Genetic Diversity in Puerto Rico and Its Implications for the Peopling of the Island and the West Indies

Miguel G. Vilar,[1] Carlalynne Melendez,[2] Akiva B. Sanders,[1] Akshay Walia,[1] Jill B. Gaieski,[1] Amanda C. Owings,[1] Theodore G. Schurr,[1]* and The Genographic Consortium

[1]*Department of Anthropology, University of Pennsylvania, Philadelphia, PA 19104-6398*
[2]*National Geographic Society, IBM and the Waitt Family Foundation (Genographic Project Grant), and the University of Pennsylvania. Liga Guakia Taina-Ke, Humacao, Puerto Rico*

ABSTRACT   Puerto Rico and the surrounding islands rest on the eastern fringe of the Caribbean's Greater Antilles, located less than 100 miles northwest of the Lesser Antilles. Puerto Ricans are genetic descendants of pre-Columbian peoples, as well as peoples of European and African descent through 500 years of migration to the island. To infer these patterns of pre-Columbian and historic peopling of the Caribbean, we characterized genetic diversity in 326 individuals from the southeastern region of Puerto Rico and the island municipality of Vieques. We sequenced the mitochondrial DNA (mtDNA) control region of all of the samples and the complete mitogenomes of 12 of them to infer their putative place of origin. In addition, we genotyped 121 male samples for 25 Y-chromosome single nucleotide polymorphism and 17 STR loci. Approximately 60% of the participants had indigenous mtDNA haplotypes (mostly from haplogroups A2 and C1), while 25% had African and 15% European haplotypes. Three A2 sublineages were unique to the Greater Antilles, one of which was similar to Mesoamerican types, while C1b haplogroups showed links to South America, suggesting that people reached the island from the two distinct continental source areas. However, none of the male participants had indigenous Y-chromosomes, with 85% of them instead being European/Mediterranean and 15% sub-Saharan African in origin. West Eurasian Y-chromosome short tandem repeat haplotypes were quite diverse and showed similarities to those observed in southern Europe, North Africa and the Middle East. These results attest to the distinct, yet equally complex, pasts for the male and female ancestors of modern day Puerto Ricans. Am J Phys Anthropol 155:352–368, 2014.   © 2014 Wiley Periodicals, Inc.

The Caribbean region is of great interest to anthropologists because it bore witness to pre-Columbian human migrations and interactions for several millennia, and European colonization and forced movement of African peoples to the area over the past several centuries. Like a series of stepping-stones connecting the continents of North and South America (Fig. 1), these islands and their living inhabitants are key to unraveling the complex processes and timing of the peopling of the region. In the center, and at the crossroads of the Caribbean, rests the island of Puerto Rico. Puerto Rico is the smallest of the islands that comprise the Greater Antilles, which extend west through Cuba, although it is also geographically larger than all of the neighboring Lesser Antilles to the east and south of it. As a consequence, the pre-colonial history of Puerto Rico is closely linked to that of both Antilles archipelagos and the rest of the Caribbean islands.

For six millennia, Puerto Rico was home to various indigenous groups. The last of these was the Taíno, an Arawakan speaking horticultural people who occupied the island and neighboring islets at the time that Christopher Columbus arrived on his second voyage to the Americas in 1493 (Morisson, 1939). Because of its location and size, Puerto Rico was one of the first islands to be colonized by Spain and later transferred, along with Cuba, Guam and the Philippines, to the United States by way of the Treaty of Paris of 1898 that ended the Spanish-American war (Musicant, 1998). Today, Puerto Rico remains a Spanish-speaking United States protectorate, with a deep genetic connection and a growing cultural tie to its indigenous Caribbean past.

Most anthropologists agree that multiple waves of peoples traversed and settled the Caribbean (Rouse, 1992), contributing to the diverse cultural and linguistic traditions noted by historians in the late 15th century (Las Casas, 1992). Undoubtedly, the ancestry of Caribbean peoples was further complicated by the arrival of migrants from Europe, Africa and Asia over the last 500 years (Rogozinski, 1999). The origins of these indigenous populations, the timing and directionality of pre-colonial migrations, and the biological and cultural connections to different continental and source populations are issues that are still not well understood.

Fig. 1. A map of the Caribbean Basin with an inset map of Puerto Rico. The sampling locations are marked with filled circles.

Pre-colonial history of the Caribbean

Early Holocene remains from sites in the islands of Cuba and Trinidad, on either end of the Caribbean island archipelago, suggest that human entry into the region may have occurred as early as 7,000–8,000 years ago (Haag, 1964; Willey, 1956; Rouse, 1992; Wilson, 2007). Cuba represents the northwestern extent of the Caribbean chain of islands, and is approximately 100–120 miles from both the continental extensions of the Yucatan Peninsula in Mexico and the Florida Keys. It has been speculated that either one of these areas could have been an early staging ground for expansion into the Caribbean regions for indigenous populations (Keegan, 1995; Wilson, 2007; Fitzpatrick and Ross, 2010).

On the other end of the Caribbean island chain is Trinidad. Parts of Trinidad are located less than 10 miles from the Venezuelan coast, and the island was at one time attached to the mainland. Geologists suggest that Trinidad fully separated from the South American continent between 11,000 and 13,000 years before present (YBP) (Bryan, 1973), thus implying that the island's first inhabitants may have walked there from mainland South America. However, the earliest direct evidence of human occupation on the island is dated to 8,000 YBP, with the remains of the Banwari Culture representing the earliest evidence of human occupation on any Caribbean island (Rouse, 1992).

The earliest evidence of human occupation on Puerto Rico, the most distant Caribbean island from any continental landmass, did not occur for another 2,000 years (Alegria, 1965). The oldest archaeological site on Puerto Rico (Angostura) dates to approximately 6,000 YBP and is thought to be associated with the Ortoiroid culture, a pre-ceramic culture with lithic ties to the Lesser Antilles, specifically Trinidad and northern South America (Rouse, 1992). Related sites have also been found on the Virgin Islands to the east (Rouse, 1952, 1992).

Sites belonging to the distinct Casimiroid culture, which may have emerged in Mesoamerica (Belize and Honduras; Rouse, 1992), have been unearthed on the islands of Cuba and Hispaniola (the island that comprises the modern countries of Dominican Republic and Haiti), and dated to the mid-Holocene (4,000 YBP) (Alegria, 1965; Granberry and Vescelius, 2004). Casimiroid sites have also been uncovered on Mona Island (midway between Hispaniola and Puerto Rico), as well as on the southwestern region of Puerto Rico (Ortiz, 1976), although the Puerto Rican site may only date to 2,000 YBP (Pantel, 1976; Lundberg, 1989). Thus, the island of Puerto Rico may represent an early point of interaction between the Ortoiroid and Casimiroid pre-ceramic cultures of the Caribbean (Rouse, 1992).

A few millennia later, populations originating in South America and possibly ancestral to the Taíno inhabitants

of the Greater Antilles are thought to have migrated north from the Amazon Basin and through the Orinoco River Valley, reaching the West Indies by way of the Guyanas and Venezuela around 2,000 YBP (Rouse, 1964). During this time, Saladoid ceramic cultures reached Hispaniola and Puerto Rico and began expanding into the region (Allaire, 1997; Rouse, 1992), with distinct ceramic traditions arising in the Greater Antilles and the Leeward Islands.

The Taíno arose during this process of transculturation and population interactions around 1,200 YBP. Distinguished by their ceramics, the Taíno practiced horticulture and had a strict hierarchical structure of chiefdoms governed by local leaders known as caciques (Rouse, 1992). They were the dominant culture of the Caribbean when Europeans arrived in the late 15th century CE, and likely numbered in the hundreds of thousands across dozens of islands at that time (Rouse, 1992). In fact the Spanish identified distinct Taíno chiefdoms and caciques across the entire Caribbean region, from Cuba and the Bahamas to Jamaica and the Virgin Islands (Rouse, 1992; Hoogland and Hofman, 1999). Given their distinct cultural features, the Taíno may represent a new wave of migration into the region, as opposed to a cultural shift by an earlier group of Caribbean inhabitants (Rouse, 1992). Their linguistic and cultural ties to other Arawakan-speaking groups in South America (Kaufman, 1994; Campbell, 1997; Granberry, 2013) further indicates that the Taíno likely arrived from that region. Archaeologists have also theorized that the Taíno maintained connections with continental populations in South America during the centuries following their arrival on the islands (Haag, 1964; Rouse, 1992; Fitzpatrick and Keegan, 2007). Thus, at the time of European contact, the Taíno culture may have represented centuries of interactions among populations from the north coast of South America and the Antilles.

Around 800 YBP, Carib-speaking groups originating from the Orinoco River region began moving north through the Lesser Antilles, coming into contact with the culturally distinct Taíno (Allaire, 1980; Forte, 2005; Granberry, 2013). By 600 YBP, the earlier settlers had likely been supplanted or assimilated in most of the Lesser Antilles by expanding Carib populations, and these newcomers were encroaching upon the Virgin Islands and Puerto Rico when the Spanish arrived in the late 15th century (Allaire, 1997). Thus, yet again, Puerto Rico appears to have been a geographic point of interaction, this time between Taíno and Carib populations, both of which were flourishing at the time of European contact. In fact, a number of different indigenous groups were living in the Caribbean at that time, including the Caribs, Taínos, and the Ciboney, the latter being a pre-ceramic culture limited to western Cuba and the southwestern peninsula of Haiti (Rouse, 1992).

Colonial history of the Caribbean

During the two centuries that followed European contact, the colonial powers of Britain, France, Holland, and Spain each laid claim to their own Caribbean colonies and maintained sovereignty over these colonies for several centuries, some even into the present (e.g., the British Virgin Islands and the Dutch Antilles) (Rogozinski, 1999). The variety of languages spoken across the region today attest to a "free for all" attitude that the European colonial powers employed in capturing and colonizing the Caribbean. During the colonization process, all indigenous Caribbean languages were lost, in part due to indigenous groups being forced to use European languages (Rogozinski, 1999). However, linguists have long recognized that indigenous words were gradually adopted by the Spanish colonists of Puerto Rico, the Dominican Republic and Cuba (Rouse, 1992; Wilson, 1997) and are now part of the Caribbean Spanish vernacular.

Furthermore, historical accounts suggest that the European powers governed their colonies in distinct ways (Rogozinski, 1999). For example, over the course of hundreds of years across British colonies such as Jamaica, African slaves replaced the enslaved native islanders, who were either decimated by armed colonists shortly after time of contact, or died from deadly pathogens introduced from Europe and Africa in the decades that followed their arrival (Rouse, 1992; Wilson, 1997; Crosby, 2003; Raudzens, 2003). By contrast, Spain laid claim to the large islands of Cuba, Hispanola, and Puerto Rico, and governed each one differently (Rogozinski, 1999). While Hispanola, and specifically the eastern half known as Dominican Republic, became the Spanish colonial capital in the Caribbean and most slaves from Africa were introduced to the region through there, much of central Cuba and Puerto Rico remained unsettled for centuries (Dietz, 1992). The varied colonial practices across the region differently influenced the cultural and biological composition of populations from each individual island.

During the next few decades, the Spanish crushed several Taíno rebellions, and by the late 16th century they claimed to have all but exterminated the indigenous population from Puerto Rico (Las Casas, 1992). However, most large Spanish settlements on the island were restricted to the coasts, including the peninsula on the north known today as San Juan, with the town of San Germán arising inland in the southwestern region. The mountainous center and the southeastern regions of the island remained mostly unsettled by colonists until the 19th century (Rogozinski, 1999).

Caribbean genetic diversity

Both the pre-Columbian migrations and European conquest and colonization of the Caribbean ultimately left their marks on the gene pools of the region's populations. Earlier genetic surveys of the circum-Caribbean have revealed the presence of Native American, African and West Eurasian mitochondrial DNA (mtDNA) haplogroups among its inhabitants (Martínez-Cruzado et al., 2001, 2005; Benn Torres et al., 2008; Mendizabal et al., 2008; Simms et al., 2010; Gaieski et al., 2011). For some of these islands, the indigenous maternal genetic ancestry exceeds that contributed by non-native populations, despite the historical claim that indigenous Caribbean cultures had been exterminated (Las Casas, 1992). In Puerto Rico (Martínez-Cruzado et al., 2005) and Cuba (Mendizabal et al., 2008), the frequencies of indigenous mtDNAs were 61 and 33%, respectively, whereas 28% of those in Dominica were indigenous in origin (Benn Torres et al., 2008). More limited work has been done in Hispaniola to ascertain the genetic composition of populations there, but estimates suggest that approximately 15% of people in the Dominican Republic and fewer than 5% of people in Haiti have indigenous mtDNAs (Torroni et al., 2001; Simms et al., 2010; Wilson et al., 2012). In

both of the countries comprising the island of Hispanola, African mtDNA lineages are more commonly found. Similarly, the majority (>98%) of mtDNAs in Bermuda and Jamaica, two former British colonies, were African and West Eurasian in origin (Gaieski et al., 2011; Deason et al., 2012). By contrast, a limited assessment of mtDNA variation in the Dutch colony of Aruba found that 80% of lineages were Amerindian in origin (Toro-Labrador et al., 2003).

