Ancient DNA evidence of population replacement following
the Aztec conquest of Xaltocan, Mexico
Ancient DNA evidence of population replacement following the Aztec conquest of Xaltocan, Mexico

by

Jaime Mata-Míguez, B.S.; B.S.

Report
Presented to the Faculty of the Graduate School of The University of Texas at Austin in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

The University of Texas at Austin
December 2011
Acknowledgements

First and foremost, I thank Deborah A. Bolnick, my advisor. This master’s report would not have been completed without her guidance. I also thank Enrique R. Rodríguez-Alegria and Lisa Overholtzer for loaning the samples and instructing me in the archaeological aspects of this project. I will always be deeply grateful to the Fundación Pedro Barrié de la Maza for its wonderful fellowship. Thanks for this opportunity.
Abstract

Ancient DNA evidence of population replacement following the Aztec conquest of Xaltocan, Mexico

Jaime Mata-Miguez, MA
The University of Texas at Austin, 2011

Supervisor: Deborah A. Bolnick

The Aztec empire emerged in AD 1428 as a result of the triple alliance among the city-states of Tenochtitlan, Texcoco, and Tlacopan. Although it is well documented that the Aztecs conquered numerous polities in the Basin of Mexico over the next 100 years, the demographic consequences of this expansion remain unclear. At the influential Otomi city-state of Xaltocan, for example, colonial documents suggest that the Aztec conquest led to a replacement of the original Otomi population, whereas archaeological finds suggest that a significant portion of the original population may have remained at the city under Aztec rule. To help resolve questions about Xaltocan’s population history during this period, I extracted ancient DNA from 21 individuals that can be divided into two temporal subpopulations (roughly predating and postdating the hypothesized replacement event). I determined mitochondrial DNA haplogroups through RFLP analyses and constructed haplotypes based on 372 bp of HVR1 sequence. Statistical analyses show significant differences between the mitochondrial composition of the two subpopulations. Altogether, the results of this study support the hypothesis that matriline at Xaltocan
underwent a significant replacement event following the Aztec conquest, and they suggest that the Aztec expansion may have had a substantial genetic impact on certain Mesoamerican populations.
# Table of Contents

List of Tables ........................................................................................................ viii

List of Figures ........................................................................................................ ix

Introduction ......................................................................................................... 1
  Using ancient DNA to test population replacement hypotheses ................. 3

Materials and Methods ....................................................................................... 6
  Skeletal samples ........................................................................................... 6
  DNA extraction and purification ................................................................. 7
  PCR amplification and analysis ................................................................. 7
  Contamination controls ............................................................................. 8
  Statistical analyses ...................................................................................... 9

Results ............................................................................................................... 12

Discussion ........................................................................................................ 15

Appendix .......................................................................................................... 19

Bibliography .................................................................................................... 29
List of Tables

Table 1. Xaltocan samples analyzed in this study.................................................. 25
Table 2. Haplogroup frequencies. ......................................................................... 26
Table 3. Genetic diversity estimates (mean ± standard deviation). ...................... 27
Table 4. Weighted means of within-haplogroup Fst values................................. 28
List of Figures

Figure 1. Map of the Basin of Mexico in pre-Hispanic times. The dotted areas represent lakes (from Rodríguez-Alegría, 2010)............................. 20

Figure 2. Network for haplogroup A. Mutations are identified with reference to the Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999). Bold circles represent founding haplotypes. The size of the circles is proportional to the number of samples. White circles represent haplotypes not found in the populations employed in this analysis. 21

Figure 3. Network for haplogroup B. Mutations are identified with reference to the Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999). Bold circles represent founding haplotypes. The size of the circles is proportional to the number of samples. White circles represent haplotypes not found in the populations employed in this analysis. 22

Figure 4. Network for haplogroup C. Mutations are identified with reference to the Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999). Bold circles represent founding haplotypes. The size of the circles is proportional to the number of samples. White circles represent haplotypes not found in the populations employed in this analysis. 23

Figure 5. Network for haplogroup D. Mutations are identified with reference to the Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999). Bold circles represent founding haplotypes. The size of the circles is proportional to the number of samples. White circles represent haplotypes not found in the populations employed in this analysis. 24
**Introduction**

The Postclassic period in Mesoamerica (AD 900-1500) was characterized by the proliferation of city-states, major population growth, commercialization of the economy, and great volumes of long-distance exchange (Smith and Berdan, 2003). Within this context, the Aztec empire emerged in 1428 as a result of the triple alliance among the city-states of Tenochtitlan, Texcoco, and Tlacopan. Over the next 100 years, the Aztecs conquered a large number of city-states throughout the Basin of Mexico (Berdan and Smith, 2003). The expansion of the Aztec empire has been the subject of much research, but the demographic consequences of this expansion remain unclear. In particular, it is not yet known whether (1) the Aztec occupation of defeated city-states consisted mostly of a social reorganization of groups already present in the Basin of Mexico, or (2) the Aztecs replaced the original inhabitants of the conquered city-states with migrants from other places (González-José et al., 2007). To help evaluate these two possible scenarios and to investigate the genetic impact of the Aztec expansion, I examined ancient DNA (aDNA) from human remains recovered at Xaltocan, a town located 35 km north of Mexico City (Fig. 1) that may have experience a population replacement when it was incorporated into the Aztec empire.

