

Genetic analysis of Paleo-Colombians from Nemocón, Cundinamarca provides insights on the early peopling of northwestern South America

Marcela Díaz-Matallana^{1,2,*}, Alberto Gómez¹, Ignacio Briceño^{1,3}, José Vicente Rodríguez⁴

¹Instituto de Genética Humana, Facultad de Medicina, Pontificia Universidad Javeriana, Bogotá, Colombia

²Instituto Nacional de Medicina Legal y Ciencias Forenses, Genética Forense, Regional Oriente - Seccional Meta, Villavicencio, Colombia

³Facultad de Medicina, Universidad de La Sabana, Chía, Colombia

⁴Laboratorio de Antropología Física, Universidad Nacional de Colombia, Bogotá, Colombia

Abstract

The Checua site within Nemocón in the eastern Andean Region of Colombia has produced evidence of hunter-gatherer activities from 9500-8700 cal yr BP. The aim of this work was to analyze genetically the Checua group according to mitochondrial DNA HVR-I sequence data. Following strict criteria of authenticity for studies of ancient DNA, extracted DNA was amplified and sequenced for each of the human remains. Twelve of the 22 human remains from Checua were successfully assessed (54.54%) in this study. Paleo-Colombian Checua revealed the Native American haplogroups A2, B2 and C1, plus the rare haplogroup D4h3a, reported for the first time in ancient Colombia. Comparison between Paleo-American and modern Colombian genetic data demonstrates the continuity of A2, B2, and C1, and a probable discontinuity of D4h3a since prehistoric times in Colombia, if not a shortage of D4h3a records in the literature and/or databases, while D1 is present in modern Colombia. The presence of D4h3a, A2, B2, and C1 haplogroups from Colombian early Holocene burials supports genetic evidence for Paleo-American dispersals downward by the Pacific coast. Our Paleo-Colombian DNA data, along with other genetic, archaeological, and paleoenvironmental published data, allow us to bring forward an *Integrated Migration Model for the early peopling of northwestern South America*, which involved different movements: (1) North-to-South along the coast; (2) West-to-East across the Andean mountain ranges; (3) Riverine discovery and riverside colonization. © 2016. Acad. Colomb. Cienc. Ex. Fis. Nat.

Key words: Checua site, hunter-gatherers occupation, Paleo-DNA, mitochondrial DNA (mtDNA), Cundinamarca, Colombia, Migration waves.

El análisis genético de paleo-colombianos de Nemocón, Cundinamarca proporciona revelaciones sobre el poblamiento temprano del Noroeste de Suramérica

Resumen

El sitio Checua localizado en Nemocón en la Región de los Andes Orientales Colombianos, ha mostrado evidencia de actividades de cazadores-recolectores desde 9500-8700 cal AP. El objetivo de este trabajo fue analizar genéticamente el grupo Checua mediante datos de secuencia de ADN mitocondrial HVR-I. Siguiendo criterios estrictos de autenticidad para los estudios de ADN antiguo, se extrajo, se amplificó y se secuenció el ADN para cada uno de los restos humanos disponibles. Doce de los 22 restos humanos de Checua se evaluaron con éxito (54,54%) en este trabajo. Los Paleo-Colombianos Checua revelaron los haplogrupos nativos americanos A2, B2 y C1, además del raro haplogrupo D4h3a, reportado por primera vez en Colombia antigua. La comparación de los datos genéticos entre los Paleo-Americanos y Colombia moderna demuestra la continuidad de los haplogrupos A2, B2 y C1, y una probable discontinuidad de D4h3a desde tiempos prehistóricos en Colombia, si no una escasez de registros de D4h3a en la literatura y/o bases de datos, mientras que D1 está presente en Colombia actual. La presencia de los haplogrupos D4h3a, A2, B2, y C1 en entierros Colombianos del Holoceno temprano apoya la evidencia genética para dispersiones de Paleo-americanos hacia abajo por la costa del Pacífico. Nuestros datos Paleo-genéticos de Colombia, junto con otros datos genéticos, arqueológicos y paleoambientales publicados, nos permiten proponer un *Modelo Integrado de Migración para el poblamiento temprano del Noroeste de Suramérica*, que involucró diferentes movimientos: (1) Norte a Sur a lo largo de la costa; (2) Oeste a Este a través de las Cordilleras Andinas; (3) Descubrimiento fluvial y colonización de sus riberas. © 2016. Acad. Colomb. Cienc. Ex. Fis. Nat.

Palabras Clave: Sitio Checua; Ocupación de cazadores-recolectores; Paleo-ADN; ADN mitocondrial (ADNmt); Cundinamarca; Colombia, olas de migración.

Introduction

Information obtained from mitochondrial DNA (mtDNA) has been used extensively to clarify the genetic and evolutionary history of hominins, ancient humans, and prehistoric and modern populations (Martínez-Cruzado, *et al.* 2005; Fehren-Schmitz, *et al.* 2010; Chatters, *et al.* 2014). Previous research has revealed five “Pan-American” mtDNA haplogroups, A, B, C, D and X, classified in the major lineages A2, B2, C1, D1 and X2a, with the minor lineages C4c, D2a, D3a and D4h3 (Tamm, *et al.* 2007; Achilli, *et al.* 2008). More recently, complete mtDNA analysis has revealed 15 sub-haplogroups, considered as founder lineages of the New World: A2*, A2a, A2b, B2, C1b, C1c, C1d*, C1d1, C4c, D1, D2a, D3, D4h3a, X2a and X2g (Merriwether, *et al.* 1995; Achilli, *et al.* 2008; Perego, *et al.* 2009). Molecular data generated by different genetic systems and computer simulations has been used to clarify dispersal routes and periods of entry into the Americas. Amongst the proposed models are: (1) *A single migratory wave*, based on mtDNA, SNPs and Y-STR data (Stone and Stoneking 1998; Zegura, *et al.* 2004), either by coastal routes based on analysis of mtDNA and STRs (Yang, *et al.* 2010), or by coastal and riverine colonization by means of computer simulation model (Fix 2005), or specifically along the Pacific coastal route, based on the analysis of autosomal STR or whole mtDNAs (Wang, *et al.* 2007; Fagundes, *et al.* 2008; Goebel, *et al.* 2008; Perego, *et al.* 2012; Llamas, *et al.* 2016). (2) *A dual migration model*, according to the geographical distribution of two rare mtDNA lineages: D4h3 and X2a (Perego, *et al.* 2009). (3) *Multiple waves of migration*, including: a) a tripartite model based on linguistic, dental and genetic evidence (Greenberg, *et al.* 1986); b) a three-wave model based on analysis of complete mtDNAs or SNPs (Tamm, *et al.* 2007; Reich, *et al.* 2012); c) a three-stage model: mtDNA data (Mulligan, *et al.* 2008). Concerning the Caribbean, an indigenous South American origin for the Caribbean Taínos has been confirmed by genetic and genomic analyses carried out on Mestizo mtDNA (Díaz-Matallana 2009; Díaz-Matallana & Martínez-Cruzado 2010; Martínez-Cruzado 2010), on Mestizo nuclear SNPs (Moreno-Estrada, *et al.* 2013), and also on ancient DNA (aDNA) studies in the Dominican Republic and Cuba (Lalueza-Fox, *et al.* 2001, 2003). Moreover, a Native American gene flow from Central America to the Caribbean (Lalueza-Fox, *et al.* 2003; Moreno-Estrada, *et al.* 2013), as well as a back migration from South to Central America in ancient times have been proposed (Díaz-Matallana & Martínez-Cruzado 2010; Moreno-Estrada, *et al.* 2013). Regarding South America, different studies have suggested: a) *One migration wave* (Manríquez, *et al.* 2011); b) *One migration wave in two stages* (Fuselli, *et al.* 2003); c) *Two migration waves* (Keyeux, *et al.* 2002); d) *An early colonization from north to south, including subsequent trans-Andean movements in southern South America* (Bodner, *et al.* 2012; de Saint-Pierre, *et al.* 2012).

Archaeological and genetic studies of modern populations indicate an early colonization of the Americas took place sometime prior to 15 Ka. A small number of aDNA results from the Americas would postpone the time of settlement by several thousands of years. The oldest samples studied (≥ 8 ^{14}C ka BP), ostensibly the most informative for studies of the peopling of the continent, usually come from single individuals, while more recent ancient samples are suitable for the regional prehistory (Raff, *et al.* 2011; Pickrell & Reich 2014; Llamas, *et al.* 2016). Several findings of human remains from the Savannah of Bogotá (Figure 1a, 1b) and from Boyacá (eastern Cordillera, Colombia) along with other relevant findings from elsewhere in the Americas (Pleistocene/Holocene) are cited in Table 1. Those already analyzed through mtDNA markers appear in bold and are italicized.

The Checua Site. The open-air Checua site extends over the top of a hill, steeper at its western end, rising about 15 feet above the flat surrounding area. It has produced evidence of prehistoric activities by hunter-gatherers in a cultural sequence (8500-3000 ^{14}C yr BP). An archaeological excavation (300 m²) was carried out in 1991 under the coordination of A.M. Groot (Groot 1992, 1995, 2006). The name Checua is derived from the subdivision of the municipality of Nemocón (Figure 1a, 1b), Cundinamarca-Colombia (eastern Andes) that contains the site. This is a region of subtropical lower-montane dry forest: 2600 m.a.s.l. average altitude, 14°C annual mean temperature, and 600-750mm average annual precipitation (Groot 1992). The area is characterized by a native plant (*Condalia thomasiana*) that is typical of the *Condalia* forest (Van der Hammen 1997). Nemocón has two major rivers: the Checua River, covering 8334 Ha (85%), and the Neusa River covering 311 Ha (15%) (Groot 1992, 1995, 2006; Minelli, *et al.* 2012). Paleoenvironmental studies suggest great changes in temperature and precipitation took place during the Quaternary period. During the last glaciations (23–18 Ka) the mean annual temperatures of the Savannah of Bogotá were about 8°C lower than at present; the lower Andean zone of Savannah –currently 2500/2600 and 2800m altitude- remained for several thousand years between 1700 and 2000m below. At that time, the weather was relatively dry: annual precipitation in the region was 40% lower than today (Van der Hammen 2011). During Early Holocene (10 – 8 Ka), the climate changed with a sudden rise in average annual temperature that increased evaporation, causing lakes and swampy areas to dry out. The biome reconstructions for the 9 - 8 Ka reveal that there was a notable expansion of mesic biomes (Marchant, *et al.* 2002; Van der Hammen 2011). Paleocological / paleoenvironmental data predict that an initial wave of human occupation could have used coastal routes as primary adaptation areas. Then around 15–12 Ka, some groups may have turned east, crossing the major river valleys of the eastern highlands (“Altiplano Cundiboyacense”), taking advantage of the relatively dry

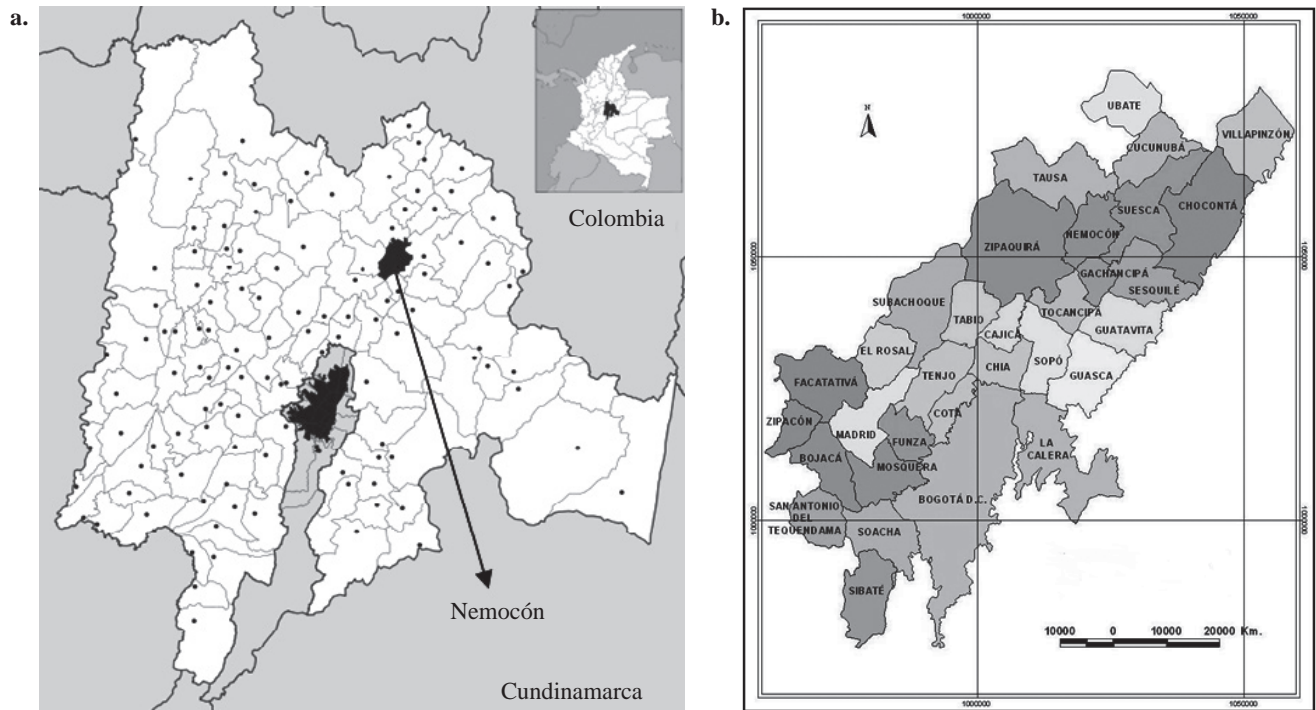


Figure 1. a. Cundinamarca Department of Colombia, Municipality of Nemocón. Source original map: Wikipedia (<https://en.wikipedia.org/wiki/Nemocón>). **b.** Savannah of Bogotá - Colombia, Municipality division. Source map: ICANH (<http://www.icanh.gov.co>). Checua remains were found on the “Finca Extremadura” (Nemocón), 7 km north of the municipality head (5°04' north latitude, 73°53' east longitude of Greenwich), at 60 km northeast of Bogotá, D.C. (Groot 1992; Neves, *et al.* 2007).

and cold environments, as well as lower water levels in fluvial systems (Dillehay, *et al.* 1999; López-Castaño & Cano-Echeverry 2011). Overall, it seems a relatively dry corridor linked Pacific coastal routes with the “Altiplano Cundiboyacense”, favoring the displacement of wildlife and humans from the Magdalena Valley towards the mountains towards the end of the Pleistocene (Van der Hammen 1986; Correal 1993).

Eight stratigraphic levels were identified at the Checua site, identifying two human occupation zones (Groot 1992; Neves, *et al.* 2007). Zone I corresponds to Stratum 4: 9500-8700 cal yr BP [8200 ± 110 ¹⁴C yr BP (Beta)]; Zone II corresponds to Stratum 5: 9100-8300 cal yr BP [7800 ± 160 ¹⁴C yr BP (Beta53924)] (Table 1). A more recent archaeological survey (Minelli, *et al.* 2012) reconstructed a stratigraphic sequence of seven levels (US7 to US1), using fragments of gastropods (snails) and charcoal samples as follows: Level US1 (3000-3500 ¹⁴C yr BP), level US3 (5837 ¹⁴C yr BP), level US4 (6461¹⁴C yr BP), and lower levels US5-7 (8500 ¹⁴C yr BP) (Lab CEDAD, Lecce-Italy) [Table 1]. Checua human remains were identified as Paleo-Americans due to anthropological analysis, showing dolichocephalic skulls (Groot 1992; Pucciarelli, *et al.* 2010). Also, found at the Checua site were ancient stone tools and bone elements closely connected with hunter-gatherer activities, as well as faunal remains (Groot 1992, 1995, 2006; Minelli, *et al.* 2012).

The aim of the present study was to analyze skeletal remains of Paleo-Colombian Checua through mtDNA markers in order to: 1) Identify possible Native American haplogroups (A2, B2, C1 and D1) in 22 Checua samples by mtDNA markers of HVR-I region; 2) Contrast the haplotypes obtained with the results from other aDNA and modern DNA studies; 3) Determine the phylogenetic relationships and genetic diversity of the individuals studied; 4) Integrate information from distinct scientific disciplines such as: human genetics (ancient and modern DNA), archaeology, anthropology, and paleoenvironment with the purpose of elucidating eventual human migration patterns for north-western South America.

