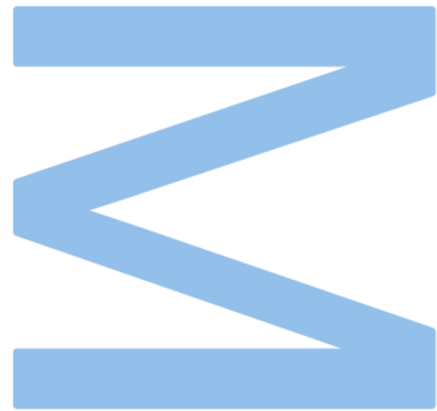
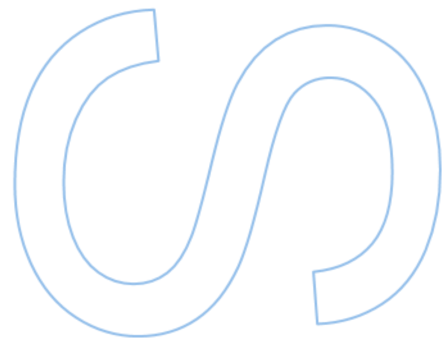
Exploring the maternal genetic landscape of Andean Mestizos from Ecuador

Karen Alejandra Garzón Salazar

Master in Forensic Genetics

Biology Department

2023

Orientador

Filipa Simão, PhD, VCU

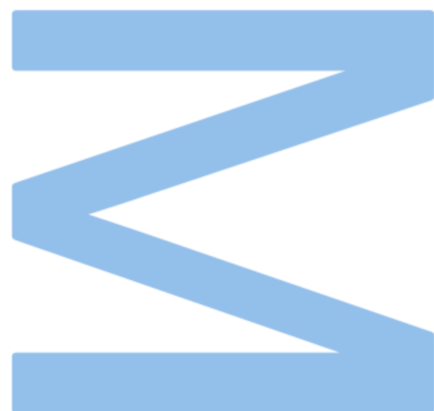
Coorientador

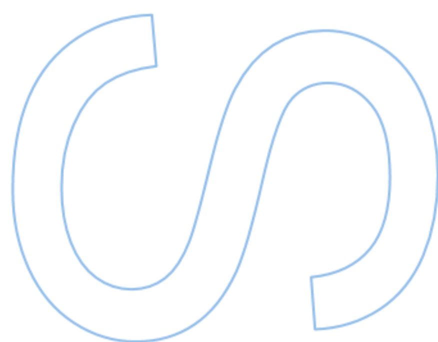
Verónica Gomes, PhD, IPATIMUP/I3S

Leonor Gusmão, PhD, UERJ

INSTITUTO DE INVESTIGAÇÃO
E INOVAÇÃO EM SAÚDE
UNIVERSIDADE DO PORTO

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO





À minha avó

Agradecimentos

Ao Porto, cidade invicta, que me acolheu de braços abertos durante estes dois anos. Foi uma travessia plena de aprendizagem, resiliência, amizade e muito carinho.

Aos meus professores orientadores que me guiaram em todos os momentos, para que eu apresentasse um projeto capaz de narrar, por intermédio da genética, um pouco da história do meu país natal.

À Filipa Simão, minha orientadora, porque a pesar da distância, sempre esteve pronta para me ajudar a resolver as minhas dúvidas e problemas. Obrigada pela transmissão de conhecimento e pelas constantes dicas de melhoria.

À Verónica Gomes, minha coorientadora, por me receber de uma forma tão gentil no laboratório e me fazer sentir em casa. Obrigada por toda a ajuda no trabalho laboratorial e *in silico* e por todos os conselhos dados e estratégias definidas para que os objetivos propostos pudessem ser cumpridos. Muito obrigada pela amabilidade e amizade, que fizeram deste ano uma experiência muito enriquecedora tanto a nível científico como a nível pessoal.

À Leonor Gusmão, minha coorientadora que, apesar da distância, sempre se mostrou disponível para me guiar, dando-me feedback de todas as dúvidas e partilhando comigo todo o seu conhecimento, e estimular o meu espírito crítico. Muito obrigada pelo apoio e disponibilidade.

Ao Germán Burgos, meu professor em Quito, por ter depositado a sua confiança em mim nesta travessia, por me impulsionar a conseguir atingir os objetivos e pelo apoio fundamental na realização deste projeto.

Ao Rodrigo Flores, meu amigo e colega de laboratório no Equador, que sempre esteve disponível para me ajudar em qualquer momento.

Ao grupo de Genética Populacional do I3S, começando pelo Prof. António Amorim, pela sua simpatia e por partilhar a sua sapiência comigo. Ao resto do grupo: Nádia, Iva, Sandra, Catarina, Sofia, Marisa, Camila e Mariana por toda a amabilidade.

Aos meus pais e à minha irmã, um agradecimento do fundo do coração, porque sem eles nada disto teria sido possível. Obrigada pela paciência, amor e apoio durante esta jornada. Obrigada por me ensinarem a ser uma pessoa forte e capaz de concretizar os objetivos a que me proponho. Obrigada por estarem presentes nos bons e maus momentos. Obrigada pela incondicionalidade.

A todos os amigos que fiz em Portugal, um agradecimento enorme por terem enchido os meus dias de aventuras e sorrisos. Começando pelas minhas queridas Inês, Rita e Ana, pela simpatia desde o primeiro dia do Mestrado, à Stephanie, ao Bruno, à Thais, ao Robson, à Johanna, à Maria, e ao Rodrigo por tanta luz, simpatia e carinho que me fizeram sentir. Obrigada por me acompanharem e me mostrarem, nesta travessia, a magia da música e cultura portuguesa.

Obrigada por tudo!

Resumo

As populações humanas da América do Sul passaram por uma série de eventos de miscigenação complexos que moldaram a sua composição genética. O Equador conserva as heranças genéticas europeias e africanas adquiridas durante o período colonial. Atualmente os Mestiços (indivíduos resultantes da mistura de Nativos, Europeus e Africanos) representam 77.4% da população do país. O presente estudo centra-se na avaliação das contribuições genéticas maternas dos Mestiços da região Andina do Equador, com o objetivo de clarificar o impacto dos eventos migratórios humanos na região. A ancestralidade materna foi explorada através da sequenciação da região controlo completa do mtDNA de 214 Mestiços Andinos do Equador. As amostras foram divididas em três subgrupos de acordo com o local de residência dos indivíduos - Norte, Centro e Sul. Foram encontradas altas diversidades haplotípicas nas três regiões: Norte (0,9993 +/- 0,0023), Centro (0,9981 +/- 0,0030) e Sul (0,9960 +/- 0,0036). Estes valores são esperados em populações miscigenadas, embora mais elevados quando comparados com países vizinhos. A maioria dos haplogrupos de mtDNA encontrados na nossa amostra é de origem Nativo Americana (A2, B2, B4, C1, D1, D4). No entanto, foram observadas diferenças na proporção de algumas linhagens Nativas entre os três subgrupos Mestiços. Os haplogrupos africanos (L1, L2, L3) e eurasiáticos (U4c1, HV, R) foram encontrados em baixas frequências nos Andes do Norte e do Sul, estando ausentes na região Central, onde a maioria das comunidades Nativas americanas está estabelecida. No entanto, as distâncias genéticas F_{ST} calculadas usando frequências de haplótipos, não revelaram diferenças estatisticamente significativas entre os três subgrupos, sendo possível concluir que, ao nível do mtDNA, não existe uma subestrutura populacional significativa entre as três regiões geográficas consideradas. No contexto da América do Sul, as nossas amostras demonstraram uma semelhança genética com outras populações dos Andes Ocidentais, principalmente devido à elevada prevalência do haplogrupo B. Esta observação sugere um percurso evolutivo distinto em comparação com as populações das regiões orientais da América do Sul.

Palavras-chave: Genética de populações, América do Sul, Equador, região Andina, grupo étnico Mestiço, linhagens femininas, mtDNA, diversidade.

Abstract

Human populations in South America experienced a series of complex admixture events that shaped their genetic makeup. Ecuador retains the European and African genetic inheritances acquired during the colonial period. Currently, Mestizos (individuals with Native, European, and African admixture) represent 77.4% of the population scattered across the country. The present study focuses on assessing the maternal genetic composition of Mestizos from the Andean region of Ecuador, aiming to clarify the impact of human migration events in the region. The maternal ancestry was explored through complete mtDNA control region sequencing of 214 Andean Mestizos from Ecuador. Samples were divided into three subgroups according to the living place of the individuals - North, Center, and South. High haplotype diversities were found in the three regions: North (0.9993 +/- 0.0023), Center (0.9981 +/- 0.0030), and South (0.9960 +/- 0.0036). These values are expected in admixed populations, although higher when compared to neighboring countries. Most of the mtDNA haplogroups found in our sample were of Native American origin (A2, B2, B4, C1, D1, D4). However, differences in the proportion of some Native lineages were observed among the three Mestizo subgroups. African (L1, L2, L3) and Eurasian (U4c1, HV, R) haplogroups were encountered at low frequencies in North and South Andes, but none were present in Central Andes, where most Native American communities are established. Nonetheless, the pairwise F_{ST} genetic distance calculated using haplotype frequencies revealed no statistically significant differences between the three subgroups, demonstrating that, at the mtDNA level, there is no significant population substructure between the three geographic regions considered. In the context of South America, our samples demonstrated a genetic similarity to other West Andean populations, primarily due to the elevated prevalence of haplogroup B. This observation suggests a distinct evolutionary path compared to the populations from the Eastern regions of South America.

Keywords: Population genetics, South America, Ecuador, Andean region, Mestizo ethnic group, female lineages, mtDNA, diversity.

Table of Contents

Table Index vii

Figure Index viii

Abbreviations ix

1. Introduction 1

1.1. Population Genetics 1

1.2. Human genetic variation 1

1.3. Mitochondrial DNA as a genetic marker 3

1.3.1. Mitochondria 3

1.3.2. Human Mitochondrial DNA 4

1.3.3. Properties of the human mtDNA 6

Maternal Inheritance 6

1.3.4. Human Mitochondrial DNA Variation 8

mtDNA Haplogroups 10

1.4. The Americas human migration history 11

1.4.1. Finding routes toward South America 13

1.5. Ecuador 14

1.5.1. The first settlers in Ecuador 15

1.5.2. Inca Empire 16

1.5.3. Spaniard Colonization 17

1.5.4. Current Ecuador: geography and demographics 18

1.5.5. Genetic Diversity of Ecuador 19

2. Aims 21

3. Methods and Materials 22

3.1. Population samples and DNA extraction 22

3.2. Mitochondrial DNA analysis 23

3.2.1. PCR Amplification 23

3.2.2. Enzymatic PCR cleanup 24

3.2.3. Sanger Sequencing 24

3.2.4. Sequencing products cleanup and electrophoresis 25

3.3. Data Analysis 26

4. Results and Discussion 27

4.1. Mitochondrial DNA sequence variation in Andean Mestizos 27

4.2. Genetic Diversity of Andean Mestizos from Ecuador 29

4.3. Haplogroup Composition 30

4.4. Population structure within the Ecuadorian Andes 34

4.5. Ecuadorian Andes in a South American context 35

4.5.1. Maternal ancestry 35

4.5.2. Native American composition 39

Conclusions 42

References 43

Anexos 63

Table Index

Table 1. Molecular diversity indexes for mtDNA control region data from the present sample and other South American populations 30

Figure Index

Figure.1 Representation of biparental and uniparental patterns of inheritance 3
Figure 2. Illustration of Mitochondrion structure.. 4
Figure 3. Representation of the circular mtDNA genome. 5
Figure 4. Pedigree representing the maternal uniparental inheritance highlighted in blue. 7
Figure 5. The basis of Sanger Sequencing.. 9
Figure 6. mtDNA sequences. 10
Figure 7. The distribution of the major mtDNA haplogroups explained by human migrations. 11
Figure 8. Map of Northeast Asia and Northwest North America, highlighting the Bering Strait (point A) 12
Figure 9. Representation of two ancient human migrations into and within South America.]. 14
Figure 10. Illustration of Ecuador's geographic natural division (Mainland Ecuador: Coast, Andes, and Amazonia; and Insular Region: The Galápagos Islands)........ 15
Figure 11. The Pre-Columbian settlements in different Ecuadorian regions. A) First human settlements in Ecuador; B) Agricultural society settlements; C) Ethnic Lordships. 16
Figure 12. Geographical distribution of the Ecuadorian samples from the Andes used in the present study.. 22
Figure 13. The workflow of mitochondrial CR typing. 23
Figure 14. Diagram with the primers used.. 25
Figure 15. Representation of mtDNA haplogroups and sub-haplogroups found in each Andean populations from Ecuador. 31
Figure 16. Median joining network of Ecuadorian Andes samples (n=214). 33
Figure 17. MDS plot based on haplotype F_{ST} genetic distances................................. 36
Figure 18. Maternal lineages found in several admixed populations from South America 37
Figure 19. Principal component analysis (PCA) representing the genetic landscape of South American countries based on mtDNA Native haplogroups frequencies 40

Abbreviations

AIM	Ancestry Informative Marker
ATP	Adenosine triphosphate
CR	Control Region
CRS	Cambridge Reference Sequence
DNA	Deoxyribonucleic acid
D-loop	Displacement loop
dNTP	Deoxynucleotide triphosphate
ddNTP	Dideoxynucleotide triphosphates
HVS-I	Hypervariable segment I
HVS-II	Hypervariable segment II
HVS-II	Hypervariable segment III
H-strand	Heavy strand
InDel	Insertion-deletion
IFC	Ice free corridor
kb	Kilobase
LHP	Length heteroplasmy
L-strand	Light strand
mtDNA	Mitochondrial DNA
MSY	Male-specific region of the Y chromosome
MPS	Massive parallel sequencing
MDS	Multidimensional scaling
NGS	Next generation sequencing
NNA	North Native Americans
SNA	South Native Americans
PCR	Polymerase chain reaction
PCA	Principal component analysis
PHP	Point heteroplasmy
RFLP	Restriction Fragment Length Polymorphism
RSRS	Reconstructed Sapiens Reference Sequence
rCRS	Revised Cambridge Reference Sequence
r-RNA	Ribosomal ribonucleic acid
SNP	Single Nucleotide Polymorphism
STR	Short Tandem Repeat
t-RNA	Transfer ribonucleic acid

VNTR	Variable Number Tandem Repeat
Y-SNP	Y chromosome Single Nucleotide Polymorphism
Y-STR	Y chromosome Short Tandem Repeat

1. Introduction

1.1. Population Genetics

Human migrations played an essential role in shaping the genetic diversity of our species. Over the years, human populations scattered across the globe have encountered different environmental conditions, food resources, and even pathogens that led to various genetic adaptations with the development of new genetic traits. Two examples of this adaptation are the lighter skin pigmentation in populations at higher latitudes (Jablonski & Chaplin, 2010), and the presence of sickle cell disease in regions with a high incidence of malaria (Mkombachepa et al., 2022). Continuous migrations were and still are responsible for inducing genetic admixture between different populations, which results in the increment of diversity.

Population genetics is the study of the genetic composition of populations and the variation within or between populations over space and time. This branch of genetics focuses on the evolutionary processes behind genetic variation, such as genetic drift, mutation, gene flow, and natural selection (Amorim, 2013; Cavalli-Sforza, 1974; Hartl & Clark, 2007; Keats & Sherman, 2014; Luikart et al., 2018). This field has a wide range of applications in several areas, namely in: (i) medicine, by elucidating how genetic variation influences therapeutic response and illness vulnerability (Giugliani et al., 2019); (ii) forensics, by disclosing DNA variants distribution across populations, essential for statistical evaluation of genetic evidence (Amorim & Budowle, 2016) and; (iii) anthropology, by interpreting human genetic data together with culture, history, and demographic dynamics, to understand ancestral behaviors (Relethford, 2017). The integration of genetic data with archaeological and linguistic evidence has been the focus of population studies that aim to track down the migration patterns of ancient and modern human populations and shed light on their origins.

1.2. Human genetic variation

The genome, which is composed of sets of nucleic acids arranged into DNA molecules, is the foundation of heredity in most living organisms. In eukaryotes, DNA is

condensed into chromosomes. Each human cell has 22 pairs of autosomes and one pair of sex chromosomes (XY in males or XX in females). Although most of the DNA is found in the nucleus of the cell, with a linear structure – nuclear DNA -, a small portion is found in the matrix of mitochondria, with a circular structure - mitochondrial DNA.

Over 99.7% of human DNA is identical in all individuals, with only a small percentage presenting variation (Butler, 2010a). Most of the genetic variation is found on the non-coding region of the DNA and can be presented in the form of insertions and deletions of one or more nucleotides (InDels), point mutations known as Single Nucleotide Polymorphisms (SNPs), and a variable number of tandem repeats (VNTRs), which are motifs of several nucleotides organized as a tandem repeat, with variations in length. This last form of variation can be distinguished as minisatellites and microsatellites depending on the size of the repeated DNA motifs. The microsatellites or Short Tandem Repeats (STRs) are variations of short motifs (less than 10 nucleotides) and show the highest power of discrimination (Kimpton et al., 1993; Sprecher et al., 1996), being the standard choice for human identification purposes.

All the above-mentioned types of polymorphism can be found scattered across the human genome. Nevertheless, the location of these variants in the genome impacts the transmission mode over generations. Unlike autosomes, the male-specific region of the Y chromosome (MSY) and the mitochondrial DNA (mtDNA) do not suffer recombination and have uniparental transmission (*Figure 1*). This means that mutations are accumulated sequentially over generations and inherited purely down the paternal or maternal line of descent. The genetic information on these regions is described as a haplotype, which is not the case for unlinked autosomal polymorphisms (H.-J. Bandelt et al., 2006; Butler, 2010b; Jobling & Tyler-Smith, 2003; Pereira & Gusmão, 2016).

The study of genetic polymorphisms located in MSY and mtDNA plays an important role in population genetics and evolutionary biology, giving indications of population demography, and allowing the reconstruction of historical patterns.

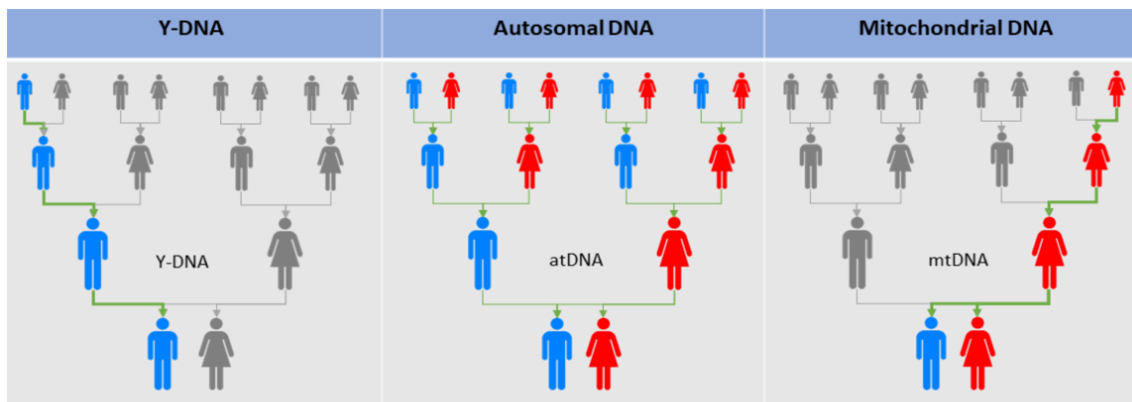


Figure.1 Representation of biparental and uniparental patterns of inheritance [Adapted from www.antoniosdnaproject.de].

1.3. Mitochondrial DNA as a genetic marker

1.3.1. Mitochondria

Mitochondria are small organelles scattered in the cytoplasm of eukaryotic cells, delimited by two specialized membranes: the inner and outer membranes, each including a lipid bilayer (*Figure 2*). They oversee cellular homeostasis, playing a part in the metabolism of amino acids, lipids, and steroids, intracellular signalling and apoptosis, and in intermediary metabolism (i.e., tricarboxylic acid and Krebs cycle). However, their prime role is energy production in the form of ATP through oxidative phosphorylation (Gray, 2001; Manoli et al., 2007).

The origin of the mitochondrion and its genome is explained by the “Endosymbiont hypothesis”. According to this hypothesis, these organelles are descendants of cells (endosymbionts) that found a way to live inside another host cell. Based on genetic evidence, the mitochondrion derives from an independent bacterium (α-Proteobacteria) (Garg et al., 2016; Manoli et al., 2007) engulfed by a proto-eukaryotic cell that required energy from its guest, developing a symbiotic relation (Goios & Alvarez, 2013; Gray et al., 1999).

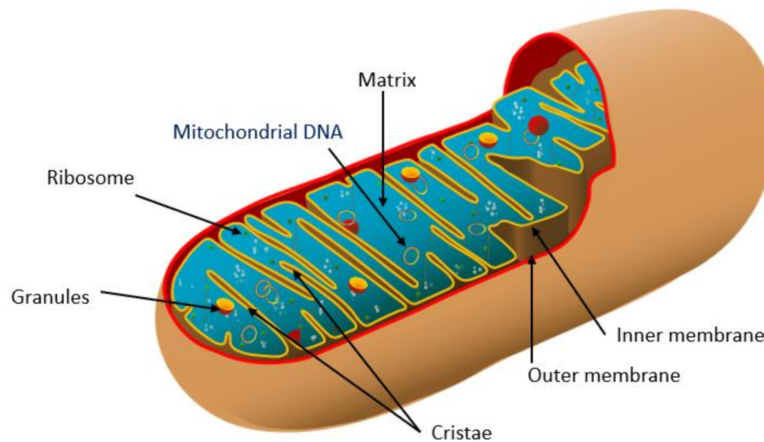


Figure 2. Illustration of Mitochondrion structure [Adapted from TeachMe Physiology].

1.3.2. Human Mitochondrial DNA

Mitochondrial DNA is a small circle double-stranded molecule located within the matrix of the mitochondria. The two strands are referred to as the heavy strand (H-strand), and the light strand (L-strand) (*Figure 3*), based on the unequal distribution of nucleotides between them. The heavy strand contains a greater proportion of guanine (G) than the light strand, which has a higher cytosine (C) content (Scheffler, 2008).

The human mtDNA has approximately 16.6 kilobases (kb) and is divided into two main regions: the coding region and the control region (CR) (*Figure 3*). The coding region of mtDNA contains genes that code for 37 products that include 13 essential proteins of the respiratory chain, 2 ribosomal RNAs (rRNAs), and 22 transfer RNAs (tRNAs). The control region is a non-coding region with 1.1 kb and supervises the regulation of transcription and replication of the molecule (Bandelt et al., 2006; Robin & Wong, 1988; Satoh & Kuroiwa, 1991). The CR is also known as a D-loop (displacement loop) since it forms a loop structure during replication. This region contains the origin of replication in the heavy strand and comprises three short regions distinguished by their high mutation rate, known as hypervariable segments (HVS-I, HVS-II, and HVS-III) (*Figure 3*).

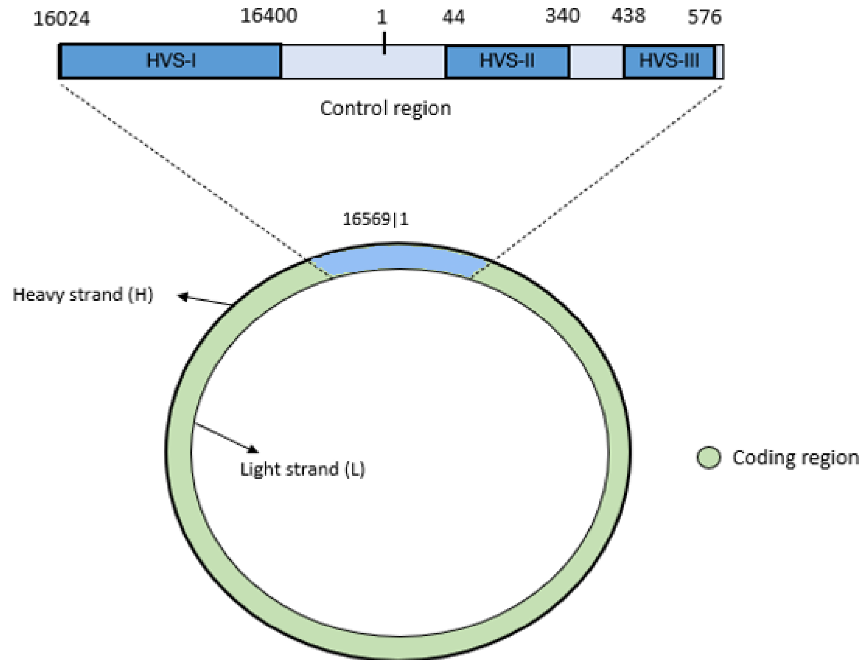


Figure 3. Representation of the circular mtDNA genome. The Heavy strand (H) is shown in the outside layer and the Light strand (L) is on the inside layer. The coding region (between positions 577 and 16023) is marked in green. Control Region (between positions 16024 and 576), at the top of the graph with its three hypervariable regions (HVS-I, HVS-II, and HVS-III), is marked in blue.

The Cambridge Reference Sequence (CRS) or "Anderson" sequence was the first human mtDNA full sequence to be reported, by the Laboratory of Frederick Sanger group in Cambridge (Anderson et al., 1981). A numerical identification was attributed to each nucleotide based on the L-strand. The positions in the CRS are numbered from 1 to 16569. Position 1 does not lie in the replication origin but falls between HVS-I and HVS-II (*Figure 3*) (Bandelt et al., 2006; Brandstätter, Niederstätter, et al., 2004). The control region extends from position 16024 to position 576. The HVS-I spans between positions 16024-16400, HVS-II between positions 44-340, and HVS-III between positions 438-576 (Brandstätter, Peterson, et al., 2004).

In 1999, the resequencing of the original placental mtDNA sample used by Anderson et al. (1981) showed some errors. A total of 11 mismatches were detected (10 substitution errors and two cytosine residues instead of one at 3106 and 3107 positions). The revised CRS (rCRS) was implemented after errors corrections and is currently used for nomenclature purposes, where every mtDNA sequence studied is reported by indicating the differences against the rCRS.

In 2012, a proposal was made to replace the rCRS with a reference sequence that represents the ancestral state of the human mtDNA sequences in modern

populations. The sequence was called Reconstructed Sapiens Reference Sequences (RSRS) and represents the more recent common ancestor of all mtDNA sequences (Behar et al., 2012). Nevertheless, there was a consensus among the scientific community to maintain the rCRS as a reference, since there are numerous works based on it and a change in the nomenclature could lead to mistakes or misinterpretations (Bandelt et al., 2013).

1.3.3. Properties of the human mtDNA

Maternal Inheritance

One of the great advantages of using mtDNA in population genetic studies relies on its exclusively maternal inheritance (Giles et al., 1980; Shitara et al., 1998). This genetic transmission mode allows tracking maternal-related lineages through time without the effects of recombination (*Figure 4*). Studies have demonstrated that sperm midpiece, which also contains mitochondria, is tagged by a proteolytic peptide (ubiquitin) for subsequent mitochondria destruction in the oocyte cytoplasm after fertilization (Sutovsky et al., 1999; Thompson et al., 2003), inhibiting paternal transmission of the mtDNA. Although uncommon cases of paternal leakage of mtDNA have been recorded (Luo et al., 2018; Schwartz & Vissing, 2002), significant recombination between maternal and paternal mtDNA is extremely unlikely (H.-J. Bandelt et al., 2006; Howell, 1997). Thus, mtDNA molecules are passed down to all maternal offspring, which means that if the mother bears a mutation, the same mutated sequence will be passed down to her offspring (Butler, 2011).

