, durante o último evento de glaciação no Pleistoceno (Figura 1) (CAVALLI-SFORZA, MENOZZI e PIAZZA, 1995; FIEDEL, 2000; BOURGEON et al., 2017). Este modelo conhecido como "Out of Beringia" (GOEBEL et al., 2008) sugere que, após o povoamento inicial da Beríngia por norte-asiáticos, esta população protoamericana se expandiu, migrando para o sul ao longo da costa e através de um corredor interno (GREENBERG et al., 1986).

Embora as rotas migratórias dos proto-americanos ainda não tenham sido totalmente definidas, uma das hipóteses é de que os grupos populacionais nortistas teriam dispersado ao longo da costa até ao Sul e, mais tarde, adentraram o continente sul-americano atravessando a Cordilheira dos Andes em diferentes latitudes (GOEBEL *et al.*, 2008). Alternativamente,

teriam ocorrido simultaneamente dois movimentos, um para a região da costa (oeste) e outro para o interior do continente (leste) o que, hipoteticamente, provocou uma separação dos grupos populacionais na região norte da América do Sul (ROTHHAMMER e DILLEHAY, 2009). Existe também debate quanto ao número de migrações envolvidas na colonização da América do Sul por povos nativos, sendo que os estudos mais recentes apontam para, pelo menos, dois grandes eventos migratórios (MORENO-MAYAR *et al.*, 2018).

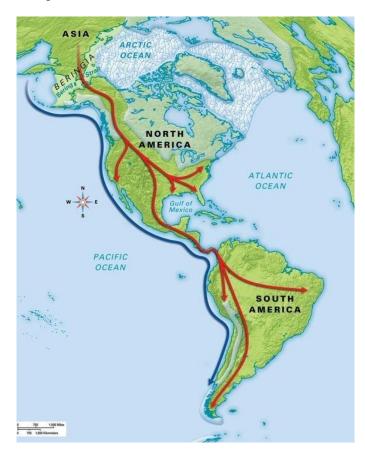


Figura 1. Entrada e dispersão do homem moderno nas Américas

Fonte: Adaptado de

http://resourcesforhistoryteachers.pbworks.com/w/page/123832947/Peopling%20of%20the%20Americas Legenda; Rotas migratórias dos proto-americanos pelo interior (em vermelho) e pela região costeira (em azul).

Hoje, a maioria dos pesquisadores apoiam a hipótese de que o assentamento humano inicial das Américas foi um processo relativamente rápido, apoiada principalmente pela datação de achados arqueológicos de populações proto-americanas na América do Norte e do Sul (ROEWER *et al.*, 2013). Por exemplo, o local Manis na América do Norte (WATERS *et al.*, 2011) tem uma idade estimada de 13.860-13.765 YBP e o sítio de Monte Verde no Chile

data de 14.220–13.980 YBP (DILLEHAY *et al.*, 2008). Outros locais de escavação na América do Sul, nas regiões do norte (Taima-Taima, Falcon, Venezuela), leste (Lagoa Santa, Minas Gerais, Brasil), oeste (Pikimachay, Vale Ayacucho, Peru) e sul (Los Toldos, Santa Cruz, Argentina), bem como na Amazônia (Pedra Pintada, Pará, Brasil), também indicam que populações humanas de alto padrão tecnológico foram espalhados por todo o continente por volta de 12.000 YBP e eram contemporâneos da cultura Clovis norte-americana (SALZANO e CALLEGARI-JACQUES, 1988; SILVERMAN e ISBELL, 2008).

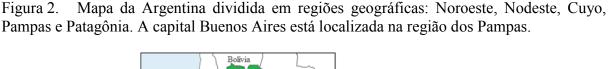
A miscigenação que marca as populações sul-americanas contemporâneas teve início com as expedições da conquista marítima europeia, no século XVI, com a chegada de grupos europeus nas Américas (DENEVAN, 2003). Além dos europeus, indivíduos africanos foram trazidos como mão de obra escrava a partir de diferentes locais da África.

Os muitos fluxos migratórios ao longo da história resultaram em uma alta diversidade genética das populações sul-americanas, compostas por diferentes grupos ancestrais, predominantemente, de três continentes: África, Europa e América (RUIZ-LINARES *et al.*, 2014; SANS, 2000; SALZANO e SANS, 2014). Logo, a composição genética e padrões de miscigenação das populações sul-americanas são altamente diversos. Dessa maneira, é importante a construção de bancos de dados genéticos que permitam capturar a alta diversidade e heterogeneidade genética presente nos diferentes grupos populacionais sul-americanos (QUINTANA-MURCI, KRAUSZ e MCELREAVEY, 2001).

Neste trabalho, foram estudadas várias populações da América do Sul. Para melhor compreensão do tema será realizada uma breve descrição dessas populações, com ênfase na história da sua formação e composição atual. Para cada uma das populações estudadas, será também apresentado um panorama dos trabalhos científicos que têm sido realizados no âmbito da genética populacional e suas principais descobertas.

1.1.1 <u>Argentina – Terra do Fogo</u>

A Argentina é o segundo maior país da América do Sul em território. Faz fronteira com o Paraguai e Bolívia ao norte, com o Brasil e Uruguai a nordeste e com o Chile a oeste. O país está dividido em 23 províncias e a sua capital, Buenos Aires, agrupadas em cinco regiões geográficas: Noroeste, Nordeste, Cuyo, Pampas, e Patagônia (Figura 2).



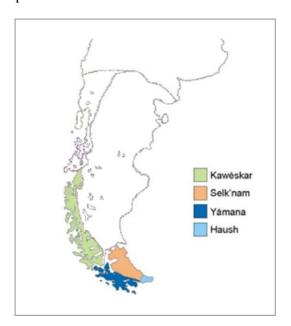


Historicamente, a Argentina recebeu duas ondas migratórias de europeus. A primeira, representada por espanhóis em sua maioria, ocorreu durante a era colonial (SEVINI et al., 2013). Desse modo, deu-se início ao primeiro episódio de miscigenação entre os nativos e migrantes ocidentais, fortalecido posteriormente pela introdução de escravos africanos no final do século 16. A segunda onda migratória de europeus, principalmente da Espanha e Itália, representados por mais de 76% do contingente migratório, chegou a Argentina entre 1856 e 1930 com o destino para as grandes cidades, dando origem a uma população urbana diversificada, com o componente ameríndio altamente diluído (PELLEGRINO, 2003; SALAS et al., 2008; BLANCO-VEREA et al., 2010; CORACH et al., 2010; SEVINI et al., 2013). Mais recentemente, a Argentina foi alvo de migrações de países vizinhos, da América do Sul, bem como de países da África, Ásia e Europa (CORACH et al., 2010).

O arquipélago da Terra do Fogo, localizado na extremidade sul da América do Sul, é formado por uma ilha principal (a Ilha Grande da Terra do Fogo) e um grupo de ilhas menores. O arquipélago é separado do continente sul-americano pelo estreito de Magalhães e dividido entre a Argentina e o Chile.

Os povos indígenas que primeiro habitaram a Terra do Fogo, há 10.000 anos, são conhecidos por Yámana (também conhecidos como Ona), Selk'nam, Manekénk (também conhecidos como Haush) e marginalmente os Kawésqar (Figura 3) (BORRERO, 1997; VILA et al., 2007; RODRÍGUEZ e HORLENT, 2016; SOTO, 2019). Os Haush concentravam-se no extremo sudeste, junto com os Selknam, enquanto os Onas viviam nas margens da Ilha Grande da Terra do Fogo e nas demais ilhas e ilhotas ao sul (RODRÍGUEZ e HORLENT, 2016).

Figura 3. Distribuição dos povos indígenas na Terra do Fogo Yámana, Selk'nam, Haush e marginalmente os Kawésqar.



A grande ilha da Terra de Fogo foi território fueguino dos Selknam até a chegada dos exploradores de ouro na década de 1880. A estes se vieram juntar grandes fazendeiros e missionários salesianos, bem como outros migrantes estimulados pela concessão de terras a empresas pecuárias (LUIZ e SCHILLAT, 1998; MATEO, 1998; BASCOPÉ, 2010; RODRIGUEZ e HORLENT, 2016).

Aos poucos, a maioria dos nativos foi sendo exterminada, devido a: (i) confrontos em prol da resistência dos Selknam à colonização do território por imigrantes, para a criação de ovelhas; (ii) perseguição dos nativos, contratados a serviço das propriedades nascentes de imigrantes ingleses e croatas; e (iii) doenças exógenas, especialmente a tuberculose e sarampo (RODRIGUEZ e HORLENT, 2016; HARAMBOUR, 2018). Os indígenas que sobreviveram aos conflitos foram forçados a se deslocar da região nativa, abandonando seus costumes e

hábitos. Além disso, sobreviventes foram transferidos para missões, onde eles tiveram que aprender ovinocultura, serração de madeira e comércios, abandonando a caça de animais (RODRIGUEZ e HORLENT, 2016).

Missionários anglicanos liderados por Thomas Bridges chegaram ao sul da ilha no território Yagán. Em 1884, o local foi estabelecido pelas autoridades argentinas como a cidade de Ushuaia, "baía que olha para o oeste" na língua Yagán (RODRIGUEZ e HORLENT, 2016). O território, anteriormente sede de missões, foi utilizado como sede política com o objetivo de facilitar o controle da população indígena local (BALDACCHINO, 2013).

O censo nacional de 1895 em Ushuaia, único assentamento indígena da Terra do Fogo neste momento, contava com 39 casas, 39 famílias e 313 habitantes indígenas (BALDACCHINO, 2013). Nos anos seguintes, o assentamento e crescimento populacional foi escasso, o que estimulou o governo nacional a desenhar um projeto para elevar a Terra do Fogo a uma colônia penal, com a primeira penitenciária estabelecida em 1902 (BALDACCHINO, 2013).

Atualmente, existem duas comunidades nativas na Terra do Fogo. A comunidade Rafaela Ishton, em Rio Grande, possui o estatuto legal e título de propriedade da comunidade, desde 2011, e é formada por membros dos povos Selknam e Haus. A segunda comunidade, embora ainda esteja em processo de conformação, reside em Ushuaia, é Yagán e foi denominada de Paiakoala (RODRIGUEZ e HORLENT, 2016).

Têm sido publicados diversos estudos com marcadores STR autossômicos e do cromossomo Y, bem como DNA mitocondrial em populações nativas e miscigenadas argentinas (BLANCO-VEREA et al., 2006; TOSCANINI et al., 2006 e 2007; BAILLIET et al., 2009; CORACH et al., 2010; PAURO et al., 2013; REGUEIRO et al., 2013; ROEWER et al., 2013; PURPS et al., 2014; TOSCANINI et al., 2016; GARCÍA et al., 2018; CASTAGNOLA et al., 2019; CAPUTO et al., 2019; PAROLIN et al., 2019; TAMBURRINI et al., 2021).

Na análise da composição genética de populações miscigenadas argentinas residentes em áreas urbanas, foram observados haplótipos majoritariamente de origem europeia (BAILLIET *et al.*, 2009; TOSCANINI *et al.*, 2016; CAPUTO *et al.*, 2019). Da mesma maneira, outros estudos com marcadores do cromossomo Y mostraram uma prevalência de linhagens europeias na Argentina, bem como na região da Patagônia, exceto nas províncias do noroeste (SALAS *et al.*, 2008; CORACH *et al.*, 2010; BAILLIET *et al.*, 2011; JURADO MEDINA *et al.*, 2015; CASTAGNOLA *et al.*, 2019; TAMBURRINI *et al.*, 2021).

A maioria das populações nativas argentinas apresenta valores maiores que 80% de ancestralidade paterna nativa (SALA *et al.*, 2010; SEVINI *et al.*, 2013; SALA e CORACH *et al.*, 2014; SALA *et al.*, 2018). No entanto, em alguns grupos nativos foi detectada a presença de linhagens masculinas europeias, embora com diferentes graus de mistura (SALAS *et al.*, 2009; TOSCANINI *et al.*, 2011), sendo predominantes nos grupos étnicos Mapuche, Diaguita e Colla, em detrimento do componente genético ameríndio (BLANCO-VEREA *et al.*, 2010).

1.1.2 Brasil – Rio de Janeiro e Maranhão

Considerado o maior país em extensão continental da América do Sul, o Brasil é dividido em cinco principais regiões geopolíticas (Norte, Nordeste, Centro-Oeste, Sudeste e Sul). O Brasil faz fronteira com todos os outros países sul-americanos, exceto Chile e Equador (Figura 4).

No início do século XVI, cerca de 500.000 europeus chegaram ao Brasil, vindos principalmente de Portugal (SALZANO e FREIRE-MAIA, 1970; NEVES e PUCCIARELLI, 1991; ALENCAR, RAMALHO e RIBEIRO *et al.*, 1996; PENA *et al.*, 2011), com a finalidade de explorar os recursos naturais e ocupar o território que pertencia aos nativos. Na época, estima-se que milhões de indígenas viviam no Brasil, com cerca de um terço habitando seu litoral, o que atualmente chega em torno de 300 mil em todo o território nacional (IBGE, 2007). O primeiro contato de mistura envolveu o acasalamento entre homens europeus e mulheres ameríndias, uma vez que a imigração de mulheres europeias durante os primeiros séculos da colonização no Brasil foi insignificante (RIBEIRO 1995; FAUSTO, 1996).

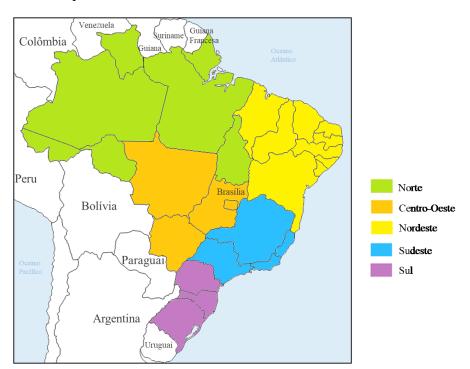


Figura 4. Mapa do Brasil dividido em regiões geopolíticas: Norte, Nordeste, Centro-Oeste, Sudeste e Sul. A capital Brasília está localizada no Centro-Oeste.

Relatos históricos mostram que os grupos dominantes na costa atlântica brasileira eram representados pelos grupos Tupiniquim e Tupinambá, ambos falantes do Tupi (FAUSTO, 2010; DA SILVA, 2019), que chegaram à costa leste depois que deixaram o noroeste da bacia amazônica há mais de 2.000 YBP (BROCHADO, 1984; SILVA *et al.*, 2020). Além destes, os não falantes Tupi, reconhecidos como Tapuias, também residiam na região do Rio de Janeiro (IBGE, 2007).

A população ameríndia sofreu uma redução populacional intensa, a partir do contato com o colonizador europeu, decorrente dos conflitos armados e epidemias. Nas décadas seguintes à conquista europeia, os escravos africanos, majoritariamente do sexo masculino, foram trazidos de diferentes regiões da África para o Brasil, como mão-de-obra nas fazendas de cana de açúcar, nas casas de família portuguesas, e posteriormente nas minas de ouro e diamantes e nas plantações de café (FAUSTO, 1996; IBGE, 2007).

Sendo o país que mais importou mão de obra africana das Américas, o Brasil recebeu nas cidades de Salvador e Rio de Janeiro o maior influxo de escravos africanos, vindo principalmente das regiões da Guiné e de Angola, respectivamente (FAUSTO, 1996; KIMURA *et al.*, 2017). De fato, estima-se que cerca de quatro milhões de escravos entraram

no país durante o período do tráfico de escravos (1538 a 1850) (IBGE, 2007; KIMURA *et al.*, 2017).

A migração europeia foi incentivada pelo governo brasileiro (que tornou-se independente no século XIX) com o objetivo de expandir os cafezais pelo Sudeste e pela necessidade de colonização efetiva da região Sul (SALZANO e FREIRE-MAIA, 1967).

Embora Portugal tenha sido a principal fonte de imigrantes europeus no Brasil, a partir de 1808 outros indivíduos europeus migraram para o país, como por exemplo, espanhóis, italianos e alemães, bem como sírios, libaneses e japoneses (SALZANO e FREIRE-MAIA 1967; CURTIN, 1969; RIBEIRO, 1995; IBGE, 2007).

O Rio de Janeiro foi a capital do Brasil durante 197 anos (1763 a 1960). A importância do local para o domínio colonial se refletiu como principal porto para a entrada de escravos e migrantes estrangeiros, bem como o transporte das riquezas brasileiras. De fato, a região foi a localização de diferentes empreendimentos econômicos que enriqueceram a colônia e futuramente a pátria independente, tais como, o comércio do pau-brasil, o transporte do ouro mineiro, e o cultivo do açúcar e do café. E, assim, indivíduos de diferentes populações transitaram servindo como mão de obra ou em busca de melhores oportunidades. Assim como no restante do Brasil, os migrantes portugueses compuseram a maioria da população fluminense. Nos anos seguintes, os suíços chegaram na região das serras fundando a cidade de Nova Friburgo, da mesma maneira que os alemães permaneceram principalmente em Petrópolis. Além destas populações, após a abolição da escravidão, migrantes italianos e espanhóis foram incentivados pelo governo brasileiro a migrar para a região Sudeste e assim contribuíram para a diversidade étnica do Rio de Janeiro (ALENCAR, RAMALHO e RIBEIRO *et al.*, 1996; IBGE, 2007).

Contrapondo o restante do Brasil, no Maranhão ocorreu um domínio colonial francês no século XVII. Este domínio foi extensível a outros territórios da linha do Equador, tendo ficado conhecido como França Equinocial. Apenas em 1615 o Maranhão foi conquistado por portugueses (MEIRELES, 2015). Vinte anos depois, a ilha de São Luiz recebeu uma onda de holandeses, na tentativa de conquistar o Maranhão, mas sem sucesso.

Diversos estudos utilizando marcadores uniparentais e autossômicos demostraram a heterogeneidade da população brasileira, resultado do cruzamento interétnico entre europeus, africanos e ameríndios (ALVES-SILVA *et al.*, 2000; SANTOS *et al.*, 2010; PALHA *et al.*, 2012; MANTA *et al.*, 2013; OLIVEIRA *et al.*, 2014; FIGUEIREDO *et al.*, 2016; RESQUE *et al.*, 2016; STANGE *et al.*, 2019; AMBROSIO *et al.*, 2020; JANNUZZI *et al.*, 2018 e 2020; PENA *et al.*, 2020). Estudos com marcadores Y-STR apontaram que, nas regiões do Sudeste,

do Centro-Oeste e do Sul, o componente europeu se sobrepõe aos restantes (PALHA *et al.*, 2012; RESQUE *et al.*, 2016), embora a região Sudeste também tenha uma contribuição de ancestralidade africana significativa (RESQUE *et al.*, 2016). Além disso, nas regiões Norte e Nordeste observa-se uma grande contribuição ameríndia e africana, respectivamente (PALHA *et al.*, 2012; RESQUE *et al.*, 2016).

Outros estudos relativos aos padrões de ancestralidade paterna mostraram uma predominância de linhagens europeias nas populações brasileiras, semelhante aos resultados de outras populações miscigenadas sul-americanas (AZEVEDO *et al.*, 2009; OLIVEIRA *et al.*, 2014; FIGUEIREDO *et al.*, 2016; RESQUE *et al.*, 2016; JANNUZZI *et al.*, 2020; SCHAAN *et al.*, 2020).

O grau elevado de miscigenação brasileiro reforça a importância de bancos de dados genéticos nacionais que capturem a diversidade genética regional. No estudo de 2012, Palha e colaboradores não observaram subestrutura para 23 *loci* Y-STR em 17 diferentes populações miscigenadas das cinco regiões do país estudadas, o que foi atribuído à alta frequência de linhagens masculinas europeias em todas as regiões. Entretanto, um estudo posterior de Oliveira e colaboradores (2014), para um conjunto menor de 17 Y-STRs, mostrou uma diferença significativa entre as populações do Sudeste e Norte. Este resultado foi corroborado pelos resultados da análise de Y-SNP.

1.1.3 Colômbia

A Colômbia é um país sul-americano que faz fronteira com a Venezuela, Equador, Peru e Brasil. O país possui 32 departamentos e um distrito capital e é subdividido em cinco regiões naturais: Caribe, Andes, Orinoquía, Amazônia, Pacífico (Figura 5) e região Insular.

A população colombiana é majoritariamente miscigenada, geneticamente diversa e multicultural. Os padrões de mistura genética que se observam atualmente resultaram não somente do encontro de grupos nativos com europeus, e em seguida populações africanas, como também de movimentos migratórios que caracterizaram os séculos XX e XXI.

Antes da chegada dos colonizadores europeus, o território colombiano foi ocupado por diversos grupos nativos independentes entre si, geralmente em guerra, com línguas e costumes distintos (RESTREPO, 1985). Dentro destes, é possível destacar os Chibchan, Carib

e Arawakan, na costa do Atlântico, Chocoan na costa do Pacífico e Paezan, Barbacoan e Quechua na região Andina Meridional (ARANGO e SANCHÉZ, 1998 e 2004; DANE, 2007).

Os Muíscas (chibchas) viviam em federações independentes e foram considerados um dos grupos nativos colombianos mais avançados, embora não tenham alcançado o título de civilização como Maias e Astecas (RESTREPO, 1985). Cultivavam milho, batata e algodão, no planalto central colombiano, próximo a capital Bogotá.

O grupo nativo Tairona também estabeleceu algum nível de organização social. São conhecidos pela construção de caminhos de pedra e cerâmicas. Viviam no extremo nordeste colombiano, na região do Caribe, ocupando um território que corresponde aos atuais departamentos de Magdalena, Guajira e Cesar, nos leitos da Serra Nevada de Santa Marta (REICHEL-DOMATOFF, 1953). Embora a princípio tenham estabelecido relações comerciais com os europeus recém-chegados, liderados por Fernando González de Oviedo, em 1498, os Tairona compuseram forças armadas resistentes a colonização espanhola, principalmente após a fundação da cidade de Santa Marta (REICHEL-DOMATOFF, 1953). Além da captura para trabalho forçado e extermínio de nativos Tairona, após os conflitos, os indivíduos sobreviventes foram obrigados a se deslocar para pontos isolados da Serra Nevada de Santa Marta (CAYCEDO, 1986). Seus descendentes são conhecidos como Koguis.

Atualmente, existem 87 grupos de populações indígenas colombianas, tais como os Kogui e Arzarios residentes nas regiões da Serra Nevada de Santa Marta (MIDDLETON *et al.*, 2000; DANE, 2007), habitando em zonas rurais e pequenas aldeias do país ou reservas indígenas, e uma pequena minoria em cidades urbanas (DANE, 2007).

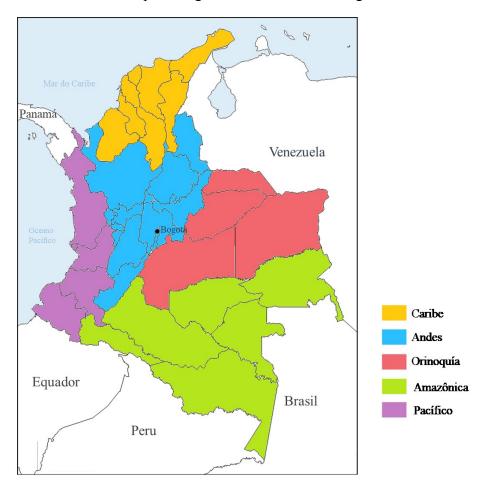


Figura 5. Mapa da Colômbia dividida em regiões naturais: Caribe, Andes, Orinoquía, Amazônica e Pacífico. A capital Bogotá está localizada na região dos Andes.

Os espanhóis estabeleceram, na costa caribenha, as duas principais bases coloniais nas Américas: Santa Marta, fundada em 1525; e Cartagena das Índias, fundada em 1533. Posteriormente, com o avanço interiorano, foram fundadas as cidades de Popayán, em 1536, e Santa Fé, atual Bogotá, em 1538.

Como em outras populações sul-americanas, o terceiro componente populacional colombiano é resultante da introdução de africanos, como mão de obra escrava, pelo colonizador europeu. As regiões que receberam os maiores contingentes de escravos na Colômbia foram as regiões do pacífico, como Chocó, logo abaixo do Panamá, e a costa Caribe. Cartagena, capital de Bolívar, foi centro do tráfico de escravos na Colômbia, durante os séculos XVII e XVIII (ROMERO et al., 2008; NOGUERA et al., 2013). Os escravos fugidos de Cartagena se estabeleceram na zona rural, fundando a comunidade negra, conhecida como Palenque de San Basílio (FRIEDEMANN e ROSSELLI, 1983; NOGUERA et al., 2013; ANSARI-POUR et al., 2015; MARTINEZ et al., 2020). Esta comunidade resistiu com sucesso a vários ataques espanhóis e tornou-se uma das primeiras comunidades de

ancestralidade africana oficialmente livres em toda a América (FERRARI, 2012; MARTINEZ *et al.*, 2020).

Os poucos estudos populacionais realizados na Colômbia revelaram diferentes padrões de mistura e alta heterogeneidade regional (ACOSTA *et al.*, 2009; ROJAS *et al.*, 2010; NOGUERA *et al.*, 2013; YUNIS e YUNIS, 2013; OSSA *et al.*, 2016; MORALES *et al.*, 2018; XAVIER *et al.*, 2015).

Estudos populacionais com marcadores do cromossomo Y mostraram que as regiões do Pacífico e da Costa Caribenha apresentam os maiores índices de ancestralidade africana (NOGUERA *et al.*, 2013; MORALES *et al.*, 2018). Os maiores índices de ancestralidade nativo americana por sua vez são observados na região da Amazônia e nas áreas rurais do sudoeste e norte colombiano (ROJAS *et al.*, 2010; OSSA *et al.*, 2016; MORALES *et al.*, 2018).

Em resumo, a ancestralidade paterna europeia foi observada predominantemente em populações urbanas (ROJAS *et al.*, 2010; NOGUERA *et al.*, 2013; MORALES *et al.*, 2018). Contrariamente, os estudos usando mtDNA em populações urbanas indicaram a predominância de ancestralidade nativa (ROJAS *et al.*, 2010; YUNIS e YUNIS, 2013). Além disso, as frequências mais baixas de ancestralidade nativo americana foram encontradas nos departamentos da costa do Caribe e na região do Pacífico, onde a ancestralidade africana é maior (YUNIS e YUNIS, 2013).

1.1.4 Equador

O Equador está localizado no noroeste da América do Sul sendo limitado ao norte pela Colômbia, ao leste e sul pelo Peru, e a oeste pelo Oceano Pacífico. O país é politicamente dividido em 23 províncias continentais e uma insular, distribuídas em 4 regiões naturais: Andes, Pacífico (Litoral), Amazônia e região insular (Ilhas Galápagos) (Figura 6).

A história populacional do Equador é resultado de vários eventos migratórios pré e pós-colombianos. A população equatoriana, multiétnica e pluricultural, é composta atualmente por três grupos étnicos principais: mestiços, nativo-americanos e afro-equatorianos.



Figura 6. Mapa do Equador dividido em regiões naturais: Andes, Pacífico, Amazônia, e a região insular (Ilhas Galápagos). A capital Quito está localizada na região andina.

A conquista espanhola promoveu uma redução nas populações ameríndias residentes, que foram acometidas por conflitos, doenças infecciosas e escravidão, estabelecendo a era colonial no Equador (ELLIOTT, 1984; AYALA, 2008). Além disso, assim como o ocorrido em outras colônias americanas, o Equador recebeu a imigração de milhões de africanos como mão de obra escrava sob domínio colonial espanhol (ELLIOTT, 1984; AYALA, 2008).

De acordo com o último senso nacional, o maior grupo étnico equatoriano corresponde aos mestiços autodeclarados, que constituem pouco mais de 72,3% da população (SÁNS, 2000; INEC, 2010). Este grupo caracteriza a população urbana encontrada em todas as cidades e são descendentes de ameríndios e europeus que chegaram durante todo o período colonial (BAEZA *et al.*, 2007; GONZÁLEZ-ANDRADE *et al.*, 2007 e 2009; SANTANGELO *et al.*, 2017; TOSCANINI *et al.*, 2018), bem como imigrantes recentes do século XX (AYALA, 2008).

O grupo étnico dos nativos é o segundo maior em tamanho populacional e representa menos de 10% da população (GONZÁLEZ-ANDRADE *et al.*, 2011; GAVIRIA *et al.*, 2013; TOSCANINI et al., 2018). Em 1470, o império inca conquistou a região equatoriana, que anteriormente era ocupada por diferentes grupos nativo-americanos. Os nativos equatorianos atuais são descendentes de pré-incas e incas que avançaram no domínio da região (AYALA, 2008).

As populações nativas ameríndias equatorianas compreendem um total de 14 grupos multiculturais que residem principalmente nas regiões andina e amazônica: Kichwa, Waorani, Secoya, Siona, A'I Cofan, Shuar, Achuar, Shiwiar, Zapara, Andoanos, Chachi, Awà, Tsa'Chila e Epera (GONZÁLEZ-ANDRADE *et al.*, 2007 e 2011; VILLAESCUSA *et al.*, 2021).

Os nativos falantes Kichwa representam cerca de 92% da população nativo-americana no Equador (Instituto nacional de estadística y censos (INEC, 2001; CODENPE, 2002; SINDEPE, 2003). A língua Kichwa inclui famílias de Línguas dos povos nativo-americanos andinos, portanto, distribuindo-se principalmente ao sul da Colômbia e Equador, Bolívia, Peru, e ao norte da Argentina e do Chile.

Os grupos nativos equatorianos não falantes Kichwa, por exemplo, são representados pelos Waoranis da Amazônia, falantes da língua Wao que aparentemente não tem relação com nenhuma outra da região (PEEKE, 1973), e os Tsáchilas (ou Colorados) residentes da costa equatoriana, que falam a língua Tsafiki da família linguística Barbacoa de nativos da Colômbia e do Equador (BENÍTEZ, 1993).

Os Waoranis são encontrados em pequenos grupos isolados geográfica e culturalmente na Amazônia, nas províncias de Pastaza, Napo e Orellana, e mantém a cultura nômade de caçadores e coletores (CODENPE, 2002; GONZÁLEZ-ANDRADE *et al.*, 2009; CARDOSO *et al.*, 2012; VILLAESCUSA *et al.*, 2021).

Por último, o terceiro maior grupo étnico corresponde aos afro-equatorianos, que são descendentes de africanos escravizados, e representam cerca de 7,2% da população equatoriana (INEC, 2001). Este grupo reside atualmente ao norte na província litorânea de Esmeraldas, e na região andina no Valle del Chota, entre as províncias de Imbabura e Carchi (BENÍTEZ, 1993; ROITMAN, 2008; GONZÁLEZ-ANDRADE *et al.*, 2009; VÁSQUEZ e SALTOS GALARZA, 2014).

Estudos populacionais e forenses foram realizados em diferentes populações do Equador e revelaram padrões de mistura complexa e uma alta heterogeneidade (BAEZA et al., 2007; GONZÁLEZ-ANDRADE et al., 2009; GAVIRIA et al, 2013; REGUEIRO et al., 2013; ROEWER et al., 2013; HOMBURGER, et al., 2015; MEZZAVILLA et al., 2015; JOTA et al., 2016; SANTANGELO et al., 2017; TOSCANINI et al., 2018; ZAMBRANO et al., 2019; DI CORCIA et al., 2021; FLORES-ESPINOZA et al., 2021; VILLAESCUSA et al., 2021).

Alguns trabalhos descreveram a composição genética de populações nativas e revelaram a predominância dos haplogrupos Q e C paternos (REGUEIRO *et al.*, 2013;

ROEWER et al., 2013; HOMBURGER, et al., 2015; MEZZAVILLA et al., 2015; JOTA et al., 2016; SANTANGELO et al., 2017; TOSCANINI et al., 2018; BURGOS et al., 2019a e 2019b; ZAMBRANO et al., 2019; DI CORCIA et al., 2021; VILLAESCUSA et al., 2021). Por outro lado, as populações mestiças apresentam uma ascendência mista de nativo-americano e europeu, com um componente europeu maior (SANTANGELO et al., 2017). Entretanto, a análise de marcadores de ancestralidade autossômica em populações mestiças revelou uma proporção de mais de 50% de ancestralidade nativa (SANTANGELO et al., 2017; ZAMBRANO et al., 2019).

Quando a composição genética por regiões equatorianas foi analisada, observou-se que o componente nativo-americano estava mais presente na região amazônica, seguida pela região andina (TOSCANINI *et al.*, 2018). Por outro lado, o componente europeu estava concentrado principalmente no Pacífico, predominando nos grupos mestiços que residem nestas regiões (SANTANGELO *et al.*, 2017; TOSCANINI *et al.*, 2018; BURGOS *et al.*, 2019a e 2019b). Por fim, a prevalência do componente africano foi encontrada no noroeste do país, especificamente nas províncias de Esmeraldas e Imbabura (TOSCANINI *et al.*, 2018).

Finalmente, ao comparar os valores de diferenciação do cromossomo X e autossômico, o trabalho de Flores-Espinoza e colaboradores (2021) revelou uma miscigenação do Equador com viés sexual entre homens europeus e mulheres nativas americanas e africanas, bem como entre homens africanos e mulheres nativas americanas (FLORES-ESPINOZA *et al.*, 2021).

1.1.5 Paraguai

O Paraguai é um país sul-americano sem litoral, que se divide em duas regiões geográficas pelo rio Paraná, compreendendo 17 departamentos (Figura 7). A região Oriental ou Leste é conhecida como Paraneña, onde vive 98% da população. A região Ocidental ou Oeste, também conhecida como Chaco, é mais extensa e compreende em torno de 2% da população atual (HEBBLETHWAITE, 2014).

O Rio Paraná foi responsável pela primeira divergência entre os nativos americanos que ocuparam o território paraguaio, os Guaranis. Estes habitavam a região ocidental e viviam principalmente da caça e coleta, enquanto àqueles residentes da região oriental usufruíam da

agricultura (VÁZQUEZ, 2006). Além dos guaranis, viviam na região Ocidental, no mínimo, cinco grupos linguísticos diferentes (JERMYN e LIN, 2010).

Após a chegada dos colonizadores espanhóis ao território paraguaio, no século XVI, teve início um processo de mistura, que ocorreu principalmente pelo casamento entre homens europeus e mulheres americanas nativas (POTTTHAST-JUTKEIT, 1997).

A princípio, o colonizador espanhol formou alianças pacíficas com os indígenas Guaranis que viviam nas margens do rio Paraguai, mas, em seguida, implementou o sistema de "encomienda", onde os indígenas eram escravizados e obrigados a trabalhar no campo. Assim, como ocorreu no processo de exploração europeia dos outros impérios coloniais da América do Sul, escravos africanos foram trazidos ao Paraguai, embora em baixa proporção (RODRIGUEZ, 1997; JERMYN e LIN, 2010).

Visto que as riquezas minerais eram escassas, os colonizadores perderam o interesse no território, de maneira que a terra paraguaia se tornou uma região isolada, geograficamente, culturalmente e politicamente separada das sul-americanas restantes (JERMYN e LIN, 2010; LAMBERT, 2012).

Outro momento importante que afetou a população paraguaia ocorreu durante a Guerra da Tríplice Aliança, travada de 1864 a 1870, entre Argentina, Brasil e Uruguai contra o Paraguai. Durante a meia década de conflitos, a população paraguaia masculina sofreu uma redução de mais da metade (WARREN, 1985). Com o objetivo de restaurar a população, a imigração foi encorajada sob incentivos políticos. Dessa maneira, imigrantes Euroasiáticos (maioritariamente Italianos, Alemães, Franceses, Espanhóis, Coreanos e Japoneses) e de outras populações sul-americanas vizinhas entraram no país.

Bolívia

Brasil

Argentina

Região Ocidental

Região Oriental

Figura 7. Mapa do Paraguai: sua subdivisão nas regiões (Ocidental e Oriental) e suas fronteiras com outros países

Legenda: Mapa do Paraguai e sua subdivisão em duas regiões (ocidental e oriental), bem como suas fronteiras continentais com os países Argentina, Brasil e Bolívia, e a capital Asunción.

Atualmente, existem poucos trabalhos publicados que descrevem a composição genética do Paraguai, e dentre eles estão estudos com populações nativas do Chaco e miscigenadas da região leste (MARTINEZ-ESPIN et al., 2003; CATANESI, 2007; VULLO et al., 2016; RIBEIRO et al., 2018). Num estudo em que se analisou a região HVSI do mtDNA em grupos nativos Ayoreo da região do Chaco, verificou-se que aproximadamente 80% das amostras pertenciam ao haplogrupo C nativo-americano (DORNELLES et al., 2004). Por outro lado, para o grupo nativo Aché do leste paraguaio, todos os haplótipos observados pertenciam aos haplogrupos B e A nativo-americanos (SCHMITT et al., 2004). Portanto, ambos os estudos sugerem uma baixa variabilidade do mtDNA, uma vez que são encontrados poucos haplogrupos em cada grupo nativo estudado. Estudos mais recentes com marcadores STR autossômicos e do cromossomo Y mostraram uma homogeneidade genética das populações da região leste do Paraguai bem como uma proximidade com populações europeias (VULLO et al., 2016; RIBEIRO et al., 2018).

1.2 Marcadores genéticos

As variações na sequência de DNA entre indivíduos surgem por mutação e resultam em polimorfismos que podem ser empregados como marcadores em diversos campos da genética. Um *locus* é considerado polimórfico quando apresenta, no mínimo, dois alelos com frequência igual ou superior a 1% na população. Estudos genéticos empregam *loci* polimórficos como material de estudo, visto que estes possibilitam a diferenciação entre indivíduos e/ou populações, podendo ser aplicados em identificação humana, estudos populacionais, clínicos e genealógicos. Entretanto, sua utilização depende da localização cromossômica, de seu mecanismo de transmissão hereditária, do tipo de variação que apresentam e da distribuição da variabilidade intra e interpopulacional (BAYE, 2011).

Um marcador genético que apresenta diferenças apenas na sequência de nucleotídeos é classificado como polimorfismos de sequência e é resultado da substituição de um ou mais nucleotídeos em uma determinada região do DNA (LIU *et al.*, 2014). Alternativamente, os polimorfismos que apresentam diferenças no comprimento de segmentos de DNA entre indivíduos são conhecidos como polimorfismos de comprimento, e sua origem advém de inserções/deleções de um ou mais nucleotídeos na sequência do DNA (BUTLER, 2009; LIU, 2018).

Os polimorfismos de sequência mais frequentes são os SNPs (*Single nucleotide polymorphisms*). Os SNPs são caracterizados pela troca, em uma posição específica no genoma, de uma única base por outra e podem ser classificados em transições e transversões (LANDER *et al.*, 2001; BUDOWLE e DAAL, 2008; LAPPALAINEN *et al.*, 2019). As transições surgem após a troca entre duas bases pirimídicas (A por G ou vice-versa) ou entre duas bases púricas (C por T ou vice-versa). Por outro lado, as transversões resultam da troca entre bases púricas e pirimídicas, isto é, de C por G ou A, de T por A ou G, ou vice-versa (ROBERT e PELLETIER, 2018).

Os STRs (do inglês, *Short Tandem Repeats*) ou microssatélites são polimorfismos de comprimento que constituem, aproximadamente, 3% do total de variações das sequências de DNA do genoma humano (ELLEGREN, 2004; BUTLER, 2010; PLANZ e HALL, 2012). Esta classe de marcadores apresenta sequências repetitivas consecutivas *in tandem*, formadas por blocos de repetição de 2 a 6 pares de bases, localizadas ao longo do genoma (Tabela 1) (GYMREK *et al.*, 2017). As sequências repetitivas são variáveis entre indivíduos e conferem um grande poder de discriminação.

Os marcadores InDel (do inglês, *Insertion deletion polymorphism*) são polimorfismos de comprimento caracterizados por inserções ou deleções de um ou mais nucleotídeos (WEBER *et al.*, 2002; LAPPALAINEN *et al.*, 2019). Tal como os SNPs, estes marcadores apresentam taxas de mutação mais baixas que as encontradas para os STRs. Logo, estes marcadores tendem a apresentar uma menor diversidade intrapopulacional (ECKERT e HILE, 2009; XIE *et al.*, 2018).

Os marcadores autossômicos sofrem recombinação durante a formação de gametas, apresentando herança biparental, a qual contem informação de ambos os progenitores. Os marcadores de linhagem encontram-se em estado haploide e não sofrem recombinação. São encontrados no genoma mitocondrial e no cromossomo Y e caracterizam-se pela transmissão da mãe para todos os filhos (mtDNA) ou do pai para filhos do sexo masculino (Cromossomo Y). Dessa maneira, o cromossomo Y e o DNA mitocondrial são segmentos genômicos de herança uniparental, paterna ou materna, respectivamente, que permitem discriminar linhagens parentais em detrimento da identificação individual (BUTLER, 2010; BUTLER, 2014; COMAS *et al.*, 2018). Logo, os marcadores autossômicos são a principal fonte de informação genética, a fim de identificação individual, bem como de determinação de relações de parentesco.

Em estudos de identificação humana são recomendados os marcadores genéticos localizados em regiões não codificantes do genoma, por motivos éticos. Desta maneira, evitase a exposição de quaisquer outros tipos de informação além da identificação de um indivíduo, como por exemplo, a detecção de variantes em genes relacionados à predisposição a doenças (QUINTANA-MURCI, KRAUSZ e MCELREAVEY, 2001; AMORIM e BUDOWLE, 2016). Além disso, os marcadores localizados em regiões não codificantes estão sujeitos a menores pressões seletivas e, por conseguinte, podem apresentar maior diversidade.

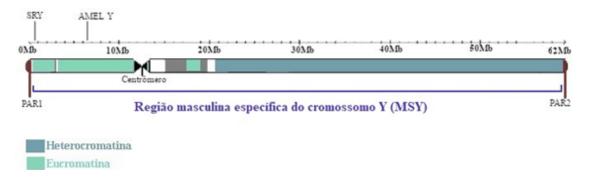
1.3 Cromossomo Y: estrutura e modo de transmissão

O cromossomo Y está presente em apenas uma cópia por célula nos indivíduos do sexo masculino com cariótipo convencional (46, XY). Embora seja o segundo menor cromossomo humano, possui em média 62 milhões de pares de bases (Mb) (MORTON, 1991; RHIE *et al.*, 2023) e é o portador dos genes responsáveis pelas características sexuais masculinas.

Em 95% do seu comprimento, o cromossomo Y não sofre recombinação com o cromossomo X, durante a meiose masculina. Esta região anteriormente era conhecida como região não-recombinante do cromossomo Y (NRY, do inglês, *Non Recombining region of the Y chromosome*). Atualmente, após a descoberta de uma elevada recombinação intracromossômica (ROZEN *et al.*, 2003; SKALETSKY *et al.*, 2003), esta região passou a designar-se por região masculina específica do cromossomo Y (MSY, do inglês, *Male Specific Y region*).

A região MSY é constituída por sequências repetitivas e é subdividida nas porções de eucromatina e heterocromatina (Figura 8). A porção de eucromatina, com 24 Mb de extensão, compreende porções do braço curto, centrômero e braço longo do cromossomo Y. Neste segmento, são encontradas sequências homólogas ao cromossomo X, sequências repetitivas e genes codificantes (SKALETSKY *et al.*, 2003; BACHTROG, 2013).

Figura 8. O cromossomo Y



Legenda: Cromossomo Y e sua extensão aproximada em megabases. Em destaque, a região MSY e as regiões PAR1 e PAR2, bem como as porções de heterocromatina e eucromatina e os genes da região de determinação sexual (SRY) e da amelogenina (AMELY).

Fonte: RÊGO, 2019. Modificada.

Em contrapartida, a porção de heterocromatina, com 40Mb de extensão, está presente em grande parte do braço longo do cromossomo Y e só possui 13 genes codificantes (QUINTANA-MURCI, KRAUSZ e MCELREAVEY 2001; ALI e HASNAIN, 2002; SINGH *et al.*, 2011; BACHTROG, 2013). No restante de sua extensão, esta zona apresenta sequências intracromossômicas extremamente repetitivas, também conhecidas como amplicônicas *in tandem* (consecutivas) ou em palíndromos.

Nas porções terminais dos braços curto (Yp) e longo (Yq) do cromossomo Y são observadas as regiões PAR1 (2,77 Mb) e PAR2 (329,5 kb) (do inglês, *Pseudoautosomal*

Regions 1 and 2), que possuem homologia a sequências do cromossomo X (RHIE et al., 2023). Durante a gametogênese masculina, essas regiões pareiam com o cromossomo X, estando sujeitas à recombinação (*crossing-over*), sendo por isso denominadas regiões pseudoautossomicas (QUINTANA-MURCI, KRAUSZ e MCELREAVEY, 2001; ALI e HASNAIN, 2002; MANGS e MORRIS, 2007; SINGH et al., 2011; BACHTROG, 2013).

Na última década, uma terceira região pseudoautossomica de 2,5 Mb foi designada PAR3. Esta região se originou de Xq21.3 quando uma região do cromossomo X sofreu duplicação e transposição em Yp11.2 (VEERAPPA *et al.*, 2013). Entretanto, a PAR3 foi encontrada apenas em alguns casos de patologia e em menos de 2% da população normal (VEERAPPA *et al.*, 2013).