Materials and Methods

Human Skeletal Remains. Samples from 22 human remains (N=22) from Checua-Nemocón, Cundinamarca, including 21 dental remains (premolars and molars), and 1 cranial (parietal) fragment (Groot 1992) belonging to the Reference Collection of the Physical Anthropology Lab (Universidad Nacional de Colombia-Bogotá) were analyzed. The Instituto Colombiano de Antropología e Historia (ICANH), pursuant to the provisions of the Culture Law 397/1997 (Act 1185/2008; Decrees 833/2002, 763/2009), considering the fulfillment of requirements, granted the *Authorization Statement on Archaeological Heritage N°4359* to: Marcela Díaz-Matallana, and the other team members (I. Briceño,

Table 1. Human Evidence from North, Meso, Central and South America (Pleistocene / Holocene)

North America	Cultural association	Age ¹⁴ C yr B.P.	Lab Code	Age range Cal yr B.P.	Material dated	Reference
New World Arctic: -Paleo-Eskimo	Saqqaq (Arctic Small Tool Tradition)	(4044 ± 31) ^a	(OxA-20656)	4170 – 3600	Hair	(Rasmussen, <i>et al.</i> 2010)
Southeast Alaska: -OYK: "Prince of Wales Island man"	Micro blade industry	9200 (correction) - (9730 ± 60) - (9210 ± 50)	(CAMS-29873) (CAMS-43990)	10300 - 11300 – 10800 [£] - 10550 – 10200 [£]	- Bone - charcoal	(Kemp, <i>et al.</i> 2007)
British Columbia, Canada: -Ancient-939	n.i.	(5710 ± 40)	(Beta-317343)	6260 – 5890	Bone	(Cui, <i>et al.</i> 2013)
Montana: - "Anzick-1" Infant	Clovis	(10,705 ± 35) ^a	(CAMS-80538)	12,707 – 12,556	Bone	(Rasmussen, <i>et al.</i> 2014)
Oregon, USA -Paisley Caves	Pre-Clovis	(12,265 ± 25) ^a	(UCIAMS-76190)	~14,300 14,595 – 14,079	Human Coprolite (cave 5)	(Jenkins, <i>et al.</i> 2008)
Santa Rosa Island, California: -Arlington Springs man	Late Pleistocene	(10,960 ± 80)	(CAMS-16810)	13,010 – 12,710	Bone	(Johnson, <i>et al.</i> 2002)
Meso America	Cultural association	Age ¹⁴ C yr B.P.	Lab Code	Age range Cal yr B.P.	Material dated	Reference
Central Mexico: - Peñón III woman - Tlapacoya I	-Early Preceramic -Early Preceramic	- (10,755 ± 75) ^a - (10,200 ± 65) ^a	n.i. n.i.	-12,770 – 12,560 -12,150 – 11,610	Bone Bone	(González, <i>et al.</i> 2005)
Quintana Roo, Mexico: - Hoyo Negro: "Naia"	Late Pleistocene	(10,976 ± 20) ^a	(Stafford / UCIAMS)	12,910 – 11,750	Tooth enamel	(Chatters, <i>et al.</i> 2014)
Central America	Cultural association	Age ¹⁴ C yr B.P.	Lab Code	Age range Cal yr B.P.	Material dated	Ref.
Turrialba Valley, Costa Rica: - Guardiría site	Early Preceramic / Clovis	n.i.	n.i.	12,000 – 10,000	Charcoal	(Cooke, <i>et al.</i> 2013)
Panamá: - Los Vampiros-1	Early Preceramic / Clovis	(11,550 ± 140)	(Beta-167520)	13,800 – 13,100 [£]	Organic sediment	(Cooke, <i>et al.</i> 2013)
South America	Cultural association	Age ¹⁴ C yr B.P.	Lab Code	Age range Cal yr B.P.	Material dated	Reference
Tocaima-Cundinamarca, Colombia: -Pubenza ^d	Pleistocene / Early Preceramic	(16,400 ± 420)	(Col.700, GrN-19.857)	21,000 – 18,500 [£]	Charcoal	(Van der Hammen & Correal 2001)
Zipacquirá-Cundinamarca, Col.: -El Abra	Early Preceramic	(12,460 ± 160)	(GrN-5556)	15,200 – 14,000 [£]	Charcoal	(Correal, <i>et al.</i> 1970)
Tocancipá, Cundinamarca, Col.: -Tibitó 1 ^d	Early Preceramic	(11,740 ± 110)	(GrN-9375)	13,800 – 13,300 [£]	Bone	(Correal 1990)
Junín, Cundinamarca, Col.: -Sueva 1	Early Preceramic	(10,090 ± 90)	(Col. 258, GrN-8111)	13,100 – 11,200 [£]	Charcoal (associated to human burial)	(Correal 1990, Neves, <i>et al.</i> 2007)
Cundinamarca, Col.: -Tequendam 1	Early Preceramic	(9740 ± 135)	(GrN-7115)	11,700 – 10,600 [£]	Charcoal (associated to human burials)	(Correal 1990, Neves, <i>et al.</i> 2007)
Gachalá-Cundinamarca, Col.: -Guavio 1	Early Preceramic	(9360 ± 45)	(GrN-8.448)	10,705 – 10,435 [£]	Charcoal (associated to Burial 8)	(Correal 1990, Neves, <i>et al.</i> 2007)
Cundinamarca, Col.: -Galindo 1	Early Preceramic	(8740 ± 60)	(GrN-16346)	10,150 – 9,500 [£]	Charcoal	(Pinto 2003)
Nemocón-Cundinamarca, Col.: -Checua (This Study)	Early Preceramic	- (8500 – 3000) ^b - (8200 ± 110) - (7800 ± 160) - (5770 ± 40) - (7850 ± 190) ^c - (5021 ± 202) ^c	(Beta / CEDAD Italy) (Beta) (Beta-53924) (Beta-278827) [♦] (Física-UNAL-EPR) (Física-UNAL-EPR)	- n.i. - 9500 – 8700 [£] - 9100 – 8300 [£] - 6660 – 6470 - n.i. - n.i.	- Charcoal; Gastropods (snails) - Charcoal (Stratum 4: Checua I) - Charcoal (Stratum 5: Checua II) - Human Bone (CH I-12) - Tooth enamel (Burial 9: CH I-09) - Tooth enamel (CH II-02)	- (Groot 1992, Minelli, <i>et al.</i> 2012) - (Groot 1992, 1995, 2006; Neves, <i>et al.</i> 2007) - (Sandoval & Almanza 2012) - (Carvajal, <i>et al.</i> 2014)
Cundinamarca, Col.: -Neusa	Early Preceramic	(8370 ± 90)	(Beta-21060)	9600 – 9100 [£]	Charcoal	(Rivera 1991)

South America	Cultural association	Age ¹⁴ C yr B.P.	Lab Code	Age range Cal yr B.P.	Material dated	Reference
Nemocón, Cundinamarca, Col.: -Nemocón 4	Early Preceramic	(7530 ± 100)	(GrN-8281)	8600 – 8000 [£]	Charcoal	(Correal 1990)
Cundinamarca, Col.: -Chía 3	Late Preceramic	(5040 ± 100)	(GrN-12122)	6000 – 5500 [£]	Human Bone	(Neves, <i>et al.</i> 2007)
Soacha-Cundinamarca, Col.: -Aguazuque	Late Preceramic	-(5025 ± 40) -(2725 ± 35)	(GrN-14.477) (GrN-14.479)	- 5894 – 5660 [£] - 2917 – 2757 [£]	Charcoal (associated to human burials)	(Correal 1990, Neves, <i>et al.</i> 2007)
Mosquera-Cundinamarca., Col.: -Vistahermosa	Late Preceramic	(3410 ± 35)	n.i.	3820 – 3572 [£]	Charcoal (associated to human burials)	(Correal 1990, Neves, <i>et al.</i> 2007)
Boyacá, Col.: -Floresta	Early Preceramic	(7950 ± 40)	(Beta-299693)	8990 – 8630	Human Bone	(Rodríguez 2011)
Santa Elena Peninsula, Ecuador: -Las Vegas	Early Preceramic	(10,000 – 6000) -(10,100 ± 130) -(8250 ± 120)	n.i. (Tx-4461) (Tx-3413)	-13,000 – 11,000 -12,200 – 11,200 [£] -9600 – 8900 [£]	-Shell -Human Bone	(Stohtert 1985)
Valle Chicama, Perú: -Huaca Prieta	Late Pleistocene / Pre-Clovis	-(12,950 ± 50)	(Beta-310274)	-14,200 – 13,300 -14,034 – 13,301	Sea lion Bone	(Dillehay, <i>et al.</i> 2012)
Venezuela: -Taima-taima ^d	Late Pleistocene / Pre-Clovis	(13,000 ± 200)	n.i.	16,200 – 14,900 [£]	Wood twigs	(Bryan, <i>et al.</i> 1978)
Lagoa Santa, Brazil: -Lapa Vermelha IV: “Luzia”	Early Preceramic	(10,030 ± 60)	(SacA17000/ Gif-12418)	11,710 – 11,243	Charcoal (near Luzia’s skull)	(Fontugne 2013)
Lagoa Santa, Brazil: -Cerca Grande 6	Early Preceramic	(8240 ± 40) ^a	n.i.	9400 – 9033 [£]	Human Bone	(González, <i>et al.</i> 2005)
Brazil: -Amazon Region	Ceramic / Pottery	4000 – 500 [£]	n.i.	n.i.	Association between human bone and the ceramic period	(Ribeiro-Dos-Santos, <i>et al.</i> 1996)
Argentinian Pampas: -Arroyo Seco 2 (AS2), Chocorí (Chc), La Tigra (LTi)	Early Preceramic	AS2 (6860 ± 60) ^a Chc (7010 ± 60) ^a LTi (7270 ± 60) ^a	n.i.	AS2: 7850 – 7550 [£] Chc: 8000 – 7700 [£] LTi: 8200 – 7950 [£]	Human Bone	(Figueiro & Sans 2007; Pucciarelli, <i>et al.</i> 2010; Llamas, <i>et al.</i> 2016: Central Andes & Mexico)
Chile: - <i>Northern: Quebrada de Acha; Camarones 14; Playa Miller 8; Morro 1, 1-6.</i> - <i>Central: Tagua-Tagua; St. Amelia.</i> - <i>Southern: Baño Nuevo.</i>	Quebrada de Acha (QA); Tagua-Tagua (TT); Baño Nuevo (BN): Early Preceramic	(9000 – 4000) ^b QA: (8790 ± 255); (8120 ± 90); (7570 ± 40). TT: (8070 – 6105) ^b BN: (8950 ± 50)	n.i.	n.i. QA: 10,600 – 9200 [£] ; 9400 – 9700 [£] ; 8430 – 8324 [£] . TT: n.i. BN: 10,250 – 9900 [£]	Human Bone	(Manríquez, <i>et al.</i> 2011)
Southern Chile -Monte Verde	Pleistocene / Pre-Clovis	n.i. (12,310 ± 40)	(Beta-239650)	14,600 – 14,200 -14,220 – 13,980	-Wood artifacts, charcoal -Seaweed <i>Mazaella</i> sp.	(Dillehay, <i>et al.</i> 2008)

^a Exact Age - AMS, ^b Stratigraphic Age, ^c Age -Electron Paramagnetic Resonance / Electron Spin Resonance (EPR / ESR), ^d Megafauna remains (mastodon) + lithic artifacts, ^e Estimated Age. Cal: Calibrated (‘calendar’) years BP.

£ Calibrations in This Study were computed in yr B.P. OxCal v4.2.4, 95.4% (Reimer *et al.* 2013). n.i.: no information. GrN: Groningen Lab, Netherlands. ♦ CHI-12 sample: The 13C/12C stable isotope ratio corresponded to -19.7‰, the 15N/14N stable isotope ratio was equal to +7.3‰. Note 1: In bold and italicized, prehistoric human remains analyzed by mtDNA markers. Note 2: The Age-EPR/ESR of two Checua samples agrees with the stratigraphy of the site.

A. Gómez, J.V. Rodríguez), letter N°130-3692. Additionally, ICANH authorized the Human Genetics Institute to send ancient samples (mtDNA) abroad, letter N°132-4349.

Authenticity Criteria for aDNA results. Initial research was conducted at Pontificia Universidad Javeriana-Bogotá (Human Genetics Institute). A second experimental phase was performed at University of Texas-Austin, USA (Anthropology). Stringent authenticity criteria were followed (Pääbo 2004; Willerslev & Cooper 2005; Bolnick, *et al.* 2012): (1) *aDNA extraction Zone and PCR reactions were setup in one building (pre-PCR).* The ancient DNA lab of the University of Texas at Austin is a restricted-access clean room dedicated to pre-PCR aDNA research with specialized

equipment, overhead UV lights ($\lambda = 254\text{nm}$), positive air pressure, and HEPA-filtered ventilation. DNA extractions and PCR set-up were performed in a laminar flow hood, with UV irradiating the entire lab for two hours after each use. (2) *Post-PCR area located in another building.* Thermal cyclers, electrophoresis, UV trans-illuminator/gel documentation systems were in this area. The flow from the post-PCR area to pre-PCR on the same day was avoided (Knapp, *et al.* 2012). (3) *Several DNA extractions and amplifications were carried out for each sample.* Extractions were carried out as many as three times, and reagents were prepared *de novo* for each extraction. Amplifications were carried out up to five times. (4) *Negative (blank) controls for*

DNA extraction, PCR and sequencing. Two or three negative controls were included in all PCRs. Negative controls served as monitors for potential contamination generated during the extraction and analyses. (5) *Decontamination of reagents and materials* [autoclave/UV light ($\lambda=254\text{nm}$)]. (6) *Essential disposable supplies and cleaning solutions.* Biosafety suit or coveralls (DuPont™ Tyvek®); two pair of gloves: latex/latex, or, latex/nitrile (Knapp, et al. 2012), facemasks, hair covers, shoe covers, and resealable plastic bags were used. Solutions such as 6% sodium hypochlorite (NaClO) were used to remove bone surface contamination (Kemp & Smith 2005), and samples were treated with both bleach and UV radiation. Solutions of 70% ethanol and DNase (Eshleman & Smith 2001) were used for cleaning micropipettes and acrylics. (7) *Reproducibility was tested at an independent ancient DNA lab* (Third experimental phase). Six samples were also tested independently at Universidad de los Andes-Bogotá, where the Anthropology (aDNA Lab: pre-PCR) and Human Genetics Lab (post-PCR) are located in separate buildings with specialized equipment for each area: starting by sample preparation, teeth/bone fragmentation, DNA extractions (salting-out), amplifications with four-overlapping pair of primers, and sequencing. For archaeological remains, reproducible amplification of the same sequence by independent laboratories using different parts of a specimen (i.e., tooth and femur) provides good support of authenticity (Willerslev & Cooper 2005). (8) *DNA typing of the researchers manipulating the samples.*

Sample Preparation and DNA Isolation. Preparing ancient samples and extraction protocol comprised Non-destructive Silica / Guanidine Thiocyanate (GuSCN) solution (Bolnick, et al. 2012), avoiding any further DNA purification. In addition, the salting-out extraction protocol fully adapted from previous studies (Jara 2007; Silva, et al. 2008), was improved for dental remains in this study (Díaz-Matallana 2015). DNA was purified with the MoBio® Power Clean kit, with silica columns for removing potential PCR inhibitors and humic substances (Lee, et al. 2010).

Amplification and Sequencing. Four-overlapping primer pairs specific for mtDNA (HVR-I) reported were implemented (Kemp, et al. 2007; Bolnick, et al. 2012); these not only generate short PCR products but also aDNA sequences of higher quality. The four HVR-I marker pairs were: (1) np 16011-16131, (2) np 16127-16229, (3) np 16210-16330, and (4) np 16250-16382. The advantage of these markers is that no molecular cloning is required, but the consensus sequence from overlapping sequences can be determined (Winters, et al. 2011). These primers have been successfully implemented on human skeletal remains from different periods including Paleo-Americans (Kemp, et al. 2007, 2009; Bolnick, et al. 2012; Mata-Míguez, et al. 2012). At the University of Texas-Austin, PCR reactions were performed in 7.4 μl containing 1.25 to 2 μl of DNA template, 1X Buffer (Invitrogen), 1X PCR Enhancer MasterAmp (Epicentre®), 0.2mM dNTPs (each), Bovine

Serum Albumin-BSA 15.6 μM (Roche®), 1.5mM MgCl₂, 0.8 μM primers, and 0.3U of Platinum Taq (Invitrogen®). At Universidad de los Andes-Bogotá, PCR reactions were adjusted to 25 μl , using Boline® reagents. An enhancer was not required, and Dimethyl sulfoxide-DMSO (Sigma®) was used instead of BSA. In this study, the HVR-I region was only analyzed and HVR-II was avoided, the obtention of mtDNA sequences of good quality and authentic results in ancient human remains represents a challenging task, also it consumes so much time, due to *Post-mortem* DNA is usually fragmented and degraded, plus the risk of contamination with human modern DNA (Pääbo, et al. 2004, Willerslev & Cooper 2005, Knapp, et al. 2012). Consequently other ancient DNA studies have focused exclusively in the HVR-I region (Kemp, et al. 2007, 2009; Manríquez, et al. 2011; Bolnick, et al. 2012; Mata-Míguez, et al. 2012; Mendisco, et al. 2014).

PCR Conditions. Initial denaturation was carried out at 94°C for 4 min, with 50 cycles (94°C for 30s, annealing temperature for 30s, 72°C for 30s) and a final extension of 72°C for 5min. The PCR (Touch Down-type) was carried out with the annealing temperature reduced 0.1°C after each cycle. It was observed that 58°C worked better to amplify all HVR-I fragments. Initially, 60 cycles were used (Bolnick, et al. 2012), but PCR was reduced to 50 cycles, showing best results in sequencing. Electrophoresis was performed in 1.5% agarose gels stained with GelRed (Biotium®) in 1X TBE buffer. Amplicons were visualized on a UV transilluminator, and a digital record was kept. At modern DNA Lab (University of Texas-Austin, Anthropology), small polyacrylamide gels stained with GelRed (Biotium®) were prepared and displayed on a gel documentation system. These gels have the advantage of requiring minimum volumes of PCR products (2-3 μl) and making it quicker to visualize bands. PCR products were purified prior to sequencing using the QIAGEN® QIAquick kit, and were later submitted to the DNA Sequencing Facility at the University of Texas-Austin, as well as to MacroGen Inc., Korea (ABI® 3730/3730XL sequencers), mediated by CorpoGen, Bogotá. Each sample was sequenced in both directions of HVR-I (light/heavy chains) as many as five times.