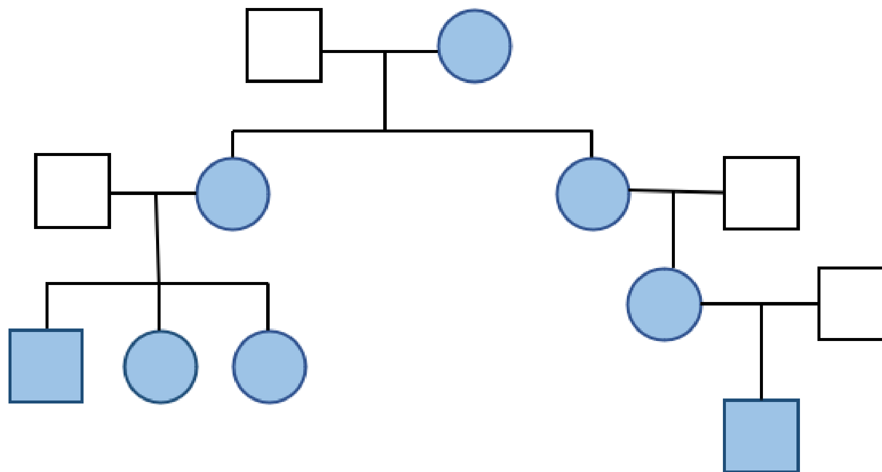


Figure 4. Pedigree representing the maternal uniparental inheritance highlighted in blue. Females are represented in circles and males in squares.

Copy Number

The human cell contains a variable range of hundreds to thousands of mitochondria. Contrarily to the nuclear genome that only carries two copies per cell, each mitochondrion can carry, on average, 4-5 copies of mtDNA molecules. As a result, thousands of copies of mtDNA can be found in each cell. The high amount of mtDNA molecules allows the success of mtDNA genotyping in degraded and/or low-quality samples, which is not always the case when using nuclear DNA (Amorim et al., 2019; Butler, 2011; Goios & Alvarez, 2013; Goodwin, 2016; Jain & Vishnu Priya, 2012).

Mutation rate

The mutation rate is higher in mtDNA than in the nuclear genome (Brown et al., 1979). Some of the reasons for this higher rate are the exposure of mitochondria to reactive oxygen species produced as by-products in oxidative phosphorylation (Hashiguchi et al., 2004), the lack of a mtDNA repair mechanism, and the low fidelity of mtDNA polymerase. It is worth noting that the mutation rate can vary depending on the tissue type, environmental exposures, and age. The mutation rate is not the same across the mtDNA molecule, with some regions being more prone to mutations than others. The

estimate of the mutation rate in the coding region is 0.017×10^{-6} substitutions per site per year. This value is 10 times higher in HVS-I and HVS-II of the D-loop (Ingman et al., 2000).

1.3.4. Human Mitochondrial DNA Variation

Analysis of mtDNA

The first studies on mtDNA sequences relied on restriction fragment length polymorphism (RFLP) analysis to identify variants, by employing endonucleases. This approach was classified as low-resolution analysis if mtDNA was digested by six enzymes, and as high-resolution analysis if it was carried out with 14 enzymes (W. M. Brown, 1980; Cann et al., 1984; Richards & Macaulay, 2001). Later, with the development of polymerase chain reaction (PCR), specific regions of mtDNA started to be amplified for sequencing or other downstream applications. The first sequencing studies were conducted by analysing HVS-I and HVS-II of the CR. Later, sequencing of the entire CR was routinely implemented, incrementing the resolution of the analysis.

Over the last decades, the Sanger sequencing method has been the most widely used sequencing technique for mtDNA typing and has been used for numerous applications namely disease diagnosis, forensic identification, and population genetics.

The Sanger sequencing uses a polymerase that incorporates fluorescently labeled dideoxynucleotides (ddNTP) chain terminators to stop the DNA synthesis at specific bases (Sanger et al., 1977). The ddNTPs lack a hydroxyl group in the 3’ end, which prevents the addition of further nucleotides. This process produces fragments of different sizes with one base difference in length (*Figures 5A and 5B*). The sequence of the original template can be rebuilt separating each fragment by size, using capillary electrophoresis. The separated fragments with the fluorescent labels incorporated are exposed to a laser beam at the end of the capillary, which emits a light detected by a photomultiplier tube. The signal obtained is then processed by specialized software to generate a DNA sequence (*Figure 5C*).

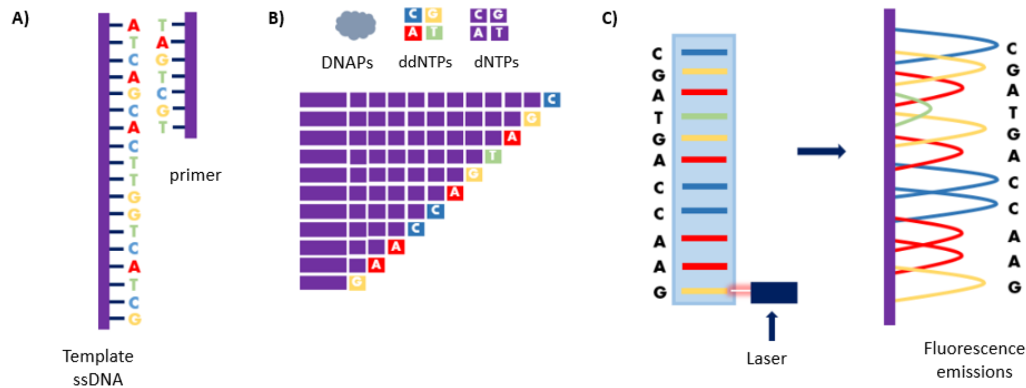


Figure 5. The basis of Sanger Sequencing. A) The amplified DNA fragment is denatured and bound to a single primer; B) The process occurs in a single reaction with the addition of fluorescently labeled ddNTPs, dNTPs, and DNA polymerase; C) The capillary electrophoresis separates the fragments by size and detects them by laser excitation [Adapted from Apolloinstitute.org].

Currently, new technologies have allowed to sequence the entire mtDNA genome more efficiently (Ingman et al., 2000; Irwin et al., 2011). New high-throughput sequencing called Next-generation sequencing (NGS) allows massive parallel sequencing (MPS) to generate a high number of sequence reads simultaneously. This process requires sample preparation where DNA is fragmented, adapters are added, and the resulting DNA fragments are amplified to create libraries. These libraries are loaded to an NGS instrument, where the fragments are sequenced in parallel through cycles of sequencing chemistry (Muzzey et al., 2015). Despite being a technique that allows obtaining a large amount of information in a simple way, it is still not yet routinely integrated in most laboratories due to the high cost.

Control region variation

The polymorphisms detected in the mtDNA control region are SNPs and InDels. The CR rich in sequence variation, containing some sites that behave as mutation hotspots (positions 146, 150, 152, 195, 16189, 16311, 16362, 16519), and others with high stability (positions 477, 493, 16108, 16219) (Galtier et al., 2006; Pereira & Gusmão, 2016; van Oven & Kayser, 2009).

The high mutation rate inherent to mtDNA can also generate different sequence variants that co-exist with the original mitotype. When more than one mtDNA type is found in one individual, the term heteroplasmy is used (Melton, 2004). This event occurs at low frequency and just one or few positions differ from the common type (Payne et al.,

2013). The presence and degree of heteroplasmy can vary among molecules in the same mitochondrion, among cells in the same tissue, among different organs within the same person, and among individuals in the same maternal line (J. B. Stewart & Chinnery, 2015). There are two types of heteroplasmy associated SNPs or InDels. The point heteroplasmy occurs when there are two or more variants at a single nucleotide position, showing an overlapping of peaks in the electropherogram (*Figure 6A*). The length heteroplasmy occurs when there is a difference in the length of similar sequences due to the insertion or deletion of one or more bases (*Figure 6B*). Length heteroplasmies are more commonly associated with InDels in poly-C stretches mostly located between positions 302-310 and 16183-16194 (Amorim et al., 2019; Ballard, 2016; T, 2004).

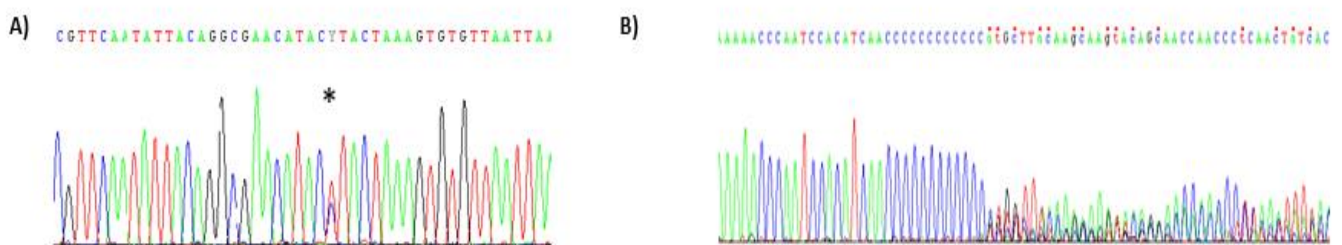


Figure 6. mtDNA sequences. A) An example of point heteroplasmy is marked by an asterisk. B) An example of length heteroplasmy.

mtDNA Haplogroups

Similar mtDNA haplotypes are grouped into haplogroups (*Figure 7*). Consequently, mitotypes that cluster to the same haplogroup share one or more distinctive polymorphisms from the ancestral mtDNA. Haplogroups are classified with letters of the alphabet, from A to Z, and were named in the order of their discovery.

The haplogroup nomenclature was initiated by Torroni et al. (1992) based on the analysis of *RFLPs*. At that time, the authors grouped the Native American mtDNA mitotypes in four basal branches, named A, B, C, and D (*Figure 7*) (Torroni et al., 1992; Wallace & Torroni, 2009). Over the years, studies on different populations around the world have contributed to the understanding of the distribution of mtDNA haplogroups and their geographic specificity. The PhyloTree (https://www.phylotree.org/) shows the phylogeny of the haplogroups and provides a comprehensive picture of worldwide human mtDNA phylogeny. All non-African haplotypes are derived from two macro-haplogroups, M and N. These two macro-haplogroups derived from the L3 macro-

haplogroup that originated in sub-Saharan Africa. Similar to L3, all the other L-clades have their origin in Africa, with different incidences all over the continent (Salas et al., 2002). Currently, some L-lineages are found at high frequencies in the American continent as a result of the transatlantic slave trade (Salas et al., 2002). European populations have mitotypes classified within the haplogroups T, U, V, W and, X (*Figure 7*) (Torroni et al., 1996). Current East Asian populations are composed of mtDNAs classified into haplogroups A, B, C, D, F, G, Y, and Z (Kivisild et al., 2002).

This pattern of mtDNA variation helped to dismiss the Multiregional Hypothesis for human origin and dispersion that pointed to human origin in different continental locations. The variability and genetic patterns found on mtDNA fit with an alternative hypothesis named “Recent African Replacement Theory” or “Out of Africa” which postulates that humans evolved as new species in sub-Saharan Africa and later moved out of Africa around 60 000 years ago (H.-J. Bandelt et al., 2006; Jobling et al., 2019; van Oven & Kayser, 2009).

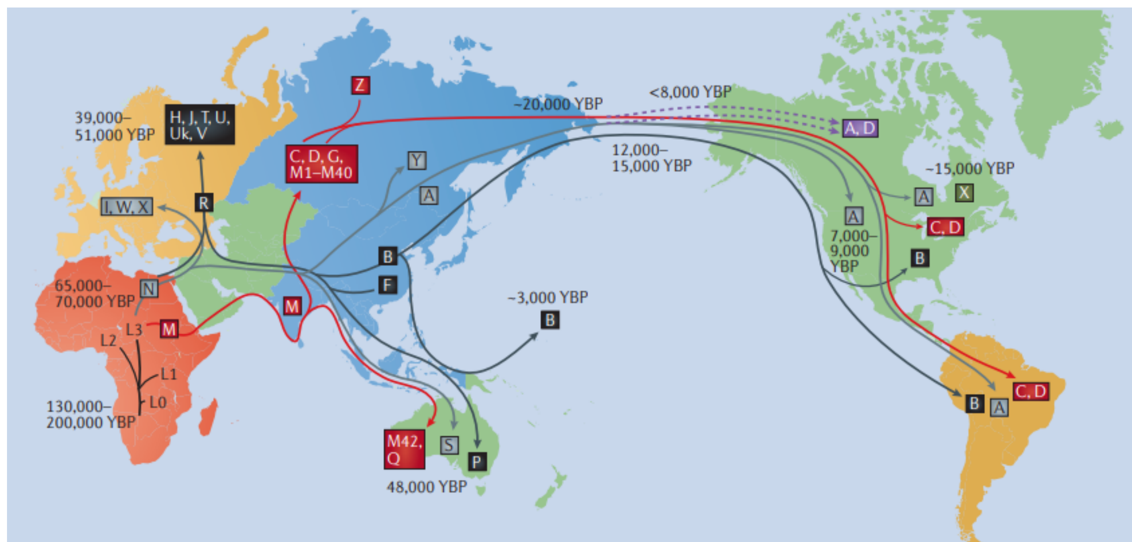


Figure 7. The distribution of the major mtDNA haplogroups explained by human migrations. YBP (Years Before Present) [Adapted from Stewart et al. (2015)].

1.4. The Americas human migration history

For decades, anthropologists, archaeologists, and geneticists have attempted to unravel the peopling of the Americas and its migration routes. The prevailing hypothesis suggests that Native American ancestors crossed Beringia, a land bridge that connected Northeast Asia and Northwest North America during the end of the Pleistocene (*Figure*

8) (Goebel et al., 2008; O'Rourke & Raff, 2010). Current recreation of sea fluctuations proved the presence of the Bering Land Bridge that linked the two regions when sea levels decreased and terrestrial ice-mass balance changed (Jakobsson et al., 2017; Lambeck et al., 2002). The theory on the existence of this land bridge immersed in the waters between Alaska (Norwest North America) and Chukotka (Northeast Russia) escalated with the discovery of mammoth remains from Pribilof and Aleutians islands (*Figure 8*) (Dall & Harris, 1892; Elias & Crocker, 2008; Hoffecker et al., 2016). The proof of existence of Beringia played an important role in the migration theories on routes and settlement of the first humans in the Western Hemisphere.

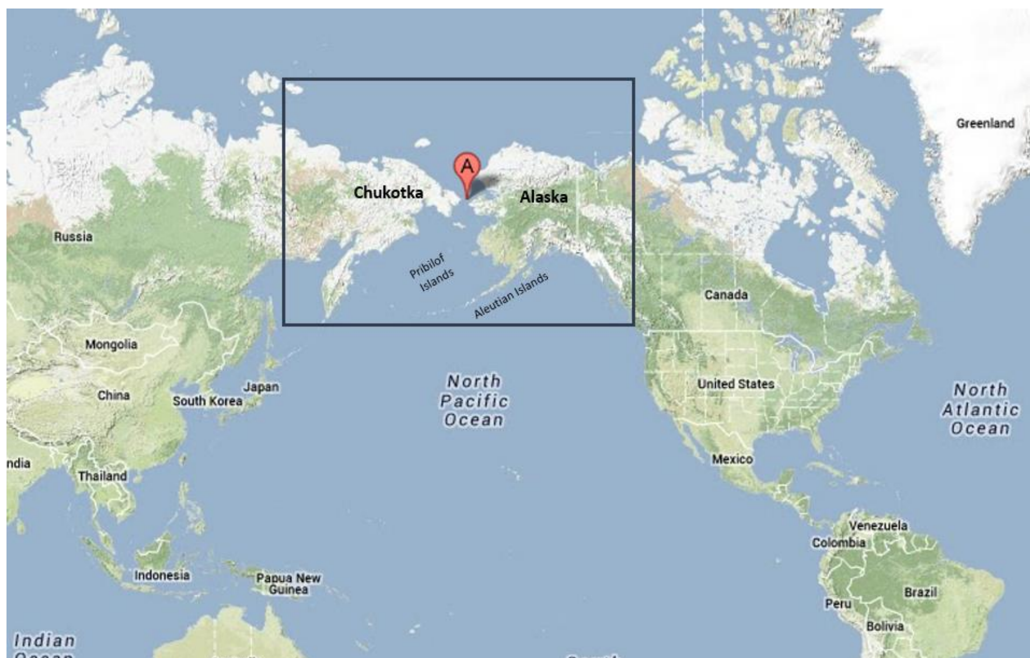


Figure 8. Map of Northeast Asia and Northwest North America, highlighting the Bering Strait (point A) [Adapted from Google maps)]

Ancient North Siberians were thought to be the first group of individuals to cross the Bering Bridge. This population probably moved across Siberia into Beringia and ultimately originated the basal American branch of people. Nonetheless, the time and place of the emergency of this branch is still discussed in the scientific community.

Several genetic studies suggest that the ancestors of Native Americans descend from this single founding population that experienced a period of isolation upon their arrival to the American continent, several millennia before their dispersal through North America (Hoffecker et al., 2014; Tamm et al., 2007). This initial isolation (described as the Beringian 'standstill model') possibly occurred in the region of the land bridge and led to the genetic differentiation of the New World settlers. This isolation led to the

emergency of new mtDNA and Y-chromosome genomes, differentiated from the Asian sister-clades. Lineages inside five maternal haplogroups (A, B, C, D, and X) and two paternal haplogroups (Q and C) have been identified in the American continent and helped to establish an ancestral link between the earliest Native Americans and Asian populations (Perego et al., 2009; Starikovskaya et al., 2005; Tarazona-Santos et al., 2001). After crossing Beringia, Native American populations in Northern North America underwent a split into two distinct branches with two migratory paths: the Northern Native American (NNA) branch and the Southern Native American (SNA) branch. The NNA branch was the ancestral component for present-day Salishan, Algonquian, and Na-Dené speakers from Canada. In contrast, the SNA lineages represent the ancestral component of contemporary populations in Central and South America (Moreno-Mayar, Potter, et al., 2018; Perego et al., 2009; Rasmussen et al., 2014; Scheib et al., 2018). The existence of these two migration paths was supported by several genetic evidence. The migration towards the South carried “pan-American” mtDNA haplogroups (A2, B2, C1, D1, and D4h3) (Schurr, 2004; Schurr et al., 1990; Torroni et al., 1993) and Asian Q-M3 and Q-L54 *(xM3) founding paternal lineages (Battaglia et al., 2013; O’Rourke & Raff, 2010; Zegura et al., 2004). The Northern migration carried X2a and C4c mtDNA haplogroups, and C-M130 and P-M45b Y-chromosomal haplogroups to North America (Kashani et al., 2012; Perego et al., 2009; Schurr & Sherry, 2004) Dulik et al., 2012). The close genetic links between ancient individuals living in North and South America in the same period supports a rapid southward movement of SNA populations (Willerslev & Meltzer, 2021).

1.4.1. Finding routes toward South America

The initial movement of Native South American (SNA) populations across the hemisphere resulted in repeated divisions within the SNA lineage as they traveled South, ultimately leading to significant variability among ancient South Americans (Moreno-Mayar, Vinner, et al., 2018; Willerslev & Meltzer, 2021). The genetic discontinuity between South Americans in the East and West of the Andes suggests that geographical and social barriers have contributed to the isolation of these populations. It is likely that the initial dispersal southward occurred along both sides of the mountain chain, and the challenge of traversing the East-West axis of the mountains has further maintained the separation over time (Gómez-Carballa et al., 2018; Nakatsuka et al., 2020; Reich et al., 2012; Rothhammer & Dillehay, 2009). Studies have revealed that populations in the

Eastern region, particularly Brazil, display lower levels of heterozygosity compared to their Western counterparts. This observation is consistent across analyses of both Y chromosome and mtDNA genomes, which confirmed that populations in the Western part of the Andes present a higher level of gene flow and larger effective sizes, leading to genetic homogenization. On the other hand, populations in the Eastern part exhibit higher rates of genetic drift and lower levels of gene flow, resulting in population differentiation. Furthermore, the examination of sub-haplogroups that are specific to certain geographic regions provides additional evidence for the hypothesis that the initial Paleoindian settlers of South America diverged into two main groups. These groups followed distinct migration routes along the Pacific and Atlantic coastlines (*Figure 9*) (Fuselli et al., 2003; Gómez-Carballa et al., 2018; Lewis et al., 2005; Tarazona-Santos et al., 2001).

Figure 9. Representation of two ancient human migrations into and within South America. One route heading the Pacific Coastline and the other heading the Atlantic Coastline. The black circles on the map represent important locations and the ages of ancient human genetic material found at those sites [Adapted from Willerslev et al. (2021)].

1.5. Ecuador

The Ecuadorian territory lies between the Pacific coast and the Amazonian plains in the West region of South America. The Andean Mountain range crosses the country

from North to South giving rise to valleys and rivers. Ecuador is bordered by Colombia in the North and Peru in the South and East. Politically, it is divided into 24 provinces distributed geographically into four natural regions: Coast or Coast Lowlands, Andes or Andean Highlands, the Amazon, and the Galápagos Islands Archipelago (*Figure 10*).

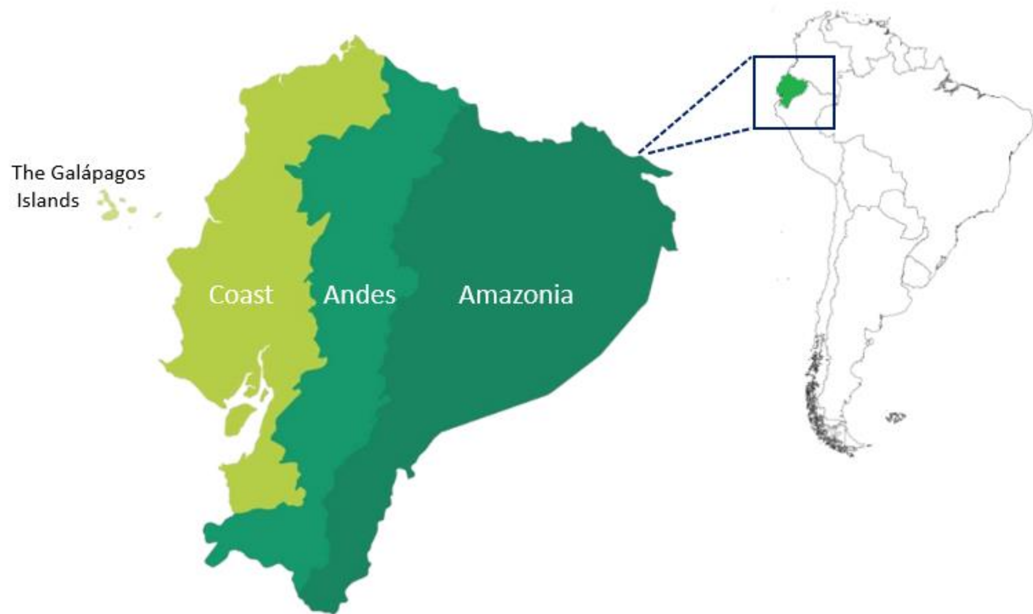


Figure 10. Illustration of Ecuador's geographic natural division (Mainland Ecuador: Coast, Andes, and Amazonia; and Insular Region: The Galápagos Islands).

1.5.1. The first settlers in Ecuador

In Ecuador, there is archaeological evidence of human settlements around 12 ka ago. The oldest vestiges such as lithic artifacts, stone axes, and arrowheads were found in the Andean valleys at El Inga, Chobsi, Jondachi, and Cubilán. On the Coast, the burial of the "Sumpa lovers" (a man and a woman embraced and covered with large stones) was discovered at Las Vegas in the Santa Elena Peninsula (*Figure 11A*). The first inhabitants were assumed to be nomadic hunter-gatherers that were organized in groups of approximately eighty people.

Later, with agricultural development, important civilizations were founded. Valdivia (*Figure 11B*) is the oldest agricultural culture in Ecuador, and it flourished in the Coastal lowlands in 3500 B.C. Over the years, the Machalilla and Chorrera cultures also emerged on the coast; Cerro Narrío, Alausí, Cotocollao on the Andes; and people of

Cueva de los Tayos in the Amazon (*Figure 11B*) (Almeida, 2000; Athens, 1980; Ayala, 2003; BORCHART de Moreno & Moreno Yánez, 1997). The agricultural societies allowed the beginning of the exchange of products, intensifying contact between people from different regions.

Over time, the most advanced cultures created social clusters: the farmers, which oversaw agricultural activities, and the chiefs or *caciques* that directed wars and trade. Moreover, the *llajtacunas* rulers exercise authority over the mentioned social clusters. Through the union of minor *caciques*, *cacicazgos*[1] were created. Some of the *cacicazgos* in the North of the Andes were Pastos, Quillacinga, and an important group of inter-connected lordships: Cochasquí, Caranqui, Otavalo, and Cayambe (Figure 11C). In the South of the Andes, Pillaro, Sigchos, Puruhá, and Cañari were established (Figure 11C). La Tolita, Manteño, Huancavilca, Punáes, and Chonos were settlements on the Coast. Towards, the Amazon lived the Quijos, Cofanes, and Jíbaros peoples (*Figure 11C*).

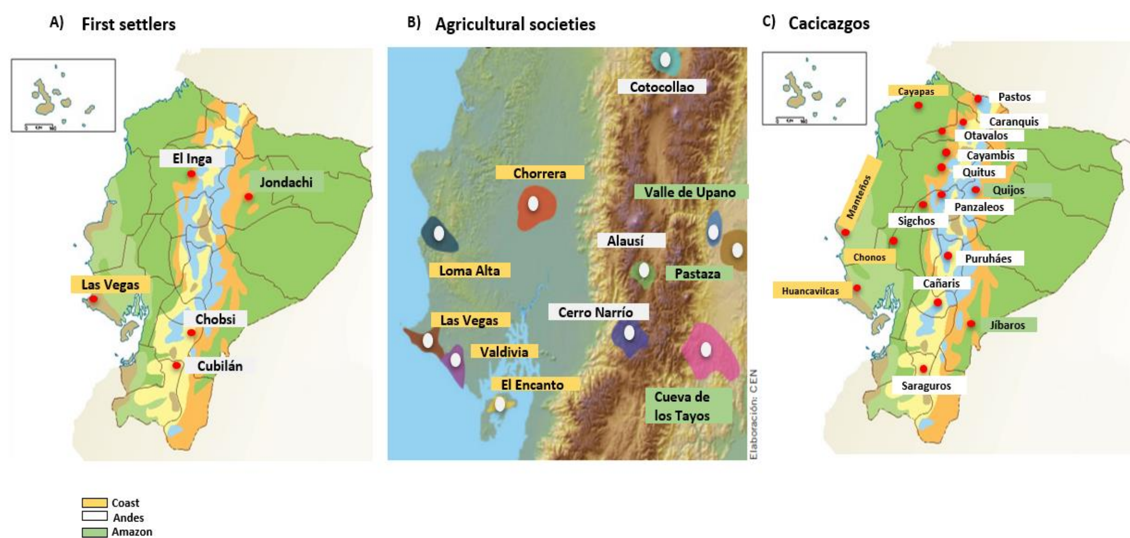


Figure 11. The Pre-Columbian settlements in different Ecuadorian regions. A) First human settlements in Ecuador; B) Agricultural society settlements; C) Ethnic Lordships. Coast settlements are marked in yellow; Andean settlements, in white; and Amazon settlements, in green [Adapted from Ayala Mora (1989)].

