Mitochondrial DNA Analysis of Ancient Peruvian Highlanders

Ken-ichi Shinoda,[1]* Noboru Adachi,[2] Sonia Guillen,[3] and Izumi Shimada[4]

[1]*Department of Anthropology, National Science Museum, Tokyo 169-0073, Japan*
[2]*Department of Anatomy and Anthropology, School of Medicine, Tohoku University, Sendai 980-8575, Japan*
[3]*Centro Mallqui Institute, Lima 11, Peru*
[4]*Department of Anthropology, Southern Illinois University, Carbondale, Illinois 62901*

KEY WORDS Inca; ancient DNA; Urubamba; Peru

ABSTRACT Ancient DNA recovered from 57 individuals excavated by Hiram Bingham at the rural communities of Paucarcancha, Patallacta, and Huata near the famed Inca royal estate and ritual site of Machu Picchu was analyzed by polymerase chain reaction, and the results were compared with ancient and modern DNA from various Central Andean areas to test their hypothesized indigenous highland origins. The control and coding regions of the mitochondrial DNA (mtDNA) of 35 individuals in this group were sequenced, and the haplogroups of each individual were determined. The frequency data for the haplogroups of these samples show clear proximity to those of modern Quechua and Aymara populations in the Peruvian and Bolivian highlands, and contrast with those of pre-Hispanic individuals of the north coast of Peru that we defined previously. Our study suggests a strong genetic affinity between sampled late pre-Hispanic individuals and modern Andean highlanders. A previous analysis of the Machu Picchu osteological collection suggests that the residents there were a mixed group of natives from various coastal and highland regions relocated by the Inca state for varied purposes. Overall, our study indicates that the sampled individuals from Paucarcancha and Patallacta were indigenous highlanders who provided supportive roles for nearby Machu Picchu. Am J Phys Anthropol 000: 000–000, 2006.

The discovery of the magnificent site of Machu Picchu in 1911, at an elevation of ca. 2,740 m above sea level atop a ridge connecting two granite peaks in the Eastern Cordillera of the Peruvian Andes by Bingham (1912, 1913), generated great professional and public interest in the site and in the Inca Empire in general. It also brought attention to the prehistory of this heretofore little-known rugged area approximately 100 km (ca. a 3-day walk) to the northwest of the Inca capital of Cuzco.

Understandably, the discovery of Machu Picchu raised broader questions, such as whether the settlement was isolated, intrusive, and/or specialized; and if not, the locations of the contemporaneous settlements in which the sustaining population lived; and whether the residents were local people. The explorations by Bingham in 1911 and 1912 focused on Machu Picchu and only identified a handful of relatively accessible sites on the Inca road to it. Bingham (1916, 1930) returned to the vicinity of Machu Picchu in 1914 (spanning 1914–1915) and attempted to answer these questions. Though the survey that aimed at locating additional Inca sites was conducted along Inca roads (Fig. 1), his fieldwork focused mainly on the excavation of the site at Patallacta (also known as Lacta Pata, ca. 2,500 m above sea level) that he had explored and mapped in 1911, prior to finding Machu Picchu. This site is impressive with respect to its size, orderly planned layout, and placement on an artificially leveled platform. The platform was carved out of a semicircular promontory overlooking the confluence of the Cusichaca and the Urubamba Rivers, a 17-km linear distance to the southeast of Machu Picchu. A more recent and comprehensive study of this site and nearby Inca sites was conducted by the Cusichaca Archaeological Project led by Kendall (1974, 1985, 1988). This study suggests that Patallacta was an important Inca local administrative center, surrounded by satellite settlements with varied functions that complemented the center. Patallacta had 115 *kanchas* or walled compounds, each composed of four buildings surrounding a central courtyard, that were organized into "four distinct residential sectors offering accommodation for social groups of varied status and structure" (Kendall, 1988).

This settlement complex included the relatively small but well-preserved "fort" of Paucarcancha, the place from which the majority of human remains that were analyzed in the present study were derived. The site is located approximately 5 km upstream (south) of Patallacta, atop terraces (3,000 m above sea level) overlooking the confluence of the Cusichaca and the Quescamayo Rivers; it presumably guarded the road to and from the Urubamba Valley (Kendall, 1985). The curving perimeter wall encloses approximately 18 rectangular structures. Based on its location relative to various major state

Grant sponsor: Ministry of Education, Science, Sports and Culture, Japan; Grant number: 13575017.

*Correspondence to: Ken-ichi Shinoda, Department of Anthropology, National Science Museum, 3-23-1 Hyakunincho, Shinjyuku-ku, Tokyo 169-0073, Japan. E-mail: shinoda@kahaku.go.jp

Received 18 February 2005; accepted 21 October 2005.

DOI 10.1002/ajpa.20408
Published online in Wiley InterScience (www.interscience.wiley.com).

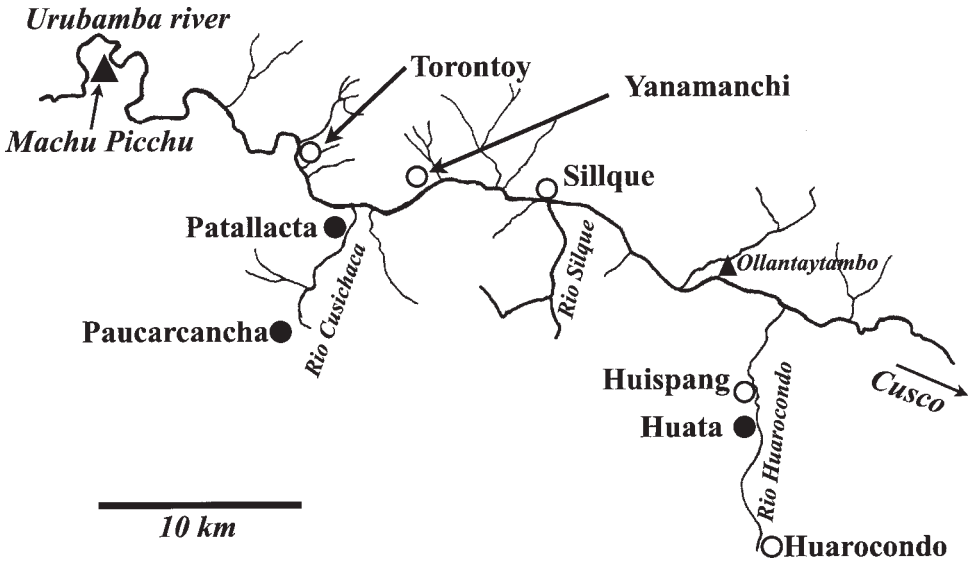

Fig. 1. Map of Machu Picchu and other archaeological sites. Circles indicate sites discovered by Hiram Bingham. Solid circles indicate sites analyzed in this study.

installations and roads and its architectural and ceramic styles, the history of Paucarcancha dates back to the reign of the Inca king Topa Inca (son of the king Pachacuti Inca Yupanqui), approximately in the late 15th century (Kendall, 1985). Based on architecture, ceramics, and other artifacts found in association, the burials that Bingham excavated at Paucarcancha and Patallacta can be assigned to the period of the Inca control of the Urubamba Valley, from ca. mid-15th to early 16th centuries (Bingham, 1913; Kendall, 1985; MacCurdy, 1923).

