

RESEARCH ARTICLE

Paternal portrait of populations of the middle Magdalena River region (Tolima and Huila, Colombia): New insights on the peopling of Central America and northernmost South America

Luz Angela Alonso Morales ^{1*}, Andrea Casas-Vargas ¹, Madelyn Rojas Castro¹, Rafael Resque ², Ândrea Kelly Ribeiro-dos-Santos³, Sidney Santos³, Leonor Gusmão⁴, William Usaquén^{1*}

1 Populations Genetics and Identification Group, Institute of Genetics, Universidad Nacional de Colombia, Bogotá, Colombia, **2** Laboratório de Toxicologia e Química Farmacêutica, Departamento de Ciências da Saúde e Biológicas, Universidade Federal do Amapá, Macapá, Brazil, **3** Human and Medical Genetics Laboratory, Institute of Biological Sciences, Federal University of Pará (Universidade Federal do Pará - UFPA), Belém, state of Pará (PA), Brazil, **4** DNA Diagnostic Laboratory (LDD), Institute of Biology, State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil

* laalonsom@unal.edu.co (LAAM); wusaquenm@unal.edu.co (WU)



 OPEN ACCESS

Citation: Alonso Morales LA, Casas-Vargas A, Rojas Castro M, Resque R, Ribeiro-dos-Santos ÂK, Santos S, et al. (2018) Paternal portrait of populations of the middle Magdalena River region (Tolima and Huila, Colombia): New insights on the peopling of Central America and northernmost South America. PLoS ONE 13(11): e0207130. <https://doi.org/10.1371/journal.pone.0207130>

Editor: Chuan-Chao Wang, Xiamen University, UNITED STATES

Received: May 9, 2018

Accepted: October 25, 2018

Published: November 15, 2018

Copyright: © 2018 Alonso Morales et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files. The Y-STR chromosomal data for all 83 donors are available in the open-access Y-STR Haplotype Reference Database (YHRD, www.yhrd.org); accession number YA004510.

Funding: This project received two grants from the Dirección de Investigación Bogotá of the Universidad Nacional de Colombia URL <http://>

Abstract

The valley of the Magdalena River is one of the main population pathways in Colombia. The gene pool and spatial configuration of human groups in this territory have been outlined throughout three historical stages: the Native pre-Hispanic world, Spanish colonization, and XIX century migrations. This research was designed with the goal of characterizing the diversity and distribution pattern of Y-chromosome lineages that are currently present in the Tolima and Huila departments (middle Magdalena River region). Historic cartography was used to identify the main geographic sites where the paternal lineages belonging to this area have gathered. Twelve municipalities were chosen, and a survey that included genealogical information was administered. Samples collected from 83 male volunteers were analyzed for 48 Y-SNPs and 17 Y-STRs. The results showed a highly diverse region characterized by the presence of 16 sublineages within the major clades R, Q, J, G, T and E and revealed that 93% (n = 77) of haplotypes were different. Among these haplogroups, European-specific R1b-M269 lineages were the most representative (57.83%), with six different subhaplogroups and 43 unique haplotypes. Native American paternal ancestry was also detected based on the presence of the Q1a2-M3*(xM19, M194, M199) and Q1a2-M346*(xM3) lineages. Interestingly, all Q1a2-M346*(xM3) samples (n = 7, with five different haplotypes) carried allele six at the DYS391 locus. This allele has a worldwide frequency of 0.169% and was recently associated with a new Native subhaplogroup. An in-depth phylogenetic analysis of these samples suggests the Tolima and Huila region to be the principal area in all Central and South America where this particular Native lineage is found. This lineage has been present in the region for at least 1,809 (+/- 0,5345) years.

www.dib.unal.edu.co/. The first one was the grant number 13425, which financed the developing of the research project, and its grantee was the Populations Genetics and Identification Laboratory, Institute of Genetics, Universidad Nacional de Colombia, Bogotá, Colombia. The second grant, with the number 15118, was assigned for supporting the master thesis project of Luz Angela Alonso, student of the Faculty of Medicine. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Similar to most South American populations, the genetic and cultural background of Colombians has been shaped by complex population dynamics: first, the peopling of Colombia by Native American inhabitants; second, the Spanish conquest and slave trade; and third, the different waves of migration and movements that characterized the XX century and the first decade of the XXI century [1,2]. Over time, Colombia has received genetic contributions from different Native American groups, European settlers and African slaves from different areas, resulting in high genetic diversity and a nonuniform composition of the current gene pool [3,4,5].

Several genetic studies have been performed in Colombia using different types of molecular markers (matrilineal, patrilineal, and autosomal) to analyze numerous populations in different regions of the country [3,5,6,7,8,9,10,11,12,13,14,15,16].

Such investigations have demonstrated that populations from urban areas usually present higher levels of non-Native admixture and uniparental-lineage diversity. In addition, these studies have shown how the proportions of Native American, African and European lineages vary throughout the country. The highest frequencies of African genetic contribution have been found on the Caribbean coasts and islands and in the Pacific Coast region. In contrast, Native lineages are found in higher proportions in Native populations, mainly from the eastern part of the country, including the Amazon and rural areas of southwest and northern Colombia. The European contribution is present throughout the country. This diversity reflects a nonuniform genetic pool in Colombian communities and regions.

In the history of Colombian peopling, three routes have been suggested: one following the Pacific coast, another traveling into the Colombian mountains (through the valleys of the Cauca and Magdalena Rivers) and the last, along the Atlantic coast [17].

In this context, the valley of the Magdalena River is the main hydrographic system in Colombia and the major economic development backbone in the country [18]. Its basin covers a surface of 257.438 square kilometers (km²) and is densely populated. Nearly 79% of the entire population of Colombia lives in the Magdalena watershed, amounting to a demographic density of 120 inhabitants km² [19]. The valley spans the final spurs of the Andes Mountain range at high altitudes to its delta in the Atlantic Ocean, in the north of the country, in low-lying areas (Fig 1).

This scenario of this region encompasses three key historical moments in the configuration of the Colombian territory: (i) the pre-Columbian Period, in which Native American groups used this region as a route of communication and source of food; (ii) the Colonial Period, in which this area served as a penetration route towards the center of the country; and (iii), the end of the XIX century, at which time the Magdalena River became the main fluvial route in the country and a transit path for the arrival of new settlers [1].

The area covered by the Tolima and Huila territory is part of the Magdalena River's mid-valley. This region was inhabited by hunting, fishing and gathering groups 16,000 years ago [23]. The inventory of the existing archaeological sites in these departments represents the passage of time, beginning with the first hunter-gatherers who occupied this region, extending to the current farming and urban communities [24].

In pre-Columbian times, the region of Tolima and Huila was characterized by high diversity of Native groups, which were greatly reduced in size during the European colonization of America. The mineral resources of the region were the original reason for this colonization, but throughout the XVIII century, the exploitation of quinine brought a new migratory wave of settlers from Spain and from other regions of Colombia, followed by agricultural modernization [2].



Fig 1. Geographic location of the Colombia and the Magdalena River in South America and the 12 sampling locations (in bold) in the departments of Tolima and Huila. Maps created with the QGIS open source software [20], map data is available from <https://geoportal.dane.gov.co/v2/?page=elementoDescargaMGN> [21] and <https://www.arcgis.com/home/item.html?id=d3d2bae5413845b193d038e4912d3da9> [22].

<https://doi.org/10.1371/journal.pone.0207130.g001>

According to the Regional Indigenous Corporation of Tolima (CRIT), from pre-Columbian times to the present day, two main Native communities have inhabited the Tolima and Huila territories. The “Pijao” ethnic group (known in Colombia for their resistance to Spanish domination) remains on their ancestral land, especially in the municipalities of Ortega, Coyaima,

Natagaima, Chaparral, and Saldaña, with approximately 17,000 members. The “Nasa” group mainly resides in the Cauca and Huila departments, with some communities in Tolima. Currently, there are approximately 138,501 Nasa people in both departments [25].

The majority of genetic studies performed in Tolima and Huila have focused on evaluating diversity levels in the “Pijao” and “Nasa” Native communities using molecular markers such as autosomal STRs, restriction fragment length polymorphisms (RFLPs) in mitochondrial DNA (mtDNA) [26,27], Y-chromosome-specific SNPs [27,28,29,30], and autosomal SNPs [27,31]. In Huila, a single study was carried out in a sample of seven non-Native individuals from the municipality of Neiva, in which autosomal SNPs, X-chromosomal STRs, mtDNA and Y-chromosomal markers were investigated [3].

Since the last century, the Tolima and Huila territories have experienced waves of violence and forced displacement (incited by armed groups). This phenomenon has had a great impact on the demography of Native and rural communities in this area. Thus far, there has been no exhaustive evaluation of the diversity of the Y-chromosome lineages that are present in this territory (not only in individuals from Native communities) using a representative sample from both departments.

To characterize the patrilineages of the current inhabitants of the Tolima and Huila regions, this study focused on phylogenetic analysis of the Y-chromosome, using a synergic combination of markers defining haplogroups, such as single-nucleotide polymorphisms (SNPs) and microsatellites (short tandem repeats [STRs]), to evaluate migration, admixture, and ancestry patterns and to reveal additional microevolutionary processes associated with population structure.

Materials and methods

Study population

This study was conducted on samples collected by the Population Genetics and Identification Group of the Universidad Nacional de Colombia in the Tolima and Huila region in 2012.

Prior to the field phase, a historic cartography was performed on both departments to identify the main geographic sites where the paternal lineages that shaped the region have gathered over time, by the analysis of demographic, economic and political variables of the territory from the colony time to the present (all details are in [S1 File](#)). From this a priori investigation, 12 locations from both departments were selected as a reference sampling places (Honda, Líbano, Ibagué, Espinal, Coyaima, and Natagaima in Tolima; and Villavieja, Neiva, Palermo, Garzón, La Plata, and San Agustín in Huila).