Em estudos genéticos, os marcadores do cromossomo Y limitam-se àqueles localizados na região MSY (JORDE *et al.*, 2000; UNDERHILL *et al.*, 2001), em razão da sua presença exclusiva em indivíduos do sexo masculino e transmissão integral aos descendentes masculinos da mesma linhagem, a menos que ocorram mutações. Portanto, as sequências presentes na porção MSY apresentam variações, devido ao acúmulo de novas mutações (SANTOS *et al.*, 2000; ALI e HASNAIN, 2002). Logo, a variabilidade obtida com o acúmulo de mutações faz com que os cromossomos Y divirjam, resultando em linhagens paternas distintas.

1.3.1 Marcadores do Cromossomo Y

Os marcadores comumente usados na análise do cromossomo Y são os Y-STRs e os Y-SNPs. O termo Y-SNP refere-se a ambos os tipos de marcadores bialélicos, SNPs e InDels, localizados em regiões específicas do cromossomo Y. Estes marcadores podem ser usados em estudos de genética de populações, na definição de linhagens masculinas (JOBLING e TYLER-SMITH, 2003). Além disso, o estudo combinado destes marcadores permite a inferência de padrões de migração em populações humanas, mediados por indivíduos do sexo masculino.

Devido ao modo de transmissão e a ausência de recombinação, os marcadores do cromossomo Y são utilizados em casos de investigação de paternidade, quando não se tem acesso ao pai alegado ou a parentes próximos. Nestas situações, é possível recorrer a outros indivíduos aparentados por via paterna, uma vez que o perfil genético do cromossomo Y é

partilhado por todos os indivíduos de uma mesma patrilinhagem, mesmo que separados por várias gerações. A análise de Y-STRs em evidências forenses é especialmente útil em casos de agressão sexual, quando há mistura de material biológico da vítima, do sexo feminino, e do agressor, do sexo masculino (JOBLING, PANDYA e TYLER-SMITH, 1997; THOMPSON et al., 2012). Nestes casos, quando usados marcadores autossômicos tende a ocorrer a amplificação preferencial do componente majoritário da mistura, geralmente correspondente ao material biológico da vítima, não sendo possível recuperar o perfil genético do agressor. Já quando se usam marcadores do cromossomo Y, apenas o DNA de origem masculina será amplificado, obtendo-se assim ao perfil correspondente ao agressor, que pode ser comparado ao do suspeito. Além disso, em casos de agressão múltipla é possível identificar o número de perpetradores do sexo masculino.

Os Y-STRs apresentam taxas de mutação elevadas, aproximadamente de 10⁻³ por geração, o que permite o estudo de eventos de diversificação populacional recentes. O haplótipo corresponde ao conjunto de informações obtidas após a análise de um grupo de Y-STRs. Quanto aos Y-SNPs, por apresentarem uma taxa de mutação muito menor (aproximadamente 10⁻⁹ por geração), permitem estabelecer linhagens masculinas mais estáveis, denominadas haplogrupos, que agrupam diferentes haplótipos que divergem por um baixo número de mutações.

1.3.1.1 Short Tandem Repeats específicos do cromossomo Y

Os STRs apresentam taxas de mutação mais elevadas que as regiões não repetitivas, e assim resultam em valores altos de diversidade intrapopulacional (ECKERT e HILE, 2009; BUGOYE *et al.*, 2018). Por estas características, os STRs são os marcadores mais utilizados em genética forense (GUSMÃO *et al.*, 2005).

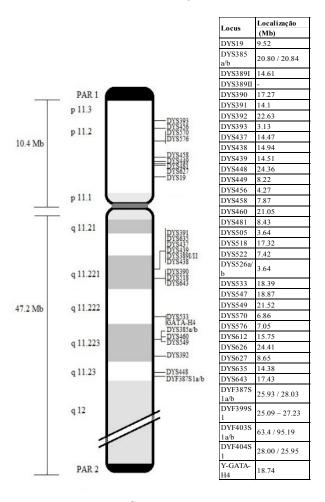
Nas últimas décadas houve um crescimento no número de marcadores do cromossomo Y (Figura 9) empregados em estudos populacionais e forenses. Isso deve-se a modernização dos métodos de análise envolvidos, que permitiu a análise simultânea de um número elevado de marcadores numa única reação multiplex (PLANZ e HALL, 2012).

Os Y-STRs podem ser observados em cópia única ou em múltiplas cópias numa sequência de DNA, ao longo da extensão do cromossomo Y. O Y-STR DYS385 corresponde a dois *loci* originados pela duplicação de uma mesma região ao longo do cromossomo. Os

alelos correspondentes a estes dois *loci* podem apresentar ou não um mesmo número de copias da unidade repetitiva.

Devido às altas taxas de mutação e ao elevado polimorfismo dos Y-STRs, aliados à rapidez e facilidade de análise, esses marcadores têm sido muito utilizados em estudos de identificação individual e investigações de parentesco, bem como em estudos populacionais (KAYSER *et al.*, 2000a e 2004; QUINTANA-MURCI, KRAUSZ e MCELREAVEY, 2001; GUSMÃO e CARRACEDO, 2003; DE SOUZA GÓES *et al.*, 2005; BALLANTYNE *et al.*, 2010 e 2012). Em estudos populacionais, os haplótipos descritos por Y-STRs permitem determinar distâncias genéticas entre populações e, consequentemente, inferir movimentos migratórios recentes (estudos micro-evolutivos) (QUINTANA-MURCI, KRAUSZ e MCELREAVEY, 2001; BALLANTYNE *et al.*, 2010).

Figura 9. Marcadores Y-STR mais utilizados em estudos de genética populacional e forense, bem como sua localização no cromossomo Y



Fonte: PEREIRA e GUSMÃO (2019).

O "haplótipo mínimo" foi o primeiro conjunto de Y-STRs sugerido pela comunidade forense europeia, em 1997, sendo formado por 9 *loci* Y-STR (DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393 e DYS385a/b). Nos anos seguintes, os avanços tecnológicos auxiliaram na criação de conjuntos mais extensos de Y-STRs para o uso forense, com melhor poder de diferenciação, e sua incorporação em kits comerciais (Tabela 1).

Nas últimas décadas, os avanços no sequenciamento massivo em paralelo (MPS) trouxeram vantagens para a análise de DNA. Em comparação com a eletroforese capilar, é possível produzir milhões de leituras da sequência desejada em um curto período de tempo e analisar grandes sequências de DNA em fragmentos de qualquer tamanho (ALVAREZ-CUBERO et al., 2017). A par do desenvolvimento dos kits descritos anteriormente, cuja determinação alélica é baseada no tamanho dos fragmentos amplificados por PCR, os Y-STRs começaram a ser incorporados em metodologias desenvolvidas para MPS, tornando possível distinguir alelos do mesmo comprimento. Por exemplo, o kit MiSeq FGxTM Forensic Genomics System permite amplificar simultaneamente até 231 loci forenses em uma única reação multiplex, que incluem o gene da Amelogenina (AMELY), 27 STRs autossômicos, 24 Y-STRs, 7 X-STRs e três classes funcionais de polimorfismos de nucleotídeo único (SNPs) (YANG et al., 2022). Em 2021, foi desenvolvido o primeiro painel de sequenciamento específico para cromossomo Y, chamado de CSYseq, com a cobertura de 202 loci Y-STR (CLAERHOUT et al., 2021). O painel inclui Y-STRs de mutação rápida, lenta e moderada capazes de individualizar parentes paternos próximos. Portanto, com o alto rendimento do MPS é possível integrar mais *loci* Y-STR em sistemas multiplex baseados em MPS.

Os dados relativos às taxas de mutação de cada Y-STR, bem como as distribuições das frequências haplotípicas em diversas populações, são disponibilizadas através de publicações científicas ou em bancos de dados populacionais e forenses. Um banco de dados online de livre acesso, denominado YHRD (*Y-Chromosome STR Haplotype Reference Database; www.yhrd.org*), foi criado com o objetivo de disponibilizar dados que contém perfis de haplótipos dos marcadores mais utilizados em populações humanas. Atualmente, a base de dados contém cerca de 349.750 perfis de 142 populações mundiais para o haplótipo mínimo, que é composto pelos *loci* DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS385 a/b (*Release* R69 de junho de 2023). O YHRD possui uma ferramenta de busca que fornece aos usuários uma estimativa das frequências dos haplótipos (Y-STRs) nas populações inseridas no banco de dados, bem como outras informações relevantes, como, por exemplo, estimativas de taxas de mutação em Y-STRs.

Tabela 1. Conjunto de marcadores Y-STR utilizados pela comunidade forense

Conjunto de marcadores	Número de <i>loci</i> Y-STRs	<i>loci</i> amplificados
Haplótipo Mínimo	9	DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393 e DYS385a/b
Haplótipo Estendido	11	DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393 e DYS385a/b, YCAIIa/b
SWGDAM	11	DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393 e DYS385a/b, DYS438, DYS439
PowerPlex®Y (Promega Corporation)	12	DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439
Yfiler® (Applied Biosystems)	17	DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393 e DYS385a/b, DYS456, DYS458, DYS439, DYS635, Y-GATA-H4, DYS437, DYS438, DYS448
PowerPlex® Y23 (Promega Corporation)	22	DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS481, DYS533, DYS549, DYS570, DYS576, DYS635, Y-GATA-H4
Yfiler® Plus (Applied Biosystems)	26	DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393 e DYS385a/b, DYS456, DYS458, DYS439, DYS635, Y-GATA-H4, DYS437, DYS438, DYS448, DYS576, DYS627, DYS460, DYS518, DYS570, DYS449, DYS481, DYF387S1, DYS533
Investigator Argus Y-28 QS Kit (Qiagen)	28	DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393 e DYS385a/b, DYS456, DYS458, DYS439, DYS635, Y-GATA-H4, DYS437, DYS438, DYS448, DYS576, DYS643, DYS460, DYS518, DYS570, DYS449, DYS481, DYS533, DYS549, DYS627, DYS518, QS1, QS2

Legenda: Conjuntos de marcadores Y-STR utilizados pela comunidade forense ao longo do tempo, com ênfase no número e nomes dos *loci* incorporados em cada conjunto. O conjunto SWGDAM tem o mesmo nome do grupo forense que o sugeriu, o *U.S Scientific Working Group on DNA Analysis Methods*.

1.3.1.2 Taxas de mutação em Y-STRs

A mutação é definida como qualquer mudança hereditária permanente, no genoma das células somáticas ou germinativas do indivíduo (VICARD e DAWID, 2004; AMORIM, 2013; PINTO *et al.*, 2014 e 2015). Mutações que ocorrem durante a gametogênese são transmitidas à descendência (PINTO *et al.*, 2014 e 2015).

Uma vez que o cromossomo Y é transmitido do pai para os filhos (homens) sem recombinação, mutações durante a meiose são a única fonte de variação genética neste cromossomo, entre os homens da mesma patrilinhagem (CLAERHOUT *et al.*, 2018).

Em genética forense, a mutação resulta em uma incompatibilidade mendeliana em que a criança e o suposto pai apresentam diferentes alelos, isto é, diferenças no número de repetições do STR em análise. Diferenças entre haplótipos do suposto pai e filho apontam exclusão da paternidade biológica ou a ocorrência de mutações espontâneas na linhagem germinativa. Em investigações de parentesco, em que o suposto pai não está disponível para análise de DNA, e estão disponíveis outros indivíduos masculinos da mesma patrilinhagem, as chances de mutação aumentam proporcionalmente ao número de meioses que os separam (AMBROSIO *et al.*, 2020). Portanto, a interpretação probabilística de perfis genéticos, bem como os critérios de exclusão, levam em conta as taxas de mutação observadas para Y-STRs (KAYSER e SAJANTILA, 2001; KAYSER *et al.*, 2004; GOEDBLOED *et al.*, 2009).

As taxas de mutação também tm grande utilidade em outras aplicações forenses que incluem diferentes membros da mesma linhagem masculina. Por exemplo, em estudos de história evolutiva as taxas de mutação são usadas para estimar a origem local e temporal de um determinado haplogrupo baseado em Y-SNP (KAYSER *et al.*, 2000a; DUPUY *et al.*, 2004). Além disso, com base nos dados de taxas de mutação dos STRs são feitas inferências na determinação do tempo do ancestral comum mais recente (TMRCA) na pesquisa genealógica (WALSH, 2001; BALANOVSKY, 2017; CLAERHOUT *et al.*, 2018; BOATTINI *et al.*, 2019).

A investigação de mutações em pares pai-filho é a abordagem mais confiável para estimativa das taxas de mutação de *loci* Y-STR. Isso deve-se a observação sem ambiguidade da mutação alélica entre pai e filho. As taxas de mutação em STRs, incluindo Y-STRs, são pequenas (cerca de uma mutação em 1.000 gerações por *locus*). Portanto, o número de pares pai-filho investigados deve ser grande o suficiente para revelar estimativas confiáveis da taxa de mutação.

O principal mecanismo molecular responsável pelas mutações que alteram o número de repetições num STR é denominado de *slippage* (que significa "derrapagem") e ocorre durante a replicação do DNA (SCHOLOTTERER, 2000; ELLEGREN, 2004). Desta maneira, a fita sintetizada passa a apresentar um tamanho diferente da fita molde, seja pela inserção ou deleção de unidades repetitivas (ELLEGREN, 2004). Se o deslize produzir um laço na fita sintetizada, ocorrerá um aumento no número de repetições. De outra maneira, se o laço ocorre na fita molde, haverá uma diminuição do número de repetições na fita sintetizada.

Vários estudos investigaram as taxas de mutação em *loci* Y-STR, tendo-se observado variabilidade entre *loci* e entre diferentes alelos do mesmo *locus* (GUSMÃO *et al.*, 2005; BALLANTYNE *et al.*, 2010; da FRÉ *et al.*, 2015; TURRINA *et al.*, 2015a; WANG *et al.*, 2016; ANTÃO-SOUZA *et al.*, 2017; CLAERHOUT *et al.*, 2018; WU *et al.*, 2018; BOATTINI *et al.*, 2019; AMBROSIO *et al.*, 2020). Tal como para STRs localizados nos autossomos, existem vários fatores que influenciam as taxas de mutação de Y-STRs. São estes: o comprimento (em pares de bases) da unidade de repetição, o tamanho médio do alelo (número de repetições), a complexidade da sequência repetitiva, e a idade do pai no momento da formação do gameta (BRINKMANN *et al.*, 1998; KRUGLYAK *et al.*, 1998; CUMMINGS e ZOGHBI, 2000; GUSMÃO *et al.*, 2005; LEE *et al.*, 2007; GOEDBLOED *et al.*, 2009; BALLANTYNE *et al.*, 2010; CLAERHOUT *et al.*, 2018).

Além disso, é geralmente aceito que mutações que envolvem o ganho ou perda de uma única repetição (*single-step*) são mais frequentes do que as que envolvem vários passos mutacionais (*multi-step*) (XU *et al.*, 2000). Um estudo recente demonstrou que a proporção de mutações *single versus multi-step* varia em função do motivo repetitivo (ANTÃO-SOUSA *et al.*, 2023).

De uma forma geral, observa-se que alelos mais longos são mais mutagênicos que os curtos, e tendem a diminuir em comprimento, enquanto o oposto é visto para alelos mais curtos (KAYSER e SAJANTILA, 2001; DUPUY *et al.*, 2004; BUDOWLE *et al.*, 2005; GUSMÃO *et al.*, 2005; DOMINGUES *et al.*, 2009; GE *et al.*, 2009; GOEDBLOED *et al.*, 2009; BALLANTYNE *et al.*, 2010; SUN *et al.*, 2012; ANTÃO-SOUZA *et al.*, 2019). No estudo de Gusmão e colaboradores (2005), observou-se maiores taxas de mutação em alelos com repetições longas do tipo TAGA e GAAA, principalmente em alelos com 11 ou mais repetições consecutivas.

Diversos estudos já observaram a correlação positiva entre a idade do pai e a ocorrência de mutação (GOEDBLOED *et al.*, 2009; BALLANTYNE *et al.*, 2010; SUN *et al.*, 2012; RALF *et al.*, 2020). Segundo o estudo de Goedbloed e colaboradores (2009), a idade

deve ser considerada ao estimar as taxas de mutação em Y-STR e ao comparar as taxas estimadas de mutação de diferentes estudos.

Resultados combinados de estudos observaram um aumento da taxa de mutação com idade, embora tendesse a diminuir para a faixa etária de 51 a 60 anos, o que pode estar associado com o baixo número de pais estudados (GUSMÃO *et al.*, 2005; AMBROSIO *et al.*, 2020).

A maioria dos Y-STRs relatados anteriormente tiveram taxas de mutação estimadas na ordem de 1×10^{-3} ou superior. Ballantyne e colaboradores (2010), ao estudar 186 Y-STRs em 2.000 duos pai-filho, identificaram 13 Y-STRs de mutação rápida (RM) com uma taxa de mutação média de $1,97 \times 10^{-2}$. Neste trabalho foi observado um aumento estatisticamente significativo de 4,4 vezes na separação por eventos de mutação entre indivíduos masculinos intimamente relacionados em relação ao Yfiler®. Atualmente, alguns destes marcadores estão disponíveis nos kits comerciais PowerPlex® Y23 e Yfiler® Plus.

No YHRD, há um copilado de estimativas de taxas de mutações por *locus*, bem como uma lista dos estudos científicos e trabalhos publicados (https://yhrd.org/pages/resources/mutation_rates). Entretanto, os dados disponíveis ainda não levam em consideração os diferentes fatores que influenciam as taxas de mutação.

1.3.1.3 Single Nucleotide Polymorphisms específicos do cromossomo Y

Os Y-SNP geralmente apresentam apenas 2 alelos. Portanto, os Y-SNPs oferecem menos informação a nível intrapopulacional que os Y-STRs.

Devido às baixas taxas de mutação, os Y-SNPs apresentam baixa recorrência, isto é, a probabilidade de uma mesma mutação ocorrer ao caso em genomas diferentes é pequena. Assim, os SNPs são capazes de definir linhagens com uma elevada especificidade geográfica e populacional (BIRD, 2012; JOBLING, 2012; WANG *et al.*, 2015a). Portanto, estes marcadores são eficientes em análises genealógicas e de ancestralidade, em estudos evolutivos e populacionais (BIRD, 2012; BUTLER, 2009; JOBLING, 2012; OLIVEIRA *et al.*, 2014; RESQUE *et al.*, 2016).

Devido à ausência de recombinação, os Y-SNPs podem ser representados em uma árvore filogenética, monofilética, representando os haplogrupos de acordo com sua ascendência. Como o Y-SNP surge por mutação a partir de um alelo ancestral, todos os

homens que possuem o alelo derivado descendem de um ancestral comum (UNDERHILL *et al.*, 2001; MIZUNO *et al.*, 2010). Um alelo é estabelecido como ancestral quando se observa compartilhamento entre as sequências humanas e de primatas, sendo frequente a comparação com o chimpanzé (JOBLING e TYLER-SMITH, 2003).

Cada ramificação da arvore é definida por pelo menos um Y-SNP. A partir da genotipagem de marcadores Y-SNP, que são agrupados de maneira hierárquica nos ramos da árvore filogenética, é possível determinar a ancestralidade paterna de um indivíduo do sexo masculino (JOBLING e TYLER-SMITH, 2000; UNDERHILL *et al.*, 2001; van OVEN *et al.*, 2014).

A primeira árvore filogenética do cromossomo Y foi construída em 2002 por um grupo de colaboradores, o qual foi denominado Consórcio do Cromossomo Y (YCC, do inglês Y Chromosome Consortium) (ISOGG.ORG, 2021). Originalmente, foram agrupados 243 polimorfismos bialélicos em 19 haplogrupos, usando letras maiúsculas que vão de A-R. Em 2008, com base em novos Y-SNPs descritos, Karafet e colaboradores atualizaram árvore filogenética do cromossomo Y, tendo sido introduzidos os haplogrupos S e T (KARAFET *et al.*, 2008).

A cada ano, novas atualizações são feitas, com a inclusão de novos Y-SNPs, uma vez que novas tecnologias e plataformas de sequenciamento permitem o descobrimento de novas variantes, bem como novos conhecimentos de ascendência entre eles, além do crescente volume de publicações científicas, incluindo estudos populacionais (JOBLING e TYLER-SMITH, 2000; KARAFET *et al.*, 2008; van OVEN *et al.*, 2014).

Entretanto, a filogenia e nomenclatura mais atual é a sugerida por van Oven e colaboradores, em 2014 (Figura 10). Esta filogenia nomeia os haplogrupos utilizando a mutação mais estável que dá origem à ramificação e agrupa 18 haplogrupos principais de A-T, seus sub-clados hierárquicos dentro de cada haplogrupo, representados por ramificações na árvore, com os marcadores mais significativos e bem resolvidos (van OVEN *et al.*, 2014). Mutações mais antigas definem os haplogrupos da base e as mutações ocorridas mais recentemente definem os sub-ramos (ou sub-clados).

A filogenia e o modo de transmissão do cromossomo Y permitem que, uma vez feita a genotipagem de marcadores Y-SNP de um determinado individuo do sexo masculino, seja possível identificar a linhagem paterna à qual ele pertence. É possível analisar as mutações que este indivíduo possui e agrupá-las de maneira hierárquica seguindo os ramos da árvore até o ancestral comum. Por exemplo, é possível saber que todo indivíduo pertencente ao haplogrupo Q1a3-M3 possui as mutações M3, M346 e M242, logo, a simples detecção de M3

implica necessariamente a presença de M242 e M346. Portanto, analisando apenas os marcadores Y-SNP mais recentes é possível identificar os ramos mais antigos sem precisar genotipar os Y-SNP que os definem, de maneira a aperfeiçoar a genotipagem, com mais agilidade e menos custo.

O padrão de distribuição dos haplogrupos em diferentes regiões geográficas revela que quanto mais ampla é a sua distribuição geográfica, mais antigo é o seu surgimento e, portanto, o aparecimento da mutação que o define. Á vista disso, alguns dos haplogrupos podem ser encontrados em áreas ou continentes específicos e outros possuem extensa representação geográfica. Essa peculiaridade da distribuição geográfica dos haplogrupos permite fazer inferências sobre a possível origem geográfica de uma amostra de interesse e explorar padrões de migração, subestrutura populacional e mistura entre diversas populações (GUSMÃO *et al.*, 2007).

Os haplogrupos A e B, ramos de base da árvore filogenética, são comumente observados em populações africanas, apoiando as evidências de que o surgimento e as migrações do homem moderno partiram do continente africano (CHIARONI *et al.*, 2009). Outro haplogrupo frequente na África é o haplogrupo E. Entretanto, este não se restringe apenas a essas regiões, e pode ser encontrado no oeste do continente asiático e mais ao sul da Europa.

Os haplogrupos C e Q são observados no continente asiático, entretanto o haplogrupo C também é encontrado na América do Norte e em nativos do Equador e o Q em todo o continente americano (CRUCIANI *et al.*, 2004; KARAFET *et al.*, 2008; CHIARONI *et al.*, 2009; MYRES *et al.*, 2011; VILLAESCUSA *et al.*, 2021).

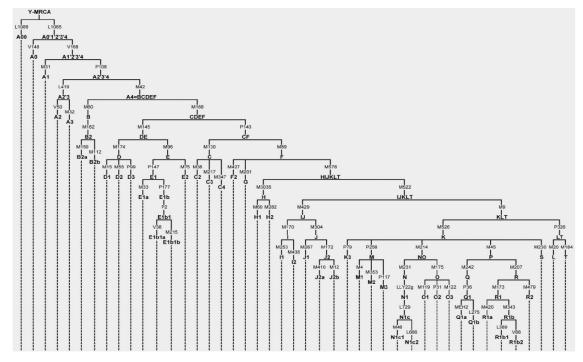


Figura 10. Árvore filogenética de haplogrupos do cromossomo Y

Legenda: Representação da árvore filogenética do cromossomo Y com destaque, em negrito, para a nomenclatura de base dos haplogrupos e os Y-SNPs que definem cada ramificação. No topo, está indicado o ancestral comum mais recente de todos os homens modernos (Y-MRCA, do inglês *Most Recent Common Patrilineal Ancestor of all modern humans*). Adaptação de van Oven e colaboradores (2014).

Os haplogrupos G, H, I, J, R, T e L são observados em diferentes populações, desde a Europa até o Médio Oriente, algumas partes da Ásia e do norte da África. Os haplogrupos R e I são encontrados em altas frequências na Europa. Enquanto o haplogrupo I é quase exclusivamente europeu, o haplogrupo R é encontrado em baixas frequências na África, Oriente Médio, Austrália, Ásia Central e muito frequente nas populações miscigenadas sulamericanas, que sofreram colonização europeia (UNDERHILL *et al.*, 2001; JOBLING e TYLER-SMITH, 2003; BELEZA *et al.*, 2006). O haplogrupo O, por sua vez, abrange a grande maioria das linhagens masculinas do leste e sudeste da Ásia (WANG e LI, 2013).

A análise combinada de marcadores Y-STR e Y-SNP é eficiente em estudos populacionais a níveis micro e macro-evolutivos, respectivamente, uma vez que suas taxas de mutação são diferenciadas.

1.3.2 <u>Programas de predição de haplogrupos com base em haplótipos Y-STR</u>

As diferenças nas taxas de mutação entre os marcadores Y-STR e Y-SNP permitem que um haplogrupo seja composto por diferentes haplótipos Y-STR, próximos entre si (BOSCH *et al.*, 1999; MUZZIO *et al.*, 2011; WANG *et al.*, 2015a). Portanto, os haplótipos STR dentro de um dado haplogrupo possuem semelhanças suficientes entre si para inferir a linhagem paterna.

Na ausência de dados de Y-SNPs, é possível utilizar *softwares* que permitem inferir o haplogrupo mais provável, com base num perfil de Y-STRs, por comparação com haplótipos depositados em bases de dados usadas como referência (WANG *et al.*, 2015a).

O software Haplogroup Predictor (http://www.hprg.com/hapest5/), disponível gratuitamente em um site online, é popular entre a comunidade científica. Este software produz estimativas baseadas nos perfis STR depositados em bases de dados de referência, consultadas pelo software, com a finalidade de predizer o haplogrupo correspondente ao perfil haplotípico consultado a partir de um algoritmo com uma abordagem Bayesiana de frequências alélicas (ATHEY, 2006; MUZZIO et al., 2011; WANG et al., 2015a).

No momento da análise, é importante indicar a região geográfica de onde provém a amostra analisada, já que as frequências dos haplótipos dependem da região considerada. O *software Haplogroup Predictor* possui as opções de região geográfica do Nordeste europeu, Leste europeu e Mediterrâneo, além da opção "Prioridades Iguais" (ATHEY, 2006).

Com o objetivo de atingir maior eficácia, o programa recomenda o uso do maior número de marcadores disponíveis. O uso de 10 a 20 marcadores são suficientes para obter probabilidades maiores que 99% para um dado haplogrupo com base no seu haplótipo (ATHEY, 2006).

A predição de haplogrupos utilizando este *software* tem limitações, já que o número de perfis STR associados a determinados haplogrupos, a estimativa das taxas de mutação e frequências alélicas dos Y-STRs utilizados podem influenciar na probabilidade apresentada. Além disso, o tamanho do banco de dados e sua população de referência são importantes para o cálculo de probabilidades. Se o banco de dados de referência for reduzido ou existirem populações pouco representadas, como por exemplo, as nativas, alguns haplogrupos poderão ser difíceis de detectar. Vale ressaltar que a predição de haplogrupos cujas populações já foram bastante estudadas é comumente mais precisa, por exemplo, as populações europeias.

Outro *software* gratuito, disponível na internet, é o NEVGEN *Y-DNA Haplogroup* (https://www.nevgen.org/) que também utiliza um algoritmo baseado na abordagem Bayesiana de frequências alélicas, mas tem o diferencial de ter em conta correlações (interdependência) entre pares de Y-STRs, na predição de sub-haplogrupos (CETKOVIC GENTULA e NEVSKI, 2015). A base de dados de referência consultada pelo NEVGEN é a do projeto *Family Tree DNA* (FTDNA) (https://www.familytreedna.com). Para o cálculo, o *software* considera apenas haplótipos que estejam associados com dados de Y-SNPs de elevada resolução e não é necessário indicar a região de origem da amostra.

2 JUSTIFICATIVA

As populações sul-americanas têm sido alvo de alguns estudos cujo objetivo é analisar haplótipos Y-STR e as linhagens paternas, com o objetivo de caracterizar a composição genética presente e seus padrões de miscigenação que permitem desenhar os padrões de migração das populações humanas na formação dessas populações.

A caracterização genética de linhagens masculinas das populações sul-americanas é mais eficiente com a analise combinada dos marcadores Y-STR, altamente discriminativos, e marcadores Y-SNP mais estáveis, com uma elevada especificadade geográfica e populacional. Portanto, o emprego destes dois tipos de marcadores em conjunto permitem determinar distâncias genéticas entre populações, e, inferir movimentos migratórios recentes, bem como podem revelar sinais de estruturação populacional a nível continental.

Tendo em vista a carência dos estudos genéticos que vêm sendo implementados em populações sul americanas, entende-se ser relevante aprofundar o conhecimento da composição genética destas populações por meio da análise combinada de um grande número de *loci* Y-STR e Y-SNP, de maneira a descrever sua composição genética, e assim prover uma base de dados destas populações afim do seu uso em genética forense e populacional.

3 OBJETIVOS

3.1 Objetivo geral

O presente trabalho teve como objetivo contribuir para o estudo das linhagens paternas de populações da América do Sul. Através da análise de marcadores uniparentais presentes no cromossomo Y, pretendeu-se efetuar uma análise comparativa entre populações da América do Sul, no que diz respeito a sua composição genética de origem paterna. Os resultados genéticos obtidos foram interpretados, tendo em conta os dados históricos disponíveis quanto à colonização, rotas de dispersão e migrações através deste subcontinente.

Para além do objetivo principal, foram feitas análises de mutação em pares pai-filho, no âmbito de um estudo colaborativo, com o objetivo de aumentar a precisão das estimativas de taxa de mutação em Y-STRs.

3.2 Objetivos específicos

De maneira a cumprir o objetivo geral, determinaram-se os seguintes objetivos específicos:

- a) Realizar a caracterização genética de linhagens masculinas com 23 loci Y-STR, em indivíduos não aparentados provenientes de populações miscigenadas do Rio de Janeiro, Maranhão, Terra do Fogo e de diferentes regiões da Colômbia e do Equador. Esta caracterização foi estendida a um conjunto de 201 pares pai/filho da população do Rio de Janeiro, para determinação de taxas de mutação;
- b) Comparar os dados haplotípicos obtidos e os disponíveis para outras populações sulamericanas, através da análise de distâncias genéticas e probabilidade de não diferenciação entre pares de populações;
- c) Determinar os haplogrupos presentes em amostras da Terra do Fogo, Paraguai, Equador e classificá-los quanto à sua origem geográfica;

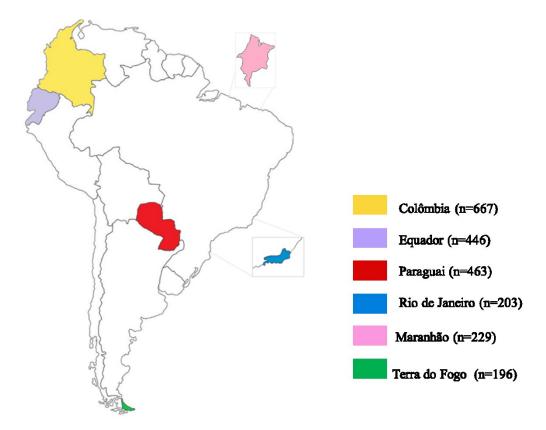
- d) Comparar os dados de haplogrupos obtidos e os disponíveis para outras populações sul-americanas, através da análise de distâncias genéticas e probabilidade de não diferenciação entre pares de populações;
- e) Analisar a contribuição de linhagens paternas provenientes de grupos ameríndios, europeus e africanos nas populações estudadas e comparar com outras populações sulamericanas já descritas na literatura;
- f) Fornecer informações pertinentes para o banco de dados de haplótipos do cromossomo
 Y;
- g) Recalcular taxas de mutação dos Y-STRs, após adição dos nossos dados aos já disponíveis na literatura.

4 MATERIAL E MÉTODOS

4.1 Amostras populacionais

Ao longo deste trabalho, a análise de marcadores do cromossomo Y foi realizada em 2.202 amostras de indivíduos não aparentados até três gerações pertencentes ao sexo masculino, provenientes de cinco países da América do Sul: Colômbia, Equador, Paraguai, Brasil (Rio de Janeiro e Maranhão) e Argentina (Terra do Fogo) (Figura 11).

Figura 11. Mapa geográfico da América do Sul, com destaque para os países de origem das 2.202 amostras deste estudo



Legenda: Representação esquemática das amostras populacionais estudadas neste trabalho e seus respectivos países de origem.

A utilização das amostras no âmbito deste projeto e o respectivo consentimento informado foram aprovados por comités de ética dos diferentes países de origem das

amostras, nomeadamente, pelo *Comité de Etica en Investigación Clínica del Instituto de Previsión Social* do Paraguai (Anexo A), pelo Comitê de Ética em *Human Research* da *Universidad de las Américas* (CEISH-UDLA 2017-0301) (Anexo B) do Equador, pelo Comitê de Ética em Investigações Biomédicas (IMBICE) (Anexo C) da Argentina, pelo Comitê de Ética da Faculdade de Ciências da *Pontificia Universidad Javeriana* (Anexo D) da Colômbia e pelo Comitê de Ética em Pesquisa do Hospital Pedro Ernesto (Anexos E e F). Além disso, foram seguidos os princípios éticos da Declaração de Helsinque de 2000, da Associação Médica Mundial (http://www.uma.net/e/policy/b3.htm).

As amostras da Colômbia foram adquiridas em colaboração com o Dr. Humberto Ossa Reyes, do Laboratório de Genética, Bogotá, Colômbia, e com a Professora Beatriz Martínez, do Laboratório de Genética Molecular, Instituto de Pesquisas Imunológicas, da Universidade de Cartagena, Cartagena, Colômbia. Parte das amostras foi extraída no âmbito de um estudo anterior (OSSA *et al.*, 2016), onde se encontram detalhes sobre a coleta, distribuição geográfica das amostras e método de extração do DNA. Quanto às amostras restantes, foram cedidas alíquotas de DNA genômico (volume final de 20 μL em uma concentração de 5 ng/μL), para este estudo.

As amostras de mestiços das três regiões equatorianas (Amazônia, Pacífico e Andes) foram obtidas em colaboração com o Dr. German Burgos Figueroa, da Faculdade de Medicina da *Universidad de las Américas*, Quito, Equador. Após coleta do material biológico, foram cedidas para este estudo alíquotas de DNA genômico (volume final de 20 μL em uma concentração de 5 ng/μL).

As amostras do Paraguai foram extraídas no âmbito de um estudo anterior (RIBEIRO et al., 2018), onde se encontram detalhes sobre a coleta, distribuição geográfica das amostras e método de extração do DNA. Este trabalho foi realizado em colaboração com o Dr. Carlos Vullo, do *Laboratorio de Genética Forense del Equipo Argentino de Antropología Forense* - EAAF em Córdoba, Argentina.

Um total de 201 pares pais-filhos do Rio de Janeiro foi selecionado com o objetivo de estimar taxas de mutação em 23 *loci* Y-STR. As relações de paternidade foram confirmadas previamente pela genotipagem de STRs autossômicos, no Laboratório de Diagnósticos por DNA (UERJ), com razões de verossimilhança acima de 10.000.

Alíquotas de DNA genômico (volume final de 20 μL em uma concentração de 5 ng/μL), extraído a partir de amostras do Maranhão, foram obtidas em colaboração com a Professora Marília Gomes do Laboratório de Diabetes, Faculdade de Ciências Médicas

(UERJ), e a Dr. Rossana Azulay, do Departamento de Endocrinologia do Hospital Universitário da Universidade Federal do Maranhão.

Neste trabalho, também foram analisadas 196 amostras coletadas no arquipélago da Terra do Fogo em colaboração com a Dra. María Laura Parolin, do *Instituto de Diversidad y Evolución Austral* (IDEAus), em Puerto Madryn, Argentina. Após a coleta do material biológico nas cidades de Rio Grande e Ushuaia, foram cedidas alíquotas de DNA genômico (volume de 20 μL a 5 ng/μL) para este estudo populacional.

4.2 Métodos de extração de DNA

A extração de DNA foi realizada por diferentes procedimentos laboratoriais, de acordo com as condições de armazenamento e quantidade de amostra. As amostras da Terra do Fogo e Maranhão, e parte das amostras da Colômbia foram extraídas pelo método de extração orgânica (Fenol Clorofórmio). As amostras do Equador, Rio de Janeiro, Paraguai e parte das amostras da Colômbia foram extraídas utilizando o reagente Chelex®. Além disso, o kit de extração PureLink™ Genomic DNA Mini Kit (*Thermo Fisher Scientific*) foi utilizado para extrair o DNA de parte das amostras da Colômbia.

4.2.1 Extração orgânica de DNA

Para cada amostra em papel Whatman[®] FTA[®] (*Sigma-Aldrich*), foi recortado um quadrado de 0,5 mm² da mancha de sangue, com o uso de um bisturi, e colocado em um tubo *eppendorf* de 1,5 mL, com a identificação correspondente, contendo 100 μL de SSC 1x (NaCl 3 M; Citrato trisódico 0,3 M). Os tubos foram incubados por 30 minutos em temperatura ambiente, com períodos de agitação manual, seguindo-se uma centrifugação a 13.000 rpm por 5 minutos (centrífuga Centrimicro 242 – Fanem).

Após o descarte do sobrenadante, foram acrescentados em cada tubo 400 μL de tampão de Extração (Tris 0,01M; EDTA Na2 0,01 M; NaCl 0,1 M; SDS 20% e H2O MiliQ), 25 μL de Proteinase K e 20 μL de Ditiotreitol (DTT). Em seguida, as amostras foram incubadas em banho-maria à 37 °C por 24 horas.

Às amostras submetidas à digestão, foram acrescentados 200 μL de Fenol/Clorofórmio/Álcool Isoamílico (25:24:1) (Invitrogen®, *Thermo Fisher Scientific*), seguida de centrifugação à 13.000 rpm por 3 minutos. A fase aquosa superior foi transferida para um novo tubo no qual foram acrescentados 200 μL de Clorofórmio/Álcool Isoamílico (24:1), seguida de nova centrifugação nas mesmas condições anteriormente citadas.

Novamente, a fase superior foi transferida para novo tubo e acrescentado 1 mL de Etanol Absoluto armazenado no frio (-20 °C) e 35 μL de Acetato de Sódio 3 M. Em seguida, os produtos foram mantidos por 24 horas à -20 °C. Após este período, foi realizada uma nova centrifugação a 12.000 g, 4 °C por 15 minutos.

Após o descarte do sobrenadante, foi adicionado em cada tubo 1 mL de etanol 70%, seguido de nova centrifugação à 12.000 g por 5 minutos. O descarte do sobrenadante foi repetido e as amostras foram colocadas em temperatura ambiente por 24 horas, para a evaporação do etanol restante. Finalmente, foram adicionados 40 μL de TE (Tris-EDTA, pH 8) em cada amostra e armazenadas a -20 °C.

4.2.2 Extração de DNA por resina Chelex®

As amostras de sangue do Rio de Janeiro, contidas em papel Whatman[®] FTA[®] (*Sigma-Aldrich*), foram extraídas utilizando-se o reagente *Chelex*® de acordo com a técnica desenvolvida por Walsh e colaboradores (1991).

Para cada amostra, foi recortado, com o uso de um bisturi, um quadrado de 0,3 mm² da mancha de sangue e depositado em um tubo *eppendorf* de 1,5 mL, com a identificação correspondente, contendo 500 μl de SSC 1x. Os tubos de extração foram incubados por 30 minutos em temperatura ambiente, com períodos de agitação manual seguindo-se uma centrifugação a 14.000 rpm por 3 minutos (centrífuga Centrimicro 242 – Fanem). Posteriormente, 400 μL do sobrenadante foram descartados e, com o auxílio de uma ponteira com a extremidade cortada, foram adicionados 500 μL de *Chelex*® 100 (100-200 *Mesh*, solução à 5%) e 400 μL de SSC 1x. Este material foi incubado a 40 °C *overnight*. No dia seguinte, o material foi centrifugado por 3 minutos a 14.000 rpm seguindo-se da retirada de 950 μL do sobrenadante e adição de 150 μL de Chelex 5%. Por fim, foi realizada uma incubação a 100 °C por 8 minutos. Após a incubação, as amostras foram centrifugadas à 14.000 rpm por 3 minutos e armazenadas a -20 °C.

4.2.3 Extração de DNA por PureLinkTM Genomic DNA Mini Kit (*ThermoFisher Scientific*)

As amostras de sangue total foram extraídas com o kit PureLinkTM Genomic DNA Mini Kit (*Thermo Fisher Scientific*). A cada 200 μL de sangue, foram adicionados 20 μL de proteinase K e 20 μL de RNase, seguido de vortex e incubação a temperatura ambiente por 2 minutos. Posteriormente, foram adicionados 200 μL do tampão de lise, seguido de uma incubação em banho seco a 55 °C por 10 minutos, seguido da adição de 200 μL de etanol 96-100%. O lisado resultante foi purificado de acordo com os seguintes passos:

- Transferência do lisado para uma coluna com tubo coletor, centrifugação a 8.300 rpm durante 3 minutos, em centrifuga não refrigerada;
- Transferência da coluna para um novo tubo coletor, adição de 500 μL de tampão de lavagem 1, preparado em etanol 96-100%, e centrifugação a 8.300 rpm durante 3 minutos, em centrifuga não refrigerada;
- Transferência da coluna para um novo tubo coletor, adição de 500 μL do tampão de lavagem 2, preparado em etanol 96-100%, e centrifugação a 14.000 rpm durante 3 minutos, em centrifuga não refrigerada e armazenado a -20 °C.

4.3 Análise de marcadores do cromossomo Y

A determinação dos haplótipos das amostras populacionais do presente trabalho foi realizada mediante o emprego dos 23 marcadores STR do cromossomo Y. A partir dos resultados haplotípicos, foi realizada a predição do haplogrupo mais provável, com o auxílio do *software* NEVGEN (https://www.nevgen.org/). Para confirmação dos haplogrupos, foi realizada a genotipagem de 66 Y-SNPs, selecionados para discriminar os principais haplogrupos que se espera encontrar em populações miscigenadas da América do Sul, nomeadamente os de origem Nativo americana, Europeia e Africana.

4.3.1 Genotipagem dos *loci* STR do cromossomo Y utilizando o kit PowerPlex® Y23

A amplificação dos 23 Y-STRs foi realizada em uma única reação de caráter *multiplex*, usando o kit PowerPlex® Y23 (*Promega*) e seguindo o protocolo disponibilizado pelo fabricante (*PowerPlex® Y23 System for Use on the Applied Biosystems® Genetic Analyzers*, 2018).

Na Tabela 2, estão indicados os 23 *loci* STR incluídos no kit usado e as respectivas informações quanto à posição cromossômica, alelos incluídos no padrão alélico, fluoróforos utilizados, e o resultado esperado para a amostra controle fornecida com o kit.

A mistura da reação de amplificação foi preparada dentro de uma capela esterilizada por luz ultravioleta de comprimento de onda de 254 nm. Com base no protocolo do kit, uma mistura de reação foi preparada multiplicando o número total de reações pelos seguintes volumes de reagentes: 0,8 μL de 5x *Master Mix* (composto de Hot Start DNA polimerase, dNTPs e MgCl₂), 0,4 μL de 10x *Primer Mix* (mistura do conjunto de 21 pares de oligonucleotídeos *forward* e *reverse*) e 3,3 μL de água ultrapura (Tabela 3). Um volume de 4,5 μL foi distribuído pelos vários tubos de reação e adicionado 0,5 μL de DNA a uma concentração de 2-5 ng/μL.

Para controlar possíveis contaminações e o sucesso da reação de amplificação, foram adicionados ao ensaio um controle positivo e um controle negativo. À reação do controle positivo adicionou-se 0,5 μL de amostra de DNA controle 2800M (10ng/μl) incluída no kit PowerPlex® Y23, cujos alelos estão descritos na tabela 2. Como controle negativo, foi preparada uma reação sem o acréscimo de DNA, sendo o volume do mesmo substituído por água ultrapura.

