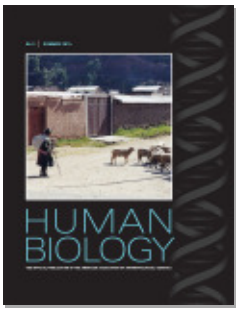
Population Genetic Structure of Traditional Populations in the Peruvian Central Andes and Implications for South American Population History

Graciela S. Cabana, Cecil M. Lewis Jr., Raúl Y. Tito, R. Alan Covey, Angela M. Cáceres, Augusto F. De La Cruz, Diana Durand, Genevieve Housman, Brannon I. Hulsey, Gian Carlo Iannacone, Paul W. López, Rolando Martínez, Ángel Medina, Olimpio Ortega Dávila, Karla Paloma Osorio Pinto, Susan I. Polo Santillán, Percy Rojas Domínguez, Meagan Rubel, Heather F. Smith, Silvia E. Smith, Verónica Rubín de Celis Massa, Beatriz Lizárraga, Anne C. Stone



Human Biology, Volume 86, Number 3, Summer 2014, pp. 147-165 (Article)

Published by Wayne State University Press

Errata

■ In issue 86.3 of *Human Biology*, "Population Genetic Structure of Traditional Populations in the Peruvian Central Andes and Implications for South American Prehistory" by Cabana et al. had the following error entered into Table 1: The population samples of Cara Cara, Puno, and San Juan del Oro were incorrectly assigned to the Peruvian *region* (province) of Cusco; their correct *region* is Puno.

Table 1. Peruvian Central Andean Populations: Sample Locations and Data Sources

Population Sample	Región	Geographic Coordinates		mtDNA		Y-Chromosome	
		Latitude	Longitude	Reference	*n*	Reference	*n*
Amantaní	Puno	–15.67	–69.71	Sandoval et al. 2013	26	Sandoval et al. 2013	26
Ancash	Ancash	–9.33	–77.56	Lewis et al. 2004; 2007	73	This study	41
Andahuaylas	Apurimac	–13.66	–73.38	This study	56	This study	35
Arequipa	Arequipa	–13.13	–71.11	Fuselli et al. 2003	22	Tarazona–Santos et al. 2001	15
Cajamarca	Cajamarca	–7.16	–78.51	Sandoval et al. 2013	19	Sandoval et al. 2013	19
Caleta Santa Rosa	Lambayeque	–7.13	–79.55	This study	31	––	–
Capachica	Puno	–15.67	–69.85	Sandoval et al. 2013	15	Sandoval et al. 2013	15
Cara Cara	Puno	–15.48	–70.35	This study	26	This study	28
Catacaos	Piura	–5.27	–80.68	This study	16	––	–
Chimú	Puno	–15.86	–69.95	Sandoval et al. 2013	16	Sandoval et al. 2013	16
Cusco	Cusco	–13.52	–71.97	Sandoval et al. 2013	36	Sandoval et al. 2013	36
Cusco–North	Cusco	–13.52	–71.98	This study	34	This study	54
Cusco–South	Cusco	–14.6	–71.25	This study	95	––	–
Huancapi	Ayacucho	–13.66	–74.05	This study	13	––	–
Huancavelica	Huancavelica	–12.93	–75.15	Sandoval et al. 2013	26	Sandoval et al. 2013	26
Islilla	Lambayeque	–5.12	–81.11	This study	18	––	–
Los Uros	Puno	–15.74	–69.93	Sandoval et al. 2013	25	Sandoval et al. 2013	25
Matsiguenga	Cusco	–12	–72.30	Mazières et al. 2008	38	Mazières et al. 2008	13
Otuzco	La Libertad	–7.91	–78.56	This study	26	––	–
Oxapampa	Pasco	–10.59	–75.40	Sandoval et al. 2013	18	Sandoval et al. 2013	18
Picota	San Martín	–8.2	–75.98	This study	21	This study	21
Picota–Centro	San Martín	–6.92	–76.33	This study	22	This study	22
Puno	Puno	–15.84	–70.02	This study	69	––	–
San Juan del Oro	Puno	–14.12	–69.12	This study	60	This study	56
San Martín de Pangoa	Junín	–11.43	–74.48	Fuselli et al. 2003	17	––	–
Santa Rosa de Yanaque	Puno	–15.95	–69.67	Sandoval et al. 2013	18	Sandoval et al. 2013	18
Santiago de Chuco	La Libertad	–8.15	–78.18	This study	21	This study	13
Taquile	Puno	–15.77	–69.68	Sandoval et al. 2013	35	Sandoval et al. 2013	35
Tayacaja	Tayacaja	–12.24	–74.34	Fuselli et al. 2003	63	Tarazona–Santos et al. 2001	44
Trujillo	La Libertad	–8.11	–79.03	This study	13	––	–
Tupe	Lima	–12.74	–75.81	This study	17	––	–

Population Genetic Structure of Traditional Populations in the Peruvian Central Andes and Implications for South American Population History

*Graciela S. Cabana,[1†] Cecil M. Lewis, Jr.,[2†] Raúl Y. Tito,[2] R. Alan Covey,[3] Angela M. Cáceres,[4] Augusto F. De La Cruz,[4] Diana Durand,[5] Genevieve Housman,[6] Brannon I. Hulsey,[1] Gian Carlo Iannacone,[7] Paul W. López,[4] Rolando Martínez,[4] Ángel Medina,[4] Olimpio Ortega Dávila,[4] Karla Paloma Osorio Pinto,[4] Susan I. Polo Santillán,[4] Percy Rojas Domínguez,[4] Meagan Rubel,[8] Heather F. Smith,[9] Silvia E. Smith,[10] Verónica Rubín de Celis Massa,[11] Beatriz Lizárraga,[4] and Anne C. Stone[6]**

ABSTRACT

Molecular-based characterizations of Andean peoples are traditionally conducted in the service of elucidating continent-level evolutionary processes in South America. Consequently, genetic variation among "western" Andean populations is often represented in relation to variation among "eastern" Amazon and Orinoco River Basin populations. This west-east contrast in patterns of population genetic variation is typically attributed to large-scale phenomena, such as dual founder colonization events or differing long-term microevolutionary histories. However, alternative explanations that consider the nature and causes of population genetic diversity *within* the Andean region remain underexplored. Here we examine population genetic diversity in the Peruvian Central Andes using data from the mtDNA first hypervariable region and Y-chromosome short tandem repeats among 17 newly sampled populations and 15 published samples. Using this geographically comprehensive data set, we first reassessed the currently accepted pattern of western versus eastern population genetic structure, which our results ultimately reject: mtDNA population diversities were lower, rather than higher, within Andean versus eastern populations, and only highland Y-chromosomes exhibited significantly higher within-population diversities

[1]Department of Anthropology, University of Tennessee, Knoxville, Tennessee.

[2]Department of Anthropology, University of Oklahoma, Norman, Oklahoma.

[3]Department of Anthropology, University of Texas, Austin, Texas.

[4]Centro de Investigacíon de Bioquímica y Nutrición, Facultad de Medicina, Universidad Nacional Mayor de San Marcos, Lima, Perú.

[5]Department of Anthropology, University of Arkansas, Fayetteville, Arkansas.

[6]School of Human Evolution and Social Change, Arizona State University, Tempe, Arizona.

[7]Laboratorio de Biología Molecular y de Genética, Instituto de Medicina Legal, Ministerio Público, Lima, Perú.

[8]Department of Anthropology, University of Pennsylvania, Philadelphia, Pennsylvania.

[9]Department of Anatomy, Midwestern University, Glendale, Arizona.

[10]Department of Pulmonary Medicine, University of Idaho, Moscow, Idaho.

[11]Laboratorio de Genómica y Biología Molecular Evolutiva, Facultad de Biología, Universidad Ricardo Palma, Lima, Perú.

*Correspondence to: Anne C. Stone, School of Human Evolution and Social Change, Arizona State University, Tempe, AZ 85287. E-mail: acstone@asu.edu.

†These associate editors of *Human Biology* were excluded from the editorial process for and any decisions by the journal about this article, to minimize conflict of interest.

KEY WORDS: MTDNA, Y-CHROMOSOME, NATIVE AMERICANS, ANDEAN POPULATIONS, AMAZONIAN POPULATIONS, PERU, PEOPLING OF SOUTH AMERICA.

Human Biology, Summer 2014, v. 86, no. 3, pp. 147–165.

compared with eastern groups. Multiple populations, including several highland samples, exhibited low genetic diversities for both genetic systems. Second, we explored whether the implementation of Inca state and Spanish colonial policies starting at about AD 1400 could have substantially restructured population genetic variation and consequently constitute a primary explanation for the extant pattern of population diversity in the Peruvian Central Andes. Our results suggest that Peruvian Central Andean population structure cannot be parsimoniously explained as the sole outcome of combined Inca and Spanish policies on the region's population demography: highland populations differed from coastal and lowland populations in mtDNA genetic structure only; highland groups also showed strong evidence of female-biased gene flow and/or effective sizes relative to other Peruvian ecozones. Taken together, these findings indicate that population genetic structure in the Peruvian Central Andes is considerably more complex than previously reported and that characterizations of and explanations for genetic variation may be best pursued within more localized regions and defined time periods.

Since at least the 1960s, characterizations of Andean peoples from a molecular perspective have been conducted peripherally, as part of larger investigative efforts in South America as a whole. These efforts have had two main goals: first, to inform debates on the peopling of South America and of the Americas writ large, and second, to contribute to broad understandings of human micro- to macroevolutionary change. Both goals were vitalized by the work of James V. Neel, Francisco M. Salzano, and their colleagues among tribal groups in Venezuela and Brazil. Thus, from the 1960s into the 1990s, researchers working in South America conducted numerous regional studies based on "classical" markers (blood groups and proteins), primarily among peoples from the Amazon and Orinoco River Basins (e.g., Matson et al. 1966; Neel 1970, 1978a, 1978b; Salzano et al. 1973, 1974, 1977; Ward et al. 1975; O'Rourke and Suarez 1985), and only occasionally among peoples in the Andean region (Best et al. 1966; Matson et al. 1966; Modiano et al. 1972).