Archaeological evidence indicates that the town of Xaltocan was founded around AD 900 by a group of Otomi-speaking people who migrated into the northern Basin of Mexico from the west (Gibson, 1964; Nichols et al., 2002; Brumfield, 2005a; Rodríguez-Alegría, 2008). At that time, Xaltocan was a relatively small island (approximately 800 x 400 m) rising 5-6 m above the bed of the now drained Lake Xaltocan (Brumfiel, 2005d; Morehart and Eisenberg, 2010).
In the 11th-14th centuries the town grew into an important city-state that collected tribute from neighboring communities (Carrasco-Pizana, 1987; Brumfield, 2005a; Brumfield, 2005c) and served as the capital of the Otomi empire (Gibson, 1964; Brumfield, 2005d; De Lucia, 2010). Xaltocan’s preeminence in the Basin of Mexico has been attributed to the availability of water, existence of abundant lake resources (i.e., fauna, plants, and salt), rich soil, productive agricultural practices, and a fruitful market network with nearby cities (Brumfield 2005c; Hodge and Neff, 2005; Rodríguez-Alegría, 2008; Morehart and Eisenberg, 2010). At its peak, Xaltocan may have had a population of nearly 5,000 individuals (Sanders et al., 1979: 151).

By the mid-13th century, Xaltocan became involved in a war against the neighboring Tepanec kingdom of Cuauhtitlan (Nichols et al., 2002; Morehart and Eisenberg, 2010). This conflict resulted in a lengthy struggle that gradually reduced Xaltocan’s dominions (Códice Chimalpopoca, 1975: 24-25). With the aid of other Tepanec allies, Cuauhtitlan finally vanquished Xaltocan in 1395 (Gibson, 1964; Carrasco-Pizana, 1987; Morehart and Heisenberg, 2010). Colonial documents assert that the original Otomi population fled Xaltocan at this time, leaving the city unpopulated (Códice Chimalpopoca, 1975: 50).

The emerging Aztec empire annexed Xaltocan in 1428, and Tenochtitlan (the capital of the empire) assigned rulers to govern the town (Carrasco-Pizana, 1987; Brumfield, 2005d; Rodríguez-Alegría, 2008). According to colonial documents, one of these rulers repopulated Xaltocan in 1435 with people who “were not the descendants of the original settlers, but came from other places” (Anales de Cuauhtitlan 1939: Par. 628, 1017-21, as cited in Hicks, 1994: 67).

However, archaeological evidence may contradict this account of the population history of Xaltocan after the city lost the war against Cuauhtitlan in 1395. While colonial documents indicate that this defeat and the subsequent Aztec conquest led to a population replacement at
Xaltocan, recent archaeological finds suggest a significant degree of population continuity after 1395 and even following incorporation within the Aztec empire. For instance, radiocarbon analyses indicate that some burials date between 1395 and 1435 (Lisa Overholtzer, personal communication), which calls into question the assertion that Xaltocan remained uninhabited during that period. In addition, during the Aztec period, houses and funerary monuments were built in the same locations as those constructed by the earlier Otomi residents (Miller, 2007). This evidence suggests that the original Otomi population may have remained at Xaltocan and maintained some of its customs after the defeat in 1395.

Using ancient DNA to test population replacement hypotheses

Ancient DNA studies can be used to help investigate hypothesized population replacements like the one that may have occurred at Xaltocan. Even though ancient skeletal remains usually contain scarce and degraded DNA (O’Rourke et al., 2000; Kaestle and Horsburgh, 2002; Pääbo et al., 2004), it is sometimes possible to extract and analyze the remaining genetic material. When DNA is sampled from burial populations that predate and postdate the hypothesized replacement event, comparisons of genetic variation between the two temporal groups can be used to evaluate the replacement hypothesis (Kaestle and Horsburgh, 2002; Smith et al., 2009). If genetic discontinuity is observed between the pre- and post-transition subpopulations, then the genetic data provide support for the hypothesized population replacement. If, however, genetic continuity is observed, then it is unlikely that a population replacement occurred. This approach has provided important insights in archaeological and genetic investigations of other hypothesized population replacements (e.g., Kaestle and Smith, 2001; Lewis Jr et al., 2007; Cabana et al., 2008; Kemp et al., 2009; Raff et al., 2010; Smith et al.,
2009; Snow et al., 2010). For instance, Smith et al. (2009) found a significant change in the frequencies of mitochondrial DNA lineages in the Aleutian Islands about 1,000 years ago, which is consistent with Hrdlicka’s (1945) hypothesis of a population replacement in the area at that time.

Ancient DNA studies have traditionally focused on mitochondrial DNA (mtDNA) because several copies of mtDNA are present in each mitochondrion and, in turn, some cells may contain thousands of mitochondria (Jobling et al., 2004; Raff et al., 2010). Since aDNA studies usually rely on scarce and degraded genetic material, the fact that each cell contains numerous copies of mtDNA substantially increases the chances of amplifying the target sequence (Smith et al., 2009; Raff et al., 2010). The mitochondrial genome, a circular molecule made up of 16,569 base pairs (bp), is maternally inherited and does not undergo recombination, thus constituting a useful tool for studying maternal ancestry (Eshleman et al., 2003). In addition, because mtDNA has a mutation rate about 10% higher than that of the nuclear genome and, within mtDNA, the 1,120-bp control region has an even higher mutation rate (Eshleman et al., 2003; Jobling et al., 2004), the mtDNA control region is especially useful for genetic studies of population history that span relatively short periods of time (Raff et al., 2010).