Tabela 2. 23 loci do kit PowerPlex® Y23

Locus	Posição cromos sômica (Mb)	Alelos incluídos no padrão alélico do kit PowerPlex® Y23	Fluoróforos	2800M DNA controle
DYS576	7.05	11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23		18
DYS389I	14.61	9, 10, 11, 12, 13, 14, 15, 16, 17		14
DYS448	24.36	14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24	Fluorescein	19
DYS389II	14.61	24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35		31
DYS19	9.52	9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19		14
DYS391	14.1	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16		10
DYS481	8.43	17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32		22
DYS549	21.52	7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17	IOE	13
DYS533	18.39	7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17	JOE	12
DYS438	14.94	6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16		9
DYS437	14.47	11, 12, 13, 14, 15, 16, 17, 18		14
DYS570	6.86	10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25		17
DYS635	14.38	15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28		21
DYS390	17.27	17, 18, 19, 20, 21, 22, 23, 24, 25, 26		
DYS439	14.51	6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17	TMR	12
DYS392	22.63	4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20		13
DYS643	17.43	6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17		10
DYS393	3.13	7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18		13
DYS458	7.87	10,11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24		17
DYS385 a/b	20.80 / 20.84	7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28	CXR	13, 16
DYS456	4.27	11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23		17
Y-GATA-H4	18.74	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18		11

Legenda: *Loci* PowerPlex® Y23, seus respectivos fluoróforos e alelos observados no padrão alélico e no DNA controle 2800M.

Tabela 3. Reagentes e volumes utilizados por amostra na preparação da mistura de reação para amplificação com o kit PowerPlex® Y23

Reagentes	Volume por amostra
Master Mix	0,8 μL
Primer Pair Mix	0,4 μL
Água	3,3 μL
Volume Total	4,5 μL

Legenda: Volume em μL de cada componente do kit PowerPlex® Y23 (*Mastex Mix, Primer Pair Mix* e água) utilizados por amostra na reação de amplificação.

A amplificação ocorreu em um termociclador Veriti™ 96-Well Thermal Cycler (*Thermo Fisher Scientific*), seguindo a termociclagem indicada no protocolo do kit e descrita na Tabela 4.

Tabela 4. Condições termocíclicas para a amplificação com o kit PowerPlex® Y23

Etapa	Ciclos	Temperatura (°C)	Тетро
Desnaturação inicial	1	96	2 minutos
Desnaturação		94	10 segundos
Anelamento	30	61	1 minuto
Extensão		72	30 segundos
Extensão final	1	60	20 minutos

Legenda: Etapas utilizadas na reação de amplificação, com número de ciclos e respectivas temperaturas e durações.

4.3.2 <u>Genotipagem dos marcadores Y-SNP utilizando o kit SNaPshotTM Multiplex (Applied BiosystemsTM)</u>

A análise dos marcadores Y-SNP foi realizada em um total de 279 indivíduos do Equador, 463 indivíduos não aparentados do sexo masculino originados da região oriental do Paraguai e 196 da Terra do Fogo.

Com base na informação disponível acerca da ancestralidade observada em populações sul-americanas, tendo em conta os dados históricos de colonização e os estudos

genéticos populacionais prévios (CORACH *et al.*, 2010; OLIVEIRA *et al.*, 2014; RESQUE *et al.*, 2016; JOTA *et al.*, 2016), foram selecionados 67 Y-SNPs que possibilitam diagnosticar os principais haplogrupos de origem europeia, africana e nativo-americana. Os Y-SNPs estudados foram agrupados em 8 ensaios multiplexes e um monoplex (Figura 12), descritos anteriormente por Brion e colaboradores (2005) [Multiplex 1], Gomes e colaboradores (2010) [Multiplex E2], Campos e colaboradores (2018) [Multiplex E1], Noguera e colaboradores (2013) [Multiplex Q], Resque e colaboradores (2016) [Multiplex R], Simão e colaboradores (2021) [Multiplex GIJ, Asian e O] e González e colaboradores (2013) [Monoplex V88].

Usando os dados de haplótipos de Y-STR, foi feita a predição dos haplogrupos com o *software* NEVGEN (https://www.nevgen.org/). Baseado nos resultados obtidos foi possível selecionar os marcadores Y-SNP relevantes para a comprovação do haplogrupo predito em cada uma das amostras deste estudo. Por exemplo, se o *software* indicou que a amostra pertence ao haplogrupo J2a, com uma probabilidade elevada (considerada maior que 90% no presente estudo), foi feita a sua genotipagem para um multiplex contendo oligonucleotídeos específicos para Y-SNPs determinantes do haplogrupo J2a ("Multiplex GIJ") (Figura 12). No entanto, se não foi determinado nenhum haplogrupo com elevada probabilidade, a amostra em questão foi analisada com um multiplex compostos por oligonucleotídeos específicos para Y-SNPs de ramos mais basais da árvore filogenética (Multiplex 1") (Figura 12), a fim de direcionar uma segunda genotipagem com Y-SNP determinantes de haplogrupos específicos).

As amostras do Rio de Janeiro, do Maranhão e da Colômbia não foram genotipadas para Y-SNPs e na determinação dos haplogrupos mais prováveis Para análise da contribuição de linhagens paternas destas populações, foram considerados os resultados da predição através do NEVGEN para probabilidades acima de 10%.

A4* (xM130,M174,M96,M213) M130 C-M130 D-M174 M174 E-M96* (xM33,P2,M85) Ela-M33 E1b-P2* (xM2 M35) Elbla-M2* (xM191,M154) E1b1a-U209* (xM154, U290) E1b1a-M154 Elb1a-U290 E1b1a-M191 * (xU154) Elb1a-U154 E1b1b-M35*(xM78, M81,M123,V6,M293,M85) E1b1b-M78 SRY10831 E1b1b-M81 E1b1b-M123 Elblb-V6 E1b1b-M293 E2-M85 F-M213*(xM9) G-M201 M170 I-M170+ (xM26) I2a-M26 J-12f2a*(xM172,M62) J-M62*(xM267) J1-M267*(xP58) J1-P58 J2-M172 KLT-M9*(xM22, M231,92R7, M70) L1-M22 N-M231 *(xTat) N-Tat O-M175*(xM119,L690,M122) M119 O1-M119 O2-L690*xM95) O2 a-M95 O3-M122*(xll 27,P201) O3 a-L127*(x002611) O3 a-JTS002611 O3 a-P2 01*(xP164) O3 a-P1 648(xM134) O3 a-M134* (xM117) O3 a-M117 P-92R7*(xM252,M173) Q-M242*(P36.2) Q1-P36+(M346) Q1b-M346*(xM3) Q1b-M3(xM19,Z19319,Z19383,M557,SA05) M19 Q1b-M19 Z19319 Q1b-Z19319*(xSA01) O1b-SA01 Q1b-Z19483 Q1b-M557 Q1b-SA05 R1-M173 R1a-SRY10831.2 R1b-P25*(xL23) R1b-M269*(xV88,L23) Multiplex 1 [1] R1b1b-V88 Multiplex GIJ [3] R1b1b-L23*(xU106,S116) Multiplex E1 [1,2] R1b1b-U106 Multiplex E2 [4] R1b1b-S116*(xU152,M529,M153,M167) Multiplex E1 e E2 R1b1b-U152 Multiplex Q [5] M529 R1b1b-M529 Multiplex R [6] M1 53 R1b1b-M153 Multiplex Asian [3] M167 (SRY2627) R1b1b-M167 Multiplex O [3] Tla-M70 Monoplex V88

Figura 12. Árvore filogenética de marcadores do cromossomo Y

[1] Brion et al. 2005; [2] Gomes et al., 2010; [3] Simão et al., 2021; [4] Ribeiro et al., 2022; [5] Noguera et al., 2013; [6] Resque et al., 2016.

Legenda: Conjunto de marcadores Y-SNP utilizados na genotipagem das amostras analisadas neste estudo. As cores indicam cada multiplex e os respectivos marcadores amplificados. Ao final de cada ramo da árvore se encontra o haplogrupo determinado por esses marcadores.

Na Tabela 5, estão organizados os marcadores de cada multiplex, bem como suas respectivas mutações, bases detectadas na eletroforese capilar e os tamanhos esperados dos fragmentos gerados.

Tabela 5. Marcadores Y-SNP e seus respectivos multiplexes (continua)

Y-SNP	Produto de SBE (bp)	Mutação	Detecção	Multiplex	
M22	18	A/G	A/G		
P25	26	C/CA	C/CA		
92R7	28	G/A	C/T		
SRY1532	30	A/G	A/G		
M70	34	A/C	T/G	MULTIPLEX 1	
M173	34	A/C	A/C		
Tat	42	T/C	A/G		
M213	45	T/C	A/G		
M9	48	C/G	C/G		
M172	18	T/G	T/G		
M170	22	A/C	T/G		
M201	34	G/T	G/T	MULTIPLEX GIJ	
M62	27	T/C	A/G	WOLTH LEA OIJ	
M267	37	T/G	A/C		
P58	40	T/C	T/C		
M174	19	T/C	A/G		
M130	21	C/T	C/T	MULTIPLEX Asian	
M231	27	G/A	G/A		
P164	21	A/G	A/G		
M122	25	A/G	T/C		
M95	27	C/T	C/T		
P201	29	T/C	A/G		
L690	32	A/G	T/C	MULTIPLEX O	
M134	32	C>del	G/A		
M119	38	T/G	A/C		
2611	42	C/T	G/A		
L127	46	G/A	G/A		

Tabela 5. Marcadores Y-SNP e seus respectivos multiplexes (conclusão)

	Produto de			iplexes (conclusão)	
Y-SNP	SBE (bp)	Mutação	Detecção	Multiplex	
P2	16	G/A	G/A		
M154	26	T/C	A/G		
M85	30	C/A	G/T		
M35	36	G/C	C/G		
M96	40	C/G	G/C		
U290	43	T/A	T/A	MULTIPLEX E1	
M2	45	A/G	T/C		
U209	49	C/T	G/A		
M191	51	T/G	T/G		
M33	54	A/C	T/G		
M174	58	G/A	G/A		
M293	21	T/G	T/G		
M81	27	C/T	C/T		
M78	34	C/T	C/T	MULTIPLEX E2	
V6	42	G/C	G/C		
M123	48	C/T	C/T		
M167	16	C/T	C/T		
L23	18	G/A	G/A		
M153	23	T/A	T/A		
M529	23	C/G	C/G		
U106	28	C/T	C/T	MULTIPLEX R	
M207	28	A/G	A/G		
U152	32	C/T	G/A		
M269	32	T/C	T/C		
S116	35	C/A	G/T		
M242	23	C/T	C/T		
P36.2	35	G/T	C/A		
M346	28	C/G	C/G		
M3	16	G/A	C/T	MULTIPLEX Q	
M19	28	T/A	A/T		
Z19319	40	C/T	G/A		
SA01	38	C/T	C/T		
Z19483	42	A/G	T/C		
M557	45	delTAC	G/T		
SA05	48	A/G	T/C		
V88	-	C/T	-	MONOPLEX V88	

Legenda: Marcadores Y-SNP de cada multiplex. Estão assinaladas as mutações, as bases detectadas na eletroforese capilar e os tamanhos esperados dos fragmentos gerados para cada marcador.

4.3.2.1 Reações de amplificação dos marcadores Y-SNP

As amplificações por PCR multiplex dos marcadores Y-SNP foram realizadas em uma cabine esterilizada por luz ultravioleta de comprimento de onda de 254 nm. Na PCR, foram utilizados 2,5 μL de tampão QIAGEN® multiplex PCR master mix 2x (*QIAGEN*), 0,5 μL de mix dos *oligonucleotídeos* na concentração de 2 μM, 1 μL de DNA a uma concentração de, aproximadamente, 2 ng, e 1 μL de água MilliQ, num volume final de 5 μL. Os conjuntos de *oligonucleotídeos forward* e *reverse* (não marcados), utilizadas em cada multiplex, encontram-se descritos na Tabela 6.

A cada PCR foram adicionados um controle negativo (reação sem o acréscimo de DNA, mas com o volume correspondente em água MilliQ) e um controle positivo, caracterizado pela adição de DNA genômico de um indivíduo previamente genotipado para um haplogrupo diferente daqueles pertencentes ao multiplex a ser analisado.

As reações de PCR foram realizadas em termociclador Veriti™ 96-Well Thermal Cycler (*Thermo Fisher Scientific*). As condições termocíclicas seguiram o protocolo descrito por Gomes e colaboradores (2010) e são apresentadas na Tabela 7.

Tabela 6. Conjuntos de oligonucleotídeos das PCR multiplexes realizadas (continua)

Y-SNP	Produto de PCR (pb)	Oligonucleotídeo forward (5'→3')	Oligonucleotídeo reverse (5'→3')	Referência	PCR Multiplex
M9	340	GCAGCATATAAAACTTTCAGG	AAAACCTAACTTTGCTCAAGC	BRION et al., 2004	
M173	172	GCACAGTACTCACTTTAGGTTTGC	GCAGTTTTCCCAGATCCTGA	BRION et al., 2004	
SRY1532	167	TCCTTAGCAACCATTAATCTGG	AAATAGCAAAAACTGACACAAGGC	BRION et al., 2004	
M213	145	GGCCATATAAAAACGCAGCA	*Oligonucleotídeo SBE reverso	BRION et al., 2004	
P25	121	GGACCATCACCTGGGTAAAGT	AGTGCTTGTCCAAGGCAGTA	BRION et al., 2004	MULTIPLEX 1
Tat	112	GACTCTGAGTGTAGACTTGTGA	GAAGGTGCCGTAAAAGTGTGAA	BRION et al., 2004	
M22	106	GCTGATAGTCCTGGTTTCCCTA	TGAGCATGCCTACAGCAGAC	BRION et al., 2004	
M70	81	TCATAGCCCACTATACTTTGGAC	CTGAGGCTGGACTATAGGG	BRION et al., 2004	
92R7	55	TGCATGAACACAAAAGACGTA	GCATTGTTAAATATGACCAGC	BRION et al., 2004	
M62	309	ACTAAAACACCATTAGAAACAAAGG	CTGAGCAACATAGTGACCCC	BRION et al ., 2004	
M267	256	CGTTGTCCCTGTGTTTCCAT	CTGTTGCCCAGGCTAGTGTC	NOGUEIRO et al ., 2010	
M172	187	TCCTCATTCACCTGCCTCTC	TCCATGTTGGTTTGGAACAG	BRION et al ., 2004	
P58	180	ACAGGAGGCCATAATGCAAC	GAGCCTCACACCTTCCTCTG	SIMÃO et al ., 2021	MULTIPLEX GIJ
M170	158	TGCAGCTCTTATTAAGTTATGTTTTCA	CCAATTACTTTCAACATTTAAGACC	BRION et al ., 2004	
M201	144	TCAAATTGTGACACTGCAATAGTT	CATCCAACACTAAGTACCTATTACGAA	BRION et al ., 2004	
12f2	90	CACTGACTGATCAAAATGCTTACAGAT	GGATCCCTTCCTTACACCTTATACA	BRION et al ., 2004	

Tabela 6. Conjuntos de oligonucleotídeos das PCR multiplexes estudadas (continuação)

Y-SNP	Produto de PCR (pb)	Oligonucleotídeo forward (5'→3')	Oligonucleotídeo reverse (5'→3')	Referência	PCR Multiplex
M85	283	TGGCATCCAATACTAGCTGATAAAC	AATGCTCACGCTTGTGTTCT	GOMES et al ., 2010	
U209	248	CCACAGGAATGCAAAAGATGTAAT	TGTGATGAGTGTCTGCCCAT	CAMPOS et al., 2018	
M35	198	GCATGGTCCCTTTCTATGGAT	GAGAATGAATAGGCATGGGTTC	BRION et al ., 2004	
M33	190	CACAACTTCATTGGCTACGG	GTTGA A GCCCCCA A GA GA GA C	GOMES et al ., 2010	
P2	180	GCTCCAGCCATCTTTTCCTTA	CTTCTCATGAGGGTTTTGGA	GOMES et al ., 2010	
M2	162	AAGTCCAGACCCAGGAAGGT	ACAGCTCCCCTTTATCCTC	GOMES et al ., 2010	MULTIPLEX E1
U174	150	TCCCTGCAGTGAAATAGTTTTG	AAATGGGAGTGTGGACTTGC	CAMPOS et al., 2018	
U290	135	CCTGGAAAGCCACTAGCAAC	GTGCAGACAAAAGCGTACCA	CAMPOS et al., 2018	
M154	130	TACTCACACAAACCAAGAAGAAACA	AACCATTGTGTTACATGGCCTA	GOMES et al ., 2010	
M191	122	AAAAATGGAGTGTTTATCAGAGCTT	CCCAGACACACAAAATATCTC	GOMES et al ., 2010	
M96	88	GTGATGTGTAACTTGGAAAACAGG	GGACCATATATTTTGCCATAGGTT	BRION et al ., 2004	
M78	235	GGATGCCTGTATGGGTTTCT	ATAGTGTTCCTTCACCTTTCCTT	GOMES et al ., 2010	
M293	230	AAAGAGATTGATCGGTGCATA	GCTGGCTAATACTTCCACAGAG	GOMES et al ., 2010	
M123	213	TGCTCTCAGGGGAAAATCTG	AGCAAAGTTGAGGTTGCACA	GOMES et al ., 2010	MULTIPLEX E2
M81	203	GCACTATCATACTCAGCTACACATCTC	AACCATTGTGTTACATGGCCTA	GOMES et al ., 2010	
V6	102	GATGGCACAGTGTTCGACAG	CTTCTCCCAAATGCCTGCT	GOMES et al ., 2010	

Tabela 6. Conjuntos de oligonucleotídeos das PCR multiplexes estudadas (continuação)

Y-SNP	Produto de PCR (pb)	Oligonucleotídeo forward (5'→3')	Oligonucleotídeo reverse (5'→3')	Referência	PCR Multiplex
SA01/ M557	370	AAGATCCCACCACTGCACTC	CTCTGGCCCCTAACAAACCT	NOGUERA et al., 2013	
M3	304	CATTAAAGCCGGTCACAGGT	CTGCCAGGGCTTTCAAATAG	NOGUERA et al., 2013	
P36.2	299	GA GGA GGGGA GA GA A A A	TTCAAACAGCCCACCAGATA	NOGUERA et al., 2013	
M19	277	TCACCAGAGTTTCAAATAG	ACAGACACAAAGGGCCAACT	NOGUERA et al., 2013	
M346	247	GGCCTGAAAATGTGGAAAGA	AGCCTGAAAATGTGGAAAGA	NOGUERA et al., 2013	MULTIPLEX Q
SA05	236	GAACCAAAGCACAGCACTCA	ATGCTCATGGCCTACACCTC	NOGUERA et al., 2013	
Z19483	211	CCATGTAGGAGGAGGCAAAA	CATCACAAAAGCCAAAAGCA	NOGUERA et al., 2013	
Z19319	163	TTTGCTGAAGTTGCCTGTCA	AGTTCCAGTCAGGGCAATCA	NOGUERA et al., 2013	
M242	155	TTGTGCAAAAAGGTGACCAA	TTTCGCTTTAAGGGCTTTCA	NOGUERA et al., 2013	
M207	322	CGTTACAACTATGGGGCAAA	TCCTCTGAAATGCCGAAT	RESQUE et al., 2016	
S116	241	TCAGTCAGGCAAATCTGAA	GGTGGAGTTGGGGCTAAAGT	RESQUE et al., 2016	
L23	229	ACACAGTGAAACCCCGTCTC	AAGATTGTGGGGACAAAGGA	RESQUE et al., 2016	
M529	228	GCCCCAAAACAACAGAATA	GGAAGCATTCAGAAGCAGGT	RESQUE et al., 2016	
U106	192	TCCTGAATAGCAAATCCCAAAG	AATGGCAGAGGTAGGAGGAAAT	RESQUE et al., 2016	MULTIPLEX R
M269	188	*Primer SBE	CTGGATGGTCACGATCTCCT	RESQUE et al., 2017	
U152	163	GAAACATTCCACGCTTGAGG	AGCCTCTTTTTGGCTTCCAT	RESQUE et al., 2016	
M167	131	GAGGCTGGGCCAAGTTAAG	CTTCCTCGGAACCACTACCA	RESQUE et al., 2016	
M153	76	TTCTCAGACACCAATGGTCCTA	TCTGACTTGGAAAGGGGAAA	RESQUE et al., 2016	

Tabela 6. Conjuntos de oligonucleotídeos das PCR multiplexes estudadas (continuação)

Y-SNP	Produto de PCR (pb)	Oligonucleotídeo forward (5'→3')	Oligonucleotídeo reverse (5'→3')	Referência	PCR Multiplex
M231	277	TTGCCTGTGCATGAAAAACT	TTTGACACCACAGAAATTACAGG	SIMÃO <i>et al</i> ., 2021	
M117	219	ATTTGGGTAGAAAAACTGCAAG	ACCAAAGGAATGCACATCTATCT	SIMÃO <i>et al</i> ., 2021	
M174	200	TCTCCGTCACAGCAAAAATG	GAAGGTCCTGGAGATGCAAA	SIMÃO <i>et al</i> ., 2021	MULTIPLEX Asian
M175	174	*Primer SBE	TCTTGCAGCATTTTCAGTTAGC	SIMÃO <i>et al</i> ., 2021	
M130	167	TGAGGCATGTGTGTTTTGGT	TGGCCAGCCTCTTATCTCTC	SIMÃO <i>et al</i> ., 2021	
M134	260	CCCTTCTTTGGCTTCTCTTTG	TTTCCCCACAACCAGACAAT	SIMÃO <i>et al</i> ., 2021	
L127	259	CCCTAAAACAAGGGGTCGAT	GGGGAAGGGATAGCATTAGG	SIMÃO <i>et al</i> ., 2021	
JTS002611	257	AGCCAACATACTCGCCAATC	GGCTTGCCCTACTGAGAACA	SIMÃO <i>et al</i> ., 2021	
M122	222	GGTATTCAGGCGATGCTGAT	GTGACTGCAAATGGTATGCAA	SIMÃO <i>et al</i> ., 2021	
M95	219	CCTTCTTGGGATCAAATGGA	GTTGTGAGGTCCTTCCCAGA	SIMÃO <i>et al</i> ., 2021	MULTIPLEX O
L690	207	CTCTCTCATGGGCTGGACAT	TGGCAGATGCAAGCTACCTA	SIMÃO <i>et al</i> ., 2021	
P201	152	TGTGCTGTGCAAGTTGTGTG	AACCCCAAATCCCAAGGTAG	SIMÃO <i>et al</i> ., 2021	
M119	95	TGGGTTATTCCAATTCAGCA	CAAAACCGCAGTGCTATTGTGT	SIMÃO <i>et al</i> ., 2021	
P164	80	*Primer SBE reverse	CCCTCTTTTCCTCCCATTC	SIMÃO <i>et al</i> ., 2021	
V88	143	TTTTCTTCCCCTCCTTGGAA	GGACTGGCCTTACAACACCAT	GONZÁLEZ et al ., 2013	MONOPLEX V88

Nota: Para o marcador M213, utilizou-se o primer SBE reverso na reação de PCR (*cacgtcgtgaaagtctgacaaTCAGAACTTAAAACATCTCGTTAC), pois o primer reverso confeccionado tinha um problema na sequência. *Este multiplex ainda não foi publicado e foi desenvolvido no Laboratório de Diagnósticos por DNA, de forma a compor a tese de Doutorado da Fernanda Miranda (BARROS, 2019).

Tabela 7. Condições termocíclicas utilizadas para amplificações dos multiplexes e monoplex deste estudo

Etapa	Ciclos	Temperatura (°C)	Tempo
Desnaturação inicial	1	95	15 minutos
Desnaturação		94	30 segundos
Anelamento	35	60	90 segundos
Extensão		72	1 minuto
Extensão final	1	72	10 minutos

Legenda: Etapas utilizadas na reação de amplificação, com número de ciclos, suas respectivas temperaturas e durações.

4.3.2.2 Eletroforese em gel de poliacrilamida

Para se verificar a eficácia de amplificação das sequencias de interesse, foi realizada eletroforese em gel de acrilamida e bis-acrilamida 29:1 a uma concentração de 10% em um volume final de 12 mL. Após a preparação das amostras, por adição de 1 μL de *SaferDye* (*Thermo Fisher Scientific*) a 2 μL do produto de amplificação, as amostras foram aplicadas no gel, bem como um marcador de tamanho molecular (*Thermo Fisher Scientific*), e submetidas à eletroforese em uma cuba vertical, a 80 volts, por 1 hora e 30 minutos, em tampão de corrida TBE 1x (89 mM Tris-base, 89 mM ácido bórico, 2 mM EDTA, pH 8).

O marcador de peso molecular, composto de uma mistura de fragmentos de DNA com comprimentos entre 50 a 800 pb, por meio de comparação, permitiu inferir os tamanhos dos fragmentos amplificados após a análise do gel em um transiluminador de luz azul (modelo Safe ImagerTM Blue-Light Transilluminator – *Thermo Fisher Scientific*).

Um exemplo dos padrões de amplificação dos marcadores dos multiplexes R e Q, mais frequentes em europeu e em nativo-americano respectivamente, resultantes da eletroforese em géis de poliacrilamida 10%, pode ser observado na Figura 13.

Os marcadores 12f2 do multiplex 2 e M175 e M117 do multiplex asiático 1, que consistem numa deleção, foram genotipados no próprio gel de poliacrilamida, pela presença ou ausência do produto de amplificação. Os restantes seguiram a genotipagem através da reação de SNaPhot.

SA01 e M557 M3 e P36.2 M19 M346 e SA05 Z19383 Z19319/M242

Figura 13. Gel de poliacrilamida da amplificação em multiplex Q de Y-SNPs

Legenda: Perfil eletroforético de uma amostra em gel de poliacrilamida 10%, de marcadores Y-SNP analisados para o multiplex Q. Os produtos de amplificação foram corados com *Safer Dye (Thermo Fisher Scientific)*.

4.3.2.3 Reações SBE

Com o objetivo de remover oligonucleotídeos e dNTPs não incorporados durante a PCR, os produtos amplificados foram purificados pela adição de 1 μL das enzimas USB[®] ExoSAP-IT[®] (*Thermo Fisher Scientific*) a 1 μL de produto de PCR. A mistura foi incubada a 37 °C por 1 hora e, em seguida, a 80 °C por 15 minutos, em termociclador VeritiTM 96-Well Thermal Cycler (*Thermo Fisher Scientific*).

Após purificação, os produtos foram preparados para a reação de SBE com a adição de 1 μL de SNaPshot[®] Multiplex Ready Reaction Mix (*Thermo Fisher Scientific*) e 2,5 μL de mix de primers SBE (Tabela 8). As reações SBE foram então submetidas às condições termocíclicas de 25 ciclos a 96 °C por 10 segundos, 50 °C por 5 segundos e 60 °C por 30 segundos em termociclador VeritiTM 96-Well Thermal Cycler (*Thermo Fisher Scientific*).

Em seguida, foi realizada a purificação dos produtos da reação de SBE, com o objetivo de remover os ddNTPs não incorporados, após a adição de 1 μL de enzima *Shrimp Alkaline Phosphatase* (SAP - *Thermo Fisher Scientific*) diretamente aos produtos de SBE, seguido de incubações a 37 °C por 1h e a 85 °C por 15 minutos.

Tabela 8. Conjuntos de oligonucleotídeos SBE das multiplexes (continua)

Oligonucleotídeo	Concentração na reação de SNaPshot	Produto de SBE (pb)	Sequência (5'→3')	Referência	PCR Multiplex
M22 SBE	0,09μΜ	18	GCCATTCCTGGTGGCTCT	BRION et al ., 2004	
P25 SBE	0,30μΜ	26	caaTCTGCCTGAAACCTGCCTG	BRION et al ., 2004	
92R7 SBE	0,375μΜ	26	aaGCATGAACACAAAAGACGTAGAAG	BRION et al ., 2004	
SRY1532 SBE	0,25μΜ	32	tctgacaaTTGTATCTGACTTTTTCACACAGT	BRION et al ., 2004	
M70 SBE	0,50μΜ	34	tctgacaaTAGGGATTCTGTTGTGGTAGTCTTAG	BRION et al ., 2004	MULTIPLEX 1
M173 SBE	0,375μΜ	34	gtctgacaaCTTACAATTCAAGGGCATTTAGAAC	BRION et al ., 2004	
Tat SBE	1,00μΜ	42	tcgtgaaagtctgacaaCTCTGAAATATTAAATTAAAACAAC	BRION et al ., 2004	
M213 SBE	0,15μΜ	45	gtgccacgtcgtgaaagtctgacaaTCAGAACTTAAAACATCTCGTTAC	BRION et al ., 2004	
M9 SBE	0,60μΜ	48	gtgccacgtcgtgaaagtctgacaaGAAACGCCCTAAGATGGTTGAAT	BRION et al ., 2004	
M172 SBE	0,80μΜ	18	AAACCCATTTTGATGCTT	BRION et al ., 2004	
M170 SBE	0,80μΜ	22	ACACAACCCACACTGAAAAAA	BRION et al ., 2004	
M62 SBE	1,20μΜ	27	ctgcaaCAATGTTTGTTGGCCATGGA	BRION et al ., 2004	MILEDI DV CH
M201 SBE	0,10μΜ	34	aagtctgacaaTAATAATCCAGTATCAACTGAGG	BRION et al ., 2004	MULTIPLEX GIJ
M267 SBE	1,20μΜ	37	TGAAAGTCTGACAACTTCCACACAAAATACTGAAMGT	NOGUEIRO et al ., 2010	
P58 SBE	1,20μΜ	40	CACGTCGTGAAAGTCTGACAATGACATTTGTGTGCTTTGC	SIMÃO et al ., 2021	

Tabela 8. Conjuntos de oligonucleotídeos SBE das multiplexes (continuação)

Oligonucleotídeo	Concentração na reação de SNaPshot	Produto de SBE (pb)	Sequência (5'→3')	Referência	PCR Multiplex
P2 SBE	0,20μΜ	16	GCCCCTA GGA GGA GA A	GOMES et al ., 2010	
M154 SBE	0,20μΜ	26	aaACATGGCCTATAATATTCAGTACA	GOMES et al ., 2010	
M85 SBE	0,16μΜ	30	CTTGTGTTCTATTAAGTGTAGTTTTGTTAG	GOMES et al ., 2010	
M35 SBE	2,00μΜ	36	CCCCCCCCCCCCCCAGTCTCTGCCTGTGTC	BRION et al ., 2004	
M96 SBE	0,12μΜ	40	CCCCCCCCCGTAACTTGGAAAACAGGTCTCTCATAATA	BRION et al ., 2004	
U209 SBE	0.67 μΜ	43	cgtcgtgaaagtctgacaaAAGACTGCAAGTTAAAATCA	CAMPOS et al., 2018	MULTIPLEX E1
M2 SBE	0,34μΜ	45	gtgccacgtcgtgaaagtctgacaaTTTATCCTCCACAGATCTCA	GOMES et al ., 2010	
U209 SBE	0.14 μΜ	49	actaggtgccacgtcgtgaaagtctgacaaTGTGGGAATTGATGGCGT	CAMPOS et al., 2018	
M191 SBE	0,50μΜ	51	ggtgccacgtcgtgaaagtctgacaaCATTTTTTTTTTTACAACTTGACTA	GOMES et al ., 2010	
M33 SBE	0,44μΜ	54	gtgccacgtcgtgaaagtctgacaaCAGTTACAAAAGTATAATATGTCTGAGAT	GOMES et al ., 2010	
M174 SBE	0.01 μΜ	58	TgactaaactaggtgccacgtcgtgaaagtctgacaaTGCATACCAGATTAACCCAT	CAMPOS et al., 2018	
M293 SBE	0,20μΜ	21	AAAGAGATTGATCGGTGCATA	GOMES et al ., 2010	
M81 SBE	2,00μΜ	27	CCCCCTAAATTTTGTCCTTTTTTGAA	BRION et al ., 2004	
M78 SBE	2,00μΜ	34	CCCCCCCCACACTTAACAAAGATACTTCTTTC	BRION et al ., 2004	MULTIPLEX E2
V6 SBE	0,32μΜ	42	gccacgtcgtgaaagtctgacaaTGCTGTGATTCCTGATGTG	GOMES et al ., 2010	
M123 SBE	0,50μΜ	48	taggtgccacgtcgtgaaagtctgacaaTTCTAGGTATTCAGGCGATG	GOMES et al ., 2010	

Tabela 8. Conjuntos de oligonucleotídeos SBE das multiplexes (continuação)

Oligonucleotídeo	Concentração na reação de SNaPshot	Produto de SBE (pb)	Sequência (5'→3')	Referência	PCR Multiplex
M3 SBE	3,00μΜ	16	TCACCTCTGGGACTGA	NOGUERA et al., 2013	
M242 SBE	0,75μΜ	23	aaAAAAAGGTGACCAAGGTGCT	NOGUERA et al., 2013	
M346 SBE	0,75μΜ	28	ctgacaaCAGCCAAGAGGACAGTAAGA	NOGUERA et al., 2013	
M19 SBE	0,50μΜ	28	tgacaaGTAGAGACATCTGAAACCCAC	NOGUERA et al., 2013	
P36.2 SBE	1,00μΜ	35	gtcgtgaaagtctgacaaCATCTATCTATCCATTATTCTCTCT	NOGUERA et al., 2013	MULTIPLEX Q
SA01 SBE	16,50μΜ	38	gtcgtgaaagtcctgacaaTTTGTCAGTGTAGAGTGG	NOGUERA et al., 2013	
Z19319 SBE	3,00μΜ	40	tcgtgaaagtctgacaaCCATCATCTCAACCTAAAATCC	NOGUERA et al., 2013	
Z19483 SBE	9,60μΜ	42	acgtcgtgaaagtctgacaaATAAGCTGTCTGGCTATTTCA	NOGUERA et al., 2013	
M557 SBE	3,20μΜ	45	tgccacgtcgtgaaagtctgacaaGAACAGGGTTGCAAACGGTA	NOGUERA et al., 2013	
SA05 SBE	4,40μΜ	48	aggtgccacgtcgtgaaagtctgacaaATGTTTCTAGGGTGAGCCTGT	NOGUERA et al., 2013	
M167 SBE	0,68μΜ	16	AAGCCCCACAGGGTGC	RESQUE et al., 2016	
L23 SBE	0,41μΜ	18	GCGA CA GA GCGA GA CTCT	RESQUE et al., 2016	
M153 SBE	0,20μΜ	23	AAAGCTCAAAGGGTATGTGAACA	RESQUE et al., 2016	
M529 SBE	0,10μΜ	23	AATAACAACCGCTCTCTCAGACA	RESQUE et al., 2016	MULTIPLEX R
M207 SBE	0,20μΜ	28	AACAAATGTAAGTCAAGCAAGAAATTTA	RESQUE et al., 2016	MULTIPLEX K
U106 SBE	0,04μΜ	28	TCTGACAATAGCAAATCCCAAAGCTCCA	RESQUE et al., 2016	
U152 SBE	0,20μΜ	32	CAAGGATAAGAAAATGAGTATTGTGAAAATA	RESQUE et al., 2016	
M269 SBE	0,20μΜ	32	gtctgacaaGGAATGATCAGGGTTTGGTTAAT	RESQUE et al., 2017	
S116 SBE	0,60μΜ	35	GAAAGTCTGACAAGAGTTGGGGCTAAAGTGAAAG	RESQUE et al., 2016	

Tabela 8. Conjuntos de oligonucleotídeos SBE das multiplexes (continuação)

Oligonucleotídeo	Concentração na reação de SNaPshot	Produto de SBE (pb)	Sequência (5'→3')	Referência	PCR Multiplex
M231 SBE	0,30μΜ	27	ACAACATTTACTGTTTCTACTGCTTTC	SIMÃO et al ., 2021	
M174 SBE	0,20μΜ	19	CACCCCTCACTTCTGCACT	SIMÃO et al ., 2021	MULTIPLEX Asian
M130 SBE	0,40μΜ	21	GGGCAATAAACCTTGGATTTC	SIMÃO et al ., 2021	
M134 SBE	0,40μΜ	32	agtctgacaaAAGAAAAGGCCCAGGAAAGTAT	SIMÃO et al ., 2021	
L127 SBE	0,50μΜ	46	cgtcgtgaaagtctgacaaTCAGAATAATAACAGAGTAATTGGCAG	SIMÃO et al ., 2021	
JST002611 SBE	0,40μΜ	42	ccacgtcgtgaaagtctgacaaCGAGGCCCTGTGCTTCCAGA	SIMÃO et al ., 2021	
M122 SBE	0,36μΜ	25	gacaaTCAGATTTTCCCCTGAGAGC	SIMÃO et al ., 2021	
M95 SBE	0,40μΜ	27	GGATAAGGAAAGACTACCATATTAGTG	SIMÃO et al ., 2021	MULTIPLEX O
L690 SBE	0,36μΜ	32	ctgacaaAGCCAGGTATTCAGAGAAGAAGCAA	SIMÃO et al ., 2021	
P201 SBE	0,24μΜ	29	gtctgacaaGTGAGAGCCAGTTAAAGCCC	SIMÃO et al ., 2021	
M119 SBE	0,40μΜ	38	gtgaaagtctgacaaTTATTCCAATTCAGCATACAGGC	SIMÃO et al ., 2021	
P164 SBE	0,30μΜ	21	AGCATTTTGGTCCCATCTTTT	SIMÃO et al ., 2021	

Legenda: Nesta tabela, estão listados os primers SBE utilizados neste estudo, suas respectivas concentrações na reação de SNaPshot, suas sequências no sentido 5'→3', os artigos de referência para cada conjunto de primer, o tamanho dos produtos de minissequenciamentos gerados e as reações de SBE multiplex às quais os marcadores foram agrupados e amplificados.

Os produtos purificados foram aplicados a cada poço de uma placa de 96 poços e adicionada uma mistura composta de 8,8 µl de *Formamida Hi-Di*® e 0,2 µL Liz-120, e submetidos à eletroforese capilar em sequenciador automático ABI Prism 3500, com polímero POP-7TM (*Thermo Fisher Scientific*). Os resultados foram analisados com o *software* GeneMapper versão 3.5 (*Thermo Fisher Scientific*).

O eletroferograma da Figura 14 exemplifica o resultado obtido após a realização de uma reação de SNaPshot dos Multiplexes R e Q, em duas amostras do Paraguai.

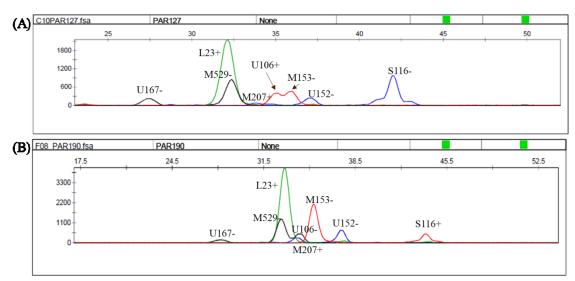


Figura 14. Eletroferograma da reação de SNaPshot do Multiplex R

Legenda: Eletroferograma da reação de SNaPshot com os marcadores Y-SNP da Multiplex R, exceto M269. Os sinais negativos (-) indicam que o alelo observado é o não mutado e os sinais positivos (+) indicam que o alelo observado é o mutado. (A) Amostra PAR127: R1b1a-U106, (B) Amostra PAR190: R1b1a-S116.

4.3.2.4 Genotipagem do marcador V88

A genotipagem do marcador V88 foi realizada por sequenciamento (GONZÁLEZ *et al.*, 2013). A PCR foi realizada conforme descrito anteriormente para os multiplexes, exceto para a temperatura de anelamento que foi aumentada para 63 °C para maximizar a especificidade de ligação do *primer*. O desempenho do PCR foi avaliado em um gel de poliacrilamida visualizado usando o método de coloração de prata. O produto de PCR foi purificado com USB® ExoSAP-IT® (*Thermo Fisher Scientific*), seguindo os mesmos volumes e condições aplicadas para os multiplexes.

A reação de sequenciamento foi preparada num volume final de 5 μL, incluindo 0,8 μL de BigDyeTM Terminator v3.1 Ready Reaction (Applied Biosystems), 1 μL de BigDyeTM Terminator v1.1, tampão de sequenciamento 5x (Applied Biosystems), 0,5 μL de *primer* V88 *foward* em uma concentração de 10 μM, 2 μL do produto purificado e água para completar.

A reação de sequenciamento foi então submetida às condições termocíclicas de uma temperatura inicial de desnaturação de 96 °C por 2 minutos, seguida de 35 ciclos a 96 °C por 30 segundos, 55 °C por 20 segundos e 60 °C por 30 segundos. Uma etapa final de extensão foi realizada a 60 °C por 10 minutos em termociclador VeritiTM 96-Well Thermal Cycler (*Thermo Fisher Scientific*).

Antes dos fragmentos sequenciados serem submetidos à eletroforese capilar em um ABI 3500 Genetic Analyzer (Applied Biosystems), foi realizada uma purificação em colunas com 750 μL de SephadexTM G-50 (GE Healthcare) a 10% e, adição do produto sequenciado purificado em 8 μL de Formamida Hi-Di. Os resultados foram analisados usando o *Software* de análise de sequenciamento v6.0 (Thermo Fisher Scientific).

4.4 Análises estatísticas

Com base nos dados obtidos, foram estimadas as frequências haplotípicas e os parâmetros de diversidade genética intrapopulacional, utilizando o *software Arlequin* v3.5.1.2 (SCHNEIDER *et al.*, 2000). Os valores de diversidade haplotípica (HD) foram calculados a partir da heterozigotia esperada, correspondendo, neste caso, à probabilidade de se encontrar dois haplótipos distintos em uma população. A este valor, foi aplicada a correção de Nei (NEI, 1987) ao efetivo amostral.O mesmo *software* foi utilizado nos cálculos de distâncias genéticas ($F_{\rm ST}$) e correspondente probabilidade de não diferenciação entre as amostras populacionais estudadas e outras populações da América do Sul previamente descritas.

Para as análises de distância genética ($F_{\rm ST}$) com base em dados haplotípicos, foi removido o marcador DYS385, uma vez que, por ser um *locus* duplicado, ocorre a amplificação de dois fragmentos que poderão ou não sobrepor-se em tamanho (ou número de unidades de repetição), o que dificulta a atribuição do *locus* (Figura 15A). Além disso, o número de repetições em DYS389I foi subtraído de DYS389II. Para o *locus* DYS389 foram amplificados dois fragmentos com tamanhos diferentes (239- 263 pb e 353-385 pb) (Figura 155B). Isto acontece, pois o primer *forward* hibridiza em duas regiões resultando em dois

produtos de amplificação com repetições de número variado, sendo que o fragmento mais curto (DYS389I) está contido no mais longo (DYS389II).

Além disso, os alelos nulos, microvariantes, duplicações e triplicações foram codificados como dados em falta e simbolizados com "?". As distâncias genéticas foram representadas no espaço bidimensional, por meio da análise de escalonamento multidimensional (MDS), através do programa Statistica *Software* v. 14.0.0.15 (TIBCO *Software Inc.*).

As frequências dos haplogrupos e as proporções de haplogrupos Europeus, Africanos e Ameríndios nas populações estudadas foram calculadas por contagem direta.

As taxas de mutação para os 23 Y-STRs, analisados em 201 pares de pai-filho, foram calculadas dividindo-se o número de mutações pelo número total de transmissões alélicas. Nas análises, foram considerados 2.225 duos pai/filho obtidos de 10 populações mundiais, bem como de trabalhos publicados anteriormente. Dessa maneira, para além dos 23 Y-STRs, foram analisados Y-STRs que pertencem ao Yfiler® Plus e outros oito que não pertencem a nenhum dos kits anteriormente mencionados (DYS388, DYS435, DYS461, DYS526a/b, DYS547, DYS612, DYS626 e Y-GATA-A10).

A correlação entre a idade paterna e a ocorrência de mutação foi avaliada por meio de testes qui-quadrado, bem como a correlação entre o comprimento dos alelos e ambas a ocorrência de uma mutação e o ganho ou perda de repetições para os 31 Y-STRs de estrutura simples. Para a análise entre a ocorrência de mutações compatíveis com ganho ou perda de repetições e o comprimento dos alelos, foi estabelecido, para cada marcador, o alelo modal, ou seja, o alelo com maior número de observações. Dessa maneira, para cada marcador, os alelos foram incluídos em uma das três categorias a seguir: i. alelo modal, ii. mais curto, e iii. alelo mais longo (relativo ao modal). Os intervalos de confiança a 95%, dos valores calculados para as taxas de mutação, foram estimados a partir do desvio padrão binomial. O poder estatístico dos resultados foi avaliado considerando um nível de significância α=0,05.

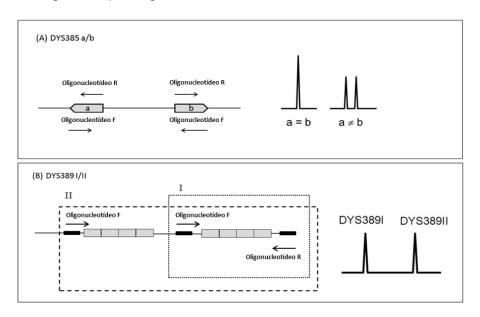


Figura 15. Representação esquemática dos *loci* DYS389 e DYS385

Legenda: Esquema dos *loci* STR: (A) DYS385, cuja região é duplicada e espaçada. Dessa forma, a amplificação com um único conjunto de oligonucleotídeos dá origem a um único pico, se a repetição "a" possuir o mesmo tamanho da "b" e dois picos se possuírem tamanhos diferentes; (B) DYS389 possui duas regiões de pareamento com o oligonucleotídeo *forward* e uma com oligonucleotídeo *reverse*. Note que DYS389II contêm a sequência que corresponde a DYS389I.