A purposive sample was performed among these 12 locations, with a total sample size of 330 people. From it, 119 samples were from men. However, as a consequence of long periods of storage before this study was performed; 36 of the 119 samples showed poor DNA quality. For this reason, to ensure good data quality, the final sample size of this study is composed by 83 male samples from the most relevant historical settlement places were selected to analyze the paternal portrait of the Tolima and Huila populations.

Unfortunately, due to both departments have suffered from serious internal violence as a consequence of the presence of illegal defense forces, the population access is limited, for this reason, we could not increase the male sampling number.

Samples were collected at the selected location's hospitals. Two inclusion criteria were used to select participants; the participants had to be at least 18 years old or older and had to have been a local area resident for at least one year. All samples were taken under written informed consent.

DNA extraction was performed from blood samples using the salting out method, following the DNA 2000 kit (CorpoGen) protocol.

This research was approved by the ethics committee of the Faculty of Science of the Universidad Nacional de Colombia.

Genotyping of Y-chromosome STRs

Seventeen Y-chromosomal STR loci were amplified using the AmpFlSTR Yfiler PCR Amplification kit, following the manufacturer's instructions (User's Manual, Applied Biosystems). Separation and detection of the amplified products were performed in an Applied Biosystems PRISM 310 genetic analyzer. Allele assignments were performed by comparison with a reference ladder included in the kit, and using the nomenclature recommended by the International Society for Forensic Genetics (ISFG) [32]. These analyses were performed in the laboratory of the Population Genetics and Identification group, of the Genetics Institute of the Universidad Nacional de Colombia.

The Y-chromosomal data for all 83 donors have been submitted to the open-access Y-STR Haplotype Reference Database (YHRD, www.yhrd.org): accession number YA004510.

Genotyping of Y-chromosome SNPs

To select the best set of SNP markers to be typed, Y-STR haplotype information was used to predict the most likely haplogroup of each sample, using an online tool (<http://www.hprg.com/hapest5/>). This software is based on a Bayesian approximation using allelic frequencies calculated from haplotype collections taken from published papers and databases [33]. This makes it possible to determine the probability of a Y-STR haplotype being found within a given haplogroup. The assigned haplogroup is the one to which a given haplotype has the highest probability of belonging.

The predicted haplogroups were then confirmed by genotyping 48 Y-chromosome SNPs, using the SNaPshot Kit (Applied Biosystems), in six previously described multiplex reactions including the markers: (i) SRY1532, M213, M9, M70, M22, Tat, 92R7, M173, and P25 [34]; (ii) M201, M170, 12f2, M26, M62, and M172 [34]; (iii) M269, L23, U106, S116, U152, M529, M153, M167(SRY2627) [35]; (iv) M12, M67, M92, M241, M410 [36], M267 [37]; (v) M3, M19, M194, M199, M242, M346 and P36.2 [12]; (vi) M96, M33, P2, M2, M154, M191, M35, M78, M81, M123, V6, M293, M85 [38].

For all multiplex reactions, PCRs were performed in a 10 μ l volume containing 5 μ l of the Qiagen multiplex PCR kit, 1 μ l of primer mix (each primer at 2 μ M), 3.5 μ l of water, and 0.5 μ l of DNA (5–20 ng/ μ l). The PCR thermocycling conditions consisted of an initial denaturation at 95°C for 15 minutes, followed by 35 cycles at 94°C for 30 seconds, 60°C for 90 seconds and 72°C for 60 seconds, and a final extension at 72°C for 10 minutes. After confirming the amplification in a polyacrylamide gel (T9%, C5), 1 μ l of the PCR product was purified using 0.5 μ l of Exo-SAP-it (USB), and incubated at 37°C for 15 minutes, followed by inactivation of the enzyme at 85°C for 15 minutes.

The SNaPshot multiplex reactions were performed in 5 μ l containing 2 μ l of SNaPshot multiplex mix (Applied Biosystems), 1.5 μ l of the purified PCR product and 1.5 μ l of a mix of the single base extension (SBE) primers. The reaction protocol consisted on 25 cycles at 95°C for ten seconds, 50°C for five seconds and 60°C for 30 seconds. The final products were purified with 1 μ l of SAP (USB) at 37°C for one hour, followed by a denaturation step at 85°C for 15 minutes. The SBE products were run in an Applied Biosystems 3130 Genetic Analyzer and analyzed using the GeneMapper software v4.0, at the Human Genetics and Medicine Laboratory, Biological Sciences Institute at the UFPA (Federal -University of Pará), Brazil.

The haplogroup nomenclature was assigned according to van Oven et al. [39]

Statistical analysis

An analysis of molecular variance (AMOVA) was performed using Arlequin software 3.5.1.2 [40] to evaluate the random distribution hypothesis for individuals among populations. Haplotype diversities were calculated according to Nei [41], using the following equation:

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

Where *n* is the sample size and *p_i* is the frequency of the *i*-haplotype.

Phylogenetic relationships between Y-chromosome haplotypes among American and Asian populations were inferred for the Native American Q lineage (S1 Table) and visualized with the median-joining algorithm [42] using the Network software 4.6.1.1 (www.fluxustechonology.com), which assumes a stepwise mutation model for STRs. In this analysis, DYS385 was excluded because it corresponds to two loci that cannot be differentiated via the applied typing method [43], while for the DYS389II-I locus, the number of repeats in DYS389I was subtracted [44]. In the network, a weight of three to 10 was assigned to each locus, following the methodology proposed by Muzzio et al. [45] and using the mutation rates reported in the Y-haplotype reference database [46].

Estimates of the time to the most recent common ancestor (TMRCA) were determined using rho statistics, implemented in the NETWORK program, employing a mean effective mutation rate for Y-STRs of 6.9×10^{-4} /locus/25 years [47]. The ancestral haplotype was inferred using the modal allele at each STR locus [48].

To compare the Y-chromosome haplogroup frequencies identified in this study with those from 30 other Colombian samples, a principal component analysis (PCA) was performed using Multiple Statistical Package (MVSP) software, version 3.1 [49].

Table 1. Y-chromosome haplogroup frequencies in the Tolima and Huila departments. The number of chromosomes per haplogroup is shown in parentheses (n).

Haplogroup-SNP	Huila	Tolima	Total
E1b1b-M78	4.76% (2)	2.43% (1)	3.61% (3)
E1b1b-M81	2.38% (1)	7.31% (3)	4.81% (4)
E1b1b-M123	2.38% (1)	0% (0)	1.20% (1)
G-M201	7.14% (3)	0% (0)	3.61% (3)
J1—M267	2.38% (1)	2.43% (1)	2.40% (2)
J2a-M410* (xM67, M92)	0% (0)	4.87% (2)	2.40% (2)
J2a-M67* (xM92)	2.38% (1)	0% (0)	1.20% (1)
Q1a2-M346* (xM3)	7.14% (3)	9.75% (4)	8.43% (7)
Q1a2-M3* (xM19, M194, M199)	9.52% (4)	7.31% (3)	8.43% (7)
R1b-U106	2.38% (1)	7.31% (3)	4.81% (4)
R1b- S116* (xU152, M529, M65, M153, M167)	28.57% (12)	34.14% (14)	31.32% (26)
R1b-U152	4.76% (2)	7.31% (3)	6.02% (5)
R1b-M529	14.28% (6)	7.31% (3)	10.84% (9)
R1b-M153	2.38% (1)	0% (0)	1.20% (1)
R1b-M167	2.38% (1)	4.87% (2)	3.61% (3)
T-M70	7.14% (3)	4.87% (2)	6.02% (5)
Total	100% (42)	100% (41)	100% (83)

<https://doi.org/10.1371/journal.pone.0207130.t001>

Table 2. Diversity statistics per haplogroup in the Tolima and Huila region.

Haplogroup-SNP	Sample size (n)	N° Different Haplotypes	N° polymorphic loci	Average N° alleles per locus	Haplotype diversity ^a
E1b1b-M78	3	2	7	1.412±0.507	NA
E1b1b-M81	4	3	3	1.176±0.393	NA
E1b1b-M123	1	1	0	1.000±0.000	NA
G-M201	3	3	10	1.824±0.809	NA
J1—M267	2	2	10	1.588±0.507	NA
J2a-M410* (xM67, M92)	2	2	8	1.471±0.514	NA
J2a1-M67* (xM92)	1	1	0	1.000±0.000	NA
Q1a2-M346* (xM3)	7	5	6	1.412±0.618	0,904
Q1a2-M3* xM19, M194, M199)	7	7	16	2.941±0.966	1,000
R1b-U106	4	3	6	1.412±0.618	NA
R1b- S116* (xU152,M529, M65, M153, M167)	26	25	16	3.176±1380	0,997
R1b-U152	5	5	10	1.706±0.686	1,000
R1b-M529	9	9	12	1.941±0.827	1,000
R1b-M153	1	1	0	1.000±0.000	NA
R1b-M167	3	3	3	1.235±0.562	NA
T-M70	5	5	13	2.353±0.996	1,000
Total	83	77			

a. Haplogroup diversities were calculated for clades with five or more samples using Nei's haplotype diversity [41].

NA: Not Applicable.

<https://doi.org/10.1371/journal.pone.0207130.t002>

Results and discussion

Y-chromosome diversity in the Tolima and Huila departments

The Y-chromosome haplotypes and haplogroups observed in Tolima and Huila are listed in [S2 Table](#). The absolute and relative haplogroup frequencies per department and for the whole region are reported in [Table 1](#). The genotyping of 48 Y-SNP markers revealed six different major clades: R, Q, J, E, G, and T, representing 16 subhaplogroups ([Table 1](#)). More than 70% of the chromosomes in the total sample came from two clades: the European haplogroup R1b was found at the highest frequency (57.83%; n = 48), and the percentage of the Native American haplogroup Q1a2 was 16.86% (n = 14). No significant differences were found when the frequency distribution of haplotypes was compared between Tolima and Huila using AMOVA ($F_{ST} = -0.00133$; $p = 0.50255 \pm 0.0016$); these two groups were therefore analyzed as one population.

