Tracing maternal Amerindian ancestry in the Aruban population through mtDNA sequencing

by

Ivenise Carrero-González

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
in
Biology

UNIVERSITY OF PUERTO RICO
MAYAGÜEZ CAMPUS
2011

Approved by:

_____	_____
Fernando Bird-Picó, Ph.D. Member, Graduate Committee	Date
_____	_____
Matías Cafaro, Ph.D. Member, Graduate Committee	Date
_____	_____
Juan C. Martínez-Cruzado, Ph.D. President, Graduate Committee	Date
_____	_____
Nannette Diffoot Carlo, Ph.D. Chairperson of the Department	Date
_____	_____
Carlos Quiñones, Ph.D. Graduate School Representative	Date

ABSTRACT

As part of a study about pre-Colombian female migrations in the Caribbean, the mitochondrial DNA of the modern Aruban population was studied. The main objective was to identify the haplotypes in the Amerindian haplogroups A2, B2, C1 and D1 and compare them with those patterns already identified in the Caribbean and South America. The HVR-I was sequenced in 144 samples of umbilical cord blood from babies born in 2006, of which 76 (54.3%) resulted of Amerindian origin. The HVR-II was sequenced in those Amerindian samples. The results obtained from HVR-I and II demonstrated that 32 (42.1%) of the 76 samples belong to the haplogroup D1, 18 (23.7%) belong to haplogroup B2, 17 (22.4%) to haplogroup A2, and 9 (11.8%) to haplogroup C1. It was shown that Aruba has the New World founder Amerindian haplotypes for the four haplogroups. These haplogroups have a nucleotide diversity (π) of 0.0042 (A2), 0.0032 (B2), 0.0018 (C1) and 0.0049 (D1). Median Network analyses show that haplogroup A2 has one derived haplotype that is present in the Dominican Republic and another one in Cuba; this suggests a migration process from the north of South America to the Greater Antilles. Haplogroup B2 has the New World founder lineage with derived haplotypes. Haplogroup C1 has the lowest diversity and frequency, and the New World founder lineage predominate in this haplogroup. Haplogroup D1 has four lineages, two of them in high frequency and very low diversity. This fact, together with the high frequency of haplogroup D1 in Aruba, suggests massive recent migrations to the island from a population with a high frequency of haplogroup D1.

RESUMEN

Como parte de un estudio de migraciones femeninas pre-colombinas en el Caribe, se estudió el ADN mitocondrial de la población moderna de Aruba. El objetivo era identificar haplotipos distintivos de los haplogrupos nativo americanos A2, B2, C1 y D1, y compararlos con aquellos patrones nativo-americanos ya conocidos de la región circumcaribeña. Se le secuenció la región hipervariable I (HVR-I) a 144 muestras de sangre umbilical de bebés nacidos en el 2006, de las cuales 76 (54.3%) han resultado ser de origen amerindio. A estas se les secuenció la región hipervariable II (HVR-II). Los resultados de HVR-I y II demuestran que 32 (42.1%) de las 76 muestras son del haplogrupo D1, 18 (23.7%) del haplogrupo B2, 17 (22.4%) del haplogrupo A2, y 9 (11.8%) del C1. Se demostró que Aruba posee los haplotipos fundadores del Nuevo Mundo para cada uno de los cuatro haplogrupos. Estos haplogrupos presentan una diversidad nucleotídica (π) de 0.0042 (A2), 0.0032 (B2), 0.0018 (C1) y 0.0049 (D1). Análisis de redes medianas demuestran que en el haplogrupo A2 existe un haplotipo derivado que está presente en la República Dominicana y otro derivado que se encuentra en Cuba; esto sugiere un proceso migratorio del norte de Sur América hacia las Antillas Mayores. El haplogrupo B2 contiene el linaje correspondiente al fundador del Nuevo Mundo con haplotipos derivados. El haplogrupo C1 tiene la menor diversidad y la menor frecuencia, y el linaje que predomina en este haplogrupo es el fundador del Nuevo Mundo. El haplogrupo D1 consiste de cuatro linajes distintos, dos de estos en alta frecuencia y muy poca diversidad. Este hecho, en combinación con la más alta frecuencia del haplogrupo D1 en Aruba, sugiere migraciones recientes masivas de poblaciones con altos niveles del haplogrupo D1 a esta isla.

ACKNOWLEDGMENTS

First, I want to thank Dr. Juan C. Martínez-Cruzado for all his guidance, support and patience since day one. I learned a lot from him through this process. I want to thank my friends in the laboratory, whom always were there when I needed it, especially Francez, for helping me in the process of reading sequences, Alondra, for all the techniques she taught me, Zamara, Liz, Leonardo, Ernesto, and everyone who made all the hours spent at the laboratory a little funnier. I want to thank Dr. Matías Cafaro and Dr. Fernando Bird, for all their advices, María Mendez and Mary Jimenez in the office of graduate studies in Biology for always being there when I needed it. I want to express my sincere gratitude to Oswald Webber who provided all the samples. I am very grateful to my parents and sister for all their unconditional support in all these years and God, for giving me all these great opportunities.

Table of Contents

Table List

Figure List

INTRODUCTION

America is the continent most recently colonized by Homo sapiens. It is accepted that humans migrated to North America from Asia through Beringia during the Pleistocene approximately 15,000 to 20,000 years ago (Fagundes et al., 2008). Migration to South America occurred approximately 14,000 years ago, but it is still not clear if it was only a single migration or multiple migration events. This uncertainty stems from the fact that South America has a distinctive pattern of genetic diversity. The tribes in the western part of South America have high levels of genetic diversity within populations and have short genetic distances between populations while eastern populations have lower genetic diversity and larger genetic distance between populations. This pattern may occur as a result of two separate migration events, one giving origin to the Amazonian population and the other one giving origin to the Andean population. Andean populations could have a larger time depth or fewer effects from bottlenecks than Amazonian populations (Lewis et al., 2007). Another explanation for this pattern is a single migration event with later Amazonian and Andean populations having different population sizes and little gene flow between them. The single migration theory is more accepted and it proposes that the first colonizers or Paleoamerindians migrated to the south crossing the Darien forest between Colombia and Panamá, reaching the Andes, the plains, and the Amazon Forest (Keyeux et al., 2002).

The Clovis tradition originated in America approximately 13,500 years ago. This tradition belonged to gatherer-hunter people who used stone-pointed spears to hunt animals. The Clovis tradition was characterized by fluted projectile points, objects with straight sides and the flutes removed from their bases, all of them constructed from different stone types. Clovis

people hunted large animals like mammoths and mastodons, although they hunted smaller animals as well. In South America there is no evidence of Clovis points, but there is evidence of Paleoamerindian remains. These sites had artifacts which include fishtail, willow-leaf and triangular stemmed points dating approximately 13,200-12,500 years ago. An example is the Cave of the Painted Rock in Monte Alegre, Brazil, which contains biological remains and rock paintings that show the use of nuts, fish, turtles and mussels. In addition, this cave has triangular points made of quartz and chalcedony that could be used by Paleoamerindians to hunt large animals (Roosevelt et al. 1996). Thus, Paleoamerindians from South and North America seemed to have different lifestyles.

In Aruba, the oldest remains found are dated to approximately 4,500 years ago; these pre-ceramic archaeological sites are from the Meso-Indian I period (5,000-1,000 years before present). These people were fisher-hunter-gatherers who subsisted primarily on seafood. It is believed that Aruban archaics traded with the ceramic cultures of Venezuela, because polished axe-heads made of volcanic material foreign to Aruba have been found associated with this culture (Oliver, 1995). The Paraguana peninsula in Venezuela is very close to Aruba, approximately 30 km away, which is why it is suggested as the most likely candidate for such trade (Toro-Labrador et al., 2003). The Aruban ceramic culture belonged to the Dabajuroid tradition, specifically the Dabajuran sub-tradition and their pottery had elements related with the Tierroid tradition of western Venezuela. It is suggested that they have a common ancestral culture that may have inhabited in Turbio, Upper Cojedes or Upper Yaracuy rivers. The Amerindians who lived in Aruba, Curaçao, Bonaire and Paraguana peninsula were known as Caquetíos by the Spaniards, and their language belonged to the Arawakan family of languages. The Arawakan family of languages was originated between the Negro and Amazon rivers about

5,500 years ago (Oliver, 1989). Some of the Proto-Arawakan speakers migrated west, upstream along the Meta River reaching the Andes or the Apure River diverging up the Cojedes River and then reached the Yaracuy River. People who reached the Yaracuy River gave rise to the Macro-Dabajuroid tradition. In Aruba, the Dabajurans were farmers and their ceramic sites have fewer shells than the pre-ceramic people. Aruban Amerindians maintained close contact with mainland cultures; this is known because approximately 700 chert flakes were found at the Tanki Flip site, but there is no chert source in Aruba. Additional evidence was found that showed avemorphic motifs painted in pottery. These avemorphic motifs originated in La Guajira peninsula of Colombia, suggesting that Aruban Amerindians traded with tribes in Venezuela and tribes in other current countries as well. Because Caquetíos from Venezuela and Guajiros from Colombia were geographically close to Aruba, both belonged to the Arawakan family, and both maintained communication with Aruban Amerindians, it is suggested that Aruban Amerindians might have originated from both of them (Toro-Labrador et al., 2003).