Data Analyses. The following softwares and statistics were used for data analysis: (1) OxCal v4.2.4: To calibrate radiocarbon dates of several human remains from the Americas [61]. (2) Sequencher® v5.0: Multiple sequence alignment with the revised Cambridge Reference Sequence - rCRS (Andrews, et al. 1999) and simultaneous visualization of electropherograms (GeneCodes Corp, Ann Arbor MI, USA). Only the mtDNA sequences that showed the best quality were used for analyses. The HVR-I segments (four overlapping fragments and repetitions) were aligned to the consensus sequence for each sample. (3) Network v4.6.1.2: Construction of phylogenetic networks through Median Joining (MJ) algorithm (Bandelt, et al. 1999) for each one of the mtDNA haplogroups (A2, B2, C1,

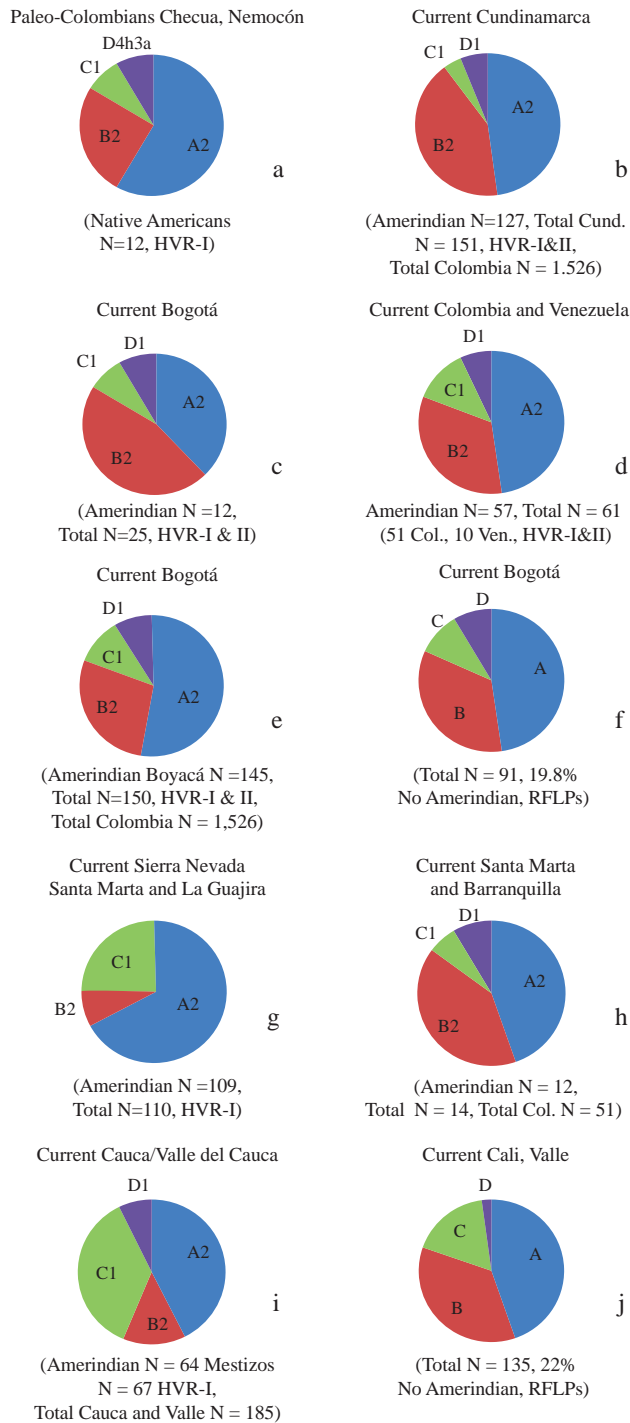


Figure 2. Paleo-DNA Vs. Neo-DNA Native American haplogroup frequencies from Colombia and Venezuela. **a)** *Paleo-Colombian Checua* (This Study) Vs. *Contemporary populations*: **b)** Cundinamarca (Yunis & Yunis 2013); **c)** Bogotá (Díaz-Matallana & Martínez-Cruzado 2010); **d)** Colombia and Venezuela (Díaz-Matallana & Martínez-Cruzado 2010); **e)** Boyacá (Yunis & Yunis 2013); **f)** Bogotá (Rodas, et al. 2003); **g)** Caribbean Coast: Arsario, Ijka, Kogi, Wayúú (Melton, et al. 2007); **h)** Santa Marta and Barranquilla (Díaz-Matallana & Martínez-Cruzado 2010); **i)** Cauca-Valle del Cauca (Salas, et al. 2007); **j)** Cali-Valle (Rondón, et al. 2008).

and D4h3a) where haplotypes of Paleo-Colombians and the Americas were compared. Haplotypes of 12 Paleo-Colombian Checua (this study) were compared with 962 ancient and modern published datasets (HVR-I) from the Americas (North, Central, Caribbean and South), for a total of 974 sequences contrasted in this way: haplogroup A2: seven (7) Paleo-Colombian Checua were compared with 414 ancient and modern mtDNAs HVR-I of the Americas; haplogroup C1: one (1) Paleo-Colombian Checua were contrasted with 383 ancient and modern mtDNAs HVR-I of the Americas; haplogroup D4h3a: one (1) Paleo-Colombian Checua were compared with 30 ancient and modern mtDNAs HVR-I of the Americas; haplogroup B2: three (3) Paleo-Colombian Checua were contrasted with 135 ancient and modern mtDNAs HVR-I of the Americas (Figure 1S, <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1650>). Each of the sequence positions of HVR-I was weighted in relation to its hyper mutability taking into account the estimated relative rates (Meyer, et al. 1999). (4) MEGA v6.06: a) Obtainment Tajima's neutrality test statistics (Tajima, et al. 1989) to analyze population expansion and selection for Paleo-Colombian Checua. b) Calculation of Mean genetic diversity measure through the Maximum Composite Likelihood method and Bootstrap method, and genetic distances (Tamura, et al. 2013). c) The Neighbor Joining (NJ) method was used to reconstruct a phylogenetic tree (Saitou & Nei 1987) through the Maximum Composite Likelihood model and Bootstrap method to obtain the complete arrangement of the Paleo-Colombian Native American haplogroups. d) Estimation of substitution patterns and rates was carried out under the Tamura-Nei model (Tamura & Nei 1993). A Maximum Likelihood (ML) estimate of Gamma parameter for Site Rates (Gamma distribution), was used to model evolutionary rate differences among sites (5 categories, [+G]: AC=AT=CG=GT, AG=CT) and to estimate nucleotide frequencies. The problem variation located around nucleotide position 16189 is usually associated with the length of heteroplasmy (Catelli, et al. 2011); therefore the 16182 and 16183 positions (B2 haplogroup) were ignored to calculate diversity indices, genetic distances, NJ phylogeny, and rates/patterns of nucleotide substitution for the Checua group. (5) DnaSp v5.10.1: Calculation of haplotype diversity and Fu's F_s neutrality test statistic (Fu 1997; Librado & Rozas 2009). (6) ANOVA statistical tests were used to differentiate haplotypes number among various ancient and modern groups (Kirkman 1996, Soper 2015). (7) MaxStat Pro® v3.6: A PCA biplot for understanding and visualizing the HVR-I variation among several ancient and modern groups was carried out.

Genetic diversity indices (MEGA6 and DnaSp5) were calculated for Checua (N=12) and for each haplogroups identified. The computations for each Occupation Zone (I and II) were avoided, since calibrated dates here confirmed overlapping layers (I: 9500-8700 / II: 9100-8300 cal yr

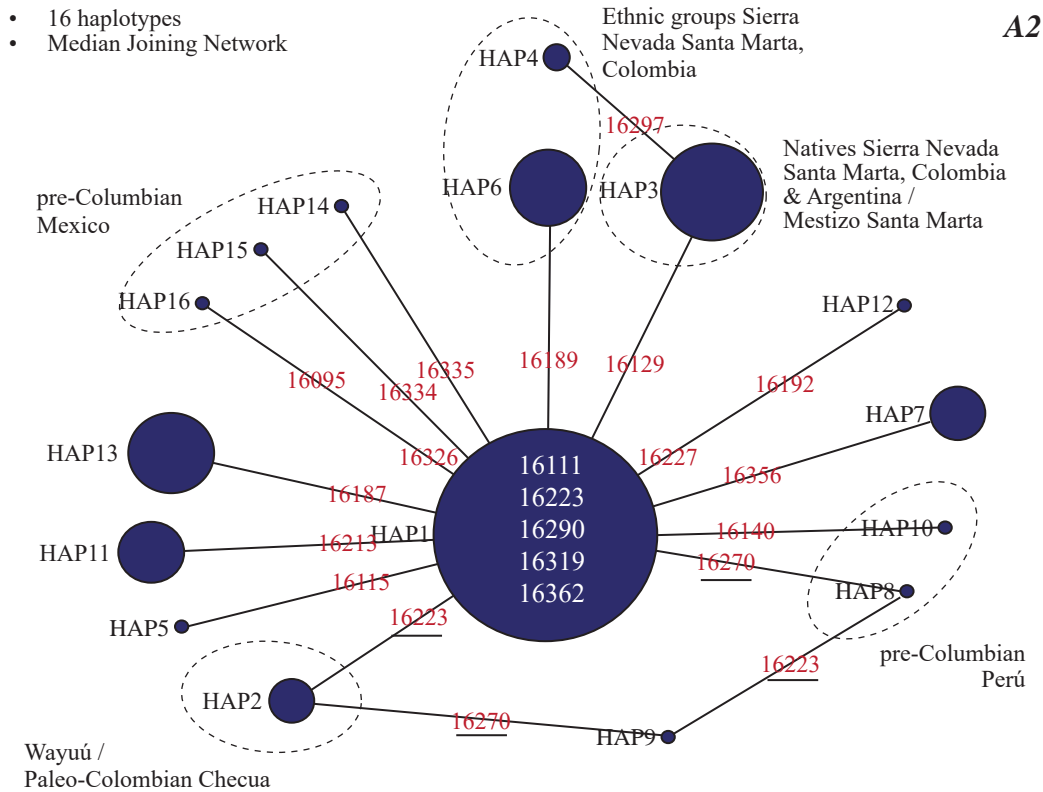


Figure 3. Phylogenetic Network of Haplogroup A2 (421 ancient and modern mtDNAs HVR-I of the Americas: 16 haplotypes) – Network v4.6.1.2. Seven (7) Paleo-Colombian Checua are represented in haplotypes 1, 2 and 5 (TS). The node size is related to its frequency. **HAP1:** Nodal Sequence of A2. **HAP1**= 5 Checua (TS), ancient Perú, Santa Marta, Col. & Venezuela, Natives Argentina & Chile, Colombians Valle, NA, CA & SA (Salas, *et al.* 2007; Díaz-Matallana & Martínez-Cruzado 2010; Fehren-Schmitz, *et al.* 2010; de Saint-Pierre 2012); **HAP2** = 1 Checua (CHII-03) (TS), Wayuú (Melton, *et al.* 2007); **HAP3** = Santa Marta, Col., Natives Argentina & Chile, Arsario, Ijka (Melton, *et al.* 2007; Díaz-Matallana & Martínez-Cruzado 2010; de Saint-Pierre 2012); **HAP4** = Kogi (Melton, *et al.* 2007); **HAP5** = 1 Checua (CHI-13) (TS); **HAP6** = Arsario, Kogi (Melton, *et al.* 2007); **HAP7** = Natives Argentina & Chile, Colombians Valle, NA, SA (Salas *et al.* 2007; de Saint-Pierre 2012]; **HAP8, HAP10** = ancient Perú (Fehren-Schmitz, *et al.* 2010); **HAP9, HAP12** = Natives Argentina (de Saint-Pierre 2012); **HAP11** = Bogotá & Venezuela, Colombians Valle, SA; **HAP13** = Nicaragua, Colombian Valle, CA, NA (Salas, *et al.* 2007; Díaz-Matallana & Martínez-Cruzado 2010); **HAP14 -16** = ancient Mexico (Mata-Míguez, *et al.* 2012). Mutations underlined mean that repeat several times. A unique reticulation or homoplasy is observed: HAP2 & HAP9 do not show np 16223, and HAP8 & HAP9 show np 16270. TS: This Study. NA, CA & SA: North, Central and South America.

BP); hence is problematic the partition of stratum to obtain estimates. It would be more appropriate such division if a great portion of the human remains were dated providing specific information; other Checua samples have been dated (Tables 1, 2).

Ethical considerations. The Scientific Investigation Committee on Medical Ethics of the Universidad Javeriana approved the research protocol of the Project N° 5176 by consensus (Minutes N°.8-2012).

Results

Identification of mtDNA Haplogroups and Haplotypes.

Out of 22 Checua samples, 12 were successfully typed in this study. Haplogroups and haplotypes are shown in Table 2. Results are listed according to Human Occupation Zones. Haplogroup frequencies were observed as follows: A2 = 58.33% (N = 7), B2 = 25% (N = 3), C1 = 8.33% (N = 1) and D4h3a = 8.33% (N = 1). Six of these ancient samples

(CHI-02, CHI-03b, CHI-8, CHI-11, CHI-13 and CHII-06) were verified in an external aDNA Lab at Universidad de los Andes-Bogotá, Colombia (Table 2).

Paleo-American DNA vs. Modern DNA. Figure 2 (a - d) shows the distribution of Native American haplogroups of Paleo-American Checua (HVR-I) obtained in this study in contrast with those from current Amerindian and Mestizo mtDNA haplogroups in Cundinamarca, Bogotá, and else where in northern South America, considering different published samplings based on sequencing results obtained from mtDNA control region (HVR-I and II). A genetic continuity between ancient and modern populations with haplogroups A2, B2, and C1 is evident, as well as perhaps a discontinuity of D4h3a (although haplogroup D1 remains relatively common in modern Colombian populations). Figure 2 (e-j) shows other distributions of Amerindian haplogroups from the Andean and adjacent lowland regions of Colombia, as well as from Mestizos and Native

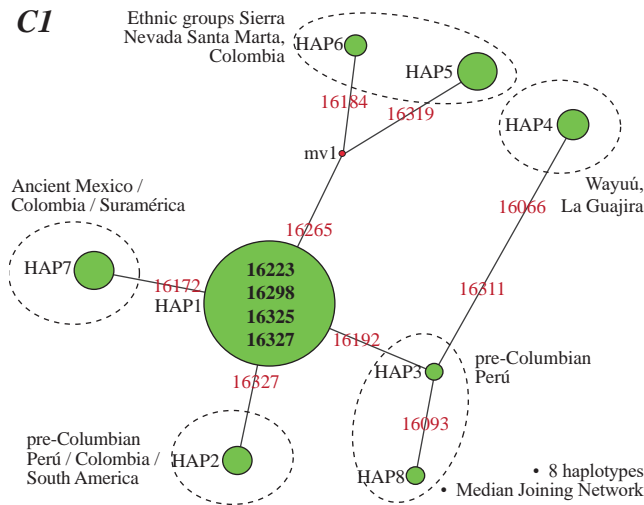


Figure 4. Phylogenetic Network of Haplogroup C1 (384 ancient and modern mtDNAs HVR-I of the Americas: 8 haplotypes) - Network v4.6.1.2. A single Paleo-Colombian Checua belongs to haplotype 1 (TS). The node size is related to its frequency. **HAP1:** Nodal Sequence of C1. **HAP1** = 1 Checua (CHI-08) (TS), ancient Perú, Wayuú, Colombian Valle, NA, CA, & SA (Melton, *et al.* 2007; Salas, *et al.* 2007; Fehren-Schmitz, *et al.* 2010); **HAP2** = ancient Perú, Colombian Valle, SA (Salas, *et al.* 2007; Fehren-Schmitz, *et al.* 2010); **HAP3, HAP8** = ancient Perú (Fehren-Schmitz, *et al.* 2010); **HAP4 - 6** = Wayuú, Arsario & Kogi (Melton, *et al.* 2007); **HAP7** = ancient Mexico (Mata-Míguez, *et al.* 2012), Colombians Valle, SA (Salas, *et al.* 2007). **Mv1:** Hypothetical ancestral node. TS: This Study. NA, CA & SA: North, Central and South America.

Americans from the Atlantic and Pacific coasts of Colombia. Most results were obtained by mtDNA sequencing (control region), though results from Bogotá and Cali were obtained by RFLPs (Keyeux, *et al.* 2002; Rodas, *et al.* 2003; Rondón, *et al.* 2007). Thus, it was not possible to discriminate subtypes of haplogroup D.

Phylogenetic Networks. Figures 3, 4, 5 show HVR-I haplotypes relationships of Paleo-Colombian Checua (this work, N=12), in contrast with individuals from other published studies for haplogroups A2, C1, D4h3a, and B2 respectively (962 ancient and modern datasets from the Americas) [See Data analyses; network B2 in Suppl. material Figure 1S].

Genetic Evidence of Early Migrations. Figure 6(a) illustrates the ancient human remains identified as belonging to haplogroup D4h3a, as well as the archaeological findings without human remains but interpreted as contributing to evidence for Paleo-American migrations into South America along the Pacific coast; a Paleo-Colombian Checua appears in this map (No. 4). Figure 6(b) illustrates ancient human remains carrying the A2, B2 and C1 haplogroups, as well as relevant archaeological sites with no human remains but interpreted as contributing to evidence for Paleo-American migrations into South America along the Pacific coast; 11 Paleo-Colombian Checua are included in this map (No. 4).

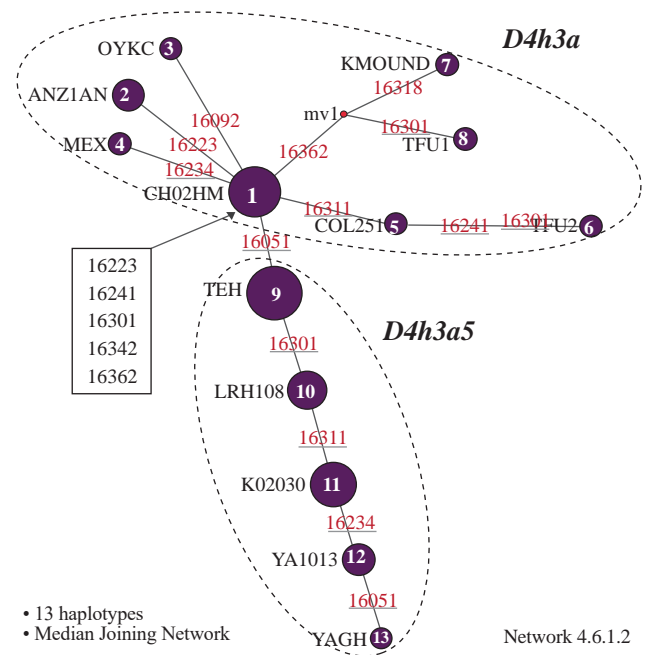


Figure 5. Phylogenetic Network of Haplogroup D4h3a (31 ancient and modern mtDNAs HVR-I of the Americas: 13 haplotypes) - Network v4.6.1.2. A single Paleo-Colombian Checua (CHI-02) belongs to haplotype 1 (TS). The node size is related to its frequency. **HAP1:** Nodal Sequence of D4h3a. **HAP1** = 1 Checua: CHI-02 (TS), Natives Chile, Mapuche, Mexican-Am (Kemp, *et al.* 2007; de Saint-Pierre 2012); **HAP2** = ANC939, Anzick-1 (Cui, *et al.* 2013; Rasmussen, *et al.* 2014); **HAP3** = OYKC; **HAP4** = Mexican (Kemp, *et al.* 2007); **HAP5** = Colombian (Bryc, *et al.* 2010); **HAP6** = post-Conquest Tierra del Fuego 2 (Kemp, *et al.* 2007); **HAP7** = Klunk Mound (1825 ± 75 ¹⁴C yr BP) [Bolnick 2005; Kemp, *et al.* 2007], 1925–1565 cal yr BP in This study (OxCal v4.2.4, 95.4%) [Reimer, *et al.* 2013], **HAP8** = post-Conquest Tierra del Fuego 1 (Kemp, *et al.* 2007); **HAP9 -13** = Natives Argentina & Chile (Kemp, *et al.* 2007; de Saint-Pierre 2012). **Mv1:** Hypothetical ancestral node. Mutations underlined mean that repeat several times. TS: This Study.