Native Americans exhibit five founding maternal lineages known as haplogroups A, B, C, D, and X (Schurr, 2004). These lineages are distinguished by specific restriction fragment length polymorphisms (RFLP) and a 9-bp deletion in the mitochondrial genome, as well as by mutations in the mtDNA control region (Schurr et al., 1990; Torroni et al., 1992; Torroni et al., 1993; Brown et al., 1998; Smith et al., 1999). Additional mutations within the control region differentiate sequences from the same haplogroup into different haplotypes, or specific sequences that can be found in one or more individuals (Kemp et al., 2005). Haplogroup and
haplotype data can serve as an important tool for studying genetic relationships among human populations, with similar haplogroup frequencies and/or shared haplotypes indicating recent common ancestry (Kaestle and Horsburgh, 2002; Bolnick and Smith, 2007).

The geographic distribution of the founding Native American haplogroups shows major trends. Broadly speaking, the frequency of haplogroup A decreases and the frequencies of haplogroups C and D increase as one moves from north to south in the Americas. On the other hand, haplogroup B shows no clear pattern other than being nearly absent in northern North America, and haplogroup X is present almost exclusively in North America (Schurr, 2004). In Mesoamerica, most populations are characterized by high frequencies of haplogroup A, lower frequencies of haplogroups B, C, and D, and the absence of haplogroup X (Torroni et al., 1994; Kemp et al., 2010). Ancient DNA studies of Late Classic-Postclassic Maya (AD 600-1500) samples from Xcaret (González-Oliver et al., 2001) and post-Classic Aztec (AD 900-1500) samples from Tlatelolco (Kemp et al., 2005) indicate the relative antiquity of this overall pattern in the region (Raff et al., 2011).

In this study, I used ancient mtDNA to evaluate the hypothesis that a population replacement took place at Xaltocan when it was abandoned after its defeat in 1395 and repopulated under Aztec rule in 1435. I report mtDNA data from human remains belonging to two temporal subpopulations (before and after 1395), and show that these subpopulations exhibited significantly different patterns of genetic variation. This study suggests that different matrilineal groups were present before and after 1395 at Xaltocan, supporting the population replacement hypothesis.
Materials and Methods

*Skeletal samples*

Skeletal remains (teeth or bone) were sampled from 21 individuals recovered during archaeological research in Xaltocan (Table 1). Excavations at this site, which covers about 68 hectares and lies almost entirely under the modern town of Xaltocan, began in 1987 (Brumfiel, 2005d; Brumfiel, 2005a). All samples that I analyzed in this study were recovered from Operation Este, a group of excavation units placed on the eastern edge of the site (Lisa Overholtzer, personal communication). The samples come from males and females who were buried outside two neighboring houses in a mound containing stratified domestic deposits (Lisa Overholtzer, personal communication) and whose ages at death span a broad range (from newborns to elders). All individuals are thought to have been members of the nuclear families that inhabited the households, but it is unlikely that they represent all of the residents (Lisa Overholtzer, personal communication).

Of the 21 individuals, 11 have been sampled for radiocarbon dating and their dates range from about AD 1345 to AD 1640 (Table 1). Approximate dates for the remaining samples were determined based on stratigraphy and the style of pottery placed in the graves (Lisa Overholtzer, personal communication). For the purpose of this study, the sampled individuals were divided into two temporal subpopulations: before 1395 (n=5) and after 1395 (n=16).
DNA extraction and purification

I extracted aDNA following Bolnick et al.’s (In press) non-destructive extraction protocol. In order to remove surface contamination, each sample was submerged in 6% sodium hypochlorite (full strength bleach) for 15 minutes (Kemp and Smith, 2005), rinsed twice with DNA-free HPCL-grade water, irradiated with 254 nm ultraviolet light (UV) for ten minutes per side, and left to dry at room temperature. Each sample was then transferred to a 15 mL polypropylene tube and incubated at room temperature in 10 mL of extraction solution (0.45 M EDTA and 0.25 mg mL\(^{-1}\) proteinase K, pH 8.0) while rocking gently overnight. Following incubation, the extraction solution was poured into a separate tube without disturbing the sample, and the tooth or bone was then rinsed three times with DNA-free HPLC-grade water and air-dried at room temperature. DNA left in the extraction solution was purified using the silica/guanidium thiocyanate method published by Rohland and Hofreiter (2007) and then eluted in 70 µL of DNA-free HPLC-grade water.

PCR amplification and analysis

I performed two sets of mtDNA analyses. First, I screened the samples for the mutations that define the Native American mitochondrial haplogroups A, B, C, and D (Schurr et al., 1990; Torroni et al., 1993). I did not screen the samples for the RFLP that defines haplogroup X, as this mitochondrial haplogroup is rare in Mesoamerican populations (Smith et al., 1999; González-Oliver et al., 2001; Schurr, 2004; Kemp et al., 2005; Kemp et al., 2010) and all samples were identified as belonging to haplogroups A-D. PCR reactions were performed following Bolnick et al. (In press). I visualized three µL of each PCR product with ethidium bromide on 6%
polyacrylamide gels to confirm amplification and conducted the screening of haplogroups A-D according to Bolnick and Smith (2007).