Figure 1. Map illustrating Aruba, the North of Venezuela and La Guajira Peninsula

At the time of colonization in Aruba by the Spaniards in 1499, approximately 450 to 600 Amerindians lived in Aruba. The population of Aruba, Curaçao and Bonaire was sent to work in mines to Hispaniola in 1513 because these islands were declared useless. Later, in 1526, approximately 150 to 200 Amerindians were returned to the three islands to work for the Spaniards on the exportation of brasilwood, divi-divi and kwihi. The majority of these Amerindians were Caquetíos, but some of them were Arawaks from other islands (Hartog et al., 1961). At this time, the Spaniards noticed the presence of Amerindians in the island. They could have been Caquetíos who had escaped from deportation hiding in caves, or Amerindians from Venezuela who escaped from the mainland when the colonization of this region began. Another deportation of Aruban Amerindians occurred upon the Dutch conquest of the Netherland Antilles in 1636. At this time, Spanish and Caquetío languages were widely spoken. Amerindians were deported to the mainland because they were considered sympathetic to the Spaniards. However, in this same year, some Amerindians were returned to Aruba to work in the breeding of horses and cattle (Hartog el al. 1961). In 1648, with the Dutch Peace Treaty, Aruba was abandoned again and in 1655 the Amerindians were considered free by the Dutch West Amerindian Company. In the 17th century the Amerindians were described as Catholics, speaking Spanish and were frequently visited by Spanish priests from the mainland. There is evidence that Aruban Amerindians had strong links with people from the mainland, for example, some of them agreed to go to Venezuela to raise the town of El Carrizal. The presence of African slaves was mentioned in 1750, because some Africans killed four residents. At that time, the only white person was the Governor of Aruba and it is possible that he was the only person who owned black slaves, meaning that the number of black people in Aruba was very low. During the 19th century the admixture increased and the Caquetío language disappeared but the trade of

Amerindian slaves, mainly Guajiro Amerindians, persisted until the abolition of slavery in 1863 (Hartog et al. 1961).

In conclusion, the Arubans are a mixed population consisting of Europeans, Africans and Amerindians. Their history suggests that their Amerindian portion is closely related to Amerindians from the Guajira and Paraguana peninsulas. Although their culture can be related to the Arawaks migrating from the Amazon, it is possible that they are the result of an admixture process between manihot-cultivating Amerindians coming from the Amazon and nomads from northwestern South America. This study focuses on the northwestern part of South America, this area where inhabited only by neotropical Amerindians. For this reason, they are going to be mentioned simply as Amerindians from this point forward.

This study is aimed to find the percentage of Amerindian heritage of the Aruban population through the maternal line and their relationship with various South American and Caribbean tribes. This research is part of a study about pre-Colombian female migrations in the Caribbean. Sequencing the Hypervariable Region I (HVR-I) of the mtDNA, the samples of Aruban population will be clustered in different haplogroups, looking for the four Native American haplogroups A2, B2, C1 and D1. The Hypervariable Region II (HVR-II) of those belonging to any Native American haplogroup was also being amplified and sequenced. The initial expectations were to find high percentage of haplogroups C1 and D1 because of Aruba's proximity to South America. They were not met.

PREVIOUS PUBLICATIONS

Mitochondrial DNA (mtDNA) is a small circular double stranded extranuclear DNA. It is haploid and uniparentally inherited, exclusively from the mother. This DNA does not recombine with the mtDNA of the father; as a result it is maintained intact through generations, conserving its original identity (Torroni et al. 1994a, Martínez-Cruzado, 2002). Variability in this DNA is introduced by mutations alone, not by recombination. These facts enable tracing the maternal lineage far back in time. The mutation rate in mitochondrial DNA is five to ten times higher than in nuclear DNA and these mutations are common deletions and base substitutions (Brown et al., 1979, Wallas et al., 2009). Deletions generally cause disease, while the majority of base substitutions are neutral and thus are one of the main objects of study by evolutionary geneticists (Excoffier and Langaney, 1989). Because mtDNA is not highly conserved and its mutation rate is rapid, it is useful for phylogeographic studies.

Mitochondrial DNA has a control region which is non-coding and regulatory. The rate of base substitutions in this region is higher than in the rest of the molecule (Soares et al., 2009). There are two hypervariable regions inside this region, HVR-I and HVR-II. The HVR-I consists of 349 base pairs and it is the most polymorphic. Genetic bottlenecks associated with historical migrations followed by population expansions were often accompanied by mutations that became unique to the derived expanded populations (Wallace et al., 2009, Henríquez et al, 2004, Hernstadt et al., 2002). These variations are inherited as haplotypes and lineages in different populations. People who share the same mutations are clustered in the same haplogroup, suggesting that they share a common ancestor. While Restriction Fragment Length Polymorphism (RFLP) was the most commonly used technique in the 1980's and early 1990's,

the DNA sequencing of the HVR-I and HVR-II, and more recently of the complete mtDNA, has been employed, resulting in much more precise phylogeographic analyses (Henríquez et al., 2004). In the analysis made by Herrnstadt et al. (2002), with 560 sequenced mtDNA coding regions from unrelated individuals, more than 99% of the sequences could be assigned unambiguously to a single mtDNA haplogroup.

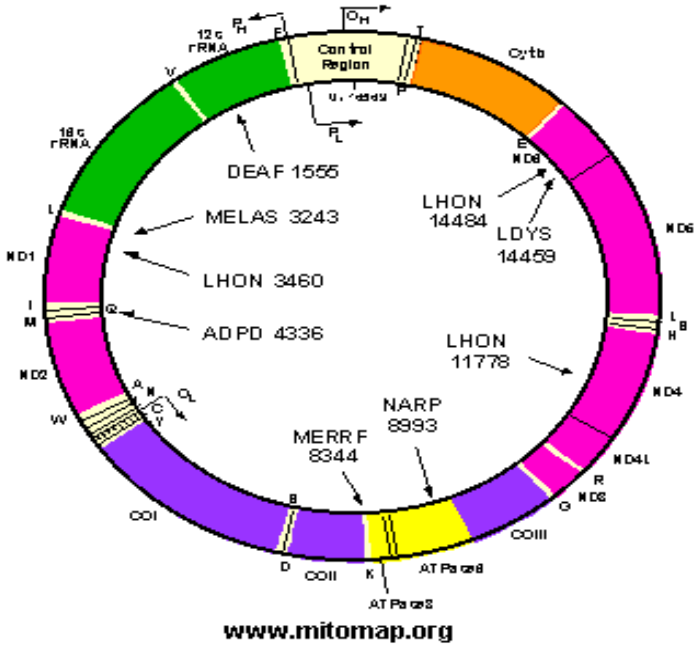

Figure 2. Mitochondrial DNA map

Mitochondrial DNA analysis made by Cann et al. (1987) with RFLP suggested that the descent of all human mtDNAs is from one woman from Africa who existed about 200,000 years ago. Outside Africa, mtDNAs fall into two clades, M and N. In Sub-Saharan Africa the most predominant haplogroup is L, but in Ethiopia the haplogroups M and N are present, pointing that modern human migration occurred from Eastern Africa. Haplogroup M has branches, which are: C, D, E, Q, Z and G and they are found commonly in South and East Asia and the Americas.

Branches of haplogroup N are: A, B, F, H, I, J, K, O, P, R, S, T, U, V, W, X and Y and they are common in West Asia and Europe (Jobling, Hurles and Tyler-Smith, 2004).

Phylogeographic analysis of mtDNA lineages from all over the world has led to the identification of mtDNA haplogroups that are specific to Africans, Europeans, or Asians/Amerindians (Alves-Silva et al., 2000). With mitochondrial genome analyses we can construct phylogenetic trees and define the timing and direction of human dispersions more precisely. The emergence of complete mtDNA sequences made it possible to reconstruct the phylogenies of African, European, Oceanian, eastern Asian, southeastern Asian, and Amerindian lineages and to gain detailed insight into human evolution and pioneer settlement processes (Derenko et al., 2007). The coexistence of different genetic lineages in southwestern Asia and the Altai-Sayan region of southern Siberia may have resulted from various migrations from diverse geographical sources at different times, beginning with the early human settlements in the Paleolithic era (Derenko et al., 2007).