Over the past 20 years, in addition to the aforementioned work led by Kendall, there has been much effort to elucidate Inca and pre-Inca occupations along the "Sacred Valley of the Inca" and further downstream along the Urubamba River past Machu Picchu (e.g., Farrington, 1995; Niles, 1984, 1988, 1999; Protzen, 1993). Bingham (1930) believed that Machu Picchu was built and occupied by various Inca kings and served as a citadel and as the final Inca capital of Vilcapampa.

On the other hand, following a report by Rowe (1990), a majority of scholars now believe that it was a royal country estate or a palace of the Inca king Pachacuti Inca Yupanqui. In fact, Machu Picchu was one of at least 18 palaces built by three generations of Inca kings (Niles, 1988, 1999).

These estates were associated with large-scale, state-directed land-reclamation and land-management projects that created thousands of irrigated agricultural terraces, making the Urubamba Valley the breadbasket for the Inca capital of Cuzco (e.g., Kendall, 1988). In addition to its relatively low elevation, as Protzen (1993) noted, the Sacred Valley with its roughly east-west orientation created a thermal belt or microenvironment that favored intensive agriculture. Large-scale maize cultivation at that site appears to have been manned, in part, by "experienced maize growers from other parts" of the Inca Empire (Von Hagen and Morris, 1998; Niles, 1984). A majority of the 68 settlements in the Inca-period settlements of diverse sizes that Kendall (1985, 1988) recorded (including Patallacta and Torontoy) along three major roads between the Santa Teresa (downstream from Machu Picchu) and Anta Rivers (just upstream from Ollantaytambo) probably served, in part, as residential settlements for both local and relocated farmers. These observations suggest that Inca-period burials at surveyed settlements near Machu Picchu may include both local and genetically distinct nonlocal individuals.

The present study aims at shedding light on the biological identities and geographical origins of Inca-period residents of the Machu Picchu-Cusichaca section of the Urubamba Valley. Eaton (1916) described the materials recovered at Machu Picchu by the projects led by Hiram Bingham of Yale University. He included a thorough report on the bones and artifacts, as well as information on the condition of the caves where the material had been located. Eaton (1916) concluded that females greatly outnumbered males at Machu Picchu, by a ratio of approximately 4:1. This is the scientific basis for Bingham's idea that the site included an *Aclla Wasi*, a House of the Chosen Women. More recent research questioned this interpretation. Reanalysis of the data of Eaton (1916) revealed that the sex ratio was 1.54:1 (Verano, 2003). Eaton (1916) also proposed that the presence of brachycephalic and dolicocephalic skulls was indicative of highland and coastal groups. The recent examination of various types of cranial deformities, and a related multivariate craniometric analysis of the Machu Picchu osteological collection by Verano (2003), suggest that the residents were a mixed group of natives from various coastal and highland regions. These findings support the earlier conclusion by Eaton (1916) and raise an intriguing question regarding whether the residents of nearby communities were also of mixed origins.

MacCurdy (1918, 1923), the physical anthropologist who participated in the aforementioned Yale University expedition of 1914–1915, studied the material recovered by Bingham from places near Machu Picchu, i.e., Paucarcancha, Patallacta, Torony, Sillque, Huispand, Huata, Huarocondo, and Yanamanchi (Fig. 1). MacCurdy (1918, 1923) also made thorough descriptions of the material studied, including statistical information. A series of measurements was included, along with photographs of the sites and bones. His summary of the study of the material from the eight sites indicated that according to the funerary pattern and characters, the individuals all belonged to the same group. They had been buried in caves and shelters. He noted the high frequency of the highland or Aymara type of deformation and the almost complete ab-

sence of the brachycephalic, coastal type. This was in great contrast to the Machu Picchu sample.

As the Inca state practiced forced relocation of populations for political-military, economic, and other purposes (Rowe, 1982), there is a question as to the provenience and biological identities of residents at the sites from which the samples were derived. While ethnohistorical documents and architectural quality and style, as well as site location, all point to Machu Picchu as a planned royal estate and ceremonial center, the connected settlements of Paucarcancha, Patallacta, and Huata lack architectural or other features diagnostic of Inca royal or ceremonial sites; rather, they show similarities to local pre-Inca antecedents. Based on these observations and their proximity to cultivable fields and water sources, we hypothesize that the burials excavated at Paucarcancha and Patallacta are those of highland agricultural populations derived from immediate or nearby highland areas. Accordingly, we expect that these burials as a group would be characterized by relatively homogeneous DNA composition, similar to modern-day highland Indians (Quechua speakers). On the other hand, the presence of individuals relocated from distant lands within the Inca Empire would yield an internally heterogenous genetic composition. We would expect a contrasting situation for the Machu Picchu.

In the current study, we present data from mtDNA control and coding region of ancient Peruvian highlanders in the Urubamba Valley, to address questions about the genetic relationships among these ancient groups and South American populations. Our purpose is to test the hypothesis that ancient highlanders were more closely related to modern highlanders than to other modern populations in the region.

MATERIALS AND METHODS

Sampling

To analyze the mitochondrial DNA of these residents, we sampled 57 teeth that had been collected by G. MacCurdy. We sampled 26 individuals from Paucarcancha, 24 from Patallacta, and seven from Huata. Although the materials from this expedition had been initially exported to the US, they were returned to Peru (Burger and Salazar, 2003) "in accordance with the original agreement with the Peruvian government," and are housed in the National Museum of Anthropology, Archaeology and History in Lima (Museo Nacional de Antoropologia Arqueologia e Historia del Peru, MNAAH).

Because the collection was returned to Peru, these skeletal remains were preserved commingled with materials from other sites. The remains were sorted out in accordance with the provenience information recorded in the Museum catalog. However, we were able to find samples recovered from only three sites, i.e., Paucarcancha, Patallacta, and Huata. There is a possibility that the remaining samples are lying unrecorded in the MNAAH without any record. Only specimens that were clearly labeled and that coincided with Macurdy's inventory numbers were used in our study.

Teeth were collected only from the mandible, to avoid the possibility of sampling from the same individual. A well-preserved tooth was extracted from each individual. Teeth, thus collected, were exported from Peru to Japan with the permission of the National Institute of Culture of Peru.