Fonte: Butler, 2010.

5 RESULTADOS E DISCUSSÃO

5.1 Análise populacional de haplótipos definidos por Y-STRs

Neste trabalho, foi analisado um total de 1.739 amostras da Argentina (Terra do Fogo), Brasil (Rio de Janeiro e Maranhão), Colômbia e Equador, para 23 Y-STRs. O Paraguai foi anteriormente genotipado e os dados haplotipicos publicados (RIBEIRO *et al.*, 2018). A tabela de haplótipos, os resultados da predição por NEVGEN da Colômbia, Maranhão e Rio de Janeiro e os resultados de haplogrupos definidos por Y-SNPs estão disponíveis em https://figshare.com/s/b5d23801320d296d0a44.

O kit PowerPlex® Y23 permitiu a obtenção de bons resultados para a maioria das amostras, com perfil completo e um bom balanço na amplificação dos 23 *loci*. Na Figura 16 é apresentado um eletroferograma representativo do resultado obtido com o kit PowerPlex® Y23, em uma das nossas amostras. Para cada um dos *loci* incluídos neste kit, um dos olinucleotídeos está marcado com um dos 4 fluoróforos: Fluorescein, JOE, TMR, CXR. O *software* de análise traduz as fluorescências emitidas em diferentes comprimentos de onda nas cores azul, verde, amarelo e vermelho. Tal como esperado para marcadores do cromossomo Y, no eletroferograma observa-se a presença de um único alelo em todos os marcadores, à exceção do marcador DYS385 que apresenta dois alelos (11 e 14). Isto deve-se à amplificação de dois *loci* com o mesmo par de olinucleotídeos. Porém, os dois alelos podem coincidir em tamanho, de maneira que, por vezes, apenas observa-se um pico no eletroferograma.

A maior parte dos alelos encontrados nos Y-STRs analisados variaram apenas no número repetições de um bloco de 3 a 6 pares de base, dependendo da estrutura do *locus* (DYS392 e DYS481 são *trirepeats*; DYS643 e DYS438 são *pentarepeats*; DYS448 é um *hexarepeat*; sendo os demais *loci tetrarepeats*).

Neste trabalho, foram encontradas microvariantes que, na sua maioria, já tinham sido previamente observadas em outras populações e depositadas no YHRD. Além disso, foram observados uma triplicação, 14 duplicações e 7 alelos nulos.

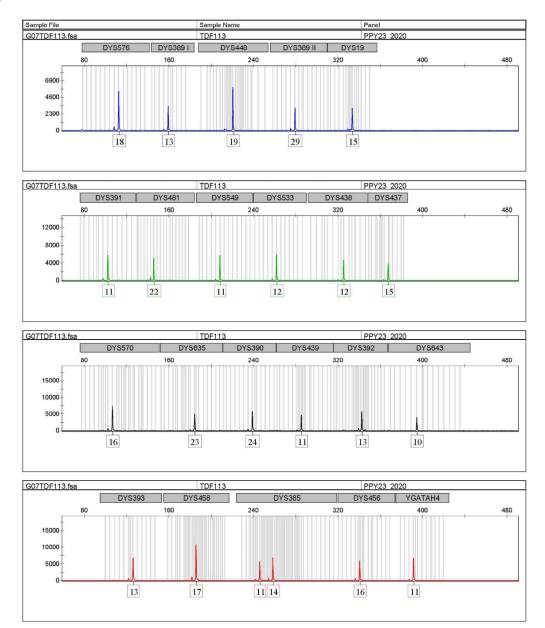


Figura 16. Eletroferograma do perfil genético da amostra TDF_113 obtido através da tipagem com o kit PowerPlex® Y23

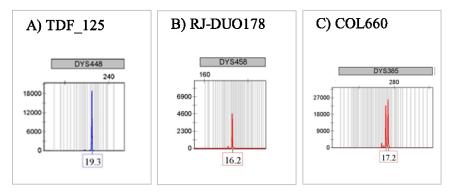
Legenda: O kit PowerPlex® Y23, que reúne 23 *loci* Y-STR, cujos oligonucleotídeos se encontram marcados com: (A) 6- Fluorescein; (B) JOE; (C) TMR;(D) CXR. O eixo X representa o tamanho dos fragmentos em pb, o eixo Y a intensidade de fluorescência (RFUs). Os picos representam os alelos designados com base no número de repetições *in tandem*.

5.1.1 Alelos microvariantes observados nas populações amostradas

Em alguns *loci*, foram detectadas microvariantes, alelos que se caracterizam pela inserção ou deleção de um número de bases diferente daquele que caracteriza a repetição. Este tipo de alelo, também denominado por intermédio ou não consensual, é classificado pelo

número de repetições completas, seguido de "." e do número de bases adicionais. Na Figura 17, podem ser visualizados três exemplos de microvariantes encontradas em três amostras deste estudo nos *loci* DYS488, DYS458 e DYS385.

Figura 17. Exemplos de microvariantes em três amostras, detectados usando o kit PowerPlex® Y23



Legenda: (A) Microvariante 19.3 em DYS448 na amostra TDF_125; (B) Microvariante 16.2 em DYS458 na amostra RJ-DUO178 (C) Microvariante 17.2 em DYS385 na amostra COL660.

O locus DYS458 foi o que apresentou o maior número de microvariantes, com alelos 16.1, 16.2, 17.2, 18.2, 19.2, 20.1 e 20.2 observados em 9 amostras do Rio de Janeiro, 9 do Maranhão, 27 da Colômbia, 7 da Terra do Fogo e 10 do Equador (Tabela 9). De forma a avaliarmos uma possível origem geográfica destes alelos, com base nos haplótipos, foram determinados os haplogrupos mais prováveis, usando o programa NEVGEN. O haplogrupo mais provável destas amostras foi o J1, com exceção de sete amostras atribuídas ao haplogrupo R1b1a-M269, duas ao E1b1b-M81, uma ao I2, uma ao T, e uma ao E1b1b-M78. Este resultado está de acordo com o reportado por Myres e colaboradores (2007), que descrevem uma associação entre os alelos não consensuais DYS458*.2 e os haplogrupos J1-M267 e R1b3-M405, embora não exista nenhuma relação relatada para os haplogrupos E1b1b-M81, E1b1b-M78, T e I2. Estas microvariantes já se encontram descritas na YHRD. Como esperado, dada a origem europeia dos haplogrupos J, R, e I2, bem como a presença de T, E1b1b-M81 e E1b1b-M78 em frequências consideráveis na Europa, estas microvariantes foram anteriormente observadas em populações sul-americanas e europeias (ALVES et al., 2007; PEREIRA et al., 2007; ROMERO et al., 2008; SÁNCHEZ-DIZ et al., 2008; TOSCANINI et al., 2008; BENTO et al., 2009; VALVERDE et al., 2012; OLIVEIRA et al., 2014; PURPS et al., 2014; CÁRDENAS et al., 2015; TURRINA et al., 2015a; MORALES et al., 2018; PAROLIN et al., 2018; TOSCANINI et al., 2018).

Três amostras da Colômbia (Tabela 9) apresentaram simultaneamente os alelos 9.2 no *locus* DYS438 e 19.2 no *locus* DYS458. Para este haplótipo, o NEVGEN indicou como haplogrupo mais provável o J1a3-Z1828. Embora o haplótipo não tenha nenhuma ocorrência em YHRD, o alelo 9.2 em DYS438 aparece em populações do leste Asiático (Release R69). Com base nesses resultados, é mais provável tratar-se de uma linhagem de origem nativoamericana, pertencente a um haplogrupo diferente do atribuído pelo NEVGEN.

O *locus* DYS385a/b foi o segundo que apresentou o maior número de microvariantes, com alelos 9.1, 10.2, 11.2, 12.2, 13.2, 14.2, 17.2 e 19.2 observados em 12 amostras da Colômbia, 1 do Rio de Janeiro e 1 do Maranhão (Tabela 9). Estes alelos foram anteriormente relatados em populações europeias, do Rio de Janeiro e da Argentina (PURPS *et al.*, 2014). O trabalho de Myres e colaboradores (2009) descreve uma correlação entre os alelos não consensuais DYS385.2 e três linhagens do cromossomo Y: D-M145, E1b1a-M2 e R1b*-M343. Entretanto, quando avalia-se a possível origem geográfica destes alelos no NEVGEN, com base nos haplótipos, várias linhagens foram atribuídas (A1a-M31, G, E1b1a-V38, I, J2, J1a3, R1b-V88 e R1b), o que sugere múltiplas origens das microvariantes em DYS385a/b.

Além dos alelos comuns, para o marcador DYS448 foram encontrados dois tipos de microvariantes: (i) uma amostra da Terra do Fogo apresentou o alelo 19.3 e outra do Equador apresentou o alelo 18.3; e (ii) numa amostra do Equador observou-se o alelo 20.2. Embora bem menos frequente que as microvariantes detalhadas anteriormente, algumas ocorrências destes alelos já foram depositadas na YHRD (Tabela 9). O alelo 18.3 também já foi relatado no Brasil (FIGUEIREDO *et al.*, 2016) e o alelo 20.2 na Holanda (PURPS *et al.*, 2014). Além disso, trabalhos populacionais relataram o alelo 19.3 em populações da China, Espanha, Estados Unidos, Países baixos e Brasil (PURPS *et al.*, 2014; FIGUEIREDO *et al.*, 2016). Para os alelos 18.3 e 20.2, o NEVGEN indicou como mais provável o haplogrupo G, e para o alelo 19.3 o haplogrupo J2a1, ambos frequentes na Europa, indicando, portanto, uma possível origem geográfica neste continente.

Finalmente, o alelo 12.1 no *locus* Y-GATA-H4 foi observado em apenas uma amostra equatoriana. Na YHRD não foi observada nenhuma ocorrência para este alelo. Entretanto, na literatura, este alelo foi descrito em Benim e na Finlândia (PURPS *et al.*, 2014). No entanto, o haplogrupo Q foi indicado pelo NEVGEN como sendo o mais provável,

apontando para uma origem nativo-americana. Em conjunto, os resultados obtidos não deixam clara a origem geográfica desta microvariante.

Tabela 9. Microvariantes encontradas nas amostras deste estudo

Locus	Código de amostra	Alelo	Número de ocorrências e frequência em YHRD
	EQM254	16.1	24 (2,3E-4)
	COL176, COL337, COL570, RJ-DUO21, RJ-DUO178 e MARA95	16.2	193 (6,6E-2)
	COL47, COL68, COL168, COL243, COL327, COL541, COL583, COL648, COL656, EQM48, EQM61, EQM78,EQ426, MARA2, MARA154, TDF_34, TDF_38,TDF_162,TDF_400 e TDF_529	17.2	1095 (3,74E-1)
DYS458	COL38, COL256, COL349, COL402, COL425, COL575, EQ261, EQM262, EQ277, EQ319, EQ338, RJ-DUO15, RJ-DUO41, RJDUO53, RJ-DUO101, RJ-DUO149, RJ-DUO160, RJ-DUO191, MARA56, MARA86, MARA91, MARA153, MARA207 e TDF_531	18.2	2023 (6,9E-1)
	COL85, COL86, COL105, COL274, COL497, COL564, MARA 146 e TDF_33	19.2	890 (3,04E-1)
	COL333 e COL505	20.1	1 (1,43E-4)
	COL431	20.2	280 (9,5E-2)
DYS438	COL85, COL105 e COL271	9.2	5 (4,8E-05)
	COL368	9.1	1 (1,43E-4)
	COL12	10.2	19 (2,71E-3)
	RJ135 e COL340	11.2	52 (7,43E-3)
DYS385	COL431	12.2	79 (1,13E-2)
D18383	COL198, COL365 e COL576	13.2	199 (2,84E-2)
	COL606	14.2	72 (1,03E-2)
	COL174, COL618, COL628 e MARA98	17.2	73 (1,04E-2)
	COL196	19.2	9 (1,29E-3)
	EQM119	18.3	3 (1,03E-3)
DYS448	TDF_125	19.3	1 (1,43E-4)
	EQM303	20.2	10 (3,4E-3)
Y-GATA-H4	EQM86	12.1	0

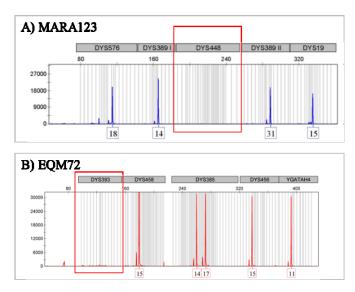
Legenda: Na tabela, são indicadas as microvariantes por *locus* nas amostras estudadas, bem como as ocorrências na base de dados YHRD para o kit PowerPlex® Y23 dentre 103.631 perfis haplotípicos, bem como a frequência alélica (Release R69).

5.1.2 Alelos nulos observados nas populações amostradas

Em alguns *loci*, foram detectados alelos nulos (também denominados alelos silenciosos), caracterizados pela ausência de picos no eletroferograma. Na Figura 18, podem

ser visualizados dois exemplos de alelos nulos encontrados em duas amostras deste estudo nos *loci* DYS448 e DYS393. Os alelos nulos resultam de uma deleção do *locus* ou da presença de mutações na sequência de anelamento do oligonucleotídeo, que impedem a amplificação.

Figura 18. Exemplos de alelos nulos em duas amostras, detectados usando o kit PowerPlex® Y23



Legenda: (A) Alelo nulo em DYS448 observado na amostra MARA123 e (B) Alelo nulo em DYS393 observado na amostra EQM72.

As deleções, duplicações e triplicações são comumente observadas em regiões amplicônicas do cromossomo Y, já que este cromossomo está sob baixa pressão seletiva (BUTLER et al., 2005; BALARESQUE et al., 2008; TURRINA et al., 2015a). As sequências palindrômicas presentes no braço longo do cromossomo Y tornam essas regiões mais suscetíveis a rearranjos estruturais que levam a deleções intracromossômicas e duplicações (SKALETSKY et al., 2003; TURRINA et al., 2015b). No braço longo do cromossomo Y, foram identificadas três regiões, denominadas "fatores de azoospermia" (AZFa, AZFb, AZFc), que desempenham um papel importante na fertilidade masculina (VOGT et al., 1996). Estas regiões estão submetidas a frequentes rearranjos estruturais, o que faz com que alguns dos Y-STRs comumente usados em genética forense, localizados nessas regiões, tendam a apresentar duplicações e/ou deleções (Figura 19).

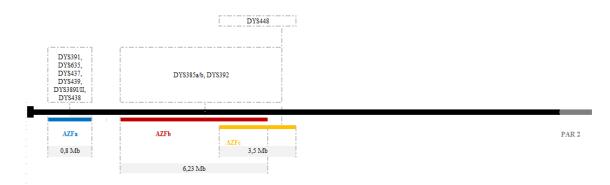


Figura 19. Regiões de azoospermia AZFa, AZFb e AZFc em Yq, e Y-STR

Legenda: Fatores de azoospermia AZFa, AZFb e AZFc no braço longo do cromossomo Y, sinalizando os Y-STRs que estão contidos nessas regiões.

No Maranhão, foram observados alelos nulos em DYS437 e DYS392, que estão localizados respectivamente em AZFa e AZFb (Tabela 10). Na YHRD, é possível observar ocorrências para estes alelos. Para os alelos nulos em DYS437 e DYS392, o NEVGEN indicou respectivamente os haplogrupos R1b e E1b1b, ambos de origem europeia. No entanto, o alelo nulo em DYS437 foi observado em afro-americanos dos EUA e em Gana (PURPS *et al.*, 2014; KOFI *et al.*, 2019).

O alelo nulo em DYS448 foi observado em uma amostra colombiana e outra maranhense (Tabela 10). Alelos nulos e duplicações no DYS448 são comuns, dada a sua localização na região AZFc. Além das ocorrências depositadas em YRHD, este alelo nulo já foi relatado na Argentina (TOSCANINI *et al.*, 2008, CAPUTO *et al.*, 2019), em nativos americanos do Equador (Kichwa Salasaka) (VILLAESCUSA *et al.*, 2021), na África do Sul (REID *et al.*, 2020), em populações europeias, tais como Inglaterra, Eslovênia, Suíça, Dinamarca e Holanda (PURPS *et al.*, 2014), e nos EUA (PURPS *et al.*, 2014). Para o alelo nulo em DYS448, o NEVGEN indicou como mais provável a linhagem europeia R1a. Entretanto, Balaresque e colaboradores (2008) identificaram deleções no DYS488 em amostras pertencentes aos haplogrupos O3e, C*, C3c; G, O2 D*, e E3b. Os dados mostram que terão ocorrido vários eventos de deleção com origens independentes, não sendo possível estabelecer uma origem geográfica única para os alelos nulos encontrados neste trabalho.

Locus	Código de amostra	Alelo	Número de ocorrências e frequência em YHRD		
DYS437	MARA120	-	11(3E-3)		
DYS392	MARA25	-	11(1E-4)		
DYS448	MARA 123 e COL345	-	319 (3E-3)		
DYS549	TDF2 e TDF387	-	9 (8,71E-5)		
DYS393	EQM72	-	6 (5,8E-5)		
DYS458	EQM490	-	30 (2,9E-4)		
DYS533	EQM192, EQM378	-	1 (1,43E-4)		
DYS576	EQM440	-	36 (3,4E-4)		
Y-GATA-H4	MARA83	344pb	0		

Tabela 10. Alelos nulos encontrados nas amostras deste estudo

Legenda: Na tabela, são indicados os alelos nulos por *locus* nas amostras estudadas, bem como as ocorrências na base de dados YHRD para o kit PowerPlex® Y23 dentre 103.280 perfis haplotípicos e respectiva frequência aélica (Release R69).

Em duas amostras fueguinas foi observado um alelo nulo em DYS549. Na YHRD já haviam sido depositados alelos nulos para este *locus*, e já foi relatado na Índia (PURPS *et al.*, 2014). Entretanto, para este alelo nulo, o NEVGEN atribuiu as linhagens europeias I1 e E1b1b, apontando para uma possível origem neste continente.

No Equador, foi observado um alelo nulo em DYS393. Uma possível origem geográfica nativo-americana (Q-M346) foi atribuída pelo NEVGEN. As ocorrências somente em populações da Argentina e Peru reforçam sua origem nativo-americana (ROEWER *et al.*, 2013; PURPS *et al.*, 2014).

Na Terra do Fogo, foi observado um alelo nulo em DYS458. Na YHRD há trinta haplotipos com este alelo nulo. O haplogrupo desta amostra foi determinado pela tipagem de Y-SNPs como R1b-S116, com origem geográfica na Península Ibérica. Entretanto, ocorrências deste alelo nulo foram relatadas em populações fora da Peninsula Ibérica (PURPS et *al.*, 2014).

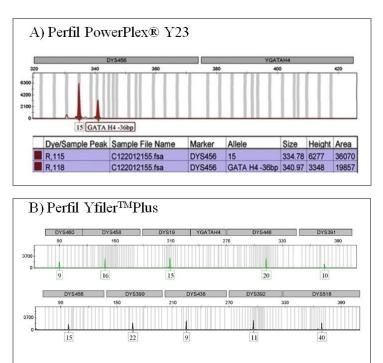
Além destas, duas amostras equatorianas apresentaram um alelo nulo em DYS533. Na literatura, não foi observada nenhuma ocorrência deste alelo, embora esteja reportado na YHRD. Além disso, o NEVGEN atribuiu a linhagem nativo-americana Q-M346 para ambos os perfis observados.

Por fim, na Terra do Fogo foi observado um alelo nulo em DYS576, cuja linhagem atribuída pela tipagem de Y-SNPs pertence ao haplogrupo europeu J2-M172. Na YHRD, já haviam sido relatados alelos nulos em DYS576 em populações europeias e asiáticas (PURPS

et *al.*, 2014), bem como no Brasil (AMBROSIO *et al.*, 2020), de maneira que sua origem provavelmente é euroasiática.

A amostra MARA83 não apresentou pico no eletroferograma, na região dos bins referentes ao Y-GATA-H4 (Figura 20A), o que poderia ser interpretado como um alelo nulo. Entretanto, um pico extra com o tamanho de 344 pb foi observado na região dos bins correspondentes ao marcador DYS456, o qual normalmente apresenta um só alelo.

Figura 20. Amostra MARA83



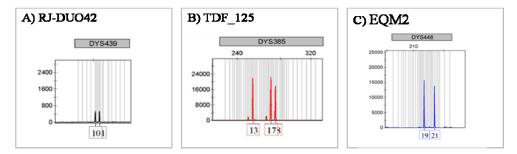
De forma a investigar se o perfil observado correspondia a (i) um alelo nulo no Y-GATA-H4 e duplicação no DYS456; ou (ii) um novo alelo no Y-GATA-H4, cujo tamanho se sobrepõe aos alelos do DYS456, esta amostra foi genotipada com o kit Yfiler® Plus, que possui olinucleotídeos diferentes. Com este outro kit, não foi encontrada duplicação no DYS456 e nenhum alelo foi observado em Y-GATA-H4, caracterizando-se como alelo nulo (Figura 20B). Em conjunto, os resultados obtidos com os dois kits indicam a ocorrência de uma deleção no Y-GATA-H4 de pelo menos 36 pb (correspondente à distância entre o alelo amplificado e o primeiro alelo reportado para este marcador). Portanto, a deleção inclui a região de anelamento de um dos olinucleotídeos incluídos no kit Yfiler® Plus, dando origem a um alelo nulo quando se utiliza este kit. Na literatura e na base de dados YHRD, não foi

observada nenhuma ocorrência deste alelo. No entanto, a linhagem mais provável indicada pelo NEVGEN para este haplótipo é a J2a1, de origem europeia.

5.1.3 Alelos extra (duplicações) observados nas populações amostradas

Em alguns *loci*, foram detectados alelos extra resultantes de duplicações (Figura 21).

Figura 21. Exemplos de alelos extra em 3 amostras, detectados usando o kit PowerPlex® Y23



Legenda: (A) Duplicação em DYS439 observada na amostra RJ-DUO042, (B) Triplicação em DYS385 observada na amostra TFD 125, e (C) Duplicação em DYS448 observada na amostra EQM2.

Geralmente, quando ocorre uma duplicação, ambas as cópias do *locus* (ou *loci*) duplicado(s) apresentam o mesmo alelo, não sendo observados alelos extra no electroferograma. Entretanto, a análise quantitativa da altura do pico no eletroferograma pode demonstrar que duas cópias estão presentes nesse *locus*. Após a duplicação, a ocorrência de mutação num dos *loci* duplicados leva ao aparecimento de um segundo alelo, possibilitando a identificação da duplicação, pela presença de dois picos com alturas semelhantes.

Muitos trabalhos já identificaram duplicações em um ou mais *loci* (DYS635, DYS437, DYS438, DYS439 e DYS389I/II) que estão localizados na região azospermica AZFa (BOSCH e JOBLING 2003; DIEDERICHE *et al.*, 2005).

A amostra TDF_17 (Terra do Fogo) apresentou duplicação para os marcadores DYS439 e DYS389I/II (Tabela 11). O mesmo perfil duplicado nos dois *loci* foi relatado na Argentina (TOSCANINI *et al.*, 2008) e apresenta ocorrências em populações americanas miscigenadas na YHRD.

TC 1 1 11	A 1 1	4 1	4	1 4 4 1
Tabela 11.	A lelos extra	encontrados nas	amostras	deste estudo
rabbia ii.	THOIDS CAHA	chicomitados nas	amosuas	acsic cstado

Locus	Código de amostra	Alelo	Número de ocorrências e frequência em YHRD
DYS439	TDE 17	10-11	9 (7.71E 5)
DYS389I/II	TDF_17	29-30	8 (7,71E-5)
DYS437	COL215	14-15	0
DYS389I	COL213	12-13	Ü
DYS439	RJ-DUO42 e RJ-DUO59	10-11	98 (9,4E-4)
DYS385	TDF_125	13-17-18	12 (1,15E-4)
	COL218 e COL546	20-21	14 (1,35E-4)
DVC 440	EQM2 e COL502	19-21	8 (7,71E-5)
DYS448	COL557	19-20	86 (8,2E-4)
	MARA59	21-22	1 (1,43E-4)
DYS456	RJDUO199	14-16	1 (1,43E-4)
DYS570	EQM152	16-17	4 (3,85E-5)
DVC10	EQM27	15-16	61 (5,88E-4)
DYS19	EQM119	14-15	26 (2,5E-4)

Legenda: Na tabela, são indicadas as duplicações e triplicações por *locus* nas amostras estudadas, bem como as ocorrências na base de dados YHRD para o kit PowerPlex® Y23 dentre 103.631 perfis haplotípicos, bem como a frequência alélica (Release R69).

Além disso, no trabalho de Purps e colaboradores (2014) foi relatada uma duplicação envolvendo vários *loci*, em uma amostra da população hispânica dos Estados Unidos (DYS439*10,11; DYS389I*13,14; DYS389II *29,30 e DYS437* 14,15), e no trabalho de Diederich e colaboradores (2005) uma mesma duplicação na Bahia e em Madri (DYS439*10,11; DYS389II *29,30 e DYS437* 14,15). A amostra Co215 (Colômbia) apresentou uma duplicação para os dois *loci* DYS437 e DYS389I. Embora não tenham ocorrências em YRHD para os dois *loci* duplicados simultaneamente, a duplicação observada no *locus* DYS437 foi anteriormente relatada na Espanha (VALVERDE *et al.*, 2012) e em hispânicos dos EUA (PURPS *et al.*, 2014).

As amostras RJ-DUO42 e RJ-DUO59 apresentaram uma duplicação para o marcador DYS439. Esta duplicação foi encontrada em perfis depositados na YHRD e anteriormente relatada em populações do Rio de Janeiro (OLIVEIRA *et al.*, 2014; AMBROSIO *et al.*, 2020).

Para todos os haplótipos com *loci* duplicados envolvendo DYS437, DYS439 e DYS389I/II, o NEVGEN indicou como mais provável o haplogrupo R1b1a-M269. Levando em conta as ocorrências destes haplótipos em populações ibéricas e sul-americanas miscigenadas, é possivel prever que a duplicação de toda uma seção do cromossomo Y, em

que estão localizados os *loci* DYS437, DYS439 e DYS389I/II, ocorreu na Península Ibérica, em algum momento no passado.

No *locus* DYS385, localizado na região AZFb (BOSCH e JOBLING, 2003), foram observados três alelos numa amostra da Terra do Fogo (Tabela 11). Para este haplótipo, o haplogrupo mais provável indicado pelo NEVGEN foi o J2a1, com sua origem e maior frequência em populações da Europa. Entretanto, na YHRD observam-se ocorrências em populações do leste asiático, de maneira que não é possível atribuir uma origem geográfica única para o haplótipo.

O *locus* DYS448 duplicado, observado no Maranhão, no Equador, e na Colômbia, está localizado na região AZFc (BALARESQUE *et al.*, 2007). Ocorrências destas duplicações podem ser observadas na YHRD. Vários estudos populacionais já relataram duplicações neste *locus* em populações sul-americanas (PALHA *et al.*, 2012; PUPS *et al.*, 2016; TOSCANINI *et al.*, 2018; AMBROSIO *et al.*, 2020; VILLAESCUSA *et al.*, 2021), europeias, por exemplo, Inglaterra e Suíça, e africanas, tais como Benim, Quênia e Nigéria (PUPS *et al.*, 2016). Entretanto, para as duplicações em DYS448 observadas neste trabalho, o NEVGEN atribuiu os haplogrupos europeu R1b e africanos E1b1a e E1a. Além disso, o trabalho de Balaresque e colaboradores (2008) indicou que esta duplicação provavelmente ocorreu em eventos independentes, dentro de linhagens europeias G, asiáticas O3e, nativo-americanas Q e africanas E1, de maneira que não é possível indicar uma origem geográfica para as duplicações deste trabalho.

Foram também observadas duplicações em DYS456 no Rio Janeiro e no Equador, e em DYS570 no Equador (Tabela 11). Além das ocorrências relatadas para estas duplicações na YHRD, Purps e colaboradores (2014) observaram a duplicação no *locus* DYS456 em Benim, no Alaska e em hispânicos dos EUA, e a duplicação no *locus* DYS570 na Holanda. Tendo em vista que o NEVGEN indicou a linhagem R1b1a-M269 como a mais provável para a duplicação em DYS456 e a linhagem J2a1 para a duplicação em DYS570, sugere-se uma possível origem europeia para estes alelos extra.

Ademais, uma duplicação no *locus* DYS19 foi observada no Equador (Tabela 11). Além das ocorrências em YHRD, estudos populacionais anteriormente relatm a duplicação 15-16 no Brasil (PALHA *et al.*, 2013; FIGUEIREDO *et al.*, 2016), Argentina (PURPS *et al.*, 2014; CAPUTO *et al.*, 2019), Colômbia (ROMERO *et al.*, 2008) e em populações euro-asiáticas (ZERJAL *et al.*, 2004; NASIDSE *et al.*, 2005; PURPS *et al.*, 2014). Entretanto, o haplótipo da amostra EQM337 não foi relatado na literatura. Capelli e colaboradores (2007) encontraram esta duplicação em amostras da Itália pertencentes ao haplogrupo europeu G2a.

Esta duplicação é também encontrada por outros autores em populações do Cáucaso e da Ásia, estando associada ao haplogrupo C3a, de origem asiática, o que indica mais de uma origem possível (ZERJAL *et al.*, 2004; NASIDSE *et al.*, 2005). Entretanto, para este haplótipo, o NEVGEN indicou como haplogrupo mais provável a linhagem G2a2a de origem europeia, confirmada pela tipagem de Y-SNPs.

Duplicações e triplicações, bem como microvariantes e deleções em determinados *loci* ao longo do cromossomo Y podem fornecer vantagens em forense no que diz respeito à força de uma coincidência entre uma amostra questionada e de referência. Além disso, na análise forense, a observação de vários picos em um determinado *locus* pode sugerir a presença de uma mistura de perfis de DNA de vários agressores, indiciando uma situação de estupro coletivo (BUTLER *et al.*, 2005).

5.1.4 <u>Diversidade haplotípica</u>

Neste trabalho, foi avaliada a diversidade genética presente nas diferentes populações estudadas e feita a sua comparação com a diversidade reportada para outras populações da América do Sul.

Nas populações deste estudo, os valores de HD se mantiveram próximos àqueles anteriormente observados para outras populações sul-americanas (PURPS *et al.*, 2014; TOSCANINI *et al.*, 2016; AMBROSIO *et al.*, 2020), exceto para o Paraguai (RIBEIRO *et al.*, 2018), em que se observa um valor ligeiramente menor (Tabela 12).

Entre as 190 amostras da Terra do Fogo estudadas, foram identificados 187 haplótipos únicos e uma HD de $0,9998 \pm 0,0006$. Desta maneira, um total de três haplótipos foi compartilhado por dois indivíduos. A população da Terra do Fogo apresentou um valor de HD semelhante ao de outras populações da região sul da Argentina (Tabela 12) (TOSCANINI *et al.*, 2016).

Tabela 12. Diversidade haplotípica para 23 *loci* Y-STR em diferentes populações miscigenadas sul-americanas

População	N	Número de Haplótipos	Diversidade haplotípica
Colômbia (este trabalho)	667	630	$0,9998 \pm 0,0001$
Equador (este trabalho)	446	416	$0,9997 \pm 0,0002$
Maranhão (este trabalho)	229	223	$0,9998 \pm 0,0004$
Rio de Janeiro (este trabalho)	201	201	$1,0000 \pm 0,0004$
Terra do Fogo (este trabalho)	190	187	$0,9998 \pm 0,0006$
Argentina (PURPS et al., 2014)	766	754	$1,0000 \pm 0,0001$
Argentina (Região Sul) (TOSCANINI et al ., 2016)	130	128	$0,9998 \pm 0,0010$
Bolívia (PURPS et al ., 2014)	56	55	$1,0000 \pm 0,0050$
Brasília (PURPS et al., 2014)	409	322	$0,9987 \pm 0,0003$
Equador (TOSCANINI et al., 2018)	270	269	$0,9998 \pm 0,0000$
Paraguai (RIBEIRO et al., 2018)	537	480	$0,9994 \pm 0,0002$
Peru (PURPS et al ., 2014)	83	72	$0,9965 \pm 0,0028$
Rio de Janeiro (PURPS et al ., 2014 e AMBROSIO et al ., 2020)	324	323	$0,99998 \pm 0,0005$
São Paulo (PURPS et al., 2014)	314	313	$1,0000 \pm 0,0003$

Legenda: Valores de diversidade haplotípica nas diferentes populações analisadas neste estudo e em trabalhos já publicados para as mesmas regiões/países. Os haplótipos foram definidos com base nos seguintes *loci*: DYS19, DYS385, DYS389I, DYS389I, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS481, DYS533, DYS549, DYS570, DYS576, DYS635, DYS643 e Y-GATA-H4.

A HD da Colômbia (N=667) é de $0,9998 \pm 0,0001$. Nesta amostra, há 630 haplótipos únicos. Um total de três haplótipos foi compartilhado por três indivíduos e 31 por dois indivíduos.

Para as amostras do Equador (N=446), a HD encontrada foi de $0,9997 \pm 0,0002$, com 416 haplótipos únicos. Consequentemente, dois haplótipos foram compartilhados por três indivíduos e quatro por dois indivíduos. A amostra do Equador apresentou uma HD menor do que a encontrada por Toscanini e colaboradores (2018) em outra amostra populacional equatoriana.

No Maranhão, foram observados cinco haplótipos compartilhados por dois indivíduos e 223 haplótipos únicos. A HD observada para esta população foi de 0,9998 \pm 0,0004. Por outro lado, na amostra do Rio de Janeiro (N=201) observou-se uma HD de 1,0000 \pm 0,0004 e nenhum haplótipo foi observado em mais de uma amostra.

O valor da HD encontrada para o Rio de Janeiro foi maior do que o relatado em uma amostra combinada de haplótipos analisados por Purps e colaboradores (2014) e Ambrosio e colaboradores (2020) (N = 324; HD = 99,99%) (Tabela 12).

Os elevados valores de HD calculados para as populações deste estudo estão de acordo com o esperado para populações sul-americanas, levando em conta a miscigenação, pelo aporte de pelo menos três continentes diferentes, resultantes dos eventos históricos como a colonização, o comércio transatlântico de escravos e as imigrações pós colonização em larga escala.

5.1.5 Análise de distâncias genéticas (F_{ST})

Com o objetivo de avaliar a diferenciação entre as populações deste estudo e outras populações da América do Sul com dados disponíveis para o kit PowerPlex® Y23, foram calculadas as distâncias genéticas F_{ST} e os respectivos valores de probabilidade de não diferenciação (p) (Tabela 13). As distâncias genéticas pareadas foram calculadas com base em 22 Y-STRs, após excluir o DYS385, por possuir 2 *loci*, cujos alelos não podem ser discriminados. Para efeitos de interpretação dos resultados, foram incluídas na análise populações de referência europeia, nativo-americana e africana, que se sabe terem contribuído para a composição atual das populações estudadas. Para além das nossas amostras, foram incluídas na análise as seguintes populações: europeias – Espanha, Portugal, Itália, Alemanha e Reino Unido (PURPS *et al.*, 2014), africanas – África do Sul_Xhosa, Norte-Benin e Yoruba (PURPS *et al.*, 2014), nativo-americanas – Bolívia e São Gabriel da Cachoeira (PURPS *et al.*, 2014), e as sul-americanas – Argentina, Bolívia, Peru, Brasília, São Paulo, (PURPS *et al.*, 2014), Rio de Janeiro (PURPS *et al.*, 2014; AMBROSIO *et al.*, 2020), Chile (TOSCANINI *et al.*, 2016), Patagônia (PAROLIN *et al.*, 2019), e Paraguai (SIMÃO *et al.*, 2021).

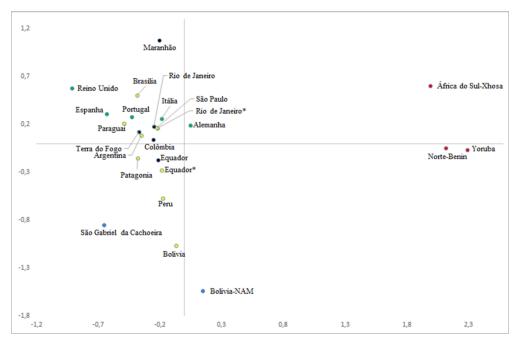
Tabela 13. Matriz das distâncias genéticas de $F_{\rm ST}$ entre todos os pares de populações, juntamente com os correspondentes valores de probabilidade de não diferenciação obtidos para 10.000 permutações

	Maranhão	Rio de Janeiro	Rio de Janeiro*	Brasília	São Paulo	São Gabriel de Cachoeira	Bolivia_NAM	Bolivia	Equador	Equador*	Peru	Colombia	Terra do Fogo	Argentina_Purps	Patagonia-Central	Paraguai	Spain	Portugal	Reino Unido	Alemanha	Itália	Norte-Benin	África do Sul_Xhosa	Yoruba
Maranhão		0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Rio de Janeiro	0,0557		0,98119	0,00000	0,74111	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,02138	0,02633	0,00446	0,0002	0,00109	0,00000	0,04564	0,0002	0,00000	0,0001	0,00000	0,00000	0,00000
Rio de Janeiro"	0,0571	-0,0014		0,00000	0,42441	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,0004	0,00861	0,00000	0,00000	0,00000	0,00000	0,00762	0,0001	0,00000	0,00000	0,00000	0,00000	0,00000
Brasília	0,0712	0,0181	0,0171		0,00000	0,00000	0,00000	0,00000	0,00000	0.00000	0,00000	0,00000	0,00000	0,00000	0,00000	0.00000	0,00000	0,00000	0,00000	0.00000	0,00000	0,00000	0,00000	0,00000
São Paulo	0,0556	-0,0008	0,0000	0,0170		0,00000	0,00000	0,00000	0,00000	0.00000	0,00000	0,01851	0,0294	0,00436	0,00000	0,0001	0,00000	0,00941	0,0001	0,00000	0,0002	0.00000	0,00000	0,00000
São Gabriel de Cachoe	0,1193	0,0748	0,0696	0,0806	0,0676		0,00000	0,00000	0,00000	0.00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0.00000	0,00000	0,00000
Bolivia_NAM	0,1689	0,1242	0,1204	0,1477	0,1194	0,0868		0,04079	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0.00000	0,00000	0,00000	0,00000	0,00000
Bolivia	0,1291	0,0810	0,0808	0,1004	0,0803	0,0635	0,0108		0,00000	0,0001	0,00376	0.00000	0.00000	0,00000	0.00000	0,00000	0,00000	0,00000	0,00000	0.00000	0,00000	0.00000	0,00000	0,00000
Equador	0,0663	0,0124	0,0125	0,0311	0,0106	0,0423	0,0729	0,0368		0,36432	0,0002	0,00000	0,0002	0,00000	0,00624	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0.00000	0,00000	0,00000
Equador*	0,0729	0,0191	0,0189	0,0370	0,0172	0,0395	0,0600	0,0274	0,0002		0,01277	0,00000	0.00000	0.00000	0,00238	0,00000	0.00000	0,00000	0,00000	0.00000	0,00000	0.00000	0,00000	0.00000
Peru	0,0957	0,0416	0,0414	0,0611	0,0395	0,0515	0,0473	0,0196	0,0126	0,0064		0,00000	0,00000	0,00000	0.00000	0,00000	0,00000	0,00000	0,00000	0.00000	0,00000	0,00000	0,00000	0.00000
Colombia	0,0580	0,0024	0,0029	0,0191	0,0018	0,0580	0,1044	0,0659	0,0059	0,0093	0,0285		0,01366	0,00139	0,00089	0,00000	0,00000	0,0004	0,00000	0,00000	0,00000	0.00000	0,00000	0,00000
Terra do Fogo	0,0591	0,0036	0,0034	0,0219	0,0029	0,0647	0,1216	0,0804	0,0096	0,0155	0,0345	0,0028		0,62043	0,0002	0,00871	0,0002	0,06019	0,00000	0.00000	0,09504	0,00000	0,00000	0,00000
Argentina_Purps	0,0606	0,0038	0,0039	0,0184	0,0027	0,0611	0,1114	0,0705	0,0067	0,0122	0,0302	0,0021	-0,0004		0,0003	0,0002	0,00000	0,00535	0,00000	0.00000	0,00000	0,00000	0,00000	0,00000
Patagonia-Central	0,0703	0,0136	0,0130	0,0264	0,0141	0,0508	0,0876	0,0456	0,0047	0,0065	0,0197	0,0072	0,0118	0,0073		0,0001	0,00000	0,00000	0,00000	0,00000	0,00000	0.00000	0,00000	0,00000
Paraguai	0,0618	0,0056	0,0061	0,0192	0,0068	0,0759	0,1391	0,0906	0,0160	0,0237	0,0466	0,0066	0,0038	0,0032	0,0131		0,00426	0,06782	0,00228	0.00000	0,00000	0.00000	0,00000	0,00000
Spain	0,0700	0,0121	0,0127	0,0245	0,0136	0,0905	0,1629	0,1076	0,0248	0,0341	0,0588	0,0122	0,0075	0,0080	0,0194	0,0022		0,00267					0,00000	
Portugal	0,0607	0,0026	0,0033	0,0197	0,0037	0,0828	0,1462	0,0983	0,0177	0,0266	0,0506	0,0064	0,0023	0,0031	0,0167	0,0014	0,0039		0,00208	0,00000	0,00000	0,00000	0,00000	0,00000
Reino Unido	0,0862	0,0324	0,0317	0,0377	0,0350	0,1289	0,2172	0,1473	0,0505	0,0628	0,0904	0,0343	0,0311	0,0301	0,0438	0,0168	0,0098	0,0179		0.00000			0,00000	
Alemanha	0,0717	0,0157	0,0144	0,0358	0,0157	0,0817	0,1201	0,0863	0,0240	0,0284	0,0459	0,0178	0,0135	0,0176	0,0292	0,0249	0,0331	0,0196	0,0453		0,00000	0,00000	0,00000	0.00000
Itália	0,0584	0,0069	0,0057	0,0196	0,0053	0,0724	0,1265	0,0911	0,0177	0,0240	0,0440	0,0086	0,0013	0,0054	0,0208	0,0096	0,0161	0,0065	0,0394	0,0153		0,00000	0,00000	0,00000
Norte-Benin	0,1998	0,1502	0,1463	0,1819	0,1441	0,2265	0,2192	0,2193	0,1605	0,1566	0,1864	0,1493	0,1772	0,1748	0,1862	0,1964	0,2218	0,1903	0,2820	0,1653	0,1627		0,00000	0,13048
África do Sul_Xhosa	0,20236	0,15203	0,14316	0,17025	0,14512	0,20155	0,20747	0,20529	0,15467	0,15032	0,17904	0,14262	0,17099	0,1642	0,17561	0,19117	0,21187	0,18738	0,27512	0,15447	0,15793	0,05701		0,00000
Yoruba	0,22104	0,17402	0,16749	0,20364	0,16613	0,23345	0,22751	0,23822	0,1785	0,17505	0,20612	0,16866	0,19888	0,19488	0,20976	0,21923	0,24561	0,21375	0,30892	0,18218	0,18287	0,00454	0,0651	

Legenda: Abaixo da diagonal, estão os valores de $F_{\rm ST}$; Acima da diagonal, estão os valores de probabilidade de não diferenciação, computados com 10.000 permutações. Os valores de p abaixo do nível de significância (0,000181, após correção de Bonferroni) estão grifados em negrito.

Os valores de distâncias genéticas ($F_{\rm ST}$) pareadas foram representados no MDS *plot* que se apresenta na Figura 22. Os resultados apontam para uma maior proximidade das populações deste trabalho às populações ibéricas e outras populações sul-americanas (Argentina, Brasília, Equador, Patagônia, Rio de Janeiro e São Paulo), como se pode observar pelo seu aglomeramento em um mesmo quadrante (Figura 22). Estes resultados estão de acordo com o esperado, já que, historicamente, a conquista da América do Sul por europeus ocorreu principalmente por aqueles oriundos da Península Ibérica.

Figura 22. MDS *plot* das distâncias genéticas $F_{\rm ST}$ pareadas, calculadas entre a Terra do Fogo, a Colômbia, o Equador, o Rio de Janeiro, o Maranhão e outras populações da América do Sul, bem como populações de referência europeia, africana e nativo-americana



Legenda: MDS *plot* das distâncias genéticas $F_{\rm ST}$ pareadas, calculadas entre a Terra do Fogo, a Colômbia, o Equador, o Rio de Janeiro e o Maranhão, destacados em preto, e outras populações de referência europeias em verde (Espanha, Portugal, Itália, Alemanha, Itália e Reino Unido), africanas em vermelho (África do Sul_Xhosa, Norte-Benin e Yoruba), nativo-americanas em azul (São Gabriel da Cachoeira e Bolívia_NAM) e sul-americanas em amarelo (Argentina, Bolívia, Brasília, Equador, Paraguai, Peru, Rio de Janeiro, Brasília e São Paulo). As populações do Equador, e Rio de Janeiro assinaladas com * são de trabalhos previamente publicados por outros autores. *Stress*=0,0788.