In order to confirm haplogroup assignment and determine the haplotype of each sample, nucleotide positions (np) 16,011-16,382 of the first hypervariable region (HVRI) of the mitochondrial genome were sequenced. For this purpose, four overlapping fragments were amplified employing the primers and protocol published in Bolnick et al. (In press). For samples exhibiting a cytosine (C) at np 16189, I also amplified one additional fragment (np 16,001-16,181) using the primers published in Kemp et al. (2009). PCR products were purified using either the QIAquick PCR purification kit (Qiagen) or the combined use of magnetic beads (Sera-Mag Magnetic Carboxylate-Modified Microparticles) and a Biomek FX robot. I submitted the fragments to the DNA Sequencing Facility at the University of Texas at Austin for direct sequencing, with sequencing performed in both directions.

**Contamination controls**

All of the pre-PCR methods described above were conducted in the ancient DNA laboratory of the University of Texas at Austin. Ancient DNA studies are methodologically challenging because aDNA is usually very degraded, present in small amounts, and may be obscured by contamination from exogenous sources (O’Rourke et al., 2000; Kaestle and Horsburgh, 2002; Pääbo et al., 2004). To prevent and detect contamination, precautions were followed at all stages of the analysis. The ancient DNA lab of the University of Texas at Austin is a restricted-access clean room that is dedicated to pre-PCR aDNA research and equipped with dedicated equipment, overhead UV lights, positive air pressure, and HEPA-filtered ventilation. The post-PCR laboratory was located on a separate floor of the building, and personnel
movement between facilities was unidirectional (from pre-PCR aDNA lab to post-PCR lab) each day. In addition, other precautions included: wearing disposable coveralls, hair covers, facemasks, sleeve covers, shoe covers, and gloves at all times in the aDNA lab, regularly decontaminating workspaces and equipment with bleach, bleaching the entire aDNA lab weekly with 3% sodium hypochlorite (50% v/v bleach), UV irradiating the entire lab for two hours after each use, irradiating both reagents and tubes with 254 nm UV light prior to use when possible, using reagents that were certified DNA-free and/or molecular grade whenever possible, performing DNA extractions and PCR set-up in a laminar flow hood, using aerosol resistant filter tips in all pre-PCR analyses, treating samples with both bleach and UV radiation (254 nm) to eliminate any surface contamination, and including negative (blank) controls at all stages of the extraction and amplification process to detect any contamination that did occur. Finally, at least two independent extractions (different bones and/or teeth) were performed for each individual, and results were confirmed through multiple amplifications.

**Statistical analyses**

Statistical analyses were performed to characterize the patterns of genetic variation at Xaltocan and test the population replacement hypothesis. All analyses of the haplotype data excluded np 16,183 (because mutations at this site are strictly dependent on the presence of a C at np 16,189; Pfeiffer et al., 1999) and insertions in poly-C stretches (due to uncertainty in the exact position of such mutations). In this study, results were considered statistically significant when $P < 0.05$. 
For the two Xaltocan subpopulations and for Xaltocan as a whole, I calculated haplogroup diversity \( (h) \), haplotype diversity \( (h) \), and nucleotide diversity \( (\pi) \) using the program Arlequin 3.5.1.2 (Excoffier and Lischer, 2010). Haplogroup diversity and haplotype diversity are defined as the probability that two randomly chosen haplogroups or haplotypes, respectively, are different in the population sample. Nucleotide diversity is the probability that two randomly chosen homologous nucleotide positions are different (Excoffier and Lischer, 2010). To calculate nucleotide diversity, I used the Kimura 2-parameters model, which allows multiple substitutions at each nucleotide position as well as different substitution rates between transitions and transversions, with the transition-transversion ratio estimated from the data (Excoffier and Lischer, 2010). Haplogroup, haplotype, and nucleotide diversities for the two subpopulations were compared using two-sample \( t \)-tests (Zar, 1999).

Second, I used Arlequin to perform two exact tests, with one based on the haplogroup data and one based on the haplotype data. These tests evaluated the null hypothesis that the haplogroup or haplotype frequencies were identical in the two Xaltocan subpopulations (indicating population continuity).

Finally, to help test the population replacement hypothesis, I compared the haplotypes in the two Xaltocan subpopulations and in Xaltocan as a whole with those in four previously studied extant populations from Mesoamerica (Table 2). The Mixtec and the Zapotec, like the Otomi, speak languages belonging to the Otomanguean language family. The Nahua-Atocpan and Nahua-Cuetzalan, on the other hand, speak Uto-Aztecan languages. Genetic distances were calculated between all population pairs to determine whether (1) the pre-1395 Xaltocan population was most similar to Otomanguean-speaking populations, and (2) the post-1395 Xaltocan population was most similar to Uto-Aztecan-speaking populations, as the population
replacement hypothesis would predict. Genetic distances were calculated following the method published by Qamar et al. (2002). Pairwise $F_{ST}$ values were computed separately for each haplogroup in Arlequin using the Kimura 2-parameters model, and then weighted to reflect the frequencies of that haplogroup in the two populations being compared. Weighted means of the within-haplogroup $F_{ST}$ values are reported.