Estimates of Genetic Diversity. Two measures of selective neutrality (Tajima's D and Fu's F_s) were calculated to determine deviations from the Neutral Theory of Molecular Evolution (Table 3). Table 4 shows the estimated Mean genetic diversity (π_T) and haplotype diversity (Hd). Mean genetic distances between and within mtDNA haplogroups, are reflected in the arrangement of the NJ phylogenetic tree, in which the four Native American haplogroups (A2, B2, C1 and D4h3a) identified in the Checua remains are exhibited (Figure 2S, <http://www.racefyn.co/index.php/racefyn/article/downloadSuppFile/328/1652>). Also pairwise distances between sequences were calculated (Table 1S, <http://www.racefyn.co/index.php/racefyn/article/downloadSuppFile/328/1659>). Table 5 displays the comparison of total mtDNA genetic diversity (π_T) in Native Americans along the Americas; the π_T values closest to those of the Paleo-Colombian Checua are shown in bold

Table 2. Identification of mtDNA haplogroups and haplotypes (HVR-I) in *Checu* remains according to Human Occupation Zones^a

Occupation Zone II: <i>Stratum 5</i> 9100 – 8300 cal yr BP ^c	Origin ^e	Haplotype	HG	Sequence Read ^d
CHII-03	COL	16111T 16290T 16319A 16362C	A2	16011-16382
CHII-04	COL	16111T 16223T 16290T 16319A 16362C	A2	16011-16382
CHII-05 T5A	COL	16111T 16223T 16290T 16319A 16362C	A2	16011-16382
CHII-06 ^b	COL	16183C 16189C 16217C	B2	16011-16382
Occupation Zone I: <i>Stratum 4</i> 9500 – 8700 cal yr BP ^c	Origin ^e	Haplotype	HG	Sequence Read ^d
CHI-03a ^g	COL	16111T 16223T 16290T 16319A 16362C	A2	16011-16382
CHI-03b ^(b, g)	COL	16111T 16223T 16290T 16319A 16362C	A2	16011-16382
CHI-11 ^b	COL	16111T 16223T 16290T 16319A 16362C	A2	16011-16375
CHI-13 ^b	COL	16111T 16115T 16223T 16290T 16319A 16362C	A2	16011-16382
CHI-01	COL	16183C 16189C 16217C	B2	16011-16382
CHI-07	COL	16183C 16189C 16217C 16244A 16274A	B2	16011-16382
CHI-08 ^b	COL	16223T 16298C 16325C 16327T	C1	16011-16367
CHI-02 ^(b, f)	COL	16223T 16241G 16301T 16342C 16362C	D4h3a	16011-16382

^a Most results were obtained at University of Texas-Austin (USA), Dept. Anthropology: aDNA Lab / Modern DNA Lab (Díaz-Matallana et al. 2014, Díaz-Matallana 2015, This Study); ^b Six Sample results were verified at Universidad de Los Andes-Bogotá DC, (Colombia), Dept. Anthropology: aDNA Lab, and Human Genetics Lab (Díaz-Matallana et al. 2014, Díaz-Matallana 2015, This Study); ^c COL= Colombia (Nemocón, Cundinamarca); ^dHVR-I Coverage Region: 16011-16382 = Total 372bp. Identification through revised Cambridge Reference Sequence – rCRS (Andrews et al. 1999); ^e Age ¹⁴C yr BP Stratum 4 and 5 in Table 1; ^fCHI-02: Sub-adult < 6years old. ^gCHI-03a and CHI-03b: Sub-adults 3-4 years and 18-24 months old respectively, calculated by femur longitude and tooth development; ■ Age yr BP of other samples analyzed in This Study (see Table 1).

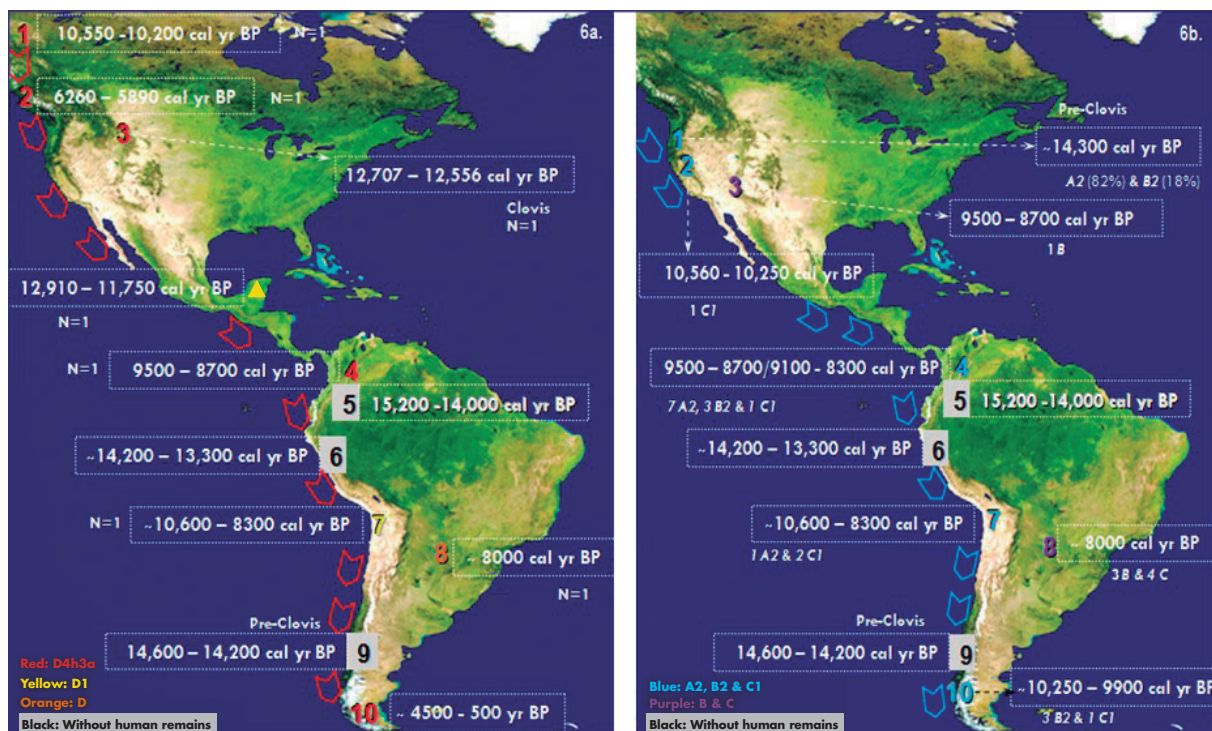


Figure 6. Genetic Evidence of early migrations along the Pacific coastal route: **a. Haplogroups D4h3a, D1 and D:** (1) OYKC; (2) ANC939; (3) Anzick-1; Yellow triangle: Naia ; (4) *Checu*: CH I-02 (TS); (5) El Abra; (6) Huaca Prieta; (7) Quebrada de Acha; (8) Argentine Pampas; (9) Monte Verde; (10) Tierra del Fuego (Faltyskova et al. 2014). **b. Haplogroups A2, B2 and C1:** (1) Paisley Caves; (2) Wizard’s Beach (Chatters et al. 2014); (3) Hourglass Cave (Mosch & Watson 1997); (4) *Checu* (TS); (5) El Abra; (6) Huaca Prieta; (7) Quebrada de Acha; (8) Argentinian Pampas; (9) Monte Verde; (10) Southern Chile. (See other Refs. cited in Table 1). TS: This Study. Source original map: http://www.mappi.net/Americas_satellite_map.jpg

Table 3. Neutrality Test Statistics

<i>Checua</i>	m	S	p_s^1	Θ^1	π^1	D^1	Fu's F_s^2
Total ^a	12	19	0.053371	0.017673	0.015109	-0.634978	0.169
HG A2 ^b	7	3	0.008197	0.003346	0.002342	-1.358415	-0.237
HG B2	3	2	0.005277	0.003518	0.003518	NAN	1.061

¹Tajima's test (Tajima 1989) was conducted in MEGA6 (Tamura, et al. 2013). ^aThis analysis involved 12 nucleotide sequences. All positions containing gaps and missing data were eliminated; there were a total of 357 positions in the final dataset. ^bThis analysis involved 7 nucleotide sequences. All positions containing gaps and missing data were eliminated; there were a total of 366 positions in the final dataset. Abbreviations: m = number of sequences, n = total number of sites, S = Number of segregating sites, $p_s = S/n$, $\Theta = p_s/a1$, π = nucleotide diversity, and D is the Tajima test statistic. Estimates for C1 and D4h3a haplogroups were excluded from these analyzes since MEGA6 calculates between at least two sequences. Tajima's D not significant $p > 0.10$.²Fu's F_s statistic (Fu 1997) was computed in DnaSp5 (Librado & Rozas 2009). Fu's F_s not significant $p > 0.02$.

Table 4. Estimates of Genetic Diversity

<i>Checua</i>	m	$d (\pi_r)^1$	S.E. ¹	Hd ²	S.D. ²
Total ^a	12	0.016	0.004	0.833	0.100
A2	7	0.002	0.001	0.524	0.209
B2	3	0.004	0.003	0.667	0.314

¹Evolutionary analyses were conducted in MEGA6 (Tamura, et al. 2013). ^aThe number of base substitutions per site from mean diversity (d) calculations for the entire population are shown (Nei and Kumar 2000). Standard error estimate(s) are shown in the second column and were obtained by a bootstrap procedure (500 replicates). Analyses were conducted using the Maximum Composite Likelihood model. The analysis involved 12 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 356 positions in the final dataset. ²Haplotype (gene) diversity (Hd), and Standard Deviation were computed in DnaSP5 (Librado & Rozas 2009).

and are italicized. The Maximum Likelihood (ML) estimate of Gamma parameter for site rates was calculated, and the estimated value of the shape parameter for the discrete Gamma Distribution was equal to 0.5689, denotes a great heterogeneity rate of nucleotide substitution among sites. Mean evolutionary rates in these categories were 0.03, 0.19, 0.52, 1.14, and 3.12 substitutions per site, the last category corresponds to transitions (AG=CT). The nucleotide frequencies were A=33.05%, T=23.24%, C=32.94%, and G=10.77%.

ANOVAs. It may be noted that pre-Columbian communities (A: Checua, Ciboney, Taino, Guane, Palpa-Perú, and this study's Checua) [Lalueza-Fox, et al. 2001, 2003; Fehren-Schmitz, et al. 2010; Casas-Vargas, et al. 2011], in spite of the smaller sample sizes (N), exhibit a greater variety of haplotypes in comparison with modern ethnic groups that have been mentioned, such as Central America (B: Ngobe, Huetar, Kuna) [Batista, et al. 1995; Melton, et al. 2013], Caribbean Colombia (C: Arsario, Ijka, Kogi) [Melton, et al. 2007], and Amazon (D: Guahibo, Gaviao, Zoro, Xavante) [Lewis, et al. 2005; Vona, et al. 2005; Melton, et al. 2007] from which a larger sample sizes have been drawn (Table 2S, <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1660>; table 3S, <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1665>; Figure 3S, <http://www.raccefyn.co/index.php/raccefyn/article/>

[downloadSuppFile/328/1653](http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1653) and figure 4S, <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1654>). ANOVA analysis for Number of haplotypes (HVR-I) showed that the *pre-Columbians* -Group A- exhibit the greatest variation, range of 95% confidence for the media: 6,489 - 10.71 ($p < 0.05$) [Suppl. Table 3S; Suppl. Figures 3S] followed by the other Groups D, B, and C. Moreover, the One-way ANOVA (Between-subjects) shows at least two of the means of haplotypes are dissimilar from each other ($p < 0.05$, $p < 0.01$), but also indicates a greater amount of observed variation within (1,826) than between (1,315) groups (Table 4S, <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1661>).

PCA Biplot of mtDNA HVR-I Data in Ancient and Modern Population groups.

To determine visual biological relationships of the *Paleo-Colombian Checua* among ancient and modern groups, a PCA biplot was constructed using Amerindian mtDNA (HVR-I) haplogroup frequency data (Figure 7). According to the obtained scores, the closest populations to Paleo-Colombian Checua (1.286) are the following: Guane (1.175), Paracas Palpa (1.709), Wayuú (1.081), Emberá (0.652), Maya (0.477), and Arsario (0.358) (Table 5S, <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1662>; table 6S, <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1663>; table 7S, <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1664> and figure 5S, <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1655>; figure 6S, <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1656> and figure 7S, <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1658>).

Discussion

Identification of mtDNA Haplogroups and Haplotypes.

In summary, 12 of the 22 individuals of the early Holocene at the Checua site yielded good DNA quality from the HVR-I sequences and clearly revealed the presence of the Native American haplogroups' diagnostic mutations (Tables 1, 2). This represents a success rate of 54.54% from the samples analyzed. Similarly, an aDNA Chilean

Table 5. Comparison of mtDNA Genetic Diversity (HVR-I) in Amerindians from North America, Mesoamerica, Central America, Caribbean and South America

Population / Community / Ethnic Group	N	Nucleotide Diversity (π_T)	Region	Ref.
Eskimo	77	0.0084	North America	(Vona, <i>et al.</i> 2005)
Cree	11	0.0104	North America	(Yang, <i>et al.</i> 2010)
<i>Shuswap-speaking</i> (territory)	31	<i>0.0156</i>	North America	(Malhi, <i>et al.</i> 2010)
<i>Nuu-Chah-Nulth</i>	63	<i>0.016</i>	North America	(Stone & Stoneking 1998)
<i>Yakima</i>	42	<i>0.015</i>	North America	(Stone & Stoneking 1998)
Cheyenne	39	0.0213	North America	(Vona, <i>et al.</i> 2005)
<i>Oneota – pre-Columbian</i>	50	<i>0.015</i>	North America	(Stone & Stoneking 1998)
Pima	44	0.0202	North America	(Vona, <i>et al.</i> 2005)
Mixtec	17	0.0096	Mesoamerica	(Yang, <i>et al.</i> 2010)
Zapotec	19	0.0114	Mesoamerica	(Yang, <i>et al.</i> 2010)
Mixe	20	0.0101	Mesoamerica	(Yang, <i>et al.</i> 2010)
<i>Xaltocan – pre-Conquest</i>	10	<i>0.017</i>	Mesoamerica	(Mata-Míguez, <i>et al.</i> 2012)
<i>Xaltocan – pos-Conquest</i>	15	<i>0.018</i>	Mesoamerica	(Mata-Míguez, <i>et al.</i> 2012)
<i>Xaltocan (full)</i>	25	<i>0.018</i>	Mesoamerica	(Mata-Míguez, <i>et al.</i> 2012)
<i>Maya</i>	34	<i>0.017</i>	Mesoamerica	(Melton, <i>et al.</i> 2007)
Ngöbe	46	0.012	Central America	(Melton, <i>et al.</i> 2013)
Quitirrisí - Huétar	29	0.010	Central America	(Melton, <i>et al.</i> 2013)
<i>Zapátón - Huetar</i>	25	<i>0.017</i>	Central America	(Melton, <i>et al.</i> 2013)
Kuna	63	0.009	Central America	(Batista, <i>et al.</i> 1995)
<i>Emberá</i>	44	<i>0.017</i>	Central America	(Kolman & Bermingham 1997)
Wounan	31	0.019	Central America	(Kolman & Bermingham 1997)
Ciboney – pre-Columbian	15	0.009	Caribbean Cuba	(Lalueza-Fox, <i>et al.</i> 2003)
Taino – pre-Columbian	19	0.008	Caribbean Dom. Rep.	(Lalueza-Fox, <i>et al.</i> 2001)
Arsario	28	0.012	Caribbean Colombia	(Melton, <i>et al.</i> 2007)
Ijka	31	0.004	Caribbean Colombia	(Melton, <i>et al.</i> 2007)
Kogi	21	0.009	Caribbean Colombia	(Melton, <i>et al.</i> 2007)
<i>Wayúú</i>	30	<i>0.016</i>	Caribbean Colombia	(Melton, <i>et al.</i> 2007)
Zenú	15	0.0054	Caribbean Colombia	(Yang, <i>et al.</i> 2010)
Santa Marta & Barranquilla – Mestizo ^a	11	0.0118	Caribbean Colombia	(Díaz-Matallana & Martínez-Cruzado 2010)
Venezuela – Mestizo ^a	8	0.0127	Caribbean Venezuela	(Díaz-Matallana & Martínez-Cruzado 2010)
Guahibo	59	0.0132	Amazon Venezuela	(Vona, <i>et al.</i> 2005)
Xavante	25	0.009	Amazon Brazil	(Melton, <i>et al.</i> 2007)
Zoró	29	0.011	Amazon Brazil	(Melton, <i>et al.</i> 2007)
Guane – pre-Columbian (Santander)	17	0.0231	Andes Colombia	(Casas-Vargas, <i>et al.</i> 2011)
<i>Checua – Paleo-Colombian</i> (Cundinamarca)	12	<i>0.016</i>	Andes Colombia	(This Study)
Bogotá DC & Villa de Leyva, Boyacá – Mestizo ^a	33	0.0105	Andes Colombia	(Díaz-Matallana & Martínez-Cruzado 2010)
<i>Cayapa</i>	30	<i>0.018</i>	Andes Ecuador	(Rickards, <i>et al.</i> 1999)
Quechua	18	0.0098	Andes Perú	(Yang, <i>et al.</i> 2010)
<i>Ancash</i>	34	<i>0.0162</i>	Andes Perú	(Fehren-Schmitz, <i>et al.</i> 2010)
Arequipa	22	0.0140	Andes Perú	(Fuselli, <i>et al.</i> 2003)
<i>Tayacaja</i>	60	<i>0.0166</i>	Andes Perú	(Fuselli, <i>et al.</i> 2003)
Paracas – pre-Columbian (Peninsula)	25	0.0105	Andes Perú	(Fehren-Schmitz, <i>et al.</i> 2010)
<i>Palpa – pre-Columbian</i> (Middle Horizon)	6	<i>0.0174</i>	Andes Perú	(Fehren-Schmitz, <i>et al.</i> 2010)
<i>Pacapaccari</i>	12	<i>0.0169</i>	Andes Perú	(Fehren-Schmitz, <i>et al.</i> 2010)
<i>Mapuche</i>	39	<i>0.017</i>	Andes Argentina	(Melton, <i>et al.</i> 2007)
Huilliche	20	0.0093	Andes Chile	(Yang, <i>et al.</i> 2010)
Africans (Sub-Saharan) ^b	20	0.024	Africa	(Nei & Kumar 2000)
Asians ^b	20	0.019	Asia	(Nei & Kumar 2000)

^a Estimates obtained from HVR-I + HVR-II. ^b Additional estimates from Africa and Asia. Note 1: In bold and italicized, π_T values closest to the Paleo-Colombian *Checua* (This Study). Note 2: The π measure does not depend on sample size (N), therefore is useful for comparing the extents of polymorphism at sequence level (ADNmt) of different populations (Nei & Kumar 2000).