Much less research into the genetic diversity of paternal lineages in Caribbean populations has been conducted. However, the available studies suggest that less than 2% of individuals from populations in the Spanish Caribbean (Cuba, Hispaniola and Puerto Rico) carry indigenous paternal lineages (Mendizabal et al., 2008), whereas >75% have European and ∼23% have African Y-chromosomes, the latter likely being introduced through the trans-Atlantic slave trade (Salas et al., 2005). The relative absence of indigenous Y-chromosomes is likely due to centuries long conflict between the colonists and indigenous men, decimation from disease, and intermarriage between Spanish men and indigenous women (Rouse, 1992; Magnus, 1995; Guitar, 2000; Martin and Wasserman, 2005; Benn Torres et al., 2008; Mendizabal et al., 2008).

Given the known archaeological and linguistic evidence and the historical circumstances surrounding the peopling of the Caribbean, we characterized the genetic diversity in Puerto Rico and compared the gene pool of Puerto Ricans to those from other Caribbean and neighboring continental populations to infer place(s) of origin for indigenous Caribbean peoples. We also assessed the relative genetic contributions of non-native populations to Puerto Ricans and other Caribbean populations to determine the impact of European colonization and the Trans-Atlantic slave trade on the region. Our findings indicate distinct, yet equally complex, pasts for the male and female ancestors of modern day Puerto Ricans.

MATERIALS AND METHODS

Sample and data collection

In November 2010 and May 2011, we undertook two expeditions to the southeastern region of Puerto Rico. During these trips, we worked with 326 participants from the island of Puerto Rico and the neighboring island of Vieques. The majority of the participants are members of the Naguake Community, which is currently seeking state recognition as an indigenous base community. After obtaining informed consent, we obtained buccal cell samples from these individuals through either Omni cheek swabs or mouthwash rinse, and also gathered family history and genealogical information through interviews. Sample and data collection were undertaken with approval from the University of Pennsylvania IRB #8 and the support of Liga Guakia Taina-Ke. Samples were extracted in the Laboratory of Molecular Anthropology at the University of Pennsylvania using Qiagen QiAmp DNA Mini Kits, following the manufacturer's protocol.

Molecular genetic analysis

mtDNA diversity. To elucidate the maternal genetic ancestry of Puerto Ricans, we examined mtDNA variation in all 326 male and female participants through the analysis of single nucleotide polymorphisms (SNPs) using custom TaqMan assays. The SNP analysis involved screening the samples for markers that identify the basal structure of the mtDNA phylogeny, as well as known mtDNA haplogroups (2758G, 4248C, 4580T, 5178A, 6371A, 7028C, 8280G, 8994A, 10034C, 10238C, 10400A, 10550C, 10873G, 11467G, 11719C, 12612G, 12705G, 13263C, 13368A, 13928T, 14766G). All SNPs were analyzed on an ABI 3130XL Gene Analyzer using the novel GenoCore multiplex kit following the manufacturer's protocol (Applied Biosystems). The samples were also screened for additional markers that defined specific haplogroups and their subgroups with custom TaqMan assays that were read on an ABI 7900 Real Time PCR analyzer (Zhadanov et al., 2010; Schurr et al., 2012a).

In addition, for each sample, we sequenced 1121 base pairs (bp) of the mtDNA control region (CR), including both hypervariable segments-I and -II (HVS1 and HVS2) (Zhadanov et al., 2010; Gaieski et al., 2011). All sequences were run on an ABI 3130XL Gene Analyzer, and aligned and read with the Sequencher 4.8 software tool (Gene Codes Corporation). Polymorphic sites were ascertained relative to the revised Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999). The combined data from coding region SNPs and CR sequences defined the maternal haplogroup and haplotypes, respectively, for each individual.

We also sequenced the complete mitogenomes for twelve Puerto Rican participants having indigenous mtDNA haplogroups (A2, B2, and C1) and distinctive HVS1 sequences. The mtDNAs were sequenced using a published method (Rieder et al., 1998), and read as described above. The resulting mitogenome sequences were compared to previously published sequences, including those presented in Phylotree Build 16 (van Oven and Kayser, 2009), Family Tree DNA (The Cuban mtDNA Group, 2013; The Puerto Rican Geographical DNA Group, 2013), and GenBank (Benson et al., 2010), in order to identify the corresponding branches of haplogroups A2, B2, and C1 to which they belonged, hence, infer the likely continental source for each haplotype.

Y-chromosome diversity. We investigated the paternal genetic ancestry of Puerto Ricans by screening 121 male samples for phylogenetically informative biallelic markers in the non-recombining region of the Y-chromosome (NRY) that define paternal haplogroups and their major subbranches. This high-resolution SNP analysis was undertaken in a hierarchical fashion according to published information (Y Chromosome Consortium, 2002; Karafet et al., 2008). All markers were screened using custom TaqMan assays, and scored on an ABI 7900HT Fast Real-Time PCR System (Zhadanov et al., 2010; Gaieski et al., 2011; Schurr et al., 2012a).

Paternal haplotypes were further defined through the analysis of 17 Y-chromosome short tandem repeats (Y-STRs) that are part of the AmpFℓSTR Y-filer PCR Amplification Kit (Applied Biosystems). A separate multiplex reaction was also used to characterize six additional fragment length polymorphisms (M17, M60, M91, M139, M175, and M186) and two additional Y-STRs (DYS388, and DYS426). PCR products were run with GeneScan 500 LIZ Size Standards and read on an ABI 3130xl Gene Analyzer. The combination of SNPs and STR alleles defined paternal lineages and haplotypes, respectively, for each male individual.

Phylogenetic analysis. To infer patterns of diversity and ages of haplogroups through coalescent time estimation, median-joining (MJ) networks were constructed for both mtDNA haplogroups A2 and C1 using Network 4.611 software (Bandelt et al., 1999). The weighting scheme suggested in Bandelt et al. (2002) was used, where fast-evolving sites were given lower weights relative to other less mutable sites. For haplogroup age estimation, a rate of 1 mutation every 16,667 years was used for the extended HVS1 segment of mtDNA [np 16,024–16,390] (Soares et al., 2009). MJ networks were also constructed and coalescence times were estimated for haplogroups A2k, A2z, A2*, C1b2 and C1b4, using a rate of one mutation every 9,058 years for the complete CR of mtDNA [np 16,024–573] (Soares et al., 2009). All indels were excluded from this analysis.

We also inferred patterns of diversity and estimated coalescent times for Y-chromosome haplogroups with Network 4.611 (Bandelt et al., 1999). MJ networks based on 17 Y-STRs were constructed for haplogroups E1b1b, E1b1a, G2a, I, J, and R1b, the most commonly occurring paternal haplogroups among Puerto Rican males. For coalescent time estimates, we used a rate of one mutation every 453 years, which was estimated by taking the inverse per generation mutation rate of each locus multiplied by the number of loci and by generation time, or 25 years (Chandler, 2006).

Statistical analysis. Due to the uniqueness of indigenous mtDNA lineages in Puerto Rico when compared to other neighboring Caribbean and continental populations and the paucity of shared derived haplotypes among these groups, we employed haplogroup frequencies to measure the inter-population genetic distances. We estimated F_{ST} genetic distances between Puerto Ricans (this study), Cubans (Mendizabal et al., 2008), Dominicans [Dominican Republic] (Bryc et al., 2010), Trinidadians (Benn Torres et al., 2008), Dominicans [Dominica] (Benn Torres et al., 2008), Bermudians (Gaieski et al., 2011), Yucatec Mayans (Sandoval et al., 2012), Cherokee and Muskogean-speaking indigenous populations from the southeastern United States (Bolnick and Smith, 2003; Bolnick et al., 2006), African-Americans from Philadelphia (Stefflova et al., 2009), Venezuelans (Lander et al., 2008), Senegalese (Stefflova et al., 2009), Angolans (Plaza et al., 2004), and Spanish (Richards et al., 1998), using Arlequin 2.0 (Excoffier et al., 2005). These populations were chosen because of their geographic proximity to Puerto Rico and/or because they potentially represented source populations for the island's genetic diversity. The F_{ST} values were plotted in two dimensions using multi-dimensional scaling (MDS) in SPSS v. 17.0 (SPSS Inc., 2008).

Similarly, we estimated F_{ST} values from Y-chromosome haplogroup frequencies to assess the paternal genetic affinities of Puerto Ricans and comparative populations, using Arlequin 2.0 (Excoffier et al., 2005). The comparative populations included Cubans (Mendizabal et al., 2008), Bermudians (Gaieski et al., 2011), Trinidadians (Benn Torres et al., 2008), Yucatec Mayans (Sandoval et al., 2012), Venezuelans (Martinez et al., 2007), Tunisians and Algerians (Arredi et al., 2004), Angolans (Coelho et al., 2009), Moroccans (Semino et al., 2004), Senegalese (Hassan et al., 2008) and Portuguese and Spanish (Gonçalves et al., 2005). These populations were chosen for the same reasons listed above for the compara-

tive analysis of mtDNA data. The resulting F_{ST} values were plotted in two-dimensions using MDS.

RESULTS

mtDNA diversity in Puerto Rico

Analysis of the mtDNAs from 326 male and female participants yielded 98 distinct HVS1 haplotypes and 153 CR haplotypes from eighteen different haplogroups (Supporting Information Tables S1–S3). Native American haplogroups (A2, B2, C1, and D1) (Schurr, 2004; Schurr and Sherry, 2004) comprised most of these lineages, with haplogroups A2 and C1 accounting for 51.8% of the population, and haplogroups B2 and D1 for 4.3 and 0.9%, respectively. Haplogroup X2a mtDNAs were not present among the Puerto Rican samples. Sub-Saharan African haplogroups L0, L1, L2, and L3 accounted for 25.8% of the participants, while West Eurasian and North African haplogroups (H, J, K, R0a, T, U2, U5a, U5b1, U6b1, W, and X2) accounted for the remaining 17.2%.

Native American haplogroups. Haplogroup A2 was the most common maternal lineage in Puerto Rico, accounting for 29.8% of the participants and being represented by 11 distinct HVS1 haplotypes (Table 1). Outside of the A2 founder haplotype (#1), one of the two most commonly occurring A2 haplotypes (#7) had two additional transitions at nucleotide pairs (np) 16,083 and 16,256. Named A2z in Phylotree (Build 16; van Oven and Kayser 2009), haplotypes belonging to this lineage also had a transition at np 214 in the HVS2. To date, A2z has not been reported in any other population with the exception of Cubans (Mendizabal et al., 2008; The Cuban mtDNA Group, 2013), and generically labeled Hispanic Americans from a forensic genetic analysis (Just et al., 2008). The latter group likely represents individuals of either Cuban or Puerto Rican descent living in the United States, although the exact ethnicity of these individuals is unclear.

Among Cubans, a variety of A2z haplotypes has been observed, with some of the most common forms possessing an additional T16126C mutation (also seen in three Puerto Ricans, haplotype #8; Table 1), or a G16274A mutation (not seen in Puerto Ricans) (Mendizabal et al., 2008; The Cuban mtDNA Group, 2013). In this regard, the published data from Cuba (Mendizabal et al., 2008) showed a slight discrepancy with other mtDNA data sets in that some putative A2z samples lacked the T16083C mutation. This transition defines the haplogroup, and appears in all A2z haplotypes in our study, in those from genome-wide studies on Hispanic populations (Just et al., 2008; Bryc et al., 2010), and in Cuban and Puerto Rican samples from Family Tree DNA projects (The Cuban mtDNA Group, 2013; The Puerto Rican Geographical DNA Group, 2013). That being said, it is possible that some A2z haplotypes simply lack the T16083C mutation.