In order to determine the phylogenetic relationships among haplotypes within each haplogroup, I created median-joining networks (Bandelt et al., 1999) with the software Network 4.6.0.0 (www.fluxus-engineering.com). Mutations were identified with reference to the Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999). Following Perego et al. (2010), I considered founding haplotypes as haplotypes that were already present in the founding Paleo-Indian population(s) who entered the Americas about 16,000 years ago. Haplotypes exhibiting additional mutations when compared to the founding haplotypes were considered derived. Because some of the sequences reported by Kemp et al. (2010) were missing nps 16,011-16,042, only nps 16043-16382 were used when creating the networks. Preliminary network constructions showed high levels of reticulation. I resolved this problem by following Kemp et al. (2009), who applied a default weight of 10 to all sites and then down-weighted some polymorphic sites depending on their relative mutation rates as estimated by Meyer et al. (2003). Hence, in this study sites with more than 4-fold higher rates (16051, 16126, 16129, 16189, 16223, 16278, and 16311) were down-weighted to 4, sites with 3-fold higher rates (16093, 16148, 16172, 16187, 16192, and 16293) were down-weighted to 5, and sites with 2-fold higher rates (16111, 16241, 16247, 16256, 16274, 16292, and 16300) were down-weighted to 6.
Results

I successfully recovered and analyzed DNA from 20 of the 21 individuals sampled (a success rate of 95.2%). This success rate is particularly high compared to other studies of ancient mtDNA from Native American burial populations. For instance, Kemp et al. (2009) had a success rate of 33% in their study of an ancient population in Conchopata, Peru. The haplogroup and haplotype data for each individual are given in Table 1, and haplogroup frequencies for all populations included in this study are given in Table 2. Of the 20 Xaltocan samples that yielded analyzable DNA, 10 belong to haplogroup A (50%), 4 to haplogroup B (20%), 1 to haplogroup C (5%), and 5 to haplogroup D (25%). The observed frequencies are consistent with previous studies of Central American populations, which generally exhibit a predominance of haplogroup A, lower frequencies of haplogroups B, C, and D, and an absence of haplogroup X (Torroni et al., 1994; González-Oliver et al., 2001; Kemp et al., 2005; Kemp et al., 2010). Moreover, the Xaltocan samples share 1, 2, 1, and 2 haplotypes with Zapotec, Mixtec, Nahua-Cuetzalan, and Nahua-Atocpan populations, respectively (Kemp et al., 2010). Overall, these results indicate that the data reported here make phylogeographic sense.

When the Xaltocan samples are split into two temporal subpopulations (before 1395 and after 1395), the frequencies of haplogroups A and C increase over time (from 20% to 60% and from 0% to ∼7%, respectively) whereas the frequency of haplogroup D decreases over time (from 60% to ∼13%) and the frequency of haplogroup B (20%) does not change. The exact test based on the haplogroup data shows that haplogroup frequencies do not differ significantly between the two temporal subpopulations ($P = 0.174$), probably because of small sample sizes. The exact test based on the haplotype data, on the other hand, indicates that haplotype
frequencies are significantly different between the two subpopulations ($P = 0.007$). This result suggests that the null hypothesis of population continuity should be rejected.

Genetic diversity estimates are reported in Table 3. The pre-1395 subpopulation shows slightly greater haplogroup diversity than the post-1395 subpopulation, whereas the post-1395 subpopulation exhibits somewhat greater haplotype and nucleotide diversity than the pre-1395 subpopulation. However, $t$-tests indicate that none of these differences are statistically significant (haplogroup diversity $P = 0.77$; haplotype diversity $P = 0.87$; nucleotide diversity $P = 0.66$).

Genetic distances ($F_{ST}$) between populations are reported in Table 4. As the population replacement hypothesis would predict, the Xaltocan post-1395 population is more similar to the Nahua-Atocpan ($F_{ST} = 0.209$) and Nahua-Cuetzalan ($F_{ST} = 0.201$) populations than to the Mixtec ($F_{ST} = 0.215$) or Zapotec ($F_{ST} = 0.228$) populations. However, the pre-1395 Xaltocan population is more similar to the Nahua-Atocpan ($F_{ST} = 0.076$) population than to the Mixtec ($F_{ST} = 0.177$) or Zapotec ($F_{ST} = 0.181$) populations, in conflict with the hypothesized linguistic affiliations. The highest $F_{ST}$ (0.571) corresponds to the comparison between Xaltocan pre-1395 and Xaltocan post-1395 subpopulations, indicating that these subpopulations are more different from each other than either is from any other population included in this analysis. Interestingly, when taken as whole, Xaltocan is more similar to the Nahua-Atocpan ($F_{ST} = 0.165$) and Nahua-Cuetzalan ($F_{ST} = 0.157$) populations than to the Mixtec ($F_{ST} = 0.184$) or Zapotec ($F_{ST} = 0.191$) populations.