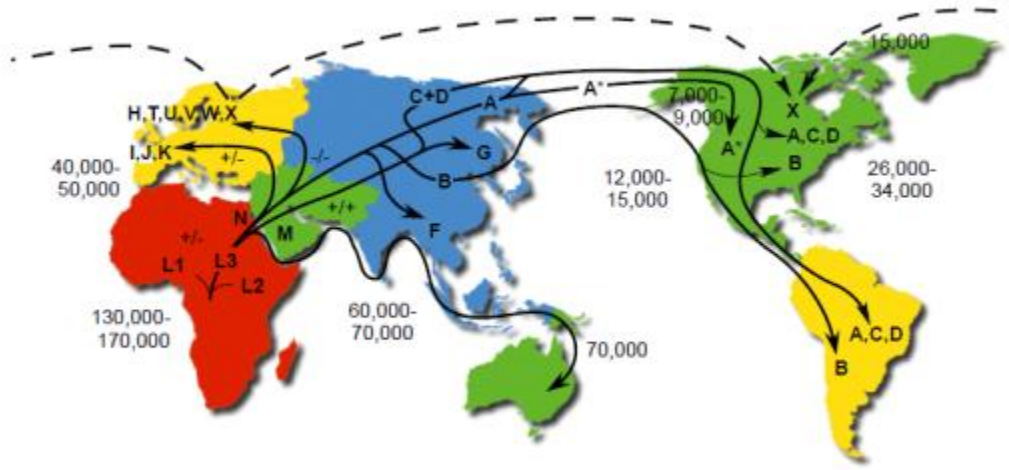

Figure 3. World map showing human migration routes from Africa to America

Mitochondrial DNA (mtDNA) has been commonly used to understand the peopling of the Americas. Molecular evidence suggests that ancestral Native American populations may have emerged from Siberia, since several maternally and paternally inherited genetic lineages present in both Siberia and the Americas appear to have evolved in that region of northern Asia (Derenko et al., 2007). It has been found that extant Native American populations exhibit almost exclusively five mtDNA haplogroups: A, B, C, D, and X. Haplogroups A–D are found all over the New World and are frequent in Asia, supporting a northeastern Asian origin of these lineages (Wallas et al., 2009, Torroni et al., 1993, Fagundes et al., 2008). These lineages were first described by RFLP. Later, sequence data showed a correlation between these lineages and particular mutations in the HVR-I of the mtDNA genome (Torroni et al., 1993). In the case of haplogroup X, it is presently found in the New World at a relatively low frequency and only in North America, it is uncommon in West Eurasians, and it is almost absent in Siberia, but there is evidence suggesting that this haplogroup together with the other four main mtDNA haplogroups, was part of the gene pool of a single Native American founding population (Fagundes et al., 2008). In South America and the Caribbean, ancient skeletal remains older than the time of Columbus' arrival and samples of present day populations are used to trace migration processes. The vast majority of ancient DNA studies have been based on the analysis of mtDNA because it has a better chance of recovery, since a cell with two copies of the nuclear genome can contain several hundred copies of the mtDNA genome. These analyses confirm that the most predominant haplogroups in South America and the Caribbean are A, B, C and D for Native Americans (Lalueza-Fox et al. 2001, Torroni et al 1993, Alves-Silva et al. 2000, Fuselli et al 2003). The founder lineages of the Americas, A2, B2, C1 and D1, have been identified by RFLP analysis and the complete sequencing of the mtDNA. Lineage A2 is characterized by the

presence of HaeIII restriction site at nucleotide position (np) 663 and the absence of HaeIII restriction site at np 16517. Lineage B2 is characterized by a 9 base pair intergenic deletion in COII/tRNALys together with the presence of HaeIII restriction site at np 297 and 16517. Lineage C1 is characterized by the absence of HincII restriction site at np 13259 and the presence of restriction sites for the enzymes: AluI at np 13262 and 10397, DdeI at np 10394 and HaeIII at np 16517. Lineage D1 is characterized by the absence of AluI restriction site at np 5176, and DdeI np 10,394 and AluI np 10397 site gains (Torroni et al. 1992, Easton et al. 1996, Kolman et al. 1996). The HVR-I and HVR-II characteristic mutations these lineages are: A2 (16111, 16362, 64, 146, and 152 or 153), B2 (16189, 16217 and 499), C1 (16325, 290del and 291del), and D1 (16325) (Achilli et al. 2009, Torroni et al. 1993).

Despite the dramatic reduction in size of the Amerindian populations after the European conquest and colonization of the Americas, the mtDNA lineages found in present-day Native American populations are thought to represent the ancestral lineages (Keyeux et al., 2002). A high frequency of haplogroups A and B is observed in North America, Central America and the Andes in South America. In Antioquia, Colombia, a region right next to Panama, haplogroups A and B are the most predominant, resembling the pattern of Central America (Carvajal-Carmona et al., 2000). In the northwest of Colombia haplogroups C and D increase progressively from north to south, being very common from Venezuela as far as to Patagonia displacing completely the haplogroups A and B in the Yamanas tribes at South America (Henríquez et al., 2004). In Venezuela, the most predominant haplogroups are A and C, but in the town of Macuquita, which belongs to the Falcon State, the most predominant haplogroup is D as well as in Aruba (Castro et al., 2009).

In the Caribbean, the most predominant haplogroups are A and C as well as in Venezuela (Martínez et al., 2007), suggesting that migration to the Caribbean may have occurred from this continental region. As an example, from most of the linguistic, biological, and cultural evidence, it is apparent that the Ceramic culture that arrived in Puerto Rico 2,500 years before present originated in South America, in or close to the territory occupied today by the Warao in Venezuela. The Taíno language belongs to the Arawakan family of languages. From the distribution of this family of languages, Arawakan speakers have been inferred to have originated in the South American lowlands of the Amazon Basin. The haplogroup distribution in the Puerto Rican Taínos mtDNAs is highly structured, with haplogroups A and C accounting for 88% of all mtDNAs. The highly structured haplogroup distribution suggests that the Taínos had reduced genetic diversity. This scenario is consistent with the repeated water crossings over the Lesser Antilles that a colonization pattern originating in the Orinoco delta would entail (Martínez-Cruzado et al., 2001). The scarcity and geographic distribution of haplogroup D mtDNAs in Puerto Rico suggest that haplogroup D may have been imported during colonial times. It is known that, for centuries, one of the main Amerindian slave harbors in the Americas was located in Coro, northwestern Venezuela. It provided plenty of Amerindian slaves to Aruba in the 19th century, where haplogroup D is predominant, and it is possible that it provided slaves to Puerto Rico during that period (Martínez-Cruzado et al., 2005). In Cuba, the haplogroup A2 is the most predominant but unlike Puerto Rico, its lineages cannot be related to any continental region, even though other haplogroups suggest that the origins of Native Americans in this country, the Taínos and the Ciboneys are related to South America (Lalueza-Fox et al., 2003; Mendizabal et al., 2008). From the only scientific article on Aruba, it seems as if the distribution of Native American mtDNA haplogroups is highly structured being the haplogroup D the most

predominant. Moreover, their genetic diversity is low suggesting that most Aruban Amerindians originated from one or few closely related women. These observations imply that despite the intense movement of Amerindians in and out of Aruba during historical times, most Aruba Amerindian mtDNAs originate from one or few sibling tribes (Toro-Labrador et al., 2003).

OBJECTIVES

- To extract mitochondrial DNA from umbilical cord blood samples of the Aruban population.
- To amplify and sequence a region overlapping most of the HVR-I of the control region of mitochondrial DNA and compare it with the CRS (Cambridge Reference Sequence) to look for specific base haplotypes distinctive of Native American haplogroups.
- To amplify and sequence the HVR-II of those mtDNAs previously determined to belong to a Native American haplogroup.
- To obtain information about the possible origins of Aruban Caquetíos.
- To estimate demographic parameters of Aruba and study the phylogenetic relationship of Aruban mtDNA lineages with those of other regions in South America and the Caribbean as part of a study about pre-Colombian female migrations to the Caribbean.

EXPECTATIONS

The results of Toro-Labrador et al. (2003) showed that 13 of their 16 Aruban samples had Native American origins, being the haplogroup D1 the most predominant followed by haplogroup C1. I thus expect to find high percentage of haplogroups D1 and C1. Another good reason to expect a high frequency of haplogroup D1 is the fact that this haplogroup is very common in the coast of Venezuela, especially in the state of Falcon, and there is evidence suggesting that migration to the Caribbean occurred from South America, probably from Venezuela. In addition, I expect low genetic diversity because Aruba and the other islands of the Caribbean were the most recent lands to be colonized in America. The founder effect could affect the genetic pool of the Aruban Native Americans resulting in a highly structured distribution of the mtDNAs of this population. It is reasonable to expect a high percentage of Native American mtDNAs because the majority of the population in America results of the admixture of European colonizer men and Native American colonized women. For this reason, I anticipate to find more than 50% of my samples clustering in haplogroups D1 and C1.

METHODS

Samples:

Umbilical cord blood samples of 144 Aruban born in three months on 2006 were obtained with informed consent by Dr. Oswald Wever at the Dr. Horacio Oduber Hospital in Oranjestad, Aruba.

Experimental procedure:

DNA extraction

First, 50 µl were utilized of each sample and 1.25 µl of Proteinase K (20 mg/ml) was added. The samples with Proteinase K were left rotating overnight at 37 ºC. The purification of DNA was followed by adding 51.25 µl of the organic phase of phenol-chloroform-isoamyl alcohol, vortexing, centrifuging for 3 minutes at 14,000 rpm, taking the aqueous phase and transferring to another microtube. Then, 51.25 µl of chloroform were added, followed by vortexing and centrifugation for 1 minute at 14,000 rpm. The upper aqueous phase was transferred to a new microtube. The next step was to add 1/10 volume of sodium acetate 3 M pH 5.4 and 2.33 times the volume of ice-cold 100% ethanol, and mix by inversion. These solutions were left overnight at -20 ºC. The next day, the samples were centrifuged for 15 minutes at 14,000 rpm, the supernatant was discarded and 500 µl of 70% ice-cold ethanol was added. The samples were centrifuged for 5 minutes at 14,000 rpm. The supernatants were discarded, and the pellets were left air-drying from excess ethanol for at least 1-2 hours. The dried pellets were resuspended in 100 µl of TE buffer (10 mM Tris-HCl, 1 mM Na2EDTA, pH 8.0) to preserve the DNA.

Amplification of HVR I and II of mitochondrial DNA (PCR)

The amplification reaction mixture contained 0.4 µl of sample DNA, 3 µl of 25 mM dNTPs, 2.5 µl of NEB's (New England BioLabs) PCR Standard Buffer 10X, 1 µl of 25 mM MgCl2, 1 µl of Bovine Serum Albumin at 10 mg/ml, 1.3 µl of 20 µM L16154 primer (CCATAAATACTTGACCACCTG) and 1.3 µl of 20 µM H34 primer (ACCAAATGCATGGAGAGCTCC), or 1.3 µl of 20 µM L15829 primer

(CCTCCCTAAGACTCAAGG) and 1.3 µl of 20 µM H16345 primer (GGGACGAGAAGGGATTTGAC) for HVR-I; 1.3 µl of 20 µM L16491 primer (GGGGTAGCTAAAGTGAACTG) and 1.3 µl of 20 µM H501 primer (GTGTGTGCTGGGTAGGATG) for HVR-II, and 0.50 µl of Taq DNA polymerase (5 U/µl) plus 14 µl of distilled water for a total volume of 25 µl. The PCR conditions included an initial step at 94 º for 30 sec, 34 cycles consisting of: denaturation at 94 º for 30 sec, annealing at 53 º for 1 min and elongation at 72 º for 70 sec, and a final step of elongation at 72 º for 5 min.