Authentication methods

No modern DNA-based studies had been performed previously in the area subsumed by our ancient DNA study. Standard contamination precautions, such as separation of pre- and post-polymerase chain reaction (PCR) experimental areas, use of disposable laboratory ware and filter-plugged pipette tips, treatment with DNA contamination removal solution (DNA-OFF^{TM}, TaKaRa, Otsu, Japan), ultraviolet irradiation of equipment and benches, negative extraction controls, and negative PCR controls, were employed in the present study. DNA-based experiments were performed by two authors (N.A. and K.-I.S.) independently, in different laboratories for cross-validation.

Extraction and purification of DNA

Tooth samples were dipped in a 13% bleach solution for 15 min, rinsed several times with DNase-/RNase-free distilled water (Invitrogen, Carlsbad, CA), and allowed to air-dry. Moreover, using a dental drill, the outer surface of samples was removed to a depth of 1 mm. Next, samples were again rinsed with DNase-/RNase-free distilled water and allowed to air-dry. After samples were completely dry, they were irradiated in an ultraviolet cross-linker for 30 min and turned over several times. Next, samples were pulverized in a mill (Multi-beads Shocker MB400U, Yasui Kikai, Osaka, Japan). The powdered samples (0.3–0.5 g) were decalcified with 0.5 M EDTA (pH 8.0) (Invitrogen) at 4°C for 5–7 days, and then rinsed three times with DNase-/RNase-free distilled water.

Two different methods were employed for the extraction of DNA from the powdered samples. In the first method, decalcified samples were lysed in 3 ml of Buffer ASL (Qiagen, Hilden, Germany) with 150 μl of 20 mg/ml Proteinase K (Invitrogen) at 55°C overnight. DNA was extracted from the lysate using a QIAamp® DNA Stool Mini Kit (Qiagen), in accordance with the technical manual. In the second method, DNA was extracted in two steps, using a GENECLEAN kit for ancient DNA (Bio 101). The pulverized tooth (powder) was placed in a 15-ml conical tube, and 5 ml of an overnight soaking solution were added for Proteinase K digestion. Samples were rotated and incubated at 37°C for 12–15 hr. The supernatant was used for DNA extraction with the kit. Approximately 100 μl of extracted DNA solution were obtained. The eluted DNA was amplified by PCR without further processing. DNA extraction was done only once, and if the following PCR amplification was not successful, no further extraction was carried out.

Amplification and sequencing of HVR 1, HVR 2, and coding region 10382–10465

A segment of hypervariable region (HVR) 1 (nucleotide positions 16209–16402, relative to the revised Cambridge reference sequence (CRS); Andrews et al., 1999), HVR 2 (128–267), and a segment of the coding region (10382–10465) that covers a part of the NADH dehydrogenase 3 and $tRNA^{Arg}$ gene were sequenced for all samples. The mtDNA sequence can be tentatively assigned to respective haplogroups according to specific mutations ob-

TABLE 1. Primers used for sequence analysis

Primer	Sequence (5′–3′)	Nucleotide positions	Annealing temperature (°C)	Reference
M13-L127	(−21M13)[1] AGCACCCTATGTCGCAGTAT	128–267	46	Adachi et al. (2003)
M13-H268	(M13 reverse)[2] TGTTATGATGTCTGTGTGG			
M13-L10381	(−21M13) CTGGCCTATGAGTGACTACA	10382–10465	50	Adachi et al. (2004)
M13-H10466	(M13 reverse) TGTAAATGAGGGGCATTTGG			
M13-L16208	(−21M13) CCCCATGCTTACAAGCAAG	16209–16402	50	Nata et al. (1999)
M13-H16403	(M13 reverse) ATTGATTTCACGGAGGATGG			

[1] 21M13 sequence represents 5′-TGTAAAACGACGGCCAGT-3′.
[2] M13 reverse sequence represents 5′-AACAGCTATGACCATG-3′.

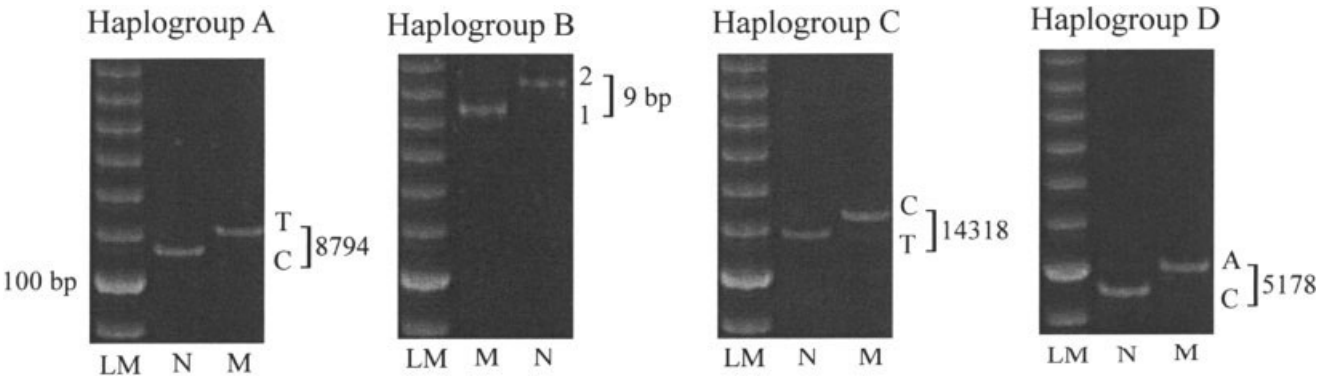

Fig. 2. APLP band patterns of diagnostic regions for haplogroups A, B, C, and D. Lane N indicates that no nucleotide change was observed. Lane M indicates mutation at site. Lane LM indicates 10-bp ladder marker (Invitrogen).

served in the HVR 1 region. Further characterization of haplogroup status was by means of other specific mutations in HVR 2 and the coding region. This coding region sequence includes the detection site of macrohaplogroups M and N. Each of sites 10398 and 10400 is one of the defining sites for macrohaplogroups M and N (Quintana-Marci et al., 1999). The known Native American haplogroups A and B are included in N, while C and D pertain to M. Detection of this coding region provides a firmer basis for haplogroup assignment. If haplogroup assignments derived from HVR 1 and coding region sequences did not accord with each other, the result was rejected. Primers used to amplify the regions described above are listed in Table 1.