The haplogroup diversity of the Tolima and Huila region was 0.8692. Haplotype diversity was evaluated using the information obtained from 17 STR loci, revealing a high diversity rate of 0.9680, characterized by 93% (n = 77) unique haplotypes. Shared haplotypes were observed for haplogroups E1b1b-M78, E1b1b-M81, Q1a2-M346* (xM3), R1b-U106, and R1b-S116* (xU152, M529, M65, M153, M167) ([Table 2](#)). The haplotype diversity per haplogroup was estimated for clades with at least five samples and was found to range from 0.904 for Q1a2-M346* (xM3) to 1.000 for R1b-U152, R1b-M529, T-M70, and Q1a2-M3* (xM19, M194, M199) lineages. This diversity analysis showed high genetic heterogeneity within haplogroups, with no signs of important founder events in the Tolima and Huila region.

Native American background of the Tolima and Huila populations

Most studies conducted in admixed Colombian populations have identified a male European contribution ranging from 58% to 94% and representation of Native lineages ranging from 1% to 38% [7,11,27,50].

The present study revealed a 16.86% frequency of Y-chromosomes belonging to haplogroup Q (Table 1). Individuals carrying the M3-derived allele, which characterizes the Native American founder haplogroup within Q1a2 [51], were found in seven samples, at a frequency of 8.43%, and all of these individuals exhibit unique haplotypes (Table 2).

Haplogroup Q1a2-M346* (xM3) was identified in the same proportion as haplogroup Q1a2-M3. This haplogroup has been found in samples from Argentina, Chile, Bolivia, some studies from North America, the northern part of South America, and Siberia [52,53].

Using a phylogenetic approach, the possibility of past minor founder effects was evaluated in the Native American lineages. In the Q1a2-M3* (xM19, M194, M199) network (Fig 2), 310 haplotypes belonging to six Colombian and eight Latin American populations were included (S1 and S3 Tables). Fig 2 shows that the seven sampled haplotypes are dispersed within the network, with no coincidence between them or with the other haplotypes analyzed. This finding reflects the high diversity among M3 Native American chromosomes at the haplotype level as well as the need to improve the analysis of new SNP mutations within this haplogroup to extensively characterize lineages and identify possible settlement routes of M3 carriers.

As illustrated in Table 2, the Q1a2-M346* (xM3) clade exhibited a diversity value of 0.904, with five different haplotypes, which are distinct from each other due to six polymorphic loci. A detailed analysis showed that all seven individuals carried allele 6 at the DYS391 locus. This allele is present in a very low frequency worldwide and can be used to extensively characterize the Q1a2-M346* (xM3) lineages. Among the 197,102 minimal haplotypes in the YHRD database (2018-03-06), only 334 show this allele (0.169%), which is mainly present in Asian and Latin American samples. In this study, the presence of this allele was confirmed via direct sequencing, as shown in S1 Fig.

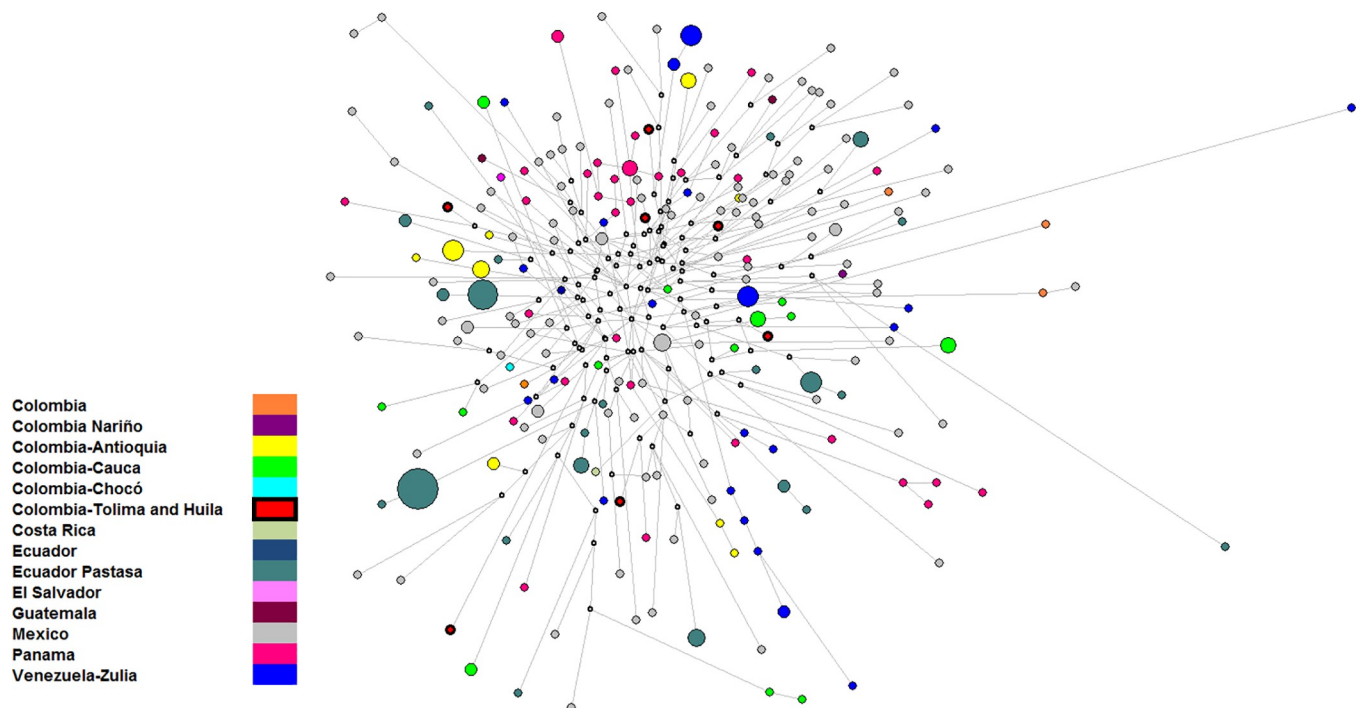


Fig 2. Median-joining network of haplotype Q1a2-M3 in 14 samples from Latin American population based on 15 STR markers. In the figure, the circles represent haplotypes, with areas proportional to their frequencies, and the colors indicate the original population. The median vectors (absent or extinct haplotypes) are shown in white.

<https://doi.org/10.1371/journal.pone.0207130.g002>

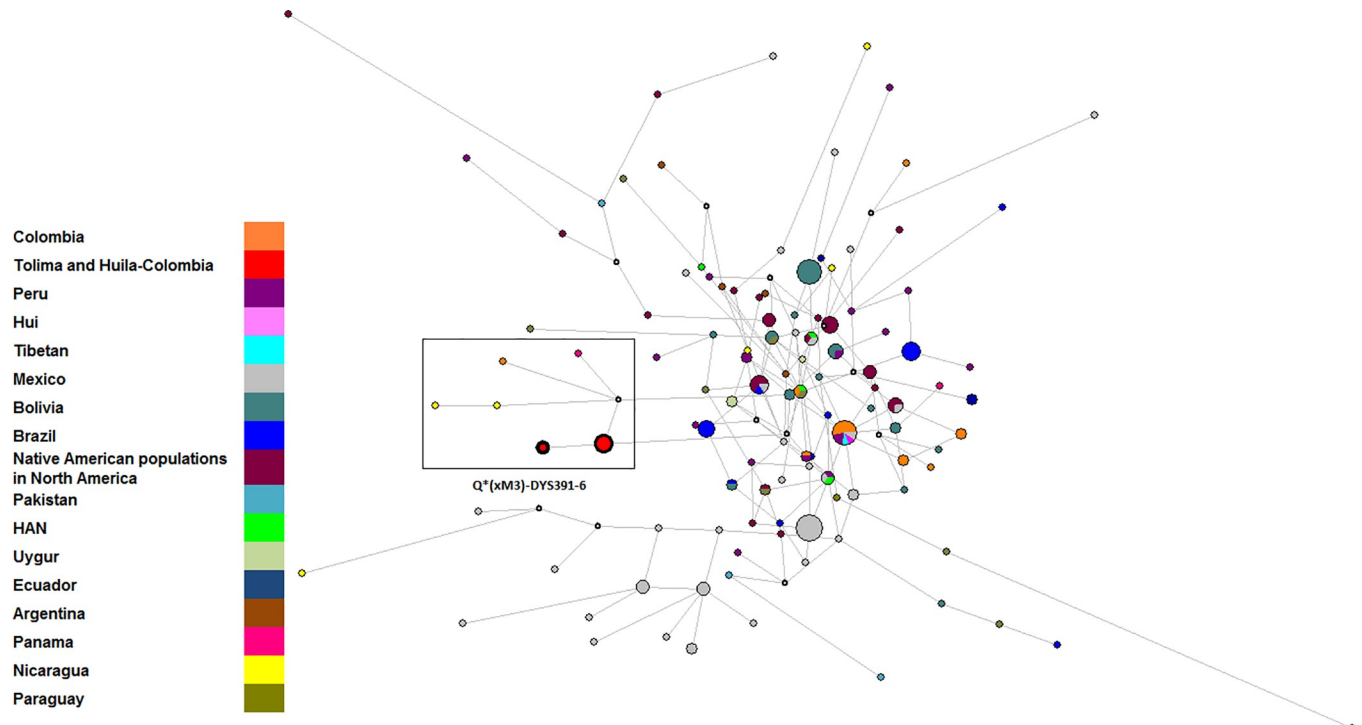


Fig 3. Median-joining network of haplogroup Q-M242*(xM3) using seven STRs (DYS19, *DYS389 I and II-I*, *DYS390*, *DYS391*, *DYS392*, and *DYS393*) to compare 17 Asian and American populations. The circles represent haplotypes, with areas proportional to their frequencies, while the colors indicate the population. The median vectors (absent or extinct haplotypes) are shown in white.