PCR Product Purification

Purification of the PCR product was undertaken using the QIA quick PCR Purification Kit as directed by the manufacturer (Qiagen).

To complete the purification process an ethanol precipitation was made. First 5 µl of 3M sodium acetate pH 5.4, 1 µl of glycogen at 20 mg/ml and 131 µl of ice-cold ethanol 100% were added to the 50 µl purification product and left overnight at -20ºC after gentle mixing. Afterward, centrifugation at 14,000 rpm was made for 15 minutes, the supernatant was discarded, 0.5 ml of ice-cold ethanol 70% was added, and another centrifugation at maximum velocity for 5 minutes was made. The supernatant was discarded and the pellet was air dried. Dried pellets were resuspended in 15 µl of Elution Buffer (10 mM Tris-Cl pH 8.0).

Sequencing HVR I and II of mitochondrial DNA

For sequencing, 10 ng of the amplicons were used. Strands light (L) and heavy (H) were sequenced for HVR-I and Light Strand (L) for HVR-II. The reactants used were: 0.8 µl of 1 µM primer, 0.5µl of Sequence Buffer 5x and 0.5 µl of Big Dye Terminator v 3.1 (Applied

Biosystems) in a 5 µl volume reaction. The primers used for HVR-I were L16154, H34, L15829, or H16345. For HVR-II, only L16491 was used. Sequence reactions were subjected to capillary electrophoresis in an ABI 3130 Genetic Analyzer at the Department of Biology of the University of Puerto Rico at Mayagüez. A group of samples were sent to sequencing to Nevada Genomics Center.

Reading and comparison of the sequences

- The program Sequencher 4.1 was used to read and compare the sequences with the CRS (Cambridge Reference Sequence)
- The program Chromas 1.62 (Technelysium Pty Ltd. ©) was used to edit the sequences.
- The samples were identified as Amerindian using the basic haplotypes shown in Tables 1 and 2.

Table 1. HVR-I basic haplotypes of Amerindian haplogroups

Haplogroups	Basic haplotypes				
A2	16111	16223	16290	16319	16362
B2	16189	16217			
C1	16223	16298	16325	16327	
D1	16223	16325	16362		

Table 2. HVR-II basic haplotypes of Amerindian haplogroups

Haplogroups	Basic haplotypes						
A2	64	73	146	152	153	235	263
B2	73	263					
C1	73	249(del)	263	290(del)	291(del)		
D1	73	263					

Data Analysis

Program Network 4.6.0.0

Median Joining networks (Bandelt et al., 1999), with HVR I and II were constructed for each Amerindian haplogroup, A2, B2, C1 and D1 using NETWORK 4.6.0.0 (fluxus-engineering.com). These networks were used to identify the different lineages in each haplogroup. Lineages, were identified under the criteria that one haplotype separated from the others by more than two mutations belonged to a different lineage. The program Network was also used to calculate the parameter ρ and the time elapsed since The Most Recent Common Ancestor (TMRCA). The positions used to reach the time estimates were 16051-16400 and 68-263 combined, with a mutation rate of one base substitution for every 9,058 years (Soares et al., 2009). These networks were compared with networks of different populations: Venezuela, Colombia, French Guiana, Brazil, Mexico, Honduras, Cuba, Dominican Republic and Puerto Rico.

MEGA 5

The program MEGA 5 (Tamura et al., 2011) was used to align the sequences of each haplogroup and to calculate genetic diversity measures and neutrality test statistics in haplogroups and lineages.

The Tajima's test of neutrality was used to study whether a population is neutral (their mutations occur randomly) or if a population expansion occurred. The parameters used to calculate diversity measures were: nucleotide diversity, π; expected level of diversity, θ; number of segregating sites, S; number of segregating sites divided by number of base pairs, Ps; and the average number of mutational changes between the root haplotype and all the samples, ρ.

Arlequin 3.0

The program Arlequin 3.0 (Excoffier et al., 2005) was used to calculate Heterozygosity (H), the probability that two alleles, in this case haplotypes, drawn at random will be different from each other (Nei, 1987).

RESULTS AND DISCUSION

According to the basic haplotypes in the HVR I and II, 54.3% of the Aruban samples belong to an Amerindian haplogroup. The African ancestry is the second largest in Aruba, with 32.8% of the samples belonging to this group. All African major haplogroups, L1, L2 and L3, as well as paragroup L0, were found. The European ancestry is the lowest, with 12.9%.

The most predominant Amerindian haplogroup is D1; encompassing 42.1% of the Amerindian samples (Table 3). These results are in concordance with those of Toro-Labrador et al. (2003); they found haplogroup D1 as the most predominant haplogroup, but in greater proportion (69.2%).

Table 3. Amerindian Haplogroups Distribution in Aruba

Haplogroup	Percentage
A2	22.4%
B2	23.7%
C1	11.8%
D1	42.1%

The Tajima-D test was performed for these four haplogroups. This test measures neutrality in a population; if the value is not significantly different from zero, the population is under neutrality. Positive values indicate population subdivision or balancing selection and negative values indicate positive selection or population expansion (Jobling, Hurles and Tyler-Smith, 2004). Significant values were determined as in Tajima et al. (1989) with a confidence interval of 95%. The diversity measures π, θ, S, Ps, H and ρ were calculated (Table 4). Low diversity values are expected because New World populations have been affected by the founder effect; this means that a small group of people from a population in Siberia migrated and formed a new population of small size, and thus of lower genetic diversity (Jobling, Hurles and Tyler-Smith, 2004).

Population comparisons with other countries were performed only with HVR-I, because most publications do not present data for HVR-II.

Table 4. Genetic diversity measures and neutrality test results, and time since TMRCA

Haplogroups	Frequency	π	S	θ	Ps	H	ρ	Age	S.D.	Tajima's D
A2	0.224	0.0042	21	0.0066	0.0211	0.9890	1.71	15528	3362	-1.549
B2	0.237	0.0032	17	0.0053	0.0172	0.9417	1.19	10756	3444	-1.582
C1	0.118	0.0018	6	0.0026	0.0063	0.5238	0.86	7764	3167	-1.524
D1	0.421	0.0049	16	0.0040	0.0160	0.7097	3.28	29721	10580	0.751

Table 5. Time since TMRCA and rho estimations for each lineage

Lineage	ρ	S.D	Age	S.D.
A2				
A2-I	3	1.732	27174	15689
A2-II	2	1.414	18116	12810
A2-III	1	1	9058	9058
A2-IV	1.091	0.426	9881	3862
B2				
B2-I	1	0.385	9058	3484
B2-II	1.667	0.745	15097	6751
C1				
C1-I	0.333	0.236	3019	2135
C1-II	0.667	0.333	6039	3019
D1				
D1-I	0.308	0.243	2787	2203
D1-II	0.118	0.083	1066	753
D1-III	0.187	0.108	1698	980
D1-IV	3	1.732	27174	15689

Haplogroup A2

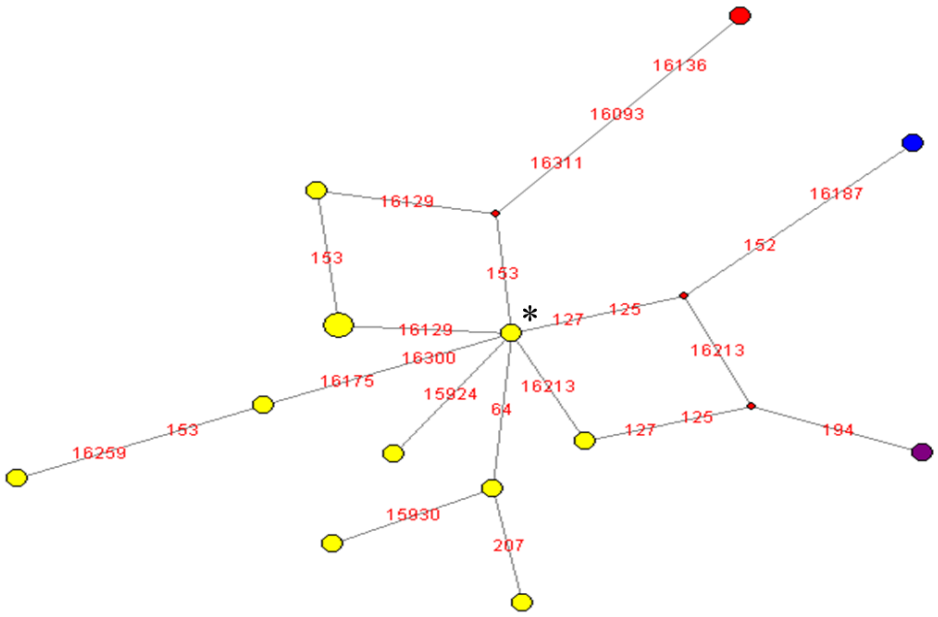

Figure 4. Median Joining Network of Haplogroup A2 for HVR I – II. (15900-16475 + 1-450) Lineages A2-I (red), A2-II (blue), A2-III (purple), and A2-IV (yellow) are highlighted. *16111-16223-16290-16319-16362-64-73-146-153-235-263

Haplogroup A2 is one of the most diverse among the four haplogroups ($\theta = 0.0066$). The network shows two reticulations, consistent with an old age in Aruba (Figure 4). These reticulations can be due to gene flow between populations sharing young alleles regardless of a more ancient fission from a common ancestor (Jobling, Hurles and Tyler-Smith, 2004). There are many mutations separating lineages from each other. Lineages A2-I, II and III are monohaplotypic and may thus represent very recent, even post-Columbian migrations. By contrast, all haplotype pairs in A2-IV but two are separated by a single transition. The absence of the intermediate haplotypes in the latter may be due to a sampling effect or to genetic drift. Genetic drift can occur through a bottleneck or a founder event due to migrations from another population to Aruba. In the latter case, lineage A2-IV would be containing two actual lineages. A bottleneck may have occurred when all Aruban Amerindians during the Spaniards colonization were sent to other Caribbean islands, and several years later many were sent back to Aruba, losing haplotypes in the process. The Tajima's D-test has no significant negative value. This means that a recent, strong population expansion did not occur. Their most common recent ancestor is dated to approximately 15528 years ago, being one of the most ancient haplogroups in Aruba. Lineage A2-IV may be the native lineage in Aruba, and their most recent common ancestor is dated to approximately 9881 years ago (Table 5). The antiquity of this haplogroup is observed in the expected diversity θ, which has one of the largest values.