It was not until the late 1980s that researchers began to focus explicitly on the genetic history and structure of Andean peoples. Rothhammer and Silva (1989) were the first to argue for a distinct Andean pattern of biological variation as ascertained via craniometric isoline maps; this interpretation was supported by follow-up studies of craniometric and classical marker data (Rothhammer and Silva 1990, 1992; Cavalli-Sforza et al. 1994). Beginning in the early 2000s, spatial analyses of molecular data revealed that Andean populations contain a pattern of high within- and low between-population genetic variation compared with populations in the Amazon and Orinoco River Basins, which appear to harbor much higher levels of between-group variation (Luiselli et al. 2000; Simoni et al. 2000; Tarazona-Santos et al. 2001; Fuselli et al. 2003).

Subsequently, a distinction between "west" (i.e., the Andean mountain range and surroundings regions) and "east" (i.e., Amazonian lowlands and surrounding regions) has prevailed in discussions of continental South American population genetic variation (e.g., Pucciarelli et al. 2006; Lewis et al. 2007b; Wang et al. 2007; Lewis and Long 2008; Lewis 2009; Yang et al. 2010). This dichotomized representation of genetic variation in South American populations, in turn, has had important implications for the reconstruction of the continent's early population history. Whereas some researchers view the contrasting patterns as the outcome of two distinct initial colonizing events (Luiselli et al. 2000; Rodríguez-Delfín et al. 2001), others posit that the two regions experienced differing population histories only after the initial entry into South America through the Panamanian isthmus (Tarazona-Santos et al. 2001; Fuselli et al. 2003; Rothhammer and Dillehay 2009; Yang et al. 2010; Bodner et al. 2012).

However, the above characterization of Andean diversity may be incomplete, for three reasons. First, it derives from a small number of population samples, most of which contain few individuals. Until recently, only 14 western Andean populations have been represented in the scholarly literature; these populations are spread over more than 4,000 km, ranging from Ecuador to Peru, northwest Argentina, and northern Chile. Many of the post-1990s molecular studies are based on repeated genotyping efforts of an even smaller set of population samples (e.g., Luiselli et al. 2000;

Simoni et al. 2000; Tarazona-Santos et al. 2001; Fuselli et al. 2003; Battilana et al. 2006; Scliar et al. 2012; Roewer et al. 2013). Second, most relevant studies are based on either mtDNA or Y-chromosome polymorphisms, which, in isolation, contain limited information about amounts and patterns of diversity (see Lewis and Long 2008).

A third weakness of the current perspective is that views on the nature and causes of population genetic diversity within the western Andean region of South America remain underdeveloped. This is despite the fact that archaeological and ethnohistorical studies attest to complex and varying human-environment interactions through time and space in the Andes (Sandweiss et al. 2001; Wernke 2007; Gosling and Williams 2013). Specifically, early patterns may have been most recently overlaid by Inca state practices and, later, by Spanish colonial-driven policies. For example, Inca rulers routinely moved individuals (males and females) and entire families, often throughout locales in the mountainous highlands [e.g., Matienzo 1967 (1567); Cieza de León 1985 (c. 1550); Wightman 1990], but also throughout the coast and areas to the south of the Lake Titicaca basin (Covey and Elson 2007). Subsequent European invasions commencing in the 1530s precipitated the rapid decline of many indigenous coastal communities, whereas highland communities had higher initial rates of survival and were used for long-distance service (tributary) activities (Cook 1980). The Spanish conquest reinforced a frontier between the highlands and the Amazon Basin, while colonial policies indirectly encouraged population mixture and migration within the Andean region.

The expected cumulative outcome of these phenomena on present-day indigenous population diversity is an overall maintenance of genetic variation combined with a pronounced lack of population structure throughout the region, particularly in the highland regions. It is therefore possible that the region's extant pattern of population genetic variation is best explained by events from the last half millennium, instead of or in addition to early migratory phenomena.

Given the current paucity of genetic data representing Andean indigenous peoples, we may have an unclear picture of the structure of Andean population genetic diversity, including the role of postpeopling events in shaping this structure (Pucciarelli et al. 2006; Dillehay 2009). The present study reexamines Andean genetic diversity using both mtDNA and Y-chromosome data from a large sample of traditional populations from within the Peruvian Central Andes, a region that includes the coast and highlands of Peru, as well as the tropical forest that begins on the eastern slopes of the Andes and descends to the Amazon Basin (Von Hagen and Morris 1998: 14).

Our study contributes 17 population samples from multiple ecological zones (highlands, coast, and lowlands) associated with broad cultural patterns revealed by archaeological, ethnohistorical, and ethnographic studies (see D'Altroy 2015). We investigate the region's population genetic structure to test two hypotheses: (1) In response to previous characterizations of population genetic diversity, we tested the hypothesis that populations throughout the Peruvian Central Andes in general, and in highland areas in particular, show little or no population genetic structure in the form of high within- and low between-population genetic variation relative to eastern regions of South America. (2) We tested whether extant patterns of variation within the Peruvian Central Andes can be parsimoniously explained by events from the last half millennium. Given that both Inca and Spanish policies were strongly implemented among all individuals (i.e., both males and females), primarily within highland communities, we explored (a) whether the pattern of genetic structure differs in highland versus coastal and lowland communities in the Peruvian Central Andes and (b) whether any sex biases in patterns of variation can be detected.

Materials and Methods

Population Samples

This study is based on multilocus genetic data from samples collected through an international collaborative effort within the Republic of Peru between 2001 and 2005. The senior author (A.C.S.) moved from the University of New Mexico (UNM) to Arizona State University (ASU) during the course of the project. Thus, this project was approved by institutional review boards from both UNM and ASU, as well as from the Universidad Nacional Mayor de San Marcos and the Universidad Ricardo Palma, Lima, Peru. Biological samples in the form of buccal

FIGURE 1. Map of Peru and eastern South America (*inset*) with sampling locations. For the Peruvian Central Andes, coastal samples are designated with stars, highland samples with circles, and lowland samples with squares. Eastern samples from the Amazon and Orinoco River Basins are designated with Xs.

cells or blood spots were collected with informed consent from 611 individuals from 17 populations distributed throughout 9 of 25 *regiones* (first-level political and administrative subdivisions) and three ecological contexts in Peru (Figure 1).

The Andes range defines a succession of ecozones running from the Pacific coastal desert, across areas of montane valleys and high tundra, and onto the humid eastern slope that descends to the Amazon Basin. Although mindful of the considerable local environmental diversity of the Andean region, we have found it useful to group our samples into three regions: coastal (areas lying to the west of the Andes), highland (areas lying in the mountain valleys), and lowland (areas lying in the humid tropics of the eastern escarpment).

Populations were assigned sample names according to the city, village, or locale from which individual biological samples were taken (Table 1, Figure 1). We adopt a geographically based nomenclature given that the samples reflect relatively vulnerable populations, therefore warranting a level of conservatism in community identification (see Obregón-Tito 2013). Each individual in the sampled populations was from one of three ecological contexts: coast (n = 78), highlands (n = 430), and lowlands (n = 103). Individuals in the coastal sample were from Caleta Santa Rosa (n = 31), Catacaos (n = 16), Islilla (n = 18), and Trujillo (n = 13). Individuals in the highlands sample were from Ancash (n = 73), Andahuaylas (n = 56), Cara Cara (n = 26), Cusco North (n = 34), Cusco South (n = 95), Huancapi (n = 13), Otuzco (n = 26), Puno (n = 69), Santiago de Chuco (n = 21), and Tupe (n = 17). The lowlands sample includes individuals from Picota (n = 21), Picota Centro (n = 22), and San Juan del Oro (n = 60). All population samples were rural indigenous communities with the exception of Ancash, which comprises recent local rural migrants to the cities of Huaraz, Chancay, and Lima or from the small town of Yungay in the *región* of Ancash. Inclusion criteria included unrelated individuals

Table 1. Peruvian Central Andean Populations: Sample Locations and Data Sources

Population Sample	Región	Geographic Coordinates		mtDNA		Y-Chromosome	
		Latitude	Longitude	Reference	*n*	Reference	*n*
Amantaní	Puno	–15.67	–69.71	Sandoval et al. 2013	26	Sandoval et al. 2013	26
Ancash	Ancash	–9.33	–77.56	Lewis et al. 2004, 2007b	73	This study	41
Andahuaylas	Apurimac	–13.66	–73.38	This study	56	This study	35
Arequipa	Arequipa	–13.13	–71.11	Fuselli et al. 2003	22	Tarazona-Santos et al. 2001	15
Cajamarca	Cajamarca	–7.16	–78.51	Sandoval et al. 2013	19	Sandoval et al. 2013	19
Caleta Santa Rosa	Lambayeque	–7.13	–79.55	This study	31	—	—
Capachica	Puno	–15.67	–69.85	Sandoval et al. 2013	15	Sandoval et al. 2013	15
Cara Cara	Cusco	–15.48	–70.35	This study	26	This study	28
Catacaos	Piura	–5.27	–80.68	This study	16	—	—
Chimú	Puno	–15.86	–69.95	Sandoval et al. 2013	16	Sandoval et al. 2013	16
Cusco	Cusco	–13.52	–71.97	Sandoval et al. 2013	36	Sandoval et al. 2013	36
Cusco North	Cusco	–13.52	–71.98	This study	34	This study	54
Cusco South	Cusco	–14.6	–71.25	This study	95	—	—
Huancapi	Ayacucho	–13.66	–74.05	This study	13	—	—
Huancavelica	Huancavelica	–12.93	–75.15	Sandoval et al. 2013	26	Sandoval et al. 2013	26
Islilla	Lambayeque	–5.12	–81.11	This study	18	—	—
Los Uros	Puno	–15.74	–69.93	Sandoval et al. 2013	25	Sandoval et al. 2013	25
Matsiguenga	Cusco	–12	–72.30	Mazières et al. 2008	38	Mazières et al. 2008	13
Otuzco	La Libertad	–7.91	–78.56	This study	26	—	—
Oxapampa	Pasco	–10.59	–75.40	Sandoval et al. 2013	18	Sandoval et al. 2013	18
Picota	San Martín	–8.2	–75.98	This study	21	This study	21
Picota Centro	San Martín	–6.92	–76.33	This study	22	This study	22
Puno	Cusco	–15.84	–70.02	This study	69	—	—
San Juan del Oro	Cusco	–14.12	–69.12	This study	60	This study	56
San Martín de Pangoa	Junín	–11.43	–74.48	Fuselli et al. 2003	17	—	—
Santa Rosa de Yanaque	Puno	–15.95	–69.67	Sandoval et al. 2013	18	Sandoval et al. 2013	18
Santiago de Chuco	La Libertad	–8.15	–78.18	This study	21	This study	13
Taquile	Puno	–15.77	–69.68	Sandoval et al. 2013	35	Sandoval et al. 2013	35
Tayacaja	Tayacaja	–12.24	–74.34	Fuselli et al. 2003	63	Tarazona-Santos et al. 2001	44
Trujillo	La Libertad	–8.11	–79.03	This study	13	—	—
Tupe	Lima	–12.74	–75.81	This study	17	—	—

over 18 years of age with all four grandparents from the designated sample area.