Haplotype median-joining networks show the molecular relationships among the haplotypes within each haplogroup. In haplogroup A (Fig. 2), the pre-1395 sample exhibits a unique derived haplotype, whereas the post-1395 Xaltocan samples share two derived haplotypes with modern Nahua individuals. In this haplogroup, the pre-1395 Xaltocan sample is in a clade with Mixtec speakers, and the post-1395 Xaltocan samples are present in clades with Mixtec,
Zapotec, Nahua-Atocpan, and Nahua-Cuetzalan speakers. In haplogroup B (Fig. 3), the pre-1395 Xaltocan samples share a derived haplotype with two modern Zapotes and one modern Nahua-Atocpan individual, while the post-1395 Xaltocan samples share a derived haplotype with a modern Mixtec individual and are present in clades with both Mixtec and Nahua-Atocpan speakers. In haplogroup C (Fig. 4), the one post-1395 Xaltocan sample exhibits a unique derived haplotype. In haplogroup D (Fig. 5), the three pre-1395 Xaltocan samples exhibit a derived haplotype that was not found in the Otomanguean or Aztec samples studied by Kemp et al. (2010), and the post-1395 Xaltocan samples belong to the founding haplogroup D haplotype. Significantly, the two temporal subpopulations from Xaltocan share no haplotypes, and the pre-1395 haplotypes and post-1395 haplotypes are always found in different clades, indicating that there are no close relationships between the haplotypes present at Xaltocan before and after 1395.
Discussion

This study provides the first genetic test of the hypothesis of population replacement at Xaltocan between 1395 (defeat in the war against Cuauhtitlan) and 1435 (repopulation under Aztec rule). This investigation also represents the first research into how the expansion of the Aztec empire affected the genetic makeup of Mesoamerican populations. The results reported here are thought to be authentic because (1) the haplogroup and haplotype data for each individual were confirmed through multiple amplifications and multiple independent extracts, (2) stringent procedures were followed during all the stages of the analysis to prevent and detect contamination, (3) haplogroup frequencies as well as the presence of certain haplotypes accord with those previously reported in populations from the same region, and (4) the haplotypes of all Xaltocan individuals differ from that of the author as well as those of other researchers working in the aDNA laboratory.

Unlike previous aDNA studies testing population replacement hypotheses (e.g., Kaestle and Smith, 2001; Smith et al., 2009), which relied on samples spanning thousands of years, this study uses samples that are narrowly distributed in time. Because each Xaltocan subpopulation contains samples that span only a few generations, the subpopulations defined here should accurately represent the genetic makeup of the sampled households during the two periods of interest.

The results of this study support the hypothesis that matriline s at Xaltocan underwent a population replacement after 1395 for three reasons. First, the two subpopulations have significantly different haplotype frequencies and are unlikely to have been sampled from the same population. Second, each subpopulation exhibits distinct haplotypes that are present in different clades. Because the samples included in the temporal subpopulations span such a short
period of time from an evolutionary perspective (about 300 years from the earliest to the most recent radiocarbon dates), different recent ancestries are much more likely to account for the observed differences than the occurrence of new mutations. It is also unlikely that genetic drift could account for the observed differences between the temporal subpopulations. Even though genetic drift may significantly change haplotype frequencies in a relatively short period of time (especially if population size is small), this evolutionary force alone cannot be responsible for the appearance of new haplotypes in a population, as was observed in this study. Third, the highest $F_{ST}$ was found when comparing the pre-1395 and the post-1395 Xaltocan subpopulations. This result suggests a substantial genetic distance between the two Xaltocan subpopulations, as the population replacement hypothesis would predict.

Genetic comparisons between the Xaltocan subpopulations and the modern Mesoamerican populations are only partly in agreement with the expectations of the population replacement hypothesis. In several cases, the observed genetic distances are consistent with the expectations of the hypothesis based on linguistic affiliations because (1) the pre-1395 Xaltocan population is more similar to the Otomanguean populations than to the Nahua-Cuetzalan (Uto-Aztecan language family) population and (2) the post-1395 Xaltocan subpopulation is more similar to Uto-Aztecan populations than to Otomanguean populations. In spite of their linguistic closeness, however, the pre-1395 Xaltocan subpopulation is not more similar to the Otomanguean populations than to the Nahua-Atocpan (Uto-Aztecan language family) population. Networks also agree only in part with the expectations of the population replacement hypothesis. On the one hand, some post-1395 Xaltocan samples belonging to haplogroup A share a derived haplotype with one modern Uto-Aztecan speaker, as the population replacement hypothesis would predict. In haplogroup B, however, the expectations of the hypothesis are not
met because one pre-1395 Xaltocan sample shares a derived haplotype with one Uto-Aztecan speaker, and two post-1395 Xaltocan samples share a derived haplotype with one Otomanguean individual. In haplogroup D, the post-1395 Xaltocan samples and one Otomanguean individual also exhibit the same sequence, but this result is not in disagreement with the expectations of the population replacement hypothesis because such a sequence represents a founding haplotype (which does not necessarily indicate close ancestral relationships). Importantly, networks show that there are not clear phylogenetic patterns distinguishing modern Otomanguean from modern Uto-Aztecan populations. Thus, definitive differences between ancient Otomanguean and ancient Uto-Aztecan populations are not expected either.

