# Ancient DNA Perspectives on American Colonization and Population History

Jennifer A. Raff,[1,2]* Deborah A. Bolnick,[3] Justin Tackney,[1] and Dennis H. O'Rourke[1]

[1]*Department of Anthropology, University of Utah, Salt Lake City, UT*
[2]*Department of Medicine, Endocrinology Division, Feinberg School of Medicine, Northwestern University*
[3]*Department of Anthropology, University of Texas, Austin, TX*

*KEY WORDS* Americas; ancient DNA; genetics

*ABSTRACT* Ancient DNA (aDNA) analyses have proven to be important tools in understanding human population dispersals, settlement patterns, interactions between prehistoric populations, and the development of regional population histories. Here, we review the published results of sixty-three human populations from throughout the Americas and compare the levels of diversity and geographic patterns of variation in the ancient samples with contemporary genetic variation in the Americas in order to investigate the evolution of the Native American gene pool over time. Our analysis of mitochondrial haplogroup frequencies and prehistoric population genetic diversity presents a complex evolutionary picture. Although the broad genetic structure of American prehistoric populations appears to have been established relatively early, we nevertheless identify examples of genetic discontinuity over time in select regions. We discuss the implications this finding may have for our interpretation of the genetic evidence for the initial colonization of the Americas and its subsequent population history. Am J Phys Anthropol 146:503–514, 2011. 

Since the early reviews by O'Rourke et al. (2000a) and Kaestle and Horsburgh (2002), ancient DNA (aDNA) studies have continued to shed light on regional prehistory, the distribution of pathogens and disease in ancient America, and the effects of evolutionary forces on Native American population genetic structure. A few studies have also attempted to use aDNA analyses to address larger questions of inter-continental colonizations and/or population movements (e.g., Kemp et al., 2007; Rasmussen et al., 2010). In this article, we have two primary aims. First, we review the findings of ancient mitochondrial DNA (mtDNA) studies published over the last several years. Second, we draw on these studies to compare the published mtDNA haplogroup frequencies of ancient and modern Native American populations in order to extend the earlier analyses of O'Rourke et al. (2000b) and provide a comprehensive geographic and temporal picture of Native American genetic diversity. Documentation of the extent and patterning of ancient genetic variation relative to modern populations will yield important insights into the evolutionary forces that gave rise to the modern Native American gene pool and permit more informed hypotheses regarding the origin and history of populations in this hemisphere.

Although the timing, geographic routes, and number of initial colonizations into the Western Hemisphere are still contested, it is widely agreed that all genetic components of Native American populations—except those due to more recent demographic collapse and admixture from post-1492 European contact—trace their origins to a sampling of Asian populations before the last glacial maximum (LGM) (Tamm et al., 2007; Wang et al., 2007; Fagundes et al., 2008; González-José et al., 2008; Mulligan et al., 2008; Schroeder et al., 2009; Yang et al., 2010). Following the LGM, founder effects and other evolutionary processes further shaped Native American genetic variation during the expansion into the American continents. The archaeological evidence, hypothetical demographic models of colonizing populations, and genetic signatures of the peopling of the Americas observed in modern populations have been reviewed recently elsewhere (Goebel et al., 2008; Kemp and Schurr, 2010; O'Rourke and Raff, 2010; Perego et al., 2010).

Ancient DNA studies are uniquely situated to investigate the extent and patterns of genetic diversity in past populations and the results of gene flow, genetic drift, and population bottlenecks across time and space. aDNA studies provide important benchmarks not only for evaluation of coalescent models in genome evolution but also for testing general demographic models of the development of genetic diversity in colonizing populations, and for beginning to critically evaluate alternative scenarios bearing on the dispersal of populations into the Americas. As a starting point for evaluating alternative migration scenarios, we compare patterns of genetic variability in ancient and contemporary populations in the Americas. Since the vast majority of aDNA studies have focused on mtDNA, because of its high copy number per cell, maternal inheritance, and rapid evolution, we also

Additional Supporting Information may be found in the online version of this article.

Jennifer A. Raff and Deborah A. Bolnick contributed equally to this manuscript.

*Correspondence to: Jennifer A. Raff, Department of Medicine Endocrinology Division, Feinberg School of Medicine, Northwestern University, Chicago IL 606-3008. E-mail: jenny@northwestern.edu

Received 25 March 2011; accepted 7 July 2011

DOI 10.1002/ajpa.21594
Published online 13 September 2011 in Wiley Online Library (wileyonlinelibrary.com).

focus our attention on variation in this genome. Although a few aDNA studies have also evaluated Y-chromosome haplotypes, too few such studies exist to be informative in a continental survey. Accordingly, we do not include Y-chromosome diversity in our review.