Another set of very common A2 haplotypes in Puerto Ricans (#11–12) differed from the founder sequence in having T179C and A385G mutations in the HVS2 (Supporting Information Table S2). Previously observed in Puerto Ricans (Martínez-Cruzado, 2010), we have provisionally designated them as belonging to haplogroup A2*. A2* mtDNAs cannot be distinguished from the A2 founder haplotypes using HVS1 data alone because they

TABLE 1. *Native American HVS1 sequences in Puerto Ricans and comparative populations*

#	Hg	HVS1 sequence (16,000+)	PRO	CUB	MAY	ELS	SVG	TRI	VEN	COL	GUY	BRZ	SEU
1	A2	111-223-290-319-362	42	6	24	23		6	63	6	1	4	6
2	A2	111-129-223-290-319-362	16		1			1	26				
3	A2	111-189-223-290-319-362	1		1	4				1			
4	A2	111-218-223-290-319-362	2	3									
5	A2	111-223-234-290-319-362	2		1								
6	A2	111-223-259-290-319-362	1						1				
7	A2	083-111-223-256-290-319-362	15	3									
8	A2	083-111-126-223-256-290-319-362	3	2	1								
9	A2	083-111-189-223-256-290-319-362	1										
10	A2	097-098-111-189-223-290-319-320-362	4									2	
11	A2	223-290-319-362	10	1									
12	B2	092-189-217-249-312-344	14	1									
13	C1	223-298-325-327	43	6	23		10	1	8	6	1	3	6
14	C1	223-295-298-325-327	2							1			
15	C1	223-298-299-325-327	4						1				
16	C1	223-248-298-299-325-327	1									1	
17	C1	223-298-325-327-355	1										
18	C1	093-223-298-325-327	1	2								3	
19	C1	189-223-298-325-327	1						1				
20	C1	086-189-223-298-325-327	1										
21	C1	086-189-223-278-298-325-327	16						1				
22	C1	077T-086-129-183-189-223-278-298-325-327	1										
23	C1	221-223-261-298-325-327	1										
24	D1	223-325-362	3	4	8				1		2	13	1
		Totals	186	28	59	27	10	8	102	16	2	26	13

Note: Population Abbreviations abbreviations are as follows: PRO = Puerto Ricans; CUB = Cubans; MAY = Mayans; ELS = Salvadorans; SVG = Vincentians; TRI = Trinidadians; VEN = Venezuelans; COL = Colombians; GUY = Guyanese; BRZ = Brazilians; SEU = Native Americans from the southeastern USA.

have no derived mutations in that portion of the mtDNA genome. This fact limits the number of comparative data sets that can be searched for this otherwise unique Caribbean haplotype in other populations. For this reason, we have begun further analysis of this haplogroup through mitogenome sequencing GenoChip analysis (Elhaik et al., 2013). A related haplotype (#4), defined by an additional C16218T transition in the HVS1 (Table 1), has not been reported outside of Puerto Rico and Cuba (Mendizabal et al., 2008; The Cuban mtDNA Group, 2013). It also lacks the C64T mutation in HVS2 that is otherwise found in most Mesoamerican and South American A2 haplotypes (Kemp, 2006; Tamm et al., 2007; Achilli et al., 2008) (Supporting Information Table S2).

A third A2 haplotype found in Puerto Rico (#5) differed from the founder sequence by a C16234T mutation (Table 1). This haplotype could possibly be present in Mexican-Americans based on an analysis of complete mtDNA genome sequences (Kumar et al., 2011), although the defining mutations for this haplotype, which currently define haplogroup A2f3 in Phylotree (Build 16), exist outside of the CR (G5585A, A9156G, G11914A, C14275T, and G15323A). Similar to the two previously described A2 lineages, this haplotype has not been identified among any South American populations (e.g., Alves-Silva et al., 2000; Salas et al., 2008; Lander et al., 2008).

Most of the remaining A2 haplotypes occurring at moderate frequency in Puerto Ricans have not been observed in indigenous Mesoamerican or North American populations, but appear in various groups from continental South America (Alves-Silva et al., 2000; Salas et al., 2005; Lander et al., 2008). A number of Puerto Rican haplotypes belonging to haplogroup A2k were identified through a combination of CR sequencing, TaqMan SNP assays and complete genome sequencing (#17–23) (Supporting Information Tables S2 and S4), with A2k itself being defined by the T3202C mutation. While rare in Cubans (<1%; Mendizabal et al., 2008; The Cuban mtDNA Group, 2013), A2k haplotypes have been observed among Amazonian populations (e.g., Tamm et al., 2007; Achilli et al., 2008), suggesting a genetic connection between Puerto Rico and South America.

Although HVS1 haplotype #2 has been found throughout the Americas (Table 1), we did not observe any exact matches to its full mitogenome sequence outside of Puerto Rico in the current published literature (Supporting Information Table S4). This observation, along with the polymorphic nature of the G16129A mutation, makes identifying the exact origin of this haplotype difficult. Nonetheless, haplotype #2 is one of the most common A2 mtDNAs in northern South America (e.g., Gomez-Carballa et al., 2012), although less frequent in Mesoamerica (e.g., Kemp, 2006; Salas et al., 2009; Sandoval et al., 2009).

Haplogroup C1 was the second most common haplogroup in our Puerto Rican data set, accounting for 22.1% of mtDNA lineages and seven distinct HVS1 haplotypes (Table 1; Supporting Information Tables S1–S3). Within this haplogroup, C1b accounted for the majority of the haplotypes (20.8%). Among Puerto Ricans, this haplogroup is comprised of the two major sub-groups C1b2 and C1b4, which are defined by the A263G and G143A mutations in the HVS2, respectively (Supporting Information Table S2). An unnamed subgroup of C1b defined by the A16299G mutation and the subgroup C1b11, defined by the C16295T mutation, also occurred among Puerto Ricans, but at low frequency (~1%) (Table 1; Supporting Information Tables S2 and S3). Nearly all C1b branches have been previously identified among

indigenous South American populations (Alves-Silva et al., 2000; Salas et al., 2008; Lander et al., 2008; Gomez-Carballa et al., 2012). Furthermore, C1b as a whole is less common than C1c and C1d (the other indigenous sub-groups of C1) among Mexican, Cuban, and native North American populations (e.g., Bolnick and Smith, 2003; Kemp, 2006; Sandoval et al., 2009).

Haplogroup C1c was identified in several Puerto Rican participants. This haplogroup is defined by two mutations occurring outside of the CR (G1888A and G15930A). We discovered the G15930A mutation in haplotype #55 (Supporting Information Table S2) through extended CR sequencing, and confirmed the presence of the G1888A mutation in this and other putative C1c samples (haplotype #54 and 55; Supporting Information Table S2) using a custom TaqMan assay. However, none of our Puerto Ricans samples belonged to haplogroup C1d, which is common in South American (Alves-Silva et al., 2000; Salas et al., 2008; Lander et al., 2008), Cuban (Mendizabal et al., 2008), Lesser Antillean (Benn Torres et al., 2008) and Mexican-American (Kumar et al., 2011) populations.

As previously seen in Puerto Ricans (Martínez-Cruzado et al., 2001, 2005), haplogroups B2 and D1 represented a minority of the indigenous mtDNAs in our sample. Both B2 and D1 were each represented by a single HVS1 haplotype (#12 and #24, respectively) (Table 1). Interestingly, the sole Puerto Rican B2 haplotype (B2b3) was previously observed in Native Mexicans, Mexican-Americans, and Amazonian Indians (Alves-Silva et al., 2000; Kemp, 2006; Sandoval et al., 2009; Mandarino, 2010; Kumar et al., 2011), as well as Cubans (Mendizabal et al., 2008). By contrast, the D1 mtDNA had the HVS1 founder haplotype, making its exact place of origin and time of arrival difficult to infer.

African mtDNA haplogroups. African haplogroups comprised 27.6% of the Puerto Rican mtDNAs, with L haplogroups encompassing the major of them (25.8%). Among these lineages, haplogroups L1b and L3e were the most common, accounting for 7.1 and 3.7% of the Puerto Rican mtDNAs, respectively. In addition, L0a, L1c, L2a, L2c, L3b, L3d, and L3f were also detected in the Puerto Rican data set.

Although observed in African-Americans from the eastern United States (Stefflova et al., 2009), as well as populations from Jamaica and other former British colonies (e.g., Benn Torres et al., 2008; Gaieski et al., 2011; Deason et al., 2012), haplogroups L3e and L0a are also commonly found in Brazilian populations (Alves-Silva et al., 2000; Bandelt et al., 2001; Hünemeier et al., 2007; Gonçalves et al., 2007), and are thought to reflect the arrival of Africans originating in the former Portuguese colonies of Angola and Mozambique. While present in only 3.7% of Puerto Rican participants, L3e was represented by seven distinct HVS1 haplotypes, all belonging to either subgroups L3e1 and L3e2b (Bandelt et al., 2001) (Supporting Information Tables S2 and S3). This diversity of L3e haplotypes could possibly represent the arrival of Africans on Portuguese slave ships during the early colonial settlement of the island, as Portugal was the dominant trans-Atlantic slave trader in the 16th and 17th century (Salas et al., 2005).

By contrast, L1b accounted for 7.1% of Puerto Rican haplotypes, but consisted of only five different HVS1 haplotypes (Supporting Information Tables S2 and S3).

L1b is commonly seen in African Americans in the United States (6%) (Stefflova et al., 2009) and in Afro-Caribbean populations from former British colonies, such as Jamaica (25%; Deason et al., 2012). This lineage is also one of the most common haplogroups among Senegalese and other West African populations (Rosa et al., 2004; Salas et al., 2005; Stefflova et al., 2009). Given its high frequency and limited diversity, L1b may have been introduced into Puerto Rico during the late stages of colonization and slavery while the British controlled the trade in the 18th and 19th centuries (Rawley et al., 2005).

However, a recent analysis of autosomal diversity has suggested that short autosomal fragments of African origin in Puerto Rico are likely to have originated in the Senegambia region, rather than West-Central Africa (Moreno-Estrada et al., 2013). These results imply that the slave trade was initially more intensive in the Senegambia region, where L1b mtDNAs are more common, and then shifted south in the African continent in the 18th and 19th centuries (Curtin, 1969; Lovejoy 2000). If correct, this scenario implies that L3e haplotypes in Puerto Rico came with African women captured in southeastern Africa in the latter stages of the slave trade.

In addition, five mtDNAs belonged to haplogroup U5b1b (Rando et al., 1998; Achilli et al., 2004; Rosa et al., 2004), and were represented by a single haplotype (#93; Supporting Information Table S3). This lineage appears in a number of West African populations, including the Fulbe. Its presence there has been suggested to possibly reflect a Berber expansion into West Africa (Rosa et al., 2004) or instead a prehistoric back-migration to sub-Saharan Africa from Eurasia (Cruciani et al., 2002). Its presence in about 1.5% of all Puerto Ricans has been considered a product of the African slave trade (Martínez-Cruzado et al., 2005).

In addition, one mtDNA belonged to haplogroup U6b1 (#94; Supporting Information Table S3). U6b1 is found only in the Canary Islands and on the Iberian Peninsula (Maca-Meyer et al., 2003). Current work suggests that North African settlers likely brought this lineage to the Canarian archipelago several thousand years ago. Therefore, its dispersal into the Iberian Peninsula probably occurred somewhat earlier during the Capsian period in North Africa (Maca-Meyer et al., 2003).

West Eurasian mtDNA haplogroups. West Eurasian haplogroups accounted for 16% of Puerto Rican mtDNAs. Among these, haplogroups U, J, and H were the most commonly occurring, accounting for 5.2, 3.7 and 3.7%, respectively, of the participants (Supporting Information Tables S2 and S3). These haplogroups are commonly found in Western European countries, such as Spain (Richards et al., 1998; Crespillo et al., 2000; Simoni et al., 2000), and thus are likely to have originated there and reached Puerto Rico with the migration of Spanish women to the island during the colonization period. However, given the recent influx of North Americans into Puerto Rico (Rogozinski, 1999), some haplogroups of European origin may have arrived in Puerto Rico via immigration from the United States. Haplogroups K, R0a, T, W, and X2b also occurred at low frequency among Puerto Ricans (Supporting Information Tables S2 and S3). Although West Eurasian in origin, these haplogroups are generally less common in Spain
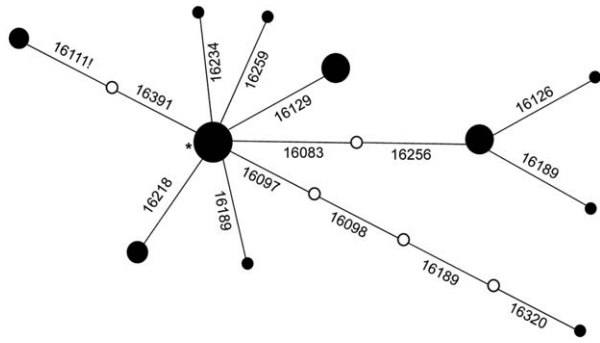
Fig. 2. A MJ network of A2 HVS1 haplotypes. The asterisk (*) denotes the founder haplotype (16111, 16223, 16290, 16319, 16362). The node size reflects the number of individuals having the same haplotype. Filled circles represent haplotypes observed in this study, while open circles represent inferred haplotypes not seen in the Puerto Rican samples.
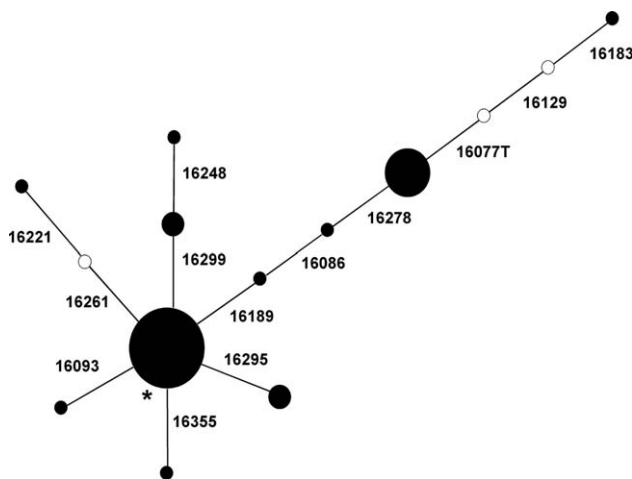
Fig. 3. A median-joinng network of C1 haplotypes. The asterisk (*) denotes the founder haplotype (16,223, 16,298, 16,325, 16,327). The node size reflects the number of individuals having the same haplotype. Filled circles represent haplotypes observed in this study, while open circles represent inferred haplotypes not seen in the Puerto Rican samples.

and Western Europe than other parts of the Mediterranean region (Richards et al., 1998; Crespillo et al., 2000; Simoni et al., 2000). Thus, they may be indicative of maternal ancestry from other West Eurasian source populations, such as central or northern Europe, or perhaps North Africa and the Middle East.