While this study suggests that matrilines at Xaltocan underwent a population replacement after 1395, the data presented here do not provide a definitive answer to the question of population replacement at Xaltocan for three reasons. First, these samples were recovered exclusively from two neighboring households and a single group of archaeological units (Operation Este), so the haplogroup and haplotype frequencies reported here might not be representative of the entire population at Xaltocan. The most common haplotypes, for example, might reflect the presence of maternally related individuals in these households rather than a high frequency of those haplotypes in the Xaltocan population as a whole. It is therefore possible that the replacement of matrilines observed in this study was a localized event rather than a more widespread change affecting the entire town. The analysis of samples from other archaeological units at Xaltocan in the near future will help to clarify whether the replacement event observed in this study was a more widespread phenomenon at Xaltocan. Second, this study has a relatively small sample size (especially for the pre-1395 Xaltocan subpopulation). For this reason, the likelihood that the haplogroup and haplotype frequencies reported here are representative of the
entire population at Xaltocan is further decreased, and the statistical power of the analyses may be relatively low. The addition of samples from other archaeological units in upcoming investigations may provide a more representative estimate of Xaltocan’s genetic makeup in ancient times and will increase statistical power. Finally, it is important to note that mtDNA provides information exclusively about the matrilineal history at Xaltocan. While the results from this study are suggestive, additional data on Y-chromosome and/or autosomal markers are needed to obtain a more complete picture of Xaltocan’s population history after 1395 and following its incorporation into the Aztec empire in 1428.
Figure 1. Map of the Basin of Mexico in pre-Hispanic times. The dotted areas represent lakes (from Rodríguez-Alegria, 2010).
Figure 2. Network for haplogroup A. Mutations are identified with reference to the Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999). Bold circles represent founding haplotypes. The size of the circles is proportional to the number of samples. White circles represent haplotypes not found in the populations employed in this analysis.
Figure 3. Network for haplogroup B. Mutations are identified with reference to the Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999). Bold circles represent founding haplotypes. The size of the circles is proportional to the number of samples. White circles represent haplotypes not found in the populations employed in this analysis.
Figure 4. Network for haplogroup C. Mutations are identified with reference to the Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999). Bold circles represent founding haplotypes. The size of the circles is proportional to the number of samples. White circles represent haplotypes not found in the populations employed in this analysis.
Figure 5. Network for haplogroup D. Mutations are identified with reference to the Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999). Bold circles represent founding haplotypes. The size of the circles is proportional to the number of samples. White circles represent haplotypes not found in the populations employed in this analysis.
Table 1. Xaltocan samples analyzed in this study.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Age at death$^1$</th>
<th>Radiocarbon dates (AD)$^2$</th>
<th>Haplogroup</th>
<th>Haplotype$^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-1395</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E14.3</td>
<td>3 months</td>
<td>1365 ± 65</td>
<td>D</td>
<td>223T, 292T, 325C, 362C</td>
</tr>
<tr>
<td>E14.4</td>
<td>7 months</td>
<td>NA</td>
<td>D</td>
<td>223T, 292T, 325C, 362C</td>
</tr>
<tr>
<td>E14.5</td>
<td>10 months – 1 year</td>
<td>NA</td>
<td>D</td>
<td>223T, 292T, 325C, 362C</td>
</tr>
<tr>
<td>E14.7</td>
<td>9 months – 1 year</td>
<td>1345 ± 55</td>
<td>B</td>
<td>182C, 183C, 189C, 217C, 295T</td>
</tr>
<tr>
<td>E25.2</td>
<td>9 months – 1 year</td>
<td>1365 ± 65</td>
<td>A</td>
<td>111T, 223T, 290T, 319A, 335G</td>
</tr>
<tr>
<td><strong>Post-1395</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E6.1</td>
<td>12 years</td>
<td>1435 ± 35</td>
<td>B</td>
<td>182C, 183C, 189C, 217C, 357C</td>
</tr>
<tr>
<td>E7.1</td>
<td>5 years</td>
<td>1390 ± 60</td>
<td>A</td>
<td>93C, 111T, 136C, 223T, 290T, 311C, 319A, 362C</td>
</tr>
<tr>
<td>E8.2</td>
<td>More than 50 years</td>
<td>NA</td>
<td>A</td>
<td>93C, 111T, 136C, 223T, 290T, 311C, 319A, 362C</td>
</tr>
<tr>
<td>E8.4</td>
<td>1 – 3 months</td>
<td>1375 ± 65</td>
<td>A</td>
<td>93C, 111T, 136C, 223T, 290T, 311C, 319A, 362C</td>
</tr>
<tr>
<td>E10.1</td>
<td>20 – 35 years</td>
<td>NA</td>
<td>A</td>
<td>172C, 223T, 290T, 319A, 362C</td>
</tr>
<tr>
<td>E10.2</td>
<td>More than 50 years</td>
<td>1640 ± 310</td>
<td>A</td>
<td>93C, 111T, 136C, 223T, 290T, 311C, 319A, 362C</td>
</tr>
<tr>
<td>E14.1</td>
<td>8 – 9 years</td>
<td>NA</td>
<td>B</td>
<td>183C, 189C, 217C, 258C, 260.1T$^4$</td>
</tr>
<tr>
<td>E14.2</td>
<td>1 month</td>
<td>1525 ± 105</td>
<td>A</td>
<td>111T, 192T, 223T, 290T, 319A, 362C</td>
</tr>
<tr>
<td>E14.6</td>
<td>35 – 50 years</td>
<td>1520 ± 100</td>
<td>B</td>
<td>182C, 183C, 189C, 217C, 357C</td>
</tr>
<tr>
<td>E30.A</td>
<td>NA</td>
<td>NA</td>
<td>D</td>
<td>223T, 325C, 362C</td>
</tr>
<tr>
<td>E30.3</td>
<td>30 – 35 years</td>
<td>NA</td>
<td>C</td>
<td>172C, 223T, 298C, 325C, 327T</td>
</tr>
</tbody>
</table>

---

$^1$ Data from Lisa Overholtzer (personal communication). NA, not available.