A nossa amostra do Equador apontou proximidade com outras amostras previamente estudadas do Equador (TOSCANINI *et al.*, 2016), não tendo sido encontradas diferenças significativas entre elas. Da mesma maneira, a amostra do Rio de Janeiro não apresenta distancias genéticas significativas com uma amostra estudada anteriormente da mesma

população (PURPS *et al.*, 2014; AMBROSIO *et al.*, 2020). Além disso, a população da Terra do Fogo não apresentou distâncias significativas com a Argentina e a Patagônia (PURPS *et al.*, 2014; PAROLIN *et al.*, 2019).

A população do Maranhão, embora esteja no aglomerado com as outras populações sul-americanas mais próximas das populações europeias, apresenta-se a uma certa distância destas e possui diferenças significativas com todas as amostras analisadas (Figura 22 e Tabela 13). Este distanciamento sugere que esta população carrega um componente ancestral, provavelmente europeu, que poderá estar presentes em populações não usadas nesta análise, como por exemplo da França e Holanda que historicamente impactaram na colonização deste Estado.

Longe do aglomerado observado no MDS *plot* estão as populações nativo-americanas, e mais próxima delas as populações miscigenadas da Bolívia e do Peru, o que provavelmente deve-se ao maior componente nativo presente nestas populações. A população nativo-americana de São Gabriel da Cachoeira é a única brasileira que se distancia das demais, devido ao grande aporte do componente nativo-americano. Além disso, São Gabriel da Cachoeira e os nativos da Bolívia estão relativamente distantes entre si (Figura 22), reforçando que o componente nativo andino da Bolívia é diferente do amazônico encontrado em São Gabriel da Cachoeira. Por outro lado, as populações africanas estão próximas umas das outras e relativamente longe das demais, apontando para uma menor contribuição africana nas populações estudadas. Além do mais, todas as amostras deste estudo evidenciaram diferenças significativas com as populações nativo-americanas e africanas.

A partir das distâncias genéticas observadas, conclui-se que as populações sulamericanas estão mais próximas das europeias, seguidas pelas nativo-americanas e africanas, exceto Bolívia e Peru que estão mais próximas das populações nativo-americanas.

5.2 Análise populacional de haplogrupos definidos por Y-SNPs

Com o objetivo de determinar a ancestralidade paterna de três das populações incluídas neste trabalho (Equador, Paraguai e Terra do Fogo), foi realizada a genotipagem de marcadores Y-SNP, agrupados de maneira hierárquica nos ramos da árvore filogenética. Devido as baixas taxas de mutação, em comparação com os Y-STRs, os Y-SNPs permitem a definição de haplogrupos com elevada especificidade geográfica.

5.2.1 <u>Diversidade de haplogrupos</u>

Uma elevada diversidade de haplogrupos do cromossomo Y foi observada nas populações equatoriana, paraguaia e fueguina (Tabela 14). Os valores encontrados estão de acordo com o esperado, levando em consideração a mistura de linhagens europeias, nativo-americanas e africanas, resultante dos eventos históricos na formação das populações atuais miscigenadas. Valores semelhantes foram descritos em outras populações miscigenadas sul-americanas, exceto a Bolívia e o estado brasileiro de Alagoas que apresentam uma diversidade de haplogrupos menor que as demais populações (Tabela 14).

O valor de diversidade de haplogrupos mais baixo encontrado na Bolívia pode ser atribuído ao menor grau de miscigenação desta população, com um maior componente nativo-americano. Entretanto, o valor de HD observado em Alagoas muito provavelmente está associado ao fato deste estudo populacional ter genotipado menos Y-SNPs (por exemplo, não foram determinados os subtipos do haplogrupo R), subestimando a diversidade de haplogrupos presente na população, em comparação com outros trabalhos.

Tabela 14. Diversidade de haplogrupos em populações miscigenadas sul-americanas, bem como suas referências.

População	Diversidade	Referência
Equador	$0,8674 \pm 0,0008$	Este trabalho
Paraguai	$0,849 \pm 0,0006$	Este trabalho
Terra do Fogo	$0,8886 \pm 0,0017$	Este trabalho
Bolívia	$0,6788 \pm 0,0497$	VULLO et al ., 2014
Brasil (Sudeste)	$0,7356 \pm 0,0007$	OLIVEIRA et al., 2014
Brasil	$0,856 \pm 0,007$	RESQUE et al., 2016
Brasil (Sudeste)	$0,9323 \pm 0,0002$	JANNUZZI et al., 2020
Brasil	$0,8032 \pm 0,0020$	JOERIN-LUKE et al., 2022
Alagoas	$0,6735 \pm 0,0321$	AZEVEDO et al., 2009
Espírito Santo	$0,7794 \pm 0,0229$	FIGUEIREDO et al ., 2016
Chile	0,7438	FLORES-ALVARADO et al., 2022
Chile	$0,8067 \pm 0,0013$	REYES-MADRID et al., 2022
Colômbia	$0,7310 \pm 0,0038$	ACOSTA et al., 2009
Colômbia	$0,801 \pm 0,025$	NOGUERA et al., 2013
Colômbia	0,8692	ALONSO MORALES et al., 2018
Equador	$0,8359 \pm 0,0025$	VILLAESCUSA et al., 2020

Nota: No trabalho de Azevedo *et al.* (2009) não foram analisados os subtipos do haplogrupo R, o que muito provavelmente explica o valor de HD menor que os demais.

5.2.2 <u>Frequência e origem dos haplogrupos encontrados</u>

Dos 67 haplogrupos possíveis de serem classificados com os 67 Y-SNPs estudados, somente 29 foram observados no Equador, 27 no Paraguai e 18 na Terra do Fogo. As frequências dos haplogrupos detectados estão apresentadas na Tabela 15. Os resultados do estudo das populações do Paraguai e da Terra do Fogo foram descritos nos artigos "The Ancestry of Eastern Paraguay: A Typical South American Profile with a Unique Pattern of Admixture" e "Tierra Del Fuego: What Is Left from the Precolonial Male Lineages?", publicados na revista Genes (Anexos G e H).

Nas populações estudadas, foi observado um predomínio de linhagens paternas europeias, em detrimento das linhagens nativo-americanas ou africanas (CORACH *et al.*, 2010; NOGUERA *et al.*, 2013; VULLO *et al.*, 2014; RESQUE *et al.*, 2016; GARCÍA *et al.*, 2018). Este efeito é semelhante ao relatado em outros estudos para populações miscigenadas sul-americanas e atribuído aos fortes efeitos da colonização europeia, levando a uma diminuição drástica ou mesmo ao desaparecimento de muitas comunidades nativas, especialmente no que se refere aos homens.

O haplogrupo mais frequente em todas as populações foi o R1b1a-S116 (Tabela 15). A maior frequência do haplogrupo R1b1a-S116 também foi anteriormente observada nas populações do Brasil e Colômbia (RESQUE *et al.*, 2016; MORALES *et al.*, 2018). O segundo e terceiro haplogrupos mais frequentes no Equador foram Q-M3* e R1b1a-M529; no Paraguai, J2-M172 e R1b1a-M529; e na Terra do Fogo o G-M201 e J2-M172, respectivamente (Tabela 15). Os haplogrupos com menores frequências no Paraguai são C-M130, E1a-M33, O3a-M117 e Q1a2-Z19319*(xSA01), e no Equador são E1a-M33, E1b1b-M35, E2-M85, E1b1a-U174, K-M9*(xTat, 92R7, M70) e R1b1a-M167. Por outro lado, os haplogrupos observados em menores frequência na Terra do Fogo são E1b1b-M81 e E1b1a-U209, R1b1a-L23 e R1b1a-U106 (Tabela 15).

Os haplogrupos E1b1b-M78, R1a, R1b1a-M153, T1a-M70, E1b1a-M2, E1b1b-M123, bem como as sublinhagens de I-M170 e J-12f2 foram observados em todas as populações estudadas (Tabela 15).

Tabela 15. Frequência dos haplogrupos no Equador, Paraguai e Terra do Fogo

Haplogrupo	Equador	Paraguai	Terra do Fogo	Ancestralidade
C-M130	0,00%	0,22%	0,00%	ASIAT
E1 a-M33	0,36%	0,22%	0,00%	AFR
E1 b1 a-M2	0,00%	0,65%	0,00%	AFR
El bl a-U174	0,36%	0,00%	0,00%	AFR
E1 b1 a-U209*(xM154, U290)	1,79%	0,00%	0,00%	AFR
E1 b1 a-U290	0,00%	0,00%	0,51%	AFR
E1 b1 b-M35*(xM78, M81,M123,V6,M293)	1,79%	0,65%	0,51%	AFR
E1 b1 b-M78	4,30%	3,89%	5,10%	AFR/EUR
E1 b1 b-M81	3,94%	3,24%	1,53%	EUR
E1 b1 b-M123	3,58%	0,43%	3,06%	EUR
E2-M85	0,36%	0,00%	0,00%	AFR
G-M201	3,94%	5,18%	8,16%	EUR
I-M170	4,30%	8,86%	6,63%	EUR
I2a-M26	2,15%	0,00%	2,04%	EUR
J-12f2a*(xM62, M172)	3,70%	0,00%	3,57%	EUR
J1-M267	0,00%	1,30%	0,00%	EUR
J1-P58	0,00%	1,51%	0,00%	EUR
J2-M172	6,81%	6,91%	12,24%	EUR
L1-M22	0,00%	0,43%	0,00%	EUR
O3a-P164(xM134)	0,00%	0,43%	0,00%	ASIAT
O3a-M117	0,00%	0,22%	0,00%	ASIAT
K-M9*(xTat, 92R7, M70)	0,36%	0,00%	0,51%	EUR
Q1b-M346*(xM3)	0,72%	0,00%	0,00%	NAM
Q1b-M3	15,41%	4,97%	3,06%	NAM
Q1b-M3*(xM19, Z19319, Z19483, M557, SA05)	2,15%	0,00%	0,00%	NAM
Q1b-Z19483	0,72%	0,22%	2,04%	NAM
R1a-SRY10831.2	0,72%	1,73%	3,06%	EUR
R1b-M269*(Xv88, L23)	0,00%	0,00%	1,02%	EUR
R1b1a-L23*(xU106, S116)	1,08%	2,59%	0,51%	EUR
R1b1a-U106	1,08%	4,10%	0,51%	EUR
R1b1a-S116*(xU152, M529, M153, M167)	29,39%	34,34%	25,00%	EUR
R1b1a-M153	2,51%	1,30%	2,04%	EUR
R1b1a-M167	0,36%	1,94%	2,55%	EUR
R1b1a-U152	3,94%	4,54%	7,14%	EUR
R1b1a-M529	3,23%	7,78%	6,12%	EUR
R1b1b-V88	0,00%	0,00%	1,02%	AFR
T1a-M70	1,08%	2,38%	2,04%	EUR

Legenda: Frequência relativa (%) dos haplogrupos observados no Equador, Paraguai e Terra do Fogo, bem como a ancestralidade associada.

5.2.2.1 Haplogrupos de origem europeia

O haplogrupo R foi o mais frequente em todas as populações, com diferentes linhagens incluídas dentro dos macro-haplogrupos R1a e R1b.

Atualmente, a distribuição de R1a na Europa mostra um gradiente crescente de frequência e variância oeste-leste com picos entre os falantes finno-úgricos e eslavos (ROSSER *et al.*, 2000, SEMINO *et al.*, 2000; ROZHANSKII e KLYOSOV, 2012), diminuindo lentamente para o sul da Europa (10% nos albaneses, 8% em gregos e 7% em turcos) e abruptamente para o oeste (3% em italianos) (ROSSER *et al.*, 2000, SEMINO *et al.*, 2000; PERICIĆ *et al.*, 2005b). A mutação SRY10832.1 foi observada nas populações sulamericanas, em frequências menores de 5% (CORACH *et al.*, 2010; NOGUERA *et al.*, 2013), semelhante ao observado nas populações estudadas deste trabalho.

O R1b-M269 é o haplogrupo mais frequente na Europa ocidental (REGUEIRO *et al.*, 2015; SOLÉ-MORATA *et al.*, 2017; VILLAESCUSA *et al.*, 2017; GRUGNI *et al.*, 2018; HERNANDEZ *et al.*, 2019), apresentando elevadas frequências em toda a Península Ibérica, particularmente maiores que 90% em populações do País Basco e dos Pirenéus (ALONSO *et al.*, 2005; GONÇALVES *et al.*, 2005; BELEZA *et al.*, 2006; NOGUEIRO *et al.*, 2010; MARTINIANO *et al.*, 2013; SAIZ *et al.*, 2019). Nas populações estudadas, o haplogrupo ancestral R1b-M269*(xV88, L23) apareceu apenas na Terra do Fogo.

O haplogrupo R1b1a-L23*(xU106, S116) é característico do Médio Oriente, com frequência superior a 10% em populações do Cáucaso, Turquia e em algumas populações do sudeste da Europa. Na Europa Ocidental ele é observado em frequências menores, no excedendo os 5% (MYRES *et al.*, 2011). Nas populações estudadas, o haplogrupo R1b1a-L23*(xU106, S116) foi observado em baixas frequências (>4%), à semelhança do que tem sido relatado para outras populações do Brasil e da Bolívia (VULLO *et al.*, 2014; RESQUE *et al.*, 2016; JANNUZZI *et al.*, 2020; SCHANN *et al.*, 2020; JOERIN-LUKE *et al.*, 2022).

Os subtipos do haplogrupo R1b1a-L23 (nomeadamente, R1b1a-U106, R1b1a-U152, R1b1a-S116, R1b1a-M167, R1b1a-M529 e R1b1a-M153) são os mais frequentes na Europa Ocidental (ZEI *et al.*, 2003; ALONSO *et al.*, 2005; BOATTINI *et al.*, 2013; SOLÉ-MORATA *et al.*, 2017; VILLAESCUSA *et al.*, 2017; GRUGNI *et al.*, 2018; MORALES *et al.*, 2018; HERNANDEZ *et al.*, 2019), com as maiores frequências relatadas para populações na Irlanda (MOORE *et al.*, 2006). Este haplogrupo possui também elevadas frequências em toda a Península Ibérica, particularmente maiores que 90% em populações do País Basco e

dos Pirenéus (ALONSO *et al.*, 2005; BELEZA *et al.*, 2006; NOGUEIRO *et al.*, 2010; MARTINIANO *et al.*, 2013; SAIZ *et al.*, 2019). Além das populações europeias, tem sido também relatado o seu predomínio em populações miscigenadas sul-americanas, como o Brasil, Bolívia, Colômbia e Argentina, sendo a sub-linhagem R1b1a-S116 a mais frequente (CORACH *et al.*, 2010; ROJAS *et al.*, 2010; NOGUERA *et al.*, 2013; OLIVEIRA *et al.*, 2014; RESQUE *et al.*, 2016; PEREZ-BENEDICO *et al.*, 2016; CRIOLLO-RAYO *et al.*, 2018; MORALES *et al.*, 2018; VILLAESCUSA *et al.*, 2021; JOERIN-LUKE *et al.*, 2022). A sub-linhagem R1b1a-S116 foi também a mais frequente nas amostras da Terra do Fogo, do Paraguai e do Equador, em que apareceu com frequências de mais de 25%.

As sub-linhagens R1b1a-M167 e R1b1a-M153 foram observadas nas populações deste estudo com frequências inferiores a 3%. As frequências mais elevadas para estes dois haplogrupos foram descritas em populações do País Basco e dos Pirenéus (ALONSO *et al.*, 2005; LÓPEZ-PARRA *et al.*, 2009; MYRES *et al.*, 2011; SOLÉ-MORATA *et al.*, 2017).

Por outro lado, o haplogrupo R1b1a-U152 foi observado nas nossas amostras com valores aproximados de frequência entre de 4%-7%. Este haplogrupo é mais frequente na Suíça, Itália, França e Polônia Ocidental. Na Península Ibérica, o haplogrupo R1b1a-U152 mostra uma distribuição caracteristica com picos de frequência na região costeira e na região alpina (BOATTINI *et al.*, 2013; VALVERDE *et al.*, 2015; GRUGNI *et al.*, 2018).

O sub-haplogrupo R1b1a-U106 foi observado com frequências de 0,5% a 4% nas populações estudadas. Este haplogrupo é mais frequente na Europa Central e Oriental (MYRES *et al.*, 2011; CRUCIANI *et al.*, 2011; HERNANDEZ *et al.*, 2019).

Dentro do paragrupo E, foram detectadas as seguintes linhagens: E1b1b-M78, E1b1b-M81 e E1b1b-M123. A linhagem E1b1b-M78 tem uma ampla distribuição, incluindo Europa, Oriente Médio e Norte, e Leste da África (BOSCH *et al.*, 2001; UNDERHILL *et al.*, 2001; CRUCIANI *et al.*, 2002; ARREDI *et al.*, 2004; CINNIOGLU *et al.*, 2004; GONÇALVES *et al.*, 2005; REGUEIRO *et al.*, 2006; FADHLAOUI-ZID *et al.*, 2014). Esta é a linhagem E mais comum na Europa (CRUCIANI *et al.*, 2004; SEMINO *et al.*, 2004) com um pico de frequência centrado no sul e sudeste da Europa (~13% no sul da Itália e 17% – 27% nos Balcãs), e atinge a média de 4%-10% em poloneses, russos, croatas, húngaros, herzegovinianos e bósnios. A sub-linhagem E1b1b-M78 foi observada nas amostras da Terra do Fogo, do Paraguai e do Equador em frequências de até 5%, semelhante aos valores relatados para outras populações sul-americanas (CORACH *et al.*, 2010; NOGUERA *et al.*, 2013; VULLO *et al.*, 2014; RESQUE *et al.*, 2016; MORALES *et al.*, 2018).

Por outro lado, o haplogrupo E1b1b-M81 foi observado nas nossas amostras com valores aproximados de frequência entre de 1,5%-4%. No noroeste da África, aproximadamente 75% das linhagens do cromossomo Y pertencem ao haplogrupo E1b1b-M81 (BOSCH *et al.*, 2001) e sua frequência declina no continente em direção ao leste (SANCHEZ *et al.*, 2003; ARREDI *et al.*, 2004; CRUCIANI *et al.*, 2004; LUIS *et al.*, 2004). Nas populações ibéricas, este haplogrupo foi observado em frequências variáveis (1-12%) e está quase ausente, ou pouco frequente em outras partes da Europa Ocidental (BOSCH *et al.*, 2001; CRUCIANI *et al.*, 2004; SEMINO *et al.*, 2004). Na América do Sul, foram observadas frequências de até 6% para a mutação M81, próximo ao observado nas amostras deste estudo (CORACH *et al.*, 2010; NOGUERA *et al.*, 2013; VULLO *et al.*, 2014; RESQUE *et al.*, 2016; MORALES *et al.*, 2018).

Independentemente das mutações M78 e M81 terem sua origem no leste e norte da África, respectivamente, a presença destas linhagens na América do Sul, muito provavelmente, é o resultado da colonização europeia, uma vez que, as frequências encontradas na Península Ibérica assemelham-se áquelas das populações sul-americanas (CRUCIANI et al., 2004; SEMINO et al., 2004; ALONSO et al., 2005; CRUCIANI et al., 2007; ADAMS et al., 2008; CAPELLI et al., 2009; AMBROSIO et al., 2010; TOSCANINI et al., 2011; NOGUERA et al., 2014; VULLO et al., 2014).

A linhagem E1b1b-M123 foi encontrada em frequências em torno de 3% nas populações estudadas, exceto no Paraguai em que se observa menos de 1%. Esta mutação encontra-se espalhada pelo Oriente Próximo e também é observada no Norte da África, Península Arábica, no sul da Europa e em baixas frequências na Europa Ccidental (CRUCIANI *et al.*, 2004; SEMINO *et al.*, 2004; ABU-AMERO *et al.*, 2009). Na Peninsula Ibérica, frequências em torno de 1-3% foram observadas, bem como em outras populações da América do Sul, e, portanto, assemelham-se aos valores observados neste trabalho (BELEZA *et al.*, 2006; CORACH *et al.*, 2010; MARTINIANO *et al.*, 2013; MORALES *et al.*, 2018; NOGUERA *et al.*, 2013).

O G-M201 é um haplogrupo europeu pouco frequente. Está presente em toda a Europa, Ásia e Mediterrâneo, com frequências de apenas 2% a 4% (DI GIACOMO *et al.*, 2003; CINNIOGLU *et al.*, 2004; FLORES *et al.*, 2004; ALONSO *et al.*, 2005; GONÇALVES *et al.*, 2005; BATTAGLIA *et al.*, 2009; LÓPEZ-PARRA *et al.*, 2009; SIMS *et al.*, 2009; BOATTINI *et al.*, 2013; VOSKARIDES *et al.*, 2016; GRUGNI *et al.*, 2018; SAIZ *et al.*, 2019). A linhagem G-M201 também é observada na região do Cáucaso, que abrange o leste europeu e oeste asiático (REGUEIRO *et al.*, 2006; BALANOVSKY *et al.*, 2011;

YUNUSBAYEV *et al.*, 2011). A mutação M201 foi encontrada em baixas frequências na Colômbia, Brasil, Bolívia e Argentina (SIMS *et al.*, 2009; NOGUERA *et al.*, 2013; SEVINI *et al.*, 2013; VULLO *et al.*, 2014; CÁRDENAS *et al.*, 2015; PEREZ-BENEDICO *et al.*, 2016; RESQUE *et al.*, 2016; MORALES *et al.*, 2018) que são superadas por aquelas encontradas no Equador (3,4%), Paraguai (5,2%) e Terra do Fogo (9%).

O haplogrupo I-M170 tem uma notável especificidade continental, pois é generalizado na Europa e está praticamente ausente em outros lugares (FRANCALACCI e SANNA, 2000; HAMMER *et al.*, 2001; JOBLING e TYLER-SMITH, 2003; ROOTSI *et al.*, 2004; ALONSO *et al.*, 2005; PERICIĆ *et al.*, 2005a; GONÇALVES *et al.*, 2005; CSÁNYI *et al.*, 2008; LOPEZ-PARRA *et al.*, 2009; MARTINIANO *et al.*, 2013; GRUGNI *et al.*, 2018; HERNANDEZ *et al.*, 2019). Nas populações sul-americanas a mutação M170 foi encontrada em média de 3-9% na Argentina, na Bolívia, na Colômbia e no Brasil (CORACH *et al.*, 2010; NOGUERA *et al.*, 2013; VULLO *et al.*, 2014; RESQUE *et al.*, 2016), valores que se assemelham aqueles da população paraguaia e fueguina deste trabalho.

O subtipo I2a-M26 é encontrado em baixas frequências na Europa Ocidental (>5%), sua subclade I1b2 (M26) (SEMINO *et al.*, 2000; BOSCH *et al.*, 2001; CAPELLI *et al.*, 2003; FRANCALACCI *et al.*, 2013; HERNANDEZ *et al.*, 2019), exceto na Sardenha (41%) e em Castilla na Espanha (FLORES *et al.*, 2004; SEMINO *et al.*, 2004). Nas populações estudadas, o sub-haplogrupo I2a-M26 foi observado com frequências de ~2% No Equador e na Terra do Fogo.

As linhagens do Haplogrupo J-12f2 são encontradas em altas frequências no Oriente Médio, Norte da África, Europa, Ásia Central, Paquistão e Índia (HAMMER *et al.*, 2000 e 2001; UNDERHILL *et al.*, 2001; SEMINO *et al.*, 2002; BEHAR *et al.*, 2004; CINNIOGLU *et al.*, 2004; DIGIACOMO *et al.*, 2004; SEMINO *et al.*, 2004; BELEZA *et al.*, 2006; SENGUPTA *et al.*, 2006; ADAMS *et al.*, 2008; MORELLI *et al.*, 2010; CHENNAKRISHNAIAH *et al.*, 2013). Nas Américas, baixas frequências do clado J foram encontradas na Argentina, Bolívia e Colômbia (CORACH *et al.*, 2010; SEVINI *et al.*, 2013; NOGUERA *et al.*, 2014; VULLO *et al.*, 2014; CÁRDENAS *et al.*, 2015; MORALES *et al.*, 2018). Entretanto, as frequências relatadas na Argentina aproximam-se daquelas observadas na Terra do Fogo e no Equador (CORACH *et al.*, 2010). Por outro lado, no Paraguai foram observadas apenas linhagens J-12f2 derivadas.

O sub-haplogrupo J1-M267*, observado apenas no Paraguai, caracteriza populações africanas e árabes, com sua maior frequência no Oriente Médio, no norte da África e Etiópia, e menor na Europa e no sudoeste asiático (ARREDI, *et al.*, 2004; SEMINO *et al.*, 2004;

TOFANELLI *et al.*, 2009; FADHLAOUI-ZID *et al.*, 2011 e 2014; ENNAFAA *et al.*, 2011). A presença desta linhagem em Portugal (que atinge 7% no Sul) foi associada à colonização judaica na Península Ibérica (GONÇALVES *et al.*, 2005). Na América do Sul, este haplogrupo havia sido anteriormente relatado na Colômbia (MORALES *et al.*, 2018).

O clado J1-P58, derivado da linhagem J1-M267, apresenta a frequência máxima na Península Arábica (CHIARONI *et al.*, 2010; TRIKI-FENDRI *et al.*, 2016) e a sua distribuição foi associada à expansão démica de agricultores neolíticos do crescente fértil na Península Arábica (CHIARONI *et al.*, 2010). A mutação foi observada na Espanha e em Portugal (REGUEIRO *et al.*, 2015). Por outro lado, uma frequência de 1,5% foi observada no Paraguai.

A linhagem J2-M172, por sua vez, é mais comum no Levante (ZALLOUA et al., 2008; ABU-AMERO et al., 2009), com frequências que diminuem para o oeste nos países do norte africano e para o leste na Península Arábica (CINNIOGLU et al., 2004; SEMINO et al., 2004; ZALLOUA et al., 2008; EL-SIBAI et al., 2009), embora também seja encontrada no litoral do Mediterrâneo e Cáucaso (DIGIACOMO et al., 2004; SEMINO et al., 2004; BALANOVSKY et al., 2011). Este haplogrupo representa 17% do pool masculino português, uma frequência menor do que a observada na Itália ou na Grécia (SEMINO et al., 2000; DIGIACOMO et al., 2003; BELEZA et al., 2006; GRUGNI et al., 2018), mas superior à do nordeste da Europa (SEMINO et al., 2000; WELLS et al., 2001). Nas populações sulamericanas, a mutação M172 foi observada na Argentina e na Bolívia (CORACH et al., 2010; VULLO et al., 2014), bem como em frequências menores que 10% nas amostras do Equador, Paraguai e Terra do Fogo.

O paragrupo K-M9*(xTat, 92R7, M70) muito provavelmente tem sua origem na Ásia, embora sua presença nas populações sul-americanas seja considerada de ancestralidade europeia. Atualmente, a mutação é amplamente distribuída do Sudeste Asiático à Oceania, embora suas sub-linhagens sejam encontradas em todos os continentes (KARAFET *et al.*, 2010; van OVEN *et al.*, 2011). No Equador e na Terra do Fogo, o haplogrupo K-M9*(xTat, 92R7, M70) teve frequências menores que 1%, e, portanto, menos do que o observado em outras populações sul-americanas (2%-4%) (ACOSTA *et al.*, 2009; CORACH *et al.*, 2010; NOGUERA *et al.*, 2013; VULLO *et al.*, 2014; RESQUE *et al.*, 2016; FLORES-ALVARADO *et al.*, 2022; JOERIN-LUKE *et al.*, 2022; REYES-MADRID *et al.*, 2022).

O haplogrupo L1-M22, por sua vez, foi observado apenas no Paraguai. A maioria dos cromossomos Y pertencentes à linhagem L são encontrados no subcontinente indiano e no Paquistão (WELLS *et al.*, 2001; LACAU *et al.*, 2012), embora estejam presentes em

baixas frequências no Oriente Médio, na Ásia Central, no norte da África e na Europa, ao longo da costa do Mediterrâneo (UNDERHILL *et al.*, 2001; CRUCIANI *et al.*, 2002; JOBLING e TYLER-SMITH, 2003; BEHAR *et al.*, 2004; CINNIOGLU, 2004; KARAFET *et al.*, 2005; SENGUPTA *et al.*, 2006). Embora de origem asiática, a presença dessa linhagem em baixa frequência nas populações sul-americanas pode ser explicada como proveniente da Europa. Na Península Ibérica, a linhagem L1-M22 ocorreu em 0,4-1,2% (Hernandez *et al.*, 2019). Em populações costeiras do Mediterrâneo europeu foram relatadas frequências próximas a 1% (DI GAETANO *et al.*, 2009; BATTAGLIA *et al.*, 2009).

O haplogrupo T1a-M70 provavelmente teve origem no Oriente Próximo e de lá se espalhou pelo Mediterrâneo e pela África Oriental (NOGUEIRO *et al.* 2010; MENDEZ *et al.*, 2011). Este haplogrupo é notavelmente raro, exibindo uma frequência global de cerca de 1% (KING *et al.*, 2007), estando presente na Europa e, particularmente, em Portugal e Itália (ALONSO *et al.*, 2005; BELEZA *et al.*, 2006; NOGUEIRO *et al.*, 2010; MARTINIANO *et al.*, 2013; GRUGNI *et al.*, 2018). Em estudos com populações sul-americanas, foram encontradas frequências de aproximadamente 6% da mutação M70 na Bolívia e na Colômbia (NOGUERA *et al.*, 2013; VULLO *et al.*, 2014; MORALES *et al.*, 2018). No Equador, no Paraguai e na Terra do Fogo, foram encontradas frequências menores, em torno de 2% do haplogrupo T1a-M70.

5.2.2.2 Haplogrupos de origem nativo-americana

O haplogrupo Q-M3 é classificado como nativo-americano. Este clado atinge altas frequências (70-100%) em nativos da América do Sul (BORTOLINI *et al.*, 2003; SEIELSTAD *et al.*, 2003; BISSO-MACHADO *et al.*, 2011; JOTA *et al.*, 2011; MALYARCHUK *et al.*, 2011; TOSCANINI *et al.*, 2011, 2016; BATTAGLIA *et al.*, 2013; LARDONE *et al.*, 2013; NOGUERA *et al.*, 2013; REGUEIRO *et al.*, 2013; RASMUSSEN *et al.*, 2014; SALA e CORACH, 2014; TORRES *et al.*, 2015; MORALES *et al.*, 2018), mas é encontrado em frequências menores em populações miscigenadas sul-americanas. A linhagem Q1b-M3 mais difundida parece ser mais diferenciada na região andina, onde novas sub-linhagens foram observadas (BATAGLIA *et al.*, 2013). Nas populações miscigenadas da Terra do Fogo e Paraguai, a mutação M3 aparece com frequência de 5%, concordante com o observado em outras populações da América do Sul, à exceção da Bolívia (CORACH *et al.*,

2010; NOGUERA et al., 2013; VULLO et al., 2014; RESQUE et al., 2016; MORALES et al., 2018).

A mutação derivada Z19319 do haplogrupo Q1b-M3, foi anteriormente encontrada em comunidades indígenas andinas do Peru e da Bolívia (JOTA *et al.*, 2016). Segunto estes autores, a mutação Z19319 provavelmente se originou no Holoceno (~ 9.000 YBP) na região norte dos Andes Centrais (Cajamarca, Peru). Esta mutação tem sido relacionada a períodos pré-cerâmicos I e II, quando o cultivo da mandioca, da abóbora e da batata doce estava sendo inicializados (JOTA *et al.*, 2016). Neste trabalho, a referida mutação foi detectada no Equador, na Terra do Fogo e no Paraguai, com frequências de no máximo 2%.

No Paraguai, também foi observado o paragrupo Q1b-M346. Este paragrupo ancestral é amplamente distribuído na Eurásia do Norte, e também foi identificado em populações da América do Sul (BORTOLINI *et al.*, 2003; KARAFET *et al.*, 2008; BAILLIET *et al.*, 2009; JOTA *et al.*, 2011; BISSO-MACHADO, BORTOLINI e SALZANO 2012; VULLO *et al.*, 2014; PEREZ-ENEDICO *et al.*, 2016; MORALES *et al.*, 2018) e detectada neste estudo com a frequência de aproximadamente 1% no Equador.

5.2.2.3 Haplogrupos de origem asiática

Embora o componente nativo-americano presente na América do Sul pertença majoritariamente ao haplogrupo Q, foram também descritas linhagens mais raras dentro do haplogrupo C (SCHURR e SHERRY, 2004; ZEGURA et al. 2004; GEPPERT et al., 2015). Dentro do haplogrupo C-M130, foi detectada uma única amostra na população paraguaia. O software de predição NEVGEN indicou como provável haplogrupo desta amostra a linhagem C1a2-V20 para o perfil haplótipico fornecido. Os nativos da América do Sul que pertencem ao haplogrupo C, possuem a mutação M217, que define a sub-linhagem C3. Esta linhagem também é observada na região central e nordeste da Ásia (ZEGURA et al., 2004; ZHONG et al., 2010; ROEWER et al., 2013; GEPPERT et al., 2015; MEZZAVILLA et al., 2015; PINOTTI et al., 2019). Uma vez que apenas haplótipos pertencentes a linhagem C3 são classificados como nativo-americanos, a linhagem C-M130 permaneceu como asiática.

A presença das mutações estritamente asiáticas P-164 e M117, pertencentes ao haplogrupo O, no Paraguai é uma raridade, anteriormente nunca relatada em populações sulamericanas. O haplogrupo O, caracterizado por uma deleção de 5 pb conhecida como M175, é

o haplogrupo dominante entre os homens em todo Leste e Sudeste Asiático (JIN *et al.*, 2003; KAYSER *et al.*, 2003; KARAFET *et al.*, 2005; HAMMER *et al.*, 2006; KARAFET *et al.*, 2010; ZHONG *et al.*, 2010; CHEN *et al.*, 2011; WANG *et al.*, 2015b; WEN *et al.*, 2019). Remotamente, também é observado na Oceania (KAYSER *et al.*, 2010), e está presente em Madagascar (HURLES *et al.*, 2005).

O haplogrupo O3a-P164*(xM134) é bastante frequente em Taiwan, sul do Vietnã, Tailândia, Filipinas, Indonésia e Japão (TREJAUT *et al.*, 2014). Sua distribuição ao longo do sudeste asiático provavelmente ocorreu durante o Neolítico (SU *et al.*, 2000; TREJAUT *et al.*, 2014). Adicionalmente, o haplogrupo O3a-M117 é encontrado na maioria das populações do leste e sudeste da Ásia, bem como populações falantes de Tibeto-Burman no sudoeste da China e da Índia (SHI *et al.*, 2005; XUE *et al.*, 2006; GAYDEN *et al.*, 2007; REDDY *et al.*, 2007; CAI *et al.*, 2011; YAN *et al.*, 2011; WANG *et al.*, 2013). Portanto, a mutação M117 tem sido associada à expansão das populações sino-tibetanas (CAI *et al.*, 2011).

5.2.2.4 Haplogrupos encontrados de origem africana

O sub-clado E1a-M33, que pertence a umas das linhagens mais frequentes da África, o haplogrupo E, tem uma distribuição bastante limitada no oeste subsaariano (CRUCIANI *et al.*, 2002). No Paraguai e no Equador, foram observadas frequências relativamente semelhantes àquelas nas populações brasileiras e colombianas (NOGUERA *et al.*, 2013; RESQUE *et al.*, 2016; MORALES *et al.*, 2018).

Os valores encontrados no Paraguai da linhagem E1b1a-M2 aproximam-se daqueles observados na Argentina (CORACH e al., 2010). Esta linhagem é considerada um marcador da expansão Bantu (UNDERHILL *et al.*, 2001; BELEZA *et al.*, 2005; COELHO *et al.*, 2009), e é o haplogrupo mais comum na África Subsaariana, com picos de frequência no Oeste (cerca de 80%) e África Central (cerca de 60%) (CRUCIANI *et al.*, 2002; TROMBETTA *et al.*, 2011). Por outro lado, no Norte da África a mutação M2 é observada em frequências menores (geralmente abaixo de 10%) (CRUCIANI *et al.*, 2002; SEMINO *et al.*, 2004; WOOD *et al.*, 2005).

Dentro da E1b1a-M2, os sub-haplogrupos E1b1a-U174 e E1b1a-U209 são predominantes em grupos de língua Bantu na África central e meridional (SIMS *et al.* 2007; di FILLIPO *et al.*, 2011; MONTANO *et al.*, 2011; SANCHEZ-FADDEEV *et al.*, 2013;

LARMUSEAU *et al.*, 2015). Enquanto na Nigéria e no Gabão a mutação U174 prevalece, nos Camarões a mutação U209 predomina (MONTANO *et al.*, 2011). No Equador, as mutações U174 e U209 foram observadas em baixas frequências, assim como foi relatado anteriormente no Brasil (JOERIN-LUKE *et al.*, 2022).

O sub-haplogrupo E1b1a-U290, observado apenas na Terra do Fogo nas amostras estudadas, é encontrado em grandes proporções na África Ocidental e Central (MONTANO *et al.*, 2011; ANSARI-POUR *et al.*, 2013; LARMUSEAU *et al.*, 2015) e em frequências mais baixas na África Austral (ANSARI-POUR *et al.*, 2013). Na América do Sul, esta linhagem é comumente observada em populações afro-americanas (SIMS *et al.*, 2007; ANSARI-POUR *et al.*, 2015).

O sub-haplogrupo E1b1b-M35 é comum no noroeste da África, na África oriental, no Oriente Médio (BOSCH *et al.*, 2001; SEMINO *et al.*, 2002; ARREDI *et al.*, 2004; CRUCIANI *et al.*, 2004; REGUEIRO *et al.*, 2006) e com raras ocorrências na Europa (HAMMER *et al.*, 1998; SEMINO *et al.*, 2000; CRUCIANI *et al.*, 2004). Frequências em torno de 1% foram observadas nas populações miscigenadas sul-americanas (CORACH *et al.*, 2010; NOGUERA *et al.*, 2013; RESQUE *et al.*, 2016; MORALES *et al.*, 2018). Nas populações deste estudo foram encontradas frequências entre 0,5%-2%.

O haplogrupo E2-M85 também foi descrito como um marcador de expansão da língua Bantu na África Central, Sul, Ocidental e Oriental em frequências variando de 2 a 8%. (CRUCIANI *et al.*, 2002; BERNIELL-LEE *et al.*, 2009). Frequências semelhantes ou aproximadas àquela encontrada no Equador para a mutação M85 foram observadas na Colômbia, no sudeste brasileiro e no Espírito Santo (NOGUERA *et al.*, 2013; OLIVEIRA *et al.*, 2014; FIGUEIREDO *et al.*, 2016; JANNUZZI *et al.*, 2020).

O haplogrupo R1b1b-V88 foi observado apenas na Terra do Fogo. Embora a linhagem R1b-M269 seja de origem europeia, a sub-linhagem portadora da mutação V88 foi encontrada concentrada na parte centro-oeste do continente africano, onde podem ser detectadas frequências tão altas quanto 95% (CRUCIANI *et al.*, 2010). Sendo assim, acreditase que o haplogrupo R1b1b-V88 tenha se originado em África, após mutação em R1b-M269, a qual retornou à África por migração desde o Oriente Médio, em tempos pré-históricos, junto com a disseminação dos grupos Chadic, da família linguística afro-asiática (CRUCIANI *et al.*, 2010). Dessa maneira, o haplogrupo R1b1b-V88 foi observado em frequências altas em Sahel central (norte de Camarões, norte da Nigéria, Chade e Níger, Guiné Equatorial) e também foi relatado em frequências baixas no noroeste da África (CRUCIANI *et al.*, 2010; GONZALEZ *et al.*, 2013; HABER *et al.*, 2016; D'ATANASIO *et al.*, 2018). Fora do continente africano,

duas sub-linhagens raras de R1b1b-V88 foram observadas no Oriente Próximo e no sul da Europa (particularmente na Sardenha) (FRANCALACCI *et al.*, 2013; CRUCIANI *et al.*, 2010).

A presença da mutação V88 nas populações sul-americanas é considerada de ancestralidade africana. Na América do Sul, a linhagem V88 foi observada no Equador com a frequência de 2% (VILLAESCUZA *et al.*, 2021) e agora na Terra do Fogo.

5.2.2.5 Previsão dos haplogrupos mais prováveis nas populações da Colômbia, Rio de Janeiro e Maranhão

Com o *software* NEVGEN, foram previstos 27 haplogrupos diferentes no Maranhão, Rio de Janeiro e Colômbia, baseado nos resultados dos haplótipos genotipados com o kit PowerPlex® Y23 (Tabela 16). Para a população do Maranhão, 89 amostras foram anteriormente analisadas com marcadores Y-SNP, e os resultados estão publicados no artigo "Male lineages in Brazilian populations and performance of haplogroup prediction tools" (Anexo I).

Em termos de frequência, o haplogrupo mais encontrado em todas as populações foi o R1b1a, seguido do E1b1a-V38 (Tabela 16). Os restantes haplogrupos apresentaram frequências médias inferiores a 5% entre as três populações, com frequências variáveis entre elas. Entretanto, não foi possível predizer o haplogrupo para algumas amostras uma vez que a probabilidade menor que 10% não foi considerada (5,5% na Colômbia, 4,93% no Rio de Janeiro e 7,42% no Maranhão).

Devido à taxa de erro associada à determinação de haplogrupos com base em perfis haplotípicos, os resultados apresentados para os haplogrupos preditos nas amostras da Colômbia, Brasil e Maranhão correspondem apenas a uma aproximação da frequência real nas amostras estudadas. Para uma melhor definição dos haplogrupos e uma determinação mais precisa da sua origem geográfica, estes devem ser confirmados pela genotipagem dos Y-SNPs (JOBLING e TYLER-SMITH, 2003; KARAFET *et al.*, 2008; van OVEN *et al.*, 2014).

Tabela 16. Frequência de haplogrupos atribuídos pelo software NEVGEN

Haplogrupo	Colômbia	Rio de Janeiro	Maranhão	Ancestralidade
A1a-M31	0,30%	0,00%	0,44%	AFR
В	0,00%	0,00%	0,44%	AFR
B2a1-M218	0,30%	0,99%	0,44%	AFR
E1 a-M132	0,00%	0,00%	0,44%	AFR
E1 b1 a-V38	9,30%	14,29%	8,73%	AFR
E1 b1 b-M123	1,95%	0,99%	1,31%	EUR
E1 b1 b-M78	2,55%	1,48%	2,62%	AFR/EUR
E1 b1 b-M81	1,95%	5,91%	2,62%	EUR
E2-M75	0,30%	0,99%	0,00%	EUR
E1b1b-V1515	0,00%	0,49%	0,44%	AFR
E1 b1 b-V22	0,45%	0,49%	0,00%	AFR
E1 b1 b-L67	0,60%	0,00%	1,31%	AFR
E1 b1 b-V68	0,00%	0,00%	0,44%	EUR
G	2,40%	1,48%	3,06%	EUR
H2	0,15%	0,00%	0,00%	EUR
I1	1,95%	2,46%	3,49%	EUR
12	6,30%	3,45%	3,06%	EUR
J1	3,30%	2,96%	2,18%	EUR
J2	4,65%	5,42%	5,24%	EUR
L	0,00%	0,49%	0,00%	EUR
N	0,00%	0,00%	0,44%	EUR
Q	8,25%	0,49%	2,62%	NAM
R1a	2,40%	1,48%	1,31%	EUR
R1b	44,08%	50,74%	50,22%	EUR
R1b1b-V88	0,45%	0,00%	0,00%	AFR
T	2,85%	0,49%	1,75%	EUR
Não suportado (<10%)	5,50%	4,93%	7,42%	-

Legenda: Frequência relativa (%) dos haplogrupos preditos com o *software* NEVGEN na Colômbia e no Brasil (Rio de Janeiro e Maranhão), bem como a ancestralidade associada.

5.3 Ancestralidade paterna das populações da América do Sul

5.3.1 Ancestralidade das populações estudadas

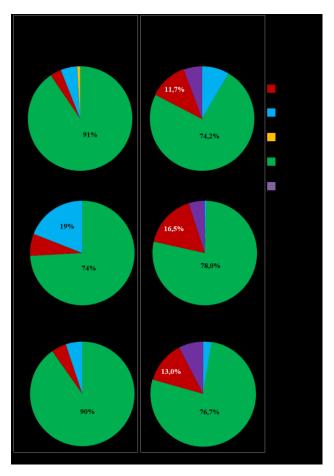
Na Figura 23, é possível se observar as proporções de ancestralidades atribuídas aos haplogrupos detectados no Equador, no Paraguai e na Terra do Fogo, bem como a ancestralidade dos haplogrupos preditos nas amostras do Brasil e da Colômbia.