1.5.2. Inca Empire

At the end of the 15th century, Andean civilizations faced the conquest of the Incas, a group originally established in the South of current Peru. The Incas started their

[1] A cacicazgo can be described as a form of indigenous chiefdom or regional leadership structure that existed prior to European colonization.

expansion into the North and created the largest empire in South America called the *Tahuantinsuyo*. Their territory was delimited in the South by the Bolivian Plateau, Argentina, and the Maule River in Chile, and in the North by Ecuador and the South of Nariño province in present-day Colombia.

The conquest of Ecuadorian territories by the Incas started in the South Andean region. The Inca emperor used a combined tactic of military action against the tribes. *Paltas* and Cañaris, Native populations that were located in the South of Ecuador (present-day Cuenca city) were dominated with almost no resistance. The Incas later advanced in the North direction and attacked the Native people of the Center of the Andes. In the North of the Andes, the Natives fought fiercely against the Inca invasion for approximately fifteen years.

The Incas also made incursions to the Ecuadorian Coast but with less domination of the local groups. It seems they controlled Puna Island and had a maritime alliance with *Manteños*. Towards the Amazon, the invasions did not take place. The Inca Empire had an enormous influence on Ecuadorian culture, with local people adopting their social and political elements. The Inca invasion persisted for a hundred years and ended with the arrival of Spaniards in Ecuadorian territory (Espinoza Soriano, 1987; Rostworowski, 2014; Salomon, 1980).

1.5.3. Spaniard Colonization

The arrival of Christopher Columbus in America in 1492 marked the beginning of a new era for the American continent. In the middle of 1534, the Spanish arrived at Quito encountering the Inca Empire ruled by Atahualpa. Initially, the Inca resistance against Spanish invaders was fierce. However, the opposition did not last long as the Inca Empire was already weakened by the previous internal crisis and civil war between their Inca leaders (Ayala, 2003; Parry, 1968). Through a combination of military force and alliances with local indigenous groups who were opposed to the Inca ruler, the Spanish were able to gain control over the Incas. At the time of the Spanish invasion, several ethnolinguistic groups inhabited the Ecuadorian regions. The Quechua-people, one of the most widespread indigenous groups, lived throughout the Andes, and other groups like Shuar, Achuar, and Huaorani lived in the Amazon rainforest (Blum et al., 2023; Enrique Ayala Mora, 2002). Although some Native communities remained isolated and

away from the Europeans, preserving, until today, their language and costumes, many others disappeared during this period.

Similar to what happened in almost all European colonies in America, enslaved Africans were taken to Ecuador as forced laborers in mining, agriculture, and domestic services. They were kidnapped from several African regions, brought on ships to the Caribbean, and sold as merchandise. Many escaped enslaved Africans found refuge in Esmeraldas (the Northern coast of Ecuador). The Pacific coast of Esmeraldas witnessed frequent shipwrecks involving black captives, and those who survived would then enter this region. In the Andes, the largest concentration of enslaved Africans was in the Chota valley in the Northern Andes (Andrien, 1995; Bossano, 1974; Rueda, 2001). Although historic records on the origin of the enslaved Africans that arrived in Ecuador are scarce, some evidence indicates that they were taken from different parts of West and Central Africa, reflecting the broad scope of the transatlantic slave trade (Enrique Ayala Mora, 2002; García Salazar, 1989).

The Spanish occupation in Ecuador was predominantly focused on the Andean region, where a significant number of indigenous people were subjugated. However, certain indigenous populations sought refuge in the Amazon rainforest and the coastal region of Esmeraldas. Apart from Esmeraldas, indigenous populations from the Pacific coast were largely devastated by the European settlers, resulting in significant losses. Indigenous groups in the Amazon region were not so affected by the European invasion, due to their isolated location.

1.5.4. Current Ecuador: geography and demographics

Ecuador was home to different human communities that have inhabited the territory for thousands of years (Paz-y-Miño & Burgos, 2015). Currently, it has a population of around 18 million individuals, grouped into five major categories according to self-identification: Mestizos (77.4%), Afroecuadorians (4.9%), Montubios (7.7%), Indigenous (7.7%), Whites/Others (2.3%) (INEC-Instituto Nacional de Estadisticas y Censos, 2022). The mestizos represent the majority of Ecuadorians and are individuals who have predominantly Native American and European admixture. The sex-biased admixture between Spanish men and Native American women during the Spanish colonization is still reflected in the current Ecuadorian population (González-Andrade et al., 2007). In addition, the Montubios is a recently recognized ethnic group represented by countryside Mestizos scattered in the Coastal region. The Afro-Ecuadorians

communities, present in the rural areas of the Northern provinces of Esmeraldas were built mostly by runaway enslaved Africans who sought refuge away from European colonizers. Contrarily, in Imbabura the African community was created by enslaved individuals introduced there for agricultural work (Moya Torres, 1997; Vásquez S. et al., 2013). Fourteen different Native ethnic groups are currently recognized by the Ecuadorian government and are scattered across the country, despite having a higher population density in the Andes region. The most representative group is the Kichwa (Enrique Ayala Mora, 2002).

1.5.5. Genetic Diversity of Ecuador

Some studies using uniparental and biparental genetic markers have been performed to characterize the genetic diversity in Ecuadorian populations. Data on autosomal ancestry informative markers (AIMs) confirmed the trihybrid composition of self-declared Mestizos, with a major contribution of Native American followed by European ancestries. Compared to other South American countries, low African ancestry has been reported in Ecuadorian Mestizos (Nagar et al., 2021; Santangelo et al., 2017; Zambrano et al., 2019)

For the Y chromosome, most studies were focused on Mestizo populations. Baeza et al. (2007) analysed 15 Y-STRs in a sample of 120 individuals from the city of Quito, reporting statistical parameters of forensic relevance. Gaviria et al. (2013) reported 11 Y-STR profiles from 206 Ecuadorian individuals, without specifying their ethnolinguistic affiliation. The most recent study by Toscanini et al. (2018) reported 415 haplotypes of Mestizo individuals from three Ecuadorian regions, using 23 Y-STRs. In addition to Mestizos, a study from González-Andrade et al. (2009) included a sample of Kichwas, Afro-descendants and Waoranis, reporting 12 Y-STRs. Likewise, some studies were conducted with Native American individuals to uncover lineages and reveal phylogenetic characteristics of the most common haplogroup found, denoted as Q. A study by Roewer et al. (2013) focused on a sample of Waoranis, Natives of the Pastaza province, reported haplogroup C2, which is highly characteristic of this population (Jota et al., 2016; Mezzavilla et al., 2015; Roewer et al., 2013; Villaescusa et al., 2021). The studies exclusively involving Ecuadorian mestizo populations report a significant contribution of European chromosomes, ranging from 60% to 70%. This is followed by a contribution of Native chromosomes around 25% to 30%, and a recent study reported nearly 40% Native chromosomes (Burgos et al., 2019). Finally, a percentage of 4 to 6% of African chromosomes has been found.

Investigations on Ecuadorian maternal ancestries were focused on HVS-I and HVS-II in Native American communities, namely Cayapa (Rickards et al., 1999), Shuar (Leone et al., 2019), Kichwa (Baeta et al., 2012), and Waorani (Baeta et al., 2009; Cardoso et al., 2012). Differences in the distribution of the Native haplogroups A, B, C, and D were encountered among these populations. The absence of lineages outside haplogroup B in the Shuar population, pointed to a high isolation of these individuals. The same was observed for Waorani Native groups, where 94% of the mitotypes belonged to haplogroup A2. Higher diversity was observed in Cayapa and Kichwa Natives, with samples distributed among several Native macro haplogroups.

More recently, studies on Mestizo individuals from Ecuador revealed that more than 90% of the population has a Native American maternal background. Although some studies include data for the entire mtDNA control region (Baeta et al., 2012; Burgos et al., 2019), or for mitogenomes (Brandini et al., 2018), the low number or samples typed and the lack of information on the geographic origin of the samples, prevent a deeper understanding of the maternal inheritance of the mestizo population of Ecuador.

2. Aims

The admixture events between Europeans, Native Americans, and Africans occurred differently across the three natural regions of Ecuador Mainland. Presently, limited information is available regarding the variability of maternal lineages within the Ecuadorian territory, particularly in the Andean region that harbours numerous Native American groups. The primary objective of this research is to uncover the maternal genetic composition of the unexplored Andean region of Ecuador and to infer how demographic and historical processes have contributed to the diversity of Mestizo people. The analysis will focus on mitochondrial DNA to trace maternal lineages, providing insights into population origins, migrations, and genetic interactions.

To achieve the main aim of this study, the goals proposed were:

- Determine the maternal lineages in Mestizo populations living in the three Andean region – North, Central, and South -, through the analysis of the mtDNA control region;
-Evaluate if there is a genetic differentiation or stratification among the three Andean subpopulations;
- Understand the genetic composition of the Ecuadorian Andes in the context of other South American admixed and Native populations.

3. Methods and Materials

3.1. Population samples and DNA extraction

Blood or mouth swabs samples were collected from 214 unrelated and self-identified Mestizo individuals from three Ecuadorian Andean regions: North (n=76), Center (n=66), and South (n=72) (*Figure 12*). All participants signed an informed consent approved by the Ethics Committee in Human Research from the Universidad de las Américas, Quito-Ecuador (CEISH-UDLA 2017-0301).

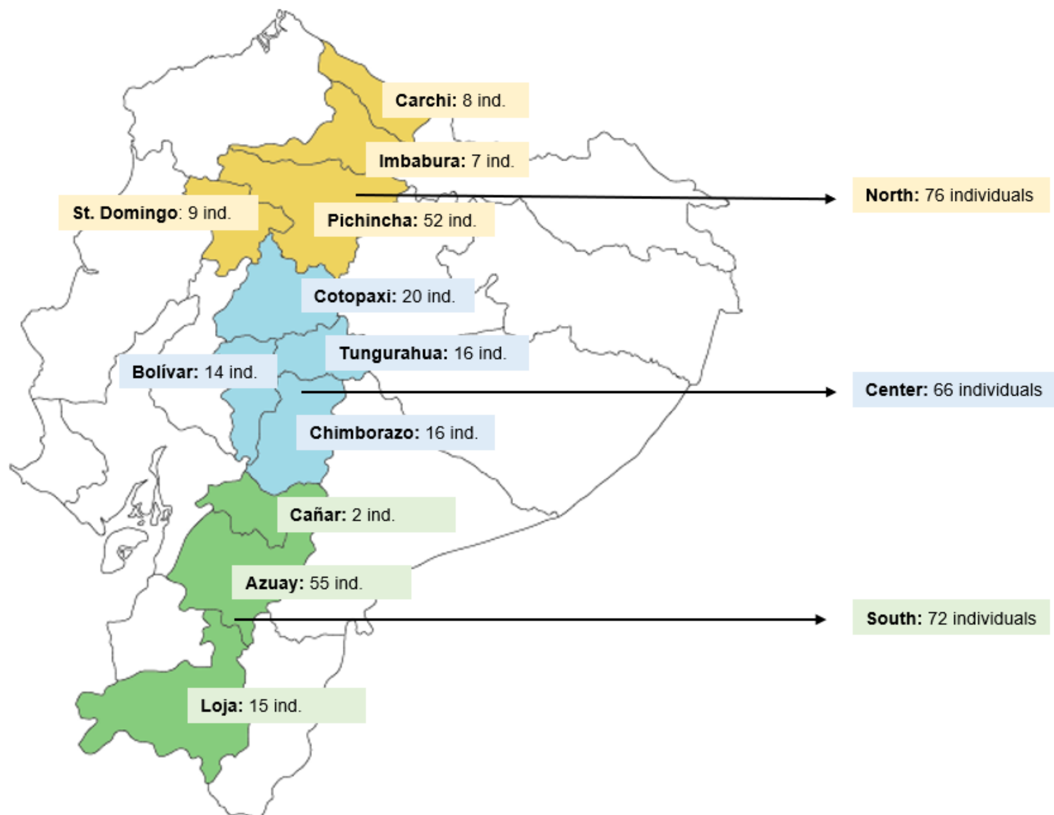


Figure 12. Geographical distribution of the Ecuadorian samples from the Andes used in the present study. Northern provinces are represented in yellow, Central in blue, and Southern in green.

The DNA extraction was conducted at *Laboratorios de Investigación of Dirección General de Investigación y Vinculación* (DGIV) of the *Universidad de las Americas*, Quito, Ecuador (DGI), using the salting-out method described by Miller et al. (1988), for blood, and the phenol-chloroform protocol described by Sambrook et al. (1989), for

swabs. The quantification of DNA was performed using a NanoDrop™ 2000/2000c spectrophotometer from Thermo Fisher Scientific, and the samples were subsequently diluted to a concentration of 5ng/µL. DNA samples were sent to the *Instituto de Investigação e Inovação em Saúde* (i3S), to carry out this collaborative work.

3.2. Mitochondrial DNA analysis

The control region of mtDNA (16024-576) was amplified by PCR and sequenced using a fluorescent dye terminator sequencing technology (Sanger et al., 1977) following the steps described in *Figure 13*.

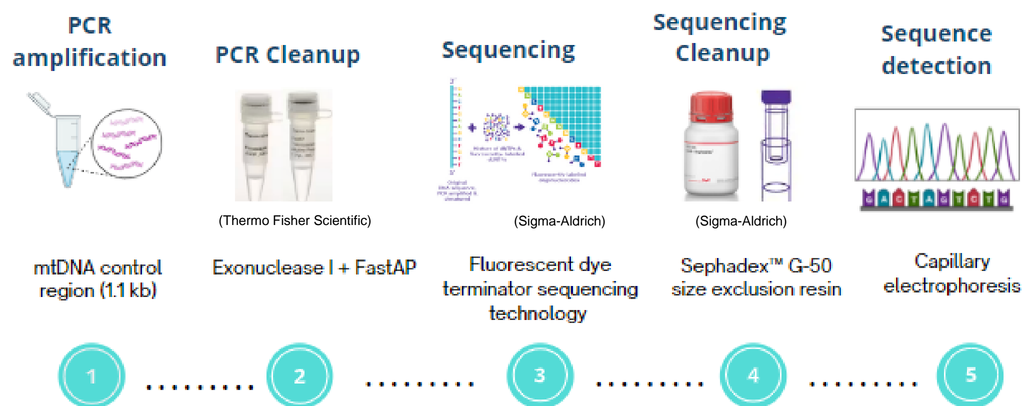


Figure 13. The workflow of mitochondrial CR typing.

3.2.1. PCR Amplification

The full mtDNA control region was amplified by PCR in a final volume of 10 µL. The reaction mix contained 1 µL of primer mix (primers L15978 and H649 at 2 µM) (*Appendix 1*), 5 µL of 2X QIAGEN Multiplex PCR Master Mix (© QIAGEN), 0.5 µL of extracted DNA (5ng/µL), and 3.5 µL of sterilized water to complete the final reaction volume. A negative control was included in each PCR set.

The PCR was carried out in a T100 Thermal Cycler (Bio-Rad Laboratories, Inc). Amplification started with an initial denaturation at 95 °C for 15 minutes, followed by 35 cycles of 3 temperature steps: denaturation (94 °C for 30 seconds), annealing (58°C for 90 seconds), extension (72°C for 90 seconds). A final extension was conducted at 70°C for 10 minutes.

The amplification of the fragment of interest was confirmed by electrophoresis in a polyacrylamide gel matrix (T9%, C5%). Each well was loaded with 1.2 μL of the amplification products and, in each gel, a GRS 100bp DNA ladder (GRiSP, Lda.) was included for size estimation. DNA fragments were visualized in the gel by a silver staining method described by Budowle et al. (1991) with slight modifications on the ethanol and nitric acid solution timing (10 minutes on ethanol instead of 5 minutes and 5 minutes in nitric acid solution instead of 3 minutes). Only samples with positive amplification, resulting in the presence of a band with the expected size, proceeded to the next step.

3.2.2. Enzymatic PCR cleanup

Alkaline phosphatase (FastAP) and Exonuclease I enzymes (Thermo Fisher Scientific) were used to purify PCR products by removing the remaining dNTPs and primers left on the reaction. FastAP dephosphorylated the dNTPs, and Exonuclease I degraded the residual primers. In this purification step, 2 μL of a mixture of Exonuclease I and FastAP (1:4) was added directly to 5 μL of the PCR product. Reaction was incubated at 37°C for 30 minutes followed by a 15 minutes at 85 °C for enzymatic inactivation.

3.2.3. Sanger Sequencing

The Big-Dye Terminator v3.1 kit (Applied Biosystems®) was used to perform cycle sequencing of the amplified fragments. Independent reactions in the forward and reverse directions were set up for each sample. As a sequence strategy, forward primers L15978 and L16555 were used first. Reverse primers were further employed in samples that needed complementary sequence information due to length heteroplasmy in polycytosine stretches (Irwin et al., 2009). The primer H017 was used when heteroplasmy was present in the HVS-I C-stretch, mostly between 16184-16189 positions, and primer H649 for sequences with length variants between 303-309 and 311-315 positions in HVS-II. Another poly-C length variant spanning between positions 568-573 results in the failure of H649 sequences, so primers L314, H408, H484, or H309 were used in these cases. Primers L16209 and H159 were used to fill a gap between the end of HVS-I and the beginning of HVS-II regions (*Figure 14*).

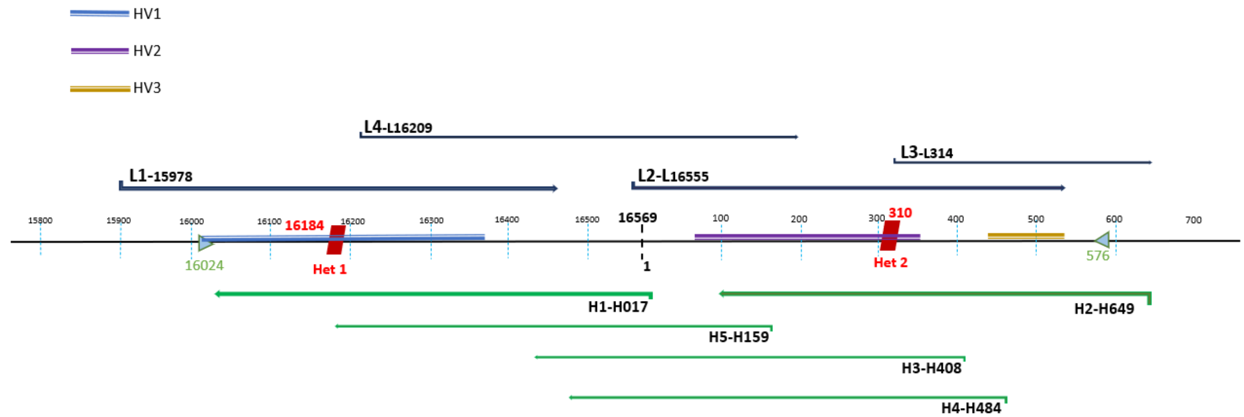


Figure 14. Diagram with the primers used. Primers are named "L" due to the interaction with the light chain and H primers with the heavy chain of mtDNA. For each primer, the mtDNA CR region that is covered is represented. The most common heteroplasmic regions are shown in red.

Reactions were prepared in a final volume of 5 µL, containing 0.5 µL of the primer at 2.5 µM, 0.8 µL of Big Dye Terminator v3.1 Ready Reaction Mix, 1 µL of 5X Sequencing Buffer, 1.2 µL of deionized water, and 1.5 µL of the PCR purified product. The reactions were carried out in a T100 Thermal Cycler (Bio-Rad Laboratories, Inc). The sequencing reaction started with an initial denaturation at 96°C for 2 minutes, followed by 35 cycles of 3 temperature steps: denaturation (96 °C for 15 seconds), annealing (55°C for 9 seconds), extension (60°C for 2 minutes). Finally, an extra extension step was conducted at 60°C for 10 minutes. Adjustments to the sequencing protocol were performed for the reaction using the H639 primer (38 instead of 35 cycles, 56°C instead of 55°C annealing temperature, and 58°C instead of 60°C extension temperature).

3.2.4. Sequencing products cleanup and electrophoresis

A clean-up step was performed to remove unincorporated dNTPs and ddNTPs, and salts that can disturb the base calling. A total of 5 µL of sequence products were purified using columns filled with 750 µL of Sephadex™ G-50 size exclusion resin at 10% (GE Healthcare). The purified products were mixed with 8 µL of Hi-Di™ formamide (Thermo Fisher Scientific) before the separation and detection by capillary electrophoresis on an ABI 3500 Series Genetic Analyser (Applied Biosystems®). Samples setup on the genetic analyser was done by the personal at the genomics platform-GenCore at i3S.

3.3. Data Analysis

The sequence output files (.ABI) generated by the genetic analyser were imported using Sequencing Analysis Software 6 (Applied Biosystems®) to determine the quality of the sequences. Afterward, all the sequences for each sample were aligned and compared with the rCRS (Andrews et al., 1999) using the SeqScape® Software 3 (Applied Biosystems®). The differences between the sequences and the rCRS (insertions, deletions, or substitutions) were exported by the Mutation Report tool and each haplotype was further assigned into an haplogroup using EMPOP database (Parson & Dür, 2007).

The Haplosearch tool (Fregel & Delgado, 2011) was used to convert haplotype data into DNA sequences. This tool also aligned all sequence data in order to perform further analyses. The Arlequin v. 3.5.2.2 Software (Excoffier & Lischer, 2010) was used to calculate statistical parameters, namely the number of shared and unique haplotypes, haplotype diversity, mean number of pairwise differences, and the number of segregation sites (Excoffier & Lischer, 2010). The same software was used to calculate pairwise F_{ST} genetic distances and to perform analyses of molecular variance (AMOVA) using haplotype frequencies. Genetic distances were plotted in Multi-Dimensional Scaling (MDS) using the Statistica 10 software (TIBCO Software Inc.). This software was also used to compute Principal component analysis (PCA). Median-joining networks were designed using the Network v5.0 software (http://www.fluxusengineering.com).

4. Results and Discussion

The haplotypes and corresponding haplogroups for the 214 Andean Mestizos samples considered in this study are available in *Appendix 3*.

4.1. Mitochondrial DNA sequence variation in Andean Mestizos

Within the 1.1kb fragment of the mtDNA control region analysed, a total of 199 polymorphic sites was detected. Among North Andes haplotypes, 144 polymorphic sites were identified, while 119 were observed for the Center, and 133 for the South haplotypes. The majority of the polymorphisms were substitutions, although a significant proportion of deletions and insertions in Poly-C regions was also detected.

To effectively check the quality of mtDNA data a network was built using the quasi-median network tool from EMPOP (Parson & Dür, 2007). The filter EMPOPall_R12 was used to remove all the mutations previously described and deposited in the database so that only yet unobserved mutations would be highlighted. Most of the haplotypes were grouped in a single cluster, except for 8 haplotypes with variants not previously reported on the EMPOP database (*Appendix 4*), all of Native-American ancestry. These positions were double-checked and confirmed to be in fact polymorphisms. The absence of these new variants on the EMPOP database does not necessarily indicate that they have not been previously reported. However, it highlights the under-representativeness of Native lineages in publicly available databases and the importance to increase the study of South American populations.

A high frequency of heteroplasmies was observed in our data, a phenomenon that warrants attention in the interpretation of mtDNA sequences since it has a potential impact in genetic variation, inheritance patterns, and evolutionary insights. Indeed, the frequency of heteroplasmy varies across populations, which could be attributed to distinct mutation rates among different maternal lineages or specific group-related artefacts (Irwin et al., 2009).

Point heteroplasmy (PHP) was identified in approximately 3% of Central Andean individuals (n=2), and 7% of South Andean individuals (n=5) (*Appendix 5*). No instances of point heteroplasmy were encountered in the Northern Andean population. The sites

with PHP were located within the HVS-I and HVS-II. Some of these sites correspond to previously reported hotspots in the HVS-II region, namely at positions 152 and 195 (Irwin et al., 2009; Soares et al., 2009; Melton, 2004). These heteroplasmies have been previously reported in several populations with no geographic association. Heteroplasmies were also observed outside mutational hotspots, at positions 185, 214, 568, and 16399. A search on “EMPOP SNP query tool” showed that two of these PHP were only reported in a specific haplogroup. Namely, 568Y was only observed inside haplogroup J1c8a, and 16399R in haplogroup U6a7a. In this study, however, 568Y and 16399R heteroplasmies were detected in haplotypes belonging to the Native haplogroups A2+(64), and D4h3a6, respectively, showing that they are not haplogroup specific. Their absence in EMPOP may be explained by the lack of representativeness of Native American haplogroups and/or populations on this database.

In HVS-I region, length heteroplasmy (LHP) was detected with frequencies of 42% in the North, 48% in the Center, and 63% in the South. Previous studies suggested that the frequency of LHP may be correlated with the number of cytosines present in the poly C-stretches (Melton, 2004). In the HVS-I region, transversions at positions A16182C, A16193C, a transition at position T16189C, together with C insertions at position 16193 can lead to the formation of stretches with more than 12Cs, often resulting in LHP. West Eurasian populations typically exhibit a frequency of 14% for LHP, when there is a transition at position 16189. However, populations in which the T16189C substitution is a mutation signature of a specific haplogroup may display higher frequencies of LHP. The high occurrence of LHP in our population is explained by the prevalence of haplogroup B, which has this mutation signature (Parson et al., 2014) (*Appendix 6*).

An even higher percentage of LHP was observed in the HVS-II region, with frequencies of 55% in the North, 63% in the Center, and 75% in the South. In HVS-II, rCRS has a poli-C fragment between positions 305 and 315, which is defined by 12 Cs interrupted by a T at position 310. Nonetheless, most mitotypes have C insertions before and/or after 310T position. In most cases, this leads to some degree of LHP (Cavelier et al., 2000). Generally, LHP is more frequent and more pronounced when there are two or more C insertions after position 310 making it challenging to interpret the sequence downstream.

Additionally, low frequencies of LHP have been reported in AC repetitions at positions 523-524, and in the C-stretch at positions 568-573. In certain populations, when the dinucleotide AC achieved 5 or more repeat units (Chung et al., 2005; Szibor et al.,

2007) LHP was barely observed. Pronounced heteroplasmy was noticed as the numbers of repeats increased. In our study, we found LHP in one individual from the North and one from the Center with five AC repeat units. Regarding the C-stretch at positions 568-573, LHP was observed in our samples when there was an insertion of one or more cytosines at position 573, consistent with findings from other studies where LHP was detected with C-stretch beyond 6 Cs (Brandstätter, Peterson, et al., 2004).

It is worth mentioning that the percentage of heteroplasmy can vary depending on the tissue type, as well as with the sequencing methodology (Parson et al., 2014). In previous studies, a variation in the frequency of LHP was observed within the same individual between saliva, blood, and hairs samples, with each tissue displaying a different number of cytosines (Naue et al., 2015; Stewart et al., 2001).

The presence of LHP can potentially lead to interpretation problems when analysing mtDNA profiles. Therefore, laboratories must establish internal guidelines for representing and reporting LHP. For standardization purposes, the International Society for Forensic Genetics (ISFG) recommends reporting the dominant LHP variant (Berger et al., 2011; Parson et al., 2014). In this study, complete mtDNA CR data, including heteroplasmic positions, was used to assess the diversity of our populations. However, for population comparisons, InDels at positions 309, 315, 523, 524, 573, and 16196 were excluded to prevent differences due to the criteria used in assigning these positions or due to the use of sequencing methodologies with different heteroplasmy detection limits.