<https://doi.org/10.1371/journal.pone.0207130.g003>

To identify patterns and matches between the unconventional Q1a2-M346*(xM3) *DYS391*6* samples from the Tolima-Huila territory, a phylogenetic analysis was performed, using 193 previously reported Q-M242*(xM3) haplotypes from Asian and American populations (Fig 3 and S1 and S4 Tables). As shown in Fig 3, the seven Q1a2-M346*(xM3) haplotypes clustered with nine previously reported haplotypes (two from Nicaragua, one from Colombia, five from the Ngöbé-Native group of Panama, and one from non-Native individuals in Panama), all of which exhibit the allele *DYS391*6* (S4 Table). This cluster forms a separate branch in the network, suggesting a closer evolutionary relationship among these Central and northernmost South American samples.

New insights into the Peopling of Central America and the Northernmost Region of South America

The above analysis revealed high diversity in the Tolima and Huila Native American lineages. The Y-chromosomes from this study belonging to Q1a2-M3*(xM19, M194, M199) and Q1a2-M346*(xM3) did not match samples from East Asia or from the American continent, including Northern, Central, and Southern populations.

The results obtained from Q1a2-M346*(xM3)-*DYS391*6* Y-chromosomes fit a model of isolation after colonization, which most likely erased the footprints of this lineage dispersion route from modern Native populations. This finding could indicate already extinct haplotypes or very low-frequency haplotypes in the current Native communities that have not been reported previously.

Nevertheless, *DYS391*6* is not a private allele of a subhaplogroup within Q-M242*-(xM3); it has also been reported in individuals carrying the C-M130, O2-M95, Q-M242, Q-M3, and

Q-M120 haplogroups, who mainly belong to the Han ethnic group, from Southeast Asian populations [54]. However, it is important to highlight that all DYS391*6 samples from Latin America with available Y-SNP data were classified within the Q-M242 clade [28,30,55,56, 57,58, and this study].

For this reason, a broad search was conducted using the YHRD database (<http://yhrd.org/>) [46], and 334 haplotypes with this allele were found, including 300 from Asia (mostly in China), three from Europe, four from Australia, and one from the United States of America, while the remaining 24 were from Latin America. Among these 24 haplotypes, only six came from a Native population (Ngöbe in Panama), while two came from Nicaragua, one from Maracay-Venezuela, and the remaining 15 from admixed Colombian populations (Santander = 3, Bogota = 3, Valle del Cauca = 3, San Andres Island = 2, Antioquia = 1, Cundinamarca = 1, Nariño = 1, and Risaralda = 1).

Furthermore, related studies have obtained data from Central and South American samples. Battaglia et al. [57] reported three additional haplotypes, one from Colombia and two from Panama; Grugni et al. [58] added four samples from Panama; Jota et al. [30] detected the same allele in eight Coyaima individuals belonging to the “Pijao ethnic group” from the Tolima department; and Franco-Candela and Barreto [28] found an additional five haplotypes with this allele in another sample from the Coyaima Native group.

Nevertheless, many of the reported DYS391*6 samples were not assigned to a haplogroup; however, their haplotype information is valuable for assessing similarities and differences among these chromosomes. To evaluate the similarity among the DYS391*6 haplotypes found in the Asian and Latin American data, a phylogenetic network was created (S5 Table and S2 Fig). Based on this analysis, a different origin was identified for Asian and Latin American samples as well as a close connection between Central and South American DYS391*6 haplotypes.

In 2013, Battaglia and collaborators [57] suggested that all Native South American chromosomes classified as Q1a2-M346*(xM3)/Q-M242*(xM3) should be allocated to the Q1a2-L54*(xM3) haplogroup, representing the second largest Pan-American founder Y-chromosome lineage, including those carrying allele 6. Similarly, to identify new informative Y-SNPs of subclades within haplogroup Q1a2-L54*(xM3), Jota et al. [30] analyzed 1841 Native individuals from South America within this haplogroup, including eight Coyaima samples from Tolima. Based on these Coyaima data, two new Q1a2-L54*(xM3)-derived sublineages were identified, Q-SA03 (n = 7) and Q-SA02 (n = 1), both of which were found exclusively in this Native Colombian group.

The Coyaima Cariban linguistic family is one of the tribes belonging to the Pijao Native American group, which is only found in the Tolima department. When 15 Y-STRs were used to compare the haplotypes from our Q1a2-M346*(xM3)-DYS391*6 samples with the Coyaima data reported by Jota et al. [30] classified as Q-SA03 and Q-SA02, complete matches were found between four individuals. The same analysis was performed for the Coyaima sample from Franco-Candela and Barreto [28], in which all individuals were reported as Q-M242*(xM3). In this case, there was not a complete haplotype match. However, the samples showed differences in only one or two microsatellites. Therefore, a close genetic relationship between our data and the other two publications was confirmed (S6 Table).

The phylogenetic relationships reported to date between all people in the Central and South American populations with allele six at DYS391 who were typified as Q-M242*(xM3) were further analyzed using YFiler profiles (S6 Table). In this network, 13 different nodes representing a total sample of 25 individuals were found; the results are shown Fig 4, and the genetic information is provided in S6 Table.

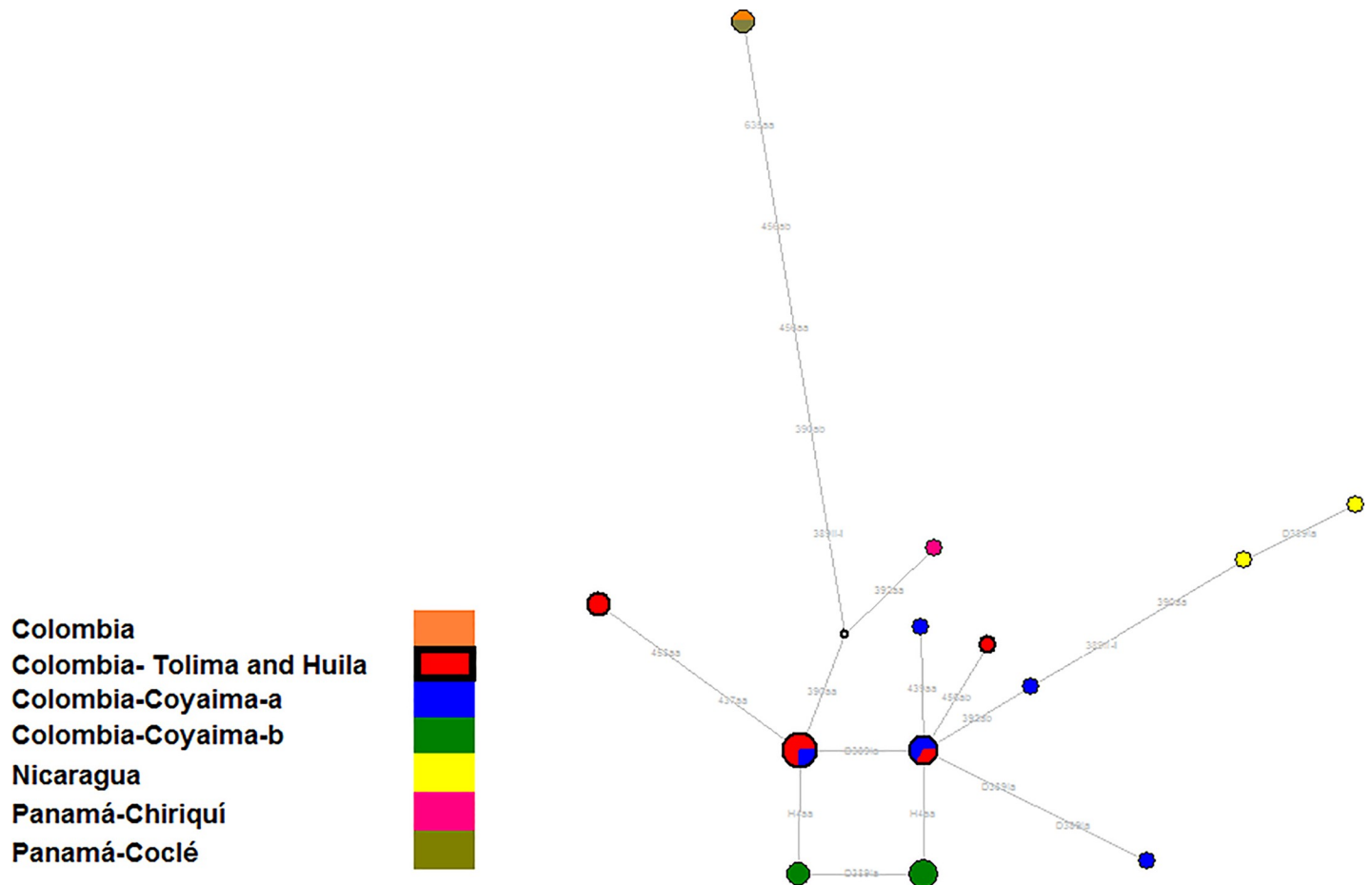


Fig 4. Median-joining network of 25 Q-M242*(xM3)-DYS391*6 South American Y-chromosomes using 15 loci STR. The circles represent haplotypes, with areas proportional to their frequencies; the numbers of haplotypes in each node are given in the figure mutation positions are shown; and the colors indicate the population of origin. Coyaima-a* represents the data reported by Jota et al. [30]. Coyaima-b* represents the data reported by to Franco-Candela and Barreto [28]. The median vectors (absent or extinct haplotypes) are shown in white.

<https://doi.org/10.1371/journal.pone.0207130.g004>

From this analysis, there were three main findings: (i) the Tolima and Huila region is characterized by the highest frequency of Q1a2-M346*(xM3)-DYS391*6 chromosomes (21 out of 25 haplotypes); (ii) 50% of these 25 haplotypes are different (N = 13); and (iii) the haplotype variability of these characteristic chromosomes is high in the Tolima and Huila samples, with nine of 13 Y-haplotypes differing.

These results indicate the middle Magdalena River region as the principal zone within all of Central and South America where this particular Native lineage is found. It is possible that all of the collected samples belong to the new sublineages Q1a2-SA02 and Q1a2-SA03 reported by Jota et al. [30]. However, to confirm this hypothesis, it would be necessary to genotype L54 and these two new Y-SNPs in all Central and South American samples carrying the *DYS391*6* allele. The time to the most recent common ancestor (TMRCA) for these 25 Q-M242*(xM3)-DYS391*6 Latin chromosomes is estimated to be 1,809 years (+/- 0,5345 years).