One of the haplotypes (16111-16175-16223-16290-16300-16319-16362) is found in Cuba (Mendizabal et al., 2008), the Dominican Republic and Honduras, but it is not present in South America (Martínez-Cruzado: Congreso Latinoamericano de Antropología Biológica, Bogotá, 2010). Haplotype 16111-16129-16223- 16290-16319-16362 is found in the Dominican Republic and Puerto Rico (Martínez-Cruzado: Congreso Latinoamericano de Antropología

Biológica, Bogotá, 2010) but is especially common among Chibcha-speaking tribes in northern Colombia (Melton et al., 2007). Because this haplogroup is probably the most ancient in Aruba, this could suggest that the migration occurred from northwestern South America to the Caribbean. The Puerto Rican Taíno language belonged to the Arawakan subfamily of northwestern South America, close to Aruba (Granberry and Vescelius 2004), indicating that in this case, there is a relationship between language and genetics. In Puerto Rico, mutation 16259 is part of a haplotype sequence that defines an archaic sequence (Martínez-Cruzado, 2010) that is not found in South America, but it is found in Honduras (Martínez-Cruzado: Congreso Latinoamericano de Antropología Biológica, Bogotá, 2010). Unlike Puerto Rico, where mutation 16259 is part of a root haplotype, in Aruba this mutation is derived and suggests a non-informative parallelism. The haplotype 16111-16213-16223-16290-16319-16362 is found in Cuba and Colombia (Mendizabal et al., 2008; Salas et al., 2008). The mutation 16213 is found in Guahibo Amerindians in Venezuela (Vona et al., 2005), sharing an ancient lineage with Aruba. According to Kemp et al. (2010), who made a study with North, Central and South American samples, this latter haplotype is only found in South America. Thus, the presence of this and the aforementioned haplogroup A2 haplotypes in the Caribbean is suggestive of extensive migration processes from South America to all the Greater Antilles. The fact that Cuba and Aruba have this haplotype, suggests a migration route from South America to the Caribbean for haplogroup A2.

Haplogroup B2

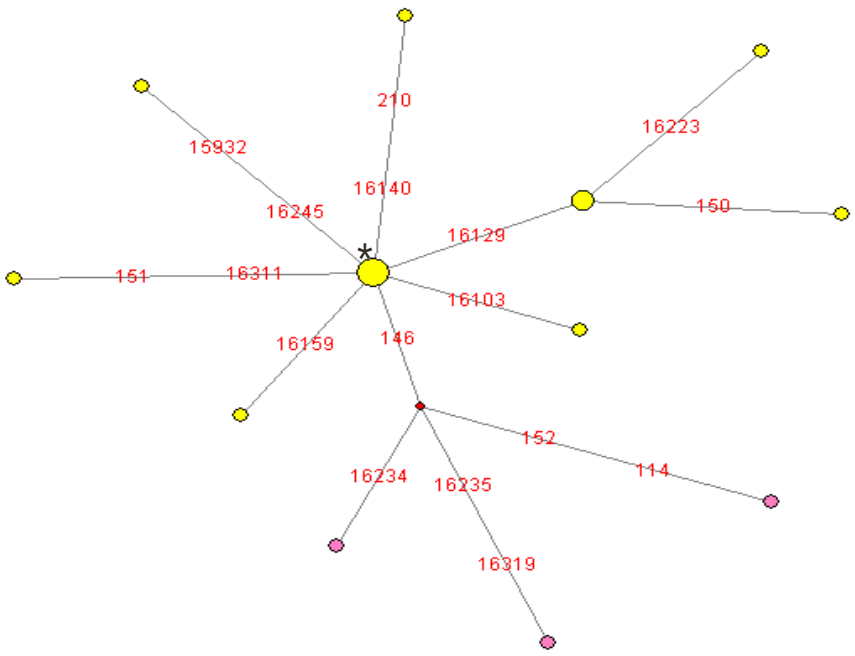

Figure 5. Median Joining Network of Haplogroup B for HVR I – II. (15900-16450 + 1-396)
Lineages B2-I (yellow) and B2-II (pink)
*16189-16217-73-263

Haplogroup B2 shows considerable diversity as well ($\theta = 0.0053$), as most haplotypes in this haplogroup are found each in a single instance (Figure 5). Its Tajima's D-test value is negative, but according to Tajima et al. (1989), it is not a significant value. Despite this value is not significant, the star-like topology in the network suggests a slight founder effect followed by a population expansion. This means that a small group of people migrated from a larger population taking with them a small number of haplotypes and the population size increased with just those haplotypes. Their most common recent ancestor is dated to about 10756 years ago, is one of the youngest among the four haplogroups. It can be concluded that this haplogroup represent relatively recent migrations in Aruba. In spite of the fact that the most common recent ancestor seems recent; it is one of the haplogroups with the highest diversity, according to

parameter θ. This is shown in the median network and with the θ, which reflect considerable diversity. Lineage B2-I may represent the native lineage for this haplogroup in Aruba. Its most recent common ancestor is dated about 9058 year ago, being one of the older lineages in the island. For this reason, it can be suggested that haplogroup B may represent one of the original Amerindian populations in Aruba.

The haplotype A16183C-16189-16217-16235-16319 is found only in Colombia (Salas et al., 2008). Another haplotype very common in Colombia and South America and less frequent in the rest of the Americas is A16183C-16189-16217, which is the founder of the New World and it is found in Aruba as well. Colombia is close to Aruba, there is evidence that Aruban Amerindians traded with Colombian Amerindians (Oliver, 1989), and the share of haplotypes may suggest that migrations from Colombia to Aruba and vice versa could have occurred. In several of the samples, the HVR II could be read up to position 500, showing that some of them have a deletion at position 498. This deletion is found in most haplogroup B mtDNAs in Colombia (Díaz-Matallana and Martínez-Cruzado, 2010), and in almost all of them in Honduras (Martínez-Cruzado: Congreso Latinoamericano de Antropología Biológica, Bogotá, 2010). However, it is not part of the founder haplotype of the New World, and its presence in some Aruban samples where it could be detected expands the geographic region in which this lineage is present. Interestingly, it is not present in the Antilles or the Amazon, and may represent a population expansion in a region encompassing Colombia and Central America.

Haplogroup C

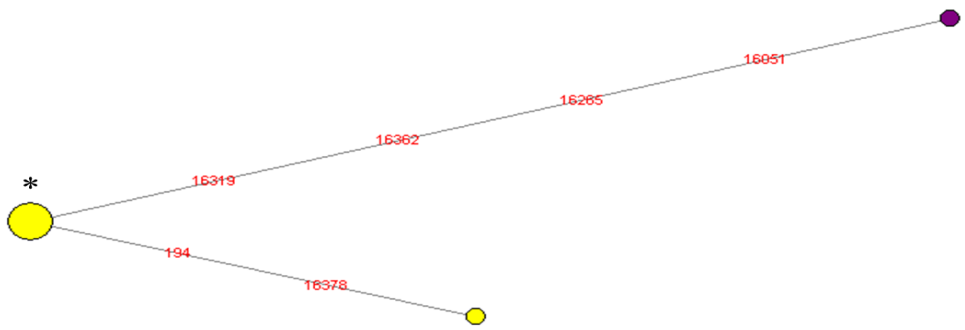

Figure 6. Median Joining Network of Haplogroup C1 for HVR I – II. (15900-16450 + 1-396)
Lineages C1-I (yellow) and C1-II (purple)
*16051-16223-16298-16325-16327-73-249del-263-290del-291del

Haplogroup C1 has the least diversity among the four haplogroups with $\theta = 0.0026$. This is shown in the median network (Figure 6). This haplogroup has two lineages with a grand total of only three haplotypes; it is also the haplogroup with the lowest frequency. The Tajima's D-test resulted with a non-significant value (Table 4), indicating that this haplogroup did not have a population expansion. Their most common recent ancestor is dated to about 7764 years ago, but lineage C1-I, which has the higher frequency between both lineages, is dated about 3019 years ago. This lineage represents a recent migration for this haplogroup.

The majority of the samples have the root haplotype 16051-16223-16298-16325-16327, which probably represent sub-haplogroup C1d, a sub-haplogroup dispersed in South America. This root haplotype was found in archaic remains in Argentina and Uruguay (Figueiro et al., 2009). One haplogroup C1 haplotype 16298-16325-16327-16362 from a sample sequenced from 16090 to 16400 has been found in the Dominican Republic (Lalueza-Fox et al., 2001) and the Guahibo Amerindians from Venezuela (Vona et al., 2005). Guahibos from Venezuela live in the

Amazonian States of this country. The mutation at position 16051 is also found in lineage C1b, which is very common in the New World and it is found in Dominican Republic. In order to confirm if Aruban haplogroup C belong to the lineage C1d, the mutation 493 has to be identified.