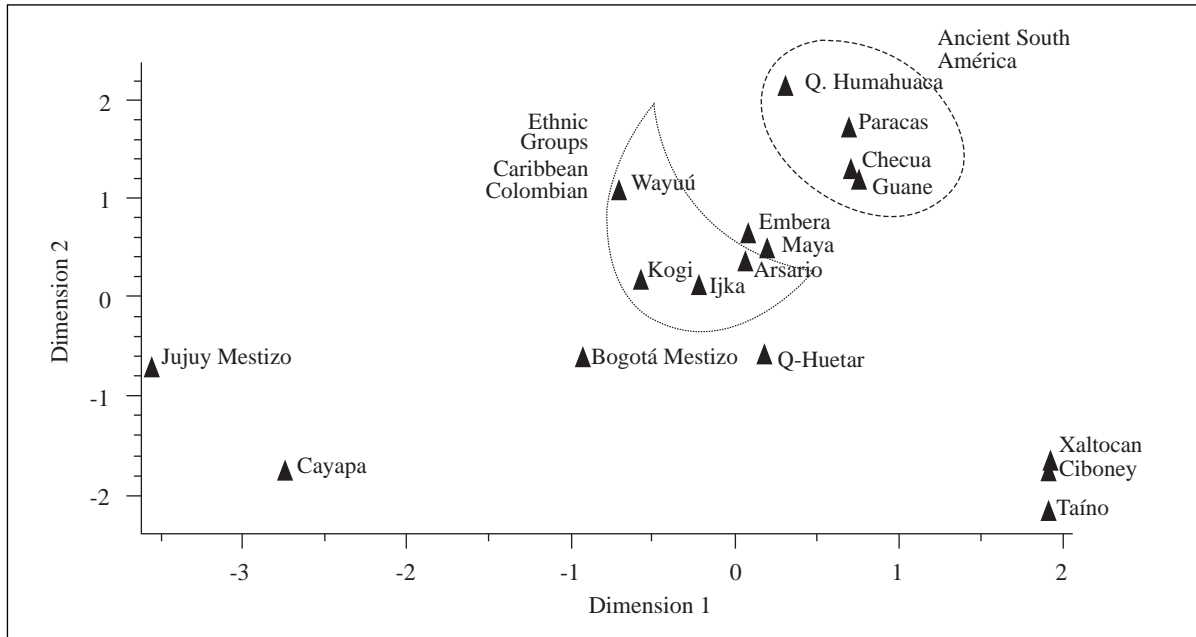


Figure 7. PCA biplot of mtDNA (HVR-I) data in Ancient and Modern Groups - MaxStat Pro v3.60. The mtDNA variation among Paleo-Colombian Checua (This Study) and other ancient and modern groups is shown (Kolman & Bermingham 1997; Rickards, *et al.* 1999; Lalueza-Fox, *et al.* 2001, 2003; Melton, *et al.* 2007; Díaz-Matallana & Martínez-Cruzado 2010; Fehren-Schmitz, *et al.* 2010; Casas-Vargas, *et al.* 2011; Mata-Míguez, *et al.* 2012; Cardoso, *et al.* 2013; Melton, *et al.* 2013; Mendisco, *et al.* 2014). Checua is related to other ancient and modern South American groups; also it is close to Emberá and Maya Amerindians from Central and Mesoamerica, respectively [See Suppl. Tables 5-7, and Suppl. Figs. 5-7].

study obtained positive results from 30 out of 55 analyzed samples (Table 1), a success rate of 54.54% (Manríquez, *et al.* 2011). Lower success rates were achieved in a RFLP study of the Arroyo Seco 2 site, Argentina (Table 1), with 8 out of 23 individuals (34.7%) [Figueiro & Sans 2007]; as well as in an analysis of the Peruvian Conchopata population [Wari: AD 600-800, post-Wari: AD 1100-1400], with 16 out of 49 individuals (33%) [Kemp, *et al.* 2009]. Haplogroups observed in Checua remains, listed from highest to lowest frequency, were: A2 (N = 7), B2 (N = 3), C1 (N = 1) and D4h3a (N = 1) (see Table 2). The X2a haplogroup was not detected in Checua samples, which is not surprising since X2a is restricted to northern North America and Europe (Perego, *et al.* 2009, Rasmussen, *et al.* 2015). The observed haplogroups can be distinguished according to human occupation zones [I = Stratum 4, and II = Stratum 5 (Tables 1, 2)]. Indeed, in the earliest Zone I haplogroups, A2 (N = 4), B2 (N = 2), C1 (N = 1) and D4h3a (N = 1) were detected, while in Zone II only the predominant haplogroups A2 (N = 3) and B2 (N = 1) were detected. Native American haplogroups found in the earliest Checua Zone I, A2, B2, C1 and D4h3a, are considered founding New World lineages (Tamm, *et al.* 2007), and are simultaneously present in this particular site. Previously, the prevalence of lineage A into the New World supported the one migration hypothesis, and this lineage varied independently of any associated linguistic group (Merriwether, *et al.* 1995). Likewise, a

Panamanian mtDNA study (N=1565) performed by Perego, *et al.* (2012) found that 83% clustered into pan-American lineages, most commonly represented by haplogroup A2 (51%). A2 subclades confirmed the key role of the Pacific coastal path in the peopling of the Americas and support the Native American continuity of mtDNA genomes in Panama prior to 10 ka. Similarly, another mtDNA analysis from Nicaragua and Costa Rica (N=131) with 94.6% assigned to the Amerindian A2 and B2 haplogroups (Melton, *et al.* 2013). In line with mentioned studies, for Paleo-Colombian Checua (N=12), A2 and B2 were the most frequent haplogroups (83.3%), while A2 was prevalent with 58.3% (see first paragraph: pag. 16).

The presence of mtDNA haplogroup D4h3a in the early Holocene human remains from Checua in this study provides new evidence in supporting Paleo-American migrations along the Pacific coastal route proposed in previous works (Kemp, *et al.* 2007; Tamm, *et al.* 2007; Wang, *et al.* 2007; Achilli, *et al.* 2008; Dillehay, *et al.* 2008; Fagundes, *et al.* 2008, O'Rourke 2009; Bodner, *et al.* 2012; de Saint-Pierre, *et al.* 2012; Dillehay, *et al.* 2012; Rasmussen, *et al.* 2014; Llamas, *et al.* 2016). In ancient Chilean populations (Manríquez, *et al.* 2011) were identified the founder haplogroups A, B, C and D sustaining one wave model. Perego, *et al.* (2009), based on a mtDNA analysis proposed a Pacific coastal route (lineages A, B, C and D) from Siberia to South America around 20-15 kya, followed by a second

migration (lineage X) in North America once the ice-free corridor appeared; which is compatible with the genomic analysis based on the geographical distribution of two rare lineages: D4h3 (Pacific of North and South America), and X2a (northern North America). Three distinct haplotypes of A2 were detected in seven Paleo-Colombian Checua; two different haplotypes of B2 were evident in three individuals; while C1 and D4h3a were represented by one haplotype each. The presence of the same haplotype A2 in five individuals (three in Occupation I, and two in Occupation II), suggests that the same maternal line could have been present at this site during 400-600 years approximately, according to radiocarbon / calibrated dates of each stratum (Tables 1, 2), this means that Checua occupants of Zone II (A2) could be descendants from the Zone I, associated to their pairwise genetic distances (0.000) [Suppl. Table 1S]. However, in future studies the inclusion of radiocarbon dating for each individual, the analysis of HVR-II, as well as whole mtDNA genomes of these five A2 individuals might discriminate better among these individuals sharing the same A2 haplotype (Table 2; main node A2 in Figure 3).

Paleo-DNA vs. Neo-DNA.

Figure 2 (a-j), show the distribution of Native American mtDNA haplogroups found in the Paleo-American Checua in contrast to those found in other modern DNA studies in Colombia. These demonstrate a genetic continuity of the Native American haplogroups A2, B2 and C1, and probably a genetic discontinuity of sub-haplogroup D4h3a (detected in Checua) in Colombia throughout time, but also could be a shortage of publications and/or records in databases of D4h3a from Colombia, given that the most common haplogroup in modern Colombia is D1 [Melton, *et al.* 2007 N=110; Salas, *et al.* 2008 N=185; Díaz-Matallana & Martínez-Cruzado 2010 N=51 Col.; Yunis & Yunis 2013 N=1,526; Xavier, *et al.* 2015 N=154 modern total N=2,026; see Figure 2 (a-j)]. After searching in the mtDNA database EMPOP v3/R11 (www.empop.org) [Parson & Dür 2007] for the literal D4h3a haplotype (HVR-I) obtained in this study, in the results of the Americas is reported a frequency of 18 / 34,617 records in this database (HVR-I: 16024-16365) as follows: California-USA (2), Ohio-USA (1), Chile (12), Argentina (3). Therefore most of the D4h3a records come from Chile and Argentina, which is consistent with Perego, *et al.* (2009), de Saint-Pierre, *et al.* (2012), and Cardoso, *et al.* (2013) [See Figure 5]. Future ancient DNA studies of Colombian populations might possibly find more evidence of D4h3a.

Currently, similar haplogroup frequencies to those found in the Checua remains (Figure 2a) were observed in several studies from northern South America (Figures 2b – f, h). Current mtDNA analysis on the ethnic groups of Caribbean Colombia showed haplogroups A2, B2, C1 (not D1) (Figure 2g). Nonetheless, Mestizos from the Caribbean, Colombia (Figure 2h) showed haplogroups A2, B2, C1, D1. A study on pre-Columbian Guane, Santander-

Colombia (N=17) (Casas-Vargas, *et al.* 2011), found haplogroups A2, B2, D1 (not C1). In contrast, in the Paleo-Colombian Checua with 12 of 22 individuals analyzed, the Native American haplogroups A2, B2, C1, and D4h3a were identified (This study).

Whilst, haplogroup frequencies of populations from the Colombia's Pacific Coast (Figure 2i and 2j) are dissimilar from those found in the samples from Checua, as well from other mentioned studies, such populations contain the haplogroups A2, B2, C1 and D1 (A2 and C1 the most frequent) in this zone. Such distribution is shared with the ethnic groups described for Colombia's Atlantic coast (Figure 2g), Venezuela (Vona, *et al.* 2005; Díaz-Matallana & Martínez-Cruzado 2010), and Puerto Rico [Martínez-Cruzado, *et al.* 2005; Díaz-Matallana & Martínez-Cruzado 2010; Martínez-Cruzado 2010]. Presently, the D4h3a clade is more frequent in South America, especially in the Southern Cone (O'Rourke 2009; de Saint-Pierre, *et al.* 2012, Rasmussen, *et al.* 2014). A Colombian Mestizo exhibited the D4h3a haplotype, though it was simply identified as haplogroup D (Bryc, *et al.* 2010). In the Caribbean, D4h3a has so far not been reported [Lalueza-Fox, *et al.* 2001, 2003 (N=19, N=47); Martínez-Cruzado 2010 (N=122); Vilar, *et al.* 2014 (N=326), Caribbean total N=514].

D4h3a is present at 1% (approximately) of current Native American populations, but has been mostly observed in ancient Native Americans (≥ 8 ^{14}C ka BP) from mtDNA sequence data (Kemp, *et al.* 2007; Perego, *et al.* 2009; Rasmussen, *et al.* 2014). Overall, haplogroup frequencies obtained in this work are consistent with other studies (Kolman & Bermingham 1997; Mesa, *et al.* 2000, Keyeux, *et al.* 2002; Rodas, *et al.* 2003; Lewis, *et al.* 2007, Watkins, *et al.* 2012, Usme-Romero, *et al.* 2013). Although, there is some variation in the haplogroup frequencies or estimates by population, often associated with geographic location, language or cultural affiliation (O'Rourke, *et al.* 2000, Emery, *et al.* 2015). Native American haplogroups A2, B2, C1, and D1 are present in the current Colombian territory. However, haplogroup C4c was also detected in modern Caribbean Colombia: Ijka 72 (Tamm, *et al.* 2007; Malhi, *et al.* 2010).

Previous modern DNA studies of Colombia based on RFLPs (Mesa, *et al.* 2000, Keyeux, *et al.* 2002; Rodas, *et al.* 2003; Rondón, *et al.* 2007; Usme-Romero, *et al.* 2013) are limited, as they could not detect point mutations from mtDNA control region for A2, B2, C1 and D1 haplogroups nor discriminate sub-haplogroups. Previous ancient DNA studies from Brazil and Colombia (Ribeiro-Dos-Santos, *et al.* 1996; Monsalve, *et al.* 1996; Fernández 1999), such as: Ribeiro-Dos-Santos, *et al.* (1996) analyzed Brazilian samples (Table 1), obtaining haplogroups A and D. Fernández (1999), analyzed a single Checua sample (T9: Stratum 4), and classified as haplogroup C; also, three Aguazuque samples (Table 1), amongst others, were determined as B. Monsalve, *et al.* (1996), studied six

Muisca mummies (AD 470-1849) using RFLPs, detecting major haplogroups A, C, D. Most of these studies, however, predate the emergence of recognized authenticity criteria for aDNA (**Cooper & Poinar** 2000). Though, it is striking that T9 sample was identified as C (**Fernández** 1999), and CHI-08 sample in this study was determined as C1, suggesting that the earlier result may have been valid.

Phylogenetic Networks for Haplogroups A2, B2, C1 and D4h3a.

Network A2: Figure 3 was constructed using 421 American mtDNAs. Haplotype 1 is the central node defined by mutations 16111, 16223, 16290, 16319 and 16362. Hap.1 consists of 274 individuals, including Paleo-Colombian Checua (This study), pre-Columbian Peruvians (**Fehren-Schmitz, et al.** 2010), Chilean Native Americans (**de Saint-Pierre, et al.** 2012), admixed Pacific Coast-Colombia, as well as on the whole America (**Salas, et al.** 2007), Colombian and Venezuelan Mestizos (**Díaz-Matallana & Martínez-Cruzado** 2010), and also constitute the central node A2 (HVR-I) of Puerto Rico (**Martínez-Cruzado** 2010). This supports the notion that Hap.1-A2 is widely distributed in the Americas, also indicates a genetic continuity of haplogroup and haplotypes from the Paleo-American and pre-Columbian periods to current indigenous and Mestizo populations. Haplotype 2 is found in a Paleo-Colombian Checua (CHII-03), also in Wayuú (**Melton, et al.** 2007), Hap.2 is enclosed at bottom-left. Haplotype 5 is represented by a Paleo-Colombian Checua (CHI-13). In general, a phylo-geographical sense can be observed in the A2 network; also its star-like type suggests a remarkable demographic expansion.

Network C1: Figure 4 was assembled with 384 American mtDNAs. Haplotype 1 is the central node with mutations 16223, 16298, 16325 and 16327, composed of 348 individuals, including: Wayuú (**Melton, et al.** 2007), Paleo-Colombian Checua (CHI-08), Colombians (**Bryc, et al.** 2010), admixed Pacific Coast-Colombia, as well as on the whole America (**Salas, et al.** 2007), plus pre-Columbian Peruvians (**Fehren-Schmitz, et al.** 2010). Hap.1 is part of the main node of the C1 lineage of Puerto Rico (**Martínez-Cruzado** 2010). A South American link has also been found with a Venezuelan Mestizo (Caracas) [whole mtDNA genome VE6: C1b, a lineage found so far mostly in South America, but not in Central America], and shared with the ancient CII lineage of Puerto Rico, it was estimated that it arrived in Puerto Rico around to 2800±1600 or 2731±1931 yr BP (**Díaz-Matallana & Martínez-Cruzado; Martínez-Cruzado** 2010). Recently an Amerindian South American origin for the Caribbean was confirmed (**Moreno-Estrada, et al.** 2013). The C1b lineage has been reported in another Venezuelan (**Martínez-Cruzado** 2010), in Puerto Rico (**Vilar, et al.** 2014), and in ancient Argentina (**Mendisico, et al.** 2014). This evidence shows that Hap. 1-C1 is widely distributed in the Americas, and it shows a genetic continuity both at haplogroup and haplotype level, as it has been found

in Paleo-Americans, pre-Columbians and Mestizos. A particular phylo-geographical sense can be observed in the C1 network combined with its star-like shape indicative of demographic expansion.