Network analysis of mtDNA data. The MJ networks constructed from haplogroup A2 and C1 HVS1 sequences showed a distinct pattern (Figs. 2 and 3). While having large central nodes, both networks also exhibited several branches with high frequency nodes often separated from the central haplotype by two or more mutations. The topology of the networks suggested that these haplogroups arrived on the island from elsewhere as sets of already derived haplotypes, rather than reaching the island as founder types and evolving over time on the island. This pattern also suggested that the A2 and C1 lineages likely arrived in Puerto Rico at different times.

To better infer the pattern of population diversification (haplogroup evolution and dispersal) in the Caribbean, we constructed individual networks for five of the most common sub-haplogroups in Puerto Rico (A2k, A2z, A2*, C1b2, and C1b4) using complete CR haplotypes (Fig. 4a–e) and estimated divergence dates for them (Table 2). This analysis revealed divergence dates for 9,510 YBP for A2k and 905 YBP for A2k1. While the estimate for A2k predates human settlement of the Caribbean, that for A2k1 is consistent with the arrival of ancestral Taíno populations in the Greater Antilles. The A2k estimate further implies both the antiquity of this lineage in South America and arrival in the Caribbean as a diverse set of haplotypes. The dates of 4,291 YBP for A2z and 1,906 YBP for A2* were somewhat earlier than the arrival of ancestral Taíno populations from South America. Interestingly, neither of these haplogroups has been observed in South America, although they are otherwise common in Cuba. In fact, these haplotypes appear similar to those in native populations from North America that lack the C64T mutation in the HVS2 (Helgason et al., 2006; Vilar et al., 2011) (Supporting Information Tables S2 and S3).

The estimated divergence time for haplogroup C1b4 was 3,019 YBP, whereas that for haplogroup C1b2 was 647 YBP (Table 2). The age of C1b4 was consistent with the earliest archaeological evidence for people on the island (possibly the Ortoroid cultural period). By contrast, the younger age of C1b2 suggested that it arrived later during the Taíno expansion into the Caribbean, or possibly even arriving with Carib populations that were expanding through the Lesser Antilles during the last millennia. However, C1b2 haplotypes have not been found in Carib descendant populations in the Lesser Antilles or northern South America (Benn Torres et al., 2008; Mandarino, 2010).

Statistical analysis of mtDNA data. Given the low number of HVS1 haplotypes shared between Puerto Ricans and other populations (Table 1), the analysis of genetic variation was carried out comparing the Puerto Rico data set to other Caribbean and continental source populations based on haplogroup frequencies. The MDS plot of pairwise F_{ST} estimates revealed that Puerto Ricans were most similar to Cubans and Venezuelans, who also have high frequencies of indigenous haplotypes (Fig. 5). In addition, populations from former British and Spanish colonies grouped together with African populations, in large part because of having significant frequencies of haplogroup L mtDNAs. Overall, the MDS plot showed a clear separation of the three ancestral source populations for admixed groups in the Americas (West Eurasian, Amerindian and African) in the corners of a triangle. The positions of admixed groups within that triangle thus reflect the relative genetic contributions of these three source populations.

Heterozygosity estimates based on indigenous lineages showed that Puerto Ricans were less diverse than Cubans (Supporting Information Table S5). This result, along with the high frequency of haplotypes shared between the two islands (Table 1), may be indicative of the directionality of population movements from west to east across the Greater Antilles, with the Puerto Rican haplotypes being a subset of those present in Cuba. The pattern may also be reflective of genetic drift occurring in Puerto Rico, which is geographically less than a tenth
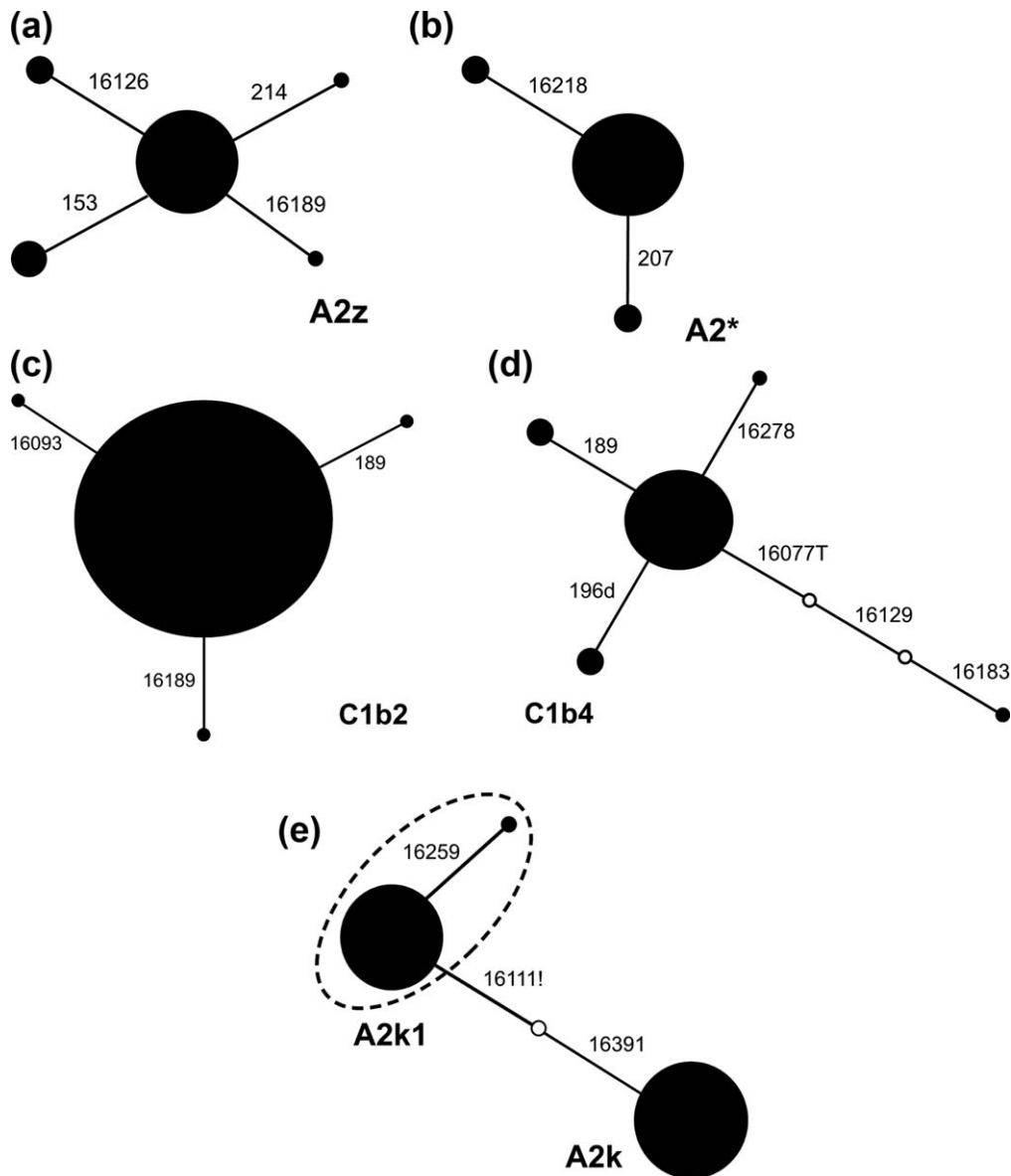
Fig. 4. MJ networks of CR sequences for different mtDNA haplogroups observed in Puerto Ricans. (**a**) A2z (**b**) A2* (**c**) C1b2 (**d**) C1b4 (**e**) A2k. The node size reflects the number of individuals having the same haplotype. Filled circles represent haplotypes observed in this study, while open circles represent inferred haplotypes not seen in the Puerto Rican samples.

TABLE 2. *Coalescence Times times for mtDNA haplogroups in Puerto Rico based on CR sequences*

Haplogroup	Coalescence time (YBP)
A2k	9,510 ± 6,420
A2k1	905 ± 905
A2*	1,906 ± 1,348
A2z	4,291 ± 2,477
C1b2	647 ± 373
C1b4	3,019 ± 1,423

the size of Cuba. However, Puerto Ricans and Venezuelans had the smallest (non-significant) F_{ST} value, while sharing a number of HVS1 haplotypes between them (#1, 2, 6, 13, 15, 19, 21, and 24; Table 1). These data clearly point to northern South America as a significant source of genetic lineages for the island.

Y-chromosome diversity in Puerto Rico

Approximately 84% of the paternal lineages in Puerto Ricans were of European or West Eurasian origin (Table 3; Supporting Information Tables S6 and S7). Of these, 44.6% belonged to haplogroup R1b, which occurs commonly among Western European populations (Myres et al., 2010). The second most common haplogroup was E1b1b (17.3%), which is also found in Europe but has a more Mediterranean and North African distribution than R1b (Cruciani et al., 2007). In addition, haplogroups G2a (6%), I2 (4%), J2 (5%), and T (2%) were found at varying frequencies among Puerto Rican men, with these also having a North African and Mediterranean geographic focus (Arredi et al., 2004; Semino et al., 2004; Cruciani et al., 2007; Rootsi et al., 2012), while haplogroup I1, common in central and northern Europe (Lappalainen et al.,

2006) occurred in 5% of Puerto Ricans. Approximately 15.7% of participants belonged to a single, yet diverse sub-Saharan haplogroup E1b1a (Rosa et al., 2007). No indigenous Y-chromosome haplogroups (e.g., Q-M3, Q-L54; Dulik et al., 2012a,b) were identified in the 121 male participants.

Network analysis of Y-chromosome haploty-pes. The MJ networks were constructed from Y-STR haplotypes belonging to the six major Y-chromosome haplogroups for Puerto Rican men (E1b1a, E1b1b, G2a, I1, J2, R1b, and T). The resulting networks had long branches with a number of missing intermediate nodes and no clear central node, suggesting that the haplotypes did not originate on the island (see Fig. 6 as an example). The coalescence dates for most haplogroups were also much greater than 500 years (data not shown), suggesting that the Y-chromosome diversity within each network pre-dated the arrival of European and African peoples to the island. Together, these results likely reflect the consequences of the arrival and settlement of men from many different regions during the European colonization of Puerto Rico over the course of the past several centuries.

Based on Y-STR networks, the only haplogroup that may have diversified to some extent on Puerto Rico was G2a1b. Haplogroup G2a1b is defined by having two alleles at microsatellite DYS19, a marker otherwise

Fig. 5. A MDS plot of F_{ST} estimates based on mtDNA haplogroup frequencies. The stress value for the plot was 0.18.

characterized by just one allele in all other haplogroups. The coalescence estimate for Puerto Rican G2a1b was approximately 480 ybp, suggesting it might have arrived during the early stages of European colonization. By and large, G2a1b is found in the Mediterranean (including Corsica, Sardinia, and southern Italy) (Keller et al., 2012), suggesting that populations from this region may have been the sources of some of the earliest male settlers to Puerto Rico and the Spanish Caribbean.

Statistical analysis of Y-chromosome data. The F_{ST} analysis of Y-chromosome haplogroup frequencies for Puerto Ricans, comparative Caribbean and putative source populations indicated that Puerto Ricans were again most genetically similar to Cubans and Venezuelans, although also very close to Portuguese and Spanish populations (Fig. 7). This pattern was not unexpected given the high frequency of West Eurasian haplogroups in these populations. At the same time, Puerto Rican men showed a pattern of haplogroup diversity (high frequencies of E1b1b and J2) more similar to that seen in North African populations (Semino et al., 2004; Cruciani et al., 2007) than a typical modern Spanish population (Myres et al., 2010). Furthermore, some of their Y-STR haplotypes matched those present in Semitic speaking populations from Morocco and other parts of North Africa (Semino et al., 2004). Such results suggest a connection between Puerto Rico and North Africa, one possibly associated with the recruitment of men from those regions (or the Canary Islands) by the Spanish to colonize the Caribbean.