Haplogroup D

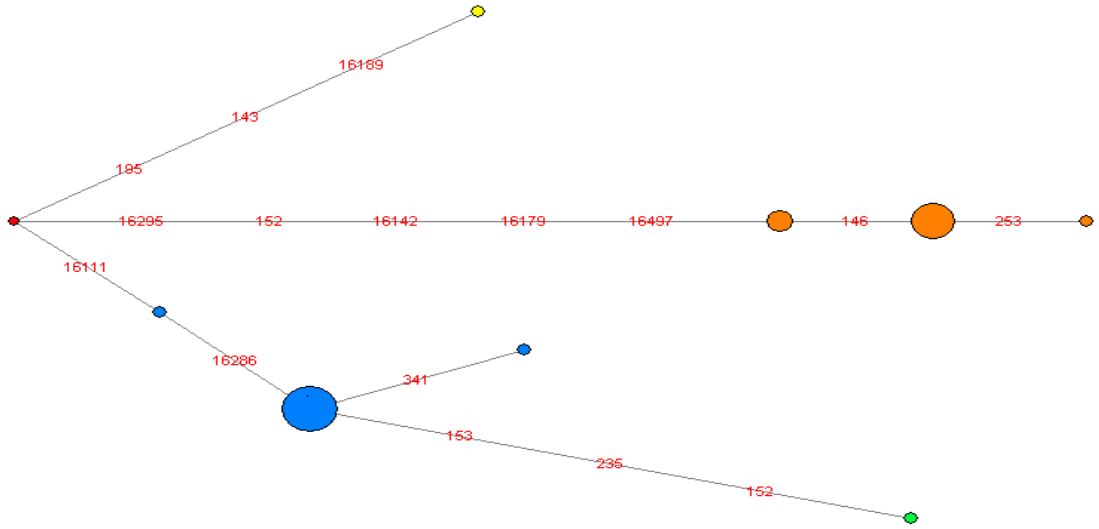

Figure 7. Median Joining Network of Haplogroup D1 for HVR I – II. (15900-16450 + 1-396)
Lineages: D1-I (orange), D1-II (blue), D1-III (green) and D1-IV (yellow)
*16111-16223-16325-16362-73-146-152- 263

Haplogroup D1 is the predominant haplogroup in the Aruban population, but it has a low genetic diversity (θ = 0.0040). It has four lineages, but the majority of the samples were clustered in D1-I and D1-II (Figure 7). Both lineages are separated from others by several mutations and have low genetic diversity as haplogroup C1, suggesting recent migrations for these haplogroups. The difference is that the lineages in haplogroup D1 are in high frequency, suggesting a massive recent migration. The Most Recent Common Ancestor is dated to about 29721 years. It is interesting to notice that this is the most ancient among the four haplogroups.

The Tajima's D-test has a small positive non-significant value. This positive value could happen because the majority of the population is divided into two major lineages and this can resemble a population subdivision. This may also be the reason behind such an old estimate for this haplogroup. The haplogroup age is derived from ρ, an estimator designed for star-like phylogenies where Tajima's D values are expected to be negative (Forster et al., 1996). The very old age estimated for this haplogroup is thus an artifact. The most recent common ancestor was calculated for each lineage, D1-I is dated approximately 2787 years ago and D1-II is dated about 1066 years ago. These ages suggest the recent migration for this haplogroup, because they are one of the youngest among the four haplogroups.

The lineage D1-I has been found in Puerto Rico and may also be found in the Apalaí Amerindians in Brazil (they sequenced until position 16368 and Aruban samples have the mutation 16497) (Mazieres et al., 2007). This may suggest a migration route from South America to the Antilles. Haplogroup D1 has no diversity in Puerto Rico, but it has several haplotypes in Aruba. Therefore, it can be suggested that this lineage is older in South America and came from South America to this island. However, lineage D1-I is not only rare, but has a very limited geographic distribution in Puerto Rico, circumscribed to a region with a local harbor and signs of high immigration rate, as attested by an African mtDNA frequency higher than its surroundings (Martínez-Cruzado et al., 2005; Martínez-Cruzado, 2010). It is thus possible that the presence of D1-I in Puerto Rico reflects only post-Columbian migrations. Unlike Aruban Amerindians who spoke an Arawak language, the Apalaí Amerindian language belongs to the Karib family of languages. The fact that Aruban and Apalaí Amerindians do not share the same family of languages could mean that their genetic relationship is more dependent of geographical proximity.

The lineage D1-II is in high frequency, and it was found in northwestern Venezuela, specifically in the town of Lara (Castro de Guerra, Congreso Latinoamericano de Antropología Biológica, Colombia 2010). It is worth noticing that Macuquita town in this Peninsula is the one that grouped together with Aruba in a principal component analysis made by Castro de Guerra et al. (2009) based on RFLP data. These facts suggest that the lineage D1-II might have come from Paraguana Peninsula in Venezuela. These results are in accordance with Toro-Labrador et al. (2003) and Castro de Guerra et al. (2009). They suggested a relationship between Aruban and Venezuelan Amerindians.

CONCLUSION

Aruba has a strong maternal Amerindian ancestry, being haplogroup D the predominant among the four haplogroups. It is remarkable that this haplogroup does not represent the original population in Aruba. Haplogroups A and B are the most ancient in this island, and better represent the original population. Both haplogroups have lineages that may have arrived simultaneously and that may have been represented by more than one founder haplotype. Both haplogroups could be affected by genetic drift due to the movement of Aruban Amerindians back and forth by the Spaniards and then the Dutch colonization. The haplogroup A is related with Venezuela and Colombia, this haplogroup is very common in Venezuela, therefore, it can be suggested that most of the migrations related to haplogroup A occurred from this country. The relationship between Aruba and Venezuela can be demonstrated with archeological evidence. The ceramic cultures in Aruba and Venezuela shared similar elements in their pottery, and it is

believed that they shared a common ancestral culture. The ancestral culture may have existed between the Turbio, Upper Cojedes, and Upper Yaracuy rivers then expanding their tradition to Venezuela, Aruba, Curacao, and Bonaire (Oliver, 1989). By contrast, haplogroup B is more common in Colombia and share several haplotypes with Aruba. There is archeological evidence that Aruban and Colombian Amerindians maintained close contact. For example, avemorphic motifs found in pottery in Aruba are not related with Aruba and Venezuela cultures, but it is related with La Guajira peninsula in Colombia (Oliver, 1989). In conclusion, the most ancient Aruban haplotypes maintain a close relationship with the Northwest South American region of La Guajira and Northwest Venezuela

Haplogroups C and D seem to represent recent migrations in Aruba. The latter haplogroup is predominant in this population, suggesting massive recent migrations rich in haplogroup D. These migrations could represent pre-Colombian migrations or it may have occurred at the time the Spaniards transported the Aruban Amerindians to the Hispaniola, leaving the island empty. When the Spaniards returned to Aruba, they found Amerindians in the island (Versteeg et al., 1995). During that time, people from South America may have come to the island. When Aruba was conquered by the Dutch, they deported the Amerindians to mainland and then the Dutch stimulated migrations from mainland to the island because they were good at breeding horses (Hartog et al., 1961). In the case of the haplogroup D, it can be implied that a big group of Amerindians belonging to this haplogroup migrated from the mainland at some time in history and stayed in Aruba. Aruba shares lineage D-II with the town of Lara in Venezuela and lineage D-I with Apalaí Amerindians in Brazil. Thus, both haplogroup D lineages are within a region that is Amazonian or whose ceramic culture suggests close contacts with Amazonian groups it may suggest that the majority of the Aruban modern population came from Venezuela

and the Amazonian part of Brazil. In the case of haplogroup C, Aruba does not share any haplotypes with Puerto Rico, and share only one low frequency haplotype with Dominican Republic. It can be concluded that no direct relationship is found for this haplogroup among these islands, but it may have close relationship with South America.

The haplogroups C and D share haplotypes with non-Arawakan Amerindians, but they are geographically close to Aruba, for example the Amazonian part of Brazil and Venezuela. In these cases the family of languages is not related with the genetics of this haplogroups. The strongest relationship is between genetics and geographical proximity.

The modern population in Aruba constituted approximately 25% of immigrants that were established in 1924 with the arrival of the refinery. Those immigrants came from the British islands in the Caribbean. Another 25% are immigrants from Colombia, Venezuela, Peru and the Dominican Republic. In general terms, because of this, haplotypes with low frequency and farther from native lineages can be expected. This is reflected in the Median Joining Networks, where very recent haplotypes and lineages are shown. For example, lineages A2-I, A2-II and A2-III in haplogroup A may represent post-Columbian migrations. These migrations may have occurred in the 1900's. The migration of the lineage B2-I may also have occurred in this period. All haplogroup C might represent migrations in the 1900's as well, because of the low frequency and diversity. In the case of haplogroup D, both lineages of which are in high frequency, it is less probable that the migration occurred in that century. It is more possible that the migrations occurred at the time of Spaniards or Dutch colonization in order to have sufficient time to disperse the haplogroup in the island at a time the population of the island had been drastically reduced.

The results of this work demonstrated a strong relationship between Aruba and South America, and some relationship with the Caribbean. It brings evidence that at least one migration event involving one haplogroup D1 lineage occurred from South America to the Caribbean. In the case of Aruba, despite the intensive movement of Amerindians back and forth to Aruba, it is still possible to trace their heritage in the modern population.

RECOMMENDATIONS

- To sequence the HVR II in haplogroup B2 from position 400 to 500 in order to verify if all the samples have the deletion in the position 498. This will give a better understanding of whether this haplogroup came from Colombia or another country.