A 1-μl aliquot of the extract was used as the template for PCR. Amplifications were carried out in a total reaction volume of 10 μl containing 0.25 units of Taq DNA polymerase (HotStarTaq™ DNA polymerase, Qiagen), 0.2 μM of each primer, 200 μM of dNTPs, and 0.1 μg/μl of bovine serum albumin (Amersham Pharmacia Biotech, Uppsala, Sweden) in 1 × PCR buffer (Qiagen). The conditions for PCR included incubation at 95°C for 15 min; 40 cycles at 94°C for 20 sec, 46°C–50°C for 20 sec, and 72°C for 15 sec; and final extension at 72°C for 1 min.

Following PCR, the products were purified with a MinElute™ PCR Purification Kit (Qiagen) and subjected to direct sequencing. Sequence reactions were prepared using 21M13 and M13 reverse primers and a DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Pharmacia Biotech). All sequencing reactions were analyzed using a 377 or 310 DNA Sequencer with SeqEd software (Applied Biosystems, Foster City, CA), and the sequence of each region that did not contain primer regions was determined and compared with the revised CRS.

Amplified product-length polymorphisms analysis

Several methods are available for the detection of single-nucleotide polymorphisms (SNPs). Restriction fragment-length polymorphism (RFLP) analysis is one of the most widely used techniques for detecting known mutation sites in the coding region of mtDNA. Recently, a simple and rapid amplified product-length polymorphisms (APLP) method was developed for the analysis of mtDNA polymorphisms (Umetsu et al., 2001). The principle of this method is based on an attachment of a noncomplementary sequence to the 5′-end of one of two allele-specific primers. The use of such primers permits the amplification of two size-different products distinguishing two alleles (Fig. 2). Using APLP, diagnostic polymorphic sites can be examined directly. Therefore, ambiguity with respect to the actual mutation site, which is one of the problems of the RFLP method caused by insufficient enzymic digestion, can be avoided.

It is generally agreed that most of the mtDNA from Native Americans could be traced to one of four maternal lineages present in the founders of New World populations. Further, these lineages can be defined by three restriction-site polymorphisms and a 9-bp deletion (Wallace et al., 1985). In the present study, three SNPs in the coding region (np 5178, 8794, and 14318) and a 9-bp repeat variation in the noncoding cytochrome oxidase II/$tRNA^{Lys}$ intergenic region (9 bp) were analyzed. The primers used for analyses are listed in Table 2. The constitution of the PCR reaction mixture was the same as described above. Thermal conditions were as follows: incubation at 95°C for 15 min; 40 cycles at 94°C for 10 sec, 52°C–54°C for 10 sec, and 72°C for 5 sec; and final extension at 72°C for 1 min. Each region was examined

TABLE 2. Primers used for APLP analysis

Site	Primer	Sequence[1] (5′–3′)	Product size (genotype)	Annealing temperature (°C)	Reference
5178	5178C	gTCGCACCTGAAgCAAGC	96 bp (C type)	52	Umetsu et al. (2001)
	5178A	tgatcaaCGCACCTGAAACAAGA	101 bp (A type)		
	5178R	attGCAAAAAGCaGGTTAGCG			
8794	8794C	aCCTCGGACTtCTGCCTC	107 bp (C type)	52	Umetsu et al. (2001)
	8794T	attggaCCTCGGACTCtTGCCTT	112 bp (T type)		
	8794R	aCAGCGAAAGCCTATAATCAC			
14318	14318T	CCTTCATAAATTATTCAGCTTCCaACACTAT	110 bp (T type)	54	Present study
	14318C	aaaaagctaCATAAATTATTCAGCTTCCTACtCTAC	115 bp (C type)		
	14318R	TTAGTGGGGTTAGCGATGGA			
9 bp	9bp-F	(-21M13)[2] ACAGTTTCATGCCCATCGTC	143 bp (1; deletion)	50	Present study
	9bp-R	(M13 reverse)[3] CTAAGTTAGCTTTACAGTGGG	152 bp (2; normal)		

[1] Noncomplementary nucleotides are written in lower-case letters.
[2] -21M13 sequence represents 5′-TGTAAAACGACGGCCAGT-3′.
[3] M13 reverse sequence represents 5′-AACAGCTATGACCATG-3′.

independently, using the monoplex PCR method to maximize the robustness of PCR.

A 1-μl aliquot of the PCR product was separated by electrophoresis in an 8-cm native polyacrylamide gel (10% T, 5% C) containing 1 × TBE buffer (pH 8.0) with running buffer (0.5 × TBE, pH 8.0). DNA bands were detected by ultraviolet irradiation after staining with ethidium bromide (Fig. 2).

Data analysis

With improved knowledge of the global mtDNA tree in recent years, an understanding of the structure of mtDNA data and assigning the mtDNA type to a place in the global mtDNA tree have been simplified. Control-region motifs were identified for a majority of the major haplogroups and their subhaplogroups (Alves-Silva et al., 2000; Bandelt et al., 2001; Kivisild et al., 2002; Kong et al., 2003; Macaulay et al., 1999; Maruyama et al., 2003; Quintana-Murci et al., 1999; Yao et al., 2002, 2003).

Therefore, we assigned each mtDNA to haplogroups according to the HVR 1, HVR 2, and coding-region data, using the data and classification tree described above, such that each sample was allocated to the smallest named haplogroup to which it belonged. If the haplogroup had further characterized subhaplogroups, an asterisk was attached to the name of the haplogroup to indicate that the haplogroup status could not be identified further (Table 3). Since several segments of the same mtDNA were analyzed independently, meticulous care was taken to avoid artificial recombination caused by potential sample crossover. After assigning the mtDNAs to relevant haplogroups, we classified them further into maternal lines, based on the nucleotide changes observed in the control and coding regions.

To elucidate biological relationships between hypothesized ancient highlanders from the communities of Paucarcancha, Patallacta, and Huata on the one hand, and Native American mtDNA on the other, 1,063 contemporary Native Americans from 16 populations and 36 pre-Hispanic north coast community haplogroup data were obtained from published data (Torroni et al., 1993, 1994; Batista et al., 1995; Santos et al., 1994; Kolman et al., 1995; Rickards et al., 1999; Merriwether et al., 1995; Rodriguez-Delfin et al., 2001; Lewis et al., 2005; Baillet et al., 1994; Ginther et al., 1993; Shimada et al., 2004). Haplogroup frequencies of these populations were compared with the ancient individuals under study.

Population differentiation tests (Raymond and Rousset, 1995) were computed by means of the Arlequin 2.000 computer program (Schneider et al., 2000).

RESULTS

Throughout the study, both negative extraction controls and negative PCR controls showed negative results consistently. Of the 57 specimens examined, 35 specimens were successfully analyzed (Table 3). The first method that used the stool kit recovered DNA from 31 samples. On the other hand, the second method that used the ancient DNA kit recovered DNA from 35 samples. Thus, the ancient DNA kit was observed to be the superior method for recovering DNA. However, suspected false-positive results stemming from contamination with contemporary DNA and other questionable data (e.g., Kolman and Tuross, 2000) were obtained while using this method. Following this criterion, two of the sequences were omitted.