Network D4h3a: Figure 5 was constructed with 31 American mtDNAs. Haplotype 1 is the main node, characterized by mutations 16223, 16241, 16301, 16342 and 16362, found in six individuals: a Mexican (**Kemp, et al.** 2007), a Paleo-Colombian Checua (this study: CHI-02); Native Chileans (**de Saint-Pierre, et al.** 2012). Haplotype 2 contains Paleo-Americans: ANC939 (**Cui, et al.** 2013), and Anzick-1 (**Rasmussen, et al.** 2014). Haplotype 3 is the Paleo-American OYKC (**Kemp, et al.** 2007). Haplotypes 1 to 8, belong to D4h3a lineage, delimited at the top. D4h3a also been reported in 22% Cayapa-Ecuador (**Rickards, et al.** 1999; **Malhi, et al.** 2010), in Brazilian-Mestizo (**Kemp, et al.** 2007), in Quechua-Perú (**Fuselli, et al.** 2003), and in modern California, USA (**Perego, et al.** 2009). D4h3 originated in eastern China, subsequently its distribution through the Americas, mostly in South America (**Perego, et al.** 2009, **Rasmussen, et al.** 2014). To date, D4h3a has not been reported in the Caribbean, only D1 has been detected [**Lalueza-Fox, et al.** 2001, 2003 (N=19, N=47); **Martínez-Cruzado** 2010 (N=122); **Vilar, et al.** 2014 (N=326), the Caribbean total N=514]. In Colombia, D4h3a was detected in a Paleo-Colombian Checua (this study, Hap.1) and in a Mestizo: Haplotype 5, nevertheless it was simply classified as haplogroup D (**Bryc, et al.** 2010). After searching in the mtDNA database EMPop v3/R11 (www.empop.org) [Parson & Dür 2007] for the literal D4h3a haplotype (HVR-I) obtained in this study, in the results of the Americas is reported a frequency of 18 / 34617 records in this database (HVR-I: 16024-16365) as follows: California-USA (2), Ohio-USA (1), Chile (12), Argentina (3). Thus most of the D4h3a records come from Chile and Argentina, which is consistent with **Perego, et al.** (2009), **de Saint-Pierre, et al.** (2012), and **Cardoso, et al.** (2013) [See Figure 5].

Haplotypes 9 to 13 belong to D4h3a5 lineage (N=17), demarcated at the bottom, including: several Natives from Chile and Argentina (**de Saint-Pierre, et al.** 2012). D4h3a and D4h3a5 lineages are fully distinguished, D4h3a5 exhibits np 16051 and is restricted to the Southern Cone (**de Saint-Pierre, et al.** 2012; **Rasmussen, et al.** 2014). Furthermore, it has a pattern of star-like phylogeny.

Network B2: was assembled from 138 American mtDNAs. Haplotype 2 is the central node, with mutations 16183, 16189 and 16217, and consists of 115 individuals, including: Paleo-Colombian Checua (CHI-01 and CHII-06), a pre-Columbian mummy, Santander-Colombia (**Díaz-Matallana, et al.** 2010; **Díaz-Matallana** 2015), pre-Columbian Peruvians (**Fehren-Schmitz, et al.** 2010), admixed Pacific Coast-Colombia, as well as overall America (**Salas, et al.** 2010). This shows that Hap.2 is widely distributed in the Americas, also indicates a genetic continuity seen in Paleo-Americans, pre-Columbians, and contemporary Mestizos. Haplotype 3, found in a Paleo-

Colombian Checua (CHI-07). Haplotype 1 is found in 20 individuals: Ijka and Wayuú (Melton, *et al.* 2007), Colombian Mestizos (Díaz-Matallana & Martínez-Cruzado 2010), also a pre-Columbian Peruvian (Fehren-Schmitz, *et al.* 2010). Haplotypes 3 and 9, found in a Paleo-Colombian Checua (This study) and in a Colombian Mestizo - Pacific Coast (Salas, *et al.* 2007); they are grouped on the right. A phylo-geographical sense can be observed in the B2 network combined with its star-like phylogeny diagnostic of population expansion (Suppl. Figure 1).

Genetic and Archaeological Evidence of Early Migrations.

Figure 6 (a, b) shows a spatio-temporal distribution of Paleo-Americans, providing information on migrations throughout America based on studies of mtDNA haplogroups A2, B2, C1 and D4h3a from different sites. Results obtained from Paleo-Colombian Checua (Díaz-Matallana, *et al.* 2014; Díaz-Matallana 2015; This study), ancient Chile (Manríquez, *et al.* 2011; Faltyskova, *et al.* 2014, de la Fuente, *et al.* 2014, Llamas, *et al.* 2016), central Andes (Perú, Bolivia, México and Argentinian Pampas) [Llamas, *et al.* 2016], provide the most up-to-date data with implications for the early peopling of South America. Additionally, recent findings from America, such as NAIA (Chatters, *et al.* 2014), ANC939 (Cui, *et al.* 2013), Anzick-1 (Rasmussen, *et al.* 2014), among others are in Figure 6. Along with recent findings such as Huaca Prieta, Perú (Dillehay, *et al.* 2012), Monte Verde-Chile dated to 14,600-14,200 cal yr BP (Dillehay, *et al.* 2008) [Table 1], has become a key reference with regards to the American colonization along the Pacific coast (Tamm, *et al.* 2007; Fagundes, *et al.* 2008; Perego, *et al.* 2009; Bodner, *et al.* 2012; Llamas, *et al.* 2016). Previously, a Pacific coastal route was proposed as the initial entry to America on linguistic and archaeological grounds (Gruhn 1994). Figure 6a shows the distribution of D4h3a and D1 Native American haplogroups, providing evidence for early migrations extending along the Pacific coast from Alaska to Tierra del Fuego, corroborating the spatial distribution of whole mtDNA lineages described by Perego, *et al.* (2009), in which haplogroup D4h3a likewise suggests a Pacific coastal route. D4h3a is uncommon in North America, being most common in South America (de Saint-Pierre, *et al.* 2012, Cardoso, *et al.* 2013; Rasmussen, *et al.* 2014), whilst this work provides the first instance from Colombia. Figure 6b, shows a combined distribution of the A2, B2 and C1 Native American haplogroups, and also suggests an early migration into South America along the Pacific coast from Oregon (USA) to southern Chile. Taken together, these distributions correlate with the roles proposed for mtDNA haplogroups A2, B2, C1 and D4h3a as founder lineages of the New World (Tamm, *et al.* 2007; Achilli, *et al.* 2008). The mtDNA distribution reported by Perego, *et al.* (2009) confirms a strong geographical structure of mtDNA and autosomal diversity in the Americas (O'Rourke and Raff 2010).

The paucity of early data from the migration route into South America along Colombia's Pacific coast (1300km length) is not surprising, this region has been characterized by extremely complex geological dynamics (tectonic plates movement causing significant seismic activity), dramatic changes in sea level, and elevated environmental humidity since the Holocene throughout the Darién Gap and Chocó, Colombia. These conditions are very different from those in Ecuador, Perú and Chile (Stohtert 1985; Dillehay, *et al.* 2008, 2012). Moreover, the socio-cultural conflict in this area has discouraged archaeological research, contributing to the scarcity of preceramic remains (Van der Hammen 2006; López-Castaño & Cano-Echeverry 2011). Similarly, no conclusive evidence of human activity from the Pleistocene-Holocene transition has yet been found along Colombia's Atlantic coast (Correal 1990, Aceituno, *et al.* 2013).

On the contrary, zones situated between the Colombia's mountain ranges, as well within the mid-and high-altitude Andean ecosystems, extensive plateaus, have revealed significant densities of preceramic populations (Table 1). Many findings have been reported in the "Altiplano Cundiboyacense" and Savannah of Bogotá (Aceituno, *et al.* 2013), the second (2600 m.a.s.l., Figure 1), offers ideal conditions for the preservation of ancient evidence (Correal 1990). The wide corridor of Colombia's Magdalena River valley is an area of outstanding potential for investigating the settlement of early inhabitants, along its length have revealed different stages of human occupation (Dillehay 1999; López-Castaño & Cano-Echeverry 2011; Aceituno, *et al.* 2013). It is not unlikely that human evidence predating 12,5 ¹⁴C ka BP will eventually be found in Colombia (the earliest site is Pubenza, without human remains).

Anthropological analysis of three Paleo-American skulls from the Argentine Pampas (Table 1), concluded that these skulls are morphologically alike to others of the Early Holocene (Lagoa Santa-Brazil; Tequendama, Checua and Aguazuque-Colombia; Lauricocha-Perú; and Mexico). Archaeological evidence suggests that fishery resources were able to support relatively large human populations along the Pacific coasts of the Americas even before the development of agriculture (Melton, *et al.* 2007; Dillehay, *et al.* 2012). The prehistoric hunting tool "boleadora" found in Checua, previously detected about at Monte Verde (12,5 ¹⁴C Ka BP), in eastern Brazil (11,5 ¹⁴C Ka BP), and in the Southern Cone (Minelli, *et al.* 2012), provides further evidence for early American settlers.

Based on the combined evidence of the present study's Paleo-DNA data, along with findings from several genetic and archaeological studies, computer simulation models, and paleo-environmental studies mentioned here, we propose an *Integrated Migration Model for Northwestern South America*. This model consists of: (I) a North-to-South route along the Pacific coast; (II) Sequential movements across the Andes from West-to-East (the evidence for early human activity in the Colombian Andean region is concentrated in

the easternmost of the three north-south mountain ranges); (III) The discovery and utilization of the regions river systems (including the Magdalena, Bogotá, Checua, and Neusa rivers) as resources for food and travel (Correal 1990,1993; Gruhn 1994; Dillehay 1999; Keyeux, et al. 2002; Fix 2005; Dillehay, et al. 2008, 2012; Perego, et al. 2009; López-Castaño & Cano-Echeverry 2011; Bodner, et al. 2012; de Saint-Pierre, et al. 2012; Minelli, et al. 2012; Aceituno, et al. 2013). Further archaeological, genetic or genomic studies should permit the refinement of this model.

Estimates of Genetic Diversity

Regarding Tajima's neutrality test for the Checua remains (Table 3), the value of segregating sites observed where $S > \pi$ is indicative of either a population expansion or positive selection, both for the 12 individuals, and for haplogroups A2 and B2. These observations are related to the negative value of the Tajima's D ; although statistically not significant D values were detected ($p > 0.10$), population expansion is more noticeable for A2 (-1.358415) and is also evident in comparison with total Tajima's D (-0.634978), explained by the higher frequency of A2, identified in 7 of 12 individuals. According to Fu's F_s statistic, the only negative value was observed in the haplogroup A2 (-0.237) which agrees with Tajima's D value, thus it is indicative of population expansion, although statistically not significant ($p > 0.02$). Whereas F_s -values for full Checua group (0.169) and for haplogroup B2 (1.061) were positive, these values suggest population differentiation, as would be expect from a recent population bottleneck or from over dominant selection. Likewise, previous studies on Central and northern South American populations (Chibchan-speaking) were shown to have positive F_s -values (Melton, et al. 2013).

Concerning mtDNA Haplotype diversity (H_d) measures (Table 4), full Paleo-Colombian Checua showed an H_d value = 0.833, followed by the haplogroups B2 and A2 values (0.667 and 0.524). Other Central and northern South American populations exhibited similar H_d values, such as: Emberá (0.942), Maya (0.930), Zapatón-Huetar (0.806), Quitirrisí-Huetar (0.709), Guane (0.875), Wayuú (0.788), Arsario (0.725) [Kolman & Bermingham 1997; Casas-Vasgas, et al. 2011; Melton, et al. 2007, 2013].

By comparing mtDNA nucleotide diversity (π_T) in Native Americans (Table 5), the closest π_T values to those estimated for the *Paleo-Colombian Checua* (0.016) (This Work); were: Shuswap-speaking territory (0.0156); Nuuchah-Nulth (0.016); Yakima (0.015); Oneota, pre-Columbian (0.015); pre-Conquest (0.017) and post-Conquest Xaltocan (0.018); Maya (0.017); Zapatón-Huetar (0.017); Emberá (0.017); Wayuú (0.016); Cayapa (0.018); pre-Columbian Palpa (0.0174); Pacapaccari (0.0169); Ancash (0.0162); Tayacaja (0.0166); and Mapuche (0.017). Overall, these closest π_T values reflect a molecular trend between North America (Canada, USA), Mesoamerica (ancient Mexico, Maya), Central America (Panamá, Costa Rica), Caribbean Colombia (La Guajira), and South America (Colombia,

Ecuador, ancient / modern Perú, and Argentina). Such shared molecular trend is quite consistent with other investigations based on different molecular markers, as follows: *genetic and genomic mtDNAs* (Melton, et al. 2007; Tamm, et al. 2007; Fagundes, et al. 2008; Díaz-Matallana & Martínez-Cruzado 2010, Yang, et al. 2010, Watkins, et al. 2012; Llamas, et al. 2016), *autosomal-STRs* (Wang, et al. 2007), *chromosome X-STRs* (Yang, et al. 2010), and *nuclear-SNPs microarrays or Ancestry Informative Markers-AIMs* (Bryc, et al. 2010; Reich, et al. 2012, Watkins, et al. 2012). In summary, these studies highlight the importance of the coast and the Pacific coast on the early American colonization (Tamm, et al. 2007; Wang, et al. 2007; Fagundes, et al. 2008; Yang, et al. 2010; Reich, et al. 2012; Llamas, et al. 2016), as well as a dynamic bidirectional gene flow between Central and South America in prehistoric times (Melton, et al. 2007; Díaz-Matallana & Martínez-Cruzado 2010; Yang, et al. 2010; Reich, et al. 2012; Melton, et al. 2013; Moreno-Estrada, et al. 2013).

There are marked similarities with regards to certain haplotypes of A and C [Figure 3: Hap2 and Figure 4: Hap1], as well as close π_T values, between the Paleo-Colombian Checua (this study) and the modern Wayuú (Melton, et al. 2007). The Wayuú speak an Arawakan language (Aikhenvald 1999), and has been assigned to an Equatorial-Tucanoan linguistic stock (Ruhlen 1991), they are usually considered distinct from Chibchan-speaking populations (e.g. Muisca) [Melton, et al. 2007]. However, the Checua remains also have π_T values similar to those of the Panamanian Emberá, who speak a Chocoan language (Aguirre-Licht 2006). Another study suggested that the Chocoan and Chibchan families may be related (Constenla-Umaña 2005), which would be compatible with Ruhlen's assignment (Ruhlen 1991) of the Emberá language to the Chibchan-Paezan linguistic stock.

The PCA biplot allows visualizing biological relationships among ancient and modern population groups (Figure 7); it was possible to identify clearly: (a) an ancient South American Andes cluster (Guane, Checua, Palpa, and Q.Humahuaca) in the top right ellipse of the plot. (b) The contemporary northern South American cluster (Arsario, Ijka, Kogi, and Wayuú) is demarcated in a semi-circle. (c) In the middle of the ancient cluster and the modern cluster pointed out, there is a mixture of Meso / Central American (Maya, Emberá) and South American populations, such depiction have been observed before (Melton, et al. 2007, 2013). This mixed distribution coincides with other genetic diversity indices: Tajima's D and F -values (e.g. population expansion of A2), H_d and π_T values, also it is related to phylogenetics networks; therefore it reiterates an active interaction between Central and South America since prehistoric times. (d) Mestizo Bogotá is in the bottom of the modern cluster close to Q-Huetar, these share similar scores and π_T value. (e) Ciboney is nearer to Xaltocan (full) than Taíno in the lower right of the plot; this relation is in

agreement with a recent study (Moreno-Estrada, *et al.* 2013). Overall distribution of populations in the PCA was in line with their geographic locations [Suppl. Tables 5-7, and Suppl. Figures 5-7].

Conversely, the π_T value of Paleo-Colombian Checua (=0.016) [This work] appears some what lower than that of the pre-Columbian Guane (=0.0231) [Casas-Vargas, *et al.* 2011], although this variation may be due to the distinct HVR-I coverage ranges used in each study, for Guane was np 16217-16362, with most polymorphisms in this region (A, B and D haplogroups). In contrast, the range for Checua is np 16011-16382, with improved chances of measuring the π_T value; the same issue occurs for the Hd measure. Nonetheless, the PCA biplot yielded similar scores between Checua and Guane; without forgetting that both ancient groups inhabited the eastern Andes, Colombia (Figures 7).

Conclusion

For the first time D4h3a was reported in ancient Colombia (This study: Paleo-Colombian Checua N=12/22), and not detected in previous aDNA studies such as Muisca period (Monsalve, *et al.* 1996 N=6 RFLP), Herrera period (Silva, *et al.* 2008 N= 6/11), and pre-Columbian Guane (Casas-Vargas, *et al.* 2011 N=17). There is a genetic continuity of the A2, B2, C1 haplogroups, and presumably a discontinuity of D4h3a in Colombia since prehistoric times, but also could be a shortage of publications and/or records in databases of D4h3a from Colombia, while D1 (HVR-1) is more common in modern populations [Melton, *et al.* 2007 N=110; Salas, *et al.* 2008 N=185; Díaz-Matallana & Martínez-Cruzado 2010 N=51 Col.; Yunis & Yunis 2013 N=1,526; Xavier, *et al.* 2015 N=154 modern total N=2,026; see Figure 2 (a-j)]. It seems not unlikely that future ancient DNA studies of Colombian human remains may find more evidence of D4h3a. D4h3a was also present in ancient populations from the Southern Cone, while A2, B2, C1, D1 and D4h3a haplogroups are also found in current populations. D4h3a mtDNA haplogroup in early Holocene human remains from South America, as shown in this study, supporting early migrations into South America along the Pacific Coast, as does the distribution of A2, B2, and C1 haplogroups in Paleo-Americans. Based on the combined evidence of the present study's ancient DNA data, along with findings from other genetic and archaeological studies, computer simulation models, and paleo-environmental studies previously

mentioned, we propose an *Integrated Migration Model for Northwestern South America*. This model consists of: (1) a North-to-South route along the Pacific coast; (2) sequential movements across the Andes from West-to-East (the early human evidence in Colombia is concentrated in the eastern Andean range); (3) the Riverine finding and colonization next to the river. We found an unexpected close relationship between a Paleo-Colombian Checua (eastern highlands Colombia: This work) and the Wayúú (La Guajira Peninsula, Colombia), in terms of certain haplotypes A and C (Figures 3, 4), with equivalent total nucleotide diversity. Similarities in nucleotide diversity values were also found between Paleo-Colombian Checua and the Emberá (Panamá). Shared values of total mtDNA nucleotide diversity among Paleo-Colombian Checua, pre-Columbians and modern Natives (Table 5), other diversity indices cited, the PCA clusters observed, in conjunction with the results of various other studies on distinct molecular markers, indicate a dynamic two-way gene flow between Central and South America in ancient times. It may be hoped –or even expected–that new archaeological data will provide further clues that help establish a better understanding of when and how humans moved across the Americas.