- To trace the ancestry of the mothers and grandmothers of the samples studied in Aruba to to have a better idea of where they come from.

- To make mtDNA Control Region sequencing in Bonaire and Curacao, in order to study their haplogroups patterns and compare them with Aruba to see if their patterns are similar.

- To sequence Amazonian populations in Brazil and southern Venezuela in order to have a better understanding of the beginning of the first migrations from this area to the northwestern part of South America and the Caribbean.

CITED LITERATURE

Achilli, A., Perego, U.A., Bravi, C.M., Coble, M.D., Kong, Q.P., Woodward, S.R., Salas, A., Torroni, A., Bandelt, H.J. (2008) "The Phylogeny of the Four Pan-American MtDNA Haplogroups: Implications for Evolutionary and Disease Studies." *Plos One*, Vol. 3, No. 3, e1764

Alves-Silva, J., da Silva-Santos, M., Guimaraes, P., Ferreira, A., Bandelt, H-J., Pena, S., Ferreira-Pradro, V. (2000) "The Ancestry of Brazilian mtDNA Lineages." *American Journal of Human Genetics*, Vol. 67, pp. 444-461

Brown,W.M., Matthew, G. Jr., Wilson, A.C. (1979) "Rapid evolution of animal mitochondrial DNA". *Genetics*, Vol. 76, No. 4, pp. 1967-1971

Carvajal-Carmona, L., Soto, I. D., Pineda, N.; Ortíz-Barrientos, D., Duque, C., Ospina-Duque, J., McCarthy, M., Montoya, P., Álvarez, V. M., Bedoya, G., Ruiz-Linares, A. (2000) "Strong Amerind/White Sex Bias and a Possible Sephardic Contribution among the Founders of a Population in Northwest Colombia". *The American Journal of Human Genetics*, Vol. 67, pp. 1287–1295

Castro de Guerra, D., Figueroa-Pérez, C., Helen Izaguirre, M., Rodríguez-Larralde, A., Guerra-Castro, E., Martínez-Méndez, D., Pujol, F. (2009) "Mitochondrial diversity in Northwest Venezuela. Implications for probable prehispanic migratory routes". *Acta Biológica Colombiana*, Vol. 14, No. 1, pp. 173-184

Derenko, M., Malyarchuk, B., Grzybowski, T., Denisova, G., Dambueva, I., Perkova, M., Dorzhu, C., Luzina, F., Lee, H.K., Vanecek, T., Villems, R., Zakharov, I. (2007) "Phylogeographic Analysis of Mitochondrial DNA in Northern Asian Populations". *The American Journal of Human Genetics*, Vol. 81, No. 5, pp. 1025–1041

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 Bogotá Colombia*, Vol. 51, No.3, pp. 241-272

Easton, R.D., Merriwether, A., Crews, D.E., Ferell, R.E. (1996) "mtDNA Variation in the Yanomami: Evidence for Additional New World Founding Lineages." *American Journal of Human Genetics*, Vol. 59, pp. 213-225

Excoffier, L. G., Laval, Schneider, S. (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47-50.

Excoffier, L., Langaney, A. (1989) "Origin and Differentiation of Human Mitochondrial DNA". *The American Journal of Human Genetics*, Vol. 44, pp. 73-85

Fagundes, N., Kanitz, R., Eckert, R., Valls, A., Bogo, M.R., Salzano, F.M., Smith, D.G., Silva Jr.,W.A., Zago, M.A., Ribeiro-dos-Santos, A. K., Santos, S., Petzl-Erler, M.L., Bonatto, S.L. (2008) "Mitochondrial Population Genomics Supports a Single Pre-Clovis Origin with a Coastal Route for the Peopling of the Americas". *The American Journal of Human Genetics*, Vol. 82, pp. 583-592

Figueiro, G., Ackermann, E., Hidalgo, P., Politis, G., Sans, M. (2009) "Antigüedad y dispersión del linaje C1d-16051G en America del Sur." *Actas de las Novenas Jornadas Nacionales de Antropología Biológica*, pp. 105

Fuselli, S., Tarazona-Santos, E., Dupanloup, I., Soto, A., Luiselli, D., Pettener, D. (2003) "Mitochondrial DNA Diversity in South America and the Genetic History of Andean Highlanders." *Molecular Biology and Evolution*, Vol. 20, No.10, pp. 182-1691

Granberry, J., Vescelius, G.S. (2004) "Languages of the Pre-Columbian Antilles". The University of Alabama Press, Tuscaloosa, pp. 54-57

Hartog, J., Verleun, J.A. (1961) "Aruba past and present, from the time of the Indians until today." D.J. DeWit Publisher, Aruba

Henríquez, H., Moraga, M., Llop, E., Rothhammer, F. (2004) "Caracterización genético molecular de habitantes de Caleta Paposo, último reducto Chango en Chile". *Revista Médica de Chile*, Vol. 132, No. 6, pp. 663-672

Herrnstadt, C., Elson, J.L., Fahy, E., Preston, G., Turnbull, D.M., Anderson, C., Ghosh, S.S., Olefsky, J.M., Beal, M.F., Davis, R.E., Howell, N. (2002) "*Reduced-Median-Network Analysis of Complete Mitochondrial DNA Coding-Region Sequences for the Major African, Asian, and European Haplogroups*". The American Journal of Human Genetics, Vol.70, No.5, pp. 1152-1171

Jobling, M. A., Hurles, M., Tyler-Smith, C. (2004). "Colonization of the Old World and Australia" in *Human Evolutionary Genetics: Origins, Peoples & Disease*. Garland Publishing, New York, pp. 289-290

Jobling, M. A., Hurles, M., Tyler-Smith, C. (2004). "Into New Found Lands" in *Human Evolutionary Genetics: Origins, Peoples & Disease*. Garland Publishing, New York, pp. 341-352

Keyeux, G., Rodas C., Gelvez N., Carter D. (2002) "Possible migration routes into South America deduced from mitochondrial DNA studies in Colombian Amerindian populations". *Human Biology*, Vol. 74, No. 2, pp. 211-233

Kolman, C.J., Sambuughin, N., Bermingham, E. (1996) "Mitochondrial DNA Analysis of Mongolian Populations and Implications for the Origin of New World Founders." *Genetics*, Vol. 142, pp. 1321-1334

Lalueza-Fox, C., Luna Calderon, F., Calafell, F., Morera, B., Bertranpetit, J. (2001) "MtDNA from extinct Tainos and the peopling of the Caribbean". *Annals of Human Genetics*, Vol. 65, pp. 137-151

Lewis, C.M., Lizárraga, B., Tito, R.Y., López, P.W., Iannacone, G.C., Medina, A., Martínez, R., Polo, S.I., De La Cruz A.F., Cáceres, A.M., Stone, A.C., (2007) "Mitochondrial DNA and the peopling of South America". *Human Biology*, Vol. 79, No. 2, pp. 159-178

Martínez-Cruzado, J.C. (2002) "El uso del ADN mitocondrial para descubrir las migraciones precolombinas al Caribe: Resultados para Puerto Rico y expectativas para la República Dominicana". *KACIKE: Journal of Caribbean Amerindian History and Anthropology* [Revista Electrónica], Edición Especial.

Martínez-Cruzado, J.C., Toro-Labrador, G., Ho-Fung, V., Estévez-Montero, M.A., Lobaina-Manzanet, A., Padovani-Claudio, D.A., Sánchez-Cruz, H., Ortiz-Bermúdez, P., Sánchez-Crespo, A. (2001) "Mitochondrial DNA analysis reveals substantial Native American ancestry in Puerto Rico". *Human Biology*, Vol. 73, No. 4, pp. 491-511

Martínez-Cruzado, J.C., Toro-Labrador, G., Viera-Vera, J., Rivera-Vega, M. Y., Startek, J., Latorre-Esteves, M., Román-Colón, A., Rivera-Torres, R., Navarro-Millán, I.Y., Gómez-Sánchez, E., Caro-González, H. Y., Valencia-Rivera, P. (2005) "Reconstructing the Population History of Puerto Rico by Means of mtDNA Phylogeographic Analysis". *American Journal of Physical Anthropology*, Vol. 128, pp. 131-155

Martínez, H., Rodríguez-Larralde, A., Helen-Izaguirre, M., Castro de Guerra, D. (2007) "Admixture Estimates for Caracas, Venezuela, Based on Autosomal, Y-Chromosome, and mtDNA Markers". *Human Biology*, Vol. 79, No. 2, pp. 201-213

Mazieres, S., Guitard, E., Crubezy, E., Dugoujon, J.M., Bortolini, M.C., Bonatto, S.L., Hutz, M.H., Bois, E., Tiouka, F., Larrouy, G., Salzano, F.M. (2007) "Uniparental (mtDNA, Y-chromosome) Polymorphisms in French Guiana and Two Related Populations-Implications for the Region's Colonization." *Annals of Human Genetics*, Vol. 71, pp. 1-12

Mendizabal, I., Sandoval, K., Berniell-Lee, G., Calafell, F., Salas, A., Martínez-Fuentes, A., Comas, D. (2008) "Genetic origin, admixture, and asymmetry in maternal and paternal human lineages in Cuba". *Evolutionary Biology*, Vol. 8, pp.1-10

Melton, P.E., Briceño, I., Gómez, A., Devor, E.J., Bernal, J.E., Crawford, M.H. (2007) "Biological relationship between Central and South American Chibchan speaking populations: evidence from mtDNA". *American Journal of Physical Anthropology*, Vol. 133, No. 1, pp. 753-770

Nei, M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York

Oliver, J.R. (1989) “The archeological, linguistic and ethno-historical evidence for the expansion of Arawakan into northwestern Venezuela and northeastern Colombia.” Ph.D. thesis, University of Illinois, Urbana-Champaign.

