Ancient DNA and Genetic Relations at a 4000-year-old Archaeological Site (CA-CCO-548) in the California Delta

Naomi L. Martisius

Abstract
Ancient Northern California shows a distinctive burial pattern between 4500 and 2500 years ago, with coastal peoples burying their dead in flexed position and Sacramento Valley peoples burying their dead in extended position with the head pointing west. It has been hypothesized that Penutian speakers brought with them into the Sacramento Valley not only a distinctive burial style but also a unique genetic composition. Previous research with ancient mitochondrial DNA (aDNA) has suggested a common ancestry among coastal populations, while inland groups show a separate common ancestry. Recent excavations at a site near Brentwood, California, between the coast and Sacramento Valley, revealed over 500 burials from 3000-4000 years ago, with a mix of flexed and extended positions. Ten of each were tested for aDNA to determine haplogroups; these were also divided by gender. This aDNA analysis is an important strategy in increasing our understanding of ancient Californians and migrations.

Introduction
In many societies around the world, burial customs provide important information about individuals while they were alive. Such information may indicate, for example, the gender or the status of the individual within society. As well, burial customs vary greatly among different cultures.

Archaeologists have often focused on burial style to demarcate what they interpret as ancient cultures, and attempt to decode such styles to draw inferences about the status, social position, and gender of individuals, as well as their cultural affiliation. As early as the 1930s archaeologists working in the interior of Northern California, in the Central Valley, noted an important difference between deeply buried skeletons and those in more shallow burial settings. In particular, the deeper burials were commonly arranged in an extended position with the head pointed west, while shallow burials were flexed and oriented in a range of directions (Lillard et al. 1939). The earlier complex, with extended position, gradually came to be known as belonging to the “Windmiller” culture and was shown to date between 4500 and 2500 years ago, while the later complex came to be known as belonging to the “Berkeley” culture (Beardsley 1954).

Later research showed that the earlier Windmiller burial style was not present in other locations in Northern California. For example, Gerow (1968) showed that contemporaneous cultures in the San Francisco Bay Area did not bury their dead in the extended position with the head pointing west. There, burials were generally flexed and in a range of orientations.

Subsequent research has interpreted the changing burial styles in the interior as indicating the replacement of the older Windmiller culture with later Berkeley cultures. In particular, anthropologists hypothesize that Penutian speakers migrated into the Sacramento Valley from a location somewhere to the north and west, bringing not only a distinctive burial style but also a unique genetic composition.

By contrast, populations in the Bay area remained in place. Support for these notions comes from previous mitochondrial DNA (mtDNA) research, which suggests common ancestry among coastal populations, with inland groups showing separate ancestry.

This research seeks to test these ideas using a recently excavated site located between the Central Valley and the Bay area. The site, CA-CCO-548, or the Marsh Creek site, contains a mixture of both Windmiller-style and Berkeley-style burials. Figure 1 shows both burial styles as excavated at the Marsh Creek site. Radiocarbon dating shows that these interment styles are contemporaneous at the site. This paper tests the hypothesis that the site contains a mixture of people with Central Valley (i.e., Windmiller) and Bay area (i.e., Berkeley) affiliations. I compare mtDNA extracted from both extended and flexed burials and examine whether the data better fit a Central Valley-like or Bay area-like distribution of mitochondrial haplogroups.

A second hypothesis concerns post-marital residence patterns. Patrilocality should cause females to be more genetically variable than males. Ethnographically, tribes in this area were patrilocal, but the antiquity of patrilocality is unknown. If we assume that historic tribes in the region were the same groups that replaced Windmiller populations, and that post-marital residence patterns are conservative over time, the presence of a matrilocal pattern in the ancient population may support the population replacement hypothesis. Figure 2 shows the location of the Marsh Creek site, and other regional sites to which it will be compared in this study.
Figure 1: Burial 78 (extended; haplogroup C) and Burial 79 (flexed). Photo modified from Randy Wiberg, 2010.
Figure 2: Map of study area showing location of CCO-548 and other archaeological sites.