A majority of the mtDNA of these specimens fully represented haplogroup motifs; therefore, we could safely assign them to the relevant haplogroups. This implied that the authenticity of data obtained from these specimens was supported by established, reliable, modern human mtDNA data. However, in the DNA analysis of ancient samples, the possibility that the original sequences may have changed due to postmortem damage to the DNA (Thomas et al., 2003) must also be considered. Thus, it is not advisable to assume the accuracy of all the base sequences determined in the investigation under discussion. It should be clearly understood that such a limitation would inevitably occur in the analyses of the scarce DNA that remains in the ancient samples. In fact, the sequence of some nucleotide positions remained uncertain in some specimens (Table 3).

In the present study, the base sequences in the control regions of 35 individuals were successfully determined. The sequences in 195 base pairs of HVR 1and 139 base pairs of HVR 2 were determined, and mutations were observed in 30 portions and 14 portions, respectively (Table 3). These individuals were classified into 27 haplotypes according to their sequences. The rate of successful recovery and determination of mtDNA sequence varied a good deal from one site to the next. For example, of the 26 individuals from Paucarcancha and 24 from Patallacta, 16 and 17 were successfully typed, resulting in high DNA

TABLE 3. Nucleotide changes observed in ancient Peruvian highlanders analyzed in present study

Site and specimen number	Haplogroup	Maternal line	Mutations in segments[1]			APLP analysis[3]			
			16209–16402 (16000+)	128–267[2]	10382–10465 (10000+)	5178	8794	14318	9 bp
Paucarcancha									
195	A*	A*-1	223 ***290 319*** 362	146 ***235***	CRS	·	***T***	·	2
208	A*	A*-1	223 ***290 319*** 362	146 ***235***	CRS	·	***T***	·	2
216	A*	A*-2	217 223 266 ***290 319*** 343T 362	146 153 ***235*** 260	CRS	·	x	·	2
192	B4*	B4*-1	***217*** 272 362	CRS	CRS	·	·	·	***1***
213	B4*	B4*-2	***217*** 289	143	CRS	·	·	·	***1***
198	B4*	B4*-2	***217*** 289	143	ND	·	·	·	***1***
203	B4*	B4*-3	***217***	146 215	CRS	·	x	·	***1***
210	B4*	B4*-4	***217*** 228 379N	214	CRS	·	·	·	***1***
212	B4*	B4*-5	214 ***217*** 262	231N	CRS	·	·	·	***1***
214	B4*	B4*-6	***217*** 278	146 215	CRS	·	·	·	***1***
227	B4*	B4*-7	***217*** 357	143	CRS	·	·	·	***1***
233	B4*	B4*-8	***217*** 362	CRS	CRS	·	·	·	***1***
230	B4a	B4a-1	***217 261*** 319	CRS	CRS	·	·	·	***1***
193	C*	C*-1	223 ***298*** 325 ***327***	146 ***249d***	398 ***400***	·	·	***C***	2
204	C*	C*-1	223 ***298*** 325 ***327***	146 ***249d***	398 ***400***	·	·	***C***	2
211	C*	C*-2	223 ***298*** 325 ***327***	***249d***	ND	·	·	***C***	2
Patallacta									
680	B4*	B4*-2	***217*** 289	143	CRS	·	·	·	***1***
978	B4*	B4*-3	***217***	146 215	CRS	·	·	·	***1***
681	B4*	B4*-9	***217*** 296N 321 363 390	214 234	CRS	·	·	·	***1***
686	B4*	B4*-10	***217***	152	CRS	·	·	·	***1***
689	B4*	B4*-10	***217***	152	CRS	·	·	·	***1***
687	B4*	B4*-11	***217***	CRS	CRS	·	·	·	***1***
974	B4*	B4*-11	***217***	CRS	CRS	·	·	·	***1***
981	B4*	B4*-12	***217*** 268 348 378 379	x	CRS	·	·	·	***1***
989	B4*	B4*-13	***217*** 294	143 210	CRS	·	·	·	***1***
677	B4*	B4*-14	***217***	152, 204	CRS	·	·	·	***1***
683	B4a	B4a-2	***217 261***	CRS	CRS	·	·	·	***1***
976	B4a	B4a-3	***217 261N*** 357	143	CRS	·	·	·	***1***
678	B*	B*-1	***217*** 381	CRS	***398***	·	·	·	***1***
682	C*	C*-1	223 ***298*** 325 327	146 195 ***249d***	398 ***400***	·	·	***C***	2
975	C*	C*-3	223 246N ***298*** 325 327 373	x	398 ***400***	·	·	***C***	2
676	C*	C*-1?	223 ***298N*** 325N 327	x	398 ***400***	·	·	***C***	2
977	D*	D*-1	325 362N	CRS	398 ***400***	***A***	·	·	2
Huata									
899	C*	C*-1	223 ***298*** 325 ***327***	x	398 ***400***	·	·	***C***	2
897	C*	C*-4	223 ***298*** 325 ***327***	x	392 ***400***	·	·	***C***	2

[1] All polymorphic sites are numbered according to revised CRS (Andrews et al., 1999). CRS denotes that sequence of segment is identical to revised CRS, and N indicates not determined. Suffix T indicates transversion, and d indicates deletion. Deletions are recorded at last possible site. x indicates that PCR product was not obtained at region. Diagnostic polymorphisms are indicated by bold italic type.

[2] Nucleotide change at position 263 in segment 128–267 was observed in all specimens, and therefore is omitted.

[3] Diagnostic polymorphisms are emphasized by bold italic type. Dot indicates that no nucleotide change was observed. x indicates that PCR product was not obtained in region.

recovery and sequencing rates of 61.5% and 70.8%, respectively. In contrast, of seven individuals from the Huata, only two (or 28.6%) were successfully sequenced.

Haplogroup distribution for the total sample was as follows: 8.6% A, 65.7% B, 22.9% C, and 2.9% D. Haplogroup frequencies of contemporary Amerindian populations and ancient north coast samples are also shown in Table 4. F-statistics from haplogroup frequencies among regional populations are shown in Table 5. An exact test of differentiation between each pair of populations revealed statistically significant differences except between the ancient highlanders and contemporary central Andean population (significant F_{st} $P = 0.180 \pm 0.054$).

To investigate the relationships among the satellite communities of the royal estate of Machu Picchu, mtDNA sequences of Paucarcancha and Patallacta were compared. Haplogroup frequencies of Paucarcancha and Patallacta are shown in Table 6. Genetic diversity results for these two sites are shown in Table 7. Mean numbers of pairwise differences and nucleotide diversity are slightly larger in the Paucarcancha.