DISCUSSION

Maternal genetic diversity in Puerto Rico

This study is one of the first to examine genetic diversity in specific municipalities in the southeastern region (Maunabo, Yabucoa, Las Piedras, San Lorenzo) of Puerto Rico, as well as Vieques Island. Based on our analysis, the indigenous mtDNA HVS1 haplotypes from this region of Puerto Rico were generally similar in type and frequency to those found in the central and western portions of the island (Martínez-Cruzado et al., 2005; Martínez-Cruzado, 2010). Although Vieques had a higher frequency of African haplogroups (42%) and lower frequency of indigenous haplogroups (45%) compared to Puerto Rico (24 and 59%, respectively), the majority of HVS1 haplotypes from Vieques (82%) were also observed on the mainland.
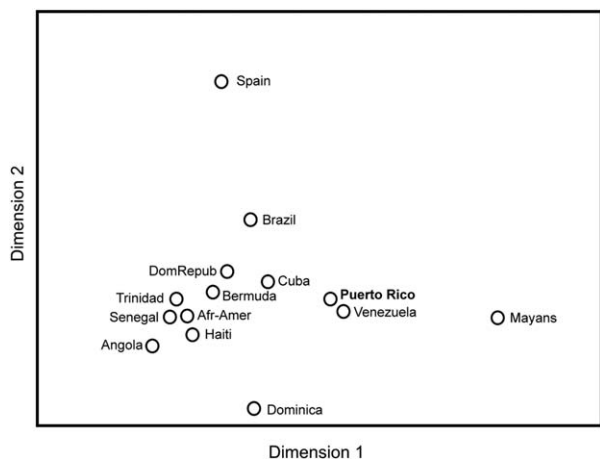
TABLE 3. Y-chromosome haplogroup frequencies among Puerto Rican men

Hg	Geographic ancestry	N	Frequency (%)		
			Total (n = 121)	SE Puerto Rico (n = 110)	Vieques (n = 11)
E1b1a	Sub-Saharan African	19	15.7 (19)	15.5 (17)	18.2 (2)
E1b1b	West Eurasian	21	17.4 (21)	14.5 (16)	45.5 (5)
G2a	West Eurasian	7	5.8 (7)	5.5 (6)	9.1 (1)
I1	West Eurasian	6	5.0 (6)	5.5 (6)	0
I2	West Eurasian	5	4.1 (5)	4.5 (5)	0
J2a	West Eurasian	6	5.0 (6)	5.5 (6)	0
R1a	West Eurasian	1	0.8 (1)	0.9 (1)	0
R1b	West Eurasian	54	44.6 (54)	46.4 (51)	27.3 (3)
T	West Eurasian	2	1.7 (2)	1.8 (2)	0
	Total	121			

Note: The numbers in parentheses indicate the number of haplotypes in each haplogroup.
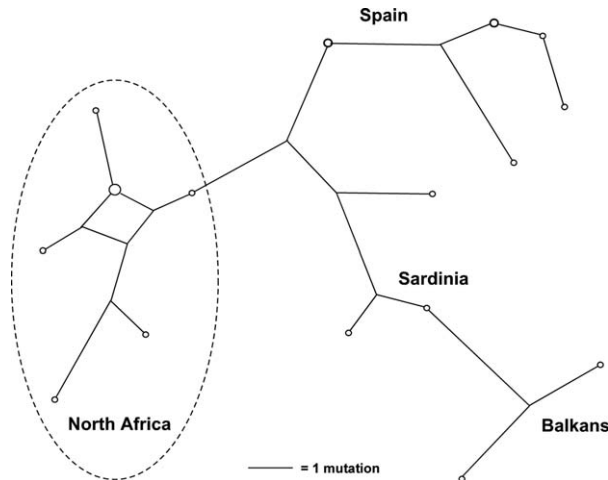
Fig. 6. A MJ network of haplogroup E1b1b Y-STR haplotypes. The node size reflects the number of individuals having the same haplotype. The key indicates the branch length of a single STR allele difference between two haplotypes. The likely source populations for the different branches are labeled in the network.

Fig. 7. A MDS plot of F_{ST} estimates based on Y haplogroup frequencies. The stress value for the plot was 0.14.
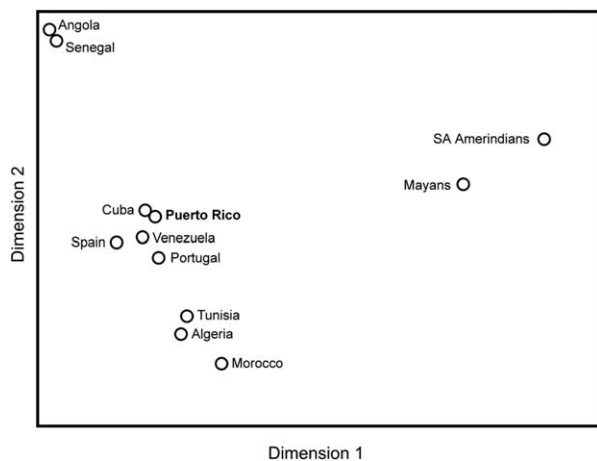
The apparent genetic continuity between the two inter-visible islands was intriguing. Historical records have suggested that Taínos (Borinqueños) occupied the island of Puerto Rico and Caribs occupied the island of Vieques at the time of European contact, with Caribs often raiding the eastern Puerto Rico coast (Fewkes, 2009). Whether or not they represented separate genetic populations, Taíno rebellions against Spanish colonization and slavery in the early 16th century led to cessation of hostilities and possibly some mixing between Taíno and Carib populations on Puerto Rico as they responded to Spanish settlement and conquest on the island (Figueredo, 1978). This mixing could possibly be reflected in the genetic similarity of individuals living on Puerto Rico and Vieques.

We also found no clear regional patterns of genetic variation in eastern Puerto Rico in our data set, in contrast to previous studies that suggested some local differences based on C1 haplotypes frequencies in west-central Puerto Rico (Martínez-Cruzado et al., 2005; Martínez-Cruzado, 2010). While it is possible that, because of its relatively small size, regional patterns of

variation did not emerge on Puerto Rico prior to European contact, historical records have hinted at differences between Taíno populations occupying eastern and western Puerto Rico (Fewkes, 2009). In this regard, any genetic substructure once present on the island may have been lost when Europeans arrived, perhaps in part as a consequence of Taíno rebellions and relocations in response to Spanish colonial domination (Figueredo, 1978) or the result of disease epidemics and enslavement, which hugely affected the demography of Taíno populations (Rouse, 1992; Crosby, 2003; Raudzens, 2003). Furthermore, Puerto Rico has had a system of highways for several decades, making modern cross-island movement in the 20th century, hence, contact between individuals, quite easy.

Conversely, efforts to identify the geographic provenience of individual haplotypes originating outside of Puerto Rico were quite informative. As the Caribbean was one of the last regions to be settled by indigenous groups in the Americas (Wilson, 2007; Fitzpatrick and Ross, 2010), we anticipated being able to trace Puerto Rican indigenous maternal ancestry to certain continental regions, even after 520 years of European colonization.

The search for the origin of the two of the most common A2 haplotypes (A2z and A2*) in Puerto Rico took us to Cuba but no further. The high frequency and diversity of A2z haplotypes in Cuba suggests that this lineage originated in Cuba, the largest Caribbean island, and began diversifying shortly after its arrival there ~4,000 YBP, reaching Puerto Rico by around 2,000 YBP. As noted earlier, haplogroup A2* follows the same distribution. These patterns suggest that A2z and A2* arrived in Cuba in the mid-Holocene, and were later brought to Puerto Rico from the west, prior to the arrival of Taíno groups. Further support for this view is the absence of these haplogroups in the Lesser Antilles (Benn Torres et al., 2008) and South America (e.g., Alves-Silva et al., 2000; Lander et al., 2008; Salas et al., 2008; Gomez-Carballa et al., 2012).

Two possible entry points into the Caribbean via Cuba are Florida and the Yucatan peninsula. Since A2 is quite common in Mexico (Kemp, 2006; Sandoval et al., 2009; Schurr et al., 2012b; Vilar et al., 2013), and Mexico was the largest population center in North America during the late Holocene (Haines and Steckel, 2000), it would seem to be the likeliest source of A2 haplotypes for the Greater Antilles. However, despite the fact that indigenous Mesoamerican groups exhibit well over a hundred distinct A2 HVS1 haplotypes (Kemp, 2006; Salas et al., 2009; Melton et al., 2013; Vilar et al., 2013), only one of these haplotypes appears in our Puerto Rican data set. In addition, the Puerto Rican A2z and A2* haplotypes are missing the C64T mutation in HVS2, which is present in the vast majority of A2 types in Mesoamerica (Achilli et al., 2008; Sandoval et al., 2009; Vilar et al., 2011; Schurr et al., 2012b; Vilar et al., 2013).

By contrast, A2 haplotypes from United States and Canada often lack the C64T mutation, and Inuit populations from the Arctic (haplogroups A2a and A2b) lack it all together (Achilli et al., 2008; Vilar et al., in review). These data raise the possibility of different colonization routes being used to colonize Cuba and the Caribbean. In this regard, a recent paper suggested that populations may have migrated across the Arctic coast to reach and settled the American east coast (O'Rourke and Raff, 2010), and eventually moved south from Florida to the Bahamas and the Greater Antilles. This model is generally consistent with the age of and patterns of diversity

in A2z and A2* in Puerto Rico and Cuba. However, given the limited number and low resolution of indigenous mtDNA data from indigenous Floridians and the American Southeast (Huoponen et al., 1996; Bolnick and Smith, 2003), which undoubtedly were reshaped by the population losses and dislocations resulting from the 1539-42 De Soto expedition (Galloway, 2006) and others following it, we can only speculate about Florida possibly being a place of origin for A2z and A2* at this time.

Nevertheless, the presence of distinct A2 haplogroups in native populations from the eastern United States and Canada, such as A2i and A2f1 [both have the C64T mutation] (Huoponen et al., 1996; Bolnick and Smith, 2003; Owings et al., 2012), and their absence in the Caribbean (Mendizabal et al., 2008; The Cuban mtDNA Group, 2013), indicates a genetic discontinuity between the two regions. In fact, to date, only a single A2 HVS1 derived haplotype appears in both the Caribbean and North America. Defined by a T16249C mutation in addition to the A2 motif, it appears at low frequencies in Cubans (Mendizabal et al., 2008; The Cuban mtDNA Group, 2013) and Chickasaw and Choctaw groups from the southeastern United States (Bolnick and Smith, 2003), while being absent in Puerto Rico (Martínez-Cruzado et al., 2010, this study). Based on this evidence, it is likely that these A2 haplotypes arose through recurrent T16249C mutations.

The presence of C1b subhaplogroups in Puerto Rico suggests a strong genetic link to South America. Haplogroup C1b4 in Puerto Rico dates to 3019 YBP, a time frame coincident with the earliest peopling of the region from the south by the Ortoroid culture (Rouse, 1992). Haplogroup C1b4 has also been found among modern Venezuelan populations (Lander et al., 2008; Gomez-Carballa et al., 2012). In addition to being one of the most common lineages among Puerto Ricans (Martínez-Cruzado, 2010, this study), C1b2 has also been identified in native populations of Brazil (Cardenas et al., 2013). In fact, an earlier analysis of Puerto Rican C1b samples (Martínez-Cruzado et al., 2005) detected the C1b2-defining G7013A polymorphism in them, one previously identified in several South American populations through restriction fragment length polymorphism (RFLP) analysis (Kraho, Macushi, Marubo, Yanomami; Torroni et al., 1993). Furthermore, the fact that C1b2 may also represent the main haplotypes observed in ancient Taíno samples from the Dominican Republic (Lalueza-Fox et al., 2001) suggests that it could potentially represent a Taíno maternal lineage (Achilli et al., 2008; Schurr, 2010). Thus, these genetic data support a direct link between the Puerto Rican Taíno and South American Arawakan-speaking tribes from southern Venezuela and northern Brazil.

Several other subhaplogroups follow the same pattern and therefore may be associated with the Arawakan expansion(s). These include Puerto Rican A2k haplotypes (#17–23), C1b11 (#36), and C1b haplotypes with an additional A16299G mutation (#34, 35) (Supporting Information Table S2), all of which appear in South American populations (Tamm et al., 2007; Achilli et al., 2008; Salas et al., 2008; Lander et al., 2008; Gomez-Carballa et al., 2012). The distribution of these lineages conforms to the well-studied pattern of expansions of indigenous people to the West Indies through the Orinoco River Valley and the northern coast of South America. Other A2 and C1 lineages in Puerto Ricans are infrequent and less diverse and may be associated with smaller pre-colonial movements, or possibly post-colonial

transfers from the South American mainland associated with the post-colonial local slave trade across the Caribbean (Salas et al., 2005; Benn Torres et al., 2008).

As these data indicate, Puerto Rico has always been involved in the larger population dynamics of the Caribbean Basin as also seen in the archeological record (Righter et al. 1997, 2004, Siegel, 2005; Hofman et al., 2008). However, mtDNA diversity in Puerto Rico probably also reflects the impact of European colonization, when Taíno, Carib, and indigenous groups from other parts of the Americas were killed, relocated and even sold as slaves (Crosby, 2003; Raudzens, 2003; Salas et al., 2005). Because these historical events are less than 450-years-old, any haplotypes that were spread throughout the region during colonial times would not likely have evolved new mutations in Puerto Rico, hence, would appear as single derived types at the tips of branches in the networks. This may be the case for the B2 and D1 haplotypes, along with a few other individual A2 and C1c haplotypes, that show limited diversity on the island.

Paternal genetic diversity in Puerto Rico

The Y-chromosome results were somewhat unexpected and reflect a pattern of genetic replacement more extreme than seen for the mtDNA data. Nearly 85% of the male haplotypes were of West Eurasian origin and approximately 15% of sub-Saharan African origin, while none were clearly indigenous. This pattern strongly contrasts with that found for maternal lineages on the island, suggesting very distinct maternal and paternal histories for Puerto Ricans.