To conduct our tests, we sequenced the mtDNA first hypervariable region (HVI) and determined the mtDNA haplogroup of 611 individuals, and we typed 10 Y-chromosome short tandem repeats (Y-STRs) and determined the Y-chromosome haplogroup of 272 males. These data were combined with published data for an additional 374 individuals distributed throughout 15 populations in the Peruvian Central Andes (Table 1, Figure 1). A comparative sample of 366 individuals from 10 populations from the Amazon and Orinoco River Basin regions (henceforth also designated as "eastern") was compiled from published sources (Table 2, Figure 1).

Laboratory Methods

Each laboratory conducted DNA extractions on a subset of the total sample using a standard phenol/chloroform method (Green and Sambrook 2012). mtDNA haplogroup assignments were conducted at Universidad Nacional Mayor de San Marcos and the Universidad Ricardo Palma on nonoverlapping population samples. mtDNA sequencing of all population samples took place at UNM/ASU, and Y haplogroup assignments and STR fragment analysis took place at ASU.

Table 2. Amazon and Orinoco River Basin Populations: Sample Information and Data Sources

Population Sample	mtDNA HVI		Y-Chromosome Q1a2a1a1-M3	
	Reference	*n*	Reference	*n*
Gavião	Ward et al. 1996	28	Tarazona-Santos et al. 2001	17
Apalaí	Mazières et al. 2008	102	Mazières et al. 2011	28
Guahibo	Vona et al. 2005	59	—	—
Karitiana	Zheng et al. 2012	24	Tarazona-Santos et al. 2001	8
Suruí	Fuselli et al. 2003	22	Tarazona-Santos et al. 2001	5
Ticuna	Yang et al. 2010	12	Tarazona-Santos et al. 2001	32
Wai Wai	Bonatto and Salzano 1997	25	Tarazona-Santos et al. 2001	5
Xavánte	Bonatto and Salzano 1997	25	Tarazona-Santos et al. 2001	5
Yanomamö	Hunley et al. 2008	40	—	—
Zoró	Ward et al. 1996	29	Tarazona-Santos et al. 2001	4

mtDNA

HVI was amplified using one of four PCR primer sets: L15996 and H16401 (Vigilant et al. 1989), L15926 (Kocher et al. 1989) and H16498 (DiRienzo and Wilson 1991), L16055 and H16410 (Handt et al. 1996), or L15936 (5'-CTTGTAAACCGGAGATGAAA-3') and H16498 (DiRienzo and Wilson 1991). PCR products were purified using shrimp alkaline phosphatase (Affymetrix/USB) and Exonuclease I (Affymetrix/USB) and sequenced in forward and reverse directions using the BigDye protocol (Applied Biosystems) on an Applied Biosystems 3730 capillary sequencer.We typed four major Native American founder haplogroups (A–D) through the identification of diagnostic HVI polymorphisms. Haplogroup assignments were verified through PCR amplification of characteristic mtDNA markers followed by restriction enzyme typing (*Hae*III, *Hinc*II, or *Alu*I), or by the presence/absence of the COII/tRNA^{Lys} nine-base-pair deletion, as described previously (Stone and Stoneking 1998). When mtDNA sequences were not diagnostic of Native American haplogroups A–D, we sequenced additional sections of the mtDNA coding region using primers as described in Ramos et al. (2009, table 1, primer sets 2, 8, and 9). Haplogroups for these samples were defined using PhylotreeMT (mtDNA tree build 16; van Oven and Kayser 2009).

Y-Chromosome

Samples were typed for 10 Y-STRs: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, and DYS439 using the Powerplex Y kit (Promega), and DYS388 and DYS426 typed using primers and PCR conditions from Kayser et al. (1997) and Jobling et al. (1997), respectively. STR fragment analysis was conducted either on an ABI 377 gel sequencer or an Applied Biosystems 3730 capillary sequencer and processed using Applied Biosystem's GeneScan and Genotyper software.

All samples were initially screened for M242, M3, and M19 SNPs that are diagnostic of Y haplogroups Q*, Q1a2a1a1, and Q1a2a1a1a, respectively. Any sample within the Q* paragroup was subsequently typed for the M346 SNP that is diagnostic of haplogroup Q1a2. If a sample did not fall into Y-chromosomal haplogroup Q, haplogroups were predicted via a Bayesian approach (Athey 2006) implemented in the Haplogroup Predictor web interface (www.hprg.com/hapest5). Haplogroup predictions were verified through direct sequencing of diagnostic SNPs. All SNP primer sets were derived from Karafet et al. (2008), except M242, M3, M415, and M420. M242 was PCR-amplified using the primer set 5'-CCTTGCTGTCTAGTTCACTAG-3' and 5'-AATACCTTACCTAGAACAACTC-3' and typed using the *BsiHKA*I restriction enzyme (New England BioLabs). M3 was amplified using mismatch primers from Santos et al. (1999) and typed using the *Mfe*I restriction enzyme (New England BioLabs). Primer sets for M415 and M420 are from Myres et al. (2011) and Hinds et al. (2005), respectively.

Analytical Methods

Analyses of mtDNA and Y-chromosome variation of traditional populations in the Peruvian Central Andes were limited to Native American founder lineages. Y-chromosome lineages use a notation of lineage-based name followed by the diagnostic SNP (e.g., "Q1a2a1a1-M3"). Y haplogroup designations are based on the International Society of Genetic Genealogy 2014 Y-DNA Haplogroup Tree (version 9.7).

mtDNA sequences were aligned to the revised Cambridge Reference Sequence (Andrews et al. 1999) and edited using SeqMan Pro (DNASTAR Lasergene 11). We excluded nucleotide positions (np) 16182 and 16183 because these positions are dependent on the presence of a C mutation at np 16189 (Pfeiffer et al. 1999). Analyses also excluded any cytosine inserts between np 16193– and np 16194, as these may be present due to individual sequence length heteroplasmy (see Irwin et al.

2009) or to differences in sequencing methods among labs. To facilitate comparisons with other published data sets, we cropped all mtDNA HVI sequences from np 16084 to 16364 and limited our Y-chromosome analyses to six STRs (DYS19, DYS389I, DYS389B, DYS390, DYS391, DYS393) within Q1a2a1a1-M3. DYS389B was calculated by subtracting DYS389I from DYS389II. Additionally, Y-chromosomes from three eastern population samples (Gavião, Suruí, Zoró) were combined into a single sample to replicate the analysis in Tarazona-Santos et al. (2001); however, we found that our results were robust compared with alternative groupings, such Gavião only, Gavião with either Suruí or Zoró, or all three separately.

To test the hypothesis that populations throughout the Peruvian Central Andes contain patterns of high within- and low between-population genetic diversity, we assessed distributions of population diversity by estimating average gene diversity within mtDNA HVI and Y-STR haplotypes (π and D estimates, respectively; Nei 1987). We further assessed levels of population genetic differentiation within ecozones through the estimation of Φ_{ST} between pairs of populations (Excoffier et al. 1992). Because previous claims of "high" within- and "low" between-population variation in the Peruvian Central Andes are based on comparisons with the same or similar measures in eastern South American populations, we conducted analyses on available population samples from the Amazon and Orinoco River Basins (see Table 2). Furthermore, because previous analyses focused exclusively on highland populations, we also conducted these tests using only Peruvian highland samples versus eastern samples. These estimates were generated using Arlequin, version 3.5.1.3 (Excoffier and Lischer 2010). Nonparametric randomization tests, in which the observed difference in average π or D estimates of populations within regions ("west" or "east") was compared with a distribution of estimates obtained from 5,000 random permutations of the original data with replacement, were executed in Microsoft Excel, version 14.4.3.

We used two approaches to ascertain the extent to which events from the last millennium could account for extant genetic structure in the Peruvian Central Andes: (a) we explored relative patterns of the apportionment of variation among populations in different ecozones using analyses of molecular variance (AMOVA; Excoffier et al. 1992), and (b) we investigated potential effects of sex-biased gene flow on patterns of genetic variation, in two ways: genetic and geographic distances, and differences in genetic structure of males and females. We assessed correlations between genetic and geographic distances via Mantel tests (Mantel 1967; Smouse et al. 1986) using PASSaGE 2 (Rosenberg and Anderson 2011). Mantel tests were statistically evaluated using Bonferroni-corrected probabilities. For mtDNA, Φ_{ST} genetic distances were estimated using the Tamura-Nei evolutionary model (Tamura and Nei 1993) with the gamma distribution parameter alpha (α) set at 0.26 (Meyer et al. 1999). For Y-STRs, distances were estimated using a stepwise mutation evolutionary model (R_{ST}; Slatkin 1995). Genetic distance analyses were performed using Arlequin, version 3.5.1.3 (Excoffier and Lischer 2010). Great circle geographic distances were calculated from the longitude and latitude coordinates for each sample site using PASSaGE 2. We assessed differences in the genetic structure of males and females using the parameter Nv, calculated as $(1/\Phi_{ST}) - 1$ (Cavalli-Sforza and Bodmer 1971). Different Nv values between the two genetic systems reflect differences in migration rates and/or effective population sizes, as mutation rate may be considered negligible (see Destro-Bisol et al. 2004; Bolnick et al. 2006).