Background

Archaeologists, anthropologists, and ethnographers, among others, have collected vast amounts of data on the prehistory of California, and have shown that ancient California had complex societies with multiethnic groups cooperating together. Archaeological evidence shows that California was occupied as early as 10-13,000 years ago (Connolly et al 1995; Erlandson 2002; Johnson et al 2000; Rick and Erlandson 2000; Rick et al 2001); this date suggests a coastal migration into the area, since no other sites with comparable dates have been found inland. Some DNA evidence is also congruent with this probability; similar genetic frequencies among coastal communities extending along a great portion of the California coast lend support to an early coastal migration (Eshleman 2004). Sites all over California show times of bountiful resources with periods of resource depressions in between. Californians left no ecosystem unaltered, and used a variety of hunting and gathering techniques depending on the resources available. Broad trade-networks were established between coastal and inland peoples with intermarriage and gene transfer likely as well. The diversity of California’s ecology necessitated different types of settlements across the landscape, as did California’s long period of habitation along with the variety of population density and technological innovations (Arnold 2004).

It has been long suggested that population movements are correlated with the spread of languages. It may be too difficult to determine what languages the earliest inhabitants of the Americas spoke, but precursors to known languages in
California may have arisen as early as 6000 to 4000 BC during a time of global warming (see Antevs 1952). Recognizable language groups – Hokan, Penutian, Yukian, and Uto-Aztecan – have their roots in the time period directly following this period, from 4000 to 2000 BC. Michael Moratto (1984) suggests that Hokan speakers were the predominant population in California during this period, while in the ensuing period, from 2000 BC to 1 AD, Penutian and Uto-Aztecan language speakers displaced many Hokan speakers. Moratto based his hypothesis on the presence of Windmiller sites in the Sacramento Valley beginning about 2000 BC. According to Moratto’s hypothesis, these Windmiller peoples should be the first Penutian speakers.

For a long while, anthropologists believed that Windmiller sites were not villages but mortuary mounds instead. C.W. Meighan worked on SJo-68 as a young student and offered several lines of evidence to support this conclusion (Meighan 1987). He noticed that burials were very close together during different time periods, and believed this was indicative of a cemetery because graves would be disturbed if they were directly inside a village. The lack of domestic types of artifacts also supports that this site was not a village. He also examined the physical nature of the site and made comparisons with other sites of the same age as the Windmiller Culture to support his conclusions that SJo-68 was exclusively a burial mound (Meighan 1987).

Recent excavations at the Marsh Creek site (CA-CCO-548) near Brentwood, California, between the coast and the Sacramento Valley revealed the presence of a habitation village and numerous artifacts, as well as over 500 burials. These burials exhibit a mix of flexed and extended positions, and their dates range from 3000 to 4000 years ago, contemporaneous with SJO-68 and other Windmiller sites. Based on these observations, Nathan Stevens and colleagues (2009) challenge the long-held understanding that Windmiller burial mounds are not ancient villages. They find evidence to determine cultural practices from the Marsh Creek site, and show a difference between the Early Horizon cultures and the Early Holocene cultures. There seems to have been a peak in long-distance trading of eastern obsidian with the Windmiller culture, and evidence shows a distinct, specialized diet of small mammals and small seeds that is less diverse than the diet of earlier cultures. Bartelink and colleagues (2010) employed stable isotope analysis to show that the inhabitants of Marsh Creek have carbon isotope ratios consistent with a mostly terrestrial diet, and nitrogen isotope values consistent with consumption of a heavy vegetarian component of C3 plants. Previous research by Bartelink (2006) also shows that the isotope values at other sites from the Central Valley that are classified as Windmiller are consistent with those found at CCO-548. On the other hand, isotopic data from a contemporaneous site in the Bay area, CAALA-307, the West Berkeley Shellmound, show a diet focused on high trophic level marine foods that is distinct from diets found further to the east (Bartelink 2009). In any case, the Marsh Creek site clearly shows distinct cultural practices, which would exclude it from being labeled as a burial mound.

Both ancient and modern mtDNA studies have been useful in understanding the population movements of ancient Californians. mtDNA is a small circular genome that is distinct from the nuclear genome. It is found inside cells, but outside the nucleus in the mitochondria. The high copy number of mtDNA allows for more readily obtainable samples compared with nuclear DNA. This is very
important, since DNA degrades over time. Secondly, mtDNA is only inherited through the mother, so maternal ancestry can be traced far back in time. Finally, mtDNA is not highly conserved, allowing mutations to accumulate fairly rapidly and leading to distinct lineages that are easily traceable.