Visto que os haplogrupos definidos por Y-SNPs são geograficamente específicos, as linhagens paternas observadas nas populações do Equador, Paraguai e da Terra do Fogo foram consideradas, respectivamente, como sendo de origem: 74%, 91% e 90% europeia (haplogrupos E1b1b-M123, E1b1b-M78, E1b1b-M81, G-M201, I-M170*, I2a-M26, J-12f2a*, J1-M267*, J1-P58, J2-M172, K-M9*, L1-M22, R1a-SRY10831.2, R1b-M269*, R1b1a-L23*, R1b1a-M529, R1b1a-S116*, R1b1a-U152, R1b1a-M153, R1b1a-M167, R1b1a-M106, T1a-M70), 7%, 3% e 5% africana (haplogrupos E1a-M33, E1b1a-M2*, E1b1a-U209*, E1b1a-U290, E1b1a-U174, E1b1b-M35*, E1b1b-M78, R1b1b-V88) 19%, 5% e 5% nativo-americana (haplogrupos Q1b-M346*, Q1b-M3* e Q1b-Z19483*) e, apenas no Paraguai, 1% asiática (haplogrupos C-M130, O3a-P164* e O3a-M117).

Por outro lado, de acordo com a predição de haplogrupos com base no haplótipo, para as populações da Colômbia e do Brasil (Rio de Janeiro e Maranhão) foram observadas as ancestralidades, respectivamente: 74,2%, 78% e 76,7% europeia (haplogrupos E1b1b-M123, E1b1b-M81, E2-M75, E1b1b-V1515, E1b1b-V22, E1b1b-V68, E1b1b-L67, G, H2, I1, I2, J1, J2, L, N, R1a, R1b, e T), 11,7%, 16,5% e 13% africana (haplogrupos A1, A1a-M31, B, B2-M218, E1b1b-L67, E1a-M132, E1b1b-M78, E1b1b-V1515, E1b1a-V38, E1b1b-V22, E1b1b-L67 e R1b1b-V88), e 8,4%, 0,5% e 2,7% nativo-americana (haplogrupo Q).

A linhagem E1b1b-M78 foi considerada parte africana e europeia, uma vez que a mutação M78 é comumente observada na Europa. Entretanto, para as amostras analisadas com o *software* NEVGEN, o haplogrupo E1b1b-M78 foi eliminado do cálculo da ancestralidade.

Figura 23. Classificação das amostras miscigenadas sul-americanas, segundo a ancestralidade dos haplogrupos



Legenda: Ancestralidade das amostras miscigenadas sul-americanas deste trabalho. Quando não foi possível predizer o haplogrupo para algumas amostras com o NEVGEN uma vez que a probabilidade menor que 10% foi observada, a ancestralidade foi considerada não suportada.

É possível observar que todas as populações estudadas apresentam maior ancestralidade europeia. Isso deve-se muito provavelmente a predominância de homens migrantes europeus, e por sua vez ao acasalamento preferencial de mulheres nativo-americanas com homens colonizadores europeus (SALAS *et al.* 2008; CORACH *et al.* 2010; ROJAS *et al.* 2010). Entretanto, o Equador é a população que apresenta o valor mais baixo de ancestralidade europeia e, em contrapartida, possui o valor mais alto de ancestralidade nativo-americana. Isso se deve, muito provavelmente, a sua localização próxima a região andina, onde as linhagens nativo-americanas foram mais preservadas.

Alternativamente, no Rio de Janeiro, no Maranhão e na Colômbia foram observadas frequências mais altas de ancestralidade africana do que nativo-americana. Este resultado corrobora trabalhos anteriores que apresentaram valores de frequências do componente

africano semelhantes no Nordeste brasileiro, bem como no Rio de Janeiro e na Colômbia (ACOSTA *et al.*, 2009; NOGUERA *et al.*, 2013; FIGUEIREDO *et al.*, 2016; RESQUE *et al.*, 2016). Isto é resultado do grande contingente de africanos que trabalharam como mão de obra no litoral brasileiro durante a colonização europeia, e de maneira semelhante na região do Pacífico e Caribe, na Colômbia, justificando a preservação das linhagens africanas.

Além disso, quando compara-se a frequência de haplogrupos não suportados pelo software NEVGEN, a Maranhão apresenta pelo mais dificuldade em ter a ancestralidade de seus haplótipos preditos do que o Rio de Janeiro e Colômbia. Isto pode sugerir que esta população possui linhagens paternas, muito provavelmente nativo-americanas, que não foram anteriormente observadas e inseridas na base de dados populacionais.

5.3.2 Panorama da América do Sul

As populações sul-americanas atuais são caracterizadas como miscigenadas, como resultado da troca genética ocorrida principalmente entre os povos de três continentes diferentes, nomeadamente os nativo-americanos, europeus e africanos. Esta mistura genética é o resultado da relação entre homens europeus e mulheres nativas ou africanas. Portanto, é relevante o estudo genético populacional dos homens modernos sul-americanos de maneira a elucidar a possível origem, a contribuição e a distribuição das linhagens masculinas (BORTOLINI *et al.*, 2003; JOTA *et al.*, 2016; ROEWER *et al.*, 2013; TOSCANINI *et al.*, 2011).

Muitos estudos buscam compreender a composição genética de populações nativas e as linhagens do cromossomo Y que permitem desenhar os padrões de migração das populações humanas na formação das populações sul-americanas. Entretanto, dados relativos a populações miscigenadas ainda são escassos. Portanto, é pertinente aprofundar o conhecimento da composição genética das populações miscigenadas sul-americanas atuais com marcadores do cromossomo Y.

Na América do Sul, são observados diferentes padrões de ancestralidade paterna (Figura 24).

O componente europeu predomina em todas as regiões da América do Sul, exceto na Bolívia. A ancestralidade europeia excede 90%, por exemplo, no Paraguai, na Terra do Fogo,

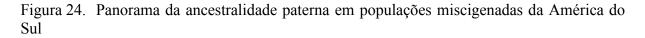
na Argentina e na região sul do Brasil (CORACH et al., 2010; RESQUE et al., 2016; SIMÃO et al., 2021).

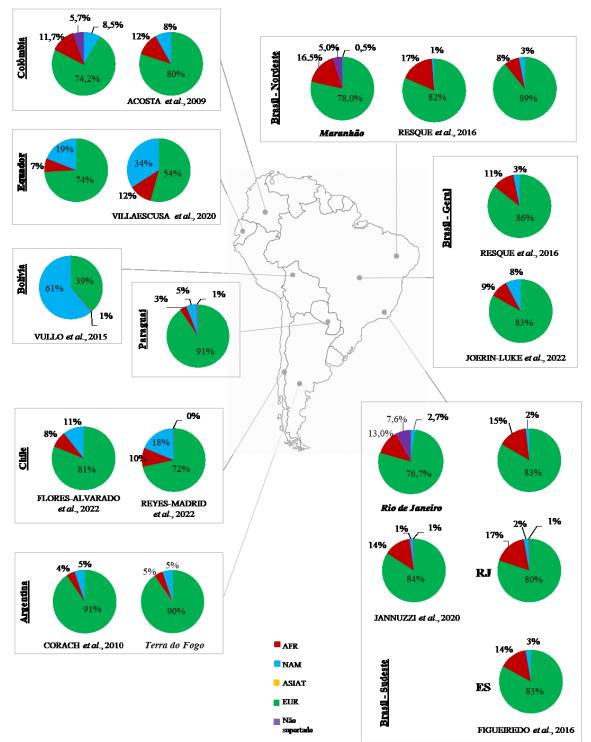
O componente nativo-americano apresenta maiores frequências em populações localizadas na região andina, como por exemplo, na Bolívia (NAM=61%), em que claramente as linhagens nativo-americanas se preservaram mais em comparação com as outras populações sul-americanas (VULLO *et al.*, 2015). Além disso, no Equador (NAM=34%), no Chile (NAM=18%) e na Argentina (NAM=5%), os valores de ancestralidade nativo-americana superam a africana (CORACH *et al.*, 2010; VILLAESCUSA *et al.*, 2021; REYESMADRID *et al.*, 2022). No Brasil o componente ancestral nativo-americano é menor, se comparado às demais populações sul-americanas, exceto para a região Norte (NAM=8%) em que se preservaram linhagens nativo-americanas amazônicas.

Alternativamente, o componente africano apresenta suas maiores frequências em populações sul-americanas do Nordeste (AFR=17%) e sudeste (AFR=15%) do Brasil, no Espírito Santo (AFR=15%), no Chile, no Equador e na Colômbia (ACOSTA *et al.*, 2009; NOGUERA *et al.*, 2013; FIGUEIREDO *et al.*, 2016; RESQUE *et al.*, 2016; VILLAESCUSA *et al.*, 2021; REYES-MADRID *et al.*, 2022) que historicamente receberam um elevado influxo de migrações africanas durante a época colonial. A maior frequência da ancestralidade paterna africana foi observada em Bolívar, na Colômbia (AFR=24%) concordando os dados históricos acerca da migração de populações africanas para a o Caribe colombiano.

Em contraste com os resultados apresentados, análises do mtDNA em populações sul-americanas revelam diferenças em termos de ancestralidade das linhagens maternas e paternas, refletindo em um viés sexual.

No Chile (GÓMEZ-CARBALLA *et al.*, 2016; FLORES-ALVARADO *et al.*, 2021), na Colômbia (1000 GENOMES PROJECT), e no Paraguai (SIMÃO *et al.*, 2018; SIMÃO *et al.*, 2021), por exemplo, os valores de linhagens maternas nativo americanas superam 80%. Por outro lado, na Bolívia (TABOADA-ECHALAR *et al.*, 2013), foi observada uma frequência de mais dsdae 98% de linhagens nativo americanas. O componente nativo-americano diminui para 61% na Argentina, bem como o componente europeu é observado em 37% (BOBILLO *et al.*, 2010). No Brasil, o componente europeu supera em 6% a frequência do componente nativo-americano para mtDNA (33% nativo-americano, 39% europeu e 28% africano) (ALVES-SILVA *et al.*, 2000).





Legenda: Populações miscigenadas da América do Sul disponíveis na literatura, bem como as populações deste estudo indicadas em itálico.

Cada uma das populações estudadas no presente estudo possui uma história distinta dos movimentos migratórios atuais para trás, passando pela conquista europeia e a ocupação do território por nativos americanos.

A Terra do Fogo primeiramente ocupada pelos nativos Yagán, Selknam, Manekénk e marginalmente os Kawésqar foi ocupada por conquistadores europeus do século XIV, especialmente espanhóis. Além disso, a população recebeu uma segunda onda migratória europeia a partir de 1884 com a venda de lotes de terra privados.

O Rio de Janeiro e o Maranhão que *a priori* eram ocupados por populações nativas pertencentes ao grupo linguístico Tupi e Tapuias (não falantes Tupi) recebeu um grande número de migrantes europeus e escravos africanos. O porto localizado no Rio de Janeiro foi um dos responsáveis pela chegada de migrantes portugueses e de africanos e pelo transporte das riquezas produzidas na colônia, tais como a cana de açúcar, as minas de ouro e diamantes e o café. Por outro lado, no Maranhão o primeiro domínio colonial europeu foi francês, além do contingente holandês que tentou conquistar esta região sem sucesso.

A população colombiana atual é resultado da miscigenação entre os diversos grupos nativos, tais como os Chibchan residentes na costa do Atlântico, os Chocoan na costa do Pacífico e os Quechua residentes na região Andina, os colonizadores espanhóis e os africanos. Relativo a este último grupo, o maior contingente migratório se estabeleceu na região do Pacífico e Caribe. Em Cartagena, por exemplo, que foi o centro do tráfico de escravos durante o período colonial, estabeleceu-se na zona rural uma importante comunidade negra nomeada Palenque de San Basílio.

Após eventos pre e pós colombianos, a população do Equador é caracterizada pelos três grupos étnicos mestiços, indígenas e afro-equatorianos. Comparado às outras populações sul-americanas, os equatorianos miscigenados possuem o maior contingente nativo-americano, principalmente nas regiões amazônicas.

Por fim, o Paraguai, que, embora tenha a princípio sido alvo do colonialismo europeu assim como outras populações sul-americanas, se manteve particularmente isolado dos seus vizinhos, uma vez que não haviam muitas riquezas minerais. A segunda onda migratória europeia para o Paraguai ocorreu após a Guerra do Paraguai (1864-1870) com o objetivo de restabelecer a densidade populacional do país, já que, durante os anos de conflito mais da metade do contingente masculino foi eliminado.

Embora a incursão europeia do período colonial e de migrações recentes, bem como a perda por guerras doenças dos nativos-americanos, tenha influenciado na predominancia de linhagens europeias nestas populações, aquelas próximas das regiões andinas possuem algum

grau de preservação das linhagens nativo-americanas. Portanto, os resultados encontrados para a ancestralidade das linhagens paternas, bem como a análise de distâncias genéticas preliminares condizem com os dados históricos acerca da formação destas populações, como resultado da mistura entre africanos, nativo-americanos e em maior parte europeus.

Em conclusão, foram fornecidos dados relevantes acerca de marcadores Y-SNPs para as populações miscigenadas brasileira, colombiana, equatoriana, paraguaia e fueguina, de maneira a ampliar a caracterização das populações sul-americanas (BELEZA *et al.*, 2006; ROMERO *et al.*, 2008; SÁNCHEZ-DIZ *et al.*, 2008; OLIVEIRA *et al.*, 2014; PURPS *et al.*, 2014; TOSCANINI *et al.*, 2018) uma vez que muitas delas, carecem de estudos que descrevam sua composição genética, e assim prover uma base de dados destas populações afim do seu uso em genética forense e populacional.

5.4 Determinação de taxas de mutação em Y-STRs em pares pai-filhos

A estimativa adequada dos parâmetros de mutação é essencial em estudos evolutivos e filogenéticos. Entretanto, como a mutação é um evento raro, é necessário a procura de sua ocorrência e coleta dos dados dentre um número de indivíduos muito elevado. Dessa maneira, os duos pai-filho do Rio de Janeiro deste trabalho foram integrados num estudo colaborativo sobre mutações de repetições curtas em tandem do cromossomo Y (Y-STRs), organizado pelo Grupo de Trabalho de Língua Espanhola e Portuguesa da Sociedade Internacional de Genética Forense (GHEP-ISFG). O estudo colaborativo teve como objetivo reunir estrategicamente vários laboratórios, de maneira a recrutar uma maior quantidade de dados genéticos. Portanto, dados de 2.225 duos pai-filho, de 10 populações em todo o mundo: Argentina (N= 414), Brasil (N= 201), Colômbia (N= 222), Dinamarca (N= 96), Equador (N= 102), Groenlândia (N= 104), Paguistão (N= 110), Portugal (N= 509), Espanha (N= 250) e Emirados Árabes Unidos (N=216), e dados de 44 trabalhos publicados anteriormente, correspondendo a 25.729 duos, foram coletados e analisados. Os resultados referentes ao estudo colaborativo das taxas de mutação estão organizados no artigo "Microsatellites' Mutation Modeling Through the Analysis of the Y-Chromosome Transmission: results of a GHEP-ISFG Collaborative Study" (anexo J), publicado na revista Forensic Science International: Genetics.

As taxas de mutação foram estimadas para 33 Y-STRs, dentre eles 27 Y-STRs incluídos em um (ou ambos) kits PowerPlex® Y23 ou Yfiler® Plus, bem como para 8 outros

Y-STRs não incluídos nos kits mencionados anteriormente (DYS388, DYS435, DYS461, DYS526a/b, DYS547, DYS612, DYS626 e Y-GATA-A10).

Um total de 467.073 transferências alélicas foram então analisadas para os 33 Y-STRs, e 1.863 mutações foram observadas. As taxas médias de mutação variaram entre 0,0005 (para DYS438 e DYS643) e 0,0170 (para DYS547) (Tabela 17), o que fortalece a recomendação sobre o uso de estimativas específicas por marcador.

5.4.1 <u>Modelo mutacional passo a passo (Single Step mutation model)</u>

Segundo Kimura e Ohta (1978), o comprimento dos STRs muda por uma ou mais unidades repetitivas a cada evento mutacional - SMM (*Stepwise Mutation Model*). O modelo de um único passo ocorre quando há a perda ou ganho de apenas uma repetição a cada evento mutacional. A mudança no comprimento do alelo que envolve mais de uma unidade repetitiva denominada de múltiplo passos são muito mais raras.

Das 1.863 mutações observadas, 1.786 eram compatíveis com mutações de um passo e 77 com mutações de múltiplos (Tabela 17), sendo a primeira 23,2 vezes mais frequente que as demais. Entretanto, esta razão pode ser altamente variável entre os marcadores, variando entre 1,25 para DYS438 e 98 para DYS449. Estudos anteriores já revelaram que independentemente do modo de transmissão genética considerado, as mutações de etapa única, com ganho ou perda de uma única repetição, são mais frequentes que aquelas de várias etapas (WEBER e WONG, 1993; XU *et al.*, 2000).

Tabela 17. Estimativa das taxas de mutação para 33 Y-STRs e intervalos de confiança correspondentes (continua)

Marcadores	Obs	ervações (N)	Taxa de	CI (95%)	Núm	Um passo/						
	Muts	Transferência alélica	mutação		1	2	3	4	6	>1*	NC	Multipasso
DYS19	57	26.372	0,0022	0,00164-0,00280	56	0	0	0	0	0	1	-
DYS385a/b	112	46.659	0,0024	0,00198-0,00289	104	7	1	0	0	0	0	13
DYS388	2	3.612	0,0006	0,00007-0,00200	2	0	0	0	0	0	0	-
DYS389I	69	26.154	0,0026	0,00205-0,00334	69	0	0	0	0	0	0	-
DYS389II-I	91	26.113	0,0035	0,00281-0,00428	88	2	1	0	0	0	0	29,33
DYS390	63	25.103	0,0025	0,00193-0,00321	62	1	0	0	0	0	0	62
DYS391	66	25.789	0,0026	0,00198-0,00325	64	2	0	0	0	0	0	32
DYS392	14	24.266	0,0006	0,00032-0,00097	13	0	0	1	0	0	0	13
DYS393	31	24.332	0,0013	0,00087-0,00181	31	0	0	0	0	0	0	-
DYS437	22	21.018	0,001	0,00066-0,00158	22	0	0	0	0	0	0	-
DYS438	10	21.033	0,0005	0,00023-0,00087	5	3	0	1	0	0	1	1,25
DYS439	110	21.111	0,0052	0,00428-0,00628	110	0	0	0	0	0	0	-
DYS448	20	16.577	0,0012	0,0074-0,00186	19	1	0	0	0	0	0	19
DYS449	99	8.480	0,0117	0,00950-0,01420	98	0	1	0	0	0	0	98
DYS456	81	17.892	0,0045	0,00360-0,00562	80	0	1	0	0	0	0	80
DYS458	126	17.920	0,007	0,00586-0,00837	122	3	1	0	0	0	0	30,5
DYS460	23	5.605	0,0041	0,00260-0,00615	23	0	0	0	0	0	0	-
DYS481	25	5.504	0,0045	0,00294-0,00670	22	3	0	0	0	0	0	7,33
DYS518	116	7.795	0,0149	0,01231-0,01782	105	5	3	2	0	0	1	10,5
DYS526a	15	4.425	0,0034	0,00190-0,00559	11	4	0	0	0	0	0	2,75
DYS526b-a	50	4.401	0,0114	0,00843-0,01498	49	1	0	0	0	0	0	49
DYS533	9	6.975	0,0013	0,00059-0,00245	8	1	0	0	0	0	0	8
DYS547	69	4.053	0,017	0,01327-0,02150	67	0	0	1	0	1	0	33,5
DYS549	10	2.617	0,0038	0,00183-0,00702	10	0	0	0	0	0	0	-
DYS570	87	9.976	0,0087	0,00699-0,01075	80	6	1	0	0	0	0	11,43
DYS576	140	9.876	0,0142	0,01194-0,01671	135	3	2	0	0	0	0	27
DYS612	66	4.056	0,0163	0,01261-0,02066	60	6	0	0	0	0	0	10
DYS626	41	4.441	0,0092	0,00663-0,01250	39	1	0	0	1	0	0	19,5

Tabela 17.	Estimativa das	taxas de	mutação	para	33	Y-STRs e	e intervalos	de	confiança
correspondente	es (conclusão)								

Marcadores	Obs	ervações (N)	Taxa de	CI (95%)	Núm	Um passo/						
wai cadoi es	Muts	Transferência alélica	mutação	1	1	2	3	4	6	>1*	NC	Multipasso
DYS627	133	8.028	0,0166	0,01389-0,01960	127	5	1	0	0	0	0	21,17
DYS635	59	18.813	0,0031	0,00239-0,00404	59	0	0	0	0	0	0	-
DYS643	1	1.978	0,0005	0,00001-0,00281	1	0	0	0	0	0	0	-
Y-GATA-A10	4	1.026	0,0039	0,00106-0,00995	4	0	0	0	0	0	0	-
Y-GATA-H4	42	17.611	0,0024	0,00172-0,00322	41	1	0	0	0	0	0	41
Total	1.863	469.611	-	-	1.786	55	12	5	1	1	3	23,2

Legenda: Número de mutações compatíveis com o ganho ou perda de 1 a 6 repetições e as proporções correspondentes entre o número de mutações de etapa única e múltipla. Os dados incluem tanto os gerados no presente estudo quanto os obtidos na literatura (BLANCHI et al., 1998; LESSIG e EDELMANN, 1998; KAYSER et al., 2000b; DUPUY et al., 2001; TSAI et al., 2002; DUPUY et al., 2004; KURIHARA et al., 2004; BALLARD et al., 2005; BERGER et al., 2005; BUDOWLE et al., 2005; GUSMÃO et al., 2005; TURRINA, ATZEI e de LEO, 2006; DOMINGUES, et al., 2007; HOHOFF et al., 2007; KUMAGAI et al., 2007; LEE et al., 2007; PONTES et al., 2007; sen SHI et al., 2007; DECKER et al., 2008; PADILLA-GUTIÉRREZ et al., 2008; SÁNCHEZ-DIZ et al., 2008; SOARES-VIEIRA et al., 2008; GE, et al., 2009; GOEDBLOED, et al., 2009; KIM et al., 2009; VIEIRA-SILVA, et al., 2009; LAOUINA et al., 2013; BALLANTYNE et al., 2014; ROBINO et al., 2015; WANG et al., 2016; ANTÃO-SOUSA et al., 2017; ADNAN et al., 2018; BUGOYE, MULIMA, e MISINZO, 2018; MERTOGLU et al., 2018; YANG et al., 2018; WU et al., 2018; PETROVIC et al., 2019; YUAN et al., 2019; ZHANG et al., 2019; AY et al., 2019; AMBROSIO et al., 2020; FU et al., 2020; LIN et al., 2020; BREDEMEYER ROEWER e WILLUWEIT, 2021). NC: Não compatível com alterações envolvendo um número inteiro de repetições.

5.4.2 <u>Variação da taxa de mutação com a idade do pai</u>

É comumente aceito que alelos de pais mais velhos são mais propensos à mutação alélica durante sua transmissão aos filhos, embora os trabalhos ainda não tivessem fornecido valores para esta correlação. Informações sobre a idade do pai no momento do nascimento do filho foram coletadas para 84.715 transmissões alélicas. Os resultados revelaram que a taxa de mutação e a idade do pai estão positivamente correlacionadas. A classe etária de 51-60 anos foi a que apresentou maior taxa de mutação (Tabela 18). Portanto, os dados analisados neste trabalho suportam uma correlação positiva entre a idade do pai e a ocorrência da mutação, e estão de acordo com relatos anteriores (BALLANTYNE *et al.*, 2010; SUN *et al.*, 2012; CLAERHOUT *et al.*, 2018; RALF *et al.*, 2020).

Considerando as classes dicotômicas, foram encontradas diferenças estatisticamente significativas entre os indivíduos com idade <30 e >31 (p-valor = 0,00302) e <40 e >41 (p-valor = 0,00018), e uma diferença quase significativa foi alcançada entre indivíduos com

idade <50 e >51 (*p*-valor = 0,05830) (Tabela 18). Em todos os casos, os indivíduos mais velhos foram associados a maiores taxas de mutação.

Tabela 18. Número de transmissões alélicas pai-filho analisadas considerando intervalos de idade paterna (no momento do nascimento do filho) e a taxa de mutação correspondente

	Classe de idade paterna		Número de transmissões alélicas	Taxa de mutação	Intervalo de confiança (95%)			
	<21	31	7.995	0,00388	0,00264-0,0055			
	21-30	112	39.408	0,00284	0,00234-0,00342			
10 anos	31-40	96	25.995	0,00369 0,00557	0,00299-0,00451			
10 a	41-50	45	8.076		0,00407-0,00745			
	51-60	14	2.407	0,00582	0,00318-0,00974			
	>61	4	834	0,0048	0,00131-0,01223			
sas	<31	143	47.403	0,00302	0,00254-0,00355			
dicotômicas	>30	159	37.312	0,00426	0,00363-0,00498			
cotí	<41	239	73.398	0,00326	0,00286-0,00370			
ss di	>40	63	11.317	0,00557	0,00428-0,00712			
Classes	<51	284	81.474	0,00349	0,00309-0,00392			
	>50	18	3.241	0,00555	0,00329-0,00878			

5.4.3 <u>Variação da taxa de mutação com o número de repetições</u>

É geralmente aceito que alelos mais longos são mais propensos a mutação do que os mais curtos, e que alelos mais longos tendem a perder repetições. Portanto, o número de transferências de alelos, mutações e mutações compatíveis com ganho ou perda de uma ou mais repetições foram analisados para cada marcador, considerando as seguintes categorias para os alelos paternos observados: alelo modal, mais curto ou mais longo que o alelo modal (Tabela 19).

Considerando o subconjunto de 62.378 alelos paternos mais curtos que o modal, 241 sofreram mutação, o que resultou em uma taxa global de mutação de 3,7E-03. Esse valor sobe para 9,8E-03 se consideradas as 621 mutações observadas entre as 63.115 transferências envolvendo alelos paternos maiores que o modal.

Os alelos mais longos mostraram-se 2,55 vezes mais propensos à mutação do que os mais curtos, e essa diferença mostrou forte significância estatística (p=1,36E-37). Além disso, a maioria das mutações envolvendo alelos mais curtos (241) envolveu o ganho de repetições (161), e essa diferença entre ganhos e perdas apresentou significância estatística (p=1,82E-04). O ganho de repetições em alelos curtos já foi anteriormente relatado (BALLANTYNE *et al.*, 2010; SUN *et al.*, 2012; ANTÃO-SOUSA *et al.*, 2019).

Por outro lado, 394 das 621 mutações envolvendo alelos paternos mais longos corresponderam à perda de repetições, sendo esta diferença entre ganhos e perdas também com significância estatística (*p*=8,4E-06). Anteriormente, alguns trabalhos já apontaram que alelos mais longos tendem a perder repetições (BALLANTYNE *et al.*, 2010; SUN *et al.*, 2012; ANTÃO-SOUSA *et al.*, 2019).

Portanto, a tendência dos alelos mais curtos ganharem repetições e dos mais longos perderem repetições mostrou-se altamente significativa (p=8,92E-15).

Tabela 19. Número de transferências de alelos (Transf) e mutações (Mut) observadas compatíveis com alterações envolvendo um número inteiro de repetições, bem como número de mutações compatível com ganhos ou perdas, considerando as seguintes categorias para o alelo paterno: alelo modal, mais curto ou mais longo que o alelo modal.

			Alelo			Tama	nho do a	alelo pater	no ana	lisado	relativo	ao modal))		
Marcadores	Transf	Mut	moda	Mais	curto q	ie o mod	lal		Moo	dal		Mais	longo	que o m	odal
			l	Transf	Mut	Perda	Ganho	Transf	Mut	Perda	Ganho	Transf	Mut	Perda	Ganho
DYS19	11.688	26	14	1.463	2	0	2	5.333	10	1	9	4.892	14	7	7
DYS388	101	1	12	19	1	0	1	73	0	0	0	9	0	0	0
DYS389I	10.616	37	13	2.426	6	1	5	6.011	14	4	10	2.179	17	15	2
DYS389II-I	10.593	48	16	886	4	0	4	5.456	10	0	10	4.251	34	23	21
DYS390	11.825	29	24	4.643	5	1	4	5.174	15	8	7	2.008	9	9	0
DYS391	11.827	35	10	632	1	0	1	6.775	10	1	9	4.420	24	14	10
DYS392	11.810	7	13	5.537	1	1	0	4.794	3	0	3	1.479	3	1	2
DYS393	10.577	17	13	2.131	2	0	2	6.848	8	4	4	1.598	7	5	2
DYS437	10.057	15	14	37	0	0	0	4.720	3	0	3	5.300	12	8	4
DYS438	10.119	6	12	6.340	4	3	1	3.435	2	2	0	344	0	0	0
DYS439	10.078	63	12	4.330	13	3	10	4.258	18	6	12	1.490	32	23	9
DYS448	8.756	13	19	1.083	1	0	1	3.436	6	4	2	4.237	6	6	0
DYS449	5.538	69	30	1.769	12	4	8	952	11	8	3	2.817	46	25	21
DYS456	8.757	53	15	1.213	3	1	2	4.005	14	2	12	3.539	36	25	11
DYS458	8.765	79	17	3.884	23	8	15	2.620	24	12	12	2.261	32	18	14
DYS460	3.411	14	11	1.797	5	1	4	1.466	5	4	1	148	4	4	0
DYS481	3.705	21	22	275	0	0	0	1.221	5	1	4	2.209	16	6	10
DYS518	5.493	93	38	1.425	14	4	10	1.166	15	7	8	2.902	64	32	32
DYS526a	3.119	14	14	890	6	3	3	1.000	4	2	2	1.229	4	2	2
DYS526b-a	3.095	38	22	842	6	2	4	773	8	1	7	1.480	24	12	12
DYS533	3.705	5	12	1.718	1	0	1	1.714	2	1	1	273	2	2	0
DYS547	3.259	57	48	1.223	21	8	13	778	9	4	5	1.258	27	22	5
DYS549	1.317	7	12	240	0	0	0	491	3	2	1	586	4	4	0
DYS570	7.015	64	17	899	7	0	7	2.060	18	6	12	4.056	39	25	14
DYS576	6.914	108	18	3.111	20	3	17	2.191	35	20	15	1.612	53	37	16
DYS612	3.288	57	36	884	21	5	16	933	6	3	3	1.471	30	15	15
DYS626	3.237	35	30	1.241	6	3	3	610	2	1	1	1.386	27	20	7
DYS627	5.663	89	21	2.811	28	10	18	1.152	15	9	6	1.700	46	29	17
DYS635	9.635	35	23	4.424	20	16	4	4.141	9	6	3	1.070	6	3	3
Y-GATA- A10	874	4	15	316	2	1	1	433	2	2	0	125	0	0	0
Y-GATA-H4	9.848	22	12	3.889	6	2	4	5.173	13	9	4	786	3	2	1
Média geral	214.685	1.161	-	62.378	241	80	161	89.192	299	130	169	63.115	621	394	237

CONCLUSÕES

Os resultados obtidos neste trabalho contribuíram com dados haplotípicos para a caracterização das linhagens paternas de populações miscigenadas da América do Sul, permitindo o aprofundamento e detalhamento da sua composição genética.

Após a análise comparativa entre as populações da América do Sul, foi detectado um padrão de ancestralidade masculina semelhante, em que se observa uma grande influência europeia, com destaque para os valores maiores de 90% na Terra do Fogo e no Paraguai, seguida da africana e, em menor proporção, nativa americana, exceto para o Equador em que se observa uma ancestralidade nativo-americana maior que a africana. A análise das distâncias genéticas revelou proximidade entre as populações estudadas e populações europeias e outras sul-americanas relatadas anteriormente na literatura. Na mesma análise, foi observada uma menor proximidade entre as populações estudadas e as populações nativo-americanas da Bolívia e do Brasil, e africanas da África do Sul-Xhosa, Benin e Youruba, destacando o impacto da colonização europeia.

Além disso, a predominância do haplogrupo europeu R1b1a-S116 em todas as linhagens masculinas das populações analisadas do Equador, Terra do Fogo e Paraguai, comumente observado em maiores frequências em países da região da Península Ibérica, reforça, os dados históricos a cerca da predominância da contribuição de colonizadores da Península Ibérica.

As populações estudadas apresentaram valores elevados de HD, como o esperado para populações sul-americanas, levando em conta a miscigenação pelo aporte de pelo menos três continentes diferentes. E, portanto, os marcadores PowerPlex Y23 proporcionaram uma elevada discriminação intra-populacional nas populações do Rio de Janeiro e da Terra do Fogo, a serem usados como base de dados no âmbito forense.

Finalmente, nossos resultados das taxas de mutação estimadas para 201 duos paifilho fornecem evidências quantificadas que suportam as premissas geralmente aceitas de que (i.) mutações de uma etapa são mais comuns do que as de várias etapas, mas mostraram que a magnitude da diferença é altamente variável entre os marcadores analisados, (ii.) os alelos longos são mais propensos à mutação do que os curtos (~ duas vezes), (iii.) os alelos longos são mais propensos a perder repetições (~ duas vezes), enquanto os alelos curtos mostram a tendência oposta (~ duas vezes) e (iv) a taxa de mutação e a idade do pai estão positivamente correlacionadas.

REFERÊNCIAS

ABU-AMERO, K. et al. Saudi Arabian Y-chromosome diversity and its relationship with nearby regions. BMC Genet, v. 10, p. 59, 2009.

ACOSTA, M. A. *et al.* The genetic male component of two South-Western Colombian populations. Forensic Sci Int Genet, v. 3, n. 2, p. e59-61, 2009.

ADAMS, S. M. *et al.* The genetic legacy of religious diversity and intolerance: Paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. Am J Hum Genet, v. 83, p. 725-736, 2008.

ADNAN, A. et al. Mutation analysis at 17 Y-STR loci (Yfiler) in father-son pairs of male pedigrees from Pakistan. Forensic Sci Int Genet, v. 36, p. e17-e18, 2018.

ALENCAR, F.; RAMALHO, L.C.; RIBEIRO, M.V.T. **História da sociedade brasileira**. 18 ed. Rio de Janeiro: Ao Livro Técnico, 467 p., 1996.

ALI, S.; HASNAIN, S. E. **Molecular Dissection of the human Y-chromosome**. Gene, v. 191, p. 4195-4206, 2002.

ALONSO, S. *et al.* The place of the Basques in the European Y-chromosome diversity landscape. Eur J Hum Genet, v. 13, p. 1293-1302, 2005.

ALVAREZ-CUBERO, M. J. et al. Next generation sequencing: an application in forensic sciences? Ann Hum Biol, v. 44, n. 7, p. 581-592, 2017.

ALVES, C. *et al.* Population data for Y-chromosome haplotypes defined by 17 STRs (AmpFISTR YFiler) in Portugal. Forensic Sci Int, v. 171, n. 2-3, p. 250-255, 2007.

ALVES-SILVA, J. *et al.* **The ancestry of Brazilian mtDNA lineages**. Am J Hum Genet, v. 67, p. 444-61, 2000.

AMBROSIO, B. *et al.* The Andalusian population from Huelva reveals a high diversification of Y-DNA paternal lineages from haplogroup E: Identifying human male movements within the Mediterranean space. Ann Hum Biol, v. 37, n. 1, p. 86-107, 2010.

AMBROSIO, I. B. *et al.* Mutational data and population profiling of 23 Y-STRs in three Brazilian Populations. Forensic Sci Int Genet, v. 48, p. 102348, 2020.

AMORIM, A. **Population Genetics**. *In*: Maloy, S.; Hughes, K. (Eds.). **Brenner's Encyclopedia of Genetics**. 2 ed. New York: Elsevier, p. 407-411, 2013.

AMORIM, A.; BUDOWLE, B. Handbook of Forensic Genetics: Biodiversity and Heredity in Civil and Criminal Investigation. 1 ed. World Scientific Publishing Europe Ltd, 652 p., 2016.

ANSARI-POUR, N.; PLASTER, C. A.; BRADMAN, N. Evidence from Ychromosome analysis for a late exclusively eastern expansion of the Bantu-speaking people. Eur J Hum Genet, v. 21, n. 4, 423-429, 2013.

ANSARI-POUR, N. et al. Palenque de San Basilio in Colombia: genetic data support an oral history of a paternal ancestry in Congo. Proc R Soc B, v. 283, p. 20152980, 2015.

ANTÃO-SOUZA, S. *et al.* Mutation rates and segregation data on 16 Y-STRs: an update to previous GHEP-ISFG studies. Forensic Sci Int Genet Suppl Ser, v. 6, p. e601-e601, 2017.

ANTÃO-SOUZA, S. *et al.* **Mutation in Y STRs: Repeat motif gains vs. losses.** Forensic Sci Int Genet Suppl Ser, vol. 7, n. 1, p. 240-242, 2019.

ANTÃO-SOUSA, S. *et al.* The sequence of the repetitive motif influences the frequency of multistep mutations in Short Tandem Repeats. Sci Rep, v. 13, n. 1, p. 10251, 2023.

ARANGO, R.; SANCHÉZ, E. Los pueblos indígenas de Colombia 1997, desarrollo y territorio. Departamento Nacional de Planeación, Unidad Administrativa Especial de Desarrollo Territorial, 334 p., 1998.

ARANGO, R; SÁNCHEZ, S. G. Los pueblos indígenas de Colombia en el umbral del nuevo milenio: Población, cultura y territorio: bases para el fortalecimiento social y económico de los pueblos indígenas. 1 ed. Bogotá: Departamento Nacional de Planeación, 553 p., 2004.

ARREDI, B. *et al.* A Predominantly Neolithic Origin for Y-Chromosomal DNA Variation in North Africa. Am J Hum Genet, v. 75, p. 338-345, 2004.

AYALA, M. **Resumen de Historia del Ecuador.** 3 ed. Quito: Corporación Editora Nacional, 58 p., 2008.

ATHEY, T. W. Haplogroup Prediction from Y-STR Values Using a Bayesian-Allele-Frequency Approach. J Genet Genealogy, v. 2, n. 2, p. 34-39, 2006.

AZEVEDO, D. A. *et al.* **Analysis of Y chromosome SNPs in Alagoas, Northeastern Brazil**. For Sci Int Genet Sup Ser, v. 2, n. 1, p. 421-422, 2009.

AY, M. et al. Genetic characterisation of 13 rapidly mutating Y-STR loci in 100 father and son pairs from South and East Turkey. Ann Hum Biol v. 45, n. 6-8, p. 506–515, 2019.

BACHTROG, D. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nat Rev Genet, v. 14, n. 2, p. 113-24, 2013.

BAEZA, C. et al. Population data for 15 Y-chromosome STRs in a population sample from Quito (Ecuador). Forensic Sci Int, v. 173, p. 214-219, 2007.

BAILLIET, G. *et al.* **Brief Communication: Restricted Geographic Distribution for Y-Q* Paragroup in South America**. Am J Phys Anthropol, v. 140, p. 578-582, 2009.

BAILLIET, G. et al. Antecedentes y nuevos aportes en el estudio del Cromosoma Y en poblaciones humanas sudamericanas. BAG J Basic Appl Genet, v. 22, p. 1-9, 2011.

BALLARD, D. J. et al. A study of mutation rates and the characterisation of intermediate, null and duplicated alleles for 13 Y chromosome STRs. Forensic Sci Int, v. 155, n. 1, p. 65-70, 2005.

BALANOVSKY, O. *et al.* Parallel evolution of genes and languages in the Caucasus region. Mol Biol Evol, v. 28, p. 2905-2920, 2011.

BALANOVSKY, O. Toward a consensus on SNP and STR mutation rates on the human **Y-chromosome.** Hum Genet, v. 136, p. 575-590, 2017.

BALARESQUE, P. *et al.* Dynamic nature of the proximal AZFc region of the human Y chromosome: multiple independent deletion and duplication events revealed by microsatellite analysis. Hum Mutat, v. 29, p. 1171-1180, 2008.

BALLANTYNE, K. N. *et al.* **Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications**. Am J Hum Genet, v. 87, n. 3, p. 341-353, 2010.

BALLANTYNE, K. N. *et al.* A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. Forensic Sci Int Genet, v. 6, n. 2, p. 208-218, 2012.

BALLANTYNE, K. N. et al. Toward male individualization with rapidly mutating y-chromosomal short tandem repeats. Hum Mutat, v. 35, n. 8, p. 1021-1032, 2014.

BALDACCHINO, G. The Political Economy of Divided Islands Unified Geographies, Multiple Polities. 1 ed., Palgrave MacMilla, 284 p., 2013.

BARROS, F. M. de M. Composição genética de populações do Sudeste Asiático: análise dos principais haplogrupos da região e contribuição para o entendimento das teorias das rotas de migração. 2019. 111 f. Tese (Doutorado em Biociências) - Instituto de Biologia Roberto Alcântara Gomes, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil, 2019.

BASCOPÉ, J. Sentidos coloniales I. El oro y la vida salvaje en Tierra del Fuego 1880–1914. Magallania, v. 38, n. 2, p. 5-26, 2010.

BATTAGLIA, V. *et al.* Y-chromosomal evidence of the cultural diffusion of agriculture in Southeast Europe. Eur J Hum Genet, v. 17, p. 820-830, 2009.

BATTAGLIA, V. *et al.* The First Peopling of South America: New Evidence from Y-Chromosome Haplogroup Q. PLoS ONE, v. 8, n. 8, p. e71390, 2013.

BAYE, T. M. Inter-chromosomal variation in the pattern of human population genetic structure. Hum Genomics, v. 4, p. 220, 2011.

BEHAR, D. M. et al. Contrasting patterns of Y chromosome variation in Ashkenazi Jewish and host non-Jewish European populations. Hum Genet, v. 114, p. 354-365, 2004.

BELEZA, S. *et al.* **The genetic legacy of western Bantu migrations**. Hum Genet, v. 117, p. 366-375, 2005.

BELEZA, S. *et al.* Micro-phylogeographic and demographic history of Portuguese male lineages. Ann Hum Genet, v. 70, p. 181-194, 2006.

BELLWOOD, P.; NESS, I. **The Global Prehistory of Human Migration.** 1 ed. Wiley-Blackwell, 448 p., 2014.

BENÍTEZ, L. Culturas ecuatorianas: aver v hoy. 1 ed. Abya-Yala, 231 p., 1993.

BENTO, A. M. et al. Distribution of Y-chromosomal haplotypes in the Central Portuguese population using 17-STRs. Forensic Sci Int Genet, v. 4, n. 1, e35-6, 2009.

BERGER, B. A. et al. Y-STR typing of an Austrian population sample using a 17-loci multiplex PCR assay. Int J Legal Med, v. 119, n. 4, p. 241-246, 2005.

BERNIELL-LEE, G. et al. Genetic and Demographic Implications of the Bantu Expansion: Insights from Human Paternal Lineages. Mol Biol. Evol, v. 26, n. 7, p. 1581-1589, 2009.

BIRD, S. C. Towards Improvements in the Estimation of the Coalescent: Implications for the Most Effective Use of Y Chromosome Short Tandem Repeat Mutation Rates. PLoS ONE, v. 7, n. 10, p. e48638, 2012.

BISSO-MACHADO, R. *et al.* **Distribution of Y-chromosome Q lineages in native Americans**. Am J Hum Biol, v. 23, n. 4, p. 563-566, 2011.

BISSO-MACHADO, R.; BORTOLINI, M. C.; SALZANO, F. M. Uniparental genetic markers in South Amerindians. Genet Mol Biol, v. 35, n. 2, p. 365-387, 2012.

BLANCHI, N. O. *et al.* Characterization of Ancestral and Derived Y-Chromosome Haplotypes of New World Native Populations. Am J Hum Genet, v. 63, n. 6, p. 1862-1871, 1998.

BLANCO-VEREA, A. *et al.* **Analysis of Y chromosome lineages in native South American population**. Int Congr Ser, v. 1288, p. 222-224, 2006.

BLANCO-VEREA, A. et al. Y-chromosome lineages in native South. American population. Forensic Sci Int Genet, v. 4, n. 3, p. 187-193, 2010.

BREDEMEYER, S.; ROEWER, L.; WILLUWEIT, S. Next generation sequencing of Y-STRs in father-son pairs and comparison with traditional capillary electrophoresis. Forensic Sci Res, v. 7, n. 3, p. 484-489, 2021.

BRINKMANN, B. *et al.* **Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat.** Am J Hum Genet, v. 52, n. 6, p. 1408-1415, 1998.

BRION, M. *et al.* Hierarchical analysis of 30 Y-chromosome SNPs in European populations. Int J Médico Legal, v. 119, p. 10-15, 2005.

BROCHADO J. P. An ecological model of the spread of pottery and agriculture into Eastern South America. 1984. Tese (Doutorado). University of Illinois at Urbana-Champaign, Illinois, Estados Unidos, 1984. Não paginado.

BOATTINI, A. *et al.* Estimating Y-Str Mutation Rates and Tmrca Through Deep-Rooting Italian Pedigrees. Sci Rep, v. 9, p. 9032, 2019.

BOATTINI, A. *et al.* Uniparental Markers in Italy Reveal a Sex-Biased Genetic Structure and Different Historical Strata. PLoS One, v. 8, n. 5, p. e65441, 2013.

BOBILLO, M. C. *et al.* Amerindian mitochondrial DNA haplogroups predominate in the population of Argentina: towards a first nationwide forensic mitochondrial DNA sequence database. Int J Legal Med, v. 124, n. 4, p. 263-268, 2010.