Last, to visualize patterns captured by the above analyses, we used principal coordinate (PCO) plots of Φ_{ST} and R_{ST} genetic distances, as above. PCO analyses were carried out using QIIME Pipeline python scripts (Caporaso et al. 2010) and visualized using KiNG, version 2.21 (Chen et al. 2009).

Results

mtDNA and Y-Chromosome Haplogroup/ Haplotype Distributions

mtDNA

The vast majority (99.3%) of individuals in the study sample belonged to Native American founder haplogroups A–D (Table 3). Three population samples (Huancapi, Picota, and Picota Centro) exhibited a 100% frequency of haplogroup B, and several others contained high frequencies of haplogroup B, such as Cara Cara (96%) and Cusco

Table 3. mtDNA and Y-Chromosome Native American Founder Haplogroup Counts (Frequencies) in Peruvian Central Andean Population Samples

Population Sample	mtDNA Haplogroup					Y-Chromosome Haplogroup			
	A	B	C	D	*n*	Q*-M242[a]	Q1a2-M346[a]	Q1a2a1a1-M3	*n*
Ancash	5 (.069)	35 (.486)	18 (.25)	15 (.205)	73	2 (.08)	2 (.08)	23 (.92)	25
Andahuaylas	9 (.164)	28 (.509)	13 (.236)	5 (.091)	55	2 (.09)	2 (.09)	20 (.9)	22
Caleta Santa Rosa	3 (.08)	12 (.414)	7 (.241)	7 (.241)	29				
Cara Cara		25 (.962)		1 (.038)	26			28 (1)	28
Catacaos	8 (.533)	1 (.067)	5 (.333)	1 (.067)	15				
Cusco North	3 (.088)	22 (.647)	6 (.176)	3 (.088)	34				
Cusco South	2 (.021)	74 (.779)	12 (.126)	7 (.074)	95	1 (.021)	1 (.021)	42 (.875)	48
Huancapi		13 (1)			13				
Islilla	3 (.167)	12 (.067)	1 (.056)	2 (.111)	18				
Otuzco	4 (.154)	5 (.192)	8 (.308)	9 (.346)	26				
Picota		21 (1)			21			21 (1)	21
Picota Centro		22 (1)			22			22 (1)	22
Puno	2 (.029)	46 (.667)	15 (.217)	6 (.087)	69				
San Juan del Oro	14 (.233)	34 (.567)	6 (.1)	6 (.1)	60			55 (1)	55
Santiago de Chuco	2 (.095)	12 (.571)	3 (.143)	4 (.190)	21			10 (1)	10
Trujillo		3 (.231)	3 (.231)	7 (.538)	13				
Tupe		11 (.647)	6 (.353)		17				

[a] Individuals with the Q1a2-M346 SNP also share the Q*-M242 SNP, as per Bailliet et al. (2009).

North (78%). The rest of the population samples contained varying proportions of lineages A–D. Four individuals exhibited nonfounder mtDNA lineages, specifically, African mtDNA haplogroup L haplotypes (Salas et al. 2004; Behar et al. 2008; van Oven and Kayser 2009; Soares et al. 2012), and likely represent recent (post-AD 1492) admixture. These four individuals were distributed in three population samples (Andahuaylas, Catacaos, and Caleta Santa Rosa; see Supplemental Appendix S1).

We additionally identified a geographically uneven distribution of a single mtDNA haplotype (16189C, 16217C, 16242T, and 16324C) among our sampled populations (and in one published sample, the Tayacaja; see Table 4). Several of the lowland samples demonstrated either 100% (Picota and Picota Centro) or a very high percentage (e.g., San Juan del Oro, 43%) of this haplotype; one of our highland samples (Huancapi) also demonstrated 100% of this haplotype, and another (Cara Cara) was at 96%.

Table 4. Distribution of mtDNA B Haplotype (16189C, 16217C, 16242T, 16324C) in the Peruvian Central Andes

Population Sample	Ecological Zone	*n*	Haplotype Count	Haplotype Frequency
Ancash	Highlands	68	1	0.015
Cara Cara	Highlands	26	25	0.962
Cusco South	Highlands	95	14	0.147
Huancapi	Highlands	13	13	1.000
Otuzco	Highlands	26	1	0.038
Picota	Lowlands	21	21	1.000
Picota Centro	Lowlands	22	22	1.000
San Juan del Oro	Lowlands	60	26	0.433
Santiago de Chuco	Highlands	21	5	0.238
Tayacaja	Highlands	63	1	0.016

Y-Chromosome

Sixteen biallelic polymorphisms on the Y-chromosome were characterized in nine population samples, resulting in 12 Y-chromosome haplogroup lineages (see Supplemental Appendix S2). Approximately 92% of males characterized exhibited one of two Native American founder haplogroups: Q1a2-M346, comprising 1.4% of the total sample, and Q1a2a1a1-M3, comprising 90% of the total sample. Other, presumably nonfounder, Y haplogroups were R1b1-M415 (9.2%), E-M96 (1.1%),

Table 5. mtDNA (Haplogroups A–D) HVI Summary Statistics

Population Sample	*n*	No. Haplotypes	Nucleotide Diversity (π)		Φ_{ST}
			Within Population	Average within Population within Ecogeographic Region	
Peruvian Central Andes				*0.0186*	*0.1883*
Highlands	*730*			*0.0192*	*0.1522*
Amantaní	26	8	0.0098		
Ancash	73	37	0.0247		
Andahuaylas	55	35	0.0258		
Arequipa	22	17	0.0205		
Cajamarca	19	14	0.0249		
Capachica	15	11	0.0219		
Cara Cara	26	2	0.0021		
Chimú	16	8	0.0183		
Cusco	36	24	0.0250		
Cusco North	34	26	0.0260		
Cusco South	95	35	0.0197		
Huancapi	13	1	0.0000		
Huancavelica	26	16	0.0264		
Los Uros	25	5	0.0147		
Otuzco	26	19	0.0281		
Puno	69	41	0.0212		
Santa Rosa de Yanaque	18	10	0.0131		
Santiago de Chuco	21	10	0.0248		
Taquile	35	4	0.0036		
Tayacaja	63	39	0.0266		
Tupe	17	9	0.0250		
Lowlands	*176*			*0.0134*	*0.3189*
Matsiguenga	38	9	0.0092		
Oxapampa	18	7	0.0227		
Picota	21	1	0.0000		
Picota Centro	22	1	0.0000		
San Juan del Oro	60	13	0.0259		
San Martín de Pangoa	17	10	0.0226		
Coast	*75*			*0.0236*	*0.1170*
Caleta Santa Rosa	29	20	0.0259		
Catacaos	15	7	0.0213		
Islilla	18	10	0.0249		
Trujillo	13	11	0.0223		
Amazon and Orinoco River Basins				*0.0320*	*0.2115*
Apalaí	102	16	0.0253		
Gavião	28	7	0.1813		
Guahibo	59	12	0.0179		
Karitiana	24	4	0.0055		
Suruí	22	3	0.0037		
Ticuna	12	4	0.0149		
Wai Wai	25	8	0.0221		
Xaványte	25	4	0.0108		
Yanomamö	40	17	0.0204		
Zoró	29	8	0.0179		

J-M304 (2.6%), G-M201 (1.8%), R1b-M343 (0.7%), I-M170 (0.7%), and R1a1-SRY10831.2 (0.4%). Many of these nonfounder haplogroups were distributed within the Ancash sample, followed by the Andahuaylas sample, and then were sparingly present in the rest of the samples, except for Picota and Picota Centro, which contained none (see Supplemental Appendix S2). We additionally screened R1b-M415 samples for R1b1a1-M73 (Malyarchuk et al. 2011) to investigate the possibility that these R1b lineages, typically attributed to European male admixture, might instead be of Asian origin (Bortolini et al. 2003; Kemp and Schurr 2010; Lell et al. 2002), but no sample had the diagnostic M73 SNP.

The Q1a2a1a1-M3 "modal haplotype," 13.10.17.24.10.13, identified in previous Y-chromosome studies among Native South Americans (Bianchi et al. 1998; Tarazona-Santos et al. 2001) was absent in our sample. The most frequent (22%) Q1a2a1a1-M3 haplotype, 13.14.18.23.10.14, occurred mostly in lowland samples (Picota, Picota Centro, and San Juan del Oro) and, furthermore, was the only haplotype found in the Picota and Picota Centro samples (excluding individuals with missing data).

Patterns of Within- and Between-Population Genetic Variation

Prior characterizations of western Andean population genetic variation assume a strong distinction from eastern South American populations. In our reanalysis, the contrast between "west" and "east" is not as stark as reported previously.

In the case of mtDNA, the relative comparison between west and east is inconsistent with previous reports (i.e., Fuselli et al. 2003; Cabana et al. 2006). The western Andes showed lower, not higher, estimates of within-population diversity. In our study, average within-population π estimates for the entire Peruvian Central Andes and the Peruvian highlands are 0.0186 and 0.0192, respectively, relative to 0.0320 in the eastern region (Table 5). Comparisons of the Peruvian highland and coast with eastern populations are statistically significant at $P \leq 0.05$, but the Peruvian lowland–eastern South America comparison is not ($P = 0.37$). In terms of population differentiation, the Peruvian Central Andes as a whole ($\Phi_{ST} = 0.1833$) and its highland and coastal areas ($\Phi_{ST} = 0.1522$ and 0.1170, respectively) have lower estimates than do the Amazon/Orinoco River Basins ($\Phi_{ST} = 0.2155$). In other words, throughout the Peruvian Central Andes, and especially within the highlands, approximately 18% of the variation is apportioned among populations, versus 21% in eastern South America; these values are not substantially different from one another. The Peruvian lowland ecozone is exceptional, exhibiting relatively high population differentiation estimates ($\Phi_{ST} = 0.3189$).