The absence of indigenous Y-chromosomes in Puerto Rico is consistent with data from other parts of the Caribbean. To date, none have been found among Cubans (Mendizabal et al., 2008) or other Caribbean populations (Benn Torres et al., 2008). While future analyses may eventually reveal indigenous paternal haplogroups in the Anglophone or Spanish Caribbean, our data reflect a complete replacement of indigenous paternal lineages with West Eurasian and African haplogroups through Spanish colonization, engagement in the trans-Atlantic slave, and colonial control over the past 500 years. Should ancient DNA research be undertaken with pre-Columbian samples from Puerto Rico, such as those from the Dominican Republic (Lalueza-Fox et al., 2001) or Cuba (Lalueza-Fox et al., 2003), indigenous Y-chromosomes might be identified in past Taíno populations.

As noted, the vast majority of male lineages are West Eurasian in origin. However, the Puerto Rican male gene pool is not an exact replica of that of Spain, the main colonial power on the island. While nearly 100% of Puerto Rican surnames are Spanish in origin, many have Muslim and even Sephardic Jewish family histories, as opposed to Castilian ones (Boyd-Bowman, 1956). Likewise, the high frequency of haplogroups E1b1b, G2a, I2, and J2 in Puerto Rican men suggests paternal ancestry from Spain, North Africa, the Mediterranean and even the Middle East. Interestingly, certain individual Y-STR haplotypes from Puerto Ricans were close matches to those from various locations including Spain, Sardinia, Algeria, and Morocco (Cruciani et al., 2007). These observations suggest that the diversity of Puerto Rico Y-chromosome haplotypes reflects the broad geographic ancestry of the men who sailed on Spanish ships and settled on the island.

The haplotype diversity revealed in the E1b1b Y-STR network (Fig. 6) also supports the interpretation that

male immigrants to the Caribbean had diverse ancestries, and likely arrived from different places in western and southern Europe, North Africa, and the Middle East over the course of the last 400 years. This distribution is consistent with colonial documents, which indicate that many of the ships embarking to the Caribbean often picked up men from the Canary Islands, the Azores and Madeira to serve on the vessels, and ultimately disperse throughout the colonies (Fernández-Armesto, 1982). Today, Puerto Rico, as well as other former Spanish colonies in the Caribbean and continental Americas, has strong historical ties to North Africa, and men from these regions apparently carried haplotypes common to those regions to the New World.

The rest of the male lineages belonged to haplogroup E1b1a. E1b1a is the most common paternal lineage in West and Central Africa, and represents approximately 90% of the men in those regions (Rosa et al., 2007; Veeramah et al., 2010). Thus, this haplogroup is likely to have been contributed to Puerto Rican populations by men of African descent through the trans-Atlantic slave trade (and perhaps through more recent immigration to the island). An earlier study found a similar E1b1a frequency for Puerto Ricans, in contrast with Dominicans and Cubans, which had 8% and 10% of E1b1a, respectively (Bryc et al., 2010). These data generally support the view that Puerto Rico was something of a backwater colony, leading fewer men with West Eurasian ancestry to settle there compared to Cuba and the Spanish colonial center of Hispañola.

Autosomal variation in Puerto Rico

While the Y-chromosome perspective on Puerto Rican genetic history argues for a greatly diminished influence of indigenous males on the island, analysis of autosomal SNP variation in Puerto Ricans reveals indigenous genetic ancestry among many of these individuals. Our initial survey of autosomal genetic variation in Puerto Rican individuals with the GenoChip (Elhaik et al., 2013) revealed 10–20% indigenous ancestry informative markers in all individuals, even those lacking indigenous mtDNA and Y-chromosome lineages. These results are consistent with other studies of autosomal variation in Puerto Rican populations (e.g., Via et al. 2011; Kidd et al., 2012; Gravel et al., 2013), as well as our analysis of admixed circum-Caribbean populations (e.g., Bermuda; Gaieski et al., 2012; Elhaik et al., 2013). They further demonstrate that autosomal data can be effectively used to reconstruct Caribbean colonization history, especially when coupled with data from the haploid mtDNA and Y-chromosome.

ACKNOWLEDGMENTS

The participants of the Genographic Consortium are arranged by surname alphabetically in the following list: Syama Adhikarla[a], Christina J. Adler[b], Elena Balanovska[c], Oleg Balanovsky[c], Jaume Bertranpetit[d], Andrew C. Clarke[e], David Comas[d], Alan Cooper[b], Clio S. I. Der Sarkissian[b], ArunKumar GaneshPrasad[a], Wolfgang Haak[b], Marc Haber[d,f], Angela Hobbs[g], Asif Javed[h], Li Jin[i], Matthew E. Kaplan[j], Shilin Li[i], Begoña Martínez-Cruz[d], Elizabeth A. Matisoo-Smith[e], Marta Melé[d], Nirav C. Merchant[j], R. John Mitchell[k], Laxmi Parida[h], Ramasamy Pitchappan[a], Daniel E. Platt[h], Lluis Quintana-Murci[l], Colin Renfrew[m], Daniela R. Lacerda[n], Ajay K. Royyuru[h], Fabrício R. Santos[n], Himla Soodyall[g], David F. Soria Hernanz[o], Pandikumar Swamikrishnan[p], Chris Tyler-Smith[q], Arun Varatharajan Santhakumari[a], Pedro Paulo Vieira[r], R. Spencer Wells[o], Pierre A. Zalloua[f], and Janet S. Ziegle[s]

[a]Madurai Kamaraj University, Madurai, Tamil Nadu, India; [b]University of Adelaide, South Australia, Australia; [c]Research Centre for Medical Genetics, Russian Academy of Medical Sciences, Moscow, Russia; [d]Universitat Pompeu Fabra, Barcelona, Spain; [e]University of Otago, Dunedin, New Zealand; [f]Lebanese American University, Chouran, Beirut, Lebanon; [g]National Health Laboratory Service, Johannesburg, South Africa; [h]IBM, Yorktown Heights, NY, USA; [i]Fudan University, Shanghai, China; [j]University of Arizona, Tucson, AZ, USA; [k]La Trobe University, Melbourne, Victoria, Australia; [l]Institut Pasteur, Paris, France; [m]University of Cambridge, Cambridge, United Kingdom; [n]Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; [o]National Geographic Society, Washington, DC, USA; [p]IBM, Somers, NY, USA; [q]The Wellcome Trust Sanger Institute, Hinxton, United Kingdom; [r]Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil; and [s]Applied Biosystems, Foster City, CA, USA.

LITERATURE CITED

Achilli A, Perego UA, Bravi CM, Coble MD, Kong QP, Woodward SR, Salas A, Torroni A, Bandelt HJ. 2008. The phylogeny of the four Pan-American mtDNA haplogroups: implications for evolutionary and disease studies. PLoS One 3:e1764.

Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V, Moral P, Dugoujon J-M, Roostalu U, Loogväli E-L, Kivisild T, Bandelt H-J, Richards M, Villems R, Silvana Santachiara-Benerecetti A, Semino O, Torroni A. 2004. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. Am J Hum Genet 75:910–918.

Alegria R. 1965. On Porto Rican archaeology. Am Antiq 3:246–249.

Allaire L. 1980. On the historicity of Carib migrations in the Lesser Antilles. Am Antiq 45:238–245.

Allaire L. 1997. Anse Trabaud: rapport 1997. In: Giraud JP, editor. Le néolithique de la Martinique dans son contexte antillais. (Projet collectif de recherche) SRA Fort de France, Martiniques) (The sites of Norman Estate, Anse des Pères, and Hope Estate with a contribution to the 'La Hueca problem'), Vol. 4. Leiden: Archaeological Studies Leiden University. p 19–22.

Alves-Silva J, Santos MS, Guimaraes PEM, Ferreira ACS, Bandelt HJ, Pena SDJ, Ferreira Prado V. 2000. The ancestry of Brazilian mtDNA lineages. Am J Hum Genet 67: 144–161.

Anderson S, Bankier AT, Barrell BG, DeBruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlick DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG. 1981. Sequence

and organization of the human mitochondrial genome. Nature 290:457–465.

Andrews RM, Kubacka I, Chinnery PF, Lightowlers N, Turnbull DM, Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet 23:147.

Arredi B, Poloni ES, Paracchini S, Zerjal T, Fathallah DM, Makrelouf M, Pascali VL, Novelletto A, Tyler-Smith C. 2004. A predominantly Neolithic origin for Y-chromosomal DNA variation in North Africa. Am J Hum Genet 75:338–345.

Bandelt HJ, Alves-Silva J, Guimaraes PEM, Santos MS, Brehm A, Pereira L, Coppa A, Larruga JM, Rengo C, Scozzari R, Torroni A, Prata MJ, Amorim A, Prado VF, Pena SDJ. 2001. Phylogeography of the human mitochondrial haplogroup L3e: a snapshot of African prehistory and Atlantic slave trade. Ann Hum Genet 65:549–563.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48.

Bandelt H-J, Quintana-Murci L, Salas A, Macaulay V. 2002. The fingerprint of phantom mutations in mitochondrial DNA data. Am J Hum Genet 71:1150–1160.

Benn Torres J, Bonilla C, Robbins CM, Waterman L, Moses TY, Hernandez W, Santos ER, Bennet F, Aiken W, Tullock T, Coard K, Hennis A, Wu S, Nemesure B, Leske CM, Freeman V, Carpten J, Kittles RA. 2008. Admixture and population stratification in African Caribbean populations. Ann Hum Genet 72(Pt 1):90–98.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2010. GenBank. Nucleic Acids Res 38(Suppl 1):D46–D51.

Bolnick DA, Bolnick DI, Smith DG. 2006. Asymmetric male and female genetic histories among Native Americans from eastern North America. Mol Biol Evol 23:2161–2174.

Bolnick DA, Smith DG. 2003. Unexpected patterns of mitochondrial DNA variation among Native Americans from the southeastern United States. Am J Phys Anthropol 122:336–354.

Boyd-Bowman P. 1956. The regional origins of the earliest spanish colonists of America. PMLA 71:1152–1172.

Bryan A. 1973. Paleoenvironments and cultural diversity in late Pleistocene South America. Quaternary Res 3:237–256.

Bryc K, Velez C, Karafet T, Moreno-Estrada A, Reynolds A, Auton A, Hammer M, Bustamante CD, Ostrer H. 2010. Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. Proc Nat Acad Sci USA 107:8954–8961.

Campbell L. 1997. American Indian languages: the historical linguistics of native America. New York: Oxford University Press.

Cardenas MMSG, Ribeiro-dos-Santos Â, Santos S, Mansur AJ, Pereira AC, Fridman C. 2013. Assessment of the relationship between self-declared ethnicity, mitochondrial haplogroups and genomic ancestry in Brazilian individuals. PLoS One 8: e62005.

Chandler JF. 2006. Estimating per-locus mutation rates. J Genet Geneal 2:27–33.

Coelho M, Sequeira F, Luiselli D, Beleza S, Rocha J. 2009. On the edge of Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. BMC Evol Biol 9: Article 80.

Crespillo M, Luque JA, Paredes M, Fernandez R, Ramirez E, Valverde JL. 2000. Mitochondrial DNA sequences for 118 individuals from northeastern Spain. Int J Legal Med 114: 130–132.

Crosby AW. 2003. The Columbian exchange: biological and cultural consequences of 1492. Westport: Praeger Publishers.

Cruciani F, La Fratta R, Trombetta B, Santolamazza P, Sellitto D, Colomb EB, Dugojong JM, Crivellaro F, Benincasa T, Pascone R, Moral P, Watson E, Melegh B, Barbujani G, Fuselli S, Vona G, Zagradisnik B, Assum G, Brdicka R, Kozlov AI, Efremov GD, Coppa A, Novelletto A, Scozzari R. 2007. Tracing past human male movements in northern/eastern Africa and western Eurasia: new clues from Y-chromosomal haplogroups E-M78 and J-M12. Mol Biol Evol 24:1300–1311.

Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olckers A, Modiani D, Holmes S, Destro-Bisol G, Coia V, Wallace DC, Oefner PJ, Torroni A, Cavalli-Sforza LL, Scozzari R, Underhill PA. 2002. A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. Am J Hum Genet 70: 1197–1214.

Curtin P. 1969. The Atlantic slave trade. Madison: University Of Wisconsin Press.

Deason ML, Salas A, Newman SP, Macaulay VA, A Morrison EY, Pitsiladis YP. 2012. Interdisciplinary approach to the demography of Jamaica. BMC Evol Biol 12:24.

Dietz JL. 1992. Historia económica de Puerto Rico. Torres Rivera Y, trans from 1989. Puerto Rico: Ediciones Huracán.

Dulik MC, Owings AC, Gaieski JB, Vilar MG, Andre A, Lennie C, Mackenzie MA, Kritsch I, Snowshoe S, Wright R, Martin J, Gibson N, Andrews TD, Schurr TG, Genographic Consortium. 2012a. Y-chromosome analysis reveals genetic divergence and new founding native lineages in Athapaskan- and Eskimoan-speaking populations. Proc Natl Acad Sci USA 109: 8471–8476.