DISCUSSION

Haplogroup profile of individuals examined in the present study

We found that haplogroup B was the most frequent among skeletal samples analyzed in the Inca-period residents of the Urubamba Valley, followed by haplogroups C, A, and finally D. The most distinctive feature of the haplogroup profile of individuals examined in the present study is the high frequency of haplogroup B (65.7%; 23 of 35 individuals; Tables 3 and 4). Classifying individuals into maternal lines resulted in haplogroup B having at least 18 different lines in 23 individuals. In other words, the high frequency of haplogroup B is not caused by the concentration of individuals on a specific maternal line.

Haplogroup B is the common haplogroup in contemporary Central Andean populations. When the haplogroup profile of these ancient residents of the Urubamba Valley was compared with that of other South American populations, the former showed a clear proximity to the modern Central Andean populations that are distributed primarily in the Peruvian and Bolivian highlands (Table 4). This finding is not surprising, considering the highland location of the study area.

On the other hand, the ancient highlanders considerably differ from individuals of the ancient north coast community in terms of mtDNA haplogroup frequency. Various lines of archaeological evidence indicate intimate cultural interactions between the ancient north coastal populations and contemporaneous Ecuadorian and Colombian populations (Shimada, 1995, 1999; Shimada et al., 1997, 2000). Relatively high frequencies of haplogroup A and low frequencies of haplogroup C in ancient north coast populations hint at their linkage with modern lower Central American and North Andean populations. However, in the cases of samples from ancient north coast communities and ancient highlanders, there is the possibility that the frequency for each haplogroup was biased, because these samples were taken only from the few sites considered to have been related by blood relationship. Yet an exact test of differentiation (Raymond and Rousset, 1995) at least shows that the ancient highlanders bear a similarity to contemporary Central

TABLE 4. Frequency distribution of mtDNA haplogroups in ancient Peruvian highlanders, ancient north coast community (Sicán and Sipán), and contemporary Amerindian populations

Population	Haplogroups and their frequencies[1]					References
	A	B	C	D	Others	
Ancient highlanders	8.6% (3)	65.7% (23)	22.9% (8)	2.9% (1)	0% (0)	Present study
Ancient north coast community	19.4% (7)	22.2% (8)	5.6% (2)	30.6% (11)	22.2% (8)	Shimada et al. (2004)
Contemporary lower Central America (total)	58.2% (213)	29.5% (108)	3.0% (11)	2.2% (8)	7.1% (26)	Torroni et al. (1993, 1994), Batista et al. (1995), Santos et al. (1994), Kolman et al. (1995), Rickards et al. (1999)
Contemporary Central Andean (Quechua)	10.0% (7)	50.0% (35)	14.3% (10)	21.4% (15)	4.3% (3)	Merriwether et al. (1995), Rodriguez-Delfin et al. (1999)
Contemporary Central Andean (Ancash)	9.1% (3)	51.5% (17)	18.2% (6)	21.2% (7)	0% (0)	Lewis et al. (2004)
Contemporary Central Andean (Aymara)	6.4% (11)	67.4% (116)	12.2% (21)	14.0% (24)	0% (0)	Merriwether et al. (1995)
Contemporary Central Andean (Atacameños)	14.3% (9)	71.4% (45)	9.5% (6)	4 .8% (3)	0% (0)	Baillet et al. (1994), Merriwether et al. (1995)
Contemporary Central Andean (Total)	8.9% (30)	63.0% (213)	12.7% (43)	14.5% (49)	0.9% (3)	
Contemporary Southern Andean (Mapuche)	7.9% (11)	30.8% (43)	27.1% (38)	30.0% (42)	4.3% (6)	Ginther et al. (1993), Horai et al. (1993), Baillet et al. (1994)
Contemporary Southern Andean (Peheuncheˆ)	2.0% (2)	9.0% (9)	37.0% (37)	52.0% (52)	0% (0)	Merriwether et al. (1995)
Contemporary Southern Andean (Huilliche)	4.2% (5)	28.8% (34)	18.6% (22)	48.3% (57)	0% (0)	Baillet et al. (1994), Merriwether et al. (1995)
Contemporary Southern Andean (total)	5.0% (18)	24.0% (86)	27.1% (97)	42.2% (151)	1.7% (6)	
Contemporary Andean (total)	6.9% (48)	42.8% (299)	20.0% (14)	28.8% (201)	1.6% (11)	

[1] Numbers in parentheses indicate number of individuals who were assigned to relevant haplogroups.

TABLE 5. Pairwise F_{st} values between each pair of populations

	Ancient highlanders	Ancient north coast	Lower Central America	Central Andean
Ancient north coast	0.1956			
Lower Central America	0.2633	0.1883		
Central Andean	0.0064*	0.1689	0.2528	
Southern Andean	0.1869	0.0679	0.2866	0.1648

* F_{st} *P*-value between ancient highlanders and contemporary Central Andean population is 0.180 ± 0.054 (standard error).

TABLE 6. Frequency distribution of mtDNA haplogroups in ancient Peruvian highlanders excavated from Paucarcancha and Patallacta sites

	Haplogroups and their frequencies[1]				
Population	A	B	C	D	Others
Paucarcancha	18.8% (3)	62.5% (10)	18.8% (3)	0.0% (0)	0.0% (0)
Patallacta	0.0% (0)	75.0% (12)	18.8% (3)	6.3% (1)	0.0% (0)

[1] Numbers in parentheses indicate number of individuals who were assigned to relevant haplogroups.

TABLE 7. mtDNA HV 1 haplotype diversity parameters of Paucarcancha and Patallacta[1]

	n	k	Pw	π	Pairwise F_{st}
Paucarcancha	16	12	4.475	0.023	0.0222
Patallacta	17	13	3.059	0.016	

[1] n, sample size; k, number of different sequences; Pw, mean number of pairwise differences; π nucleotide diversity.

Andean people on the basis of sharing similar haplogroups.

As described above, morphological analysis of the Machu Picchu residents suggests that they were a mixed group of natives from various coastal and highland regions. The mixed composition is likely to reflect the pervasive Inca practice of recruiting male and female retainers and specialists from subjugated populations and relocating them in accordance with royal and/or state purposes. Thus, the haplogroup profile of Machu Picchu individuals is expected to be different from that of residents of their rural hinterlands that we examined in the present study. This question should be solved directly bythe DNA analysis of skeletons excavated from the Machu Picchu site.