BORTOLINI, M. *et al.* **Y-Chromosome Evidence for Differing Ancient Demographic Histories in the Americas**. Am J Hum Genet, v. 73, p. 524-539, 2003.

BOURGEON, L.; BURKE, A.; HIGHAM, T. Earliest Human Presence in North America Dated to the Last Glacial Maximum: New Radiocarbon Dates from Bluefish Caves. PLoS One, v. 12, p. e0169486, 2017.

BORRERO, L. A. The origins of ethnographic subsistence patterns in Fuego Patagonia. *In*: MCEWAN, C.; BORRERO L. A.; PRIETO A. (Eds.) Patagonia: Natural history, prehistory and ethnography at the uttermost end of the earth. London: British Museum Press, p. 60-81, 1997.

BOSCH, E. *et al.* Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. Am J Hum Genet, v. 65, p. 1623-1638, 1999.

BOSCH, E. *et al.* High-resolution analysis of human Y chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian Peninsula. Am J Hum Genet, v. 68, p. 1019-1029, 2001.

BOSCH, E.; JOBLING, M. Duplications of the AZFa region of the human Y chromosome are mediated by homologous recombination between HERVs and are compatible with male fertility. Hum Mol Genet, v.12, n. 3, p. 341-347, 2003.

BUDOWLE, B. et al. Twelve short tandem repeat *loci* Y chromosome haplotypes: genetic analysis on populations residing in North America. Forensic Sci Int, v. 150, p. 1-15, 2005.

BUDOWLE, B.; DAAL, A. **Forensically relevant SNP classes.** Biotechniques, v. 44, n. 5, p. 603-608, 2008.

- BUGOYE, F. C.; MULIMA, E.; MISINZO, G. Analysis of Mutation Rate of 17 Y-Chromosome Short Tandem Repeats *Loci* Using Tanzanian Father-Son Paired Samples. Genet Res Int, p. 8090469, 2018.
- BURGOS, G. et al. A look of paternal ancestry in a sample of Ecuadorian "MESTIZO" population analyzed through PowerPlex Y23. Forensic Sci Int Genet Suppl Ser, v. 7, n. 1, 534-536, 2019a.
- BURGOS, G. et al. An approach to maternal ancestry in a sample of Ecuadorian "mestizo" population by sequencing the control region of mtDNA. Forensic Sci Int Genet Suppl Ser, v. 7, n. 1, 537-538, 2019b.
- BUTLER, J. M. Advanced Topics in Forensic DNA Typing: Interpretation. 1 ed. Academic Press, 588 p., 2014.
- BUTLER, J. M *et al.* Chromosomal Duplications Along the Y-Chromosome and Their Potential Impact on Y-STR Interpretation. J Forensic Sci, v. 50, p. 853-859, 2005.
- BUTLER, J. M. Fundamentals of Forensic DNA Typing. 1 ed. Academic Press, 500 p., 2009.
- BUTLER, J. M. Fundamentals of Forensic DNA Typing. 2 ed. Elsevier Academic Press, 520 p., 2010.
- CAI, X. et al. Human Migration through Bottlenecks from Southeast Asia into East Asia during Last Glacial Maximum Revealed by Y Chromosomes. PLoS ONE, v. 6, n. 8, p. e24282, 2011.
- CAMPBELL, M. C.; TISHKOFF, S. A. African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. Annu Rev Genomics Hum Genet, v. 9, p. 403-433, 2008.
- CAMPOS, A. Caracterização Genética das Linhagens Masculinas Presentes na População da Nigéria. 2018. Dissertação (Mestrado). Faculdade de Ciências da Universidade do Porto, Porto, Portugal, 2018. Não paginado.
- CAPELLI, C. *et al.* **A Y chromosome census of the British Isles**. Curr Biol, v. 13, p. 979-984, 2003.
- CAPELLI, C. *et al.* Phylogenetic evidence for multiple independent duplicationevents at the DYS19 locus. Forensic Sci Int Genet, v. 1, p. 287-290, 2007.
- CAPELLI, C. *et al.* Moors and Saracens in Europe: Estimating the medieval North African male legacy in southern Europe. Eur J Hum Genet, v. 17, p. 842-852, 2009.
- CAPUTO, M. et al. Demand for larger Y-STR reference databases in ethnic melting-pot countries: Argentina as a test case. Int J Legal Med, v. 133, p. 1309-1320, 2019.
- CARDOSO, S. *et al.* Genetic uniqueness of the Waorani tribe from the Ecuadorian Amazon. Heredity, v. 108, p. 609-615, 2012.

CASTAGNOLA, J. et al. Inferring the genetic structure of Northwestern Argentina by uniparental SNP typing. Forensic Sci Int Genet Suppl Ser, v. 7, p. 306–309, 2019.

CATANESI, C. I. *et al.* **Geographic structure in Gran Chaco Amerindians based on five X-chromosome STRs.** Hum Biol, v. 79, p. 463-474, 2007.

CAVALLI-SFORZA, L. L.; MENOZZI, P.; PIAZZA, A. The history and Geography of Human Genes. J Asian Studes, v. 54, p. 432, 1995.

CAVALLI-SFORZA, L. L.; FELDMAN, M. W. The application of molecular genetic approaches to the study of human evolution. Nat Genet, v. 33, p. 266-275, 2003.

CAYCEDO, A. U. **De los Taironas los Kogi: Una interpretación del cambio cultural.** Boletín Museo del Oro, n. 17, p. 32-43, 1986.

CETKOVIC GENTULA, M.; NEVSKI, A. NEVGEN Y-DNA **Haplogroup Predictor**. Disponível em: http://www.nevgen.org/>. Acesso em: 3 maio, 2020.

CHEN, Z. et al. Brief Communication: Y-chromosome Haplogroup Analysis Indicates That Chinese Tuvans Share Distinctive Affinity With Siberian Tuvans. Am J Phys Anthropol, v. 144, n. 3, p. 492-497, 2011.

CHIARONI, J.; UNDERHILL, P. A.; CAVALLI-SFORZA, L. L. Y chromosome diversity, human expansion, drift, and cultural evolution. PNAS, v. 106, n. 48, p. 20174-20179, 2009.

CLAERHOUT, S. *et al.* A game of hide and seq: Identification of parallel Y- STR evolution. Eur J Hum Genet, v. 27, p. 637-646, 2018.

CLAERHOUT, S. *et al.* **CSYseq:** The first Y-chromosome sequencing tool typing a large number of Y-SNPs and Y-STRs to unravel worldwide human population genetics. PLoS Genet, v. 17, n. 9, p. e1009758, 2021.

CLARCK, A. G.; HARTL, D. L. **Principles of Population Genetics.** 4 ed. Oxford University Press Inc, p. 545, 2007.

CÁRDENAS, J. M. *et al.* The multiethnic ancestry of Bolivians as revealed by the analysis of Y-chromosome markers. Forensic Sci Int Genet, v. 14, p. 210-218, 2015.

CHENNAKRISHNAIAH, S. *et al.* Indigenous and Foreign Y-chromosomes Characterize the Lingayat and Vokkaliga Populations of Southwest India. Gene, v. 526, n. 2, p. 96-106, 2013.

CHIARONI, J. et al. Y chromosome diversity, human expansion, drift, and cultural evolution. Proc Natl Acad Sci USA, v. 106, p. 20174-20179, 2009.

CHIARONI, J. et al. The emergence of Y-chromosome haplogroup J1e among Arabic-speaking populations. Eur J Hum Genet, v. 18, p. 348-353, 2010.

- CINNIOGLU, C. *et al.* Excavating Y-chromosome haplotype strata in Anatolia. Hum Genet, v. 114, p. 127-148, 2004.
- COELHO, M. et al. On the edge of Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. BMC Evol Biol, v. 9, p. 80, 2009.
- COMAS, D.; LUISELLI, D.; RICKARDS, O. Human population genetics of the Mediterranean. Ann Hum Biol, v. 45, n. 1, p. 1-4, 2018.
- CONDEPE. **Sistema de Indicadores de Nacionalidades y Pueblos indígenas del Ecuador (SIDENPE)**, 2002. Disponível em: < http://www.codenpe.gov.ec/>. Acesso em: 15 julho. 2022.
- CORACH, D. *et al.* Inferring Continental Ancestry of Argentineans From Autosomal, Y-chromosomal and Mitochondrial DNA. Ann Hum Genet, v. 74, p. 65-76, 2010.
- CRIOLLO-RAYO, A. A. *et al.* Native American gene continuity to the modern admixed population from the Colombian Andes: Implication for biomedical, population and forensic studies. Forensic Sci Int Genet, v. 36, p. e1-e7, 2018.
- CRUCIANI, F. *et al.* A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. Am J Hum Genet, v. 70, p. 1197-1214, 2002.
- CRUCIANI, F. *et al.* Phylogeographic Analysis of Haplogroup E3b (E-M215) Y Chromosomes Reveals Multiple Migratory Events Within and Out Of Africa. Am J Hum Genet, v. 7, n. 5, p. 1014-1022, 2004.
- CRUCIANI, F. *et al.* Tracing Past Human Male Movements in Northern/Eastern Africa and Western Eurasia: New Clues from Y-Chromosomal Haplogroups E-M78 and J-M12. Mol Biol Evol, v. 24, n. 6, p. 1300-1311, 2007.
- CRUCIANI, F. *et al.* **Human Y chromosome haplogroup R-V88:** a paternal genetic record of early mid Holocene trans-Saharan connections and the spread of Chadic languages. Eur J Hum Genet, v. 18, n. 7, p. 800-807, 2010.
- CRUCIANI, F. *et al.* Strong intra- and inter-continental differentiation revealed by Y chromosome SNPs M269, U106 and U152. Forense Sci Int Genet, v. 5, n. 3, p. e49-52, 2011.
- CSÁNYI, B. *et al.* Y-Chromosome Analysis of Ancient Hungarian and Two Modern Hungarian-Speaking Populations from the Carpathian. Ann Hum Genet, v. 72, p. 519-534, 2008.
- CUMMINGS, C. J.; ZOGHBI, H. Y. Fourteen and counting: unraveling trinucleotide repeat diseases. Hum Mol Genet, v. 9, p. 909-916, 2000.
- CURTIN, P. D. **The Atlantic slave trace: a census**. 1 ed. The University of Wisconsen Press, Milwaukee, 338 p., 1969.

DANE. **Colombia una nación multicultural:** Su diversidad étnica. Departamento Administrativo Nacional de Estadística, Dirección de Censos y Demografía, 45 p., 2007.

da SILVA, R. F. O Rio Antes do Rio. 4 ed. Relicário, 472 p., 2019.

D'ATANASIO, E. *et al.* The peopling of the last Green Sahara revealed by high-coverage resequencing of trans-Saharan patrilineages. Genome Biol, v. 19, p. 20, 2018.

DECKER, E. et al. Analysis of mutations in father-son pairs with 17 Y-STR loci. Forensic Sci Int Genet., v. 2, n. 3, p. e31-5, 2008.

DENEVAN, W. M. The Native Population of the Americas in 1492 reconsidered. Revista de Indias, v. 63, n. 227, p. 175-188, 2003.

de SOUZA GÓES, A. C. *et al.* **Population and mutation analysis of 17 Y-STR** *loci* **from Rio de Janeiro (Brazil)**. Int J Legal Med, v. 119, n. 2, p. 70-76, 2005.

DI CORCIA, T. et al. Uniparental Lineages from the Oldest Indigenous Population of Ecuador: The Tsachilas. Genes (Basel), v. 12, n. 8, p. 1273, 2021.

DIEDERICH, M, *et al.* A case of double alleles at three Y-STR *loci*: forensic implications. Int J Legal Med, n. 119, p. 223-225, 2005.

DI GAETANO, C. et al. Differential Greek and northern African migrations to Sicily are supported by genetic evidence from the Y chromosome. Eur J Hum Genet, v. 17, p. 91-99, 2009.

DI GIACOMO, F. *et al.* Clinal patterns of human Y chromosomal diversity in continental Italy and Greece are dominated by drift and founder effects. Mol Phylogenet Evol, v. 28, p. 387-395, 2003.

DILLEHAY, T. D. *et al.* Monte Verde: seaweed, food, medicine, and the peopling of South America. Sci, v. 320, p. 784-786, 2008.

DOMINGUES, P. M. et al. Sub-Saharan Africa descendents in Rio de Janeiro (Brazil): population and mutational data for 12 Y-STR *loci*. Int J Legal Med, v. 121, p. 238-241, 2007.

DORNELLES, C. L. *et al.* **Mitochondrial DNA and Alu insertions in a genetically peculiar population: the Ayoreo Indians of Bolivia and Paraguay**. Am J Hum Biol, v. 16, p. 479-488, 2004.

DUPUY, B. M. *et al.* **Y-chromosome variation in a Norwegian population sample**. Forensic Sci Int, v. 117, n. 3, p. 163-173, 2001.

DUPUY, B. M. *et al.* **Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within** *loci***. Hum. Mutat, v. 23, p. 117-124, 2004.**

ECKERT, K. A.; HILE, S. E. Every microsatellite is different: Intrinsic DNA features dictate mutagenesis of common microsatellites present in the human genome. Mol Carcinog, v. 48, n. 4, p. 379-388, 2009.

ELLEGREN, H. Microsatellites: simple sequences with complex evolution. Nat Rev, v. 5, p. 435-445, 2004.

ELLIOTT, J. H. The Spanish Conquest and settlement of America In: Bethell L, editor. The Cambridge History of Latin America. Cambridge University Press, p. 147-206, 1984.

EL-SIBAI, M. *et al.* Geographical structure of the Y-chromosomal genetic landscape of the Levant: a coastal-inland contrast. Ann Hum Genet, v. 73, n. 6, p. 568-581, 2009.

ENNAFAA, H. *et al.* **Mitochondrial DNA and Y-chromosome microstructure in Tunisia.** J Hum Genet, v. 56, p. 734-741, 2011.

FADHLAOUI-ZID, K. *et al.* **Genetic structure of Tunisian ethnic groups revealed by paternal lineages**. Am J Phys Anthropol, v. 146, p. 271-280, 2011.

FADHLAOUI-ZID, K. *et al.* **Sousse: extreme genetic heterogeneity in North Africa**. J Hum Genet, p. 1-9, 2014.

FAUSTO, B. História do Brasil. 4 ed. Edusp, 639 p., 1996.

FAUSTO, C. Os índios antes do Brasil. 4 ed. Rio de Janeiro: Zahar, 94 p., 2010.

FERRARI, L. San Basilio de Palenque (Colombia): un performance de libertad. *In*: MAGLIA, G.; SCHWEGLER, A. (Eds.). Palenque (Colombia): oralidad, identidad y resistencia. Un enfoque interdisciplinario. Bogota: Instituto Caro y Cuervo & Universidad Javeriana, p. 57-84, 2012.

FIEDEL, S. J. The peopling of the New Word: present evidence, new theories and futuredirection. J Archael Research, v. 8, p. 39-103, 2000.

FIGUEIREDO, R. F. *et al.* Male-specific contributions to the Brazilian population of **Espirito Santo**. Int J Legal Med, v. 130, n. 3, p. 679-81, 2016.

FLORES, C. *et al.* **Genetic structure of Iberian Peninsula revealed by Y chromosome analysis.** Eur J Hum Genet, v. 12, n.10, p. 855-863, 2004.

FLORES-ALVARADO, S. et al. Ancestry and admixture of a southernmost Chilean population: The reflection of a migratory history. Am J Hum Biol, v. 34, p. e23598, 2022.

FLORES-ESPINOZA, R. *et al.* **Investigating genetic diversity in admixed populations from Ecuador**. Am J Phys Anthropol, v. 176, n. 1, p. 109-119, 2021.

FORSTER, P. Ice Ages and the mitochondrial DNA chronology of human dispersals: a review. Philos Trans R Soc Lond B Biol Sci, v. 359, p. 255-264, 2004.

- FRANCALACCI, P.; SANNA, D. History and geography of human Y-chromosome in Europe: a SNP perspective. J Anthropol Sci, v. 86, p. 59-89, 2008.
- FRANCALACCI, P. et al. Low-Pass DNA Sequencing of 1200 Sardinians Reconstructs European Y-Chromosome Phylogeny. Sci, v. 341, n. 6145, p. 565-569, 2013.
- FRIEDEMANN, N; ROSSELLI, C.P. Lengua y sociedad en El Palenque de San Basilio. Bogotá: Instituto Caro y Cuervo, 300 p., 1983.
- da FRÉ, N. N. *et al.* Genetic data and de novo mutation rates in father-son pairs of 23 Y-STR *loci* in Southern Brazil population. Int J Legal Med, v. 129, p. 1221-1223, 2015.
- FU, J. et al. Assessing 23 Y-STR loci mutation rates in Chinese Han father-son pairs from southwestern China. Mol Biol Rep, v. 47, n. 10, p. 7755-7760, 2020.
- GARCÍA, A. *et al.* Genetic variation in populations from central Argentina based on mitochondrial and Y chromosome DNA evidence. J Hum Genet, v. 63, n. 4, p. 493-507, 2018.
- GAVIRIA, A. et al. Characterization and haplotype analysis of 11 Y-STR loci in ecuadorian population. Forensic Sci Int Genet Suppl Ser, v. 4, p. 310-311, 2013.
- GAYDEN, T. *et al.* The Himalayas as a Directional Barrier to Gene Flow. Am J Hum Genet, v. 80, 2007.
- GE, J. et al. Mutation rates at Y chromosome short tandem repeats in Texas populations. Forensic Sci Int Genet, v. 3, p. 179-184, 2009.
- GEPPERT, M. et al. Identification of new SNPs in native South American populations by resequencing the y chromosome. Forensic Sci Int Genet, v. 15, p. 111-114, 2015.
- GOEBEL, T.; WATERS, M. R.; O'ROURKE, D. H. The late Pleistocene dispersal of modern humans in the Americas. Sci, v. 319, n. 5869, p. 1497-502, 2008.
- GOEDBLOED, M. et al. Comprehensive mutation analysis of 17 Ychromosomal short tandem repeat polymorphisms included in the AmpFSTR® Yfiler® PCR amplification kit. Int J Legal Med, v. 123, p. 471-482, 2009.
- GONÇALVES, R. *et al.* **Y-chromosome Lineages from Portugal, Madeira and Açores Record Elements of Sephardim and Berber Ancestry**. Ann Hum Genet, v. 69, p. 443-454, 2005.
- GONZÁLEZ, M. et al. The genetic landscape of Equatorial Guinea and the origin and migration routes of the Y chromosome haplogroup RV88. Eur J Hum Genet, v. 21, n. 3, p. 324-331, 2013.
- GONZÁLEZ-ANDRADE, F. et al. Sex-specific genetic admixture of Mestizos, Amerindian Kichwas, and AfroEcuadorans from Ecuador. Hum Biol, v. 78, p. 51-78, 2007.

GONZÁLEZ-ANDRADE, F. et al. Y-STR variation among ethnic groups from Ecuador: Mestizos, Kichwas, Afro-Ecuadorians and Waoranis. Forensic Sci Int Genet, v. 3, p. e83-e91, 2009.

GONZÁLEZ-ANDRADE, F. *et al.* The genetic ethnography of the Native Amerindians of Ecuador. Int J Med Biol Front, v. 17, p. 117-136, 2011.

GOMES, V. *et al.* **Digging deeper into East African human y chromosome lineages**. Hum Genet, v. 127, n. 5, p. 603-613, 2010.

GÓMEZ-CARBALLA *et al.* Revealing latitudinal patterns of mitochondrial DNA diversity in Chileans. Forensic Sci Int Genet, v. 20, p. 81-88, 2016.

GREENBERG, J. et al. The settlement of the Americas: a comparison of linguistic, dental, and genetic evidence. Curr Anthropol, v. 4, p. 477-497, 1986.

GRUGNI, V. *et al.* Reconstructing the genetic history of Italians: new insights from a male (Y-chromosome) perspective. Ann Hum Biol, v. 45, n. 1, p. 44-56, 2018.

GYMREK, M. et al. Interpreting short tandem repeat variations in humans using mutational constraint. Nat Genet, v. 49, n. 10, p. 1495-1501, 2017.

GUSMÃO, L. *et al.* **Mutation rates at Y chromosome specific microsatellites.** Human Mutat, v. 26, p. 529-528, 2005.

GUSMÃO, L.; BRIÓN, M.; GOMES, I. Chapter 30 The human Y chromosome malespecific polymorphisms and forensic genetics. 6 ed. Handbook of Analytical Separations, p. 969-1000, 2007.

GUSMÃO, L.; CARRACEDO, A. Y Chromosome-Specific STRs. Profiles in DNA, p. 3-6, 2003.

HABER, M. *et al.* **Chad Genetic Diversity Reveals an African History Marked by Multiple Holocene Eurasian Migrations**. Am J Hum Genet, v. 99, n. 6, p. 1316-1324, 2016.

HAMMER, M. F. *et al.* Out of Africa and back again: Nested cladistic analysis of human Y chromosome variation. Mol Biol Evol, v. 15, p. 427-441, 1998.

HAMMER, M. *et al.* **Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes.** Proc Natl Acad Sci, v. 97, p. 6769-6774, 2000.

HAMMER, M. F. *et al.* **Hierarchical patterns of global human Y-chromosome diversity**. Mol Biol Evol, v. 18, p. 1189-1203, 2001.

HAMMER, M. F. *et al.* **Dual origins of the Japanese: common ground for hunter-gatherer and farmer Y chromosomes**. J Hum Genet, v. 51, p. 47-58, 2006.

HAMMER, M. F. *et al.* **Jewish and middle-eastern non-Jewish populations share a common pool of Y chromosome biallelic haplotypes.** Proc Natl Acad Sci USA, v. 97, p. 6769-6774, 2000.

HARAMBOUR, A. R. Los Prohombres y los extintos. Patrimonio, identidad e historiografia regional en Magallanes. Cuadernos de Historia, v. 48, p. 57-88, 2018.

HEBBLETHWAITE, M. Paraguay. 2 ed., Bradt Travel Guides, 384 p., 2014.

HERNANDEZ, C. L. *et al.* Paternal lineages in southern Iberia provide time frames for gene flow from mainland Europe and the Mediterranean world. Ann Hum Biol, v. 46, bn. 1, p. 63-76, 2019.

HOHOFF, C. *et al.* **Y-chromosomal microsatellite mutation rates in a population sample from northwestern Germany.** Int J Legal Med, v. 121, n. 5, p. 359-363, 2007. HOMBURGER, J. R. *et al.* **Genomic Insights into the Ancestry and Demographic History of South America.** PLoS Genetics, v. 11, p. e1005602, 2015.

HURLES, M. E. *et al.* The dual origin of the Malagasy in Island Southeast Asia and East Africa: evidence from maternal and paternal lineages. Am J Hum Genet, v. 76, p. 894-901, 2005.

IBGE. Instituto Brasileiro de Geografia e Estatística. Brasil: 500 anos de povoamento. Rio de Janeiro: IBGE, 2007. Disponível em: < https://celp.fflch.usp.br/brasil-500-anos-de-povoamento/>. Acesso em: 21 março. 2021.

INEC. National Institute of Statistics and Censuses. VI Population Census and V Housing Census of the Ecuador, 2001. Disponível em: < http://www.inec.gov.ec/>. Acesso em: 20 julho. 2022.

ISOGG.ORG. **International Society of Genetic Genealogy.** Disponível em: http://isogg.org/. Acesso em: 21 março. 2023.

JANNUZZI, J. et al. Genetic characterization of Rio de Janeiro for different Y-STR sets. Int J Legal Med, v. 132, n. 5, p. 1313-1315, 2018.

JANNUZZI, J. et al. New insights on intercontinental origins of paternal lineages in Northeast Brazil. BMC Evol Biol, v. 20, n. 1, p. 15, 2020.

JERMYN, L.; LIN, Y. J. Paraguay. 2 ed. Cavendish Square Publishing, 144 p., 2010.

JIN, H. J. *et al.* **Y-chromosomal DNA haplogroups and their implications for the dual origins of the Koreans.** Hum Genet, v. 114, p. 27-35, 2003.

JOBLING, M. A. **Y-chromosomal SNP haplotype diversity in forensic analysis.** Forensic Sci Int, v. 118, n. 2-3, p. 158-162, 2001.

JOBLING, M. A. **The impact of recent events on human genetic diversity**. Philos Trans R Soc Lond B Biol Sci, v. 367, n. 1590, p. 793-799, 2012.

JOBLING, M. A.; PANDYA, A.; TYLER-SMITH, C. The Y chromosome in forensic analysis and paternity testing. Int J Legal Med, v. 110, p. 118-124, 1997.

JOBLING, M. A.; TYLER-SMITH, C. New uses for new haplotypes. Trends Genet, v. 16, n. 8, p. 356-362, 2000.

JOBLING, M. A.; TYLER-SMITH, C. The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet, v. 4, p. 598-612, 2003.

JOBLING, M. A.; HURLES, M.; TYLER-SMITH, C. Human evolutionary genetics: origins, peoples & disease. 1 ed. Garland Science, 458 p., 2013.

JORDE, L. B. *et al.* **The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data.** Am J Hum Genet, v. 66, n. 3, p. 979-988, 2000.

JOTA, M. S. *et al.* **A New Subhaplogroup of Native American Y-Chromosomes From the Andes.** Am J Phys Anthropol, v. 146, p. 553-559, 2011.

JOTA, M. S. *et al.* **New native South American Y chromosome lineages**. J Hum Genet, p. 1-11, 2016.

JURADO MEDINA, L. S. *et al.* En búsqueda de la herencia americana en poblaciones mestizas análisis comparativo entre muestras masculinas de Patagonia, Noroeste y Cuyo. *In*: Proceedings of the XII Jornadas Nacionales de Antropología Biológica. Corrientes, Argentina, p. 82, 2015.

KARAFET, T. M. *et al.* Balinese Y-chromosome perspective on the peopling of Indonesia: Genetic contributions from pre-Neolithic hunter-gatherers, Austronesian farmers, and Indian traders. Hum Biol, v. 77, p. 93-114, 2005.

KARAFET, T. M. *et al.* New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res. v. 18, n. 5, p. 830-838, 2008.

KARAFET, T. M. *et al.* **Major east-west division underlies Y-chromosome stratification across Indonesia.** Mol Biol Evol, v. 27, p. 1833-1844, 2010.

KAYSER, M. *et al.* **Melanesian origin of Polynesian Y chromosomes.** Curr Biol, v. 10, p. 1237-1246, 2000a.

KAYSER, M. *et al.* Characteristics and frequency of germline mutations at microsatellite *loci* from the human Y chromosome, as revealed by direct observation in father/son pairs. Am J Hum Genet, v. 66, n. 5, p. 1580-1588, 2000b.

KAYSER, M. et al. Reduced Y-Chromosome, but Not Mitochondrial DNA, Diversity in Human Populations from West New Guinea. Am J Hum Genet, v. 72, p. 281-302, 2003.

KAYSER, M. *et al.* **A comprehensive survey of human Y-chromosomal microsatellites.** Am J Hum Genet, v. 74, n. 6, p. 1183-1197, 2004.

KAYSER, M. The human genetic history of Oceania: near and remote views of dispersal. Curr Biol, v. 20, p. 194-201, 2010.

KAYSER, M.; SAJANTILA, A. Mutations at Y-STR *loci*: implications for paternity testing and forensic analysis. Forensic Sci Int, v. 118, p. 116-121, 2001.

KIM, S. H. *et al.* Population genetics and mutational events at 6 Y-STRs in Korean population. Forensic Sci Int Genet, v. 3, n. 2, p. e53-4, 2009.

KIMURA, A. E; OHTA, T. Stepwise mutation model and distribution of allelic frequencies in a finite population. Pro Natl Acad Scis USA, v. 75, p. 2868-2872, 1978.

KIMURA, L. *et al.* **Inferring paternal history of rural African-derived Brazilian populations from Y chromosomes**. Am J Hum Biol, v. 29, n. 2, p. e22930, 2017.

KING, T. E. *et al.* **Thomas Jefferson's Y chromosome belongs to a rare European lineage**. Am J Phys Anthropol, v. 132, p. 583-589, 2007.

KRUGLYAK, S. *et al.* Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. Proc Natl Acad Sci, v. 95, p. 10774-10778, 1998.

KOFI, A. B. *et al.* **Population data of 23 Y chromosome STR** *loci* for the five major human subpopulations of Ghana. Int J Legal Med, v. 134, p. 1313-1315, 2020.

KURIHARA, R. *et al.* **Mutations in 14 Y-STR** *loci* **among Japanese father-son haplotypes.** Int J Legal Med, v. 118, n. 3, p. 125-131, 2004.

KUMAGAI, R. *et al.* **Haplotype analysis of 17 Y-STR** *loci* in a Japanese population. Forensic Sci Int, v. 172, n. 1, p. 72-78, 2007.

LACAU, H. *et al.* **Afghanistan from a Y-chromosome perspective**. Eur J Hum Genet, v. 20, p. 1063-1070, 2012.

LAMBERT, P. **The Paraguay Reader: History, Culture, Politics.** 1 ed. Duke University Press Books, 488 p., 2012.

LARMUSEAU, M. H. *et al.* **The Paternal Landscape along the Bight of Benin- Testing Regional Representativeness of WestAfrican Population Samples Using Y-Chromosomal Markers**. PLoS ONE, v. 10, n. 11, p. e0141510, 2015.

LANDER, E. S. *et al.* **Initial sequencing and analysis of the human genome.** Nat, v. 409, n. 6822, p. 860-921, 2001.

LAOUINA, A. et al. Mutation rate at 17 Y-STR loci in 'Father/Son' pairs from moroccan population. Leg Med, v. 15, n. 5, p. 269-271, 2013.

LAPPALAINEN, T. *et al.* **Genomic Analysis in the Age of Human Genome Sequencing**. Cell, v. 177, n. 1, p. 70-84, 2019.

- LARDONE, M. C. *et al.* Greater prevalence of Y chromosome Q1a3a haplogroup in Y-microdeleted Chilean men: a case-control study. J Assist Reprod Genet, v. 30, n. 4, p. 531-538, 2013.
- LEE, H. Y. et al. Haplotypes and mutation analysis of 22 Y-chromosomal STRs in Korean father-son pairs. Int J Legal Med, v. 121, p. 128-135, 2007.
- LESSIG, R.; EDELMANN, J. Y chromosome polymorphisms and haplotypes in west Saxony (Germany). Int J Legal Med, v. 111, n. 4, p. 215-218, 1998.
- LIN, H. *et al.* Analyzing genetic polymorphism and mutation of 44 Y-STRs in a Chinese Han population of Southern China. Leg Med (Tokyo), v. 42, p. 101643, 2020.
- LIU, H. *et al.* Chromatin structure is distinct between coding and non-coding single nucleotide polymorphisms. BMC Mol Biol, v. 15, n. 22, Não paginado, 2014.
- LIU, Y. Genetic Diversity and Disease Susceptibility. 1 ed. Intechopen, 164 p., 2018.
- LLAMAS, B. *et al.* Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. Sci Adv, v. 2, p. e1501385, 2016.
- LÓPEZ, S.; VAN DORP, L.; HELLENTHAL, G. **Human Dispersal Out of Africa: A Lasting Debate**. Evol Bioinform Online, v. 11, n. 2, p. 57-68, 2015.
- LÓPEZ-PARRA, A. M. *et al.* In search of the Pre and Post Neolithic Genetic Substrates in Iberia: Evidence from Y Chromosome in Pyrenean Populations. Ann Hum Genet, v. 73, p. 42-53, 2009.
- LUIZ, M. T.; SCHILLAT, M. Tierra del Fuego, Materiales para el estudio de la Historia Regional. L.O.L.A. 1 ed. Literatura da América Latina, 3335 p., 1998.
- LUIS, J. R. *et al.* The Levant versus the Horn of Africa: evidence for bidirectional corridors of human migrations. Am J Hum Genet, v. 74, p. 532-544, 2004.
- MALYARCHUK, B. *et al.* Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup Q1a. J Hum Genet, v. 56, p. 583-588, 2011.
- MANGS, A. H.; MORRIS, B. J. The Human Pseudoautosomal Region (PAR): Origin, Function and Futur. Curr Genomic, v. 8, n. 2, p. 129-36, 2007.
- MANTA, F. S. *et al.* Analysis of genetic ancestry in the admixed Brazilian population from Rio de Janeiro using 46 autosomal ancestry-informative indel markers. Ann Hum Biol, v. 40, p. 94-98, 2013.
- MARTIN, P. S. **Discovery of America**. Sci, v. 179, p. 969-974, 1973.
- MARTINS, A. M. **Análise genético-histórica de haplótipos do cromossomo Y humano no nordeste brasileiro**. 2007. Dissertação (Mestrado em Genética). Universidade Federal de Pernambuco, Pernambuco, Brasil, 2007. Não paginado.

MARTINEZ, B. *et al.* Searching for the roots of the first free African American community. Sci Rep, v. 10, p. 20634, 2020.

MARTINEZ-ESPIN, E. et al. Paraguayan population data on the fifteen STR loci included in the PowerPlex 16 kit. J Forensic Sci, v. 48, p. 253-255, 2003.

MARTINIANO, R. et al. Y-chromosome diversity in central Portugal reveals signatures of ancient maritime expansions. Anthrop Anz, v. 70, n. 4, p. 355-367, 2013.

MATEO, M. B. La tierra de los fuegos: historia, geografia, sociedad, economia. 1 ed. La Municipalidad de Porvenir Tierra Del Fuego, p. 221, 1998.

MATSUMURA, H. *et al.* Craniometrics Reveal "Two Layers" of Prehistoric Human Dispersal in Eastern Eurasia. Sci Rep, v. 9, n. 1, p. 1451, 2019.

MENDEZ, F. L. *et al.* Increased Resolution of Y Chromosome Haplogroup T Defines Relationships among Populations of the Near East, Europe, and Africa. Hum Biol, v. 83, n. 1, p. 39-53, 2011.

MIDDLETON, D. al. Analysis of the distribution of HLA-A alleles in populations from five continents. Hum Immuno, v. 161, p. 1048-1052, 2000.

MIZUNO, N. et al. A forensic method for the simultaneous analysis of biallelic markers identifying Y chromosome haplogroups inferred as having originated in Asia and the Japanese archipelago. Forensic Sci Int Genet, v. 4, p. 73-79, 2010.

MEIRELES, M. História do Maranhão. 5 ed. São Luís: Edições AML, 2015. Não paginado.

MERTOGLU, E. *et al.* Estimation of the Y-chromosomal short tandem repeat (Y-STR) mutation rates in Turkey. Turkish J Biochem, v. 43, n. 2, p. 142-150, 2018.

MEZZAVILLA, M. et al. Insights into the origin of rare haplogroup C3* Y chromosomes in South America from high-density autosomal SNP genotyping. Forensic Sci Int Genet, v. 15, p. 115-120, 2015.

MONTANO, V. et al. The Bantu expansion revisited: a new analysis of Y chromosome variation in Central Western Africa. Molecular Ecology, v. 20, n. 13, 2693-2708, 2011.

MOORE, L.T. *et al.* **A Y-Chromosome Signature of Hegemony in Gaelic Ireland**. Sou J Hum Genet, v. 78, p. 334-338, 2006.

MORALES, L. A. *et al.* Paternal portrait of populations of the middle Magdalena River region (Tolima and Huila, Colombia): New insights on the peopling of Central America and northernmost South America. PLoS One, v. 13, n. 11, p. e0207130, 2018.

MORELLI, L. *et al.* Comparison of Y-Chromosome Variation in Sardinia and Anatolia Is More Consistent with Cultural Rather than Demic Diffusion of Agriculture. PLoS One, v. 5, n. 4, p. e10419, 2010.

MORENO-MAYAR, J. V. *et al.* Early human dispersals within the Americas. Sci, v. 362, n. 6419, p. eaav2621, 2018.

MORTON, N. E. **Parameters of the human genome**. Proc Natl Acad Sci USA, v. 88, n. 17, p. 7474-7476, 1991.

MYRES, N. M. *et al.* **Y-chromosome short tandem repeat DYS458.2 non-consensus alleles occur independently in both binary haplogroups J1-M267 and R1b3-M405**. Croat Med J, v. 48, n. 4, p. 450-9, 2007.

MYRES, N. M. *et al.* Y-chromosome short tandem repeat intermediate variant alleles DYS392.2, DYS449.2, and DYS385.2 delineate new phylogenetic substructure in human Y-chromosome haplogroup tree. Croat Med J, v. 50, n. 3, p. 239-49, 2009.

MYRES, N. M. et al. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. Eur J Hum Genet, v. 19, p. 95-101, 2011.

MUZZIO, M. *et al.* **Software for Y-haplogroup predictions: A word of caution**. Int J Legal Med, v. 125, n. 1, p. 143-147, 2011.

NASIDZE, I. *et al.* **Genetic evidence for the Mongolian ancestry of Kalmyks.** Am J Phys Anthropol, v. 128, p. 846-854, 2005.

NOGUEIRO, I. *et al.* **Phylogeographic Analysis of Paternal Lineages in NE Portuguese Jewish Communities**. Am J Phys Anthropol, v. 141, p. 373–381, 2010.

NEVES, W. A.; PUCCIARELLI, H. M. Morphological affinities of the first Americans an exploratory analysis based on early South American human remains. J Hum Evol, v. 21, p. 261-273, 1991.

NOGUERA, M. C. *et al.* Colombia's racial crucible: Y chromosome evidence from six admixed communities in the Department of Bolivar. Ann Hum Biol, v. 41, p. 453-459, 2013.

OLIVEIRA, A. M. *et al.* Male Lineage Strata of Brazilian Population Disclosed by the Simultaneous Analysis of STRs and SNPs. Forensic Sci Int Genet, v. 13, p. 264-268, 2014.

OSSA, H. *et al.* **Outlining the Ancestry Landscape of Colombian Admixed Populations.** PLoS ONE, v. 11, n. 10, p. e0164414, 2016.

PADILLA-GUTIÉRREZ, J. R. *et al.* Population data and mutation rate of nine Y-STRs in a mestizo Mexican population from Guadalajara, Jalisco, México. Leg Med (Tokyo), v. 10, n. 6, p. 319-320, 2008.

PALHA, T. *et al.* Disclosing the Genetic Structure of Brazil through Analysis of Male Lineages with Highly Discriminating Haplotypes. PLoS ONE, v. 7, n. 7, p. e40007, 2012.

PAROLIN, M. L. *et al.* **Population genetic analysis of 23 Y-STR** *loci* in Central **Argentine Patagonia.** Int J Legal Med, v. 133, n. 3, p. 777-779, 2019.

- PAURO, M. *et al.* **Analysis of Uniparental Lineages in Two Villages of Santiago del Estero, Argentina, Seat of Pueblos de Indios in Colonial Times.** Hum Biol, v. 85, n. 5, p. 699-719, 2013.
- PLANZ, J. V.; HALL, T. A. Hidden Variation in Microsatellite *Loci*: Utility and Implications for Forensic DNA. Forensic Sci Rev, v. 24, p. 27-42, 2012.
- PELLEGRINO, A. La migración internacional en América Latina. Tendencias y per les de los migrantes. 1 ed. Chile: Conferencia Hemisférica sobre Migración Internacional, Derechos Humanos y Trata de Personas, 36 p., 2003.
- PEEKE, C. **Preliminary grammar of Auca**. Oklahoma: The Summer Institute of Linguistics-Norman, 145 p., 1973.
- PENA, S. D. J. *et al.* **DNA tests probe the genomic ancestry of Brazilians**. BrazcJ Med Biol Research, v. 42, n. 10, p. 870-876, 2009.
- PENA, S. D. *et al.* The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. PLoS One, v. 6, p. e17063, 2011.
- PENA, S. D.; SANTOS, F. R.; TARAZONA-SANTOS, E. Genetic admixture in Brazil. Am J Med Genet C Semin Med Genet, v. 184, n. 4, p. 928-938, 2020.
- PEREIRA, R. W. et al. Haplotype diversity of 17 Y-chromosome STRs in Brazilians. Forensic Sci Int, v. 171, p. 226-236, 2007.
- PEREIRA, V; GUSMÃO, L. Análise de Marcadores STRs de Cromossoma Y. *In:* MÁRQUEZ, M. C.; CABALLERO, P. B. Genética forense: Del laboratorio a los tribunales, 2019, 1 ed. Ediciones Díaz de Santos, S.A., p. 542. Capítulo 9. 2019.
- PEREZ-BENEDICO, D. *et al.* **Mayans: a Y chromosome perspective.** Eur J Hum Genet, v. 24, p. 1352-1358, 2016.
- PERICIĆ, M. *et al.* Review of Croatian genetic heritage as revealed by mitochondrial **DNA** and **Y** chromosomal lineages. Croat Med J, v. 46, p. 502-513, 2005a.
- PERICIĆ, M. *et al.* **High-resolution phylogenetic analysis of southeastern Europe traces major episodes of paternal gene flow among Slavic populations**. Mol Biol Evol, v. 22, p. 1964-1975, 2005b.
- PETROVIC, V. M. *et al.* Assessment of mutation rates for PPY23 Y chromosome STR *loci* in Serbian father-son pairs. Forensic Sci Int Genet, v. 39, p. e5-e9, 2019.
- PINOTTI, T. *et al.* Y Chromosome Sequences Reveal a Short Beringian Standstill, Rapid Expansion, and early Population structure of Native American Founders. CurrBiol, v. 29, p. 149-157, 2019.
- PINTO, N.; GUSMÃO, L.; AMORIM, A. Mutation and mutation rates at Y chromosome specific Short Tandem Repeat Polymorphisms (STRs): A reappraisal. Forensic Sci Int Genet, v. 9, p. 20-24, 2014.

PINTO, N. et al. Theory and statistics of mutation rates: A mathematical framework reformulation for forensic applications. Forensic Sci Int Genet, v. 5, p. e131-e132, 2015.

PONTES, M. L. *et al.* **Allele frequencies and population data for 17 Y-STR** *loci* **(AmpFlSTR Y-filer) in a Northern Portuguese population sample.** Forensic Sci Int, v. 170, n. 1, p. 62-67, 2007.

POTTTHAST-JUTKEIT, B. The creation of the "mestizo family model": The example of **Paraguay**. History of the Family, v. 2, n. 2, p. 123-139, 1997.

POWERPLEX® Y23 SYSTEM FOR USE ON THE APPLIED BIOSYSTEMS® GENETIC ANALYZERS. Disponível em: http://www.promega.com.br/products/genetic-identity/geneticidentity-workflow/str-amplification/powerplex-y23system/?catNum=DC2305/. Acesso em: 18 agosto de. 2018.

PURPS, J. et al. A global analysis of Y-chromosomal haplotype diversity for 23 STR *loci*. Forensic Sci Int Genet, v. 12, p. 12-23, 2014.

QUINTANA-MURCI, L.; KRAUSZ, C.; MCELREAVEY, K. The human Y chromosome: function, evolution and disease. Forensic Sci Int, v. 118, p. 169-181, 2001.

RAGHAVAN, M. et al. Genomic evidence for the Pleistocene and recent population history of native Americans. Sci. v. 349, p. aab3884, 2015.

RALF, A. *et al.* **Identification and characterization of novel rapidly mutating Y-chromosomal short tandem repeat markers**. Hum Mutat, v. 41, n. 9, p. 1680-1696, 2020.

RAMACHANDRAN, S. *et al.* **Genetics and Genomics of Human Population Structure**. Vogel and Motulsky's Human Genetics: Problems and Approaches, p. 589-615, 2010.

RASMUSSEN, M. et al. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. Nat, v. 506, p. 225-229, 2014.

REDDY, B. M. et al. Austro-Asiatic tribes of Northeast India provide hitherto missing genetic link between South and Southeast Asia. PLoS ONE, v. 2, p. e1141, 2007.

RÊGO, J. J. D. do. **Marcadores do cromossomo Y no estudo de linhagens paternas da população brasileira**. 2019. 140 f. Tese (Doutorado em Biociências) - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil, 2019.

REGUEIRO, M. *et al.* **Iran: tricontinental nexus for Y-chromosome driven migration.** Hum Hered, v. 61, p. 132-143, 2006.

REGUEIRO, M. *et al.* On the Origins, Rapid Expansion and Genetic Diversity of Native Americans From Hunting-Gatherers to Agriculturalists. Am J Phys Anthropol, v. 150, n. 3, p. 333-48, 2013.

REGUEIRO, M. *et al.* From Arabia to Iberia: A Y chromosome perspective. Gene, v. 564, p. 141-152, 2015.

REICHEL-DOLMATOFF, G. Contactos y câmbios culturales em la Sierre Nevada de Santa Marta. Rev Colomb Antropologia, v. 1, 121 p., 1953.

REID, K. M.; HEATHFIELD., L. M. Allele frequency data for 23 Y-chromosome short tandem repeats (STRs) for the South African population. Forensic Sci Int Genet, v. 46, p.102270, 2020.

RESTREPO, V. Los chibchas antes de la conquista española. Editorial Bogotá: Imp. de La Luz, p. 238, 1895.