4.2. Genetic Diversity of Andean Mestizos from Ecuador

In the three subpopulations studied, large numbers of unique haplotypes were observed: 74 in the North (97.4%), 62 in the Center (93.9%), and 66 in the South (91.7%). The absence of founder effects was confirmed by high haplotype diversities in the three subpopulations (*Table 1*). However, it is noteworthy that the Southern region exhibits the lowest haplotype diversity among all regions.

Similar high diversity values were reported for other admixed populations from South America (*Table 1*), namely in the Andean region (Norte de Santander, Santander, and Cundinamarca) of Colombia, in two Southeast Brazilian cities and the Argentina North, Center and South divisions. High values of diversity are expected in South America as a result of the multiple historic migration waves, including the initial peopling

of the Americas and the subsequent arrivals of Europeans and Africans that introduced new genetic lineages and increased overall diversity. The diversity values found in our admixed populations were higher than that reported in Kichwas from Ecuador, a Native American population from Amazonia. This is also expected since Native American populations were usually subjected to drift events caused by low population sizes, and restricted geneflow due to isolation, both promoting endogamy.

Table 1. Molecular diversity indexes for mtDNA control region data from the present sample and other South American populations

Population	n	K	H	Study
Andes North	76	74	0.9993 +/- 0.0023	This study
Andes Center	66	62	0.9981 +/- 0.0030	This study
Andes South	72	64	0.9961 +/- 0.0033	This study
Paraná_Brazil	122	108	0.9976 ± 0.0016	(Poletto et al., 2019)
Rio de Janeiro_Brazil	205	184	0.9994 ± 0.0006	(Simão et al., 2018)
North_Argentina*	97	82	0.9947 +/- 0.0028	(Bobillo et al., 2010)
Center_Argentina*	193	173	0.9977 +/- 0.0012	(Bobillo et al., 2010)
South_Argentina*	47	43	0.9954 +/- 0.0059	(Bobillo et al., 2010)
Andes_Colombia	67	54	0.9932 ± 0.0039	(Castillo et al., 2019)
Kichwa Amazon_Ecuador	65	13	0.8029±0.0373	(Baeta et al., 2012)

n=Sample size
k=Number of unique haplotypes
H=haplotypic diversity
*Molecular diversity indexes were recalculated using the haplotype data from the study

4.3. Haplogroup Composition

The presence of a wide range of distinct genetic lineages from diverse continental origins within the present Andean populations corroborates the complex history of ancestral contributions and genetic admixture. This is illustrated by the high levels of haplogroup diversity identified within each population (*Figure 15*).

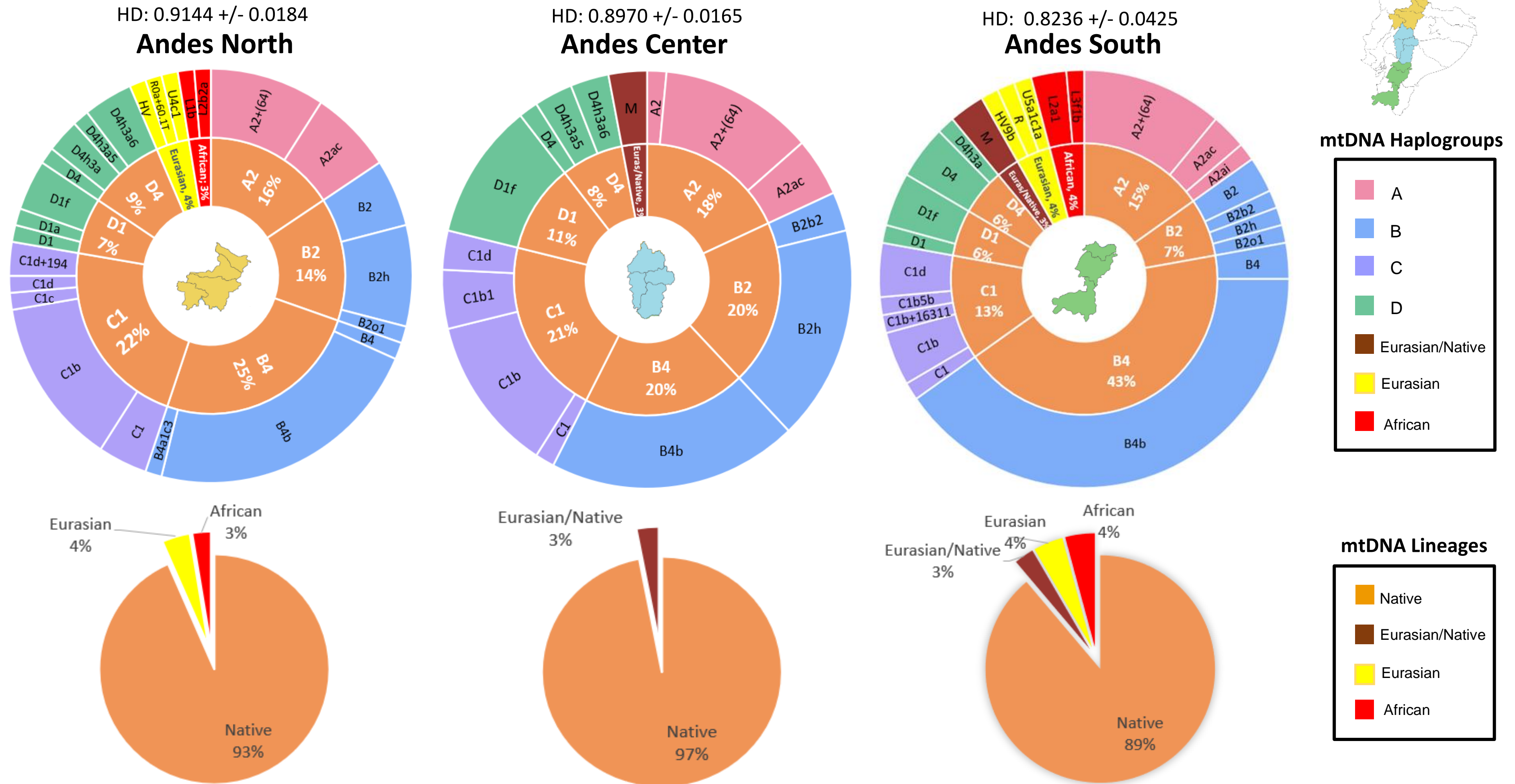


Figure 15. Representation of mtDNA haplogroups and sub-haplogroups found in each Andean populations from Ecuador. HD means Haplogroup Diversity.

The majority of the mtDNA haplogroups found in our sample were of Native American origin, accounting for 93.5% in the North, 97% in the Center and 88.8 % in the South. The prevalence of Native American lineages in the Central Andes was expected, considering that it is in this area that most Native American communities are established, followed by the North Andes region. Native American lineages were represented by haplogroups A2, B2, B4, C1, D1, and D4. However, differences in the frequency distributions of the Native lineages were observed among the three Mestizo subgroups (*Figure 15*).

A notable disparity was evident in the proportions of B4/B2 haplogroups among the three population samples. The North and Central Andes regions displayed a similar haplogroup distribution, whereas the South Andean region exhibited a substantial prevalence of B4 (43%) compared to B2 (7%). The high frequency of B4 in the Andes South could be behind the lower genetic diversity within this population (*Table 1, Figure 15*). An elevated frequency of a particular haplogroup may be due to recent shared heritage among individuals, whenever they have the same or close haplotypes. A median-joining network constructed with all the haplotypes found in Ecuadorian samples showed a high number of shared haplotypes within haplogroup B (*Figure 16*).

African (L1, L2, L3) and Eurasian haplogroups (U4c1, HV, R) were encountered at low frequencies in North and South Andes regions but they were absent in the Central Andes. As previously stated, the Central region of Andes comprises the greater number of Native communities, and a low prevalence or absence of non-Native lineages was expected. The occurrence of African haplogroups in Andes North could be anticipated due to the presence of African settlements historically dating back to the colonial period (Moya Torres, 1997; Vásquez S. et al., 2013).

Deciphering whether the trace of Eurasian haplogroups in Ecuadorian populations are remnants from the colonial era, or the outcome of World War II migrations, namely the migration of Jews from Germany, Austria, Italia, and Czechoslovakia in the XX century (Gil Blanco & Canela Ruano, 2018) or more recent migrations, is a challenging task.

It is noteworthy mentioning that Central and South Andes had two haplotypes assigned to Eurasian macro-haplogroup M, which could be further classified either as M33, D1 or C1 sub-haplogroups, according to the EMPOP (*Figure 15*). These haplogroups have a very distinct global distribution. While haplogroup M33 is more frequent in East Asia with low incidence in North America, haplogroups D1 and C1 are found scattered across the American continent, reaching the highest frequencies in

South America. Because these haplotypes could not be classified into a specific terminal branch, they were assigned to their most recent common ancestor (MRCA), which in this case was the macro-haplogroup M. The median-joining network (*Figure 16*) showed that the M haplotypes from Center and South clustered within the D and C haplogroups. These results, together with the absence of historical records of Asian migration to Ecuador, point to a probable origin of these lineages in South America. Nonetheless, an analysis of the full mitogenome would help to accurately infer the haplogroups.

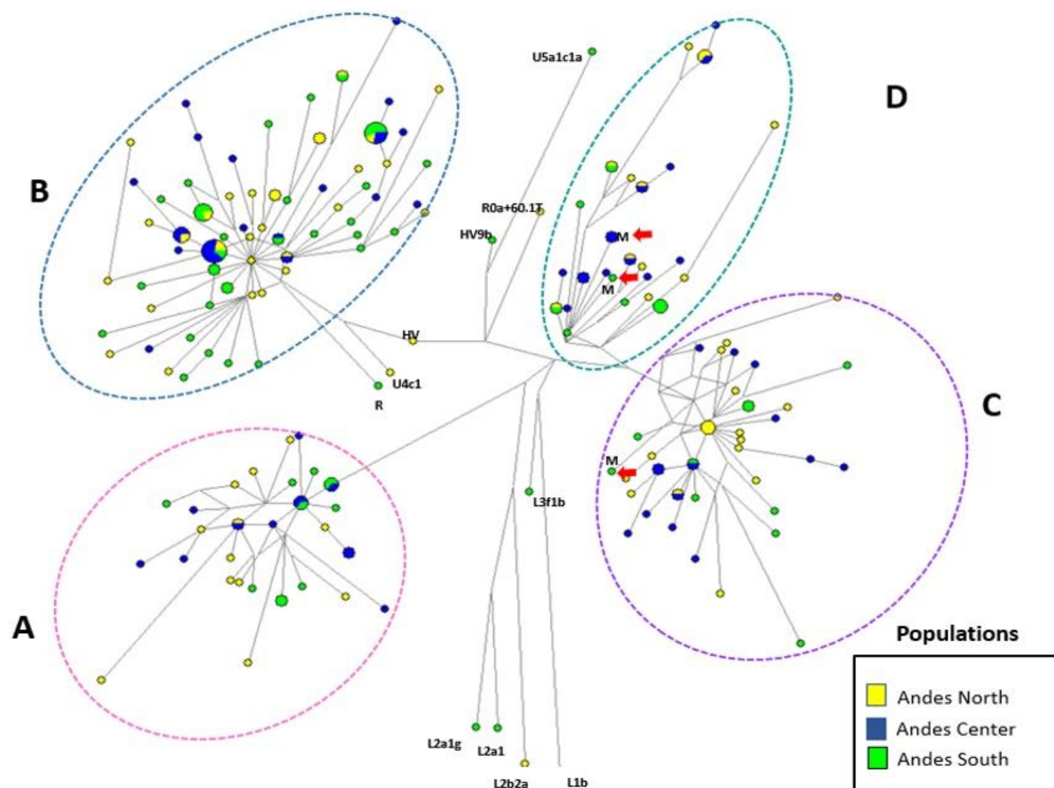


Figure 16. Median joining network of Ecuadorian Andes samples (n=214). Each population was assigned a different colour. The haplotypes are represented by circles with sizes proportional to the number of haplotypes, and branch lengths are proportional to the mutational steps. Native American haplogroups are indicated by dotted circles. Red arrows point to haplotypes classified to M haplogroup.

The median joining network (*Figure 16*) corroborates the high diversity of our population sample, without a clear geographic specificity within each clade. Haplotype sharing between North, Center, and South can be detected within all major Native haplogroups. The absence of clear star-like patterns, with main central nodes connected to less frequent haplotypes, inside of each haplogroup (Harpending et al., 1998) confirms the presence of diversified lineages without a single (or few) recent common ancestor.

Overall, the absence of important signs of genetic drift suggests that the population did not experience significant bottlenecks. Moreover, the high diversity is compatible to substantial gene flow, a process that tends to homogenize allele frequencies across subpopulations.

4.4. Population structure within the Ecuadorian Andes

Most admixed South American populations display a high diversity. However, distinct patterns of genetic structure can be found across regions/countries in the subcontinent. Numerous studies revealed a significant mtDNA heterogeneity among groups within the same country, such as Argentina that shows variation at regional level (Bobillo et al., 2010), and Brazil that presents genetic differentiation among geographically close states and among regions within the same state (dos Reis et al., 2019; Giolo et al., 2011; Poletto et al., 2019; Simão et al., 2018). Previously, substructure was also detected between nearby Andean Mestizo populations in Colombia, indicating a high degree of differentiation among them (Ribeiro, 2021).

To investigate substructure within the Andean population of Ecuador, pairwise F_{ST} genetic distances (*Appendix 7*) were calculated and an Analysis of Molecular Variance (AMOVA) was performed. The results revealed no statistically significant differences between populations from North, Central and South Andes. AMOVA results highlighted their homogeneity, with 99.73% of the genetic variation found within each Andean populations and only 0.27% attributed to differences between populations (P=0.22535+-0.00411).

The absence of discernible differences may be attributed to gene flow that has occurred over time. Factors such as intermarriage and population movements could contribute to a genetic homogenization of the populations, potentially erasing distinct genetic signatures that might have once characterized specific Native groups. Moreover, despite the occurrence of non-Native haplotypes in different proportions among populations in the Northern and Southern regions, and their absence in the central region (*Figure 15 and 16*), the differences observed are not high enough to conclude on a statistically significant differentiation between three regions.

4.5. Ecuadorian Andes in a South American context

4.5.1. Maternal ancestry

Several factors, such as geographical isolation, population origin and gene flow, as well as socio-cultural aspects, contributed to the establishment of unique genetic characteristics among different populations within South America. An analysis of pairwise genetic distances was carried out to assess the impact of geography on the genetic composition of South American populations. Comparisons were performed between our dataset and 1944 mtDNA control region sequences from several admixed and two Native American populations (*Appendix 2*). In addition, a map of haplogroup distributions was also created to correlate genetic distances and ancestry proportions.

Looking at the two-dimensional MDS plot (*Figure 17*), a considerable genetic distance was observed between the Amazonian Native population and the Andean Native and admixed populations, which can be attributed to the geographic isolation of the Amazonian population (Baeta et al., 2012). The Amazon rainforest historically acted as a barrier to genetic exchange, yet instances of contact occurred between indigenous groups and non-indigenous populations. While limited genetic exchange resulted from these interactions, the rainforest's isolation predominantly inhibited substantial admixture. Contrarily, the Andean Native populations have a complex history of gene flow from other groups due to historical interactions, motivated by their geographic location. Andean Native communities witness series of migration episodes, conquests and trade networks that played a role in shaping their genetic diversity.

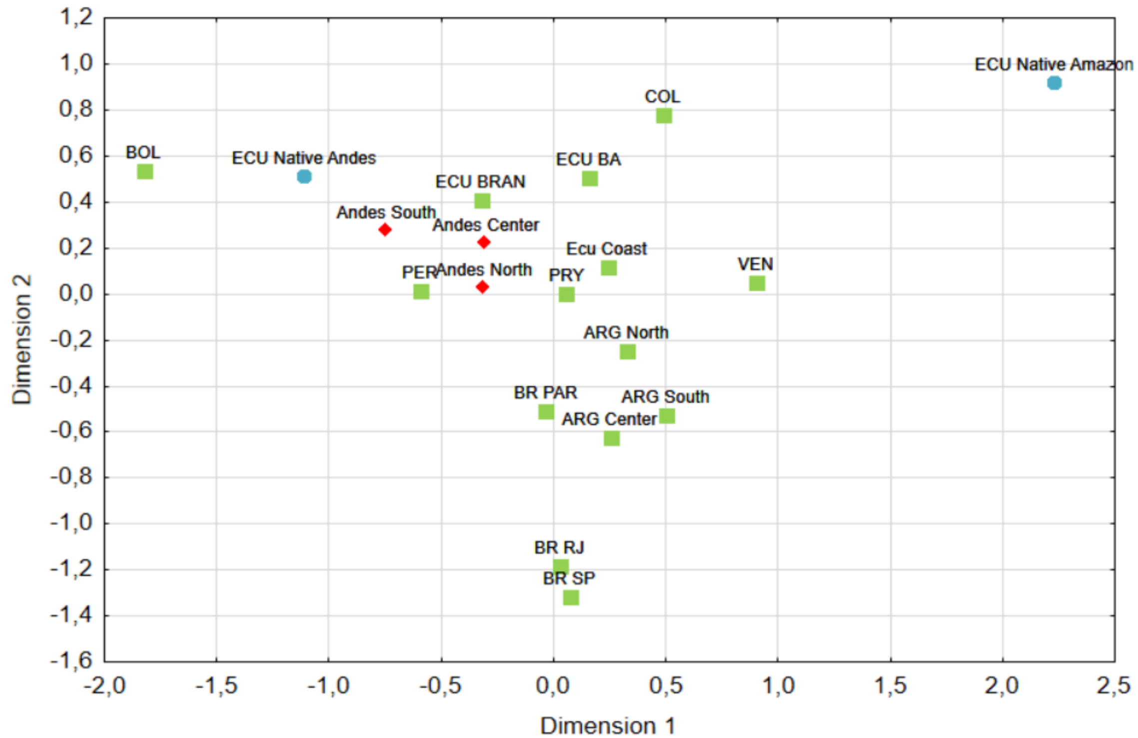


Figure 17. MDS plot based on haplotype F_{ST} genetic distances (Stress=0.08), between our Ecuadorian Andes (North, Center, and South), Ecuadorian Coast (data not published), South American admixed populations [COL: Colombia (Auton et al., 2015); PER: Peru (Auton et al., 2015); BOL: Bolivia (Taboada-Echalar et al., 2013); VEN: Venezuela (Castro De Guerra et al., 2012); PRY: Eastern region of Paraguay (Simão et al., 2021); ARG North, Center and South: Argentina (Bobillo et al., 2010) ; BR PAR: Paraná, Brazil (Poletto et al., 2019); BR RJ: Rio de Janeiro , Brazil (Simão et al., 2018); BR SP: Sao Paulo, Brazil (Prieto et al., 2011); ECU BRAN: Ecuadorian Brandini (Brandini et al., 2018); ECU BAE: Ecuadorian Baeta (Baeta et al., 2012)], and two Native populations [ECU Native Andes: Ecuadorian Kichwa Natives from the Andes (Sandoval et al., 2020), ECU Native Amazon: Ecuadorian Kichwa Natives from the Amazon (Baeta et al., 2012)]. Andes North, Andes Center and Andes South population are plotted in red; the admixed populations of South America are marked in green, and the Native American population is shown in blue.

Some population clusters are displayed on the MDS plot. The Brazilian admixed populations of São Paulo and Rio de Janeiro are grouped together, in a cluster separated from the remaining populations. This divergence can be attributed to the substantial African genetic contribution to these populations (*Figure 18*). The high frequency of African lineages is explained by the larger number of people from Africa arriving in these regions during the slave trade period (Alves-Silva et al., 2000; H. J. Bandelt et al., 2001; Secher et al., 2014). The Northeastern city of Brazil, Salvador, served as a major hub for the importation of African slaves to work on the sugarcane plantations and other labour-intensive industries in the region. However, the shift from sugar production in the Northeast to gold mining and coffee cultivation in the South led to changes in trade routes. This transformation resulted in Rio de Janeiro becoming the primary Brazilian port for the entry of most Western Central enslaved Africans (Conrad, 1983).

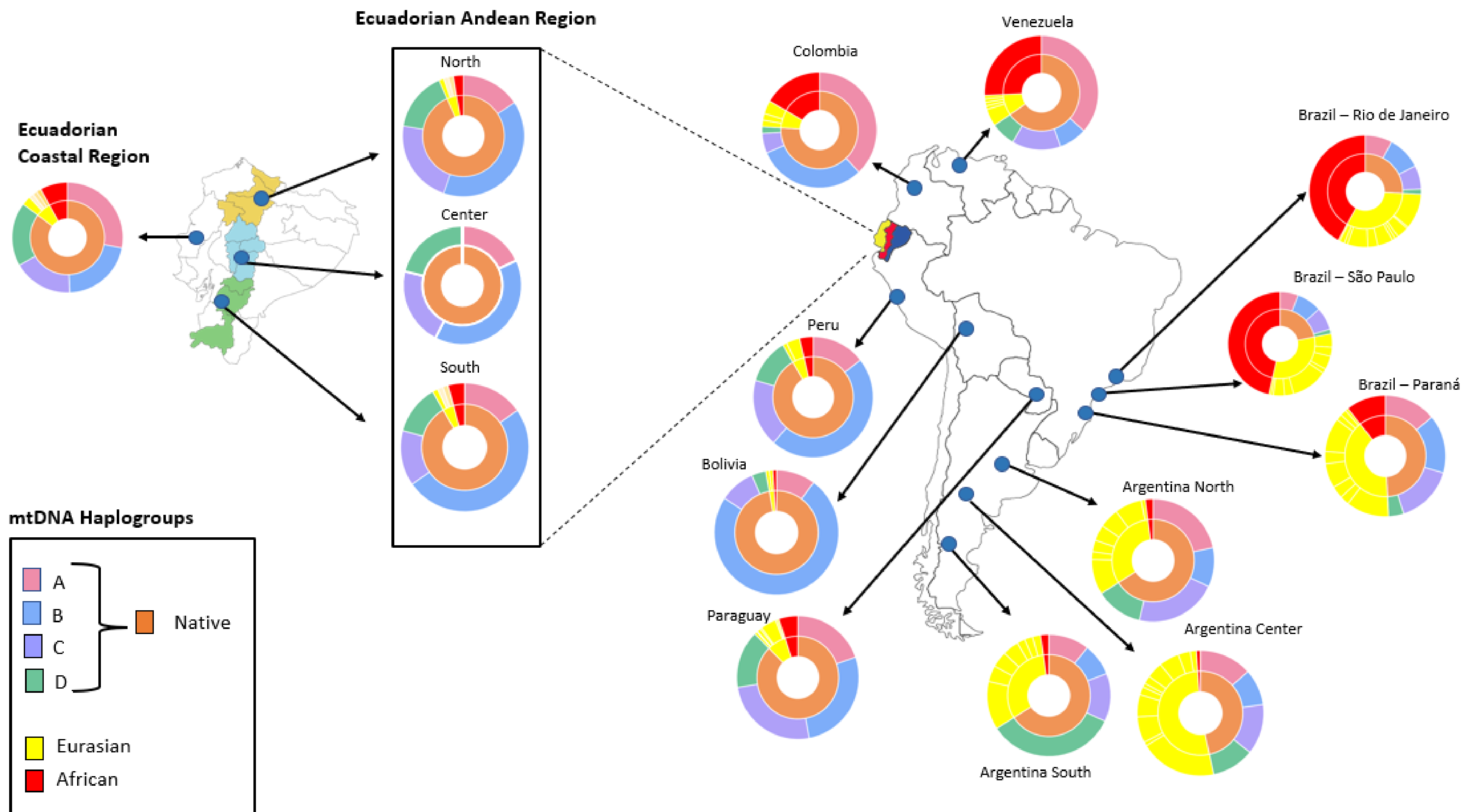


Figure 18. Maternal lineages found in several admixed populations from South America: Ecuadorian Andes (the present study); Ecuadorian Coast (data not published); Medellin, Colombia (Auton et al., 2015); Lima, Peru (Auton et al., 2015); La Paz, Bolivia (Taboada-Echalar et al., 2013); Caracas, Venezuela (Castro De Guerra et al., 2012); Eastern region of Paraguay (Simão et al., 2021); North, Center and South Argentina (Bobillo et al., 2010); Paraná, Brazil (Poletto et al., 2019); Rio de Janeiro, Brazil (Simão et al., 2018); and São Paulo, Brazil (Prieto et al., 2011).

Another cluster is observed among populations from the Southern region, particularly between Argentina and the South Brazilian state of Paraná. These populations prominently reflect the impact of Eurasian ancestry (*Figure 18*). Eurasian haplogroups are well represented in Argentina, particularly in the central part of the country, due to waves of European migration, primarily Spanish and Italians, to the industrial zones. The population of Brazil also has a significant Eurasian contribution due to migrations from Portugal, Germany, Italy, Switzerland, Spain, and other countries, resulting in a complex admixture scenario (Palencia et al., 2010; Poletto et al., 2019).

Although displaying major maternal Native American heritage and minor African and Eurasian contributions, Western South American populations show some genetic differentiation in the maternal background, which is clear by the absence of a well-defined cluster in the MDS plot. This is due to the variation on the proportion of specific Native lineages. For instance, populations in the Northwestern regions of Colombia and Northern regions like Venezuela display a greater genetic distance from Peruvian or Ecuadorian populations, likely due to the high frequency of the Native haplogroup A, in contrast to the prevalence of the Native haplogroup B in Central West Andean populations (Gómez-Carballa et al., 2018). Moreover, for the populations of Colombia and Venezuela, a significant African contribution (*Figure 18*) may also be responsible for their split from the other Western populations. The cities of Caracas in Venezuela and Cartagena in Colombia stood as prominent ports during the African slave trade, which explains the higher African heritage in these populations. However, a considerable number of individuals of African descent who initially arrived through Cartagena Port opted to establish their communities in the neighboring Chocó region (Doria et al., 2019). Nowadays, Chocoans seeking a better life are migrating to Medellin. This migration highlights the enduring African heritage in the region.

Similar haplogroup distributions were observed among Andean Ecuadorean and Peruvian populations (*Figure 18*). This genetic proximity is reflected on the MDS plot (*Figure 17*). This genetic similarity between the two countries might be attributed to historical population interactions facilitated by their geographical proximity. Furthermore, both Peru and Ecuador are home to a rich variety of indigenous populations with deep-rooted ancestral ties to the region. The presence of shared ancestral groups, such as the Quechua and several other indigenous communities, may have contributed to the genetic affinities due to cultural interactions that occurred across the Andean landscape (Blum et al., 2023).

Differences in Native American contributions are also evident between the Coast and the Andean region of Ecuador, with the coastal area showing a significant prevalence of haplogroup A (*Figure 18*). Furthermore, a substantial African ancestry is observed, in line with the presence of a noteworthy Afro-descendant settlement in one of the Northern provinces of the coast, particularly in Esmeraldas. Additionally, an influx of European haplogroups higher than the Andean region is discernible, probably due to the historical precedence of coastal contact with European colonizers, predominantly of Spanish origin.

4.5.2. Native American composition

As previously stated, Western and Eastern South American populations exhibit notable differences in the Native female contribution. To gain a deeper understanding of the genetic structure and relationships among these populations, new analyses were performed by focusing exclusively on the mtDNA lineages of Native American origin (A, B, C, D), excluding all non-Native lineages of European or African origin. A Principal Component Analysis (PCA) was performed using only the Native component to identify the haplogroups that played a role in South American populations' differentiation.

Overall, the PCA revealed differences between populations from Western and Eastern Andes. Looking at PC1, it is clear that the frequency of haplogroup B plays an important role on distinguishing the Western Andes from the remaining populations, highlighting a genetic continuity and shared ancestry within this region (*Figure 19*).