Dulik MC, Zhadanov SI, Osipova LP, Askapuli A, Gau L, Gokcumen O, Rubenstein S, Schurr TG. 2012b. Mitochondrial DNA and Y chromosome variation provides evidence for a recent common ancestry between native Americans and Indigenous Altaians. Am J Hum Genet 90:229–246.

Elhaik E, Greenspan E, Staats S, Krahn T, Tyler-Smith C, Xue Y, Tofanelli S, Francalacci P, Cucca F, Pagani L, Jin L, Li H, Schurr TG, Greenspan B, Spencer Wells RS. 2013. The GenoChip: a new tool for genetic anthropology. Genome Biol Evol 5:1021–1031.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50.

Fernández-Armesto F. 1982. The Canary Islands after the conquest: the making of a colonial society in the early sixteenth century. New York: Oxford University Press.

Fewkes JW. 2009. The aborigines of Puerto Rico and neighboring islands. Tuscaloosa: University of Alabama Press.

Figueredo AE. 1978. The Virgin Islands as a historical frontier between the Tainos and the Caribs. Rev Rev Interam 8:393–399.

Fitzpatrick SM, Keegan WF. 2007. Human impacts and adaptations in the Caribbean Islands: an historical ecology approach. Trans R Soc Edinburgh 98:29–45.

Fitzpatrick SM, Ross AH. 2010. Island shores, distant pasts. Gainesville: University Press of Florida.

Forte M. 2005. Ruins of absence, presence of Caribs: (Post) colonial representations of aboriginality in Trinidad and Tobago. Tallahassee: University of Florida Press.

Gaieski JB, Elhaik E, Owings AC, Vilar MG, Walia A, Wells RS, Schurr TG, Genographic Consortium. 2012. Genetic ancestry and admixture analysis in a Bermudian population reveals evidence of Native American origins consistent with oral histories and genealogies. Am J Hum Genet Suppl 91: 859.

Gaieski JB, Owings AC, Vilar MG, Dulik MC, Gaieski D, Gittelman RM, Lindo J, Gau L, Schurr TG, The Genographic Consortium. 2011. Genetic ancestry and indigenous heritage in a Native American descendant community in Bermuda. Am J Phys Anthropol 146:392–405.

Galloway PK, editor. 2006. The Hernando de Soto expedition: history, historiography, and 'discovery' in the Southeast. Lincoln: University of Nebraska Press.

Gomez-Carballa A, Ignacio-Veiga A, Alvarez-Iglesias V, Pastoriza-Mourelle A, Ruiz Y, Pineda L, Carracedo A, Salas A. 2012. A melting pot of multicontinental mtDNA lineages in admixed Venezuelans. Am J Phys Anthropol 147: 78–87.

Gonçalves R, Freitas A, Branco M, Rosa A, Fernandes AT, Zhivotovsky LA, Underhill PA, Kivisild T, Brehm A. 2005. Y-chromosome lineages from Portugal, Madeira and Açores: record elements of Sephardim and Berber Ancestry. Ann Hum Genet 69:443–454.

Gonçalves VF, Carvalho CMB, Bortolini MC, Bydlowski SP, Pena SD. 2007. The phylogeography of African Brazilians. Hum Hered 65:23–32.

Granberry J. 2013. Indigenous languages of the Caribbean. In: Keegan W, editor. The oxford handbook of Caribbean archeology. Oxford: Oxford University Press. p 61–69.

Granberry J, Vescelius G. 2004. Languages of the Pre-Columbian Antilles. Tuscaloosa: University of Alabama Press.

Gravel S, Zakharia F, Moreno-Estrada A, Byrnes JK, Muzzio M, Rodriguez-Flores JL, Kenny EE, Gignoux CR, Maples BK, Guiblet W, Dutil J, Sandoval K, Bedoya G, The 1000 Genomes Project, Oleksyk TK, Ruiz-Linares A, Burchard EG, Martinez-Cruzado JC, Bustamante CD. 2013. Reconstructing native American migrations from whole-genome and whole-exome data. PLoS Genet 9:e1004023.

Guitar L. 2000. Criollos: the birth of a dynamic new Indo-Afro-European people and culture on Hispaniola. KACIKE: J Caribbean Amerindian History Anthropol 1:1–17.

Haag W. 1964. A comparison of Arawak sites in the Lesser Antilles. Proceedings of the 1st ICSPCLA 1:9–37.

Haines MR, Steckel RH. 2000. A population history of North America. New York: Cambridge University Press.

Hassan HY, Underhill PA, Cavalli-Sforza LL, Ibrahim ME. 2008. Y-chromosome variation among Sudanese: restricted gene flow, concordance with language, geography, and history. Am J Phys Anthropol 137:316–323.

Helgason A, Pálsson G, Pedersen HS, Angulalik E, Gunnarsdóttir ED, Yngvadóttir B. Stefánsson K. 2006. mtDNA variation in Inuit populations of Greenland and Canada: migration history and population structure. Am J Phys Anthropol 130:123–134.

Hofman CL, Bright AJ, Hoogland ML, Keegan WF. 2008. Attractive ideas, desirable goods: examining the Late Ceramic Age relationships between Greater and Lesser Antillean societies. J Island Coastal Arch 3:17–34.

Hoogland ML, Hofman CL. 1999. Expansion of the Taino cacicazgos towards the Lesser Antilles. J Société d Américanistes 85:93–113.

Hünemeier T, Carvalho C, Marrero AR, Salzano FM, Pena J, Danilo S, Bortolini, MC. 2007. Niger-Congo speaking populations and the formation of the Brazilian gene pool: mtDNA and Y-chromosome data. Am J Phys Anthropol 133:854–867.

Huoponen K, Torroni A, Wickman PR, Sellitto D, Gurley DS, Scozzari R, Wallace DC. 1996. Mitochondrial DNA and Y chromosome-specific polymorphisms in the Seminole tribe of Florida. Eur J Hum Genet 5:25–34.

Just RS, Diegoli TM, Saunier JL, Irwin JA, Parsons TJ. 2008. Complete mitochondrial genome sequences for 265 African American and U.S. "Hispanic" individuals. Forensic Sci Int Genet 2:e45–e48.

Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res 18:830–838.

Kaufman T. 1994. The native languages of South America. In: Mosley C, Asher RE, editors. Atlas of the world's languages. London: Routledge. p 46–76.

Keegan WF. 1995. Modeling dispersal in the prehistoric West Indies. World Archaeol 26:400–419.

Keller A, Graefen A, Ball M, Matzas M, Boisguerin V, Maixner F, Leidinger P, Backes C, Khairat R, Forster M, Stade B, Franke A, Mayer J, Spangler J, McLaughlin S, Shah M, Lee C, Harkins TT, Sartori A, Moreno-Estrada A, Henn B, Sikora M, Semino O, Chiaroni J, Roosti S, Myres NM, Cabrera VM, Underhill PA, Bustamante CD, Vigl EE, Samadelli M, Cipollini G, Haas J, Katus J, O'Connor BD, Carlson MRJ, Meder B, Blin N, Meese E, Pusch CM, Zink A.. 2012. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. Nat Commun 3: Article 698.

Kemp BM. 2006. Mesoamerica and Southwest prehistory, and the entrance of humans into the Americas: mitochondrial DNA evidence. Ph.D. Dissertation, University of California, Davis CA.

Kidd JM, Gravel S, Byrnes J, Moreno-Estrada A, Musharoff S, Bryc K, Degenhardt JD, Brisbin A, Sheth V, Chen R, McLaughlin SF, Peckham HE, Omberg L, Bormann Chung CA, Stanley S, Pearlstein K, Levandowsky E, Acevedo-Acevedo S, Auton A, Keinan A, Acuña-Alonzo V, Barquera-Lozano R, Canizales-Quinteros S, Eng C, Burchard EG, Russell A, Reynolds A, Clark AG, Reese MG, Lincoln SE, Butte AJ, De La Vega FM, Bustamante CD. 2012. Population genetic inference from personal genome data: impact of ancestry and admixture on human genomic variation. Am J Hum Genet 91:660–671.

Kumar S, Bellis C, Zlojutro M, Melton PE, Blangero J, Curran JE. 2011. Large scale mitochondrial sequencing in Mexican Americans suggests a reappraisal of Native American origins. BMC Evol Biol 11: Article 293.

Lalueza-Fox C, Gilbert MTP, Martínez-Fuentes MJ, Calafell F, Bertranpetit J. 2003. Mitochondrial DNA from pre-Columbian Ciboneys from Cuba and the prehistoric colonization of the Caribbean. Am J Phys Anthropol 121:97–108.

Lalueza-Fox C, Luna Calderon F, Calafell F, Morera B, Bertranpetit. 2001. MtDNA from extinct Tainos and the peopling of the Caribbean. Ann Hum Genet 65:137–151.

Lander N, Rojas MG, Chiurillo MA, Ramirez JL. 2008. Haplotype diversity in human mitochondrial DNA hypervariable regions I-III in the city of Caracas (Venezuela). Forensic Sci Int Genet 2:e61–e64.

Lappalainen T, Koivumäki S, Salmela E, Huoponen K, Sistonen P, Savontaus ML, Lahermo P. 2006. Regional differences among the Finns: a Y-chromosomal perspective. Gene 376: 207–215.

Las Casas B, Griffin N, editor and translator. 1992. A short account of the destruction of the Indies. London: Penguin Books.

Lovejoy PE. 2000. Transformations in Slavery: a history of slavery in Africa. Cambridge: Cambridge University Press.

Lundberg ER. 1989. Preceramic procurement patterns at Krum Bay, Virgin Islands. Ph.D. Dissertation. University of Illinois, Urbana-Champaign, IL.

Maca-Meyer N, González AM, Pestano J, Flores C, Larruga JM, Cabrera VM. 2003. Mitochondrial DNA transit between West Asia and North Africa inferred from U6 phylogeography. BMC Genet 4: Article 15.

Magnus M. 1995. Spanish historians on Spanish migration to America during the colonial period. Lat Am Res Rev 30: 251–255.

Mandarino LM. 2010. Mitochondrial DNA analysis of the Brazilian Macushi population. Master's Thesis, Department of Anthropology, State University of New York at Binghamton.

Martin C, Wasserman M. 2005. Latin America and its people. New York: Allyn & Bacon.

Martinez H, Rodriguez-Larralde R, Izaguirre MH, Castro De Guerra D. 2007. Admixture estimates for Caracas, Venezuela, based on autosomal, Y-chromosome, and mtDNA markers. Hum Biol 79(2): Article 5.

Martínez-Cruzado JC. 2010. The history of Amerindian mitochondrial DNA lineages in Puerto Rico. In: Fitzpatrick SM, Ross A, editors. Island shores, distant pasts: archaeological and biological approaches to the pre-Columbian settlement of the Caribbean. Gainesville: University of Florida. p 54–80.

Martínez-Cruzado JC, Toro-Labrador G, Ho-Fung V, Estevez-Montero, M Lobaina-Manzanet A, Padovani-Claudio DA, Sanchez-Cruz H, Ortiz-Bermudez P, Sanchez-Crespo A. 2001. Mitochondrial DNA analysis reveals substantial Native American ancestry in Puerto Rico. Hum Biol 73:491–511.

Martínez-Cruzado JC, Toro-Labrador G, Viera-Vera J, Rivera-Vega MY, Startek J, Latorre-Esteves M, Román-Colón A, Rivera-Torres R, Navarro-Millán IY, Gómez-Sánchez E, Caro-González HY, Valencia-Rivera P. 2005. Reconstructing the population history of Puerto Rico by means of mtDNA phylogeographic analysis. Am J Phys Anthropol 128:131–155.

Melton PE, Baldi NF, Barrantes R, Crawford MH. 2013. Microevolution, migration, and the population structure of five Amerindian populations from Nicaragua and Costa Rica. Am J Hum Biol 25:480–490.

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. BMC Evol Bio 8: Article 213.

Moreno-Estrada A, Gravel S, Zakharia F, McCauley JL, Byrnes JK, Gignoux CR, Ortiz-Tello PA, Martínez RJ, Hedges DJ, Morris RW, Eng C, Sandoval K, Acevedo-Acevedo S, Norman PJ, Layrisse Z, Parham P, Martínez-Cruzado JC, Burchard EG, Cuccaro ML, Martin ER, Bustamante CD. 2013. Reconstructing the population genetic history of the Caribbean. PLoS Genet 9:e1003925.

Morisson SE. 1939. The second voyage of Christopher Columbus from Cadiz to Hispaniola and the discovery of the Lesser Antilles. Oxford: The Clarendon Press.

Musicant I. 1998. Empire by default: the Spanish-American war and the dawn of the American century. New York: H. Holt.

Myres NM, Rootsi S, Lin AA, Jðrve M, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EL, Pshenichnov A, Yunusbayev B, Balanovska O, Balanovska E, Rudan P, Baldovic M, Herrera RJ, Chiaroni J, Di Cristoaro J, Villems R, Kivisild T, Underhill PA. 2010. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. Eur J Hum Genet 19:95–101.

O'Rourke DH, Raff JA. 2010. The human genetic history of the Americas: the final frontier. Curr Biol 20:R202–R207.