It would be advisable in future studies related to Checua samples, the incorporation of radiocarbon dating for each individual, the analysis of HVR-II, as well as whole mtDNA genomes for these five A2 individuals sharing the same haplotype (HVR-I) might discriminate better among these A2 individuals; on the other hand, try to re-process the remaining samples with negative results obtained herein.

Cross-disciplinary research and newer genomic methods applied to aDNA studies (e.g. full control region (HVR-I and HVR-II), whole mtDNAs, nuclear genome sequencing, and next generation sequencing) should reveal a clearer radiography of the human past.

Acknowledgments

Our gratitude to D.A. Bolnick & group and C. Harkey at U.Texas-Austin, for their valuable assistance and recommendations to this work; also to C.H. Langebaek and H. Groot, both from U. de Los Andes-Bogotá, and to C.A. Barragán from U. California-Davis, for their cooperation and helpful discussions to this study. Thanks to C.E. Anderson from U. de La Sabana-Chía, for the review of English, plus the support of authors Institutions.

Supplemental materials

Figure 1S. Phylogenetic Network of Haplogroup B2 (138 ancient and modern mtDNAs HVR-I of the Americas). See Supplemental figure in: <http://www.racefyn.co/index.php/racefyn/article/download/SuppFile/328/1650>

Figure 2S. N-J Phylogenetic Tree of Paleo-Colombian Checua haplogroups. See Supplemental figure in: <http://www.racefyn.co/index.php/racefyn/article/download/SuppFile/328/1652>

Table 1S. Estimates of Pairwise distances between sequences (Checua). See Supplemental table in: <http://www.racefyn.co/index.php/racefyn/article/download/SuppFile/328/1659>

Table 2S. Number of HVR-I Haplotypes in Amerindians from Central, South America, and the Caribbean. See Supplemental table in: <http://www.racefyn.co/index.php/racefyn/article/download/SuppFile/328/1660>

Table 3S. ANOVA: Results. See Supplemental table in: <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1665>

Figure 3S. Plot the Means Group of HVR-I Haplotypes with 95% Confidence Intervals. See Supplemental figure in: <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1653>

Figure 4S. Box Plot for HVR-I haplotypes data. See Supplemental figure in: <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1654>

Table 4S. One-Way ANOVA from Summary Data. See Supplemental table in: <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1661>

Table 5S. Estimated frequencies of HVR-I Amerindian haplogroups in Ancient and Modern Groups. See Supplemental table in: <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1662>

Table 6S. Results PCA. See Supplemental table in: <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1663>

Table 7S. Scores PCA. See Supplemental table in: <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1664>

Figure 5S. PCA biplot: Scores and coefficients. See Supplemental figure in: <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1655>

Figure 6S. Final PCA biplot. See Supplemental figure in: <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1656>

Figure 7S. Comparison of HVR-I frequencies of Amerindian haplogroups in Ancient and Modern Groups. See Supplemental figure in: <http://www.raccefyn.co/index.php/raccefyn/article/downloadSuppFile/328/1658>

[See pdf file attached].

Financial support

COLCIENCIAS full Ph.D. Scholarship (2010-2014) awarded to M.D.M; Ph.D. Grant N° 5176 -Vicerectorate for Research, U. Javeriana (2012-2014); Financial aid Graduate Office - Science Faculty, U. Javeriana (2011); and Grant N° 241-2008 COLCIENCIAS (2010-2011).

Conflicts of interest

The authors declare that they have no conflict of interest.

References

- Aceituno, F.J., Loaiza, N., Delgado-Burbano, M.E., Barrientos, G. 2013. The initial human settlement of Northwest South America during the Pleistocene/Holocene transition: Synthesis and perspectives. *Quat. Int.* **301**: 23-33.
- Achilli, A., Perego, U.A., Bravi, C.M., Coble, M.D., Kong, Q.P., Woodward, S.R., *et al.* 2008. The phylogeny of the four Pan-American mtDNA haplogroups: Implications for evolutionary and disease studies. *PLoS One.* **3** (3): e1764.
- Aguirre-Licht, D. 2006. Choco Languages. In: K. Brown (editor). *Encyclopedia of Language and Linguistics*. Elsevier, Amsterdam. p. 367-381.
- Aikhenvald, A. 1999. Arawak. In: R.M.W. Dixon, A. Aikhenvald (editors). *The Amazonian Languages*. Cambridge University Press. p. 65-106.
- Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowers, R.N., Turnbull, D.M., Howell, N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23** (2): 147.
- Bandelt, H.J., Forster, P., Rohl, A. 1999. Median-Joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16** (1): 37-48.
- Batista, O., Kolman, C., Bermingham, E. 1995. Mitochondrial DNA diversity in the Kuna Amerinds of Panama. *Hum. Mol. Genet.* **40**: 921-929.
- Bodner, M., Perego, U.A., Huber, G., Fendt, L., Röck, A.W., Zimmermann, B., *et al.* 2012. Rapid Coastal spread of First Americans: Novel Insights from South America's Southern Cone mitochondrial genomes. *Genome Res.* **22**: 811-820.
- Bolnick, D.A. 2005. *The genetic prehistory of eastern North America: Evidence from ancient and modern DNA* (Ph.D. Dissertation-Anthropology). University of California-Davis, USA.
- Bolnick DA, Bonine HM, Mata-Míguez J, Kemp BM, Snow MH, Leblanc S. 2012. Nondestructive sampling of human skeletal remains yields ancient nuclear and mitochondrial DNA. *Am. J. Phys. Anthropol.* **147**: 293-300.
- Bryan, A.L., Casamiquela, R.M., Crucent, J.M., Gruhn, R., Ochsenius, C. 1978. An El Jobo Mastodon Kill at Taima-taima, Venezuela. *Science.* **200** (4347): 1275-1277.
- Bryc, K., Véléz, C., Karafet, T., Moreno-Estrada, A., Reynolds, A., Auton, A., *et al.* 2010. Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proc. Natl. Acad. Sci. USA.* **107** (2): 8954-8961.
- Cardoso S, Palencia-Madrid L, Valverde L, Alfonso-Sánchez MA, Gómez-Pérez L, Alfaro E, *et al.* 2013. Mitochondrial DNA control region data reveal high prevalence of Native American lineages in Jujuy province, NW Argentina. *Forensic Sci. Int. Genet.* **7**: e52-e55.
- Carvajal, E., Montes, L., Almanza, O.A. 2014. Datación de restos arqueológicos encontrados en Checua (Cundinamarca-Colombia) mediante Resonancia Paramagnética Electrónica. *Rev. Acad. Col. Cienc.* **38** (147): 124-129.
- Casas-Vargas, A., Gómez, A., Briceño, I., Díaz-Matallana, M., Bernal, J.E., Rodríguez, J.V. 2011. High genetic diversity on a sample of pre-Columbian bone remains from Guane territories in northwestern Colombia. *Am. J. Phys. Anthropol.* **146**: 637-649.
- Catelli, M.I., Álvarez-Iglesias, V., Gómez-Carballa, A., Mosquera-Miguel, A., Romamini, C., Borosky, A., *et al.* 2011. The Impact of modern migrations on present-day multi-ethnic Argentina as recorded on the mitochondrial DNA genome. *BMC Genet.* **12**: 77.
- Chatters, J.C., Kenneth, D.J., Asmeron, Y., Kemp, B.M., Polyak, V., Blank, A.N., *et al.* 2014. Late Pleistocene human skeleton and mtDNA link Paleoamericans and modern Native Americans. *Science.* **344** (6185): 750-754.
- Constenla-Umaña, A. 2005. ¿Existe relación genealógica entre lenguas misumalpas y chibchenses? *Estudios de Lingüística Chibcha.* **24**: 7-85.

- Cooke R, Ranere A, Pearson G, Dickau R.** 2013. Radiocarbon chronology of early human settlement on the Isthmus of Panama (13,000-7,000 BP) with comments on cultural affinities, environments, subsistence, and technological change. *Quat. Int.* **301**: 3-22.
- Cooper, A., Poinar, H.N.** 2000. Ancient DNA: Do it Right or Not at All. *Science.* **289** (5482): 1139.
- Correal, G., Van der Hammen, T., Lerman, L.C.** 1970. Artefactos líticos de abrigos rocosos en el Abra, Colombia. *Rev. Colomb. Antropol.* **14**: 9-53.
- Correal, G.** 1990. Evidencias culturales durante el Pleistoceno y Holoceno de Colombia. *Rev. Arqueol. Americana.* **1**: 33-68.
- Correal, G.** 1993. Nuevas Evidencias culturales Pleistocénicas y Megafauna en Colombia. *Boletín de Arqueología.* **8** (1): 3-12.
- Cui, Y., Lindo, J., Hughes, C.E., Johnson, J.W., Hernandez, A.G., Kemp, B.M., et al.** 2013. Ancient DNA analysis of Mid-Holocene individuals from the Northwest coast of North America reveals different evolutionary paths for mitogenomes. *PLoS One.* **8** (7): e66948.
- de la Fuente, C., Galimany, J., Kemp, B., Judd, K., Reyes, O., Moraga, M.** 2014. *Cazadores recolectores australes: Aproximación a la historia poblacional de grupos canoeros de Patagonia y Tierra del Fuego.* XIII Congreso Asociación Latinoamericana de Antropología Biológica (ALAB). Santiago de Chile.
- de Saint-Pierre, M., Bravi, C.M., Motti, J.M., Fuku, N., Tanaka, M., Llop, E., et al.** 2012. An alternative model for the Early Peopling of Southern South America revealed by analyses of three mitochondrial DNA haplogroups. *PLoS One.* **7** (9): e43486.
- Díaz-Matallana, M.** 2009. *In search of pre-Columbian migrations to the Caribbean: Indigenous Mitochondrial DNA from northern South America* (M.Sc. Thesis - Biology). University of Puerto Rico-Mayaguez, PR. ProQuest. 88p.
- Díaz-Matallana, M., Martínez-Cruzado, J.C.** 2010. Estudios sobre ADN mitocondrial sugieren un linaje predominante en la Cordillera Oriental de Colombia y un vínculo Suramericano para los arcaicos de Puerto Rico. *Universitas Médica.* **51** (3): 241-272.
- Díaz-Matallana, M., Briceño, I., Borda-Camacho, E., Gómez, A., Bernal, J.E., Rodríguez, J.V.** 2010. *Uso de imágenes diagnósticas para el estudio de una momia proveniente de San José de Suaita, Santander (Colombia).* XI Congreso Asociación Latinoamericana de Antropología Biológica (ALAB). Facultad Ciencias Humanas, Universidad Nacional de Colombia, Bogotá.
- Díaz-Matallana, M., Briceño, I., Borda-Camacho, E., Gómez, A., Bernal, J.E., Rodríguez, J.V.** 2014. *Molecular characterization of early Holocene skeletal human remains belonging to Checua culture from Colombia-South America.* 22nd Annual Meeting, Society for Molecular Biology & Evolution (SMBE). San Juan, Puerto Rico (USA).
- Díaz-Matallana, M.** 2015. *Caracterización Genética de un grupo Paleoamericano Checua proveniente de Nemocón-Cundinamarca, Colombia: Implicaciones para el Poblamiento Temprano de Suramérica* (Tesis Ph.D. Ciencias Biológicas). Pontificia Universidad Javeriana. Bogotá. 104 p.
- Dillehay, T.D.** 1999. The Late Pleistocene culture of South America. *Evol. Anthropol.* **7** (6): 206-216.
- Dillehay, T.D., Ramírez, C., Pino, M., Collins, M.B., Rossen, J., Pino-Navarro, J.D.** 2008. Monte Verde: Seaweed, food, medicine, and the peopling of South America. *Science.* **320**: 784-786.
- Dillehay, T.D., Bonavia, D., Goodbred Jr, S.L., Pin, M., Vásquez, V., Rosales-Tham, T.** 2012. A late Pleistocene human presence at Huaca Prieta, Perú, and early Pacific Coastal adaptations. *Quat. Res.* **77**: 418-423.
- Emery, L.S., Magnaye, K.M., Bigham, A.W., Akey, J.M., Bamshad, M.J.** 2015. Estimates of Continental Ancestry vary widely among Individuals with the same mtDNA Haplogroup. *Am. J. Hum. Genet.* **96**: 183-193.
- Eshleman, J.A., Smith, D.G.** 2001. Use of DNase to eliminate contamination in ancient DNA analysis. *Electrophoresis.* **22**: 4316-4319.
- Fagundes, N.J.R., Kanitz, R., Eckert, R., Valls, A.S., Bogo, M.R., Salzano, F.M., et al.** 2008. Mitochondrial population genomics supports a single Pre-Clovis origin with a coastal route for the Peopling of the Americas. *Am. J. Hum. Genet.* **82**: 583-592.
- Faltyskova, Z., et al.** 2014. *Population history of South America: ancient DNA study of extinct people from Tierra del Fuego.* 22nd Annual Meeting, Society for Molecular Biology & Evolution (SMBE). San Juan, Puerto Rico (USA).
- Fehren-Schmitz, L., Reindel, M., Cagigao, E., Hummel, S., Herrmann, B.** 2010. Pre-Columbian population dynamics incoastal Southern Peru: A Diachronic Investigation of mtDNA patterns in the Palpa Region by ancient DNA Analysis. *Am. J. Phys. Anthropol.* **141**: 208-221.
- Fernández, C.** 1999. *La arqueología molecular aplicada a la solución de problemas prehistóricos: análisis de ADNmt en momias y restos óseos prehistóricos* (Tesis pregrado-Antropología). Universidad Nacional de Colombia, Bogotá.
- Figueiro, G., Sans, M.** 2007. Primeros resultados del análisis de ADN Mitocondrial del Sitio Arroyo Seco 2, Buenos Aires, Argentina. *Rev. Arg. Antropol. Biol.* **9** (1): 78.
- Fix, A.G.** 2005. Rapid deployment of the five founding Amerind mtDNA haplogroups via coastal and riverine colonization. *Am. J. Phys. Anthropol.* **128**: 430-436.
- Fontugne, M.** 2013. New Radiocarbon Ages of Luzia woman, Lapa Vermelha IV site, Lagoa Santa, Minas Gerais, Brazil. *Radiocarbon.* **55** (2-3): 1187-1190.
- Fu, Y.X.** 1997. Statistical tests of neutrality of mutations against population growth, hitch-hiking, and background selection. *Genetics.* **147**: 915-925.
- 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. *Mol. Biol. Evol.* **20**: 1682-1691.
- Goebel, T., Waters, M.R., O'Rourke, D.H.** 2008. The Late Pleistocene dispersal of modern humans in the Americas. *Science.* **319**: 1497-1502.
- González, J.R., Neves, W.A., Lahr, M.M., González, S., Pucciarelli, H., Hernández-Martínez, M., et al.** 2005. Late Pleistocene/Holocene craniofacial morphology in Mesoamerican Paleoindians: Implications for the peopling of the New World. *Am. J. Phys. Anthropol.* **128**: 772-780.
- Greenberg, J.H., Turner, C.G. II, Zegura, S.L.** 1986. The settlement of the Americas: a comparison of the linguistic, dental and genetic evidence. *Curr. Anthropol.* **27**: 477-497.