$^2$ Data from Lisa Overholtzer (personal communication). NA, not available. When classifying the samples for which radiocarbon dates are available as pre- or post-1395, their archaeological context was also taken into account.

$^3$ Sequences from nps 16,011-16,382. Mutations are identified with reference to the Cambridge Reference Sequence (Anderson et al., 1981; Andrews et al., 1999).

$^4$ Insertion between nucleotide positions 16,260 and 16,261.
<table>
<thead>
<tr>
<th>Population</th>
<th>Language Family&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Haplogroup N</th>
<th>Haplotype N</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-1395 Xaltocan</td>
<td>Otomanguean</td>
<td>5</td>
<td>5</td>
<td>0.200</td>
<td>0.200</td>
<td>0</td>
<td>0.600</td>
<td>This study</td>
</tr>
<tr>
<td>Post-1395 Xaltocan</td>
<td>Uto-Aztecan</td>
<td>15</td>
<td>15</td>
<td>0.600</td>
<td>0.200</td>
<td>0.067</td>
<td>0.133</td>
<td>This study</td>
</tr>
<tr>
<td>Xaltocan (all)</td>
<td></td>
<td>20</td>
<td>20</td>
<td>0.500</td>
<td>0.200</td>
<td>0.050</td>
<td>0.250</td>
<td>This study</td>
</tr>
<tr>
<td>Nahua-Cuetzalan</td>
<td>Uto-Aztecan</td>
<td>46</td>
<td>29</td>
<td>0.630</td>
<td>0.196</td>
<td>0.152</td>
<td>0.022</td>
<td>Kemp et al., 2010; Malhi et al., 2003</td>
</tr>
<tr>
<td>Nahua-Atocpan</td>
<td>Uto-Aztecan</td>
<td>50</td>
<td>44</td>
<td>0.380</td>
<td>0.400</td>
<td>0.180</td>
<td>0.040</td>
<td>Kemp et al., 2010</td>
</tr>
<tr>
<td>Zapotec</td>
<td>Otomanguean</td>
<td>85</td>
<td>72</td>
<td>0.424</td>
<td>0.224</td>
<td>0.294</td>
<td>0.058</td>
<td>Kemp et al., 2010</td>
</tr>
<tr>
<td>Mixtec</td>
<td>Otomanguean</td>
<td>67</td>
<td>65</td>
<td>0.672</td>
<td>0.209</td>
<td>0.074</td>
<td>0.045</td>
<td>Kemp et al., 2010</td>
</tr>
</tbody>
</table>

<sup>1</sup> The affiliations for pre-1395 Xaltocan and post-1395 Xaltocan are hypothesized.
Table 3. Genetic diversity estimates (mean ± standard deviation).

<table>
<thead>
<tr>
<th>Haplogroup</th>
<th>Haplogroup $h$</th>
<th>Haplotype $h$</th>
<th>Nucleotide diversity ($\pi$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-1395 Xaltocan</td>
<td>0.700 ± 0.218</td>
<td>0.700 ± 0.218</td>
<td>0.015 ± 0.010</td>
</tr>
<tr>
<td>Post-1395 Xaltocan</td>
<td>0.620 ± 0.120</td>
<td>0.838 ± 0.085</td>
<td>0.017 ± 0.001</td>
</tr>
<tr>
<td>Xaltocan (all)</td>
<td>0.679 ± 0.074</td>
<td>0.895 ± 0.052</td>
<td>0.017 ± 0.001</td>
</tr>
</tbody>
</table>
Table 4. Weighted means of within-haplogroup Fst values.

<table>
<thead>
<tr>
<th></th>
<th>Pre-1395</th>
<th>Mixtec</th>
<th>Zapotec</th>
<th>Post-1395</th>
<th>Nahua-Atocpan</th>
<th>Nahua-Cuetzalan</th>
<th>Xaltocan (all)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixtec</td>
<td>0.177</td>
<td>0.181</td>
<td>0.038</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zapotec</td>
<td>0.181</td>
<td>0.571</td>
<td>0.215</td>
<td>0.228</td>
<td>0.076</td>
<td>0.063</td>
<td>0.190</td>
</tr>
<tr>
<td>Post-1395 Xaltocan</td>
<td>0.571</td>
<td>0.215</td>
<td>0.228</td>
<td></td>
<td>0.076</td>
<td>0.063</td>
<td>0.190</td>
</tr>
<tr>
<td>Nahua-Atocpan</td>
<td>0.076</td>
<td>0.063</td>
<td>0.018</td>
<td>0.209</td>
<td>0.043</td>
<td>0.052</td>
<td>0.190</td>
</tr>
<tr>
<td>Nahua-Cuetzalan</td>
<td>0.190</td>
<td>0.043</td>
<td>0.052</td>
<td>0.201</td>
<td>0.043</td>
<td>0.052</td>
<td>0.190</td>
</tr>
<tr>
<td>Xaltocan (all)</td>
<td>0.121</td>
<td>0.184</td>
<td>0.191</td>
<td>-0.091</td>
<td>0.165</td>
<td>0.157</td>
<td></td>
</tr>
</tbody>
</table>
Bibliography


Códice Chimalpopoca: anales de Cuauhtitlan y leyenda de los soles. 1975. Mexico City, Mexico: Instituto de Investigaciones Históricas de la Universidad Nacional Autónoma de México.