A handful of highland populations exhibited low mtDNA diversity levels. For example, Huancapi and Cara Cara demonstrated extraordinarily low diversity estimates ($\pi = 0$ and 0.0021, respectively) due to their high frequencies of the mtDNA haplogroup B haplotype (16189C, 16217C, 16242T, and 16324C) described above. To date, no other reported highland Andean population sample has exhibited such low mtDNA diversity.

In contrast to the mtDNA results, Peruvian Central Andean Q1a2a1a1-M3 Y-chromosomes exhibited higher, but statistically nonsignificant, average gene diversity estimates relative to the eastern region ($D = 0.3601$ vs. 0.3007, respectively; $P = 0.16$; Table 6); the Peruvian highland area showed a much higher and significant relative value ($D = 0.3936$; $P = 0.03$). Population differentiation was lower in the Peruvian Central Andes overall ($\Phi_{ST} = 0.4147$), as well as in the highlands ($\Phi_{ST} = 0.4042$) and lowlands ($\Phi_{ST} = 0.2500$), compared with the eastern estimate ($\Phi_{ST} = 0.6265$). This indicates that the eastern region contains about 20% more variation apportioned among populations relative to the Peruvian Central Andes. Interestingly, however, highland population Y-chromosomes were much more differentiated than found previously (Tarazona-Santos et al. 2001; Yang et al. 2010). For example, Tarazona-Santos et al. (2001) report a pooled Central Andean Φ_{ST} estimate of 0.024, but our comparative analyses demonstrated a much higher estimate for highland populations, as well as those within the entire Peruvian Central Andes.

Population Genetic Structure by Ecozone

To address the second hypothesis, that variation within the Peruvian Central Andes can be explained by events from the last half millennium, we first sought to assess the degree to which highland populations were differentially affected by Inca and Spanish policies relative to lowland and coastal populations, if at all. The hypothesis is

Table 6. Y-Chromosome Q1a2a1a1-M3 Summary Statistics

Population Sample	*n*	No. Haplotypes	Average Gene Diversity among Loci (*D*)		Φ_{ST}
			Within Population	Average within Population within Ecogeographic Region	
Peruvian Central Andes				*0.3601*	*0.4147*
Highlands	*377*			*0.3936*	*0.4042*
Amantaní	26	12	0.4892		
Ancash	25	12	0.3603		
Andahuaylas	22	15	0.4658		
Arequipa	15	10	0.4778		
Cajamarca	19	14	0.5136		
Capachica	15	7	0.5063		
Cara Cara	28	11	0.0688		
Chimú	16	11	0.5319		
Cusco	36	18	0.4598		
Cusco South	43	25	0.4058		
Huancavelica	26	17	0.4010		
Los Uros	25	7	0.3978		
Santiago de Chuco	10	7	0.4474		
Santa Rosa de Yanaque	18	12	0.4553		
Taquile	35	8	0.2927		
Tayacaja	44	32	0.5021		
Lowlands	*129*			*0.2291*	*0.2500*
Matsiguenga	13	6	0.3077		
Oxapampa	18	8	0.4455		
Picota	21	3	0.0151		
Picota Centro	22	6	0.0000		
San Juan del Oro	55	13	0.3771		
Amazon and Orinoco River Basins	*50*			*0.3007*	*0.6265*
Apalaí	28	8	0.4281		
Gavião-Suruí-Zoró	26	11	0.3533		
Karitiana	8	2	0.0417		
Ticuna	32	9	0.3145		
Wai Wai	5	3	0.3667		
Xaváante	5	3	0.3000		

[a] Includes haplotypes with missing data.

supported for mtDNA but not for Y-chromosome data. AMOVA analyses for both genetic systems show minimal, often nonsignificant, differences in the apportionment of variation among ecozones. When within- and between-population apportionments of variation were analyzed independently for each ecozone, highland groups had higher amounts of within-group variation for mtDNA but not for Y-chromosome data (Table 7).

More specifically, mtDNA-based AMOVA analyses of multiple comparisons among ecological zones showed most (~67–82%) genetic variation distributed within populations; the little genetic variation that was apportioned among zones rarely reached statistical significance after 10,000 permutations (Table 7). This general lack of structure by ecological zone was apparent in the mtDNA PCO plot depicting the first two PCOs (Figure 2A) in which the first and second PCOs captured 46% and 24% of the total variance in the data, respectively. Highland, coastal, and lowland samples were intermixed. Four samples that form a somewhat

Table 7. Analyses of Molecular Variance (AMOVA)

Locus/Region	Percentage of Variation		
	Among Groups	Among Populations within Groups	Within Populations
mtDNA HVI			
Peruvian Central Andes		*18.33*	*81.17*
Highlands		15.22	84.78
Lowlands		31.89	68.11
Coast		11.70	88.3
Ecozone comparisons			
Coast-highlands-lowlands	3.62	16.89	79.49
Coast-highlands	4.32[a]	14.27	81.41
Highlands-lowlands	2.59[a]	17.53	79.88
Coast-lowlands	9.67[a]	22.98	67.35
Amazon and Orinoco River Basins		21.15	78.85
Y-chromosome Q1a2a1a1-M3			
Peruvian Central Andes		*41.47*	*58.53*
Highlands		40.42	59.58
Lowlands		25.00	75.00
Ecozone comparison			
Highlands-lowlands	8.29[a]	36.19	55.52
Amazon and Orinoco River Basins		37.35	62.65

[a] $P < 0.05$ after 10,000 permutations.

distant cluster along the first PCO axis—Cara Cara, Huancapi, Picota, and Picota Centro—all share very high frequencies of the mtDNA haplogroup B haplotype (16189C, 16217C, 16242T, and 16324C) mentioned above. A comparison of the apportionment of variation within and between populations by ecozone shows that a large proportion of genetic variation in coastal and highland zones is found within (~85–88%) rather than among (12–15%) populations, whereas in the lowlands, only 68% was distributed within, and up to 31% among populations.

Y-chromosome-based AMOVA analyses showed genetic variation also as weakly and insignificantly apportioned between highland and lowland ecozones, with a little more than half (~55%) of genetic variation distributed within populations and the rest among populations within ecozones. The Y-chromosome PCO plot visually depicts the general lack of structure between highland and lowland samples (Figure 2B). The first and second PCO captured 47% and 24% of the total variance in the data, with highland and lowland samples for the most part intermixed along both axes, with the exception of two clusters of samples. One cluster, consisting of Cara Cara, Picota, and Picota Centro, shares high frequencies of a single Y haplotype (13.14.18.23.10.14) within Q1a2a1a1-M3. A second cluster, consisting of Tayacaja and Arequipa, shares many haplotypes in common to the exclusion of most or all other Peruvian Central Andean populations. Finally, as pointed out above, much more variation is apportioned among, versus within, groups in the highlands relative to those in the lowlands.

Sex-Specific Patterns of Population Genetic Structure

A secondary hypothesis related to the potential effect of Inca and Spanish policies was an expected lack of sex bias in the extant pattern of population genetic variation throughout the Peruvian Central Andes, and in the highlands especially. This secondary hypothesis is rejected: though Mantel tests lacked evidence of distance-based patterning of either maternal or paternal genetic variation, $N\nu$ values between the two genetic systems reveal higher female effective sizes and/or gene flow in the Peruvian highlands; the converse is true for the Peruvian lowlands.

Mantel tests yielded low and statistically insignificant correlations between genetic and geographic distances for both genetic systems. For mtDNA, the Mantel statistic was low ($r = 0.1555$; critical P-value = 0.001). For the Y-chromosome, the Mantel statistic was slightly negative ($r = -0.5142$; critical P-value = 0.003). Both PCO plots depict this lack of correspondence: although pairs of spatially proximate population samples cluster together somewhat along either the first or second PCO axis, relationships among population samples generally do not readily relate to geographic location (Figure 2).

Maternally and paternally inherited loci showed opposite patterns of genetic differentiation (Φ_{ST}) among populations (Tables 5, 6). mtDNA indicates high differentiation in the lowlands ($\Phi_{ST} = 0.32$) and approximately equal levels in the highlands ($\Phi_{ST} = 0.15$) and the coast ($\Phi_{ST} = 0.12$), whereas Y-chromosomes indicate low differentiation in the lowlands ($\Phi_{ST} = 0.25$) and high differentiation in the highlands ($\Phi_{ST} = 0.40$). The Φ_{ST} values for the Peruvian Central Andes produce a ratio of mtDNA to Y-chromosome $N\nu$ of 3.05, indicating a female migration rate and/or effective size that is

more than three times that of males. In the Peruvian highlands, the ratio is even larger (Nv = 3.78), whereas in the lowlands it is lower (Nv = 0.71), suggesting the male migration rate and/or effective size is almost 1.5 times higher than that of females.

Discussion

This study presents a comprehensive investigation of population genetic variation in the Peruvian Central Andes with the purpose of contributing to the reconstruction of the region's population history. To this end, the study analyzes maternally inherited mtDNA HVI and paternally inherited Y-chromosome STRs of 17 novel population samples combined with 15 published population samples. With this large number of samples covering multiple ecological areas throughout the Peruvian Central Andes, we revisit the long-standing characterization of Andean population genetic variation as relatively unstructured compared with eastern South American populations, and we assess the extent to which state expansion and colonialism during the last half millennium might have differentially affected highland communities and/or generated a lack of sex bias in population genetic structure.