Relationship between sites

One of the main purposes of analyzing DNA from ancient burial sites is to determine whether the human remains represent related or unrelated individuals. As described above, our samples were mostly obtained from two sites (Paucarcancha and Patallacta) that represent satellite support communities of the royal estate of Machu Picchu. On comparing the haplotype profiles of individuals excavated from the two sites, an interesting observation was made. As shown in Tables 6 and 7, both populations have high frequencies of haplogroup B, and the population pairwise F_{st} was not so large. Moreover, an exact test of differentiation revealed that these differences are not statistically significant ($P = 0.153 \pm 0.039$ (standard error)). However, as shown in Table 3, at least 13 and 14 maternal lines were distinguished in Paucarcancha and Patallacta, respectively; these two sites shared only three maternal lines. The result of our mtDNA analysis suggests that the maternal structure of the Paucarcancha and Patallacta individuals was clearly different. On the basis of our mtDNA data alone, we cannot discount the distinct possibility that at least some of the individuals we analyzed had been relocated by the Inca state to Paucarcancha and Patallacta from other Peruvian and/or Bolivian highland regions. Furthermore, the frequencies of mtDNA haplotypes in the sampled populations could be biased due to the small sample size, with the true frequencies likely to be different when more individuals are typed. Further genetic studies are needed to confirm our results.

CONCLUSIONS

Our mtDNA study showed the clear proximity of ancient residents of the Urubamba Valley near Machu Picchu to modern Quechua and Aymara populations in the Peruvian and Bolivian highlands, respectively. Further, these residents were considerably different from pre-Hispanic individuals of the north coast of Peru in terms of mtDNA haplogroup frequency. Thus, the results support the hypothesis that the remains of residents of the Inca-period rural communities of Paucarcancha and Patallacta that were analyzed were native highlanders, although we cannot define their exact ethnic and geographical origins; we argue that they served roles of supporting the nearby Inca royal estate of Machu Picchu. We suspect that future ancient DNA analysis of the Machu Picchu residents will show that they were genetically distinct from the sampled individuals in the present study. However, the fact that mtDNA analysis does not provide any information regarding genetic connections along the paternal line should be considered. Moreover, characteristics of group genetics are influenced by a number of factors that include not only natural phenomena, such as genetic drift or the bottleneck effect, but also social and cultural elements, such as marriage, war, or trade. Therefore, it is impossible to clarify all the changes that a group had undergone through genetic analysis alone. In order to ascertain the fate of the inhabitants of an archaeological site, in addition to clarifying its regional context, genetic data should be viewed as one of multiple independent lines of evidence garnered by a team of archaeologists, physical anthropologists, and other complementary specialists.

ACKNOWLEDGMENTS

We are grateful to Elsa Tomasto and Hilda Vidal (curators of the Museo Nacional de Antoropologia

Arqueologia e Historia del Peru) and Japanese photographer Yutaka Yoshii for their assistance in the collection of tooth samples used in the mtDNA analysis. Research by K.-I.S. for this study was supported by Grant-in-Aid for Scientific Research 13575017 from the Ministry of Education, Science, Sports and Culture, Japan.

LITERATURE CITED

Adachi N, Dodo Y, Ohshima N, Doi N, Yoneda M, Matsumura H. 2003. Morphologic and genetic evidence for the kinship of juvenile skeletal specimens from a 2,000 year-old double burial of the Usu-Moshiri site, Hokkaido, Japan. Anthropol Sci 111:347–363.

Adachi N, Umetsu K, Takigawa W, Sakaue K. 2004. Phylogenetic analysis of the human ancient mitochondrial DNA. J Archaeol Sci 31:1339–1348.

Alves-Silva J, Santos MDS, Guimarães PEM, Ferreira ACS, Bandelt HJ, Pena SDJ, Prado VF. 2000. The ancestry of Brazilian mtDNA linages. Am J Hum Genet 67:444–461.

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

Bailliet G, Rothhammer F, Carnese FR, Bravi CM, Bianchi NO. 1994. Founder mitochondrial haplotypes in Amerindian populations. Am J Hum Genet 55:27–33.

Bandelt HJ, Lahermo P, Richards M, Macaulay V. 2001. Detecting errors in mtDNA data by phylogenetic analysis. Int J Legal Med 115:64–69.

Batista O, Kolman CJ, Bermingham E. 1995. Mitochondrial DNA diversity in the Kuna Amerinds of Panama. Hum Mol Genet 4:921–929.

Bingham H. 1912. Preliminary report of the Yale Peruvian expedition. Bull Am Geogr Soc 44:20–26.

Bingham H. 1913. In the wonderland of Peru. Natl Geogr Mag 24:387–573.

Bingham H. 1916. Further explorations in the land of the Incas. Natl Geogr Mag 29:431–473.

Bingham H. 1930. Machu Picchu, a citadel of the Incas. Report of the explorations and excavations made in 1911, 1912 and 1915 under the auspices of Yale University and the National Geographic Society. New Haven: Yale University Press.

Burger RL, Salazar LC. 2003. Preface. In: Burger R, Salazar LC, editors. The 1912 Yale Peruvian scientific expedition collection from Machu Picchu: human and animal remains, Yale University publications in anthropology 85. New Haven: Peabody Museum of Natural History, Yale University. p xiii–xv.

Eaton GF. 1916. The collection of osteological material from Machu Picchu. Memoirs of the Connecticut Academy of Arts and Sciences 5. New Haven: Tuttle, Morehouse and Taylor.

Farrington IS. 1995. The mummy, estate and palace of Inka Huayna Capac at Quispeguanca. Tawantinsuyu 1:55–65.

Ginther C, Corach D, Penacino GA, Rey JA, Carnese FR, Hutz MH, Anderson A, Just J, Salzano FM, King MC. 1993. Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial DNA sequence variation and allele frequencies of several nuclear genes. EXS 67:211–219.

Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima K. 1993. Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol Biol Evol 10:23–47.

Kendall A. 1974. Architecture and planning at the Inca sites in the Cusichaca area. Baessler Archiv 22:73–137.

Kendall A. 1985. Aspects of Inca architecture: description, function and chronology. BAR international series, 242. Oxford, British Archaeological Reports.

Kendall A. 1988. Inca planning north of Cuzco between Anta and Machu Picchu and along the Urubamba Valley. In: Saunders NJ, de Montomollin O, editors. Recent studies in pre-Columbian archaeology. Oxford: British Archaeological Reports. p 457–488.

Kivisild T, Tolk HV, Parik J, Wang Y, Papiha SS, Bandelt HJ, Villems R. 2002. The emerging limbs and twigs of the East Asian mtDNA tree. Mol Biol Evol 19:1737–1751.

Kolman CJ, Tuross N. 2000. Ancient DNA analysis of human populations. Am J Phys Anthropol 111:5–23.