RESQUE, R. *et al.* Male Lineages in Brazil: Intercontinental Admixture and Stratification of the European Background. PLoS One, v. 11, n. 4, p. e0152573, 2016.

REYES-CENTENO, H. *et al.* **Genomic and cranial phenotype data support multiple modern human dispersals from Africa and a southern route into Asia**. Proc Natl Acad Sci USA, v. 111, n. 20, p. 7248-7253, 2014.

REYES-MADRID, M. et al. An approach on the migratory processes in the north of Chile based on Y chromosome analysis. Am J Hum Biol, v. 34, n. 7, p. e23736, 2022.

RHIE, A. et al. The complete sequence of a human Y chromosome. Nat, v. 621, n. 7978, p. 344-354, 2023.

RIBEIRO, D. **O povo brasileiro: a formação e o sentido do Brasil.** 2 ed. São Paulo: Companhia das Letras, 480 p., 1995.

RIBEIRO, J. et al. Analysis of 23 Y-STRs in a population sample from eastern Paraguay. Forensic Sci Int Genet, v. 37, p. e20-e22, 2018.

ROBERT, F.; PELLETIER, J. Exploring the Impact of Single-Nucleotide Polymorphisms on Translation. Front Genet, v. 9, p. 507, 2018.

ROBINO, C. *et al.* **Development of an Italian RM Y-STR haplotype database: Results of the 2013 GEFI collaborative exercise.** Forensic Sci Int Genet, v. 15, p. 56-63, 2015.

RODRIGUEZ, J. P. The Historical Encyclopedia of World Slavery. 1 ed. ABC-CLIO, 805 p., 1997.

RODRÍGUEZ, M. E.; HORLENT, L. **Tehuelches y selk'nam (Santa Cruz y Tierra del Fuego): no desaparecimos.** 1 ed. Buenos Aires: Ministerio de Educación y Deportes, 42 p., 2016.

ROEWER, L. et al. Continent wide decoupling of Y-chromosomal genetic variation from language and geography in native South Americans. PLoS Genet, v. 9, p. e1003460, 2013.

ROITMAN, K. **Hybridity, mestizaje, and Montubios in Ecuador.** QEH Work Pap Ser, v. QEWPS165, 20 p., 2008.

ROJAS, W. et al. Genetic make up and structure of Colombian populations by means of uniparental and biparental DNA markers. Am J Phys Anthropol, v. 143, p. 13-20, 2010.

ROMERO, R. E. *et al.* A Colombian Caribbean population study of 16 Y-chromosome STR *loci*. Forensic Sci Int Genet, v. 2, p. e5–e8, 2008.

ROSSER, Z. H. *et al.* Y-chromosomal diversity within Europe is clinal and influenced primarily by geography rather than language. Am J Hum Genet, v. 67, p. 1526-1543, 2000.

ROOTSI, S. *et al.* Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in europe. Am J Hum Genet, v. 75, n. 1, p. 128-137, 2004.

ROTHHAMMER, F.; DILLEHAY, T. D. The late Pleistocene colonization of South America: an interdisciplinary perspective. Ann Hum Genet, v. 73, p. 540-549, 2009.

ROZEN, S. *et al.* **Abundant gene conversion between arms of palindromes in human and ape Y chromosome**. Nat, v. 423, n. 6942, p. 873-876, 2003.

ROZHANSKII, I. L.; KLYOSOV, A. A. Haplogroup R1a, its Subclades and Branches in Europe during the Last 9000 Years. Advances in Anthropology, v. 2, p. 139-156, 2012.

RUIZ-LINARES, A. *et al.* Admixture in Latin America: Geographic Structure, Phenotypic Diversity and Self-Perception of Ancestry Based on 7.342 Individuals. PLoS Gene, v. 10, n. 9, p. e1004572, 2014.

SAIZ, M. *et al.* Genetic structure in the paternal lineages of South East Spain revealed by the analysis of 17 Y-STRs. Sci Rep, v. 9, n. 1, p. 5234, 2019.

SALA, A. *et al.* Genetic Analysis of Six Communities of MbyáGuaraní Inhabiting Northeastern Argentina by Means of Nuclear and Mitochondrial Polymorphic Markers. Hum Biol, v. 82, n. 4, p. 433-456, 2010.

SALA, A.; CORACH, D. Analysis of admixture and genetic structure of two Native American groups of Southern Argentinean Patagonia. Mol Biol Rep, v. 41, p. 1533-1543, 2014.

SALAS, A. *et al.* **Gender bias in the multiethnic genetic composition of central Argentina**. J Hum Genet, v. 53, n. 7, p. 662-674, 2008.

SALAS, A. *et al.* Mitochondrial echoes of first settlement and genetic continuity in El Salvador. PLoS One, v. 2, p. e6882, 2009.

SALA, A. *et al.* Historical records under the genetic evidence: "Chiriguano" tribe genesis as a test case. Mol Biol Rep., v. 45, p. 987-1000, 2018.

SALZANO, F. M.; FREIRE-MAIA, N. Populações brasileiras, aspectos demográficos, genéticos e antropológicos. 1 ed. São Paulo: Companhia Editora Nacional, 177 p., 1967.

- SALZANO, F. M.; FREIRE-MAIA, N. **Problems in the human biology:** A study of Brazilian population. Wayne State University Press, Detroit, 203 p., 1970.
- SALZANO, F. M.; CALLEGARI-JACQUES, S. M. South American Indians: A case study in evolution. 1 ed. Oxford: Clarendon Press, 270 p., 1988.
- SALZANO, F. M.; SANS, M. Interethnic admixture and the evolution of Latin American populations. Genet Mol Biol, v. 37, n. 1, p. 51-170, 2014.
- SANCHEZ, J. et al. Y Chromosome analysis of the Somali population suggests the origin of the haplogroups E3b1. *In*: International Symposium: DNA Polymorphisms in Human Populations. Resumos. Paris, França, p. 5-6, 2003.
- SÁNCHEZ-DIZ, P. et al. Population and segregation data on 17 Y-STRs: results of a GEP-ISFG collaborative study. Int J Legal Med, v. 122, p. 529-533, 2008.
- SANCHEZ-FADDEV, J. *et al.* The influence of clan structure on the genetic variation in a single Ghanaian village. Eur J Hum Genet, v. 21, n. 10, p. 1134-1139, 2013.
- SANS, M. Admixture studies in Latin America: From the 20th to the 21st century. Hum Biol, v. 72, p. 155-177, 2000.
- SANTANGELO, R. et al. Analysis of ancestry informative markers in three main ethnic groups from Ecuador supports a trihybrid origin of Ecuadorians. Forensic Sci Int Genet, v. 31, p. 29-33, 2017.
- SANTOS, F. R. *et al.* A polymorphic L1 retroposon insertion in the centromere of the human Y chromosome. Hum Mol Genet, v. 9, p. 421-430, 2000.
- SANTOS, N. P. C. *et al.* Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INDEL) ancestry-informative marker (AIM) panel. Hum Mutat, v. 31, p. 184-190, 2010.
- SCHANN, A. P. *et al.* New insights on intercontinental origins of paternal lineages in Northeast Brazil. BMC Evol Biol, v. 20, n. 1, p. 15, 2020.
- SCHURR, T. G.; SHERRY, S. T. **Mitochondrial DNA and Y chromosome diversity and the peopling of the Americas: Evolutionary and demographic evidence**. Am J Hum Biol, v. 16, n. 4, p. 420-439, 2004.
- SCHNEIDER, S.; ROESSLI, D.; EXCOFFIER, *L.* Arlequin: A software for population genetics data analysis. User manual version. Genetics and Biometry Laboratory, University of Geneva, Switzerland, 2000. Não paginado.
- SCHMITT, R. *et al.* Extremely limited mitochondrial DNA variability among the Aché Natives of Paraguay. Ann Hum Biol, v. 31, p. 87-94, 2004.
- SCHOLOTTERER, C. Evolutionary dynamics of microsatellite DNA. Chomosoma, v. 6, n. 109, p. 365-71, 2000.

- SEIELSTAD, M. et al. A Novel Y-Chromosome Variant Puts an Upper Limit on the Timing of First Entry into the Americas. Am J Hum Genet, v. 73, n. 3, p. 700-705, 2003.
- SEMINO, O. *et al.* The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. Sci. v. 290, p. 1155-1159, 2000.
- SEMINO, O. *et al.* Ethiopians and Khoisan share the deepest clades of the human Y-chromosome phylogeny. Am J Hum Genet, v. 70, p. 265-268, 2002.
- SEMINO, O. *et al.* Origin, Diffusion, and Differentiation of Y-Chromosome Haplogroups E and J: Inferences on the Neolithization of Europe and Later Migratory Events in the Mediterranean Area. Am J Hum Genet, v. 74, p. 1023-1034, 2004.
- SENGUPTA, S. *et al.* Polarity and temporality of high resolution Y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. Am J Hum Genet, v. 78, p. 202-221, 2006.
- sen SHI, M. *et al.* Haplotypes of 20 Y-chromosomal STRs in a population sample from southeast China (Chaoshan area). Int J Legal Med, v. 121, n. 6, p. 455-462, 2007.
- SEVINI, F. *et al.* Analysis of Population Substructure in Two Sympatric Populations of Gran Chaco, Argentina. PLoS One, v. 8, n. 5, p. e64054, 2013.
- SHI, H. et al. Y-Chromosome Evidence of Southern Origin of the East Asian-Specific Haplogroup O3-M122. Am J Hum Genet, v. 77, p. 408-419, 2005.
- SILVA, M. A. C. *et al.* Genomic insight into the origins and dispersal of the Brazilian coastal natives. Proc Natl Acad Sci USA, v. 117, n. 5, p. 2372-2377, 2020.
- SIMÃO, F. *et al.* The maternal inheritance of Alto Paraná revealed by full mitogenome sequences. Forensic Sci Int Genet, v. 39, p. 66-72, 2019.
- SIMÃO, F. The Ancestry of Eastern Paraguay: A Typical South American Profile with a Unique Pattern of Admixture. Genes, v. 12, n. 11, p. 1788, 2021.
- SIMS, L. M.; GARVEY, D.; BALLANTYNE, J. Sub-populations within the major European and African derived haplogroups R1b3 and E3a are differentiated by previously phylogenetically undefined Y-SNPs. Human Mutation, v. 28, n. 1, p. 97-97, 2007.
- SIMS, L. M. *et al.* Improved resolution haplogroup G phylogeny in the Y chromosome, revealed by a set of newly characterized SNPs. PLoS ONE, v. 4, p. e5792, 2009.
- SIDENPE. **Sistema Integrado de Indicadores Sociales del Ecuador –SIISE**, 2003 Disponível em: < http://www.siise.gov.ec/Principal.aspx>. Acesso em: 10 junho. 2021.
- SINGH, N. P. *et al.* Epigenetic profile of the euchromatic region of human Y chromosome. Nucleic Acids Res, v. 39, n. 9, p. 3594-3606, 2011.

SILVERMAN, H.; ISBELL, W. E. **Handbook of South American Archaeology.** 1 ed. New York: Springer, 1192 p., 2008.

SKALETSKY, H. *et al.* The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nat, v. 423, n. 6942, p. 825-837, 2003.

SOARES-VIEIRA, J. A. *et al.* **Population and mutation analysis of Y-STR** *loci* in a **sample from the city of São Paulo (Brazil).** Genet Mol Biol, v. 31, n. 3, p. 651-656, 2008.

SOLÉ-MORATA, N. et al. Analysis of the R1b-DF27 haplogroup shows that a large fraction of Iberian Y-chromosome lineages originated recently in situ. Sci Rep, v. 7, p. 7341, 2017.

SOTO, B. B. **Historia de la religión selk'nam de Tierra del Fuego.** 1 ed. Grupo Editorial Círculo Rojo SL, 176 p., 2019.

STANGE, V. S. *et al.* **Stratification among European descent and admixed Brazilian populations of Espírito Santo for 27 Y-STRs**. Forensic Sci Int Genet, v. 41, p. e20-e22, 2019.

SU, B. *et al.* **Y chromosome haplotypes reveal prehistorical migrations to the Himalayas**. Hum Genet, v. 107, p. 582-590, 2000.

SUN, J. X. *et al.* A direct characterization of human mutation based on microsatellites. Nat Genet, v. 44, p. 1161–1165, 2012.

TABOADA-ECHALAR, P. *et al.* The Genetic Legacy of the Pre-Colonial Period in Contemporary Bolivians. PLoS One, v. 8, n. 3, p. e58980, 2013.

TAMBURRINI, C. *et al.* **Uniparental origins of the admixed Argentine Patagonia.** Am J Hum Biol, v. 34, p. e23682, 2021.

THE 1000 GENOMES PROJECT CONSORTIUM. A global reference for human genetic variation. Nat, v. 526, p. 68-74, 2015.

TOFANELLI, S. *et al.* **J1-M267 Y lineage marks climate-driven pre-historical human displacements.** Eur J Hum Genet, v. 17, p. 1520-1524, 2009.

TOSCANINI, U. *et al.* Genetic variability of 17 Y chromosome STRs in two Native American populations from Argentina. Int Congr., v. 1288, p. 154-155, 2006.

TOSCANINI, U. et al. Testing for genetic structure in different urban Argentinian populations. Forensic Sci Int, v. 165, p. 35-40, 2007.

TOSCANINI, U. *et al.* Y chromosome microsatellite genetic variation in two Native American populations from Argentina: Population stratification and mutation data. Forensic Sci Int Genet, v. 2, n. 4, p. 274-80, 2008.

TOSCANINI, U. *et al.* Male Lineages in South American Native Groups: Evidence of M19 Traveling South. Am J Phys Anthropol, v. 146, p. 188-196, 2011.

TOSCANINI, U. *et al.* A comprehensive Y-STR portrait of Argentinean populations. Forensic Sci Int Genet, v. 20, p. 1-5, 2016.

TOSCANINI, U. *et al.* **The geographic mosaic of Ecuadorian Y-chromosome ancestry**. Forensic Sci Int Genet. v. 33, p. 59-65, 2018.

TORRES, J. B. *et al.* Genetic diversity in the Lesser Antilles and its implications for the settlement of the Caribbean basin. PLoS One, v. 10, p. 1-27, 2015.

THOMPSON, J. M. *et al.* **Developmental validation of the PowerPlex Y23 System: A single multiplex Y-STR analysis system for casework and database samples.** Forensic Sci Int Genet, v. 7, p. 240-250, 2012.

TREJAUT, J. A. et al. Taiwan Y-chromosomal DNA variation and its relationship with Island Southeast Asia. BMC Genet, v.15, p. 77, 2014.

TRIKI-FENDRI, S. *et al.* **Genetic Structure of the Kuwaiti Population Revealed by Paternal Lineages**. Am J Hum Biol, v. 28, p. 203-212, 2016.

TROMBETTA, B. *et al.* A New Topology of the Human Y Chromosome Haplogroup E1b1 (E-P2) Revealed through the Use of Newly Characterized Binary Polymorphisms. PLoS ONE, v. 6, n. 1, p. e16073, 2011.

TSAI, L.C. *et al.* Haplotype frequencies of nine Y-chromosome STR *loci* in the Taiwanese Han population. Int J Legal Med, v. 116, n. 3, p. 179-183, 2002.

TURRINA, S.; ATZEI, R.; de LEO, D. **Y-chromosomal STR haplotypes in a Northeast Italian population sample using 17plex** *loci* **PCR assay**. Int J Legal Med, v. 120, n. 1, p. 56-59, 2006.

TURRINA, S. *et al.* Haplotype data and mutation rates for the 23 Y-STR *loci* of **PowerPlex®** Y 23 System in a Northeast Italian population sample. Int J Legal Med, v. 129, n. 4, p. 725-8, 2015a.

TURRINA, S. *et al.* **Deletion and duplication at DYS448 and DYS626 loci: unexpected patterns within the AZFc region of the Y-chromosome.** Int J Legal Med, v. 129, n. 3, p. 449-55, 2015b.

UNDERHILL, P. A. *et al.* The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. Ann Hum Genet, v. 65, p. 43-62, 2001.

van OVEN, M. *et al.* An efficient multiplex genotyping approach for detecting the major worldwide human Y-chromosome haplogroups. Int J Legal Med, v. 125, p. 879-885, 2011.

van OVEN, M. et al. Seeing the wood for the trees: a minimal reference phylogeny for the human Y chromosome. Hum Mutat, v. 35, n. 2, p. 187-191, 2014.

VALVERDE, L. et al. 17 Y-STR haplotype data for a population sample of Residents in the Basque Country. Forensic Sci Int Genet, v. 6, n. 4, p. e109-11, 2012.

VALVERDE, V. *et al.* New clues to the evolutionary history of the main European paternal lineage M269: dissection of the Y-SNP S116 in Atlantic Europe and Iberia. Eur J of Hum Genet, p. 1-5, 2015.

VÁSQUEZ, S. L.; SALTOS GALARZA, N. **Equador: su realidad.** 20 ed. Fundação 'José Peralta', 375 p., 2014.

VÁZQUEZ, F. **Territorio y población: nuevas dinámicas regionales en el Paraguay**. 1 ed. Asociación Paraguaya de Estudios de Población (ADEPO), 201 p., 2006.

VEERAPPA, A. M. et al. Copy number variation-based polymorphism in a new pseudoautosomal region 3 (PAR3) of a human X-chromosome-transposed region (XTR) in the Y chromosome. Funct Integr Genomics, v. 13, p. 285-293, 2013.

VICARD, P.; DAWID, A. P. A statistical treatment of biases affecting the estimation of mutation rates. Mutat Res, v. 547, p. 19-33, 2004.

VIEIRA-SILVA, A. *et al.* **Y-STR mutational rates determination in South Portugal Caucasian population.** Forensic Sci Int Genet Suppl Ser, v. 2, n. 1, 60-61, 2009.

VILA, A. *et al.* Investigaciones etnoarqueológicas en tierra del Fuego (1986-2006): reflexiones para la arqueología Prehistórica europea. Trabajos de Prehistoria, v. 64, n. 2, p. 37-53, 2007.

VILLAESCUSA, P. et al. Characterization of the Iberian Y chromosome haplogroup R-DF27 in Northern Spain. Forensic Sci Int Genet, v. 27, p. 142-148, 2017.

VILLAESCUSA, P. et al. A Y-chromosomal survey of Ecuador's multi-ethnic population reveals new insights into the tri-partite population structure and supports an early Holocene age of the rare Native American founder lineage C3-MPB373. Forensic Sci Int Genet, v. 51, p. 102427, 2021.

VOGT, P. H. *et al.* **Human Y Chromosome Azoospermia Factors (AZF) mapped to different subregions in Yq11.** Hum Mol Genet, v. 5, p. 933-943, 1996.

VOSKARIDES, K. et al. Y-chromosome phylogeographic analysis of the Greek-Cypriot population reveals elements consistent with Neolithic and Bronze Age settlements. Investig Genet, v. 7, p. 1, 2016.

VULLO, C. *et al.* **Association between Y haplogroups and autosomal AIMs reveals intra-population substructure in Bolivian populations.** Int J Legal Med, v. 129, n. 4, p. 673-680, 2014.

VULLO, C. *et al.* **Genetic polymorphism of 22 autosomal STR markers in Paraguay**. Forensic Sci Int Genet, v. 25, p. e16-e17, 2016.

WALSH, B. Estimating the time to the most recent common ancestor for the Y chromosome or mitochondrial DNA for a pair of individuals. Genetics, v. 158, n. 897-912, 2001.

- WALSH, P. S.; METZGER, D. A.; HIGUCHI, R. Chelex 100 as a Medium for Simple Extraction of DNA for PCR-Based Typing from Forensic Material. Bio Techniques, v. 10, p. 506-513, 1991.
- WATERS, M. R. *et al.* **Pre-Clovis mastodon hunting 13,800 years ago at the Manis site, Washington.** Sci, v. 334, p. 351-353, 2011.
- WANG, C. C. et al. Late Neolithic expansion of ancient Chinese revealed by Y chromosome haplogroup O3a1c-002611. Int J Syst Evol, v. 51, p. 280-286, 2013.
- WANG, C. C.; LI, H. Inferring human history in East Asia from Y chromosomes. Investig Genet, v. 4, n. 1, p. 11, 2013.
- WANG, C. C. *et al.* Convergence of Y Chromosome STR Haplotypes from Different SNP Haplogroups Compromises Accuracy of Haplogroup Prediction. J Genet Genomics, v. 42, p. 403-407, 2015a.
- WANG, Y. et al. Analysis of 17 Y-STR loci haplotype and Y-chromosome haplogroup distribution in five Chinese ethnic groups. Electrophoresis, v. 36, n. 20, p. 2546-2552, 2015b.
- WANG, Y. et al. Genetic polymorphisms and mutation rates of 27 Y-chromosomal STRs in a Han population from Guangdong Province, Southern China. Forensic Sci Int Genet, v. 21, p. 5-9, 2016.
- WARREN, H. G. Rebirth of the Paraguayan Republic: The First Colorado Era, 1878-1904. 1 ed. University of Pittsburgh Press, 400 p., 1985.
- WEBER, J. L. *et al.* **Human diallelic insertion/deletion polymorphisms**. Am J Hum Genet, v. 71, n. 4, p. 854-862, 2002.
- WELLS, R. S. *et al.* **The Eurasian heartland: a continental perspective on Y-chromosome diversity.** Proc Natl Acad Sci USA, v. 98, p. 10244-10249, 2001.
- WEN, S-Q. *et al.* Molecular genealogy of Tusi Lu's family reveals their paternal relationship with Jochi, Genghis Khan's eldest son Shao-Qing. J Hum Genet, v. 64, p. 815-820, 2019.
- WOOD, E. T. *et al.* Contrasting patterns of Y chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes. Eur J Hum Genet, v. 13, p. 867-876, 2005.
- WU, W. et al. Mutation rates at 42 Y chromosomal short tandem repeats in Chinese Han population in Eastern China. Int J Leg Med, v. 132, p. 1317-1319. 2018.
- XAVIER, C. et al. Admixture and Genetic Diversity Distribution Patterns of Non-Recombining Lineages of Native American Ancestry in Colombian Populations. PLoS ONE, v. 10, n. 3, p. e0120155, 2015.

- XIE, T. et al. A set of autosomal multiple InDel markers for forensic application and population genetic analysis in the Chinese Xinjiang Hui group. Forensic Sci Int Genet, v. 35, p. 1-8, 2018.
- YANG, Y. *et al.* Haplotypic polymorphisms and mutation rate estimates of 22 Y-chromosome STRs in the Northern Chinese Han father—son pairs. Sci Rep, v. 8, n.1, p. 1-6, 2018.
- YANG, R. et al. Sequence polymorphisms of forensic Y-STRs revealed by a 68-plex in house massively parallel sequencing panel. Forensic Sci Int Genet, v. 59, p. 102727, 2022.
- YTREE.FTDNA.COM. **Family Tree DNA.** Disponível em: http://ytree.ftdna.com. Acesso em: 29 maio. 2022.
- YUNIS, J. J.; YUNIS, E. J. Mitochondrial DNA (mtDNA) haplogroups in 1526 unrelated individuals from 11 Departments of Colombia. Genet and Mol Biol, v. 36, n. 3, p. 329-335, 2013.
- YUNUSBAYEV, B. *et al.* The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. Mol Biol Evol, v. 29, p. 359-365, 2011.
- XU, X. et al. The direction of microsatellite mutations is dependent upon allele length. Nat Genet, v. 24, n. 4, p. 396-399, 2000.
- XUE, Y. et al. Male demography in East Asia: a north-south contrast in human population expansion times. Genetics, v. 172, p. 2431-2439, 2006.
- YAN, S. et al. An updated tree of Y-chromosome Haplogroup O and revised phylogenetic positions of mutations P164 and PK4. Eur J Hum Genet, v. 19, p. 1013-1015, 2011.
- YUAN, L. *et al.* Mutation analysis of 13 RM Y-STR *loci* in Han population from Beijing of China. Int J Legal Med, v. 133, n. 1, 59-63, 2019.
- ZALLOUA, P. A. *et al.* **Y-chromosomal diversity in Lebanon is structured by recent historical events.** Am J Hum Genet, v. 82, p. 873-882, 2008.
- ZAMBRANO, A. K. The three-hybrid genetic composition of an Ecuadorian population using AIMs-InDels compared with autosomes, mitochondrial DNA and Y chromosome data. Sci Rep, v. 9, p. 9247, 2019.
- ZEGURA, S. L. *et al.* **High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas.** Mol Biol Evol, v. 21, p. 164-175, 2004.
- ZEI, G. et al. From surnames to the history of Y chromosomes: the Sardinian population as a paradigm. Eur J Hum Genet, v. 11, p. 802-807, 2003.

ZERJAL, T. *et al.* **The genetic legacy of the Mongols.** Am J Hum Genet, v. 72, p. 717-721, 2003.

ZHANG, J. et al. Mutation rates in father-son pairs of the 27 Y-STR loci in the Dezhou Han population from Shandong province, eastern China. J Forensic Leg Med, v. 67, p. 61-63, 2019.

ZHONG, H. *et al.* Global distribution of Y-chromosome haplogroup C reveals the prehistoric migration routes of African exodus and early settlement in East Asia. J Hum Genet, v. 55, p. 428-435, 2010.

ANEXO A – Documento da aprovação pelo Comitê de Ética (Paraguai)





COMITÉ DE ETICA EN INVESTIGACION CLINICA DEL INSTITUTO DE PREVISION SOCIAL

Asuncion, 25 de Marzo de 2015

EL Comité de Ética en Investigación Clínica del Instituto de Previsión Social <u>resuelve aprobar</u> el protocolo cuyo Título es:

Investigación de frecuencias genéticas en población paraguaya: marcadores STR en cromosoma-y (Y-STR) y ADN mitocondrial (ADNmt).

Investigador Principal: Dr. Calos Vullo

Co-Investigadores:

- 1. F. Simao
- 2. A.P. Ferreira
- 3. C. Vullo
- 4. C. Savier
- 5. G. Huber
- 6. A. Quiroz
- 7. P. Machado
- 8. V. Velazquez
- 9. E.F. Carvalho
- 10. L. Gusmao
- 11. W. Parson

Objetivo de la Investigación: Investigar las frecuencias genéticas en población paraguaya: marcadores STR en cromosoma-y (Y-STR) y ADN mitocondrial (ADNmt).

NSTITUTO DE PREVISION SOCIAL

Dra. Gladys Candoval

Dra. Gladys Candoval

INSTITUTO DE PREVISION SOCIAL

Dra. Glady S Andoval

Régulation Médica

Dra. Glady S Sandoval

Presidente Comité

Comité de Etica en Comité de Etica en Investigación Clínica Investigación, I.P.S. DM:Amado Denis Miembro del Comité

Constitución y Luis Alberto de Herrera - Tel: 223 141/3 - www.ips.gov.py

ANEXO B - Documento da aprovação pelo Comitê de Ética em *Human Research* da *Universidad de las Américas* - CEISH-UDLA 2017-0301



Quito, 12 de diciembre del 2022

Sr. Germán Burgos

ASUNTO: REVISIÓN DE PROTOCOLO DE INVESTIGACIÓN (observacional/Enmienda)

Por medio de la presente y una vez que el protocolo de investigación presentado por el (la) Sr (a). Germán Burgos que titula "Estudios de diversidad genética y ancestría en la población ecuatoriana", ha ingresado al Comité de Ética de Investigación en Seres Humanos de la Universidad de Las Américas, con fecha 7-diciembre 2022 (versión 2), y cuyo código asignado es 2022-ENM-001, luego de haber sido revisado y evaluado en dos ocasiones, dicho proyecto está APROBADO para su ejecución al cumplir con todos los requerimientos éticos, metodológicos y jurídicos establecidos por el reglamento vigente para tal efecto. Como respaldo de lo indicado, reposan en los archivos del CEISH-UDLA, tanto los requisitos presentados por el investigador, así como también los formularios empleados por el comité para la evaluación del mencionado estudio. Esta aprobación tiene un año de vigencia.

En tal virtud, los documentos aprobados sumillado del CEISH-UDLA que se adjuntan en físico al presente informe son los siguientes:

- Copia del protocolo de investigación " Estudios de diversidad genética y ancestría en la población ecuatoriana", versión 2, aprobado el 8 de diciembre del 2022 y Nro. de hojas 22.
- Documento de consentimiento informado, versión 2, aprobado 8 de diciembre del 2022, y Nro. de hojas 3.
- Listado de muestras anonimizadas que se enviarán a España.

Cabe indicar que la información de los requisitos presentados es de responsabilidad exclusiva del investigador, quien asume la veracidad, originalidad y autoría de los mismos. Así también se recuerda las obligaciones que el investigador principal y su equipo deben cumplir durante y después de la ejecución del proyecto:

- Informar al CEISH-UDLA la fecha de inicio y culminación de la investigación.
- Presentar a este comité informes periódicos del avance de ejecución del proyecto, según lo estime el CEISH-UDLA.
- Cumplir todas las actividades que le corresponden como investigador principal, así como las
 descritas en el protocolo con sus tiempos de ejecución, según el cronograma establecido en
 dicho proyecto, vigilando y respetando siempre los aspectos éticos, metodológicos y jurídicos
 aprobados en el mismo.
- Aplicar el consentimiento informado y asentimiento informado a todos los participantes, respetando el proceso definido en el protocolo y el formato aprobado.
- Al finalizar la investigación, entregar al CEISH-UDLA el informe final del proyecto.

Atentamente.

MARTHA MARIA Firmado digitalmente por FORS LOPEZ Fecha: 2022.12.12 16:35:50 -05'00'

Nombre y firma del presidente

ANA LUCIA
MARTINEZ

Nombre y firma secretario

ANEXO C - Documento da aprovação pelo Comitê de Ética em Investigações Biomédicas (IMBICE)



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

La Plata, 12 de Junio, 2018

Comité de Ética en Investigaciones Biomédicas

Dictamen

Proyecto: Dinámica poblacional en Santa Cruz y Tierra del Fuego: un aporte desde el estudio de linajes uniparentales

Investigador Responsable: Lic, Irina Florencia Velázquez, Laboratorio de Identificación Genética (IDEGEN) del IDEAus-CONICET.

Examinados la Cartilla de Información para los participantes y el formulario de Consentimiento Informado, referidos al Plan de Trabajo: "Dinámica poblacional en Santa Cruz y Tierra del Fuego: Un aporte desde el estudio de linajes uniparentales", presentado por la Lic Irina Florencia Velázquez, este Comité aprueba las correcciones introducidas en ambos documentos, de acuerdo a las indicaciones realizadas por los miembros del Comité. Por este motivo y en razón de que la formulación del proyecto contempla las recomendaciones enunciadas en el Marco Ético de Referencia para las Investigaciones Biomédicas : Principios Éticos para las Investigaciones Biomédicas en seres humanos, el Comité de Ética en Investigaciones Biomédicas del IMBICE avala la ejecución del Proyecto sometido a su evaluación.

Dra. Lidia A. Vidal Rioja

Coordinadora

Comité de Ética en Investigaciones Biomédicas - IMBICE

RENIS CE000023

Dra. Lidia A. Vidal Rioja - Coordinadora

Prof. Lic. Jorge P. Asuaje

Dr. Luis J. Couyet

Dr. Eduardo L. Tinant

ANEXO D - Documento da aprovação pelo Comitê de Ética - Pontifica Universidade Javeriana



EL COMITÉ DE INVESTIGACIÓN Y ÉTICA DE LA FACULTAD DE CIENCIAS DE LA PONTIFICIA UNIVERSIDAD JAVERIANA

CERTIFICA

Que el proyecto de investigación titulado "Delineando la estructura genética de la población colombiana mediante el análisis molecular de marcadores INDELs: un estudio de ancestralidad e identificación", cuyo investigador principal es Humberto Ossa Reyes, estudiante del Doctorado en Ciencias Biológicas de la Facultad, fue evaluado y aprobado por parte del Comité de Investigación y Ética de la Facultad de Ciencias en su sesión del 23 de Febrero de 2015, considerando la pertinencia de la investigación, el rigor metodológico, su calidad científica, la coherencia y la racionalidad del presupuesto propuesto y el cumplimiento de las normas científicas, técnicas y éticas nacionales e internacionales que rigen este tipo de investigaciones.

El proyecto implica investigación en seres humanos y se ajusta a las Normas Científicas, Técnicas y Administrativas para la Investigación en Salud establecidas en la Resolución No. 008430 de 1993 y la Resolución 2378 de 2008. La categoría de riesgo a los seres humanos que ofrece la propuesta pertenece a la de riesgo mínimo.

El consentimiento informado elaborado para este proyecto incluye los aspectos requeridos para proveer la información necesaria a las personas que se incluyan en el estudio y el investigador principal debe garantizar la obtención del documento firmado por cada uno de los participantes en el estudio.

El Comité conceptúa que, por la naturaleza de la investigación, el proyecto no tiene efectos negativos sobre el medio ambiente.

La ejecución de procedimientos del laboratorio previstos en esta investigación se regirá bajo las normas de bioseguridad del Laboratorio del Instituto de Biología Roberto Alcántara Gómez de la U. del Estado de Río de Janeiro, de la Facultad de Ciencias.

Con base en lo expresado anteriormente, el Comité de Investigación y Ética conceptúa que el proyecto cumple con todos los requisitos de calidad exigidos y en consecuencia otorga su aprobación. El respectivo concepto se consigna en el Acta N° 1 de la correspondiente sesión.

FACULTAD DE CIENCIAS – DECANATURA



Se expide esta certificación el 15 de Mayo de 2015.

CONCEPCIÓN PUERTA BULA Ph.D.

Presidente del Comité de Investigación y Ética

FACULTAD DE CIENCIAS – DECANATURA

ANEXO E - Parecer do Comitê de Ética em Pesquisa do Hospital Pedro Ernesto (Rio de Janeiro, Brasil)



UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO HOSPITAL UNIVERSITÁRIO PEDRO ERNESTO COMITÊ DE ÉTICA EM PESOUISA



Rio de Janeiro, 14 de outubro de 2009

Do: Comitê de Ética em Pesquisa Profa. Patrícia Maria C. O. Duque

Para: Coord. Prof. Elizeu Fagundes de Carvalho

O Comitê de Ética em Pesquisa do Hospital Universitário Pedro Ernesto, após avaliação, considerou o projeto (2485-CEP/HUPE – CAAE: 0067.0.228.000-09) "REGIÕES POLIMÓRFICAS DO DNA COMO FERRAMENTAS PARA ESTUDOS DE GENÉTICA DE POPULAÇÕES E FORENSE: UTILIZAÇÃO COMO MARCADORES DE ANCESTRALIDADE, MOVIMENTOS MIGRATÓRIOS, ESTUDOS SOBRE ESTRUTURAÇÃO GÊNICA DA POPULAÇÃO BRASILEIRA E INVESTIGAÇÕES CRIMINAIS E DE VÍNCULOS GENÉTICOS" aprovado, encontrando-se este dentro dos padrões éticos da pesquisa em seres humanos, conforme Resolução n.º196 sobre pesquisa envolvendo seres humanos de 10 de outubro de 1996, do Conselho Nacional de Saúde, bem como o consentimento livre e esclarecido.

O pesquisador deverá informar ao Comitê de Ética qualquer acontecimento ocorrido no decorrer da pesquisa.

O Comitê de Ética solicita a V. Sa., que ao término da pesquisa encaminhe a esta comissão um sumário dos resultados do projeto.

Profal Patricia Maria C. O. D. Membro do Comitê de Ética em Pes

Memoro do comite de Edica em Pe

ANEXO F – Documento da aprovação pelo Comitê de Ética em Pesquisa– Hospital Pedro Ernesto (Rio de Janeiro, Brasil)



UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO HOSPITAL UNIVERSITÁRIO PEDRO ERNESTO COMITÊ DE ÉTICA EM PESQUISA



Rio de Janeiro, 10 de novembro de 2010

Do: Comitê de Ética em Pesquisa Prof.: Wille Oigrnán

Para: Coord. Profa. Marília de Brito Gomes

Registro CEP/HUPE: 2769/2010 (este número deverá ser citado nas correspondências referentes ao projeto)

CAAE: 0214.0.228.000-10

O Comitê de Ética em Pesquisa do Hospital Universitário Pedro Ernesto, após avaliação, considerou o projeto, "PREVALÊNCIA E MARCADORES DEMOGRÁFICOS, CLÍNICOS, LABORATORIAIS E GENÉTICOS DAS COMPLICAÇÕES CRÔNICAS MÎCRO E MACROVASCULARES EM PACIENTES COM DIABETES TIPO 1 E . AVALIAÇÃO DOS CUSTOS DA DOENÇA" aprovado, encontrando-se este dentro dos padrões éticos da pesquisa em seres humanos, conforme Resolução n.º196 sobre pesquisa envolvendo seres humanos de 10 de outubro de 1996, do Conselho Nacional de Saúde, bem como o termo de consentimento livre e esclarecido.

O pesquisador deverá informar ao Comitê de Ética qualquer acontecimento ocorrido no decorrer da pesquisa.

O Comitê de Ética solicita a V_r;Sa., que ao término da pesquisa encaminhe a esta comissão UM sumário dos resultados do projeto.

/Hrof. Wille Oigman

Presidente do Comitê d

a

ANEXO G "The Ancestry of Eastern Paraguay: A Typical South American Profile with a Unique Pattern of Admixture" publicado na revista Genes, em 2121, disponível em: 10.3390/genes12111788.





The Ancestry of Eastern Paraguay: A Typical South American Profile with a Unique Pattern of Admixture

Filipa Simão ¹, Julyana Ribeiro ¹, Carlos Vullo ², Laura Catelli ², Verónica Gomes ^{3,4}, Catarina Xavier ⁵, Gabriela Huber ⁵, Martin Bodner ⁵, Alfredo Quiroz ⁶, Ana Paula Ferreira ¹, Elizeu F. Carvalho ¹, Walther Parson ^{5,7}, ⁶, and Leonor Gusmão ¹, ⁶

- DNA Diagnostic Laboratory, State University of Rio de Janeiro, Rio de Janeiro 20550-013, Brazil; $f. simao@campus.fct.unl.pt~(F.S.); julyanaribeiro@ymail.com~(J.R.); apfa_bio@yahoo.com.br~(A.P.F.); apfa_Bio@yahoo.com.br~(A.P.F.); apfa_Bio@yahoo.com.br~(A.P.F.); apfa_Bio@yahoo.com.br~(A.P.F.); apfa_Bio@yahoo.com.br~(A.P.F.); apfa_Bio@yahoo.com.br~(A.P.F.); apfa_Bio@yahoo.com.br~(A.P.F.); apfa_Bio@yahoo.com.br~(A$ elizeufc@hotmail.com (E.F.C.)
- DNA Forensic Laboratory, Argentinean Forensic Anthropology Team, Córdoba 14001, Argentina; cvullo@yahoo.com.ar (C.V.); marialauracatelli@gmail.com (L.C.) Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4099-002 Porto, Portugal;
- vgomes@ipatimup.pt
- Institute of Pathology and Molecular Immunology, University of Porto (IPATIMUP), 4099-002 Porto, Portugal
- Institute of Legal Medicine, Medical University of Innsbruck, 6020 Innsbruck, Austria catarina.gomes@i-med.ac.at (C.X.); gabriela.e.huber@i-med.ac.at (G.H.); martin.bodner@i-med.ac.at (M.B.) Instituto de Previsión Social, Asunción 100153, Paraguay; alfredoquiroz1974@me.com
- Forensic Science Program, The Pennsylvania State University, State College, PA 16801, USA
- * Correspondence: walther.parson@i-med.ac.at (W.P.); leonorbgusmao@gmail.com (L.G.)



ANEXO H "*Tierra Del Fuego: What Is Left from the Precolonial Male Lineages?*", publicado na revista *Genes*, em 2022, disponível em: 10.3390/genes13101712.





Article

Tierra Del Fuego: What Is Left from the Precolonial Male Lineages?

Pedro Rodrigues ^{1,2,†}, Irina Florencia Velázquez ^{3,†}, Julyana Ribeiro ⁴, Filipa Simão ⁴, António Amorim ^{1,2,5}, Elizeu F. Carvalho ⁴, Claudio Marcelo Bravi ⁶, Néstor Guillermo Basso ³, Luciano Esteban Real ³, Claudio Galli ⁷, Andrea del Carmen González ⁸, Ariana Gamulin ⁹, Romina Saldutti ¹⁰, Maria Laura Parolin ^{3,*}, Verónica Gomes ^{1,2,*}, ¹ and Leonor Gusmão ^{4,‡}

- Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4099-002 Porto, Portugal
- Institute of Pathology and Molecular Immunology, University of Porto (IPATIMUP), 4099-002 Porto, Portugal
- Instituto de Diversidad y Evolución Austral (IDEAus-CONICET), Puerto Madryn 9120, Argentina
- ⁴ DNA Diagnostic Laboratory, State University of Rio de Janeiro, Rio de Janeiro 20550-013, Brazil
- Faculty of Sciences, University of Porto, 4169-007 Porto, Portugal
- 6 Laboratorio de Genética Molecular Poblacional, IMBICE (CCT-CONICET, CIC-PBA), Universidad Nacional de La Plata (UNLP), La Plata 1906, Argentina
- Servicio de Hemoterapia, Hospital Regional de Rio Grande, Rio Grande 9420, Argentina
- 8 Servicio de Hemoterapia, CEMEP, Rio Grande 9420, Argentina
- 9 Servicio de Hemoterapia, Hospital Regional de Ushuaia, Ushuaia 9410, Argentina
- Servicio de Hemoterapia, Clínica San Jorge, Ushuaia 9410, Argentina
- * Correspondence: parolinml@gmail.com (M.L.P.); vgomes@ipatimup.pt (V.G.)
- † These authors contributed equally to this work.
- ‡ These authors contributed equally to this work.

ANEXO I "Male lineages in Brazilian populations and performance of haplogroup prediction tools", publicado na revista Forensic Science International: Genetics, em 2020, disponível em: 10.1016/j.fsigen.2019.102163.



Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen



Male lineages in Brazilian populations and performance of haplogroup prediction tools



Juliana Jannuzzia,*, Julyana Ribeiroa, Clarice Alhob, Grasielly de Oliveira Lázaro e Arãoc, Regina Cicarelli^d, Heitor Simões Dutra Corrêa^e, Suelen Ferreira^f, Cíntia Fridman^g, Verónica Gomes^{h,i}, Silvia Loiola^a, Mariana Flavia da Mota^c, Ândrea Ribeiro-dos-Santosⁱ, Carlos Antonio de Souza^k, Rossana Santiago de Sousa Azulay^l, Elizeu F. Carvalho^a, Leonor Gusmão^a

- ^a DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil

 ^b Laboratory of Human and Molecular Genetics, Pontificia Universidade Católica do Rio Grande do Sul, Porto Alegre, Brazil

 ^c Laboratório de Biologia e DNA Forense, Superintendência de Policia Técnico-Científica do Estado de Goiás, Goiás, Brazil

 ^d UNESP-Universidade Estadual Paulista, Faculdade de Ciências Farmacêuticas, Laboratório de Investigação de Paternidade-NAC, São Paulo, Brazil

- ^a UNESP-Universidade Estadual Paulista, Faculdade de Cièncias Farmacèuticas, Laboratório de Investigação de Paternidade-NAC, São Paulo, Brazil

 ^c Coordenadoria de Perícias em Biologia Molecular, POLITEC Perícia Oficial e Identificação Técnica, Mato Grosso, Brazil

 ^f Faculdade Pitágoras, São Luís, Brazil

 ^g Departamento de Medicina Legal, Étita Médica e Medicina Social e do Trabalho da Faculdade de Medicina da USP, Universidade de São Paulo (USP), São Paulo, Brazil

 ^h PATIMUP-Institute of Pathology and Molecular Immunology from the University of Porto, Portugal

 ¹ ISS-Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal

 ¹ Postgraduate Program in Genetics and Molecular Biology, Laboratory of Human and Medical Genetics, Federal University of Pará, Belém, Brazil

 ¹ Secretaria de Defesa Social Pernambuco, Brazil

 ¹ Serviço de Endocrinologia e Metabologia do Hospital Universitário da Universidade Federal do Maranhão, Maranhão, Brazil

ANEXO J "Microsatellites Mutation Modeling Through the Analysis of the Y-Chromosome Transmission: results of a GHEP-ISFG Collaborative Study" publicado na revista Forensic Science International: Genetics, em 2023, disponível em: 10.1016/j.fsigen.2023.102999.



Microsatellites' Mutation Modeling Through the Analysis of the Y-Chromosomal Transmission: results of a GHEP-ISFG Collaborative Study

Sofia Antão-Sousa a b c d 1, Leonor Gusmão a d, Nidia M. Modesti e, Sofia Feliziani e, Marisa Faustino a c, Valeria Marcucci f, Claudia Sarapura f, Julyana Ribeiro d, Elizeu Carvalho d, Vania Pereira g, Carmen Tomas g, Marian M. de Pancorbo h, Miriam Baeta h, Rashed Alghafri l, Reem Almheiri l, Juan José Builes J k, Nair Gouveia l, German Burgos m n, Maria de Lurdes Pontes o, Adriana Ibarra p...Nadia Pinto a b t