As noted in the introductory remarks, an overall pattern of high within- and low between-population variation in the Andes relative to eastern South America has been a consistent finding across multiple genetic-based studies. Our results reject a consistent pattern of relatively high within- and low between-population variation for both genetic systems. For the Peruvian Central Andes, mtDNA HVI within-population diversities were low, rather than high, relative to values for eastern populations, and only highland Y-chromosomes exhibited significantly higher within-population diversities compared with eastern Y-chromosomes. Also in contrast to previous studies, we found several population samples within the Peruvian highlands, as well as throughout the Peruvian Central Andes, with relatively low mtDNA genetic diversity estimates. These results are partly due to the fact that the characterization of eastern population structure differs from previous reports because of our inclusion of novel samples. Eastern population diversities for both genetic systems are higher than,

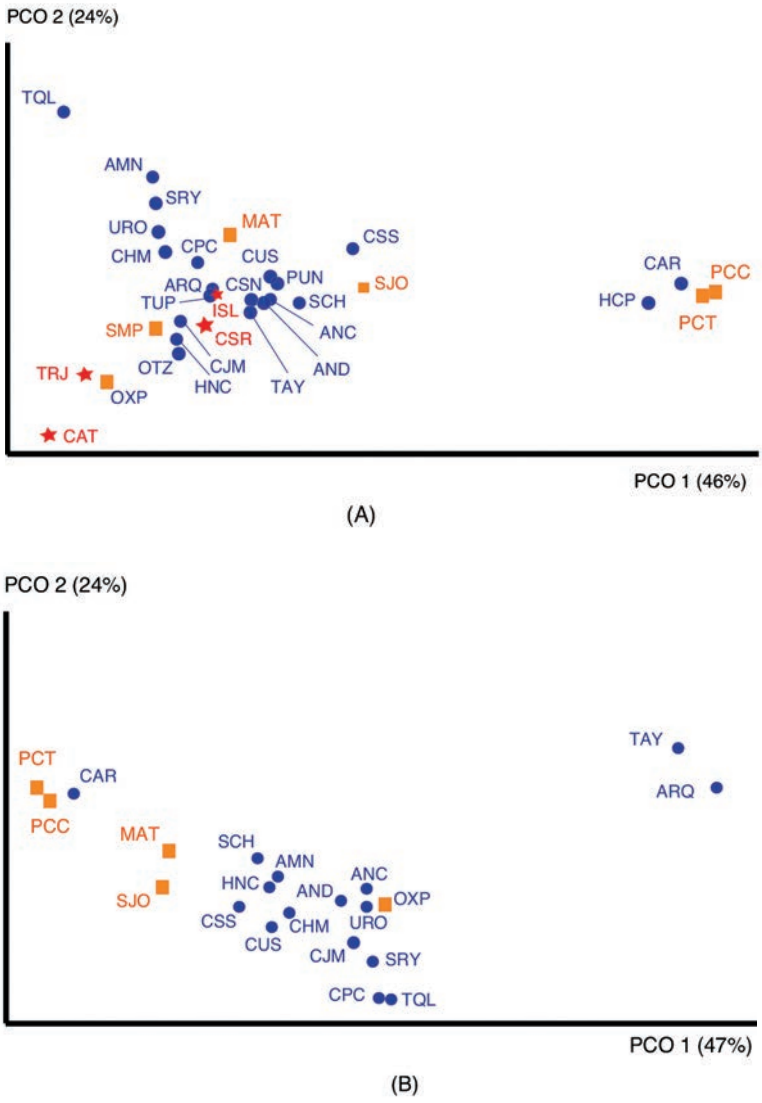


FIGURE 2. First two principal coordinates based on mtDNA HVI pairwise Φ_{ST} (A) and Y haplogroup Q1a2a1a1-M3 pairwise R_{ST} (B) distance estimates. Axes are labeled with the percentage of variance explained. Coastal population samples are designated with stars, highland samples with circles, and lowland samples with squares. Abbreviations: AMT, Amantaní; ANC, Ancash; AND, Andahuaylas; ARQ, Arequipa; CAR, Cara Cara; CAT, Catacaos; CHM, Chimú; CJM, Cajamarca; CPC, Capachica; CSN, Cusco North; CSR, Caleta Santa Rosa; CSS, Cusco South; CUS, Cusco; HCP, Huancapi; HNC, Huancavelica; ISL, Islilla; MAT, Matsiguenga; OTZ, Otuzco; OXP, Oxapampa; PCC, Picota Centro; PCT, Picota; PUN, Puno; SCH, Santiago de Chuco; SJO, San Juan del Oro; SMP, San Martín de Pangoa; TAY, Tayacaja; TQL, Taquile; TRJ, Trujillo; TUP, Tupe; URO, Los Uros.

and eastern Y-chromosomes are twice as differentiated as, previously reported (Tarazona-Santos et al. 2001; Fuselli et al. 2003; Cabana et al. 2006).

Up until this point, the characterization of a relative lack of population genetic structure in the Andean region versus eastern South America has gone unchallenged. Instead, recent work has built

on the presumed weak structure to argue against a dual founder model for the peopling of South America (Lewis and Long 2008; Lewis 2009) and to argue for the possibility of range expansion by a single founding population into the Central Andes (Yang et al. 2010; Batai and Williams 2014). However, by simply increasing the number and geographical coverage of populations sampled, we show that populations on either side of the Andean range do not have contrasting—or even opposing—patterns of genetic variation and thus cannot be easily characterized as two distinct "meta-" (Lanata and García 2005) or "mega-" (Pucciarelli et al. 2006) populations. This is the case even if we confine our definition of "western Andes" to highland regions only.

Explanations for the presumed dichotomous pattern have focused on migratory events in early South American prehistory, presumably because large-scale migrations of distinct founder populations could reasonably explain why such clear and distinctly opposing patterns of population structure were observed over such large, well-defined geographic areas several millennia later. The fact that this explanation has dominated discussions, however, means that the role of other or additional (pre)historical phenomena in shaping extant patterns of population genetic variation in the Andes has been underexplored. In order to forward that exploration, we chose to focus on the Inca imperial and Spanish colonial policies from the last half millennium because (1) they are recent phenomena relative to the region's ~15,000 years of prehistory (Dillehay 2009), and (2) they generate clear expectations of the structure of population genetic diversity.

Thus, our second hypothesis investigates population structure solely within the Peruvian Central Andean region, by ecozone and/or by genetic system. Archaeological and ethnohistorical accounts led us to expect weak signals of population genetic structure and a lack of sex bias in gene flow and/or effective size estimates, particularly in the Peruvian highlands relative to the coast and lowlands. Our results show that Peruvian highland communities do exhibit weak population genetic structure relative to other ecozones only for mtDNA data, and not for Y-chromosome data. Second, highland populations exhibit higher levels of female- versus male-mediated gene flow and/or high effective sizes relative to other regions, whereas the lowlands show the opposite pattern. Though our results provide strong evidence for high levels of gene flow and/or effective sizes among females in Peruvian highland groups, they do not fully support our expectations for both genetic systems. These results therefore provide, at best, equivocal evidence for a strong influence of combined Inca and Spanish policies on population genetic variation throughout the Andes or in highland regions only.

Population genetic structure in the Andes does not appear to be dictated by large-scale phenomena and may instead be understood via an exploration of more spatiotemporally defined processes. Archaeological, ethnohistorical, and recent paleogenetic studies of the Central Andes encourage a finer-grained analysis. These studies show early geographic and diachronic variation in the area: from the time of initial colonization, hunter-gatherers engaged in diverse subsistence practices with varying patterns of mobility, including sedentism, transhumance, and long-distance foraging—practices that probably isolated some human populations regionally (e.g., Aldenderfer 2008; Sandweiss 2008). As food production became central to the subsistence economies and social organization of most Andean societies (particularly after 3000 BC), groups increasingly engaged in contacts that spread crops, craft goods, and cultural practices within and among regions (Perry et al. 2006; Haas et al. 2013). Andean urban centers emerged around AD 400–600, signaling the presence of centralized and hierarchical societies with subsistence strategies favoring rapid population growth rates (Read and LeBlanc 2003). State expansion, in which central political control was exercised over large territories, occurred in multiple locations. In the case of Wari and Tiwanaku, state populations also established colonies. Wari and Tiwanaku disintegrated around AD 1000, replaced by agropastoral societies that used broad kin-based networks to access diverse resources. Populations grew markedly across the highlands until around AD 1400, when imperial expansion of the Inca state likely altered the region's demography. Spanish documents indicate that by the mid-sixteenth century, the region displayed diverse ethnic groups, languages, and dialects (Mannheim 1991).

Moreover, paleogenetic studies suggest that

local dynamics may have deviated from broader-scale processes over the last 500 years in the Andean region. Studies by Baca et al. (2014), Kemp et al. (2009), and Lewis et al. (2007a) present evidence for spatiotemporal continuity of populations, whereas Fehren-Schmitz et al. (2011) found evidence for both continuity and discontinuity of populations that differ between maternally and paternally inherited loci. Notably, our study's identification of an uneven geographic distribution of a single mtDNA B haplotype hints at localized instances of gene flow and drift.

Conclusion

Over the last two decades, studies of continent-level population genetic variation and structure have been executed under the presumption that either a single- or dual-founder colonization model would best explain observed patterns. Our study upends this research trend by rejecting the presumed dichotomous pattern of South American population genetic structure.

Importantly, our study contributes greater nuance to the "western" pattern. We have shown that with denser population and genomic-level sampling, the "west" reveals finer-scale patterns of population genetic variation than formerly recognized. These patterns may result from relatively localized and historically contingent processes, rather than from single major trends or causal variables. We suggest that future endeavors in the Peruvian Central Andes strive for denser geographic and genomic coverage using both ancient and modern DNA techniques to better access local (pre)histories at multiple spatial and temporal scales.

SUPPLEMENTAL APPENDICES

Complete population-level data for mtDNA HVI and Y-chromosome STRs are provided in Supplemental Appendixes S1 and S2, respectively. Supplemental Appendix S1 (available from http://goo.gl/j2bgf8) provides mtDNA haplogroup and full (uncropped) sequence data, along with four spreadsheets summarizing sequence variants by founder haplogroups A–D. Supplemental Appendix S2 (available from: http://goo.gl/Sa345l) provides Y-chromosome haplogroup and haplotype profiles based on 10 STRs. mtDNA HVI sequences are also deposited in GenBank (KM888878–KM889488).