Kolman CJ, Bermingham E, Cooke R, Ward RH, Arias TD, Guionneau-Sinclair F. 1995. Reduced mtDNA diversity in the Ngobe Amerinds of Panama. Genetics 140:275–283.

Kong QP, Yao YG, Sun C, Bandelt HJ, Zhu CL, Zhang YP. 2003. Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. Am J Hum Genet 73:671–676.

Lewis CM, Tito RY, Lizárraga B, Stone AC. 2005. Land, language, and loci: mtDNA in Native Americans and the genetic history of Peru. Am J Phys Anthropol 127:351–360.

MacCurdy V, Richards M, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, Bonné-Tamir, B, Sykes B, Torroni A. 1999. The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequence and RFLPs. Am J Hum Genet 64:232–249.

MacCurdy G. 1918. Surgery among ancient Peruvians. Art Archaeol 7:381–394.

MacCurdy G. 1923. Human skeletal remains in the Peruvian highlands. Am J Phys Anthropol 6:217–330.

Maruyama S, Minaguchi K, Saitou N. 2003. Sequence polymorphisms of the mitochondrial DNA control region and phylogenetic analysis of mtDNA lineages in the Japanese populations. Int J Legal Med 117:218–225.

Merriwether DA, Rothhammer F, Ferrell RE. 1995. Distribution of the four founding lineage haplotypes in Native Americans suggests a single wave of migration for the New World. Am J Phys Anthropol 98:411–30.

Nata M, Adachi N, Hashiyada M, Kanetake J, Ji G, Mimasaka S, Takahashi K, Sagisaka K, Funayama M. 1999. Sequence polymorphism of the D-loop of mitochondrial DNA in Japanese and Chinese population. Acta Crim Japan 65:244–254.

Niles SA. 1984. Architectural form and social function in Inca towns near Cuzco. In: Kendall A, editor. Current archaeological projects in the Central Andes—some approaches and results.British Archaeological Reports international series 210. Oxford: British Archaeological Reports. p 205–219.

Niles SA. 1988. Looking for "lost" Inca palaces. Expedition 30:56–64.

Niles SA. 1999. Shape of Inca history: narrative and architecture in an Andean empire. Iowa City: University of Iowa Press.

Protzen JP. 1993. Inca architecture and construction at Ollantaytambo. Oxford: Oxford University Press.

Quintana-Murci L, Semino O, Bandelt HJ, Passarino G, McElreavey K, Santachiara-Benerecetti AS. 1999. Genetic evidence for an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. Nat Genet 23:437–441.

Raymond M, Rousset F. 1995. An exact test of population differentiation. Evolution 49:1280–1283.

Rickards O, Martinez-Labarga C, Lum JK, De Stefano GF, Cann RL. 1999. mtDNA history of the Cayapa Amerinds of Ecuador: detection of additional founding lineages for the Native American populations. Am J Hum Genet 65:519–530.

Rodriguez-Delfin LA, Rubin-de-Celis VE, Zago MA. 1999. 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.

Rowe JH. 1982. Policies and institutions relating to the unification of the empire. In: Collier GA, Rosaldo RI, Wirth JD, editors. The Inca and Aztec states, 1400–1800. New York: Academic Press. p 93–118.

Rowe JH. 1990. Machu Picchu a la luz de documentos del siglo XVI. Historica 14:139–154.

Santos M, Ward RH, Barrantes R. 1994. mtDNA variation in the Chibcha Amerindian Huetar from Costa Rica. Hum Biol 66:963–77.

Schneider S, Roessli D, Excoffier L. 2000. Arlequuin version 2.000: a software for population genetics data analysis. Geneva: Genetic and Biometry Laboratory, University of Geneva.

Shimada I. 1995. Cultura Sicán: dios, riqueza y poder en la costa norte del Peru. Lima: Banco Continental.

Shimada I. 1999. The evolution of Andean diversity: regional formations, ca. 500 B.C.–A.D. 600. In: Salomon F, Schwartz S, editors. Cambridge history of native peoples of the Americas. Cambridge: Cambridge University Press. p 350–517.

Shimada I, Anderson KB, Haas H, Langenheim JH. 1997. Amber from 1000-year old prehispanic tombs in northern Peru. In: Vandiver PB, Druzik JR, Merkel JF, Stewart J, editors. Materials issues in art and archaeology V. Pittsburgh: Materials Research Society. p 3–18.

Shimada I, Gordus A, Jo A. 2000. Technology, iconography, and significance of metals: a multi-dimensional analysis of Middle Sicán Objects. In: McEwan C, editor. Pre-Columbian gold: technology, iconography, and style. London: British Museum Press. p 28–61.

Shimada I, Shinoda K, Farnum J, Corruccini R, Watanabe H. 2004. An integrated analysis of pre-Hispanic mortuary practices: a Middle Sican case study. Curr Anthropol 45:369–402.

Thomas M, Gilbert P, Willerslev E, Hansen AJ, Barnes I, Rudbeck L, Lynnerup N, Cooper A. 2003. Distribution patterns of postmortem damage in human mitochondrial DNA. Am J Hum Genet 72:32–47.

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

Torroni A, Chen YS, Semino O, Santachiara-Beneceretti AS, Scott CR, Lott MT, Winter M, Wallace DC. 1994. mtDNA and Y-chromosome polymorphisms in four Native American populations from southern Mexico. Am J Hum Genet 54:303–318.

Umetsu K, Tanaka M, Yuasa I, Saitou N, Takeyasu T, Fuku N, Naito E, Ago K, Nakayashiki N, Miyoshi A, Kashimura S, Watanabe G, Osawa M. 2001. Multiplex amplified product-length polymorphism analysis for rapid detection of human mitochondrial DNA variations. Electrophoresis 22:3533–3538.

Verano JW. 2003. Human skeletal remains from Machu Picchu: a reexamination of the Yale Peabody Museum's collections. In: Burger R, Salazar L, editors. The 1912 Yale Peruvian scientific expedition collection from Machu Picchu: human and animal remains. Yale University publications in anthropolology 85. New Haven: Peabody Museum of Natural History, Yale University. p 65–117.

Von Hagen A, Morris C. 1998. The cities of the ancient Andes. London: Thames and Hudson.

Wallace DC, Garrison K, Knowler WC. 1985. Dramatic founder effects in Amerindian mitochondrial DNAs. Am J Phys Anthropol 68:149–155.

Yao YG, Kong QP, Bandelt HJ, Kivisild T, Zhang YP. 2002. Phylogenetic differentiation of mitochondrial DNA in Han Chinese. Am J Hum Genet 70:635–651.

Yao YG, Kong QP, Man XY, Bandelt HJ, Zhang YP. 2003. Reconstructing the evolutionary history of China: a caveat about inferences drawn from ancient DNA. Mol Biol Evol 20:214–219.