Additionally, it is fundamental to bear in mind the frequency of this allele in Panama. As mentioned above, this population exhibits 12 reported haplotypes, six of which are found in the Ngöbe Native group (these data were not included in the comparison because less than 10 STRs were genotyped). Panama is the country with the second-highest frequency of the Q* -

DYS391*6 lineage. Grugni et al. [58] proposed that the Isthmo-Colombian area is where this clade first arose.

The Isthmo-Colombian area has been described by O'Connor and Muysken [59] as a territory that was dominated by speakers of Chibchan languages for thousands of years. These Native individuals were distributed across four noncontinuous regions in Central and South America: eastern Honduras; from southern Nicaragua to western Panama; eastern Panama and northwest Colombia; and along the Magdalena River from Cundinamarca, Colombia northward to the Caribbean Sea. Therefore, this scenario is a plausible hypothesis in light of current evidence showing that all samples carrying DYS391*6 from the continental part of Colombia come from the Andean region.

In addition, a recent study of ancient human bones belonging to a Native Colombian group belonging to the Chibcha "Muisca" family from the IX to XVI centuries A.D. in the eastern Colombian Andes identified one male carrying allele 6 at locus DYS391 [60]. Based on this archeological finding, it is possible to confirm the presence of this allele in the genetic pool of Native Chibchan Nativepeople dating back to prehistoric times.

The narrow Isthmus of Panama is a likely bottleneck that should have allowed only a comparatively small number of nomadic hunters, fishermen, and gatherers to enter the northern Andean region via the Atrato, Cauca, and Magdalena rivers and tributaries through the northern reaches of the Colombian mountain ranges belonging to the greater Andean mountain range [59,61].

After subsequent demographic events such as isolation, genetic drift, and gene flow between Native American populations, speakers of different languages, such as Chibcha and Cariban languages, could have initiated the incorporation of this new Native lineage among local settlers of the middle Magdalena River region. Their descendants then survived through colonial times and waves of violence during the XIX and XX centuries. Today, they are still present at a high frequency in the Tolima and Huila departments and are part of the Colombian genetic pool.

European lineages

The influx of male Europeans represented the greatest legacy in both the Tolima and Huila departments (83.14%). These haplotypes were distributed among 14 different subclades, within haplogroups R, E, G, J, and T (Tables 1 and 2).

Haplogroup R1b-M269 was the most frequent in our sample (57.83%; $n = 48$), which was characterized by high haplogroup diversity, including six sublineages: R1b-S116*(xU152, M529, M65, M153, M167), R1b-U152, R1b-M529, R1b-M153, R1b-U106, and R1b-M167 (Table 1). Additionally, the R1b sublineages showed a high diversity rate, with just one shared haplotype within the R1b-U106 and R1b-S116 haplogroups. The greatest numbers of polymorphic loci were found within subclades R1b-S116, R1b-U152, and R1b-M529, with 16, 10, and 12 polymorphic loci, respectively (Table 2).

These results regarding haplogroups and diversity reflect the migration dynamics during colonization and at the end of the XIX century in Tolima and Huila. These dynamics were characterized by the arrival of settlers from various parts of Europe, mainly from the Iberian Peninsula (i.e., Basque Country, Spain, Italy, and France). In that region, the presence, frequency, and diversity levels of these R1b-M269 sublineages are similar to those reported from this analysis [62,63,64,65,66].

Other frequent haplogroups common in western Europe, such as J members ($n = 5$, Table 1), were also identified in Tolima and Huila, which may be related to the above historical migratory movements. Haplogroup J has been subdivided into two major clades, J1 and J2.

Both haplogroups were detected based on the M267, M410, and M67 mutations. This haplogroup evolved in the ancient Near East and was carried into North Africa, Europe, Central Asia, Pakistan, and India [67,68,69,70].

Eurasian clades T-M70 and G-M201 were also present, at frequencies of 6% ($n = 5$) and 3.61% ($n = 3$) respectively, neither of which showed shared haplotypes (Table 2). Haplogroup T-M70 is one of the most widely dispersed paternal lineages in the world. Haplogroup G-M201 is common in the Middle East, the Mediterranean, and the Caucasus Mountains [71].

African ancestry

Haplogroup E, defined by the M96 mutation, is the most common human Y chromosome clade found in Africa [72]. Within this haplogroup, E1b1b shows the widest geographic distribution, being present in North Africa, West Asia and southern Europe [73,74,75]. Three haplogroups within the E1b1b clade were found in the Tolima and Huila territory: E1b1b-M78 ($n = 3$, 3.61%), E1b1b-M81 ($n = 4$, 4.81%), and E1b1b-M123 ($n = 1$, 1.20%) (Table 1). It is more likely that these haplogroups were introduced by Europeans, since they have been reported at relatively high frequencies in Iberian populations as a result of Muslim invasions [74].

The most common sub-Saharan African lineage, E1b1a-M2 [72], was not detected in the sample.

From the point of view of Colombian history, this is an unexpected result, because approximately 200 thousand slaves arrived in the Colombian territory through the port of Cartagena during the colonial era [76] and then dispersed via the Magdalena and Cauca Rivers to other countries or were sold in Colombian markets. One of these internal markets was Honda City in the Tolima department.

However, in the last population census of Tolima and Huila, only approximately 38 thousand people (1.8% of the total population in both places) were reported as "Afro-Colombians" [77].

Hence, the absence of this lineage in our sample might be an effect of sample size due to its low population frequency. This low representation may have resulted from the high proportion of European men present in the region, which might have reduced the number of descendants carrying haplogroup "E1b1a" in this area due to their frequency.

Genetic relationships with other colombian populations

Based on the data provided in S7 Table, PCA was performed to visualize which Colombian populations share frequencies and which haplogroups generate higher separations given their representations. These results are shown in Fig 5.

In the presented diagram, the non-Native populations cluster in the first quadrant of the plane, along with the Tolima and Huila populations (corresponding to the populations examined in this study); this group is mainly characterized by high frequencies of European male lineages, including P*-92R7(xM167) and J-12f2a(xM9). In the case of Native populations, these clades are divided into two groups. The first is located in the second quadrant of the plot, where the frequency of Q1a2a*-M3 is close to one (Arhuaco, Kogi, Arzario, Emberá, Zenú, and Ticuna). The second group is located in the third quadrant, where Q lineages other than Q-M242*-(xM3) predominate (Waunana and Ingano). These results show clustering among the Tolima and Huila samples and other admixed Colombian populations, mainly based on the high prevalence of this European lineage, illustrating that the European influence is stronger in urban, rather than rural areas.

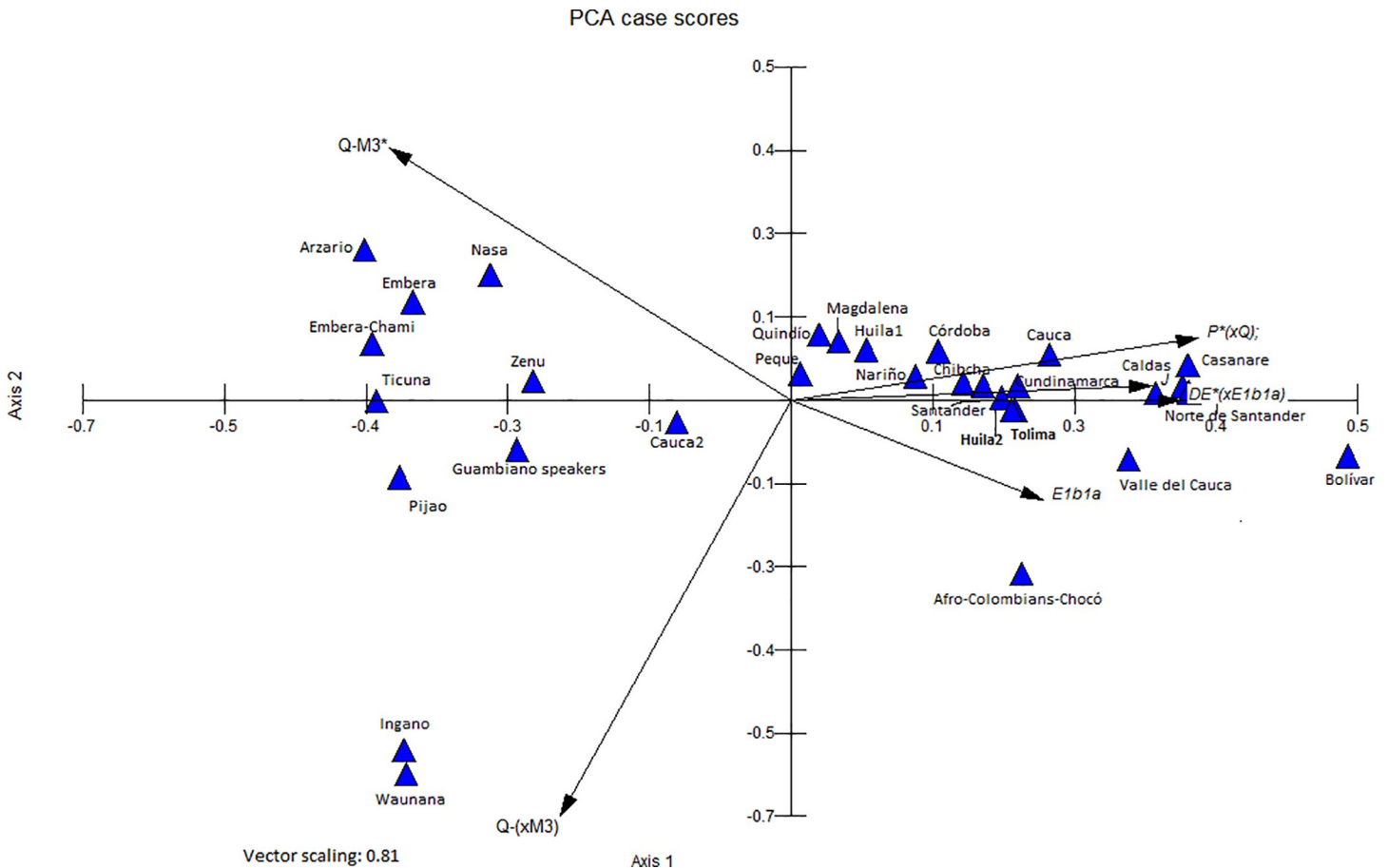


Fig 5. Principal component analysis for 32 Colombian populations using the haplogroup frequencies of Q-M242* (xM3), Q1a2-M3, P*-92R7(xM167), J-12f2a(xM9), DE*-YAP(xM2), and E1b1a-(M2). A total of 66.46% of the observed variance is explained by the first two axes represented in this figure.