Previous studies showed a high frequency of haplogroup B in Native populations, such as in Andean Quechuas of Ecuador (60%) (Sandoval et al., 2020), reaching even higher frequencies up to 70% in Peru and Bolivia (Gómez-Carballa et al., 2018; Sandoval et al., 2013; Watkins et al., 2012). The predominance of haplogroup B in Andean populations could be traced back to the earliest Native inhabitants. Andean nations were part of the Inca Empire, whose rapid territorial expansion and population growth led to the disappearance of smaller population groups. Rodriguez-Delfin et al., (2001) suggest that the distribution of haplogroup B in these Andean populations could be explained by its predominance in a possible small initial Inca population. According to these authors, genetic drift could have been the cause of the prevalence of this haplogroup and its current distribution pattern observed across the Andes.

It is important to note that the population from coastal Ecuador does not follow this trend. The lack of representation in samples from certain provinces, such as Esmeraldas and Santa Elena, may introduce a bias in the results related to their Native genetic contributions. However, a more extensive sampling effort, with a larger number of individuals, could potentially bring forth accurate insights into the distribution of haplogroups in this region.

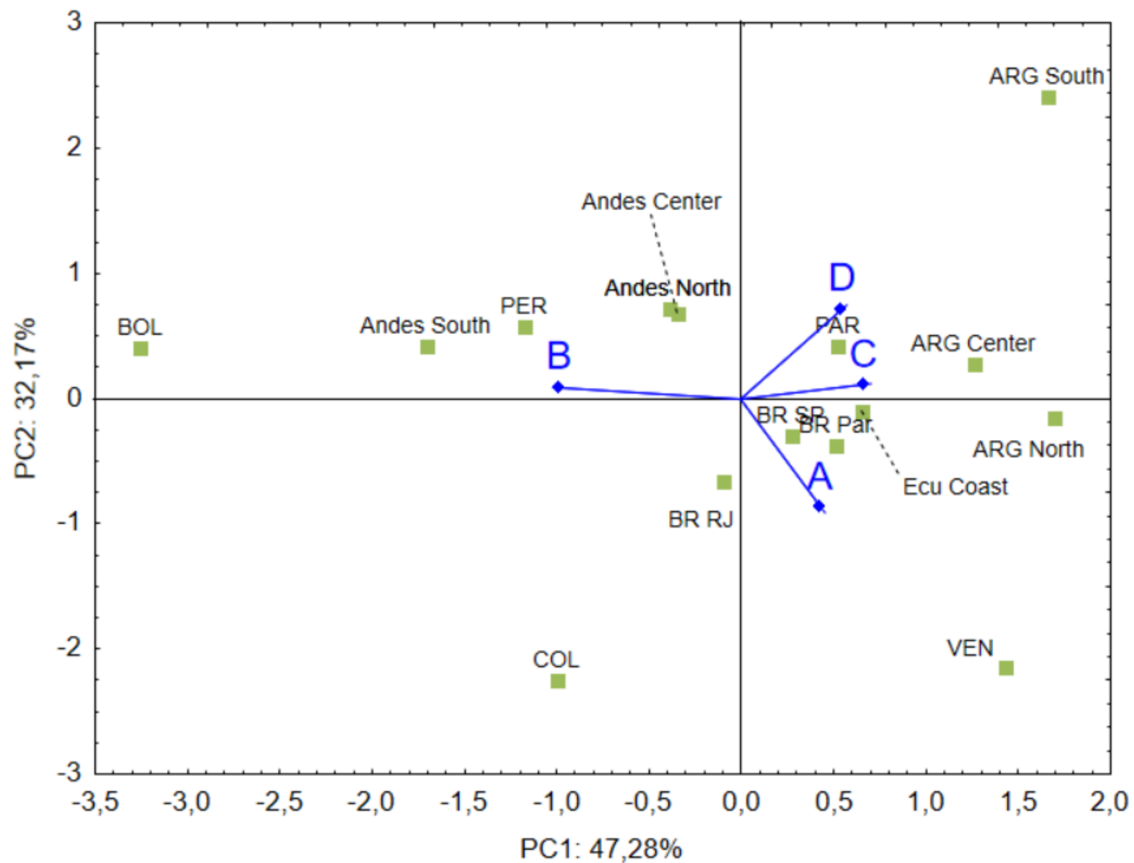


Figure 19. Principal component analysis (PCA) representing the genetic landscape of South American countries based on mtDNA Native haplogroups frequencies from Ecuadorian Andes (the present study); Ecu Coast: Ecuadorian Coast (data not published); COL: Colombia (Auton et al., 2015); PER: Peru (Auton et al., 2015); BOL: Bolivia (Taboada-Echalar et al., 2013); VEN: Venezuela (Castro De Guerra et al., 2012); PRY: Eastern region of Paraguay (Simão et al., 2021); ARG North, Center and South: Argentina (Bobillo et al., 2010) ; BR PAR: Paraná, Brazil (Poletto et al., 2019); BR RJ: Rio de Janeiro , Brazil (Simão et al., 2018); and BR SP: Sao Paulo, Brazil (Prieto et al., 2011).

For the remaining Native haplogroups - A, C and D - the PC2 highlights a North-South pattern. The populations in the Northern regions of South America, such as Colombia and Venezuela, exhibit a notable prevalence of the haplogroup A (*Figure 19*). According to some authors, these lineages might have been introduced to South America through migratory waves originating from Central America (Nuñez et al., 2010; Salas et al., 2009). This introduction could have led to a gradual decrease in the frequency of this haplogroup towards the South, particularly evident along the Caribbean coast of Colombia, Venezuela, and further in Brazil and Argentina (Rothhammer et al., 2017). On

the contrary, haplogroups C and D are characterized by significant frequency peaks in the Central and Southern regions. An upward trajectory from the Northern to the southern regions is detected for these haplogroups.

This pattern is notably discernible in the PCA, where PC2 clearly separates populations characterized by haplogroup A (Venezuela, Colombia, and Brazil) from those with haplogroups C and D (Argentina, especially South Argentina, and Paraguay).

Insights from studies involving ancient Tierra del Fuego aborigines propose a preliminary wave of migration characterized by populations of older lineage carrying haplogroups C and D towards the Southern territories. This hypothesis emerges as these aboriginal populations lack the presence of haplogroups A and B, which could have followed a subsequent expansion southward (Lalueza et al., 1997; Rodriguez-Delfin et al., 2001).

The disparities in mtDNA haplogroup frequencies between Western and Eastern Andean regions can be comprehended by considering multiple migrations of Native American ancestors to the South and two primary migration routes: one tracing the Pacific coastline and other the Atlantic coast.

Conclusions

This study provides valuable insights into the mtDNA variation among contemporary Mestizo populations in the Ecuadorian Andes. The relatively minor influence of European and African maternal lineages, coupled with a substantial Native American genetic heritage, explains the enduring preservation of Native American cultures during the colonial period. This stands in contrast to certain Eastern South American countries where mass decimation led to the extinction of indigenous lineages. Further analysis on Y-chromosomal lineages on the Andean region will shed light into the paternal contribution in these population and clarify admixture patterns.

Despite some differences detected in haplogroup distribution between the three Andean samples, there are no statistically significant haplotypic differences, demonstrating our geographic division criteria did not capture a significant maternal population substructure. Nonetheless, the statistically significant difference observed between Andes and Pacific Coast regions highlights some heterogeneity among regional divisions.

The genetic differences in mtDNA variation among Ecuador and other South American admixed populations suggest a unique evolutionary history and genetic makeup for the Andean populations. The presence of specific mtDNA lineages in the present studied populations is further evidence of complex ancestral contributions and genetic admixture in the Andean region. Overall, these findings support that Andean populations from Ecuador have undergone a separate evolutionary trajectory compared to their counterparts in Eastern regions of South America.

References

Almeida, E. (2000). *Culturas Prehispánicas del Ecuador. Viajes Chasquiguiñán.* 180. https://books.google.com/books/about/Culturas_prehisp%C3%A1nicas_del_Ecuador.html?hl=es&id=5E8UAQAAIAAJ

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

Amorim, A. (2013). Population Genetics. *Brenner's Encyclopedia of Genetics: Second Edition*, 407–411. https://doi.org/10.1016/B978-0-12-374984-0.01195-5

Amorim, A., & Budowle, B. (2016). Handbook of forensic genetics: Biodiversity and heredity in civil and criminal investigation. *Handbook of Forensic Genetics: Biodiversity and Heredity in Civil and Criminal Investigation*, 1–632. https://doi.org/10.1142/Q0023

Amorim, A., Fernandes, T., & Taveira, N. (2019). Mitochondrial DNA in human identification: a review. *PeerJ*, *7*. https://doi.org/10.7717/PEERJ.7314

Anderson, S., Bankier, A. T., Barrell, B. G., De Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., & Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature 1981 290:5806*, *290*(5806), 457–465. https://doi.org/10.1038/290457a0

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

Andrien, K. J. (1995). The Kingdom of Quito, 1690–1830: The State and Regional Development. *The Kingdom of Quito, 1690–1830*. https://doi.org/10.1017/CBO9780511529054

Athens, J. (1980). El proceso evolutivo en las sociedades complejas y la ocupación del período Tardío - Cara en los Andes Septentrionales del Ecuador. In *Otavalo, Instituto Otavaleño de Antropología* https://biblio.flacsoandes.edu.ec/shared/biblio_view.php?bibid=132242&tab=opac

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., Lander, E. S., Lee, C., Lehrach, H., … Schloss, J. A. (2015). A global reference for human genetic variation. *Nature 2015 526:7571*, *526*(7571), 68–74. https://doi.org/10.1038/nature15393

Ayala, E. (2003). *Resumen de Historia del Ecuador*.

Baeta, M., Núñez, C., González-Andrade, F., Sosa, C., Casalod, Y., Bolea, M., Zweynert, S., Vacas Cruz, O. W., González-Solorzano, J., Geppert, M., Roewer, L., & Martínez-Jarreta, B. (2009). Mitochondrial analysis revealed high homogeneity in the Waorani population—The last nomadic group of hunter-gatherers from Ecuador. *Forensic Science International: Genetics Supplement Series*, *2*(1), 313–314. https://doi.org/10.1016/j.fsigss.2009.08.025

Baeta, M., Núñez, C., Sosa, C., Bolea, M., Casalod, Y., González-Andrade, F., Roewer, L., & Martínez-Jarreta, B. (2012). Mitochondrial diversity in Amerindian Kichwa and Mestizo populations from Ecuador. *International Journal of Legal Medicine*, *126*(2), 299–302. https://doi.org/10.1007/s00414-011-0656-4

Baeza, C., Guzmán, R., Tirado, M., López-Parra, A. M., Rodríguez, T., Mesa, M. S., Fernández, E., & Arroyo-Pardo, E. (2007). Population data for 15 Y-chromosome STRs in a population sample from Quito (Ecuador). *Forensic Science International*, *173*(2–3), 214–219. https://doi.org/10.1016/j.forsciint.2006.09.011

Ballard, D. (2016). Analysis of Mitochondrial Control Region Using Sanger Sequencing. *Methods in Molecular Biology (Clifton, N.J.)*, *1420*, 143–155. https://doi.org/10.1007/978-1-4939-3597-0_12

Bandelt, H. J., Alves-Silva, J., Guimarães, P. E. M., Santos, M. S., Brehm, A., Pereira, L., Coppa, A., Larruga, J. M., Rengo, C., Scozzari, R., Torroni, A., Prata, M. J., Amorim, A., Prado, V. F., & Pena, S. D. J. (2001). Phylogeography of the human mitochondrial haplogroup L3e: a snapshot of African prehistory and Atlantic slave trade. *Annals of Human Genetics*, *65*(Pt 6), 549–563. https://doi.org/10.1017/S0003480001008892

Bandelt, H. J., Kloss-Brandstätter, A., Richards, M. B., Yao, Y. G., & Logan, I. (2013). The case for the continuing use of the revised Cambridge Reference Sequence (rCRS) and the standardization of notation in human mitochondrial DNA studies.

Journal of Human Genetics 2014 59:2, *59*(2), 66–77. https://doi.org/10.1038/jhg.2013.120

Bandelt, H.-J., Macaulay, V., Richards, M., & Chinnery, P. F. (2006). Mitochondrial DNA in Homo Sapiens. *Human Mitochondrial DNA and the Evolution of Homo Sapiens*, *18*, 3–15. https://doi.org/10.1007/3-540-31789-9_1

Battaglia, V., Grugni, V., Perego, U. A., Angerhofer, N., Gomez-Palmieri, J. E., Woodward, S. R., Achilli, A., Myres, N., Torroni, A., & Semino, O. (2013). The First Peopling of South America: New Evidence from Y-Chromosome Haplogroup Q. *PLOS ONE*, *8*(8), e71390. https://doi.org/10.1371/JOURNAL.PONE.0071390

Behar, D. M., Van Oven, M., Rosset, S., Metspalu, M., Loogväli, E. L., Silva, N. M., Kivisild, T., Torroni, A., & Villems, R. (2012). A "Copernican" Reassessment of the Human Mitochondrial DNA Tree from its Root. *American Journal of Human Genetics*, *90*(4), 675. https://doi.org/10.1016/J.AJHG.2012.03.002

Berger, C., Hatzer-Grubwieser, P., Hohoff, C., & Parson, W. (2011). Evaluating sequence-derived mtDNA length heteroplasmy by amplicon size analysis. *Forensic Science International: Genetics*, *5*(2), 142–145. https://doi.org/10.1016/J.FSIGEN.2010.10.002

Blum, F., Barrientos, C., Ingunza, A., & Poirier, Z. (2023). A phylolinguistic classification of the Quechua language family. *Indiana*, 1–20. https://doi.org/10.31235/OSF.IO/TWU6A

Bobillo, M. C., Zimmermann, B., Sala, A., Huber, G., Röck, A., Bandelt, H. J., Corach, D., & Parson, W. (2010). Amerindian mitochondrial DNA haplogroups predominate in the population of Argentina: Towards a first nationwide forensic mitochondrial DNA sequence database. *International Journal of Legal Medicine*, *124*(4), 263–268. https://doi.org/10.1007/S00414-009-0366-3/TABLES/5

BORCHART de Moreno, C., & Moreno Yánez, S. (1997). *Crónica indiana del Ecuador antiguo*. https://www.worldcat.org/title/cronica-indiana-del-ecuador-antiguo/oclc/38279508

Bossano, L. (1974). Cronología de la fundación española de Quito. In *Quito, Ed. Universitaria*. https://bibliotecamunicipalguayaquil.com/cgi-bin/koha/opac-detail.pl?biblionumber=73938&shelfbrowse_itemnumber=93836

Brandini, S., Bergamaschi, P., Fernando Cerna, M., Gandini, F., Bastaroli, F., Bertolini, E., Cereda, C., Ferretti, L., Gómez-Carballa, A., Battaglia, V., Salas, A., Semino,

O., Achilli, A., Olivieri, A., & Torroni, A. (2018). The Paleo-Indian Entry into South America According to Mitogenomes. *Molecular Biology and Evolution*, *35*(2), 299–311. https://doi.org/10.1093/MOLBEV/MSX267

Brandstätter, A., Niederstätter, H., & Parson, W. (2004). Monitoring the inheritance of heteroplasmy by computer-assisted detection of mixed basecalls in the entire human mitochondrial DNA control region. *International Journal of Legal Medicine*, *118*(1), 47–54. https://doi.org/10.1007/S00414-003-0418-Z

Brandstätter, A., Peterson, C. T., Irwin, J. A., Mpoke, S., Koech, D. K., Parson, W., & Parsons, T. J. (2004). Mitochondrial DNA control region sequences from Nairobi (Kenya): inferring phylogenetic parameters for the establishment of a forensic database. *International Journal of Legal Medicine*, *118*(5), 294–306. https://doi.org/10.1007/S00414-004-0466-Z

Brown, W. M., George, M., & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(4), 1967–1971. https://doi.org/10.1073/PNAS.76.4.1967

Budowle, B., Chakraborty, R., Giusti, A. M., Eisenberg, A. J., & Allen, R. C. (1991). Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. *American Journal of Human Genetics*, *48*(1), 137. /pmc/articles/PMC1682756/?report=abstract

Burgos, G., Simão, F., Flores-Espinoza, R., Yepez-Santos, J. I., Garzón-Salazar, A., Paz-Cruz, E., Freire-Paspuel, B., Carvalho, E. F., & Gusmão, L. (2019). An approach to maternal ancestry in a sample of Ecuadorian “mestizo” population by sequencing the control region of mtDNA. *Forensic Science International: Genetics Supplement Series*, *7*(1), 537–538. https://doi.org/10.1016/J.FSIGSS.2019.10.081

Butler, J. M. (2010a). Basics of DNA Biology and Genetics. In *Fundamentals of Forensic DNA Typing* (pp. 19–41). Elsevier. https://doi.org/10.1016/b978-0-12-374999-4.00002-3

Butler, J. M. (2010b). Lineage Markers. In *Fundamentals of Forensic DNA Typing* (pp. 363–396). Elsevier. https://doi.org/10.1016/b978-0-12-374999-4.00016-3

Butler, J. M. (2011). Advanced Topics in Forensic DNA Typing: Methodology. *Advanced Topics in Forensic DNA Typing: Methodology*, 1–680. https://doi.org/10.1016/C2011-0-04189-3

Cardoso, S., Alfonso-Sánchez, M. A., Valverde, L., Sánchez, D., Zarrabeitia, M. T., Odriozola, A., Martínez-Jarreta, B., & Pancorbo, M. M. de. (2012). Genetic uniqueness of the Waorani tribe from the Ecuadorian Amazon. *Heredity*, *108*(6), 609. https://doi.org/10.1038/HDY.2011.131

Castillo, A., Deccache, L., de Carvalho, E. F., Simão, F., & Gusmão, L. (2019). Maternal genetic characterization of a Colombian Andean population. *Forensic Science International: Genetics Supplement Series*, *7*(1), 342–344. https://doi.org/10.1016/J.FSIGSS.2019.10.004

Castro De Guerra, D., Figuera Pérez, C., Bravi, C. M., Saunier, J., Scheible, M., Irwin, J., Coble, M. D., & Rodríguez-Larralde, A. (2012). Sequence variation of mitochondrial DNA control region in North Central Venezuela. *Forensic Science International. Genetics*, *6*(5). https://doi.org/10.1016/J.FSIGEN.2011.11.004

Cavalli-Sforza, L. L. (1974). The genetics of human populations. *Scientific American*, *231*(3), 80–89. https://doi.org/10.1038/SCIENTIFICAMERICAN0974-80

Cavelier, L., Jazin, E., Jalonen, P., & Gyllensten, U. (2000). MtDNA substitution rate and segregation of heteroplasmy in coding and noncoding regions. *Human Genetics*, *107*(1), 45–50. https://doi.org/10.1007/S004390000305

Chung, U., Hwan, Y. L., Yoo, J. E., Myung, J. P., & Shin, K. J. (2005). Mitochondrial DNA CA dinucleotide repeats in Koreans: the presence of length heteroplasmy. *International Journal of Legal Medicine*, *119*(1), 50–53. https://doi.org/10.1007/S00414-004-0487-7

Conrad, R. E. (1983). *Children of God's fire: a documentary history of black slavery in Brazil*. 515.

Dall, W. H., & Harris, G. D. (1892). Correlation papers: Neocene. *Bulletin*. https://doi.org/10.3133/B84

Doria, A. M. R., Bolívar, F. J. F., Doria, A. M. R., & Bolívar, F. J. F. (2019). Marginados, pero no marginales. Negros, mulatos y sus disputas por la autonomía en Chocó, Colombia (1903-1947). *Anuario de Historia Regional y de Las Fronteras*, *24*(2), 125–149. https://doi.org/10.18273/REVANU.V24N2-2019005

dos Reis, R. S., Simão, F., dos Santos Stange, V., Garcia, F. M., Spinassé Dettogni, R., Stur, E., da Silva, A. M. Á., de Carvalho, E. F., Gusmão, L., & Drumond Louro, I. (2019). A view of the maternal inheritance of Espírito Santo populations: The contrast between the admixed and Pomeranian descent groups. *Forensic Science*

International: Genetics, *40*, 175–181. https://doi.org/10.1016/J.FSIGEN.2019.03.007

Dulik, M. C., Owings, A. C., Gaieski, J. B., Vilar, M. G., Andre, A., Lennie, C., Mackenzie, M. A., Kritsch, I., Snowshoe, S., Wright, R., Martin, J., Gibson, N., Andrews, T. D., & Schurr, T. G. (2012). Y-chromosome analysis reveals genetic divergence and new founding native lineages in Athapaskan- and Eskimoan-speaking populations. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(22), 8471–8476. https://doi.org/10.1073/PNAS.1118760109/SUPPL_FILE/ST04.DOCX

Elias, S. A., & Crocker, B. (2008). The Bering Land Bridge: a moisture barrier to the dispersal of steppe–tundra biota? *Quaternary Science Reviews*, *27*(27–28), 2473–2483. https://doi.org/10.1016/J.QUASCIREV.2008.09.011

Enrique Ayala Mora. (2002). *Ecuador: patria de todos, la nación ecuatoriana, unidad en la diversidad - Universidad Andina Simón Bolívar*.

Espinoza Soriano, Waldemar. (1987). *Los incas: economia, sociedad y estado en la era del Tahuantinsuyo /* (1a ed.). Amaru Editores.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/J.1755-0998.2010.02847.X

Fregel, R., & Delgado, S. (2011). HaploSearch: A tool for haplotype-sequence two-way transformation. *Mitochondrion*, *11*(2), 366–367. https://doi.org/10.1016/J.MITO.2010.11.001

Fuselli, S., Tarazona-Santos, E., Dupanloup, I., Soto, A., Luiselli, D., & Pettener, D. (2003). Mitochondrial DNA diversity in South America and the genetic history of Andean highlanders. *Molecular Biology and Evolution*, *20*(10), 1682–1691. https://doi.org/10.1093/MOLBEV/MSG188

Galtier, N., Enard, D., Radondy, Y., Bazin, E., & Belkhir, K. (2006). Mutation hot spots in mammalian mitochondrial DNA. *Genome Research*, *16*(2), 215. https://doi.org/10.1101/GR.4305906

García Salazar, J. (1989). *Cimarronaje en el Pacifico Sur: historia y tradición. El caso de Esmeraldas, Ecuador*. http://repositorio.uasb.edu.ec/handle/10644/5939

Garg, S., Zimorski, V., & Martin, W. F. (2016). Endosymbiotic Theory. *Encyclopedia of Evolutionary Biology*, 511–517. https://doi.org/10.1016/B978-0-12-800049-6.00191-8

Gaviria, A., Sánchez, M. E., Morejón, G., Vela, M., Aguirre, V., Burgos, G., Zambrano, A. K., Leone, P. E., & Paz-y-Miño, C. (2013). Characterization and Haplotype analysis of 11 Y-STR loci in Ecuadorian population. *Forensic Science International: Genetics Supplement Series*, *4*(1), e310–e311. https://doi.org/10.1016/j.fsigss.2013.10.158

Gil Blanco, E., & Canela Ruano, A. J. (2018). *La migración europea al Ecuador (1935-1955). La visión diplomática de la llegada de colectivos problemáticos: los judíos*. http://repositorio.flacsoandes.edu.ec/handle/10469/15780

Giles, R. E., Blanc, H., Cann, H. M., & Wallace aD., C. (1980). Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *77*(11), 6715–6719. https://doi.org/10.1073/PNAS.77.11.6715

Giolo, S. R., Soler, J. M. P., Greenway, S. C., Almeida, M. A. A., De Andrade, M., Seidman, J. G., Seidman, C. E., Krieger, J. E., & Pereira, A. C. (2011). Brazilian urban population genetic structure reveals a high degree of admixture. *European Journal of Human Genetics 2012 20:1*, *20*(1), 111–116. https://doi.org/10.1038/ejhg.2011.144

Giugliani, R., Bender, F., Couto, R., Bochernitsan, A., Brusius-Facchin, A. C., Burin, M., Amorim, T., Acosta, A. X., Purificação, A., Segal, S. L., Saraiva-Pereira, M. L., Jardim, L. B., Matte, U., Riegel, M., Cardoso-Dos-Santos, A. C., Rodrigues, G., Oliveira, M. Z. De, Tagliani-Ribeiro, A., Heck, S., … Kubaski, F. (2019). Population medical genetics: translating science to the community. *Genetics and Molecular Biology*, *42*(1 suppl 1), 312–320. https://doi.org/10.1590/1678-4685-GMB-2018-0096

Goebel, T., Waters, M. R., & O'Rourke, D. H. (2008). The Late Pleistocene Dispersal of Modern Humans in the Americas. *Science*, *319*(5869), 1497–1502. https://doi.org/10.1126/SCIENCE.1153569

Goios, A., & Alvarez, L. (2013). Mitochondrial Genetics/Evolution. *Brenner's Encyclopedia of Genetics: Second Edition*, 439–440. https://doi.org/10.1016/B978-0-12-374984-0.00956-6

Gómez-Carballa, A., Pardo-Seco, J., Brandini, S., Achilli, A., Perego, U. A., Coble, M. D., Diegoli, T. M., Alvarez-Iglesias, V., Martinón-Torres, F., Olivieri, A., Torroni, A., & Salas, A. (2018). The peopling of South America and the trans-Andean gene flow of the first settlers. *Genome Research*, *28*(6), 767–779. https://doi.org/10.1101/GR.234674.118/-/DC1

González-Andrade, F., Roewer, L., Willuweit, S., Sánchez, D., & Martínez-Jarreta, B. (2009). Y-STR variation among ethnic groups from Ecuador: Mestizos, Kichwas, Afro-Ecuadorians and Waoranis. *Forensic Science International: Genetics*, *3*(3), e83–e91. https://doi.org/10.1016/j.fsigen.2008.08.003

Goodwin, W. H. (2016). DNA: Mitochondrial DNA. *Encyclopedia of Forensic and Legal Medicine: Second Edition*, 351–358. https://doi.org/10.1016/B978-0-12-800034-2.00155-5

Gray, M. W. (2001). Mitochondria. *Encyclopedia of Genetics*, 1215–1217. https://doi.org/10.1006/RWGN.2001.0835

Gray, M. W., Burger, G., & Lang, B. F. (1999). Mitochondrial Evolution. *Science*, *283*(5407), 1476–1481. https://doi.org/10.1126/SCIENCE.283.5407.1476

Hartl, D. L., & Clark, A. G. (2007). Principles of Population Genetics, 4th edition. *Journal of Heredity*, *98*(4), 382–382. https://doi.org/10.1093/JHERED/ESM035

Hashiguchi, K., Bohr, V. A., & de Souza-Pinto, N. C. (2004). Oxidative stress and mitochondrial DNA repair: implications for NRTIs induced DNA damage. *Mitochondrion*, *4*(2–3), 215–222. https://doi.org/10.1016/J.MITO.2004.05.014

Hoffecker, J. F., Elias, S. A., & O'Rourke, D. H. (2014). Out of Beringia? *Science*, *343*(6174), 979–980. https://doi.org/10.1126/SCIENCE.1250768

Hoffecker, J. F., Elias, S. A., O'Rourke, D. H., Scott, G. R., & Bigelow, N. H. (2016). Beringia and the global dispersal of modern humans. *Evolutionary Anthropology*, *25*(2), 64–78. https://doi.org/10.1002/EVAN.21478

Howell, N. (1997). mtDNA recombination: what do in vitro data mean? *American Journal of Human Genetics*, *61*(1), 19–22. https://doi.org/10.1086/513910