Native Americans were among the first populations to have their mtDNA studied; consequently, there is a large amount of information on mtDNA haplogroups from modern tribes as well as recovered ancient samples. Within the Americas, there are five mtDNA haplogroups: A, B, C, D, and X. Haplogroups are distinct lineages characterized by single nucleotide polymorphism (SNP) mutations. Haplogroup frequency distributions across the Americas are non-random and are helpful in determining migrations of ancient peoples.

In 2002, Eshleman examined mitochondrial DNA from three different sites in the Central Valley of California with dates ranging from 1800 to 3600 years before present. There were no considerable regional differences in the frequencies of mtDNA haplogroups, which seems to signify genetic continuity in the area. There was evidence of a relationship between linguistic groups and mtDNA, which is helpful for tracing population movements. mtDNA haplogroup frequency distributions were compared between and among both ancient and modern populations in California. The Bear Creek site (aka Cecil site; CA-SJO-112) is associated with Windmiller artifacts and burial practices and has dates comparable to those of the Marsh Creek site. The Cook (CA-SOL-270) and Applegate (CA-AMA-56) sites post-date the Windmiller burials and represent Middle Horizon cemeteries. Analyses showed that the three sites have haplogroup frequencies that are relatively similar to one another, with the largest proportion of individuals belonging to haplogroup C, and haplogroups B and D occurring with less frequency. Only one modern language family group, Takic, exhibits haplogroup frequencies similar to that of the ancient Central Valley.

In sum, the previous mtDNA evidence does not support the hypotheses of an influx of Penutian peoples into Central California associated with Windmiller artifacts. Instead, the evidence suggests multiple migrations and admixture over time. There is no evidence for any relationship between these ancient populations and modern Penutian peoples. A more recent migration into the area is more consistent with the change in haplogroup frequencies (Eshleman 2002).

The San Francisco Bay Area shows different haplogroup frequencies compared with the Sacramento Valley sites. A recent study at the Yukisma cemetery site (CA-SCL-38) near Santa Clara, California, presents mtDNA results from 66 individuals. This site includes burials from the middle to late periods, which are contemporaneous with the Cook and Applegate sites. The majority of individuals at the Yukisma cemetery site belonged to haplogroup D, with smaller proportions represented for haplogroups A, B, and C (Monroe et al. 2011). This haplogroup frequency is statistically different from the Sacramento Valley sites (p<0.009; the 0.05 probability level was corrected using Bonferroni’s Correction to $\alpha = 0.017$ [Abdi 2007]), and more closely resembles that of other San Francisco Bay sites.

Geographically, the Marsh Creek site is intermediate between the San Francisco Bay and the Sacramento Valley. mtDNA analysis at this site with comparisons between other sites in the greater area will help to close the gap in our understanding of ancient California.
Materials and Methods

Twenty teeth from the Marsh Creek site were collected from twenty burials with calibrated radiocarbon dates ranging from 3100 to 3800 years before present. Half the samples came from extended burials, while the other half came from flexed burials. These two sample types were further divided into male and female categories; the sex of each individual was determined at the Marsh Creek site using morphological features. mtDNA was extracted in the Molecular Anthropology Lab (MAL) at UC Davis following the protocol outlined by Kemp et al. (2006) and Kemp and Smith (2005).

A series of steps was taken to prepare the samples for extraction in the aDNA facilities of the MAL. Each tooth had a portion of the root removed (approximately 0.08g), which was then soaked in 6% household bleach to remove any outside contaminates (Kemp and Smith 2005). After the samples were rinsed with ddH2O, they were then placed into 2 ml of molecular grade 0.5 M, pH 8.0, EDTA and mildly rocked for at least 2 days for calcium removal. The 3-day extraction process began by adding 90 ul of Proteinase K to each sample with a 65 degree Celsius incubation period of at least 4 hours, with most left overnight.

The second day consisted of a phenol-chloroform extraction, which was carried out in three steps. The first two extractions used 2ml of phenol-chloroform, while the third extraction used 2ml of chloroform. Following each extraction the samples were centrifuged for 5 minutes to separate the extracted DNA from the phenol-chloroform/chloroform. Once this process was complete, the DNA was precipitated with 100% isopropanol and 5 M ammonium acetate overnight at room temperature to ensure the removal of PCR inhibitors. On the third day, the samples were centrifuged at 4000 rpm for a half hour to pellet the DNA. Pellets were then washed with 1 ml 80% ethanol and then centrifuged again for 30 minutes. After allowing the pellet to dry for 15 minutes, the samples were resuspended in 300 ul of ddH2O and then purified with Wizard PCR Preps Purification System as the manufacturer directed. The DNA was finally eluted with 100 ul of ddH2O.