- Groot, A.M.** 1992. Checua: Una secuencia cultural entre 8,500 y 3,000 años Antes del Presente. Bogotá: Fundación Investigaciones Arqueológicas, Banco de La República; 100p.
- Groot, A.M.** 1995. Checua: Un aporte para el conocimiento del precerámico de la sabana de Bogotá. En: I. Cavellier, S. Mora (editores). *Ámbito y Ocupaciones Tempranas de la América Tropical*. Instituto Colombiano de Antropología, Bogotá.
- Groot, A.M.** 2006. Arqueología y Patrimonio: Conocimiento y Apropiación Social. *Rev. Acad. Col. Cienc.* **30** (114): 5-17.
- Gruhn, R.** 1994. The Pacific Coast route of initial entry: An overview. In: R. Bonnichsen, D.G. Steele (editors). *Method and theory for investigating the peopling of the Americas*. Oregon State University, Corvallis, USA. p. 34-44.
- Jara, N.P.** 2007. *Determinación de la estructura genética en un grupo poblacional Muisca mediante el análisis de polimorfismos en el ADNmt* (Tesis Maestría-Biología). Facultad Ciencias, Pontificia Universidad Javeriana, Bogotá.
- Jenkins, D.L., Davis, L.G., Stafford, T.W. Jr., Campos, P.L., Conolly, T.J., Cummings, L.S., et al.** 2008. Geochronology, Archaeological Context, and DNA at the Paisley Caves. In: K.E. Graf, C.V. Ketron, M.R. Waters (editors). *The Paleoamerican Odyssey*. College Station, Center for the Study of First Americans, USA. p. 485-520.
- Johnson, J.R., Stafford, T.W. Jr., Ajie, H.O., Morris, D.P.** 2002. Arlinton Springs Revisited. In: D.R. Brooks, K.C. Mitchell, H.W. Channey (editors). *Santa Barbara Museum of Natural History, California, USA*. p. 541-545.
- Kemp, B.M., Smith, D.G.** 2005. Use of bleach to eliminate contaminating DNA from the surfaces of bones and teeth. *Forensic Sci. Int.* **54**: 53-61.
- Kemp, B.M., Malhi, R.S., McDonough, J., Bolnick, D.A., Eshleman, J.A., Rickards, O., et al.** 2007. Genetic analysis of Early Holocene skeletal remains from Alaska and its Implications for the settlement of the Americas. *Am. J. Phys. Anthropol.* **132**: 605-621.
- Kemp, B.M., Tung, T.A., Summar, M.L.** 2009. Genetic continuity after the collapse of the Wari Empire: mitochondrial DNA profiles from Wari and post-Wari populations in the ancient Andes. *Am. J. Phys. Anthropol.* **140**: 80-91.
- Keyeux, G., Rodas, C., Gelvez, N., Carter, D.** 2002. Possible migration routes into South America deduced from mitochondrial DNA studies in Colombian Amerindian populations. *Hum. Biol.* **74** (2): 211-233.
- Kirkman, T.W.** 1996. Statistics to Use, ANOVA: Analysis of Variance between groups. Date Query: 21-01-2015. Available in: <http://www.physics.csbsju.edu/stats/>
- Knapp, M., Clarke, A.C., Horsburgh, K.A., Matisoo-Smith, E.A.** 2012. Setting the stage: Building and working in an ancient DNA Lab. *Ann. Anat.* **194**: 3-6.
- Kolman, C.J., Bermingham, E.** 1997. mtDNA and nuclear DNA diversity in the Chocó and Chibcha Amerinds of Panamá. *Genetics.* **147**: 1289.
- Lalueza-Fox, C., Calderon, Fl., Calafell, F., Morera, B., Bertranpetit, J.** 2001. mtDNA from extinct Tainos and the peopling of the Caribbean. *Ann. Hum. Genet.* **65**: 137-151.
- Lalueza-Fox, C., Gilbert, M.T.P., Martínez-Fuentes, A.J., Calafell, F., Bertranpetit, J.** 2003. Mitochondrial DNA from pre-Columbian Ciboneys from Cuba and the pre-historic colonization of the Caribbean. *Am. J. Phys. Anthropol.* **121** (2): 97-108.
- Lee, H.Y., Park, M.J., Kim, N.Y., Sim, J.E., Yang, W.I., Shin, K.J.** 2010. Simple and highly effective DNA extraction methods from old skeletal remains using silica columns. *Forensic Sci. Int. Genet.* **4**: 275-280.
- 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. *Am. J. Phys. Anthropol.* **127**(3): 351-60.
- Lewis, C.M., Buikstra, J.E., Stone, A.C.** 2007. Ancient DNA and genetic continuity in the South Central Andes. *Lat. Am. Antiq.* **18**: 145-160.
- Librado, P., Rozas, J.** 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics.* **25**: 1451-1452.
- Llamas, B., Fehren-Schmitz, L., Valverde, G., Soubrier, J., Mallick, S., Rohland, N., et al.** 2016. Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Sci. Adv.* **2**: e1501385.
- López-Castaño, C.E., Cano-Echeverry, M.C.** 2011. En torno a los primeros poblamientos en el Noroccidente de Sudamérica: Acercamientos desde el valle Inter-Andino del Magdalena, Colombia. *Boletín de Arqueología PUCP.* (15): 43-79.
- Malhi, R., Cybulski, J.S., Tito, R.Y., Johnson, J., Harry, H., Dan, C.** 2010. Brief Communication: mitochondrial haplotype C4c confirmed as a founding genome in the Americas. *Am. J. Phys. Anthropol.* **141** (3): 494-497.
- Manríquez, G., Moraga, M., Santoro, C., Aspíllaga, E., Arriaza, B.T., Rothhammer, F.** 2011. Morphometric and mtDNA analyses of archaic skeletal remains from south western South America. *Chungará (Arica).* **43** (2): 283-292.
- Marchant, R., Behling, H., Berrío, J.C., Cleef, A., Duivenvoorden, J., Hooghiemstra, H., et al.** 2002. Pollen-based biome reconstructions for Colombia at 3000, 6000, 9000, 12000, 15000 and 18000¹⁴C yr ago: Late Quaternary tropical vegetation dynamics. *J. Quat. Sci.* **17**: 113-129.
- Martínez-Cruzado, J.C., Toro-Labrador, G., Viera-Viera, J., Rivera-Vega, M.Y., Startek, J., Latorre-Esteves, M., et al.** 2005. Reconstructing the population history of Puerto Rico by means of mtDNA phylogeographic analysis. *Am. J. Phys. Anthropol.* **128**: 131-155.
- Martínez-Cruzado, J.C.** 2010. The history of Amerindian mitochondrial DNA lineages in Puerto Rico. In: S.M. Fitzpatrick, A.H. Ross (editors). *Island Shores, Distant Pasts: Archaeological and Biological approaches to the pre-Columbian settlement of the Caribbean*. University Press of Florida, Gainesville-Florida, USA. p. 54-80.
- Mata-Míguez, J., Overholtzer, L., Rodríguez-Alegria, E., Kemp, B.M., Bolnick, D.A.** 2012. The Genetic Impact of Aztec Imperialism: Ancient Mitochondrial DNA Evidence from Xaltocan, Mexico. *Am. J. Phys. Anthropol.* **149** (4): 504-516.
- Melton, P.E., Briceño, I., Gomez, A., Devor, E.J., Bernal, J.E., Crawford, M.H.** 2007. Biological relationship between Central and South American Chibchan speaking populations: Evidence from mtDNA. *Am. J. Phys. Anthropol.* **132**: 753-770.
- Melton, P.E., Baldi, N.F., Barrantes, R., Crawford, M.H.** 2013. Microevolution, migration, and the population structure of five Amerindian populations from Nicaragua and Costa Rica. *Am. J. Hum. Biol.* **25** (4): 480-490.
- Mendisco, F., Keyser, C., Seldes, V., Rivolta, C., Mercolli, P., Cruz, P., et al.** 2014. Genetic diversity of a late prehispanic group of the Quebrada de Humahuaca, northwestern Argentina. *Ann. Hum. Genet.* **78**: 367-380.

- Merriwether DA, Rothhammer F, Ferrel RE.** 1995. Distribution of the four founding lineage haplotypes in Native Americans suggests a single wave of migration for the New World. *Am. J. Phys. Anthropol.* **98**: 411-430.
- Mesa, N.R., Mondragón, M.C., Soto, I.D., Parra, M.V., Duque, C., Ortíz-Barrientos, D., et al.** 2000. Autosomal, mtDNA, and Y-chromosome diversity in Amerinds: Pre- and post-Columbian patterns of gene flow in South America. *Am. J. Hum. Genet.* **67**: 1277-1286.
- Meyer, S., Weiss, G., Von Haeseler, A.** 1999. Pattern of nucleotide substitution and rate heterogeneity in the Hypervariable Regions I and II of human mtDNA. *Genetics.* **152** (3): 1103-1110.
- Minelli, A., Cozzolino, M., Di Nucci, A., Guglielmi, S., Giannantonio, M., Dámore, D., et al.** 2012. The Prehistory of the Colombian Territory: The results of the Italian archaeological investigation on the Checua Site (Municipality of Nemocón, Cundinamarca Department). *J. Biol. Res.* **85** (1): 95-97.
- Monsalve, M.V., Cárdenas, F., Guhl, F., Delaney, A.D., Devine, D.V.** 1996. Phylogenetic analysis of mtDNA lineages in South American mummies. *Ann. Hum. Genet.* **60**: 293-303.
- Moreno-Estrada, A., Gravel, S., Zakharia, F., Mchauley, J.L., Byrnes, J.K., Gignoux, C.R., et al.** 2013. Reconstructing the Population Genetic History of the Caribbean. *PLoS Genet.* **9** (11): e1003925.
- Mosch, C.J., Watson, P.J.** 1997. An Ancient Rocky Mountain Caver. *J. Caves Karst Stud.* **59** (1): 10-14.
- Mulligan, C.J., Kitchen, A., Miyamoto, M.M.** 2008. Updated three-stage model for the Peopling of the Americas. *PLoS One.* **3** (9): e3199.
- Nei, M., Kumar, S.** 2000. *Molecular Evolution and Phylogenetics.* New York, USA: Oxford University Press.
- Neves, W.A., Hubbe, M., Correal, G.** 2007. Human skeletal remains from Sabana de Bogotá, Colombia: A case of Paleoamerican morphology late survival in South America? *Am. J. Phys. Anthropol.* **133** (4): 1080-1098.
- O'Rourke, D.H., Hayes, M.G., Carlyle, S.W.** 2000. Spatial and temporal stability of mtDNA haplogroup frequencies in native North America. *Hum. Biol.* **72**: 15-34.
- O'Rourke, D.H.** 2009. Human Migrations: The Two Roads Taken. *Curr. Biol.* **19** (5): 203-205.
- O'Rourke, D.H., Raff, J.A.** 2010. The Human Genetic History of the Americas: The Final Frontier. *Curr. Biol.* **20**: 202-207.
- Pääbo, S., Poinar, H., Serre, D., Jaenicke-Despres, V., Hebler, J., Rohland, N., et al.** (2004). Genetic analyses from ancient DNA. *Annu. Rev. Genet.* **38**: 645-679.
- Parson, W., Dür, A.** 2007. EMPOP--a forensic mtDNA database. *Forensic Sci Int Genet.* **1** (2): 88-92.
- Perego, U.A., Achilli, A., Angerhofer, N., Accetturo, M., Pala, M., Olivieri, A., et al.** 2009. Distinctive Paleoamerican migration routes from Beringia marked by two rare mtDNA haplogroups. *Curr. Biol.* **19** (1): 1-8.
- Perego, U.A., Lancioni, H., Tribaldos, M., Angerhofer, N., Ekins, J.E., Olivieri, A., et al.** 2012. Decrypting the mitochondrial gene pool of modern Panamanians. *PLoS One.* **7** (6): e38337.
- Pickrell, J., Reich, D.** 2014. Towards a new history and geography of human genes informed by ancient DNA. *Trends Genet.* **30** (9): 377-389.
- Pinto, M.** 2003. Galindo, un Sitio a Cielo Abierto de Cazadores-Recolectores en la Sabana de Bogotá. Bogotá: Fundación de Investigaciones Arqueológicas Nacionales, Banco de La República.
- Pucciarelli, H., Perez, I., Politis, G.** 2010. Early Holocene human remains from the Argentinean Pampas: Additional evidence for distinctive cranial morphology of Early South Americans. *Am. J. Phys. Anthropol.* **143**: 298-305.
- Raff, J.A., Bolnick, D.A., Tackney, J., O'Rourke, D.H.** 2011. Ancient DNA Perspectives on American colonization and population history. *Am. J. Phys. Anthropol.* **146**: 503-514.
- Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J.S., Albrechtsen, A., Moltke, L., et al.** 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature.* **463**: 757-762.
- Rasmussen, M., Anzick, S.L., Waters, M.R., Skoglund, P., DeGiorgio, M., Stafford, T.W., et al.** 2014. The genome of a Late Pleistocene human from a Clovis burial site in Western Montana. *Nature.* **506** (7487): 225-229.
- Rasmussen, M., Sikora, M., Albrechtsen, A., Korneliusen, T.S., Moreno-Mayar, V., Poznik, G.D., et al.** 2015. The ancestry and affiliations of Kennewick Man. *Nature.* **523**: 455-458.
- Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, et al.** 2012. Reconstructing Native American population history. *Nature.* **488** (7411): 370-374.
- Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Bronk-Ramsey, C., et al.** 2013. IntCal13 and Marine13 Radiocarbon Age Calibration Curves 0-50,000 years Cal BP. *Radiocarbon.* **55** (4): 1869-1887.
- Ribeiro-Dos-Santos, A.K.C., Santos, E.C.B., Machado, A.L., Guapindaia, M., Zago, M.A.** 1996. Heterogeneity of mitochondrial DNA haplotypes in pre-Columbian natives of the Amazon region. *Am. J. Phys. Anthropol.* **101** (1): 29-37.
- 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 lineages for the Native American populations. *Am. J. Hum. Genet.* **65**: 519-530.
- Rivera, S.** 1991. Neusa 9000 Años de Presencia Humana en el Páramo. Bogotá: Fundación de Investigaciones Arqueológicas Nacionales, Banco de la República.
- Rodríguez, J.V.** 2011. Los Chibchas: Hijos del sol, la luna y los Andes. Del mito a la historia natural. Bogotá: Universidad Nacional de Colombia, IDU-UNC. 286p.
- Rodas, C., Keyeux, G., Gelvez, N.** 2003. Mitochondrial DNA studies show asymmetrical Amerindian admixture in Afro-Colombian and Mestizo populations. *Hum. Biol.* **75** (1): 13-30.
- Rondón F, Braga Y, Barreto G.** 2007. Análisis de la diversidad y el grado de estructura genética presente en poblaciones humanas colombianas a partir del uso de marcadores RFLPs de mtDNA. *Rev. Asoc. Colomb. Cienc. Biol.* **19**: 94-103.
- Ruhlen, M.** 1991. *A guide to the world's languages, Vol. 1: Classification.* CA: Stanford University Press.
- Salas, A., Acosta, A., Álvarez-Iglesias, V., Cerezo, M., Phillips, C., Lareu, M.V., et al.** 2008. The mtDNA ancestry of admixed Colombian populations. *Am. J. Hum. Biol.* **20** (5): 584-591.
- Saitou, N., Nei, M.** 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406-425.

- Sandoval, J.A., Almanza, O.** 2012. Datación de esmalte dental prehispánico proveniente del sitio arqueológico Checua (Cundinamarca) por Resonancia Paramagnética Electrónica (EPR). *Rev. Col. Fis.* **44** (3): 248-252.
- Silva, A., Briceño, I., Burgos, J., Torres, D., Villegas, V., Gómez, A., et al.** 2008. Análisis de ADNmt en una muestra de restos óseos arcaicos del período Herrera de la Sabana de Bogotá. *Biomédica.* **28** (4): 569-77.
- Soper, D.S.** 2015. Analysis of Variance (ANOVA) Calculator - One-Way ANOVA from Summary Data. Date Query: 21-01-2015. Available in: <http://www.danielsoper.com/statcalc>
- Stone, A.C., Stoneking, M.** 1998. mtDNA Analysis of a Prehistoric Oneota population: Implications for the peopling of the New World. *Am. J. Hum. Genet.* **62**: 1153-1170.
- Stothert, K.** 1985. The preceramic Las Vegas culture of coastal Ecuador. *American Antiquity.* **50**: 613-637.
- Tajima F.** 1989. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics.* **123**: 585-595.
- Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D.G., Mulligan, C.J., Bravi, C.M., et al.** 2007. Beringian Standstill and spread of Native American founders. *PLoS One.* **9**: e829.
- Tamura, K., Stecher, G., Peterson, D., Filipinski, A., Kumar, S.** 2013. MEGA6: Molecular Evolutionary Genetics Analysis 6.0. *Mol. Biol. Evol.* **30**: 2725-2729.
- Tamura, K., Nei, M.** 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**: 512-526.
- Usme-Romero, S., Alonso, M., Hernández-Cuervo, H., Yunis, J.E., Yunis, J.J.** 2013. Genetic differences between Chibcha and Non-Chibcha speaking tribes based on mtDNA haplogroups from 21 Amerindian tribes from Colombia. *Genet. Mol. Biol.* **36** (2): 149-157.
- Van der Hammen, T.** 1986. Cambios medio-ambientales y la extinción del mastodonte en el norte de los Andes. *Revista de Antropología.* **2**: 27-33.
- Van der Hammen, T.** 1997. El Bosque de Condalia. *Caldasia.* **19** (1-2): 355-359.
- Van der Hammen, T., Correal, G.** 2001. Mastodontes en un humedal Pleistocénico en el valle del Magdalena (Colombia) con evidencias de la presencia del hombre en el pleniglacial. *Boletín de Arqueología.* **16**: 4-36.
- Van der Hammen, T.** 2006. Bases para una prehistoria ecológica amazónica y el caso de Chiribiquete. En: G. Morcote, S. Mora, C. Franky (editores). *Pueblos y Paisajes antiguos de la selva amazónica.* Universidad Nacional de Colombia, Facultad Ciencias, Taraxacum, Bogotá. p. 19-28.
- Van der Hammen, T.** 2011. Aspectos ambientales y paleo-ambientales de los enclaves secos y el caso particular del valle del Checua (Nemocón, Colombia). *Pérez Arbelaeza.* **19**: 35-47.
- Vilar, M.G., Melendez, C., Sanders, A.V., Walia, A., Gaieski, J.B., Owings, A.C., et al.** 2014. Genetic diversity in Puerto Rico and its implications for the peopling of the Island and the West Indies. *Am. J. Phys. Anthropol.* **155** (3): 352-368.
- Vona, G., Falchi, A., Moral, P., Calo, C., Varesi, L.** 2005. Mitochondrial sequence variation in the Guahibo Amerindian population from Venezuela. *Am. J. Phys. Anthropol.* **127**: 361-369.
- Wang, S., Lewis, C.M., Jakobsson, M., Ramachandran, S., Ray, N., Bedoya, G., Rojas, W., et al.** 2007. Genetic variation and population structure in Native Americans. *PLoS Genet.* **3** (11): e185.
- Watkins, W.S., Xing, J., Huff, C., Whitterspoon D.J., Zhang, Y., Perego, U.A., et al.** 2012. Genetic analysis of ancestry, admixture and selection in Bolivian and Totonac populations of the New World. *BMC Genetics.* **13**: 39.
- Willerslev, E., Cooper, A.** 2005. Ancient DNA. *Proc. Biol. Sci.* **272**: 3-16.
- Winters M, Barta JL, Monroe C, Kemp B.** 2011. To Clone or Not To Clone: Method analysis for retrieving consensus sequences in ancient DNA Samples. *PLoS One.* **6**(6): e21247.
- Xavier, C., Builes, J.J., Gomes, V., Ospino, J.M., Aquino, J., Parson, W., et al.** 2015. Admixture and Genetic Diversity Distribution Patterns of Non-Recombining Lineages of Native American Ancestry in Colombian Populations. *PLoS One* **10** (3): e0120155.
- Yang, N.N., Mazières, S., Bravi, C., Ray, N., Wang, S., Burley, M.W., et al.** 2010. Contrasting patterns of nuclear and mtDNA diversity in Native American populations. *Ann. Hum. Genet.* **74** (6): 525-538.
- Yunis, J.J., Yunis, E.J.** 2013. mtDNA haplogroups in 1526 unrelated individuals from 11 Departments of Colombia. *Genet. Mol. Biol.* **36** (3): 329-335.
- 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. *Mol. Biol. Evol.* **21** (1): 164-175.