<https://doi.org/10.1371/journal.pone.0207130.g005>

Conclusions

The location of the Tolima and Huila departments in the center of the country, characterized by the rainforest biome [78] and its connection with the rest of Colombia via the Magdalena River, has made the region one of the most important peopling areas in Colombia. This territory has received waves of migration of people from all over the world, from pre-Columbian times to the present.

Although this territory does currently not include any major Colombian city and is instead composed of many rural municipalities, the Y-chromosome gene pool found in this study is indicative of a population with high diversity and similar frequencies of paternal lineages that are very common in European regions such as Spain or the Iberian Peninsula. This diversity is a consequence of the economic and political historical importance of this region during colonial times and in the XIX century.

It is important to note the lack of Y-chromosomes belonging to haplogroup E1b1a-M2 in our sample. This finding does not preclude the possibility that there are descendants of this lineage in the Tolima and Huila departments. Men with this genetic inheritance do exist in the territory, albeit not in high frequencies. The probability of sampling these individuals is therefore lower than for other, more frequent lineages.

The analysis also revealed uncommon Native American haplogroups at high frequencies. All archeologic and genetic evidence suggests that the Q1a2-M346*(xM3)-DYS391*6 lineage

arose in prehistoric times in the Isthmo-Colombian area. However, after many generational and demographic events, this lineage currently only appears at a high density in and could even be considered near-exclusive to the Tolima and Huila departments. This hypothesis can be confirmed after analyzing more territories along the Colombian Andes in depth.

For this reason, it is important to assess new informative SNPs within the Q1a2 clade in Native and non-Native communities to obtain new insights and clues about the origin and dispersion routes that Native American groups have followed throughout time.

Supporting information

S1 Fig. Direct sequencing of the DYS391 system, showing the allele 6.
(PNG)

S2 Fig. Median-joining network of haplotypes DYS391*6 found in Asia and Latin American samples. The comparison was made using 15 STR markers. Circles represent haplotypes, with areas proportional to their frequencies; colors indicate the population of origin. The median vectors (absent or extinct haplotypes) are shown in white.
(TIF)

S1 Table. Population data sets used for Y-STR comparative analyses.
(XLSX)

S2 Table. List of Y-chromosome SNP haplogroups and STR haplotypes found in samples from the Tolima and Huila regions.
(XLSX)

S3 Table. List of the Y-STR haplotypes belonging to the haplogroup Q1a2-M3 among the Central and South American populations considered in the network analysis.
(XLSX)

S4 Table. List of the Y-STR haplotypes belonging to the haplogroup Q1a2-M242* (xM3) among the Central and South American populations considered in the network analysis.
(XLSX)

S5 Table. Y-chromosome haplotypes associated with the DYS391-6 allele among Asian and Latin American populations.
(XLSX)

S6 Table. Y-chromosome haplotypes associated with the DYS391-6 allele typified as Q-M242* (xM3) in Latin American populations.
(XLSX)

S7 Table. Absolute frequencies of Y-chromosome haplogroups and sub-haplogroups in the 30 Colombian samples included in the PCA.
(XLSX)

S1 File. Historic cartography of the Tolima and Huila departments.
(DOCX)

Acknowledgments

First, we are very grateful to the inhabitants of the Tolima and Huila departments for allowing us to gain better insight into their biological history. Furthermore, we want to thank the health centers: San Vicente de Paul Hospital in Garzón, San Sebastián Hospital and San Antonio

Hospital in La Plata, University Hospital in Neiva, San Francisco de Asís Hospital in Palermo, Arcenio Repiso Vanegas Hospital and The Culture House in San Agustín, Perpetuo Socorro Hospital in Villavieja, Pijao Salud in Coyaima, San Rafael Hospital in Espinal, San Juan de Dios Hospital, Honda Clinic, San Francisco Hospital, Minerva Clinic and Tolima Clinic in Ibagué, Regional Hospital in Libano, Natagaima Hospital, and Clinic Lab in Natagaima, for their support in the sampling stage. Additionally, we thank to the anthropologist Dobereiner Chala Aldana for his collaboration to this research in the designing and conducting of the sampling methodology. We also want to thank the members of the Human and Medical Genetics Laboratory, Institute of Biological Sciences, Federal University of Pará, and the Identification Genetics Group and the Genetics Institute of the Universidad Nacional de Colombia for their assistance in the developing of this research. Finally, we thank the Dirección de Investigación Bogotá (DIB), and the Faculty of Medicine of the Universidad Nacional de Colombia for the financial support of the project.

Author Contributions

Conceptualization: Luz Angela Alonso Morales, Madelyn Rojas Castro, Leonor Gusmão, William Usaquén.

Formal analysis: Luz Angela Alonso Morales, William Usaquén.

Funding acquisition: William Usaquén.

Investigation: Luz Angela Alonso Morales, Andrea Casas-Vargas, Madelyn Rojas Castro, Rafael Resque, Leonor Gusmão, William Usaquén.

Methodology: Luz Angela Alonso Morales, Andrea Casas-Vargas, Madelyn Rojas Castro, Rafael Resque, William Usaquén.

Project administration: William Usaquén.

Resources: Ândrea Kelly Ribeiro-dos-Santos, Sidney Santos, Leonor Gusmão.

Supervision: William Usaquén.

Validation: Leonor Gusmão, William Usaquén.

Writing – original draft: Luz Angela Alonso Morales.

Writing – review & editing: Luz Angela Alonso Morales, Andrea Casas-Vargas, Leonor Gusmão, William Usaquén.

References

1. Rivadeneira R. SIGNOS E IMÁGENES DEL POBLAMIENTO DEL RÍO GRANDE DE LA MAGDALENA [Internet]. *Revista Credencial*. 2018 [cited 28 April 2018]. Available from: <http://www.revistacredencial.com/credencial/historia/temas/signos-e-imagenes-del-poblamiento-del-rio-grande-de-la-magdalena>.
2. Pachón Castrillón X, Oliveros D, Luis W. *Geografía Humana de Colombia Región Andina Central Tomo IV Volumen II* [Internet]. 1st ed. Bogotá D.C: Instituto Colombiano de Cultura Hispánica; 2018 [cited 28 April 2018]. Available from: <http://babel.banrepcultural.org/cdm/ref/collection/p17054coll10/id/2778>
3. Rojas W, Parra MV, Campo O, Caro MA, Lopera JG, Arias W, et al. Genetic make up and structure of Colombian populations by means of uniparental and biparental DNA markers. *Am J Phys Anthropol*. 2010; 143(1):13–20. <https://doi.org/10.1002/ajpa.21270> PMID: 20734436
4. Ibarra A, Restrepo T, Rojas W, Castillo A, Amorim A, Martínez B, et al. Evaluating the X chromosome-specific diversity of Colombian populations using insertion/deletion polymorphisms. *PLoS One*. 2014; 9(1):1–10. <https://doi.org/10.1371/journal.pone.0087202> PMID: 24498042
5. Ossa H, Aquino J, Pereira R, Ibarra A, Ossa RH, Pérez LA, et al. Outlining the ancestry landscape of Colombian admixed populations. *PLoS One*. 2016; 11(10):1–15. <https://doi.org/10.1371/journal.pone.0164414> PMID: 27736937