After extraction, polymerase chain reaction (PCR) and Restriction Fragment Length Polymorphism (RFLP) were used to determine the haplogroup of each individual. Specific regions of the mtDNA that contained mutations specific to haplogroups A, B, C, D, and X were amplified in a different laboratory from the aDNA facility. PCR was carried out using a 13.5 ul cocktail of ddH2O, dNTPs, 10X PCR buffer, MgCl2, and Platinum Taq DNA polymerase with at least 1.5 ul of the DNA template. The DNA was amplified using cycling conditions described in Kemp et al. (2006). The amplified product of expected sizes was then confirmed by electrophoresis on polyacrylamide gels. At this point, samples belonging to haplogroup B could then be determined, as they exhibit a shorter fragment length due to a 9 base pair deletion. Restricting fragment length polymorphisms were then determined by digesting the DNA fragments using the following enzymes: HAE III for haplogroup A, Alu I for haplogroups C and D, and Acc I for haplogroup X. Each enzyme, along with a cocktail of ddH2O, buffer, and BSA, was added to the PCR product, which was then incubated for 9 hours at 37 degrees Celsius. This final product was then run on a polyacrylamide gel to check for haplogroup diagnostic digestion. Haplogroups A, C, and X were identified by single restriction site gains, while haplogroup D was identified by a site loss.
Each sample was extracted on at least two separate occasions to ensure confirmation of haplogroups determined. Protocols were taken to limit any outside sources of contamination at all steps: counters and instruments were bleached before and after every use; gloves were changed frequently and in between different processes; containers were kept closed at all times while not in use; items were not brought back to the aDNA facility from the PCR room without being bleached first; negative controls were used to ensure there had been no contamination.

Fisher’s exact probabilities were conducted between haplogroup frequency distributions of several ancient California sites. Fischer’s exact test was also used to compare the haplogroup frequencies between burial types at the Marsh Creek site. Calculations were performed using Genepop (Raymond and Rousset 1995).

Table 1 lists the samples included in this analysis.

<table>
<thead>
<tr>
<th>Burial #</th>
<th>Sex</th>
<th>Burial Position</th>
<th>$^{14}$C Date</th>
<th>Haplogroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>Female</td>
<td>Flexed</td>
<td>3145 ± 36</td>
<td>D</td>
</tr>
<tr>
<td>26</td>
<td>Female</td>
<td>Flexed</td>
<td>3255 ± 32</td>
<td>UE</td>
</tr>
<tr>
<td>31</td>
<td>Male</td>
<td>Extended</td>
<td>No data</td>
<td>UE</td>
</tr>
<tr>
<td>37</td>
<td>Female</td>
<td>Extended</td>
<td>3050 ± 30</td>
<td>B</td>
</tr>
<tr>
<td>78</td>
<td>Female</td>
<td>Extended</td>
<td>3010 ± 40</td>
<td>C*</td>
</tr>
<tr>
<td>81</td>
<td>Female</td>
<td>Flexed</td>
<td>3080 ± 26</td>
<td>UE</td>
</tr>
<tr>
<td>82</td>
<td>Male</td>
<td>Flexed</td>
<td>2975 ± 26</td>
<td>UE</td>
</tr>
<tr>
<td>105</td>
<td>Male</td>
<td>Flexed</td>
<td>3470 ± 36</td>
<td>D</td>
</tr>
<tr>
<td>107</td>
<td>Male</td>
<td>Extended</td>
<td>3505 ± 31</td>
<td>UE</td>
</tr>
<tr>
<td>116</td>
<td>Indeterminate</td>
<td>Extended</td>
<td>3420 ± 41</td>
<td>UE</td>
</tr>
<tr>
<td>135</td>
<td>Female</td>
<td>Extended</td>
<td>3195 ± 31</td>
<td>C</td>
</tr>
<tr>
<td>142</td>
<td>Male</td>
<td>Extended</td>
<td>3100 ± 31</td>
<td>UE</td>
</tr>
<tr>
<td>163</td>
<td>Female</td>
<td>Flexed</td>
<td>3425 ± 26</td>
<td>UE</td>
</tr>
<tr>
<td>182</td>
<td>Male</td>
<td>Extended</td>
<td>3500 ± 41</td>
<td>UE</td>
</tr>
<tr>
<td>203</td>
<td>Female</td>
<td>Extended</td>
<td>3090 ± 36</td>
<td>C</td>
</tr>
<tr>
<td>210</td>
<td>Male</td>
<td>Flexed</td>
<td>3440 ± 26</td>
<td>B</td>
</tr>
<tr>
<td>217</td>
<td>Male</td>
<td>Flexed</td>
<td>3615 ± 26</td>
<td>UE</td>
</tr>
<tr>
<td>292</td>
<td>Female</td>
<td>Flexed</td>
<td>3050 ± 41</td>
<td>C*</td>
</tr>
<tr>
<td>310</td>
<td>Female</td>
<td>Extended</td>
<td>No data</td>
<td>UE</td>
</tr>
</tbody>
</table>