ACKNOWLEDGMENTS

We are grateful to the peoples in the Peruvian Central Andes who agreed to participate in this study, and to Mariana Contreras, Leslie Castillo Pampas, and Julian Ccahuana Quispe for help with sample collection. We thank Daniel Mendoza and Gabriel Ríos of the Instituto de Ciencia y Tecnología at Universidad Ricardo Palma, Lima, Peru, for their professional support, and Keith L. Hunley, Elizabeth A. Sobel, and three anonymous referees for insightful comments. The research reported here was supported by the National Science Foundation through grants BCS-0401434 (A.C.S.), BCS-0242958 (A.C.S./C.M.L.), and SES-0309904 (G.S.C.).

Received 4 June 2014; revision accepted for publication 7 September 2014.

LITERATURE CITED

Aldenderfer, M. 2008. High elevation foraging societies. In *Handbook of South American Archaeology*, W. H. Isbell and H. Silverman, eds. New York: Springer, 131–144.

Andrews, R. M., I. Kubacka, P. F. Chinnery et al. 1999. Reanalysis and revision of the Cambridge Reference Sequence for human mitochondrial DNA. *Nat. Genet.* 23:147.

Athey, T. W. 2006. Haplogroup prediction from Y-STR values using a Bayesian-allele-frequency approach. *J. Genet. Geneal.* 2:34–39.

Baca, M., M. Molak, M. Sobczyk et al. 2014. Locals, resettlers, and pilgrims: A genetic portrait of three pre-Columbian Andean populations. *Am. J. Phys. Anthropol.* 154:402–412.

Bailliet, G., V. Ramallo, M. Muzzio et al. 2009. Brief communication: Restricted geographic distribution for Y-Q* paragroup in South America. *Am. J. Phys. Anthropol.* 140:578–582.

Batai, K., and S. R. Williams. 2014. Mitochondrial variation among the Aymara and the signatures of population expansion in the Central Andes. *Am. J. Hum. Biol.* 26:321–330.

Battilana, J., N. J. R. Fagundes, A. H. Heller et al. 2006. *Alu* insertion polymorphisms in Native Americans and related Asian populations. *Ann. Hum. Biol.* 33:142–160.

Behar, D. M., R. Villems, H. Soodyall et al. 2008. The dawn of human matrilineal diversity. *Am. J. Hum. Genet.* 82:1,130–1,140.

Best, W. R., M. Layrisse, and R. Bermejo R. 1966. Group antigens in Aymara and Quechua speaking tribes from near Puno, Peru. *Am. J. Phys. Anthropol.* 20:321–329.

Bianchi, N. O., C. I. Catanesi, G. Bailliet et al. 1998. Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. *Am. J. Hum. Genet.* 63:1,862–1,871.

Bodner, M., U. A. Perego, G. Huber et al. 2012. Rapid coastal spread of First Americans: Novel insights from South America's Southern Cone mitochondrial genomes. *Genome Res.* 22:811–820.

Bolnick D. A., D. I. Bolnick, and D. G. Smith. 2006. Asymmetric male and female genetic histories among Native Americans from eastern North America. *Mol. Biol. Evol.* 23:2,161–2,174.

Bonatto S. L., and F. M. Salzano. 1997. Diversity and age of the four major mtDNA haplogroups, and their implications for the peopling of the New World. *Am. J. Hum. Genet.* 61:1,413–1,423.

Bortolini, M. C., A. Ruiz-Linares, F. M. Salzano et al. 2003. Y-chromosome evidence for differing ancient demographic histories in the Americas. *Am. J. Hum. Genet.* 73:524–539.

Cabana, G. S., D. A. Merriwether, K. Hunley et al. 2006. Is the genetic structure of Gran Chaco populations unique? Interregional perspectives on native South American mitochondrial DNA variation. *Am. J. Phys. Anthropol.* 131:108–119.

Caporaso, J. G., J. Kuczynski, J. Stombaugh et al. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7:335–336.

Cavalli-Sforza, L. L., and W. Bodmer. 1971. *The Genetics of Human Populations*. San Francisco: Freeman.

Cavalli-Sforza, L. L., P. Menozzi, and A. Piazza. 1994. *History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.

Chen, V. B., I. W. Davis, and D. C. Richardson. 2009. KiNG (Kinemage, Next Generation): A versatile interactive molecular and scientific visualization program. *Protein Sci.* 18:2,403–2,409.

Cieza de León, P. de. 1985 [c. 1550]. *El Señorío de los Incas.* Madrid: Historia 16.

Cook, N. D. 1980. *Demographic Collapse: Indian Peru, 1520–1620*. New York: Cambridge University Press.

Covey, R. A., and C. M. Elson. 2007. Ethnicity, demography, and estate management in sixteenth-century Yucay (Cusco, Peru). *Ethnohistory* 54:303–335.

D'Altroy, T. 2015. *The Incas*. 2nd ed. Malden, MA: Wiley Blackwell.

Destro-Bisol, G., F. Donati, V. Coia et al. 2004. Variation of female and male lineages in sub-Saharan populations: The importance of sociocultural factors. *Mol. Biol. Evol.* 21:1,673–1,682.

Dillehay, T. D. 2009. Probing deeper into first American studies. *Proc. Natl. Acad. Sci. USA* 106:971–978.

DiRienzo, A., and A. C. Wilson. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 88:1,597–1,601.

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10:564–567.

Excoffier, L., P. Smouse, and J. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131:479–491.

Fehren-Schmitz, L., O. Warnberg, M. Reindel et al. 2011. Diachronic investigations of mitochondrial and Y-chromosomal genetic markers in pre-Columbian Andean highlanders from South Peru. *Ann. Hum. Genet.* 75:266–283.

Fuselli, S., E. Tarazona-Santos, I. Dupanloup et al. 2003. Mitochondrial DNA diversity in South America and the genetic history of Andean highlanders. *Mol. Biol. Evol.* 20:1,682–1,691.

Gosling, W. D., and J. J. Williams. 2013. Ecosystem service provision sets the pace for pre-Hispanic societal development in the Central Andes. *Holocene* 23:1,619–1,624.

Green, M. R., and J. Sambrook. 2012. *Molecular Cloning: A Laboratory Manual.* 4th ed. Cold Spring Harbor, NY: Cold Spring Harbor Spring Press.

Haas, J., W. Creamer, L. Huamán Mesía et al. 2013. Evidence for maize (*Zea mays*) in the Late Archaic (3000–1800 B.C.) in the Norte Chico region of Peru. *Proc. Natl. Acad. Sci. USA* 110:4,945–4,949.

Handt, O., M. Krings, R. H. Ward et al. 1996. The retrieval of ancient human DNA sequences. *Am. J. Hum. Genet.* 59:368–376.

Hinds, D. A., L. L. Stuve, G. B. Nilsen et al. 2005. Whole-genome patterns of common DNA variation in three human populations. *Science* 307:1,072–1,079.

Hunley, K. L., J. E. Spence, and D. A. Merriwether. 2008. The impact of group fissions on genetic structure in Native South America and implications for human evolution. *Am. J. Phys. Anthropol.* 135:195–205.

Irwin, J. A., J. L. Saunier, H. Niederstätter et al. 2009. Investigation of heteroplasmy in the human mitochondrial DNA control region: A synthesis of observations from more than 5000 global population samples. *J. Mol. Evol.* 68:516–527.

Jobling, M. A., A. Pandya, and C. Tyler-Smith. 1997. The Y chromosome in forensic analysis and paternity testing. *Int. J. Legal Med*. 110:118–124.

Karafet, T. M., F. L. Mendez, M. B. Meilerman et al. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res.* 18:830–838.

Kayser, M., A. Caglià, D. Corach et al. 1997. Evaluation of Y-chromosomal STRs: A multicenter study. *Int. J. Legal Med*. 110:125–133.

Kemp, B. M., and T. D. Schurr. 2010. Ancient and modern genetic variation in the Americas. In *Human Variation in the Americas: The Integration of Archaeology and Biological Anthropology*, B. M. Auerbach, ed. Carbondale, IL: Center for Archaeological Investigations, 12–50.

Kemp, B. M., T. A. Tung, and M. L. Summar. 2009. Genetic continuity after the collapse of the Wari empire: Mitochondrial DNA profiles from Wari and post-Wari populations in the ancient Andes. *Am. J. Phys. Anthropol.* 140:80–91.

Kocher, T. D., W. L. Thomas, A. Meyer et al. 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* 86:6,196–6,200.

Lanata, J. L., and A. García. 2005. Environmental corridors and early human dispersal in South America. Paper presented at the 70th annual meeting of the Society for American Archaeology, Salt Lake City, UT.

Lell, J T., R. I. Sukernik, Y. B. Starikovskaya et al. 2002. The dual origin and Siberian affinities of Native American Y chromosomes. *Am. J. Hum. Genet.* 70:192–206.

Lewis, C. M. Jr. 2009. Hierarchical modeling of genome-wide short tandem repeat (STR) markers infers Native American prehistory. *Am. J. Phys. Anthropol.* 141:281–289.

Lewis, C. M. Jr., J. E. Buikstra, and A. C. Stone. 2007a. Ancient DNA and genetic continuity in the South Central Andes. *Latin Am. Antiquity* 79:159–178.

Lewis, C. M. Jr., B. Lizárraga, R. Y. Tito et al. 2007b. Mitochondrial DNA and the peopling of South America. *Hum. Biol.* 18:145–160.

Lewis, C. M. Jr., and J. C. Long. 2008. Native South American genetic structure and prehistory inferred from hierarchical modeling of mtDNA. *Mol. Biol. Evol*. 25:478–486.

Lewis, C. M. Jr., R. Y. Tito, B. Lizárraga et al. 2004. Land, language, and loci: mtDNA in Native Americans and the genetic history of Peru. *Am. J. Phys. Anthropol.* 127:351–360.