INEC (Instituto Nacional de Estadisticas y Censos). (2022). *Resultados del Censo 2022*. http://www.ecuadorencifras.gob.ec/resultados/

Ingman, M., Kaessmann, H., Pääbo, S., & Gyllensten, U. (2000). Mitochondrial genome variation and the origin of modern humans. *Nature*, *408*(6813), 708–713. https://doi.org/10.1038/35047064

Irwin, J. A., Parson, W., Coble, M. D., & Just, R. S. (2011). mtGenome reference population databases and the future of forensic mtDNA analysis. *Forensic Science International. Genetics*, *5*(3), 222–225. https://doi.org/10.1016/J.FSIGEN.2010.02.008

Irwin, J. A., Saunier, J. L., Niederstätter, H., Strouss, K. M., Sturk, K. A., Diegoli, T. M., Brandstätter, A., Parson, W., & Parsons, T. J. (2009). Investigation of heteroplasmy in the human mitochondrial DNA control region: A synthesis of observations from more than 5000 global population samples. *Journal of Molecular Evolution*, *68*(5), 516–527. https://doi.org/10.1007/S00239-009-9227-4

Jablonski, N. G., & Chaplin, G. (2010). Human skin pigmentation as an adaptation to UV radiation. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(SUPPL. 2), 8962–8968. https://doi.org/10.1073/PNAS.0914628107/ASSET/13CF7EDB-8EA8-439A-A4E6-1ABC29AAE6DA/ASSETS/GRAPHIC/PNAS.0914628107FIG02.JPEG

Jain, M. R., & Vishnu Priya, V. (2012). Mitochondrial DNA Analysis. *International Journal of Pharmaceutical Sciences Review and Research*, *32*(2), 405–456. https://doi.org/10.1016/B978-0-12-374513-2.00014-2

Jakobsson, M., Pearce, C., Cronin, T. M., Backman, J., Anderson, L. G., Barrientos, N., Björk, G., Coxall, H., De Boer, A., Mayer, L. A., Mörth, C. M., Nilsson, J., Rattray, J. E., Stranne, C., Semiletov, I., & O'Regan, M. (2017). Post-glacial flooding of the Bering Land Bridge dated to 11 cal ka BP based on new geophysical and sediment records. *Climate of the Past*, *13*(8), 991–1005. https://doi.org/10.5194/CP-13-991-2017

Jobling, M. A., Hurles, M., & Tyler-Smith, C. (2019). Human Evolutionary Genetics : Origins, Peoples and Disease. *Human Evolutionary Genetics*. https://doi.org/10.1201/9780203487211

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

Jota, M. S., Lacerda, D. R., Sandoval, J. R., Vieira, P. P. R., Ohasi, D., Santos-Jnior, J. E., Acosta, O., Cuellar, C., Revollo, S., Paz-y-Mi'o, C., Fujita, R., Vallejo, G. A., Schurr, T. G., Tarazona-Santos, E. M., Pena, S. D. J., Ayub, Q., Tyler-Smith, C., & Santos, F. R. (2016). New native South American Y chromosome lineages. *Journal of Human Genetics*, *61*(7), 1–11. https://doi.org/10.1038/jhg.2016.26

Kashani, B. H., Perego, U. A., Olivieri, A., Angerhofer, N., Gandini, F., Carossa, V., Lancioni, H., Semino, O., Woodward, S. R., Achilli, A., & Torroni, A. (2012). Mitochondrial haplogroup C4c: a rare lineage entering America through the ice-free corridor? *American Journal of Physical Anthropology*, *147*(1), 35–39. https://doi.org/10.1002/AJPA.21614

Keats, B. J. B., & Sherman, S. L. (2014). Population Genetics. *Reference Module in Biomedical Sciences*. https://doi.org/10.1016/B978-0-12-801238-3.05486-6

Kimpton, C. P., Gill, P., Walton, A., Urquhart, A., Millican, E. S., & Adams, M. (1993). Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR Methods and Applications*, *3*(1), 13–22. https://doi.org/10.1101/GR.3.1.13

Kivisild, T., Tolk, H. V., Parik, J., Wang, Y., Papiha, S. S., Bandelt, H. J., & Villems, R. (2002). The Emerging Limbs and Twigs of the East Asian mtDNA Tree. *Molecular Biology and Evolution*, *19*(10), 1737–1751. https://doi.org/10.1093/OXFORDJOURNALS.MOLBEV.A003996

Lalueza, C., Pérez-Pérez, A., Prats, E., Cornudella, L., & Turbón, D. (1997). Lack of founding Amerindian mitochondrial DNA lineages in extinct aborigines from Tierra del Fuego-Patagonia. *Human Molecular Genetics*, *6*(1), 41–46. https://doi.org/10.1093/HMG/6.1.41

Lambeck, K., Yokoyama, Y., & Purcell, T. (2002). Into and out of the Last Glacial Maximum: sea-level change during Oxygen Isotope Stages 3 and 2. *Quaternary Science Reviews*, *21*(1–3), 343–360. https://doi.org/10.1016/S0277-3791(01)00071-3

Leone, P. E., Maldonado-Oyervide, D., Astudillo-González, O., Pérez-Villa, A., Yumiceba, V., Armendáriz-Castillo, I., García-Cárdenas, J. M., Guerrero, S., Guevara-Ramírez, P., López-Cortés, A., Zambrano, A. K., & Paz-y-Miño, C. (2019). Mitochondrial DNA study in the Shuar ethnic group from Ecuador. *Forensic Science International: Genetics Supplement Series*, *7*(1), 142–143. https://doi.org/10.1016/j.fsigss.2019.09.055

Lewis, C. M., Tito, R. Y., Lizárraga, B., & Stone, A. C. (2005). Land, language, and loci: mtDNA in Native Americans and the genetic history of Peru. *American Journal of Physical Anthropology*, *127*(3), 351–360. https://doi.org/10.1002/AJPA.20102

Luikart, G., Kardos, M., Hand, B. K., Rajora, O. P., Aitken, S. N., & Hohenlohe, P. A. (2018). *Population Genomics: Advancing Understanding of Nature*. 3–79. https://doi.org/10.1007/13836_2018_60

Luo, S., Valencia, C. A., Zhang, J., Lee, N. C., Slone, J., Gui, B., Wang, X., Li, Z., Dell, S., Brown, J., Chen, S. M., Chien, Y. H., Hwu, W. L., Fan, P. C., Wong, L. J., Atwal, P. S., & Huang, T. (2018). Biparental inheritance of mitochondrial DNA in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(51), 13039–13044. https://doi.org/10.1073/PNAS.1810946115/SUPPL_FILE/PNAS.1810946115.SD03.XLSX

Manoli, I., Alesci, S., & Chrousos, G. P. (2007). Mitochondria. *Encyclopedia of Stress*, 754–761. https://doi.org/10.1016/B978-012373947-6.00559-6

Mezzavilla, M., Geppert, M., Tyler-Smith, C., Roewer, L., & Xue, Y. (2015). Insights into the origin of rare haplogroup C3∗ y chromosomes in South America from high-density autosomal SNP genotyping. *Forensic Science International: Genetics*, *15*, 115–120. https://doi.org/10.1016/J.FSIGEN.2014.11.005

Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, *16*(3), 1215. https://doi.org/10.1093/nar/16.3.1215

Mkombachepa, M., Khamis, B., Rwegasira, G., Urio, F., Makani, J., & Luzzatto, L. (2022). High incidence of malaria in patients with sickle cell disease. *American Journal of Hematology*, *97*(10), E380–E381. https://doi.org/10.1002/AJH.26676

Moreno-Mayar, J. V., Potter, B. A., Vinner, L., Steinrücken, M., Rasmussen, S., Terhorst, J., Kamm, J. A., Albrechtsen, A., Malaspinas, A. S., Sikora, M., Reuther, J. D., Irish, J. D., Malhi, R. S., Orlando, L., Song, Y. S., Nielsen, R., Meltzer, D. J., & Willerslev, E. (2018). Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature 2018 553:7687*, *553*(7687), 203–207. https://doi.org/10.1038/nature25173

Moreno-Mayar, J. V., Vinner, L., De Barros Damgaard, P., De La Fuente, C., Chan, J., Spence, J. P., Allentoft, M. E., Vimala, T., Racimo, F., Pinotti, T., Rasmussen, S.,

Margaryan, A., Orbegozo, M. I., Mylopotamitaki, D., Wooller, M., Bataille, C., Becerra-Valdivia, L., Chivall, D., Comeskey, D., … Willerslev, E. (2018). Early human dispersals within the Americas. *Science (New York, N.Y.)*, *362*(6419). https://doi.org/10.1126/SCIENCE.AAV2621

Moya Torres, A. (1997). *Ethnos Atlas etnográfico del Ecuador*. http://repositoriointerculturalidad.ec/jspui/handle/123456789/33472

Muzzey, D., Evans, E. A., & Lieber, C. (2015). Understanding the Basics of NGS: From Mechanism to Variant Calling. *Current Genetic Medicine Reports*, *3*(4), 158. https://doi.org/10.1007/S40142-015-0076-8

Nagar, S. D., Conley, A. B., Chande, A. T., Rishishwar, L., Sharma, S., Mariño-Ramírez, L., Aguinaga-Romero, G., González-Andrade, F., & Jordan, I. K. (2021). Genetic ancestry and ethnic identity in Ecuador. *Human Genetics and Genomics Advances*, *2*(4), 100050. https://doi.org/10.1016/J.XHGG.2021.100050

Nakatsuka, N., Lazaridis, I., Barbieri, C., Skoglund, P., Rohland, N., Mallick, S., Posth, C., Harkins-Kinkaid, K., Ferry, M., Harney, É., Michel, M., Stewardson, K., Novak-Forst, J., Capriles, J. M., Durruty, M. A., Álvarez, K. A., Beresford-Jones, D., Burger, R., Cadwallader, L., … Fehren-Schmitz, L. (2020). A Paleogenomic Reconstruction of the Deep Population History of the Andes. *Cell*, *181*(5), 1131-1145.e21. https://doi.org/10.1016/J.CELL.2020.04.015

Naue, J., Hörer, S., Sänger, T., Strobl, C., Hatzer-Grubwieser, P., Parson, W., & Lutz-Bonengel, S. (2015). Evidence for frequent and tissue-specific sequence heteroplasmy in human mitochondrial DNA. *Mitochondrion*, *20*(1), 82–94. https://doi.org/10.1016/J.MITO.2014.12.002

Nuñez, C., Baeta, M., Sosa, C., Casalod, Y., Ge, J., Budowle, B., & Martínez-Jarreta, B. (2010). Reconstructing the population history of Nicaragua by means of mtDNA, Y-chromosome STRs, and autosomal STR markers. *American Journal of Physical Anthropology*, *143*(4), 591–600. https://doi.org/10.1002/AJPA.21355

O'Rourke, D. H., & Raff, J. A. (2010). The Human Genetic History of the Americas: The Final Frontier. *Current Biology*, *20*(4), R202–R207. https://doi.org/10.1016/J.CUB.2009.11.051

Palencia, L., Valverde, L., Álvarez, A., Cainé, L. M., Cardoso, S., Alfonso-Sánchez, M. A., Pinheiro, M. F., & De Pancorbo, M. M. (2010). Mitochondrial DNA diversity in a population from Santa Catarina (Brazil): predominance of the European input.

International Journal of Legal Medicine, *124*(4), 331–336. https://doi.org/10.1007/S00414-010-0464-2

Parry, J. H. (1968). The Kingdom of Quito in the Seventeenth Century. Bureaucratic Politics in the Spanish Empire. *Hispanic American Historical Review*, *48*(4), 697–699. https://doi.org/10.1215/00182168-48.4.697

Parson, W., & Dür, A. (2007). EMPOP--a forensic mtDNA database. *Forensic Science International. Genetics*, *1*(2), 88–92. https://doi.org/10.1016/J.FSIGEN.2007.01.018

Parson, W., Gusmão, L., Hares, D. R., Irwin, J. A., Mayr, W. R., Morling, N., Pokorak, E., Prinz, M., Salas, A., Schneider, P. M., & Parsons, T. J. (2014). DNA Commission of the International Society for Forensic Genetics: revised and extended guidelines for mitochondrial DNA typing. *Forensic Science International. Genetics*, *13*, 134–142. https://doi.org/10.1016/J.FSIGEN.2014.07.010

Payne, B. A. I., Wilson, I. J., Yu-Wai-Man, P., Coxhead, J., Deehan, D., Horvath, R., Taylor, R. W., Samuels, D. C., Santibanez-Koref, M., & Chinnery, P. F. (2013). Universal heteroplasmy of human mitochondrial DNA. *Human Molecular Genetics*, *22*(2), 384–390. https://doi.org/10.1093/HMG/DDS435

Paz-y-Miño, C., & Burgos, G. (2015). La Historia del Ecuador contada por los genes. *Historia de Las Ciencias En El Ecuador*, *Enero*, 27–44. https://www.researchgate.net/profile/Cesar_Paz-y-Mino/publication/287201628_La_Historia_del_Ecuador_contada_por_los_genes/links/5672cda808aee7a427435ab0.pdf?origin=publication_list

Perego, U. A., Achilli, A., Angerhofer, N., Accetturo, M., Pala, M., Olivieri, A., Kashani, B. H., Ritchie, K. H., Scozzari, R., Kong, Q. P., Myres, N. M., Salas, A., Semino, O., Bandelt, H. J., Woodward, S. R., & Torroni, A. (2009). Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. *Current Biology : CB*, *19*(1), 1–8. https://doi.org/10.1016/J.CUB.2008.11.058

Pereira, V., & Gusmão, L. (2016). *Types of Genomes, Sequences and Genetic Markers (Repeats, SNPs, Indels, Haplotypes)*. 163–191. https://doi.org/10.1142/9781786340788_0009

Poletto, M. M., Malaghini, M., Silva, J. S., Bicalho, M. G., & Braun-Prado, K. (2019). Mitochondrial DNA control region diversity in a population from Parana state—

increasing the Brazilian forensic database. *International Journal of Legal Medicine*, *133*(2), 347–351. https://doi.org/10.1007/S00414-018-1886-5/TABLES/3

Prieto, L., Zimmermann, B., Goios, A., Rodriguez-Monge, A., Paneto, G. G., Alves, C., Alonso, A., Fridman, C., Cardoso, S., Lima, G., Anjos, M. J., Whittle, M. R., Montesino, M., Cicarelli, R. M. B., Rocha, A. M., Albarrán, C., De Pancorbo, M. M., Pinheiro, M. F., Carvalho, M., … Parson, W. (2011). The GHEP-EMPOP collaboration on mtDNA population data--A new resource for forensic casework. *Forensic Science International. Genetics*, *5*(2), 146–151. https://doi.org/10.1016/J.FSIGEN.2010.10.013

Rasmussen, M., Anzick, S. L., Waters, M. R., Skoglund, P., Degiorgio, M., Stafford, T. W., Rasmussen, S., Moltke, I., Albrechtsen, A., Doyle, S. M., Poznik, G. D., Gudmundsdottir, V., Yadav, R., Malaspinas, A. S., Samuel Stockton White, V., Allentoft, M. E., Cornejo, O. E., Tambets, K., Eriksson, A., … Willerslev, E. (2014). The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature 2014 506:7487*, *506*(7487), 225–229. https://doi.org/10.1038/nature13025

Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., Parra, M. V., Rojas, W., Duque, C., Mesa, N., García, L. F., Triana, O., Blair, S., Maestre, A., Dib, J. C., Bravi, C. M., Bailliet, G., Corach, D., Hünemeier, T., … Ruiz-Linares, A. (2012). Reconstructing Native American population history. *Nature 2012 488:7411*, *488*(7411), 370–374. https://doi.org/10.1038/nature11258

Relethford, J. H. (2017). *Biological Anthropology, Population Genetics, and Race*. *1*. https://doi.org/10.1093/OXFORDHB/9780190236953.013.20

Ribeiro, B. (2021). *Genetic characterization of the maternal lineages in Andean Colombian populations* [MSc. dissertation, Universidade do Porto]. https://search.proquest.com/openview/0bfbaf4aa6411ee704f4516601045424/1?pq-origsite=gscholar&cbl=2026366&diss=y

Rickards, O., Martínez-Labarga, C., Lum, J. K., de Stefano, G. F., & Cann, R. L. (1999). mtDNA History of the Cayapa Amerinds of Ecuador: Detection of Additional Founding Lineages for the Native American Populations. *The American Journal of Human Genetics*, *65*(2), 519–530. https://doi.org/10.1086/302513

Robin, E. D., & Wong, R. (1988). Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. *Journal of Cellular Physiology*, *136*(3), 507–513. https://doi.org/10.1002/JCP.1041360316

Rodriguez-Delfin, L. A., Rubin-de-Celis, V. E., & Zago, M. A. (2001). Genetic diversity in an Andean population from Peru and regional migration patterns of Amerindians in South America: data from Y chromosome and mitochondrial DNA. *Human Heredity*, *51*(1–2), 97–106. https://doi.org/10.1159/000022964

Roewer, L., Nothnagel, M., Gusmão, L., Gomes, V., González, M., Corach, D., Sala, A., Alechine, E., Palha, T., Santos, N., Ribeiro-dos-Santos, A., Geppert, M., Willuweit, S., Nagy, M., Zweynert, S., Baeta, M., Núñez, C., Martínez-Jarreta, B., González-Andrade, F., … Krawczak, M. (2013). Continent-Wide Decoupling of Y-Chromosomal Genetic Variation from Language and Geography in Native South Americans. *PLOS Genetics*, *9*(4), e1003460. https://doi.org/10.1371/JOURNAL.PGEN.1003460

Rostworowski, M. (2014). History of Tahuantinsuyo. *Lima, Instituto de Estudios Peruanos*, 279–302.

Rothhammer, F., & Dillehay, T. D. (2009). The late pleistocene colonization of South America: An interdisciplinary perspective. *Annals of Human Genetics*, *73*(5), 540–549. https://doi.org/10.1111/J.1469-1809.2009.00537.X

Rueda, R. (2001). Zambaje y autonomía: historia de la gente negra de la provincia de Esmeraldas, siglos XVI-XVIII - Universidad Andina Simón Bolívar. In *Quito, Municipalidad de Esmeraldas /Tehis*. https://www.uasb.edu.ec/publicacion/zambaje-y-autonomia-historia-de-la-gente-negra-de-la-provincia-de-esmeraldas-siglos-xvi-xviii/

Salas, A., Lovo-Gómez, J., Álvarez-Iglesias, V., Cerezo, M., Lareu, M. V., Macaulay, V., Richards, M. B., & Carracedo, Á. (2009). Mitochondrial Echoes of First Settlement and Genetic Continuity in El Salvador. *PLOS ONE*, *4*(9), e6882. https://doi.org/10.1371/JOURNAL.PONE.0006882

Salas, A., Richards, M., De la Fe, T., Lareu, M. V., Sobrino, B., Sánchez-Diz, P., Macaulay, V., & Carracedo, Á. (2002). The Making of the African mtDNA Landscape. *American Journal of Human Genetics*, *71*(5), 1082. https://doi.org/10.1086/344348

Salomon, F. (1980). Los señores étnicos de Quito en la época de los incas. In *Otavalo, Instituto Otavaleño de Antropología*. https://biblio.flacsoandes.edu.ec/libros/9877-opac

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular Cloning. Vol 2: A Laboratory Manual*. Cold Spring Horbor Laboratory Press. http://books.google.com.mx/books?id=VLGyXwAACAAJ

Sandoval, J. R., Lacerda, D. R., Jota, M. M. S., Robles-Ruiz, P., Danos, P., Paz-Y-Miño, C., Wells, S., Santos, F. R., & Fujita, R. (2020). Tracing the genetic history of the "Cañaris" from Ecuador and Peru using uniparental DNA markers. *BMC Genomics*, *21*(7), 1–10. https://doi.org/10.1186/S12864-020-06834-1/FIGURES/8

Sandoval, J. R., Lacerda, D. R., Jota, M. S. A., Salazar-Granara, A., Vieira, P. P. R., Acosta, O., Cuellar, C., Revollo, S., Fujita, R., Santos, F. R., & Genographic Project Consortium. (2013). The Genetic History of Indigenous Populations of the Peruvian and Bolivian Altiplano: The Legacy of the Uros. *PLoS ONE*, *8*(9), 73006. https://doi.org/10.1371/JOURNAL.PONE.0073006

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(12), 5463. https://doi.org/10.1073/PNAS.74.12.5463

Santangelo, R., González-Andrade, F., Børsting, C., Torroni, A., Pereira, V., & Morling, N. (2017). Analysis of ancestry informative markers in three main ethnic groups from Ecuador supports a trihybrid origin of Ecuadorians. *Forensic Science International. Genetics*, *31*, 29–33. https://doi.org/10.1016/J.FSIGEN.2017.08.012

Satoh, M., & Kuroiwa, T. (1991). Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Experimental Cell Research*, *196*(1), 137–140. https://doi.org/10.1016/0014-4827(91)90467-9

Scheib, C. L., Li, H., Desai, T., Link, V., Kendall, C., Dewar, G., Griffith, P. W., Mörseburg, A., Johnson, J. R., Potter, A., Kerr, S. L., Endicott, P., Lindo, J., Haber, M., Xue, Y., Tyler-Smith, C., Sandhu, M. S., Lorenz, J. G., Randall, T. D., … Kivisild, T. (2018). Ancient human parallel lineages within North America contributed to a coastal expansion. *Science (New York, N.Y.)*, *360*(6392), 1024–1027. https://doi.org/10.1126/SCIENCE.AAR6851

Schurr, T. G. (2004). The peopling of the New World: Perspectives from molecular anthropology. *Annual Review of Anthropology*, *33*, 551–583. https://doi.org/10.1146/ANNUREV.ANTHRO.33.070203.143932

Schurr, T. G., Ballinger, S. W., Gan, Y. Y., Hodge, J. A., Merriwether, D. A., Lawrence, D. N., Knowler, W. C., Weiss, K. M., & Wallace, D. C. (1990). Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. *American Journal of Human Genetics*, *46*(3), 613. /pmc/articles/PMC1683611/?report=abstract

Schurr, T. G., & Sherry, S. T. (2004). Mitochondrial DNA and Y chromosome diversity and the peopling of the Americas: evolutionary and demographic evidence. *American Journal of Human Biology: The Official Journal of the Human Biology Council*, *16*(4), 420–439. https://doi.org/10.1002/AJHB.20041

Schwartz, M., & Vissing, J. (2002). Paternal inheritance of mitochondrial DNA. *The New England Journal of Medicine*, *347*(8), 576–580. https://doi.org/10.1056/NEJMOA020350

Secher, B., Fregel, R., Larruga, J. M., Cabrera, V. M., Endicott, P., Pestano, J. J., & González, A. M. (2014). The history of the North African mitochondrial DNA haplogroup U6 gene flow into the African, Eurasian and American continents. *BMC Evolutionary Biology*, *14*(1), 1–17. https://doi.org/10.1186/1471-2148-14-109/COMMENTS

Shitara, H., Hayashi, J. I., Takahama, S., Kaneda, H., & Yonekawa, H. (1998). Maternal inheritance of mouse mtDNA in interspecific hybrids: segregation of the leaked paternal mtDNA followed by the prevention of subsequent paternal leakage. *Genetics*, *148*(2), 851–857. https://doi.org/10.1093/GENETICS/148.2.851

Simão, F., Ferreira, A. P., de Carvalho, E. F., Parson, W., & Gusmão, L. (2018). Defining mtDNA origins and population stratification in Rio de Janeiro. *Forensic Science International. Genetics*, *34*, 97–104. https://doi.org/10.1016/J.FSIGEN.2018.02.003

Simão, F., Ribeiro, J., Vullo, C., Catelli, L., Gomes, V., Xavier, C., Huber, G., Bodner, M., Quiroz, A., Ferreira, A. P., Carvalho, E. F., Parson, W., & Gusmão, L. (2021). The ancestry of eastern paraguay: A typical south american profile with a unique pattern of admixture. *Genes*, *12*(11), 1788. https://doi.org/10.3390/GENES12111788/S1

Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Röhl, A., Salas, A., Oppenheimer, S., Macaulay, V., & Richards, M. B. (2009). Correcting for Purifying

Selection: An Improved Human Mitochondrial Molecular Clock. *American Journal of Human Genetics*, *84*(6), 740. https://doi.org/10.1016/J.AJHG.2009.05.001

Sprecher, C. J., Puers, C., Lins, A. M., & Schumm, J. W. (1996). General approach to analysis of polymorphic short tandem repeat loci. *BioTechniques*, *20*(2), 266–276. https://doi.org/10.2144/96202RR04

Starikovskaya, E. B., Sukernik, R. I., Derbeneva, O. A., Volodko, N. V., Ruiz-Pesini, E., Torroni, A., Brown, M. D., Lott, M. T., Hosseini, S. H., Huoponen, K., & Wallace, D. C. (2005). Mitochondrial DNA diversity in indigenous populations of the southern extent of Siberia, and the origins of Native American haplogroups. *Annals of Human Genetics*, *69*(Pt 1), 67–89. https://doi.org/10.1046/J.1529-8817.2003.00127.X

Stewart, J. B., & Chinnery, P. F. (2015). The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. *Nature Reviews Genetics 2015 16:9*, *16*(9), 530–542. https://doi.org/10.1038/nrg3966

Stewart, J. E. B., Fisher, C. L., Aagaard, P. J., Wilson, M. R., Isenberg, A. R., Polanskey, D., Pokorak, E., DiZinno, J. A., & Budowle, B. (2001). Length variation in HV2 of the human mitochondrial DNA control region. *Journal of Forensic Sciences*, *46*(4), 15059J. https://doi.org/10.1520/JFS15059J

Sutovsky, P., Moreno, R. D., Ramalho-Santos, J., Dominko, T., Simerly, C., & Schatten, G. (1999). Ubiquitin tag for sperm mitochondria. *Nature*, *402*(6760), 371–372. https://doi.org/10.1038/46466

Szibor, R., Plate, I., Heinrich, M., Michael, M., Schöning, R., Wittig, H., & Lutz-Bonengel, S. (2007). Mitochondrial D-loop (CA)n repeat length heteroplasmy: frequency in a German population sample and inheritance studies in two pedigrees. *International Journal of Legal Medicine*, *121*(3), 207–213. https://doi.org/10.1007/S00414-006-0096-8

T, M. (2004). Mitochondrial DNA Heteroplasmy. *Forensic Science Review*, *16*(1), 1–20. https://pubmed.ncbi.nlm.nih.gov/26256810/

Taboada-Echalar, P., Álvarez-Iglesias, V., Heinz, T., Vidal-Bralo, L., Gómez-Carballa, A., Catelli, L., Pardo-Seco, J., Pastoriza, A., Carracedo, Á., Torres-Balanza, A., Rocabado, O., Vullo, C., & Salas, A. (2013). The Genetic Legacy of the Pre-Colonial Period in Contemporary Bolivians. *PLOS ONE*, *8*(3), e58980. https://doi.org/10.1371/JOURNAL.PONE.0058980

Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D. G., Mulligan, C. J., Bravi, C. M., Rickards, O., Martinez-Labarga, C., Khusnutdinova, E. K., Fedorova, S. A., Golubenko, M. V., Stepanov, V. A., Gubina, M. A., Zhadanov, S. I., Ossipova, L. P., Damba, L., Voevoda, M. I., Dipierri, J. E., … Malhi, R. S. (2007). Beringian Standstill and Spread of Native American Founders. *PLOS ONE*, *2*(9), e829. https://doi.org/10.1371/JOURNAL.PONE.0000829

Tarazona-Santos, E., Carvalho-Silva, D. R., Pettener, D., Luiselli, D., De Stefano, G. F., Labarga, C. M., Rickards, O., Tyler-Smith, C., Pena, S. D. J., & Santos, F. R. (2001). Genetic Differentiation in South Amerindians Is Related to Environmental and Cultural Diversity: Evidence from the Y Chromosome. *The American Journal of Human Genetics*, *68*(6), 1485–1496. https://doi.org/10.1086/320601

Thompson, W. E., Ramalho-Santos, J., & Sutovsky, P. (2003). Ubiquitination of prohibitin in mammalian sperm mitochondria: possible roles in the regulation of mitochondrial inheritance and sperm quality control. *Biology of Reproduction*, *69*(1), 254–260. https://doi.org/10.1095/BIOLREPROD.102.010975

Torroni, A., Huoponen, K., Francalacci, P., Petrozzi, M., Morelli, L., Scozzari, R., Obinu, D., Savontaus, M. L., & Wallace, D. C. (1996). Classification of European mtDNAs From an Analysis of Three European Populations. *Genetics*, *144*(4), 1835–1850. https://doi.org/10.1093/GENETICS/144.4.1835

Torroni, A., Schurr, T. G., Yang, C. C., Szathmary, E. J. E., Williams, R. C., Schanfield, M. S., Troup, G. A., Knowler, W. C., Lawrence, D. N., Weiss, K. M., & Wallace, D. C. (1992). Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. *Genetics*, *130*(1), 153–162. https://doi.org/10.1093/genetics/130.1.153

Torroni, A., Sukernik, R. I., Schurr, T. G., Starikovskaya, Y. B., Cabell, M. F., Crawford, M. H., Comuzzie, A. G., & Wallace, D. C. (1993). mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. *American Journal of Human Genetics*, *53*(3), 591. /pmc/articles/PMC1682407/?report=abstract

Toscanini, U., Gaviria, A., Pardo-Seco, J., Gómez-Carballa, A., Moscoso, F., Vela, M., Cobos, S., Lupero, A., Zambrano, A. K., Martinón-Torres, F., Carabajo-Marcillo, A., Yunga-León, R., Ugalde-Noritz, N., Ordoñez-Ugalde, A., & Salas, A. (2018). The geographic mosaic of Ecuadorian Y-chromosome ancestry. *Forensic Science International. Genetics*, *33*, 59–65. https://doi.org/10.1016/J.FSIGEN.2017.11.011

van Oven, M., & Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human Mutation*, *30*(2). https://doi.org/10.1002/HUMU.20921

Vásquez S., Lola., Saltos Galarza, N., & Fundación "José Peralta." (2013). *Ecuador: su realidad* . 420.