Notes: UE = Unsuccessful extraction; * confirmed through two extractions
Radiocarbon dates are from Gardner, Bartelink, and Eerkens, unpublished data.
Results

To date, of the twenty samples extracted for mtDNA, only eight haplogroups could be determined with limited confirmation. These eight samples were assigned to one of the five haplogroups native to the Americas. Results are preliminary and more data is needed for confirmation.

Of the extended burials, four females were assigned to haplogroups. Burial 37 was found to belong to haplogroup B, while Burials 78, 135, and 203 belonged to haplogroup C. Of the flexed burials, four haplogroups were also assigned. Burial 14, a female, and Burial 105, a male, were found to belong to haplogroup D, while Burial 210, a male, belonged to haplogroup B and Burial 292, a female, belonged to haplogroup C. Of all these samples, only Burial 78 and Burial 292 were confirmed through two separate extractions (see Table 1). While standard aDNA protocol generally requires that samples provide double confirmation of results, we are still confident that those samples that could not be amplified for a confirmation belong to the haplogroup to which they were assigned. As no one in the aDNA laboratory belongs to the haplogroups noted here, and preventative measures against contamination were taken at all steps, the chance that the haplogroup assignments here are not endogenous to the sample is slim.

Fischer’s exact test showed that the frequency of haplogroups between the two burial styles (extended vs. flexed) at Marsh Creek (CA-CCO-548) was not significantly different. As a whole, Marsh Creek’s haplogroup frequencies were also compared with the Bear Creek site (CA-SJO-112), the Cook site (CA-SOL-270), the Applegate site (CA-AMA-56), and the Yukisima cemetery site (CA-SCL-38). Results showed that the haplogroup frequencies of Marsh Creek were not significantly different from any of the four other sites. However, the same test showed that CA-SCL-38 is significantly different from all other sites except Marsh Creek (p<0.009; the 0.05 probability level was corrected using Bonferroni’s Correction to $\alpha = 0.017$ [Abdi 2007]).

Discussion and Conclusions

With such a limited sample size at Marsh Creek, any patterns in ancestry and ancient migrations are tenuous, but Fisher’s exact test suggests an interesting pattern. The site from the San Francisco Bay (CA-SCL-38) shows haplogroup frequencies that are statistically different from the three sites in the Sacramento Valley, suggesting two separate populations with limited gene flow. Geographically, Marsh Creek is situated between these two areas; haplogroup frequencies do not seem to be statistically different from any of the populations. In other words, the Marsh Creek site seems to be intermediate between San Francisco Bay and Sacramento Valley sites not just geographically but also genetically.

The comparison between haplogroup frequencies of burial types at the Marsh Creek site shows no significant difference, but one that, with such a small sample size, yields a low degree of confidence. Also, with only two male haplogroups assigned, no residence patterns can confidently be determined. More research is needed to better understand these relationships. In the future, the mtDNA hypervariable region of these samples should be sequenced to better understand population relationships in California, and adding more individuals to the sample size will help to reduce any sampling errors. Despite its limitations this aDNA
analysis is important in closing the gap in our understanding of ancient California peoples.

Acknowledgements

First and foremost I would like to expressly thank Ramona Garibay, the most likely descendant, for allowing this research to happen. In addition, special thanks go to Randy Wiberg for site excavation, the National Science Foundation (grant #BCS-0819968) for funding the radiocarbon dating, and Dr. Jelmer Eerkens, my sponsor, for providing me with excellent guidance throughout. Also, I would like to thank Dr. David Smith for allowing me the use of his Molecular Anthropology Lab and the help of his staff and students, including Dr. Meradeth Snow, who trained, mentored, and guided me with patience through this process.

References