6. Mesa NR, Mondragon MC, Soto ID, Parra M V., Duque C, Ortiz-Barrientos D, et al. Autosomal, mtDNA, and Y-chromosome diversity in Amerinds: Pre- and Post-Columbian patterns of gene flow in South America. *Am J Hum Genet.* 2000; 67(5):1277–86. [https://doi.org/10.1016/S0002-9297\(07\)62955-3](https://doi.org/10.1016/S0002-9297(07)62955-3) PMID: 11032789
7. Carvajal-Carmona LG, Soto ID, Pineda N, Ortíz-Barrientos D, Duque C, Ospina-Duque J, et al. Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in Northwest Colombia. *Am J Hum Genet.* 2000; 67(5):1287–95. [https://doi.org/10.1016/S0002-9297\(07\)62956-5](https://doi.org/10.1016/S0002-9297(07)62956-5) PMID: 11032790
8. Bortolini M-C, Salzano FM, Thomas MG, Stuart S, Nasanen SPK, Bau CHD, et al. Y-Chromosome Evidence for Differing Ancient Demographic Histories in the Americas. *Am J Hum Genet [Internet].* 2003; 73(3):524–39. <https://doi.org/10.1086/377588> PMID: 12900798
9. Rodas C, Gelvez N, Keyeux G. Mitochondrial DNA Studies Show Asymmetrical Amerindian Admixture in Afro-Colombian and Mestizo Populations. *Hum Biol.* 2003; 75: 13–30. PMID: 12713143
10. Bedoya G, Montoya P, Garcia J, Soto I, Bourgeois S, Carvajal L, et al. Admixture dynamics in Hispanics: A shift in the nuclear genetic ancestry of a South American population isolate. *Proc Natl Acad Sci [Internet].* 2006; 103(19):7234–9. <https://doi.org/10.1073/pnas.0508716103> PMID: 16648268
11. Rojas KM, Roa M, Briceño I, Guaneme C, Gómez A. Polimorfismos de 17 marcadores STR del cromosoma-Y en una muestra poblacional del altiplano cundiboyacense. *Colomb Med.* 2011; 42(1):88–97.
12. Noguera MC, Schwegler A, Gomes V, Briceño I, Alvarez L, Uricoechea D, et al. Colombia's racial crucible: Y chromosome evidence from six admixed communities in the department of bolivar. *Ann Hum Biol.* 2014; 41(5):453–9. <https://doi.org/10.3109/03014460.2013.852244> PMID: 24215508
13. Yunis JJ, Acevedo LE, Campo DS, Yunis EJ. Geno-geographic origin of Y-specific STR haplotypes in a sample of Caucasian-Mestizo and African-descent male individuals from Colombia. *Biomedica.* 2013; 33:459–67. <https://doi.org/10.7705/biomedica.v33i3.807> PMID: 24652182
14. Alonso LA, Usaquén W. Y-chromosome and surname analysis of the native islanders of San Andrés and Providencia (Colombia). *HOMO- J Comp Hum Biol.* 2013; 64(1):71–84. <https://doi.org/10.1016/j.jchb.2012.11.006> PMID: 23290785
15. Xavier C, Builes JJ, Gomes V, Ospino JM, Aquino J, Parson W, et al. Admixture and genetic diversity distribution patterns of non-recombining lineages of native american ancestry in colombian populations. *PLoS One.* 2015; 10(3):1–13. <https://doi.org/10.1371/journal.pone.0120155> PMID: 25775361
16. Ansari-Pour N, Moñino Y, Duque C, Gallego N, Bedoya G, Thomas MG, et al. Palenque de San Basilio in Colombia: genetic data support an oral history of a paternal ancestry in Congo. *Proc R Soc B Biol Sci [Internet].* 2016; 283(1827):20152980. <https://doi.org/10.1098/rspb.2015.2980> PMID: 27030413
17. López C, Realpe J. Cambios Paisajísticos y Localización de Evidencias Tempranas en el Valle Medio del Río Magdalena [Internet]. Pereira: Universidad Tecnológica de Pereira; 2008 [cited 28 April 2018]. Available from: <http://media.utp.edu.co/centro-gestion-ambiental/archivos/cambios-paisajisticos-en-el-valle-del-rio-mag/ecologiahistorica1.pdf>
18. Bernal E. El río magdalena: escenario primordial de la patria | Revista Credencial. *Rev Credencial [Internet].* 2013 [cited 2018 Mar 26]; Available from: <http://www.revistacredencial.com/credencial/historia/temas/el-rio-magdalena-escenario-primordial-de-la-patria>
19. Restrepo JD, Syvitski JPM. Assessing the Effect of Natural Controls and Land Use Change on Sediment Yield in a Major Andean River: The Magdalena Drainage Basin, Colombia. *AMBIO A J Hum Environ [Internet].* 2006; 35(2):65–74.
20. "QGIS Development Team (2018). QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>".
21. Geoportal [Internet]. Descarga del Marco Geoestadístico Nacional (MGN). Departamento Administrativo Nacional de Estadística DANE; 2017 [cited 2018Aug21]. Available from: <https://geoportal.dane.gov.co/v2/?page=elementoDescargaMGN>
22. Brooks GG. SouthAmerica (Local Copy) [Internet]. arcgis.com. ArcGIS; 2013 [cited 2018Aug21]. Available from: <https://www.arcgis.com/home/item.html?id=d3d2bae5413845b193d038e4912d3da9>
23. Banco de la República. Tolima—Enciclopedia | Banrepcultural [Internet]. Enciclopedia.banrepcultural.org. [cited 28 April 2018]. Available from: <http://enciclopedia.banrepcultural.org/index.php/Tolima>
24. López C. Ocupaciones Tempranas en las Tierras Bajas Tropicales del Valle Medio del Río Magdalena Sitio 05-Yon-002, Yondo-Antioquia. 1st ed. Bogotá: Fundación de Investigaciones Arqueológicas Nacionales (FIAN). Banco de la República; 1999.
25. ACNUR (Oficina del Alto Comisionado de las Naciones Unidas para los Refugiados). Pueblos indígenas de Colombia—2011 [Internet]. Acnur.org. 2011 [cited 3 April 2018]. Available from: http://www.acnur.org/fileadmin/scripts/doc.php?file=fileadmin/Documentos/Pueblos_indigenas/2011/Comunidades_indigenas_en_Colombia_-_ACNUR_2011

26. Rondón F, Osorio JC, Peña ÁV, Garcés HA, Barreto G. Diversidad genética en poblaciones humanas de dos regiones Colombianas. *Colomb Med.* 2008; 39(2 SUPPL.):52–60.
27. Criollo A. Caracterización molecular de la variación genética en cuatro etnias indígenas (Pijao, Paez, Embera y Zenu) y dos poblaciones mestizas de Colombia (Tolima y Córdoba) mediante marcadores del mtDNA, NRY Y AIMs. M.Sc, Universidad del Tolima. 2012. Available from: <http://repository.ut.edu.co/handle/001/1087>
28. Franco-Candela F, Barreto G. Estructura genética de poblaciones indígenas del occidente colombiano mediante el uso de marcadores ligados al cromosoma Y. *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales.* 2017; 41(160):281. <https://doi.org/10.18257/raccefyn.476>
29. Jota MS, Lacerda DR, Sandoval JR, Vieira PPR, Santos-Lopes SS, Bisso-Machado R, et al. A new sub-haplogroup of native American Y-Chromosomes from the Andes. *Am J Phys Anthropol.* 2011; 146(4):553–9. <https://doi.org/10.1002/ajpa.21519> PMID: 21913173
30. Jota MS, Lacerda DR, Sandoval JR, Vieira PPR, Ohasi D, Santos-Júnior JE, et al. New native South American Y chromosome lineages. *J Hum Genet.* 2016; 61(7):593–603. <https://doi.org/10.1038/jhg.2016.26> PMID: 27030145
31. Porras L, Phillips C, Fondevila M, Beltrán L, Ortiz T, Rondón F et al. Genetic variability of the SNPforID 52-plex identification-SNP panel in Central West Colombia. *Forensic Science International: Genetics.* 2009; 4(1):e9–e10. <https://doi.org/10.1016/j.fsigen.2008.12.003> PMID: 19948327
32. Gusmão L, Butler J, Carracedo A, Gill P, Kayser M, Mayr W et al. DNA Commission of the International Society of Forensic Genetics (ISFG): An update of the recommendations on the use of Y-STRs in forensic analysis. *Forensic Science International.* 2006; 157(2–3):187–197. <https://doi.org/10.1016/j.forsciint.2005.04.002> PMID: 15913936
33. Athey TW. Haplogroup Prediction from Y-STR Values Using a Bayesian-Allele-Frequency Approach. *J Genet Geneal.* 2006; 2(2):34–9.
34. Toscanini U, Gusmão L, Berardi G, Gomes V, Amorim A, Salas A et al. Male lineages in South American native groups: Evidence of M19 traveling south. *American Journal of Physical Anthropology.* 2011; 146(2):188–196. <https://doi.org/10.1002/ajpa.21562> PMID: 21826635
35. Resque R, Gusmão L, Geppert M, Roewer L, Palha T, Alvarez L, et al. Male lineages in Brazil: Intercontinental admixture and stratification of the European background. *PLoS One.* 2016; 11(4):1–17. <https://doi.org/10.1371/journal.pone.0152573> PMID: 27046235
36. Gusmão A, Gusmão L, Gomes V, Alves C, Calafell F, Amorim A et al. A Perspective on the History of the Iberian Gypsies Provided by Phylogeographic Analysis of Y-Chromosome Lineages. *Annals of Human Genetics.* 2008; 72(2):215–227. <https://doi.org/10.1111/j.1469-1809.2007.00421.x> PMID: 18205888
37. Fregel R, Gomes V, Gusmão L, González A, Cabrera V, Amorim A et al. Demographic history of Canary Islands male gene-pool: replacement of native lineages by European. *BMC Evolutionary Biology.* 2009; 9(1):181. <https://doi.org/10.1186/1471-2148-9-181> PMID: 19650893
38. Gomes V, Sánchez-Diz P, Amorim A, Carracedo Á, Gusmão L. Digging deeper into East African human Y chromosome lineages. *Human Genetics.* 2010; 127(5):603–613. <https://doi.org/10.1007/s00439-010-0808-5> PMID: 20213473
39. van Oven M, Van Geystelen A, Kayser M, Decorte R, Larmuseau M. Seeing the Wood for the Trees: A Minimal Reference Phylogeny for the Human Y Chromosome. *Human Mutation.* 2014; 35(2):187–191. <https://doi.org/10.1002/humu.22468> PMID: 24166809
40. Excoffier L, Lischer H. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources.* 2010; 10(3):564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x> PMID: 21565059
41. Nei M. *Molecular evolutionary genetics.* New York: Columbia University Press; 1987.
42. Bandelt H, Forster P, Rohlf A. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution.* 1999; 16(1):37–48. <https://doi.org/10.1093/oxfordjournals.molbev.a026036> PMID: 10331250
43. Toscanini U, Gusmão L, Berardi G, Amorim A, Carracedo Á, Salas A et al. Y chromosome microsatellite genetic variation in two Native American populations from Argentina: Population stratification and mutation data. *Forensic Science International: Genetics.* 2008; 2(4):274–280. <https://doi.org/10.1016/j.fsigen.2008.03.001> PMID: 19083836
44. Bolnick D, Bolnick D, Smith D. Asymmetric Male and Female Genetic Histories among Native Americans from Eastern North America. *Molecular Biology and Evolution.* 2006; 23(11):2161–2174. <https://doi.org/10.1093/molbev/msl088> PMID: 16916941