Villaescusa, P., Seidel, M., Nothnagel, M., Pinotti, T., González-Andrade, F., Alvarez-Gila, O., M. de Pancorbo, M., & Roewer, L. (2021). 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 Science International: Genetics*, *51*. https://doi.org/10.1016/J.FSIGEN.2020.102427

Wallace, D. C., & Torroni, A. (2009). American Indian prehistory as written in the mitochondrial DNA: a review. 1992. *Human Biology*, *81*(5–6), 509–521. https://doi.org/10.3378/027.081.0602

Watkins, W. S., Xing, J., Huff, C., Witherspoon, D. J., Zhang, Y., Perego, U. A., Woodward, S. R., & Jorde, L. B. (2012). Genetic analysis of ancestry, admixture and selection in Bolivian and Totonac populations of the New World. *BMC Genetics*, *13*(1), 1–14. https://doi.org/10.1186/1471-2156-13-39/TABLES/2

Willerslev, E., & Meltzer, D. J. (2021). Peopling of the Americas as inferred from ancient genomics. *Nature 2021 594:7863*, *594*(7863), 356–364. https://doi.org/10.1038/s41586-021-03499-y

Zambrano, A. K., Gaviria, A., Cobos-Navarrete, S., Gruezo, C., Rodríguez-Pollit, C., Armendáriz-Castillo, I., García-Cárdenas, J. M., Guerrero, S., López-Cortés, A., Leone, P. E., Pérez-Villa, A., Guevara-Ramírez, P., Yumiceba, V., Fiallos, G., Vela, M., & Paz-y-Miño, C. (2019). The three-hybrid genetic composition of an Ecuadorian population using AIMs-InDels compared with autosomes, mitochondrial DNA and Y chromosome data. *Scientific Reports 2019 9:1*, *9*(1), 1–8. https://doi.org/10.1038/s41598-019-45723-w

Zegura, S. L., Karafet, T. M., Zhivotovsky, L. A., & Hammer, M. F. (2004). High-Resolution SNPs and Microsatellite Haplotypes Point to a Single, Recent Entry of Native American Y Chromosomes into the Americas. *Molecular Biology and Evolution*, *21*(1), 164–175. https://doi.org/10.1093/MOLBEV/MSH009

Anexos

Appendix 1. Primers employed in mtDNA analysis of Andean Mestizos from Ecuador.

Primer	Sequence (5' to 3')	Region to analyse
L15978*	5' CACCATTAGCACCCAAAGCT 3'	HVS-I
H017	5' TGA TAG ACC TGT GAT CCA TCG TGA 3'	HVS-I
L16536	5' CCC ACA CGT TCC CCT TAA AT 3'	HVS-II
H649*	5' TTT GTT TAT GGG GTG ATG TGA 3'	HVS-III
L314	5' CCG CTT CTG GCC ACA GCA CT 3'	HVS-III
H408	5' CTG TTA AAA GTG CAT ACC GCC A 3'	HVS-II
H484	5' TGAGATTAGTAGTATGGGAG 3'	HVS-II
L 16209	5' CCA TGC TTA CAA GCA AGT 3'	HVS-I, HVS-II
H159	5' AAA TAA TAG GAT GAG GCA GGA ATC 3'	HVS-I, HVS-II

*Primers used for initial amplification of mtDNA CR.

Appendix 2. List of populations used for genetic comparisons.

Region/City	Country	Population	N	mtDNA region	Reference
Andes North	Ecuador	Admixed	76	16024-576	This study
Andes Center	Ecuador	Admixed	66	16024-576	This study
Andes South	Ecuador	Admixed	72	16024-576	This study
Medellín	Colombia	Admixed	108	16024-576	(Auton et al., 2015)
Lima	Peru	Admixed	83	16024-576	(Auton et al., 2015)
Caracas	Venezuela	Admixed	101	16024-576	(Castro De Guerra et al., 2012)
East	Paraguay	Admixed	417	16024-576	(Simão et al., 2021)
North	Argentina	Admixed	97	16024-576	(Bobillo et al., 2010)
Center	Argentina	Admixed	193	16024-576	(Bobillo et al., 2010)
South	Argentina	Admixed	47	16024-576	(Bobillo et al., 2010)
La Paz	Bolivia	Admixed	108	16024-576	(Taboada-Echalar et al., 2013)
Paraná	Brazil	Admixed	122	16024-576	(Poletto et al., 2019)
Rio de Janeiro	Brazil	Admixed	205	16024-576	(Simão et al., 2018)
Sao Paulo	Brazil	Admixed	142	16024-576	(Prieto et al., 2011)
N/A	Ecuador	Admixed	144	16024-576	(Brandini et al., 2018)
N/A	Ecuador	Admixed	42	16024-576	(Baeta et al., 2012)
Amazon	Ecuador	Native	65	16024-576	(Baeta et al., 2012)
North Andes	Ecuador	Native	25	16024-576	(Sandoval et al., 2020)
Coast	Ecuador	Admixed	75	16024-576	(non-published)

Appendix 3. List of the haplotypes and corresponding haplogroups found in the Andean Mestizos. Haplogroups were assignment using EMPOP.

Andean Region	Sample ID	Haplogroup	Haplotype Polymorphic sites (range: 0-576; 16024-16569)
North	N1	C1b	73G 249del 263G 290del 291del 315.1C 489C 493G 523del 524del 16223T 16298C 16325C 16327T 16519C
North	N2	D4h3a6	73G 143A 152C 263G 309.1C 315.1C 489C 16185T 16192T 16223T 16301T 16342C 16362C 16463G
North	N3	B4b	73G 93G 263G 309.1C 315.1C 499A 573.1C 573.2C 16183C 16189C 16217C 16519C
North	N4	B4b	73G 152C 204C 209C 263G 309.1C 315.1C 499A 573.1C 573.2C 573.3C 573.4C 16183C 16188T 16189C 16193.1C 16217C 16362C 16519C
North	N5	B4b	73G 263G 309.1C 315.1C 499A 573.1C 573.2C 573.3C 16183C 16189C 16217C 16519C
North	N6	A2ac	64T 73G 146C 153G 235G 249G 263G 309.1C 309.2C 315.1C 523del 524del 16111T 16213A 16223T 16290T 16319A 16362C
North	N7	A2ac	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523del 524del 16111T 16213A 16223T 16290T 16294T 16319A 16362C 16519C
North	N8	A2+(64)	64T 73G 146C 235G 263G 315.1C 523del 524del 16111T 16129A 16223T 16290T 16311C 16319A 16362C
North	N9	A2+(64)	64T 73G 103.1T 103.2C 111C 146C 151T 153G 235G 263G 309del 315.1C 523del 524del 16111T 16213A 16223T 16290T 16319A 16362C 16519C
North	N10	A2+(64)+16189	64T 73G 146C 153G 235G 263G 309.1C 309.2C 315.1C 523del 524del 16111T 16183C 16189C 16193.1C 16223T 16290T 16319A 16362C
North	N11	R0a+60.1T	58C 60.1T 64T 263G 309.1C 315.1C 16126C 16362C 16519C
North	N12	B4b	73G 103A 150T 152C 263G 309.1C 315.1C 499A 16183C 16189C 16193.1C 16217C 16266T 16362C
North	N13	A2+(64)	64T 73G 125C 127C 146C 153G 263G 315.1C 499A 523del 524del 16111T 16189C 16223T 16290T 16319A 16362C 16519C
North	N14	D4h3a	73G 152C 263G 309.1C 315.1C 489C 16223T 16241G 16301T 16342C
North	N15	B4b	73G 103A 183G 263G 315.1C 499A 16182C 16183C 16189C 16217C 16234T 16266T 16519C
North	N16	C1d+194	73G 194T 249del 263G 290del 291del 309.1C 315.1C 489C 523del 524del 16051G 16111A 16129A 16162G 16189C 16223T 16234T 16298C 16325C 16327T 16519C
North	N17	D4h3a5	73G 152C 263G 315.1C 489C 573del 16223T 16241G 16311C 16342C 16362C 16445C
North	N18	B4	73G 140T 198T 263G 309.1C 309.2C 315.1C 499A 523del 524del 16182C 16183C 16189C 16217C 16295T 16301T 16311C 16519C
North	N19	B4b	73G 151T 152C 263G 309.1C 315.1C 499A 16183C 16189C 16217C 16291T 16519C
North	N20	A2+(64)	64T 73G 93G 125C 126G 127C 146C 153G 235G 263G 309.1C 315.1C 523del 524del 16111T 16126C 16213A 16223T 16290T 16319A 16320T 16362C
North	N21	D1f	73G 207A 263G 315.1C 489C 513A 524.1A 524.2C 16142T 16207G 16223T 16325C 16362C 16519C
North	N22	B2o1	73G 263G 315.1C 499A 514T 16092C 16145A 16157C 16183C 16189C 16193.1C 16217C 16274A 16294T 16519C
North	N23	C1b	73G 249del 263G 290del 291del 315.1C 489C 493G 523del 524del 16223T 16298C 16325C 16327T 16519C
North	N24	C1b	73G 249del 263G 290del 291del 309.1C 315.1C 489C 493G 523del 524del 16166del 16223T 16298C 16325C 16327T 16519C
North	N25	B2h	73G 263G 309.1C 309.2C 315.1C 499A 573.1C 16183C 16189C 16193.1C 16217C 16362C 16468C 16519C
North	N26	C1b	73G 152C 226C 249del 263G 290del 291del 315.1C 382A 489C 493G 523del 524del 16185T 16223T 16298C 16311C 16325C 16327T
North	N27	A2+(64)	73G 146C 153G 195C 235G 249G 263G 309.1C 315.1C 523del 524del 573.1C 573.2C 573.3C 573.4C 573.5C 573.6C 16111T 16213A 16223T 16290T 16319A 16362C 16519C
North	N28	B4b	73G 263G 315.1C 368G 499A 573.1C 16183C 16189C 16193.1C 16217C 16519C
North	N29	A2ac	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523del 524del 16092C 16111T 16213A 16223T 16290T 16319A 16362C 16506C
North	N30	D1	73G 196C 263G 309.1C 315.1C 489C 16223T 16234T 16325C 16362C 16519C
North	N31	C1d+194	73G 194T 249del 263G 290del 291del 315.1C 489C 523del 524del 16051G 16094C 16223T 16298C 16325C 16327T
North	N32	B2h	73G 263G 309.1C 309.2C 315.1C 499A 573.1C 573.2C 16183C 16189C 16193.1C 16217C 16468C 16519C
North	N33	B4b	73G 152C 203A 206C 263G 309.1C 315.1C 499A 16172C 16182C 16183C 16189C 16217C 16255A 16519C
North	N34	C1	73G 115C 116del 249del 263G 290del 291del 315.1C 489C 523del 524del 16223T 16298C 16325C 16327T 16390A
North	N35	HV	263G 309.1C 315.1C 16189C 16356C 16519C
North	N36	B4b	73G 103A 183G 263G 309.1C 309.2C 315.1C 499A 16182C 16183C 16189C 16217C 16234T 16266T 16319A 16519C

Andean Region	Sample ID	Haplogroup	Haplotype Polymorphic sites (range: 0-576; 16024-16569)																			
North	N37	C1d	73G	182T	249del	263G	290del	291del	315.1C	489C	16051G	16223T	16298C	16325C	16327T							
North	N38	B4b	73G	93G	263G	315.1C	408A	499A	573.1C	16183C	16189C	16217C	16519C									
North	N39	C1b	73G	249del	263G	290del	291del	309.1C	315.1C	489C	493G	523del	524del	16223T	16298C	16325C	16327T	16343G	16519C			
North	N40	B4b	73G	195C	263G	309.1C	315.1C	499A	523del	524del	573.1C	573.2C	573.3C	573.4C	573.5C	16183C	16189C	16217C	16258C	16519C		
North	N41	C1	73G	249del	263G	290del	291del	315.1C	489C	16223T	16260T	16278T	16298C	16325C	16327T	16519C						
North	N42	D4	73G	125C	127C	143A	152C	195C	263G	309.1C	315.1C	489C	16129A	16222T	16223T	16325C	16390A					
North	N43	C1c	10C	73G	146C	195C	249del	263G	290del	291del	309.1C	315.1C	489C	16223T	16298C	16325C	16327T	16526A				
North	N44	B2h	73G	263G	309.1C	315.1C	499A	573.1C	573.2C	573.3C	573.4C	573.5C	16157C	16183C	16193.1C	16217C	16468C	16519C				
North	N45	B4b	63del	66C	73G	263G	309.1C	315.1C	499A	16183C	16187T	16189C	16217C	16287T	16519C							
North	N46	B2	73G	152C	263G	309.1C	315.1C	499A	523del	524del	16092C	16183C	16189C	16193.1C	16217C	16241G	16278T	16324C	16519C			
North	N47	D4h3a6	73G	143A	152C	263G	309.1C	315.1C	489C	16129A	16185T	16192T	16223T	16241G	16301T	16342C	16362C	16463G				
North	N48	D1a	73G	125C	127C	143A	152C	195C	263G	309.1C	315.1C	489C	16223T	16325C	16362C	16519C						
North	N49	B2	73G	152C	207A	263G	315.1C	499A	16092C	16182C	16183C	16189C	16217C	16465T	16519C							
North	N50	C1b	73G	249del	263G	290del	291del	309.1C	315.1C	489C	493G	523del	524del	16126C	16223T	16298C	16325C	16327T	16365T			
North	N51	B2	73G	152C	263G	309.1C	315.1C	499A	523del	524del	16092C	16183C	16189C	16193.1C	16217C	16241G	16278T	16324C	16519C			
North	N52	B4b	73G	152C	263G	309.1C	315.1C	499A	16153A	16183C	16189C	16193.1C	16217C	16519C								
North	N53	B2h	73G	263G	309.1C	309.2C	315.1C	499A	573.1C	573.2C	573.3C	573.4C	573.5C	16157C	16183C	16189C	16193.1C	16217C	16468C	16519C		
North	N54	B4b	73G	146C	263G	315.1C	499A	16183C	16189C	16217C	16266T	16485A	16519C									
North	N55	C1b	73G	249del	263G	272G	290del	291del	309.1C	315.1C	489C	493G	523del	524del	16223T	16298C	16325C	16327T	16519C			
North	N56	L2b2a	73G	146C	150T	152C	182T	195C	198T	204C	263G	309.1C	315.1C	16114A	16129A	16213A	16223T	16278T	16354T	16390A		
North	N57	U4c1	73G	195C	263G	315.1C	499A	16179T	16356C	16519C												
North	N58	D4h3a	73G	263G	309.1C	315.1C	489C	573del	16223T	16241G	16301T	16311C	16342C	16362C								
North	N59	B2h	73G	263G	309.1C	309.2C	315.1C	499A	573.1C	573.2C	573.3C	16183C	16189C	16217C	16468C	16519C						
North	N60	B4b	73G	93G	263G	315.1C	408A	499A	573.1C	573.2C	16183C	16189C	16217C	16519C								
North	N61	B4b	73G	152C	263G	309.1C	315.1C	499A	523del	524del	16051G	16183C	16189C	16193.1C	16217C	16360T	16519C					
North	N62	C1b	73G	150T	249del	263G	290del	291del	309.1C	315.1C	489C	493G	523del	524del	16223T	16298C	16325C	16327T	16519C			
North	N63	B2h	73G	140T	263G	309.1C	315.1C	499A	524.1A	524.2C	16183C	16189C	16193.1C	16217C	16295T	16301T	16468C	16519C				
North	N64	A2+(64)	64T	73G	146C	153G	215G	235G	263G	309.1C	309.2C	315.1C	523del	524del	16111T	16223T	16290T	16319A	16362C	16519C		
North	N65	A2ac	64T	73G	146C	153G	235G	263G	279C	315.1C	523del	524del	16111T	16213A	16223T	16290T	16319A	16362C				
North	N66	D1f	73G	263G	289G	309.1C	315.1C	489C	16142T	16223T	16325C	16362C	16519C									
North	N67	L1b	73G	152C	182T	185T	188G	195C	247A	263G	315.1C	357G	523del	524del	16126C	16187T	16189C	16223T	16264T	16270T	16278T	16293G
			16311C	16519C																		
North	N68	C1b	73G	185C	195C	249del	263G	290del	291del	315.1C	470G	489C	493G	523del	524del	16178C	16223T	16298C	16325C	16327T	16519C	
North	N69	C1	73G	249del	263G	290del	291del	315.1C	489C	16223T	16298C	16325C	16327T	16390A								
North	N70	B2	73G	152C	182T	263G	309.1C	315.1C	499A	16051G	16183C	16189C	16193.1C	16217C	16519C							
North	N71	D1f	73G	239C	263G	315.1C	489C	513A	524.1A	524.2C	16142T	16207G	16223T	16325C	16362C	16519C						
North	N72	D4h3a6	73G	143A	152C	263G	309.1C	309.2C	315.1C	489C	16129A	16185T	16192T	16223T	16241G	16301T	16342C	16362C	16399R	16463G		

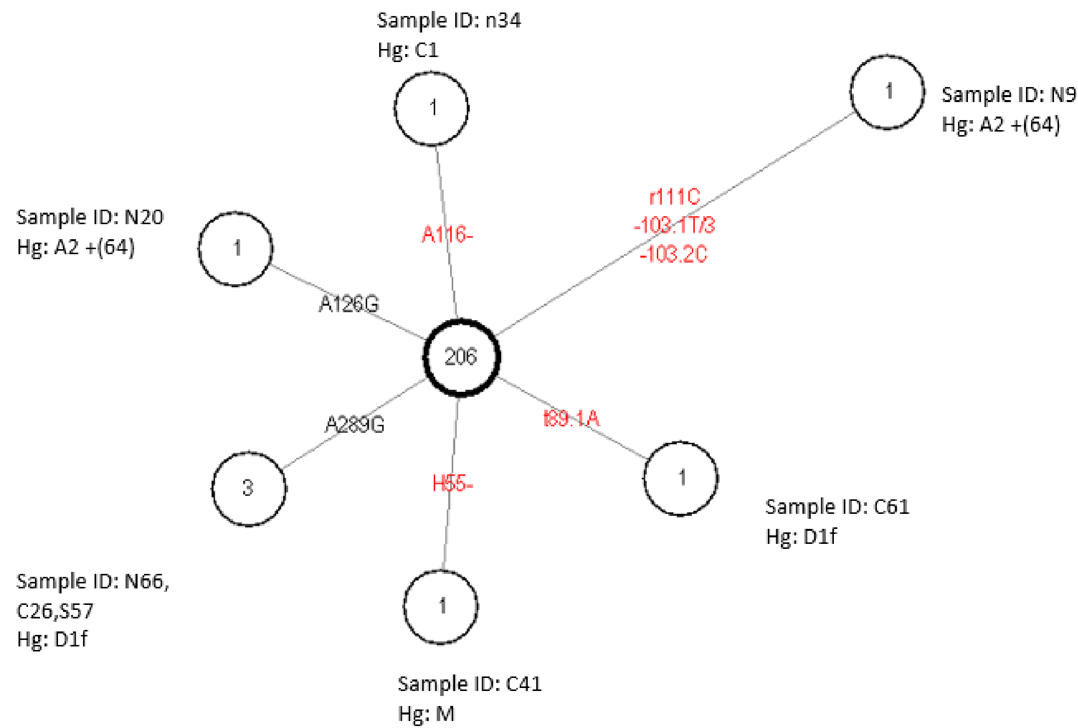
Andean Region	Sample ID	Haplogroup	Haplotype Polymorphic sites (range: 0-576; 16024-16569)																		
North	N73	A2ac	64T	73G	146C	153G	235G	263G	309.1C	309.2C	315.1C	523del	524del	16111T	16213A	16223T	16290T	16319A	16356C	16362C	
North	N74	C1b	73G	249del	263G	290del	291del	309.1C	315.1C	489C	493G	523del	524del	16223T	16298C	16325C	16327T	16354T			
North	N75	B4a1c3	73G	103A	263G	309.1C	309.2C	315.1C	499A	16183C	16189C	16194C	16195C	16217C	16298C	16352C					
North	N76	B4b	73G	263G	309.1C	309.2C	315.1C	499A	16183C	16189C	16193.1C	16217C	16266T	16519C							
Center	C1	B2b2	73G	263G	309.1C	315.1C	499A	573.1C	16145A	16183C	16189C	16193.1C	16217C	16519C							
Center	C2	B4b	73G	146C	150T	263G	309.1C	309.2C	315.1C	499A	16182C	16183C	16189C	16217C	16274A	16519C					
Center	C3	B4b	53A	73G	152C	263G	309.1C	315.1C	499A	573.1C	573.2C	573.3C	573.4C	573.5C	16183C	16189C	16217C	16519C			
Center	C4	C1d	73G	249del	263G	290del	291del	315.1C	489C	16051G	16223T	16260T	16298C	16325C	16327T	16519C					
Center	C5	A2	64T	73G	146C	153G	235G	263G	315.1C	523del	524del	16111T	16223T	16261T	16278T	16290T	16319A	16362C	16519C		
Center	C6	B2h	73G	263G	309.1C	315.1C	499A	573.1C	573.2C	573.3C	16183C	16189C	16217C	16468C	16519C						
Center	C7	B4b	73G	263G	309.1C	309.2C	315.1C	499A	16183C	16189C	16193.1C	16217C	16266T	16519C							
Center	C8	C1b	73G	152C	249del	263G	290del	291del	315.1C	489C	493G	523del	524del	16071T	16223T	16298C	16325C	16327T	16445C		
Center	C9	B2h	73G	263G	309.1C	315.1C	499A	573.1C	16183C	16189C	16217C	16468C	16519C								
Center	C10	B2b2	73G	263G	309.1C	315.1C	499A	504C	514T	16145A	16157C	16182C	16183C	16189C	16217C	16294T	16311C				
Center	C11	D4	73G	263G	309.1C	309.2C	315.1C	489C	573.1C	16153A	16223T	16243C	16325C	16362C							
Center	C12	B4b	73G	103A	150T	263G	309.1C	315.1C	499A	16183C	16189C	16217C	16256T	16266T							
Center	C13	B4b	73G	183G	263G	309.1C	315.1C	499A	16182C	16183C	16189C	16217C	16234T	16266T	16319A	16519C					
Center	C14	B2h	72C	73G	263G	309.1C	315.1C	499A	573.1C	573.2C	573.3C	573.4C	573.5C	573.6C	16157C	16182C	16183C	16189C	16217C	16468C	16519C
Center	C15	A2+(64)	64T	73G	146C	153G	215G	235G	263G	309.1C	315.1C	523del	524del	16111T	16223T	16290T	16319A	16338G	16362C	16519C	
Center	C16	B4b	73G	103A	183G	263G	315.1C	499A	16182C	16183C	16189C	16217C	16234T	16266T	16319A	16519C					
Center	C17	A2+(64)	64T	73G	146C	153G	195C	235G	263G	309.1C	315.1C	523del	524del	16111T	16223T	16290T	16319A	16362C	16519C		
Center	C18	B4b	73G	143A	263G	315.1C	499A	16183C	16189C	16217C	16289G	16362C	16519C								
Center	C19	B2h	73G	263G	309.1C	309.2C	315.1C	499A	573.1C	573.2C	573.3C	16183C	16189C	16217C	16468C	16519C					
Center	C20	D4h3a6	73G	143A	152C	263G	309.1C	315.1C	489C	16129A	16185T	16192T	16223T	16241G	16301T	16342C	16362C	16463G			
Center	C21	C1b	73G	199C	249del	263G	290del	291del	315.1C	489C	493G	523del	524del	16192T	16223T	16298C	16325C	16327T	16365T		
Center	C22	C1b1	73G	150T	249del	263G	290del	291del	309.1C	315.1C	489C	493G	523del	524del	16092C	16223T	16274A	16298C	16325C	16327T	16519C
Center	C23	B2h	73G	263G	309.1C	315.1C	499A	573.1C	573.2C	573.3C	573.4C	16157C	16183C	16189C	16193.1C	16217C	16468C	16519C			
Center	C24	D1f	73G	207A	263G	315.1C	489C	513A	524.1A	524.2C	16142T	16207G	16223T	16325C	16362C	16519C					
Center	C25	D4h3a5	73G	152C	263G	315.1C	489C	573del	16223T	16241G	16311C	16342C	16362C	16445C							
Center	C26	D1f	73G	263G	289G	315.1C	489C	16086C	16142T	16223T	16325C	16362C	16519C								
Center	C27	B4b	73G	103A	183G	263G	309.1C	309.2C	315.1C	499A	16183C	16189C	16193.1C	16217C	16234T	16266T	16519C				
Center	C28	B4b	73G	152C	263G	309.1C	315.1C	499A	16051G	16129A	16182C	16183C	16189C	16217C	16274A	16519C					
Center	C29	B2h	73G	263G	311T	315.1C	315.2C	499A	573.1C	16183C	16189C	16193.1C	16217C	16468C	16519C						
Center	C30	D1f	64T	73G	263G	315.1C	489C	16142T	16189C	16222T	16223T	16325C	16362C	16519C							
Center	C31	A2+(64)	64T	73G	146C	153G	215G	235G	263G	309.1C	315.1C	523del	524del	16111T	16223T	16290T	16319A	16338G	16362C	16519C	
Center	C32	B2h	73G	263G	315.1C	499A	573.1C	573.2C	573.3C	573.4C	16157C	16183C	16189C	16217C	16468C	16519C					