Luiselli, D., L. Simoni, E. Tarazona-Santos et al. 2000. Genetic structure of Quechua-speakers of the Central Andes and geographic patterns of gene frequencies in South Amerindian populations. *Am. J. Phys. Anthropol.* 113:5–17.

Malyarchuk, B., M. Derenko, G. Denisova et al. 2011. Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup Q1a. *J. Hum. Genet.* 56:583–588.

Mannheim, B. 1991. *The Language of the Inka since the European Invasion*. Austin: University of Texas Press.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res*. 27:209–220.

Matienzo, J. de. 1967 [1567]. *Gobierno del Perú*. Lima: Institut Français d'Études Andines.

Matson, G. A., H. E. Sutton, J. Swanson et al. 1966. Distribution of hereditary blood groups among Indians in South America II. In Peru. *Am. J. Phys. Anthropol.* 24:325–350.

Mazières, S., S. M. Callegari-Jacques, S. Goulart Crosetti et al. 2011. French Guiana Amerindian demographic history as revealed by autosomal and Y-chromosome STRs. *Ann. Hum. Biol.* 38:76–83.

Mazières, S., E. Guitard, E. Crubézy et al. 2008. Uniparental (mtDNA, Y-chromosome) polymorphism in French Guiana and two related populations—implications for the region's colonization. *Ann. Hum. Genet.* 72:145–156.

Meyer, S., G. Weiss, and A. von Haeseler. 1999. Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. *Genetics* 152:1,103–1,110.

Modiano, G., L. Bernini, N. D. Carter et al. 1972. A survey of several red cell and serum genetic markers in a Peruvian population. *Am. J. Hum. Genet.* 24:111–123.

Myres, N. M., S. Rootsi, A. A. Lin et al. 2011. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur. J. Hum. Genet.* 19:95–101.

Neel, J. V. 1970. Lessons from a "primitive" people. *Science* 170:815–822.

Neel, J. V. 1978a. The population genetic structure of an Amerindian tribe, the Yanomama. *Annu. Rev. Genet.* 12:365–413.

Neel, J. V. 1978b. Rare variants, private polymorphisms, and locus heterogeneity in Amerindian populations. *Am. J. Hum. Genet.* 30:465–490.

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

Obregón-Tito, A. 2013. *Metagenomics and Social Inclusion in Peru*. Ph.D. diss., University of Oklahoma, Department of Anthropology.

O'Rourke, D. H., and B. K. Suarez. 1985. Patterns and correlates of genetic variation in South Amerindians. *Ann.*

Hum. Biol. 13:13–31.

Perry, L., D. H. Sandweiss, D. R. Piperno et al. 2006. Early maize agriculture and interzonal interaction in southern Peru. *Nature* 440:76–79.

Pfeiffer, H., B. Brinkmann, J. Huhne et al. 1999. Expanding the forensic German mitochondrial DNA control region database: Genetic diversity as a function of sample size and microgeography. *Int. J. Legal Med.* 112:291–298.

Pucciarelli, H. M., W. A. Neves, R. González-José et al. 2006. East-west cranial differentiation in pre-Columbian human populations of South America. *Homo* 57:133–150.

Ramos, A., C. Santos, L. Alvarez et al. 2009. Human mitochondrial DNA complete amplification and sequencing: A new validated primer set that prevents nuclear DNA sequences of mitochondrial origin co-amplification. *Electrophoresis* 30:1,587–1,593.

Read, D. W., and S. A. LeBlanc. 2003. Population growth, carrying capacity, and conflict. *Curr. Anthropol.* 44: 59–85.

Rodríguez-Delfín, L. A., V. E. Rubín de Celis, and M. A. Zago. 2001. Genetic diversity in an Andean population from Peru and regional migration patterns of Amerindians in South America: Data from Y chromosome and mitochondrial DNA. *Hum. Hered.* 51:97–106.

Roewer, L., M. Nothnagel, L. Gusmão et al. 2013. Continent-wide decoupling of Y-chromosomal genetic variation from language and geography in native South Americans. *PLoS Genet.* 9:e1003460.

Rosenberg, M. S., and C. D. Anderson. 2011. PASSaGE: Pattern Analysis, Spatial Statistics and Geographic Exegesis. Version 2. *Methods Ecol. Evol.* 2:229–232.

Rothhammer, F., and T. D. Dillehay. 2009. The Late Pleistocene colonization of South America: An interdisciplinary perspective. *Ann. Hum. Genet.* 73:540–549.

Rothhammer, F. and C. Silva. 1989. Peopling of Andean South America. *Am. J. Phys. Anthropol.* 78: 403–420.

Rothhammer, F. and C. Silva. 1990. Craniometric variation among South American prehistoric populations: Climatic, altitudinal, chronological, and geographic contributions. *Am. J. Phys. Anthropol.* 82: 9–17.

Rothhammer, F. and C. Silva. 1992. Gene geography of South America: Testing models of population displacement based on archaeological evidence. *Am. J. Phys. Anthropol.* 89:441–446.

Salas, A., M. Richards, M. V. Lareu et al. 2004. The African diaspora: Mitochondrial DNA and the Atlantic slave trade. *Am. J. Hum. Genet.* 74:454–465.

Salzano, F. M., J. V. Neel, H. Gershovitz et al. 1977. Intra and inter-tribal genetic variation within a linguistic group: The Ge-speaking Indians of Brazil. *Am. J. Phys. Anthropol.* 47:337–348.

Salzano, F. M., A. G. Steinberg, and J. V. Neel. 1973. Gm and Inv allotypes of Brazilian Cayapo Indians. *Am. J. Hum. Genet.* 25:167–177.

Salzano, F. M., J. P. Woodall, F. L. Black et al. 1974. Blood groups, serum proteins and hemoglobins of Brazilian Tiriyo Indians. *Hum. Biol.* 46:81–87.

Sandoval, J. R., D. R. Lacerda, M. S. A. Jota et al. 2013. The genetic history of indigenous populations of the Peruvian and Bolivian Altiplano: The legacy of the Uros. *PLoS One* 8:e73006.

Sandweiss, D. H. 2008. Early fishing societies in western South America. In *Handbook of South American Archaeology*, W. H. Isbell and H. Silverman H, eds. New York: Springer, 145–156.

Sandweiss, D. H., K. A. Maasch, R. L. Burger et al. 2001. Variation in Holocene El Niño frequencies: Climate records and cultural consequences in ancient Peru. *Geology* 29:603–606.

Santos, F. R., A. Pandya, C. Tyler-Smith et al. 1999. The central Siberian origin for Native American Y chromosomes. *Am. J. Hum. Genet.* 64:619–628.

Scliar, M. O., G. B. Soares-Souza, J. Chevitarese et al. 2012. The population genetics of Quechuas, the largest native South American group: Autosomal sequences, SNPs, and microsatellites evidence high level of diversity. *Am. J. Phys. Anthropol.* 147:443–451.

Simoni, L., E. Tarazona-Santos, D. Luiselli et al. 2000. Genetic differentiation of South American native populations inferred from classical markers: From explorative analyses to a working hypothesis. In *America Past, America Present: Genes and Language in the Americas and Beyond*, C. Renfrew, ed. Cambridge: McDonald Institute for Archaeological Research, 125–136.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139:457–462.

Smouse, P. E., J. C. Long, and R. R. Sokal. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst. Zool.* 35:627–632.

Soares, P., F. Alshamali, J. B. Pereira et al. 2012. The expansion of mtDNA haplogroup L3 within and out of Africa. *Mol. Biol. Evol.* 29:915–927.

Stone, A. C., and M. Stoneking. 1998. mtDNA analysis of a prehistoric Oneota population: Implications for the peopling of the New World. *Am. J. Hum. Genet.* 62:1,153–1,170.

Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol.*

Biol. Evol. 10:512–526.

Tarazona-Santos, E., D. R. Carvalho-Silva, D. Pettener et al. 2001. Genetic differentiation in South Amerindians is related to environmental and cultural diversity: Evidence from the Y chromosome. *Am. J. Hum. Genet.* 68:1,485–1,496.

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

Vigilant, L., R. Pennington, H. Harpending et al. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. *Proc. Natl. Acad. Sci. USA* 86:9,350–9,354.

Vona, G., P. Moral, A. Falchi et al. 2005. Mitochondrial sequence variation in the Guahibo Amerindian population from Venezuela. *Am. J. Phys. Anthropol.* 127:361–369.

Von Hagen A, and C. Morris. 1998. *The Cities of the Ancient Andes*. London: Thames and Hudson.

Wang, S., C. M. Lewis Jr., M. Jakobsson et al. 2007. Genetic variation and population structure in Native Americans. *PLoS Genet.* 3:e185.

Ward, R. H., H. Gershowitz, M. Layrisse et al. 1975. The genetic structure of a tribal population, the Yanomama Indians. XI. Gene frequencies for 10 blood groups and the ABH-Le secretor traits in the Yanomama and their neighbors; the uniqueness of the tribe. *Am. J. Hum. Genet.* 27:1–30.

Ward, R. H., F. M. Salzano, S. L. Bonatto et al. 1996. Mitochondrial DNA polymorphism in three Brazilian Indian tribes. *Am. J. Hum. Biol.* 8:317–323.

Wernke, S. A. 2007. Negotiating community and landscape in the Peruvian Andes: A transconquest view. *Am. Anthropol.* 109:130–152.

Wightman, A. M. 1990. *Indigenous Migration and Social Change: The* Forasteros *of Cuzco, 1520–1720*. Durham, NC: Duke University Press.

Yang, N. N., S. Mazières, C. Bravi et al. 2010. Contrasting patterns of nuclear and mtDNA diversity in Native American populations. *Ann. Hum. Genet.* 74:525–538.

Zheng H.-X., S. Yan, Z.-D. Qin et al. 2012. mtDNA analysis of global populations support that major population expansions began before Neolithic time. *Sci. Rep*. 2:745–753.