Ortiz J. 1976. Excavations at the preceramic Cerillo site, Southwest Puerto Rico. In: Proceedings of the sixth international congress for the study of pre-Columbian cultures of the Lesser Antilles, Centre Universitaire, Guadelope. p 269–271.

Owings AC, Dulik MC, Zhadanov SI, Gaieski JB, Cope A, Dorey D, Schurr TG, The Genographic Consortium. 2012. Genetic variation in Mi'kmaq populations from Nova Scotia and its implications for the history of Algonquian populations in northeastern North America. Am J Phys Anthropol Suppl 147(S54):231.

Pantel AG. 1976. Excavations at the preceramic Cerillo site, southwest Puerto Rico. Sixth IACA, Pointe-à-Pitre, Guadeloupe. p 269–271.

Petersen JB, Hofman CL, Curet LA. 2005. Time and culture: chronology and taxonomy in the eastern Caribbean and the Guianas. In: Delpuech A, Hofman CL, editors. Late Ceramic Age Societies in the Eastern Caribbean. British Archaeological Reports, 1273. Oxford: Archaeopress. p 17–32.

Plaza S, Salas A, Calafell F, Corte-Real F, Bertranpetit J, Carracedo A, Comas D. 2004. Insights into the western Bantu dispersal: mtDNA lineage analysis in Angola. Hum Genet 115:439–447.

Rando JC, Pinto F, González AM, Hernández M, Larruga LM, Cabrera VM, Bandelt H-J. 1998. Mitochondrial DNA analysis of Northwest African populations reveals genetic exchanges with European, Near-Eastern, and sub-Saharan populations. Ann Hum Genet 62:531–550.

Raudzens G. 2003. Technology, disease and colonial conquests, sixteenth to eighteenth centuries. Leiden: Brill.

Rawley JA, Behrendt SD. 2005. The Transatlantic Slave Trade: a history. Lincoln: University of Nebraska Press.

Richards MB, Macaulay VA, Bandelt HJ, Sykes BC. 1998. Phylogeography of mitochondrial DNA in Western Europe. Ann Hum Genet 62:241–260.

Rieder MJ, Taylor SL, Tobe VO, Nickerson DA. 1998. Automating the identification of DNA variations using quality-based fluorescence re-sequencing: analysis of the human mitochondrial genome. Nucleic Acid Res 26:967–973.

Righter E. 1997. The ceramics, art, and material culture of the early ceramic period in the Caribbean Islands. In: Wilson S, editor. The indigenous people of the Caribbean. Gainesville: University of Florida Press. p 70–79.

Righter E, Wild KS, Lundberg ER. 2004. Late Ceramic Age developments in the Virgins Archipelago: the Puerto Rican connection. In: Late Ceramic Age Societies in the Eastern Caribbean, Delpuech A, Hofman CL, editors. BAR International Series #1273. Paris Monographs in American Archaeology #14. Oxford: Archaeopress. p 101–120.

Rodriguez-Ramos R. 2010. Rethinking Puerto Rican precolonial history. Tuscaloosa: University of Alabama Press.

Rogozinski J. 1999. A brief history of the Caribbean: from the Arawak to the present. New York: Penguin Publishing.

Rootsi S, Myres NM, Lin AA, Järve M, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Varendi K, Sahakyan H, Behar DM, Khusainova R, Balanovsky O, Balanovska E, Rudan P, Yepiskoposyan L, Bahmanimehr A, Farjadian S, Kushniarevich A, Herrera RJ, Grugni V, Battaglia V, Nici C, Crobu F, Karachanak S, Hooshiar Kashani B, Houshmand M, Sanati MH, Toncheva D, Lisa A, Semino O, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA. 2012. Distinguishing the co-ancestries of haplogroup G Y-chromosomes in the populations of Europe and the Caucasus. Eur J Hum Genet 20:1275–1282.

Rosa A, Brehm A, Kivisild T, Metspalu E, Villems R. 2004. MtDNA profile of West Africa Guineans: towards a better understanding of the Senegambia region. Ann Hum Genet 68: 340–352.

Rosa A, Ornelas C, Jobling MA, Brehm A, Villems R. 2007. Y-chromosomal diversity in the population of Guinea-Bissau: a multiethnic perspective. BMC Evol Biol 7: Article 124.

Rouse IR. 1952. Porto Rican prehistory: introduction; Excavations in the west and north. In: Scientific survey of Porto Rico and the Virgin Islands, Vol. 18, Part 3. New York: New York Academy of Sciences.

Rouse IR. 1964. Prehistory of the West Indies. Science 144:499–508.

Rouse IR. 1992. The Tainos: rise and decline of the people who greeted Columbus. New Haven: Yale University Press.

Salas A, Acosta A, Alvaerz-Iglesias V, Cerezo M, Phillips C, Lareau MV, Carracedo A. 2008. The mtDNA ancestry of admixed Colombian populations. Am J Hum Biol 20:584–591.

Salas A, Lovo-Gómez J, Álvarez-Iglesias V, Cerezo M, Lareu MV, Macaulay V, Richards MB, Carracedo A. 2009. Mitochondrial echoes of first settlement and genetic continuity in El Salvador. PLoS One 4:e6882.

Salas A, Richards M, Lareau MV, Sorbino B, Silva S, Matamoros M, Macaulay V, Carracedo A. 2005. Shipwrecks and founder effects: divergent demographic histories revealed in Caribbean mtDNA. Am J Phys Anthropol 128:855–860.

Sandoval K, Buentello-Malo L, Peñaloza-Espinosa R, Avelino H, Salas A, Calafell F, Comas D. 2009. Linguistic and maternal genetic diversity are not correlated in Native Americans. Hum Genet 126:521–531.

Sandoval K, Moreno-Estrada A, Mendizabal I, Underhill PA, Lopez-Valenzuela M, Peñaloza-Espinosa R, Comas D. 2012. Y-chromosome diversity in Native Mexicans reveals continental transition of genetic structure in the Americas. Am J Phys Anthropol 148:395–405.

Schurr TG. 2004. The peopling of the New World: perspectives from molecular anthropology. Ann Rev Anthropol 33:551–583.

Schurr TG. 2010. Coastal waves and island hopping: a genetic view of Caribbean prehistory in the context of New World colonization. In: SM Fitzpatrick, A Ross, editors. Island shores, distant pasts: archaeological and biological approaches to the pre-Columbian settlement of the Caribbean. Gainesville: University of Florida Press. p 177–198.

Schurr TG, Dulik MC, Owings AC, Zhadanov SI, Gaieski JB, Vilar MG, Ramos J, Moss MB, Natkong, F. 2012a. Clan, language, and migration history has shaped genetic diversity in Haida and Tlingit populations from Southeast Alaska. Am J Phys Anthropol 148:422–435.

Schurr TG, Sherry ST. 2004. Mitochondrial DNA and Y-chromosome diversity and the peopling of the Americas. Am J Hum Biol 16:420–439.

Schurr TG, Vilar MG, Gomez MR, Zillges H, Sanders A, Gaieski JB Owings AC, Genographic Consortium. 2012b. Genetic diversity and migration history of indigenous populations from the Central Valley of Mexico. Am J Hum Genet Suppl 91:871.

Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery N, Battaglia V. Maccioni L, Triantaphyllidis C, Shen P, Oefner PJ, Zhivotovsky LA, King R, Torroni A, Cavalli-Sforza LL, Underhill

PA, Santachiara-Benerecetti AS. 2004. Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the Neolithization of Europe and later migratory events in the Mediterranean area. Am J Hum Genet 74:1023–1034.

Siegel PE, editor. 2005. Ancient Borinquen: archaeology and ethnohistory of Native Puerto Rico. Tuscaloosa: University of Alabama Press.

Simms TM, Rodriguez CE, Rodriguez R, Herrera RJ. 2010. The genetic structure of populations from Haiti and Jamaica reflect divergent demographic histories. Am J Phys Anthropol 142:49–66.

Simoni L, Calafell F, Pettener D, Bertranpetit J, Barbujani G. 2000. Geographic patterns of mtDNA diversity in Europe. Am J Hum Genet 66:262–278.

Soares P, Ermini L, Thomson N, Mormina M, Rito T, Röhl A, Salas A, Oppenheimer S, Macaulay V, Richards MB. 2009. Correcting for purifying selection: an improved human mitochondrial molecular clock. Am J Hum Genet 84:740–759.

SPSS Inc. 2008. SPSS statistics for windows, version 17.0. Chicago: SPSS Inc.

Stefflova K, Dulik MC, Pai AA, Walker AH, Zeigler-Johnson CM, Gueye SM, Schurr TG, Rebbeck TR. 2009. Evaluation of group genetic ancestry of populations from Philadelphia and Dakar in the context of sex-biased admixture in the Americas. PLoS One 4:e7842.

Tamm E, Kivisild T, Reidla M, Metspalu M, Glenn-Smith D, Mulligan CJ, Bravi CM, Rickards, O, Martínez-Lambarga C, Khusnutdinova EK, Fedorova SA, Golubenko MV, Stepanov VA, Gubina MA, Zhadanov SI, Osipova LP, Damba L, Voevoda MO, Dipierri JE, Villems R, Malhi RS. 2007. Beringian standstill and spread of Native American founders. PLoS One 2: e829.

The Cuban mtDNA Group. 2013. Family tree geographical projects, Houston, TX. Available at: http://www.familytreedna.com/public/CubaDNAProject/default.aspx. Accessed on 10 September 2013.

The Puerto Rican Geographical DNA Group. 2013. Family tree geographical projects, Houston, TX. Available at: http://www.familytreedna.com/public/puertoricansurname/. Accessed 10 September 2013.

Toro-Labrador G, Wever OR, Martínez-Cruzado JC. 2003. Mitochondrial DNA analysis in Aruba: strong maternal ancestry of closely related Amerindians and implications for the peopling of northwestern Venezuela. Carib J Sci 39:11–22.

Torroni A, Rengo C, Guida V, Cruciani F, Sellitto D, Coppa A, Calderon FL, Simionati B, Valle G, Richards M, Macaulay V, Scozzari R. 2001. Do the four clades of the mtDNA haplogroup L2 evolve at different rates? Am J Hum Genet 69: 1348–1356.

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo C, Wallace DC. 1993. Asian affinities and continental radiation of the four founding Native American mitochondrial DNAs. Am J Hum Genet 53:563–590.

van Oven M, Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat 30:E386–E394.

Veeramah K, Connell B, Pour N, Powell A, Plaster C, Zeitlyn D, Mendell M, Weale M, Bradman N, Thomas M. 2010. Little genetic differentiation as assessed by uniparental markers in the presence of substantial language variation in peoples of the Cross River region of Nigeria. BMC Evol Biol 10: Article 92.

Via M, Gignoux CR, Roth LA, Fejerman L, Galanter J, Choudry S, Toro-Labrador G, Viera-Vega J, Oleksyk TK, Beckman K, Ziv E, Risch N, Burchard EG, Martínez-Cruzsado JC. 2011. History shaped the geographic distribution of genomic admixture of the island of Puerto Rico. PLoS One 6:e16513.

Vilar MG, Dulik MC, Owings AC, Gaieski JB, Lennie C, Zillges H, Gittelman R, Lindo J, Keating K, Schurr TG. 2014. The Genographic Consortium. Genetic variation in Inuvialuit populations from the Northwest Territories and its implications for the human colonization of the Arctic. Am. J Phys Anthropol, in review.

Vilar MG, Gomez R, Zillges H, Brooks D, Sanders A, Gaieski JB, Owings A, Meraz MA, Schurr TG, Genographic Consortium. 2013. Genetic diversity in indigenous populations from Central Mexico and its implications for the peopling of the Americas. Am J Phys Anthropol Suppl 146:278.

Vilar MG, Lum JK, Melendez C, Owings A, Gaieski JB, Schurr TG, Genographic Consortium. 2011. Analysis of mtDNA haplogroup A2 frequency, distribution and diversity among populations from North and Central America and the Caribbean. Am J Phys Anthropol Suppl 144(S52):300.

Willey G. 1956. Settlement patterns in the Caribbean area. In: Willey G, editor. Prehistoric settlement patterns in the new world (Viking Fund Publications in Anthropology, No. 23). New York: Wenner-Gren Foundation for Anthropological Research. p 165–172.

Wilson SM. 2007. The archeology of the Caribbean. Cambridge: Cambridge University Press.

Wilson SM. 1997. The legacy of the indigenous people of the Caribbean. In: Wilson SM, editor. The indigenous people of the Caribbean. Gainesville: University of Florida Press. p 206–213.

Wilson JL, Saint-Louis V, Auguste JO, Jackson BA. 2012. Forensic analysis of mtDNA haplotypes from two rural communities in Haiti reflects their population history. J Forensic Sci 57:1457–1466.

Y Chromosome Consortium. 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res 12:339–348.

Zhadanov SI, Dulik MC, Markley M, Jennings G, Gaieski JB, Elias G, Schurr TG, The Genographic Consortium. 2010. Genetic heritage and native identity of the Seaconke Wampanoag tribe of Massachusetts. Am J Phys Anthropol 142:579–589.