Oliver, J.R. (1995) “The archaeology of Maticora, western Venezuela: a prehistoric longhouse structure in site Fal-7.” Proceedings of the Fifteenth International Congress for Caribbean Archaeology, pp. 1-52

Roosevelt, A. C., Lima da Costa, M., Lopes Machado, C., Michab, M., Mercier, N., Valladas, H., Feathers, J., Barnett, W., Imazio da Silveira, M., Henderson, A., Silva, J., Chernoff, B., Reese, D.S., Holman, J.A., Toth, N., Schick, K. (1996) “Paleoamerindian Cave Dwellers in the Amazon: The Peopling of the Americas.” *Science*, Vol. 272, No. 5260, pp. 373-384

Salas, A., Acosta, A., Alvarez,-Iglesias, V., Cerezo, M., Phyllips, C., Lareu, M.V., Carracedo, A. (2008) “The mtDNA Ancestry of Admixed Colombian Populations.” *American Journal of Human Genetics*, Vol. 20, pp. 584-591

Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Rohl, A., Salas, A., Oppenheimer, S., Macaulay, W., Richards, M. (2009) “Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock”. *The American Journal of Human Genetics*, Vol. 84, No. 6, pp.740-759

Tajima, F. (1989) “Statistical Method for Testing the Neutral Mutation Hypothesis by DNA”. *Genetics*, Vol. 123, pp. 585-595

Toro-Labrador, G., Wever, O.R., Martínez-Cruzado, J.C. (2003) “Mitochondrial DNA Analysis in Aruba: Strong Maternal Ancestry of Closely Related Amerindians and Implications for the Peopling of Northwestern Venezuela”. *Caribbean Journal of Science*, Vol. 39, pp. 11-22

Torroni, A., Schurr, T.G., Yang, C., Szathmary, E., Williams, R.C., Schanfield, M.S., Troup, G.A., Knowler, W.C., Lawrence, D.N., Weiss, K.M., Wallace, D.C. (1992) “Native American Mitochondrial DNA Analysis indicates that the Amerind and the Nadene Populations were founded by Two Independent Migrations.” *Genetics*, Vol. 130, pp. 153-162.

Torroni, A., Schurr,T.G., Cabell, M.F., Brown, M.D., Neel, J.V., Larsen, M., Smith, D.G., Vullo, C.M., Wallace, D.C. (1993) “Asian Affinities and Continental Radiation of the Four Founding Native American mtDNAs”. *American Journal of Human Genetics*, Vol. 53, pp. 563-590

Torroni, A., Neel, J.V., Barrantes, R., Schurr, T.G., Wallace, D.C. (1994a) “Mitochondrial DNA "clock" for the Amerinds with its implications for timing their entry into North America.” *PNAS*, Vol. 91, No. 3, pp. 1158-1162

Versteeg, A.H., Ruiz, A.C. (1995). “Reconstructing brasilwood island: the archaeology and landscape of Indian Aruba”. *Publications of the Archaeological Museum of Aruba*, No. 6

Vona, G., Falchi, A., Moral, P., Calo, C., Varesi, L. (2005) "Mitochondrial Sequence Variation in the Guahibo Amerindian Population from Venezuela." *American Journal of Physical Anthropology,* Vol. 127, pp. 361-369

Wallas, D.C., Torroni, A. (2009) "American Indian Prehistory as Written in the Mitochondrial DNA: A Review" *Human Biology,* Vol. 81, No. 5-6, pp. 509-521

Waters, M.R., Stafford Jr. T.W. (2007) "Redefining the Age of Clovis: Implications for the Peopling of the Americas". *Science*, Vol. 315, pp. 1122-1126

APPENDIX

Table 6. HVR-I and HVR-II mutations in Haplogroup A used to construct the Median Network

Haplotype Frequency	15924	15930	16093	16111	16129	16136	16175	16187	16189	16213	16223	16259	16290	16300	16311	16319	16362	64	73	125	127	146	152	153	194	207	235	263
rCRS	A	G	T	C	G	T	A	C	T	G	C	C	C	A	T	G	T	C	A	T	T	T	T	A	C	G	A	A
1	.	.	.	T	.	.	G	.	.	.	T	.	T	G	.	A	C	T	G	.	.	C	.	G	.	.	G	G
1	.	.	.	T	.	.	.	.	.	A	T	.	T	.	.	A	C	T	G	.	.	C	.	G	.	.	G	G
1	.	A	.	T	.	.	.	.	.	.	T	.	T	.	.	A	C	.	G	.	.	C	.	G	.	.	G	G
1	.	.	C	T	.	C	.	.	.	.	T	.	T	.	C	A	C	T	G	.	.	C	.	.	.	.	G	G
1	G	.	.	T	.	.	.	.	.	.	T	.	T	.	.	A	C	T	G	.	.	C	.	G	.	.	G	G
1	.	.	.	T	.	.	G	.	.	.	T	T	T	G	.	A	C	T	G	.	.	C	.	.	.	.	G	G
1	.	.	.	T	.	.	.	.	.	A	T	.	T	.	.	A	C	T	G	C	C	C	.	G	T	.	G	G
2	.	.	.	T	A	.	.	.	.	.	T	.	T	.	.	A	C	T	G	.	.	C	.	G	.	.	G	G
1	.	.	.	T	.	.	.	.	.	.	T	.	T	.	.	A	C	T	G	.	.	C	.	G	.	A	G	G
1	.	.	.	T	.	.	.	T	C	.	T	.	T	.	.	A	C	T	G	C	C	C	C	G	.	.	G	G
1	.	.	.	T	.	.	.	.	.	.	T	.	T	.	.	A	C	T	G	.	.	C	.	G	.	.	G	G
1	.	.	.	T	.	.	.	.	.	.	T	.	T	.	.	A	C	.	G	.	.	C	.	G	.	.	G	G
1	.	.	.	T	A	.	.	.	.	.	T	.	T	.	.	A	C	T	G	.	.	C	.	.	.	.	G	G

Table 7. HVR-I and HVR-II mutations in Haplogroup B used to construct the Median Network

Haplotype Frequency	15932	16086	16103	16129	16140	16159	16189	16217	16234	16235	16245	16311	16319	73	114	146	150	152	210	263
rCRS	T	T	A	G	T	C	T	T	T	A	C	T	G	A	C	T	C	T	A	A
4	.	.	.	.	.	.	C	C	.	.	.	.	.	G	.	.	.	.	.	G
3	.	.	.	A	.	.	C	C	.	.	.	.	.	G	.	.	.	.	.	G
1	C	.	.	.	.	.	C	C	.	.	T	.	.	G	.	.	.	.	.	G
1	.	C	.	.	.	.	C	C	.	.	.	.	.	G	T	C	.	C	.	G
1	.	.	G	.	.	.	C	C	.	.	.	.	.	G	.	.	.	.	.	G
1	.	.	.	A	.	.	C	C	.	.	.	.	.	G	.	.	T	.	.	G
1	.	.	.	.	C	.	C	C	.	.	.	.	.	G	.	.	.	.	G	G
1	.	.	.	.	.	T	C	C	.	.	.	.	.	G	.	.	.	.	.	G
1	.	.	.	.	.	.	C	C	.	G	.	.	A	G	.	C	.	.	.	G
1	.	.	.	.	.	.	C	C	.	.	.	C	.	G	.	.	.	.	.	G
1	.	.	.	.	.	.	C	C	C	.	.	.	.	G	.	C	.	.	.	G

Table 8. HVR-I and HVR-II mutations in Haplogroup C used to construct the Median Network

Haplotype Frequency	**16051**	**16223**	**16265**	**16298**	**16319**	**16325**	**16327**	**16362**	**16378**	**73**	**194**	**249del**	**263**	**290del**	**291del**
rCRS	**A**	**C**	**A**	**T**	**G**	**T**	**C**	**T**	**C**	**A**	**C**	**A**	**A**	**A**	**A**
5	G	T	.	C	.	C	T	.	.	G	.	-	G	-	-
1	.	T	G	C	A	C	T	C	.	G	.	-	G	-	-
1	G	T	.	C	.	C	T	.	T	G	T	-	G	-	-

Table 9. HVR-I and HVR-II mutations in Haplogroup D used to construct the Median Network

Haplotype Frequency	**16111**	**16142**	**16179**	**16189**	**16223**	**16286**	**16295**	**16325**	**16362**	**16497**	**73**	**143**	**146**	**152**	**153**	**195**	**235**	**253**	**263**	**341**
rCRS	**C**	**C**	**C**	**T**	**C**	**C**	**C**	**T**	**T**	**A**	**A**	**G**	**T**	**T**	**A**	**T**	**A**	**C**	**A**	**A**
1	T	.	.	.	T	.	.	C	C	.	G	.	C	C	.	.	.	.	G	.
15	T	.	.	.	T	T	.	C	C	.	G	.	C	C	.	.	.	.	G	.
1	T	.	.	.	T	T	.	C	C	.	G	.	C	C	.	.	.	.	G	G
1	T	.	.	.	T	T	.	C	C	.	G	.	C	.	G	.	G	.	G	.
3	.	T	T	.	T	.	T	C	C	G	G	.	.	.	.	.	.	.	G	.
9	.	T	T	.	T	.	T	C	C	G	G	.	.	.	.	.	.	.	G	.
1	.	T	T	.	T	.	T	C	C	G	G	.	.	.	.	.	.	T	G	.
1	.	.	.	C	T	.	.	C	C	.	G	A	C	C	.	C	.	.	G	.