45. Muzzio M, Muzzio JC, Bravi CM, Bailliet G. Technical note: A method for assignment of the weight of characters. *Am J Phys Anthropol.* 2010; 143(3):488–92. <https://doi.org/10.1002/ajpa.21366> PMID: 20721942
46. Willuweit S, Roewer L. YHRD: Y-Chromosome STR Haplotype Reference Database [Internet]. Yhrd.org. 2018 [cited 28 April 2018]. Available from: <https://yhrd.org/>
47. Zhivotovsky L, Underhill P, Cinnioğlu C, Kayser M, Morar B, Kivisild T et al. The Effective Mutation Rate at Y Chromosome Short Tandem Repeats, with Application to Human Population-Divergence Time. *The American Journal of Human Genetics.* 2004; 74(1):50–61. <https://doi.org/10.1086/380911> PMID: 14691732
48. Shen P, Lavi T, Kivisild T, Chou V, Sengun D, Gefel D et al. Reconstruction of patrilineages and matrilineages of Samaritans and other Israeli populations from Y-Chromosome and mitochondrial DNA sequence Variation. *Human Mutation.* 2004; 24(3):248–260. <https://doi.org/10.1002/humu.20077> PMID: 15300852
49. MVSP Plus Version 3.1. Pentraeth, Wales, U.K: Kovach Computing Services; 2007. Available from: <http://www.kovcomp.co.uk/mvsp/MVSP3ProdInfo1.pdf>
50. Hidalgo Cerón VF. Análisis de marcadores moleculares Y-SNPs para una población bogotana y su aplicación en procesos de identificación humana. M.Sc Thesis, Universidad Nacional de Colombia. 2015. Available from: <http://www.bdigital.unal.edu.co/49833/>
51. Ruiz-Linares A, Ortiz-Barrientos D, Figuerola M, Mesa N, Munera J, Bedoya G et al. Microsatellites provide evidence for Y chromosome diversity among the founders of the New World. *Proceedings of the National Academy of Sciences.* 1999; 96(11):6312–6317. <https://doi.org/10.1073/pnas.96.11.6312>
52. Roewer L, Nothnagel M, Gusmão L, Gomes V, González M, Corach D, et al. Continent-Wide Decoupling of Y-Chromosomal Genetic Variation from Language and Geography in Native South Americans. *PLoS Genet.* 2013; 9(4). <https://doi.org/10.1371/journal.pgen.1003460> PMID: 23593040
53. Bisso-Machado R, Jota MS, Ramallo V, Paixão-Côrtes VR, Lacerda DR, Salzano FM, et al. Distribution of Y-chromosome Q lineages in native Americans. *Am J Hum Biol.* 2011; 23(4):563–6. <https://doi.org/10.1002/ajhb.21173> PMID: 21544893
54. Zhong H, Shi H, Qi X, Duan Z, Tan P, Jin L et al. Extended Y Chromosome Investigation Suggests Post-glacial Migrations of Modern Humans into East Asia via the Northern Route. *Molecular Biology and Evolution.* 2011; 28(1):717–727. <https://doi.org/10.1093/molbev/msq247> PMID: 20837606
55. Ascunce MS, González-oliver A, Mulligan CJ. Y-Chromosome Variability in Four Native American Populations from Panama. 2008; 80(3):287–302. <https://doi.org/10.3378/1534-6617-80.3.287>
56. Núñez C, Geppert M, Baeta M, Roewer L, Martínez-Jarreta B. Y chromosome haplogroup diversity in a Mestizo population of Nicaragua. *Forensic Sci Int Genet.* 2012; 6(6):4–7. <https://doi.org/10.1016/j.fsigen.2012.06.011> PMID: 22770600
57. Battaglia V, Grugni V, Perego UA, Angerhofer N, Gomez-Palmieri JE, Woodward SR, et al. The First Peopling of South America: New Evidence from Y-Chromosome Haplogroup Q. *PLoS One.* 2013; 8(8). <https://doi.org/10.1371/journal.pone.0071390> PMID: 23990949
58. Grugni V, Battaglia V, Perego UA, Raveane A, Lancioni H, Olivieri A, et al. Exploring the Y chromosomal ancestry of modern panamanians. *PLoS One.* 2015; 10(12):1–24. <https://doi.org/10.1371/journal.pone.0144223> PMID: 26636572
59. O'Connor L, Muysken P. *The Native Languages of South America.* Cambridge [u.a.]: Cambridge Univ. Pr; 2014.
60. Casas-Vargas A, Romero LM, Usaquén W, Zea S, Silva M, Briceño I, et al. Mitochondrial DNA diversity in Prehispanic bone remains on the Eastern Colombian Andes [Diversidad del ADN mitocondrial en restos óseos prehispánicos asociados al templo del sol en los andes orientales colombianos]. *Biomedica [Internet].* 2017; 37(4):1–41. <https://doi.org/10.7705/biomedica.v37i4.3377> PMID: 29373774
61. Rothhammer F, Silva C. Peopling of Andean South America. *American Journal of Physical Anthropology.* 1989; 78(3):403–410. <https://doi.org/10.1002/ajpa.1330780308> PMID: 2648861
62. López-Oarra AM, Gusmão L, Tavares L, Baeza C, Amorim A, Mesa MS, et al. In search of the pre- and post-neolithic genetic substrates in Iberia: Evidence from Y-chromosome in Pyrenean populations. *Ann Hum Genet.* 2009; 73(1):42–53. <https://doi.org/10.1111/j.1469-1809.2008.00478.x> PMID: 18803634
63. Myres NM, Rootsi S, Lin AA, Järve M, King RJ, Kutuev I, et al. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur J Hum Genet.* 2011; 19(1):95–101. <https://doi.org/10.1038/ejhg.2010.146> PMID: 20736979
64. Busby GBJ, Brisighelli F, Sanchez-Diz P, Ramos-Luis E, Martínez-Cadenas C, Thomas MG, et al. The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269. *Proc R Soc B Biol Sci [Internet].* 2012; 279(1730):884–92. <https://doi.org/10.1098/rspb.2011.1044> PMID: 21865258

65. Valverde L, Illescas MJ, Villaescusa P, Gotor AM, Garc'a A, Cardoso S, et al. New clues to the evolutionary history of the main European paternal lineage M269: Dissection of the Y-SNP S116 in Atlantic Europe and Iberia. *Eur J Hum Genet.* 2016; 24(3):437–41. <https://doi.org/10.1038/ejhg.2015.114> PMID: 26081640
66. Rey-González D, Gelabert-Besada M, Cruz R, Brisighelli F, Lopez-Soto M, Rasool M, et al. Micro and macro geographical analysis of Y-chromosome lineages in South Iberia. *Forensic Sci Int Genet.* 2017; 29:e9–15. <https://doi.org/10.1016/j.fsigen.2017.04.021> PMID: 28487219
67. Di Giacomo F, Luca F, Popa LO, Akar N, Anagnou N, Banyko J, et al. Y chromosomal haplogroup J as a signature of the post-neolithic colonization of Europe. *Hum Genet.* 2004; 115(5):357–71. <https://doi.org/10.1007/s00439-004-1168-9> PMID: 15322918
68. Cinnioğlu C, King R, Kivisild T, Kalfölu E, Atasoy S, Cavalleri GL, et al. Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet.* 2004; 114(2):127–48. <https://doi.org/10.1007/s00439-003-1031-4> PMID: 14586639
69. Sengupta S, Zhivotovsky LA, King R, Mehdi SQ, Edmonds CA, Chow C-ET, et al. Polarity and Temporality of High-Resolution Y-Chromosome Distributions in India Identify Both Indigenous and Exogenous Expansions and Reveal Minor Genetic Influence of Central Asian Pastoralists. *Am J Hum Genet [Internet].* 2006; 78(2):202–21. <https://doi.org/10.1086/499411> PMID: 16400607
70. Grugni V, Battaglia V, Hooshiar Kashani B, Parolo S, Al-Zahery N, Achilli A, et al. Ancient migratory events in the Middle East: New Clues from the Y-chromosome variation of modern Iranians. *PLoS One.* 2012; 7(7). <https://doi.org/10.1371/journal.pone.0041252> PMID: 22815981
71. Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res.* 2008; 18(5):830–8. <https://doi.org/10.1101/gr.7172008> PMID: 18385274
72. Trombetta B, D'Atanasio E, Massaia A, Ippoliti M, Coppa A, Candilio F, et al. Phylogeographic Refinement and Large Scale Genotyping of Human Y Chromosome Haplogroup E Provide New Insights into the Dispersal of Early Pastoralists in the African Continent. *Genome Biol Evol.* 2015; 7(7):1940–50. <https://doi.org/10.1093/gbe/evv118> PMID: 26108492
73. Cruciani F, La Fratta R, Torroni A, Underhill PA, Scozzari R. Molecular dissection of the Y chromosome haplogroup E-M78 (E3b1a): a posteriori evaluation of a microsatellite-network-based approach through six new biallelic markers. *Hum Mutat.* 2006; 27(8):831–2. <https://doi.org/10.1002/humu.9445> PMID: 16835895
74. Trombetta B, Cruciani F, Sellitto D, Scozzari R. A new topology of the human Y chromosome haplogroup E1b1 (E-P2) revealed through the use of newly characterized binary polymorphisms. *PLoS One.* 2011; 6(1):6–9. <https://doi.org/10.1371/journal.pone.0016073> PMID: 21253605
75. Goffredo S, Dubinsky Z. The mediterranean sea: Its history and present challenges. *Mediterr Sea Its Hist Present Challenges.* 2014; 529–547. Springer Science+Business Media Dordrecht. <https://doi.org/10.1007/978-94-007-6704-1>
76. Orobio A, Domínguez B, Cuesta E, Rodríguez E, Peña F, González M et al. Historia del pueblo Afrocolombiano [Internet]. Cali, Colombia: CEPAC Centro de Pastoral Afrocolombiana; 2013 [cited 5 April 2018]. Available from: <http://axe-cali.tripod.com/cepac/hispafrocol/3.htm>
77. Departamento Administrativo Nacional de Estadística [Internet]. Dane.gov.co. 2005 [cited 5 April 2018]. Available from: <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/censo-general-2005-1>
78. Moore J. A Prehistory of South America: Ancient Cultural Diversity on the Least Known Continent. 1st ed. University Press of Colorado; 2014