Andean Region	Sample ID	Haplogroup	Haplotype Polymorphic sites (range: 0-576; 16024-16569)
Center	C33	B4b	73G 103A 183G 263G 309.1C 309.2C 315.1C 499A 16182C 16183C 16189C 16217C 16234T 16266T 16274A 16519C
Center	C34	C1b	73G 249del 263G 290del 291del 309.1C 315.1C 489C 493G 523del 524del 16223T 16298C 16325C 16327T 16365T
Center	C35	B2h	55C 73G 263G 309.1C 315.1C 499A 573.1C 573.2C 573.3C 573.4C 16183C 16189C 16193.1C 16217C 16468C 16519C
Center	C36	A2+(64)+@153	64T 73G 143A 146C 235G 263G 309.1C 315.1C 523del 524del 529A 16111T 16134T 16223T 16290T 16319A 16362C
Center	C37	C1d	73G 249del 263G 290del 291del 309.1C 315.1C 489C 523del 524del 16051G 16093C 16223T 16298C 16325C 16327T 16519C
Center	C38	D4h3a5	73G 152C 263G 309.1C 315.1C 489C 573del 16223T 16241G 16311C 16320T 16342C 16362C 16445C
Center	C39	D1f	55C 56G 73G 185A 204C 263G 315.1C 489C 16142T 16223T 16325C 16362C 16519C
Center	C40	D4h3a6	73G 143A 151T 152C 263G 309.1C 315.1C 489C 16129A 16185T 16192T 16223T 16241G 16301T 16342C 16362C 16463G
Center	C41	M	55del 73G 263G 309.1C 315.1C 489C 16182G 16189C 16223T 16324C 16325C 16362C
Center	C42	C1b1	73G 150T 249del 263G 290del 291del 309.1C 315.1C 489C 493G 523del 524del 16092C 16223T 16274A 16298C 16325C 16327T 16352C 16519C
Center	C43	A2ac	64T 73G 146C 153G 235G 263G 279C 309.1C 315.1C 523del 524del 16111T 16213A 16290T 16295T 16319A 16362C
Center	C44	A2+(64)	64T 73G 146C 153G 235G 263G 309.1C 309.2C 315.1C 523del 524del 16111T 16223T 16319A 16362C 16519C
Center	C45	C1b	73G 249del 263G 290del 291del 309.1C 315.1C 489C 493G 523del 524del 16223T 16298C 16325C 16327T
Center	C46	A2+(64)	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523del 524del 16111T 16223T 16290T 16319A 16362C
Center	C47	D1f+16189	73G 153G 263G 309.1C 309.2C 315.1C 489C 524.1A 524.2C 16142T 16182C 16183C 16189C 16223T 16325C 16362C 16519C
Center	C48	C1b	73G 249del 263G 290del 291del 309.1C 315.1C 489C 493G 16167T 16192T 16223T 16298C 16311C 16325C 16327T
Center	C49	B2h	73G 263G 309.1C 315.1C 499A 573.1C 16183C 16189C 16217C 16468C 16519C
Center	C50	C1b	73G 189G 249del 263G 290del 291del 309.1C 315.1C 489C 493G 523del 524del 16223T 16266G 16298C 16325C 16327T 16519C
Center	C51	B4b	73G 152C 263G 309.1C 309.2C 315.1C 499A 16183C 16189C 16193.1C 16217C 16286T 16291T 16497G 16519C
Center	C52	M	73G 263G 309.1C 315.1C 489C 16182G 16189C 16223T 16324C 16325C 16362C
Center	C53	A2ac	64T 73G 146C 153G 235G 263G 279C 309.1C 315.1C 523del 524del 16111T 16192T 16213A 16223T 16290T 16319A 16362C 16519C
Center	C54	B2h	73G 263G 315.1C 499A 573.1C 573.2C 573.3C 573.4C 16157C 16183C 16189C 16217C 16468C 16519C
Center	C55	D1f	55C 56G 73G 185R 204C 263G 315.1C 489C 16142T 16223T 16325C 16362C 16519C
Center	C56	C1b	73G 249del 263G 290del 291del 309.1C 315.1C 489C 493G 523del 524del 16223T 16298C 16325C 16327T 16365T
Center	C57	B2h	73G 263G 309.1C 309.2C 315.1C 499A 573.1C 16183C 16189C 16193.1C 16217C 16468C 16519C
Center	C58	A2ac	64T 73G 146C 153G 235G 263G 279C 315.1C 523del 524del 16111T 16213A 16223T 16261T 16290T 16319A 16362C
Center	C59	A2+(64)	64T 73G 146C 153G 195C 235G 249G 263G 309.1C 315.1C 523del 524del 16111T 16213A 16223T 16290T 16319A 16362C
Center	C60	C1b1	73G 249del 263G 290del 291del 315.1C 455del 460C 463del 489C 493G 523del 524del 16223T 16274A 16298C 16325C 16327T
Center	C61	D1f	73G 89.1A 97A 105del 106del 107del 108del 109del 110del 263G 309.1C 315.1C 489C 16142T 16223T 16325C 16362C 16519C
Center	C62	C1b	73G 249del 263G 290del 291del 309.1C 315.1C 489C 493G 523del 524del 16223T 16298C 16325C 16327T 16354T
Center	C63	C1	73G 194T 249del 263G 290del 291del 309.1C 315.1C 489C 523del 524del 16223T 16234T 16243C 16298C 16325C 16327T 16519C
Center	C64	B4b	73G 94A 152C 263G 309.1C 309.2C 315.1C 499A 16051G 16129A 16182C 16183C 16189C 16217C 16274A 16519C
Center	C65	A2+(64)	64T 73G 146C 153G 235G 263G 315.1C 523del 524del 16111T 16223T 16290T 16319A 16362C 16519C
Center	C66	B4b	73G 103A 183G 195C 263G 315.1C 499A 16182C 16183C 16189C 16217C 16234T 16266T 16319A 16519C
South	S1	A2+(64)	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523del 524del 16111T 16223T 16319A 16362C 16519C
South	S2	B4b	73G 263G 309.1C 315.1C 499A 16183C 16189C 16217C 16437C 16519C

Andean Region	Sample ID	Haplogroup	Haplotype Polymorphic sites (range: 0-576; 16024-16569)																				
South	S3	D1	73G	263G	309.1C	315.1C	489C	16223T	16325C	16362C													
South	S4	B4b	73G	263G	309.1C	315.1C	499A	16173T	16182C	16183C	16189C	16217C											
South	S5	A2ac	73G	146C	153G	235G	263G	279C	315.1C	523del	524del	16093C	16111T	16213A	16223T	16290T	16319A	16362C	16519C				
South	S6	HV9b	263G	315.1C	523del	524del	16249C	16311C															
South	S7	A2+(64)	64T	73G	146C	153G	235G	263G	309.1C	315.1C	523del	524del	16111T	16223T	16319A	16362C	16519C						
South	S8	A2ac	64T	73G	146C	153G	235G	263G	315.1C	523del	524del	16092C	16111T	16147T	16213A	16223T	16290T	16319A	16362C				
South	S9	A2+(64)	64T	73G	146C	153G	235G	263G	315.1C	523del	524del	16111T	16223T	16263C	16290T	16311C	16319A	16362C					
South	S10	B4b	73G	103A	150T	263G	315.1C	499A	16183C	16189C	16217C	16266T											
South	S11	B4b	73G	152C	263G	309.1C	315.1C	499A	523del	524del	16051G	16183C	16189C	16217C	16360T	16519C							
South	S12	C1b+16311	73G	195C	249del	263G	290del	291del	309.1C	315.1C	489C	493G	508G	523del	524del	16093C	16223T	16298C	16311C	16325C	16327T	16519C	
South	S13	A2+(64)	64T	73G	146C	153G	235G	263G	309.1C	315.1C	523del	524del	16111T	16223T	16290T	16319A	16362C	16519C	16527T				
South	S14	B4b	73G	199C	263G	309.1C	309.2C	315.1C	499A	573.1C	16168T	16182C	16183C	16189C	16217C	16519C							
South	S15	B4	73G	115C	263G	315.1C	499A	523del	524del	16183C	16189C	16217C	16296T	16311C	16519C								
South	S16	B4b	73G	152C	263G	309.1C	315.1C	499A	523del	524del	16051G	16183C	16189C	16217C	16360T	16519C							
South	S17	A2+(64)	64T	73G	146C	153G	235G	263G	264T	315.1C	523del	524del	16111T	16189C	16223T	16240G	16290T	16319A	16362C	16519C			
South	S18	C1d	73G	249del	263G	290del	291del	309.1C	315.1C	489C	523del	524del	16051G	16223T	16298C	16319A	16325C	16327T					
South	S19	B2b2	73G	263G	309.1C	315.1C	499A	573.1C	16145A	16183C	16189C	16217C	16519C										
South	S20	B4b	73G	263G	309.1C	315.1C	499A	16183C	16189C	16217C	16266T	16485A	16519C										
South	S21	B4b	61T	62C	73G	146C	263G	309.1C	309.2C	315.1C	499A	16182C	16183C	16189C	16217C	16437C	16519C						
South	S22	C1b5b	73G	195Y	249del	263G	290del	291del	315.1C	489C	493G	523del	524del	16144C	16189C	16223T	16249C	16298C	16325C	16327T			
South	S23	M	73G	263G	309.1C	315.1C	489C	16209C	16325C	16362C													
South	S24	B4b	73G	146C	263G	315.1C	499A	16183C	16189C	16217C	16266T	16311C	16485A	16519C									
South	S25	B4b	73G	263G	309.1C	315.1C	499A	16182C	16183C	16189C	16217C	16437C	16519C										
South	S26	C1	73G	210G	249del	263G	290del	291del	309.1C	315.1C	327T	489C	493G	523del	524del	16086C	16183C	16189C	16223T	16298C	16323C	16325C	16327T
South	S27	L2a1	73G	143A	146C	152C	153G	195C	263G	315.1C	523del	524del	16129A	16189C	16192T	16223T	16278T	16294T	16309G	16390A			
South	S28	B4b	73G	103A	263G	309.1C	315.1C	499A	16093C	16182C	16183C	16189C	16217C	16266T	16295T								
South	S29	D4	60C	73G	194T	263G	309.1C	315.1C	489C	16223T	16325C	16362C	16519C										
South	S30	U5a1c1a	73G	151T	153G	195C	263G	315.1C	524.1A	524.2C	524.3A	524.4C	16192T	16256T	16270T	16311C	16320T	16399G					
South	S31	B4b	73G	199C	214R	263G	309.1C	309.2C	315.1C	499A	573.1C	16168T	16182C	16183C	16193.1C	16217C	16519C						
South	S32	B4b	73G	263G	309.1C	315.1C	499A	16183C	16189C	16193.1C	16217C	16437C	16519C										
South	S33	B4b	73G	263G	309.1C	315.1C	499A	573.1C	16183C	16189C	16217C	16235G	16519C										
South	S34	B4b	73G	103A	183G	263G	315.1C	499A	16182C	16183C	16189C	16217C	16234T	16266T	16319A	16519C							
South	S35	B2h	73G	263G	309.1C	309.2C	315.1C	499A	573.1C	573.2C	16183C	16189C	16217C	16468C	16519C								
South	S36	B4b	73G	183G	263G	315.1C	499A	16183C	16189C	16217C	16234T	16266T	16519C										
South	S37	B4b	73G	93G	131C	179C	195C	263G	309.1C	309.2C	315.1C	499A	16093C	16183C	16189C	16193.1C	16217C	16519C					
South	S38	D4	60C	73G	152Y	194T	263G	309.1C	309.2C	315.1C	489C	16223T	16325C	16362C	16519C								

Andean Region	Sample ID	Haplogroup	Haplotype Polymorphic sites (range: 0-576; 16024-16569)																			
South	S38	D4	60C	73G	152Y	194T	263G	309.1C	309.2C	315.1C	489C	16223T	16325C	16362C	16519C							
South	S39	B4b	73G	152C	263G	309.1C	315.1C	499A	523del	524del	16051G	16129A	16183C	16189C	16193.1C	16217C	16360T	16519C				
South	S40	B2o1	73G	263G	315.1C	499A	514T	16092C	16145A	16157C	16183C	16189C	16193.1C	16217C	16274A	16294T	16519C					
South	S41	B4b	73G	127C	263G	309.1C	315.1C	499A	573.1C	573.2C	573.3C	16182C	16183C	16189C	16217C	16324C	16519C					
South	S42	B4b	73G	103A	183G	263G	309.1C	309.2C	315.1C	499A	16182C	16183C	16189C	16217C	16234T	16266T	16319A	16519C				
South	S43	D4h3a	73G	152C	263G	309.1C	315.1C	489C	16223T	16241G	16301T	16342C										
South	S44	B4b	73G	263G	315.1C	319C	499A	16183C	16189C	16193.1C	16217C	16266T	16292A	16519C								
South	S45	B4b	73G	263G	309.1C	309.2C	315.1C	499A	573.1C	16183C	16189C	16193.1C	16217C	16235G	16519C							
South	S46	A2+(64)	64T	73G	146C	153G	235G	263G	309.1C	315.1C	523del	524del	16111T	16223T	16271C	16290T	16319A	16362C	16519C			
South	S47	A2ai	64T	73G	146C	153G	235G	263G	315.1C	523del	524del	16213A	16223T	16290T	16319A	16362C	16519C					
South	S48	C1d	73G	249del	263G	290del	291del	309.1C	315.1C	489C	523del	524del	16051G	16223T	16298C	16319A	16325C	16327T				
South	S49	B4b	73G	103A	183G	263G	309.1C	309.2C	315.1C	499A	16182C	16183C	16189C	16217C	16234T	16266T	16319A	16519C				
South	S50	L2a1g	73G	146C	152C	195C	263G	309.1C	315.1C	16131C	16189C	16222A	16223T	16225T	16234T	16278T	16294T	16309G	16390A			
South	S51	A2+(64)	64T	73G	146C	153G	235G	263G	264T	315.1C	523del	524del	16111T	16189C	16223T	16240G	16290T	16319A	16362C	16519C		
South	S52	B4b	73G	125G	152C	193G	263G	309.1C	315.1C	499A	16182C	16183C	16189C	16217C	16519C							
South	S53	B2	73G	152C	263G	309.1C	309.2C	315.1C	499A	16092C	16183C	16189C	16193.1C	16217C	16241G	16519C						
South	S54	B4	73G	263G	309.1C	315.1C	499A	573.1C	573.2C	16183C	16189C	16217C	16311C	16519C								
South	S55	L3f1b	73G	189G	200G	263G	315.1C	523del	524del	16209C	16223T	16311C	16519C									
South	S56	C1d	73G	152C	249del	263G	290del	291del	315.1C	339G	489C	523del	524del	16051G	16093C	16129A	16223T	16234T	16298C	16325C	16327T	
South	S57	D1f	73G	185A	263G	289G	315.1C	489C	16142T	16223T	16325C	16362C	16519C									
South	S58	D1f	73G	97A	105del	106del	107del	108del	109del	110del	194T	263G	309.1C	315.1C	489C	16142T	16153A	16223T	16227G	16325C	16362C	16519C
South	S59	C1b	73G	249del	263G	290del	291del	309.1C	315.1C	489C	493G	523del	524del	16223T	16298C	16325C	16327T					
South	S60	R	73G	195C	263G	310C	315del	499A	16182C	16183C	16189C	16193.1C	16217C	16234T	16356C	16362C	16390A	16519C				
South	S61	C1b	73G	249del	263G	290del	291del	309.1C	315.1C	489C	493G	523del	524del	16129A	16223T	16298C	16325C	16327T				
South	S62	B4b	73G	152C	263G	309.1C	315.1C	499A	523del	524del	16051G	16183C	16189C	16193.1C	16217C	16360T	16519C					
South	S63	D4	60C	73G	152Y	194T	263G	309.1C	309.2C	315.1C	489C	16223T	16325C	16362C	16519C							
South	S64	B4b	73G	153G	263G	309.1C	309.2C	315.1C	499A	16183C	16189C	16193.1C	16217C	16260T	16437C	16519C						
South	S65	B4b	73G	103A	183G	263G	309.1C	309.2C	315.1C	499A	16182C	16183C	16189C	16217C	16234T	16266T	16319A	16519C				
South	S66	A2+(64)	64T	73G	146C	153G	235G	263G	315.1C	523del	524del	568Y	16111T	16223T	16290T	16319A	16362C	16519C				
South	S67	M	73G	239C	263G	315.1C	489C	513A	524.1A	524.2C	16180G	16223T	16298C	16325C	16327T	16519C						
South	S68	B4b	73G	103A	195C	263G	309.1C	315.1C	499A	16183C	16189C	16193.1C	16217C	16266T								
South	S69	C1b	73G	146C	249del	263G	290del	291del	309.1C	315.1C	489C	493G	523del	524del	16223T	16234T	16298C	16325C	16327T	16519C		
South	S70	B2	73G	152C	199C	263G	309.1C	315.1C	499A	16092C	16182C	16183C	16189C	16217C	16241G	16289G	16519C					
South	S71	D1f+16189	73G	263G	309.1C	315.1C	489C	16142T	16183C	16189C	16193.1C	16223T	16325C	16362C	16519C							
South	S72	B4b	73G	152C	263G	309.1C	315.1C	499A	523del	524del	16051G	16183C	16189C	16217C	16360T	16519C						

Appendix 4. Quasi-median network of 214 haplotypes using the filter EMPOPall_R12 .

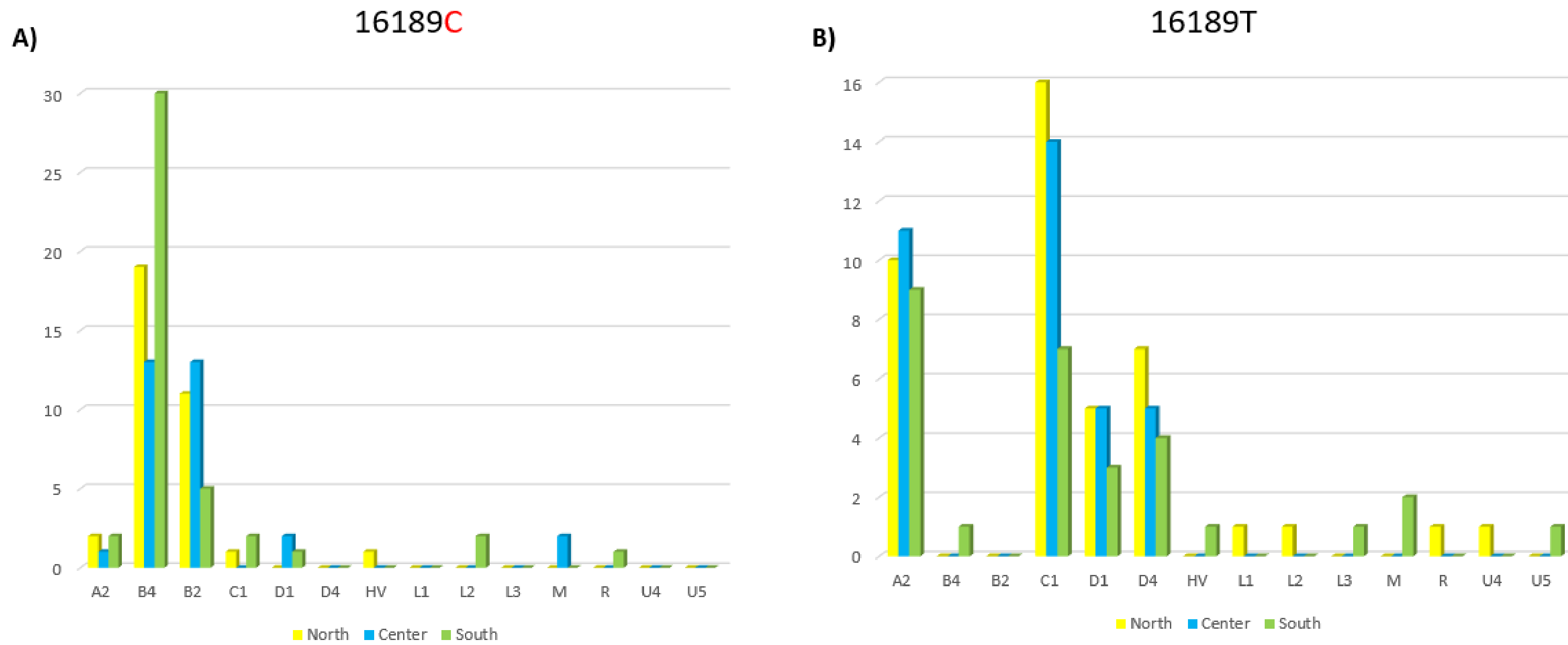


Appendix 5. HVSI and HVSII point heteroplasmies found in 7 haplotypes from Andean Mestizo population.

Point heteroplasmy	Andes	Samples	Haplogroup
16399R	North	N72	D4h3a6
568Y	South	S66	A2+(64)
185R	Center	C55	D1f
152Y	South	S38	D4
	South	S63	D4
214R	South	S31	B4b
195Y	South	S22	C1b5b

Note Y: mixture of C and T; R: mixture of A and G

Appendix 6. Figure of Point Heteroplasmy in HVS-I in Andean sequences. A) Frequency distribution of the 16189C mutated variant by haplogroup. B) Frequency distribution of 16189T wild-type variant by haplogroup.



Appendix 7. Pairwise F_{ST} genetic distances (below diagonal) and *P*-values (above diagonal) using haplotype data. Statistically non-significant values are highlighted in red ($P \geq 0.05$). After Bonferroni correction ($P \geq 0.00029$), statistically non-significant values are highlighted in blue.

	Andes North	Andes Center	Andes South	ECU Coast	VEN	COL	PER	PRY	ARG North	ARG Center	ARG South	BR PAR	BR RJ	BR SP	BOL	ECU BA	ECU BRAN	Native Amazon	Native Andes
Andes North	*	0,67944	0,10672	0,03267	0,00000	0,00000	0,05376	0,03891	0,00515	0,00000	0,00158	0,00198	0,00000	0,00000	0,00000	0,03663	0,29710	0,00000	0,03099
Andes Center	0,00000	*	0,10761	0,07970	0,00000	0,00000	0,05277	0,05970	0,00347	0,00000	0,00089	0,00010	0,00000	0,00000	0,00000	0,10286	0,41372	0,00000	0,06415
Andes South	0,00779	0,00861	*	0,00030	0,00000	0,00000	0,03920	0,00010	0,00000	0,00000	0,00010	0,00000	0,00000	0,00000	0,00020	0,01010	0,17790	0,00000	0,41352
ECU Coast	0,01355	0,01011	0,03705	*	0,02148	0,00178	0,00248	0,06465	0,08009	0,00000	0,01148	0,00069	0,00000	0,00000	0,00000	0,31155	0,02426	0,00000	0,00178
VEN	0,05370	0,05588	0,08145	0,01371	*	0,00188	0,00000	0,00000	0,00030	0,00000	0,00040	0,00000	0,00000	0,00000	0,00000	0,00683	0,00000	0,00000	0,00000
COL	0,05549	0,05719	0,05787	0,02944	0,02459	*	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,02386	0,00010	0,00000	0,00158
PER	0,01057	0,01174	0,01251	0,02734	0,07081	0,05570	*	0,00079	0,00010	0,00000	0,00010	0,00010	0,00000	0,00000	0,00000	0,01049	0,02742	0,00000	0,03178
PRY	0,00801	0,00759	0,03301	0,00646	0,03989	0,05355	0,02140	*	0,02970	0,00000	0,00436	0,00000	0,00000	0,00000	0,00000	0,00931	0,00119	0,00000	0,00079
ARG North	0,02055	0,02462	0,05311	0,00741	0,02845	0,05241	0,03801	0,00717	*	0,01277	0,03000	0,03544	0,00000	0,00000	0,00000	0,00574	0,00010	0,00000	0,00000
ARG Center	0,03365	0,04240	0,05906	0,02892	0,05087	0,06328	0,04016	0,02701	0,00877	*	0,04118	0,05950	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
ARG South	0,03236	0,03954	0,06887	0,02150	0,03789	0,07019	0,04511	0,02320	0,01424	0,00904	*	0,00693	0,00000	0,00000	0,00000	0,00168	0,00000	0,00000	0,00000
BR PAR	0,01924	0,03007	0,03764	0,02306	0,04103	0,05212	0,02939	0,02132	0,00775	0,00395	0,01657	*	0,00000	0,00000	0,00000	0,00010	0,00000	0,00000	0,00000
BR RJ	0,05581	0,07303	0,07115	0,05360	0,05336	0,07068	0,06648	0,06389	0,04466	0,03373	0,04364	0,01774	*	0,49233	0,00000	0,00000	0,00000	0,00000	0,00000
BR SP	0,05971	0,07921	0,07740	0,06034	0,05887	0,08251	0,07356	0,06869	0,04885	0,03688	0,04543	0,01927	0,00000	*	0,00000	0,00000	0,00000	0,00000	0,00000
BOL	0,07111	0,07844	0,03473	0,12434	0,17358	0,12507	0,06407	0,10691	0,14181	0,13708	0,16835	0,10729	0,13354	0,14627	*	0,00000	0,00000	0,00000	0,04613
ECU BA	0,02050	0,01281	0,03314	0,00209	0,02984	0,02466	0,03010	0,02199	0,02955	0,05026	0,04557	0,04236	0,07895	0,08994	0,11354	*	0,12058	0,00158	0,02237
ECU BRAN	0,00132	0,00000	0,00426	0,01399	0,05209	0,03889	0,01307	0,01510	0,03169	0,04866	0,04846	0,03389	0,07619	0,08364	0,06535	0,00991	*	0,00000	0,07415
Native Amazon	0,16282	0,14919	0,20568	0,09069	0,08154	0,12571	0,18268	0,11808	0,11765	0,14841	0,13889	0,15639	0,17630	0,19594	0,31305	0,07306	0,14603	*	0,00000
Native Andes	0,02973	0,02442	0,00000	0,06302	0,10810	0,07405	0,02709	0,05908	0,08576	0,09375	0,11480	0,06430	0,09332	0,10353	0,02095	0,04859	0,02064	0,24967	*

Ecuadorian Andes populations (North, Center, and South) are marked in green ; South American admixed populations [Ecuadorian Coast (data not published); COL: Colombia (Auton et al., 2015); PER: Peru (Auton et al., 2015); BOL: Bolivia (Taboada-Echalar et al., 2013); VEN: Venezuela (Castro De Guerra et al., 2012); PRY: Eastern region of Paraguay (Simão et al., 2021); ARG North, Center and South: Argentina (Bobillo et al., 2010); BR PAR: Paraná, Brazil (Poletto et al., 2019); BR RJ: Rio de Janeiro , Brazil (Simão et al., 2018); BR SP: Sao Paulo, Brazil (Prieto et al., 2011); ECU BRAN: Ecuadorian Brandini (Brandini et al., 2018); ECU BAE: Ecuadorian Baeta (Baeta et al., 2012)] are marked in orange; and Native populations [ECU Native Andes: Ecuadorian Kichwa Natives from the Andes (Sandoval et al., 2020), ECU Native Amazon: Ecuadorian Kichwa Natives from the Amazon (Baeta et al., 2012)] are marked in blue.