## ANCIENT DNA VARIATION IN THE AMERICAS

Both archaeological research and studies of genetic diversity in modern populations indicate an initial colonization of the Americas sometime before 15,000 years before present (YBP). However, the vast majority of aDNA studies on prehistoric American populations are based on samples dating to the last 5,000 years. The handful of older aDNA results in the Americas postdate the estimated colonization time(s) by several thousands of years (Smith et al., 2005; Kemp et al., 2007; Gilbert et al., 2008a; Kemp and Schurr, 2010). Moreover, the earliest samples studied, those most likely to be most informative for colonization questions, are from single individuals, rendering most standard population genetic analytical methods inappropriate. Although the presence of specific haplotypes at early points in time can be confirmed with small numbers of older samples, genetic diversity in a larger cross section of population(s) before or concurrent with the continental colonization cannot be ascertained from such data. Furthermore, ancient DNA studies focused on more recent remains investigate diversity patterns that have been shaped by thousands of years of demographic processes within the Americas. These more recent prehistoric populations are not appropriate stand-ins for the initial inferred source population, but they are informative for questions of regional prehistory that do not require great time depth. Prehistoric population movements, for example, can be reliably tracked with the larger sample sizes and tighter chronological control characteristic of many aDNA studies with samples from the past few thousand years. Such studies also provide a baseline from which to view the geographic pattern of variation in modern Native American populations, and they provide a benchmark from which to evaluate coalescent models for the origin of the American Indian/Alaska Native gene pool.

In what follows, we summarize and evaluate the results from ancient DNA studies published in the last dozen years that may inform our understanding of the colonization and population history of the Americas. We begin with a review of the studies on the earliest materials available from the Americas, followed by a summary and analysis of temporally more recent material by geographic area.

### Mitochondrial markers and terminology

Because various screening methods have been used in aDNA studies, and because there is not yet a standardized nomenclature in mitochondrial genetics (but see Kemp and Schurr, 2010), we use the following definitions to simplify our discussion and, hopefully, reduce confusion. *Haplogroup* is used to signify a group of mtDNAs that share a defined set of marker states, either in the first hypervariable region (HVS I) or the coding region of the molecule, and are therefore considered to be phylogenetically related. We use the term *haplotype* to refer to distinct subsets of mtDNAs within a haplogroup that share one or more variants in common beyond those used to define the haplogroup. There are currently fifteen recognized founding mtDNA haplotypes in the Americas (Perego et al., 2010). Sequence variants used to define haplogroups and haplotypes are summarized in O'Rourke and Raff (2010) and Van Oven and Kayser (2009). Finally, we use the term *lineage* to indicate the presence of additional variants beyond those that define haplotypes.

## METHODS

Because of the focus in early aDNA studies on a very limited number of restriction sites and a length polymorphism, the most useful comparison among all aDNA samples published over the last few years is by mtDNA haplogroup. Accordingly, mtDNA haplogroup frequencies for the ancient populations discussed here are given in Table 1, and geographic locations for all populations are shown in Figure 1. We also compiled haplogroup frequency data (from the published literature) for 98 modern Native American populations from the regions of North and South America where ancient populations have been studied (Supporting Information Table 1). To evaluate the extent of differentiation among populations grouped both geographically and temporally (usually ancient vs. modern, but additional temporal categories were added when the data permitted), we performed analyses of molecular variance (AMOVA) in the computer package Arlequin 3.5 (Excoffier and Lischer, 2010). $F_{ST}$ values were estimated for population pairs and Mantel tests were performed in Arlequin 3.5 to assess correlations between genetic distances and temporal distances (in years) to evaluate the role of time on population differentiation. If the samples in an ancient population spanned an extended time period, the average time was used in calculating temporal distances (but see our comments on the limitations of this approach below). Regional haplogroup frequencies in ancient and modern times were also calculated (Table 2), and principal components analysis (PCA) was performed using R version 2.10.1 (R Development Core Team, 2009) to illustrate genetic similarities and differences. The first two principal components are reported, representing 64.2% of the observed haplogroup variation.

Finally, haplogroup diversities ($h$) for populations (Table 1) and geographic regions (Table 2) were calculated using Arlequin 3.5. We performed a linear regression of population haplogroup diversity on time in R to investigate changes in haplogroup diversity levels over time, and an analysis of variance (ANOVA) was conducted using R to assess the influence of temporal and geographic groupings on population haplogroup diversities. We also used $t$-tests to compare the overall haplogroup diversity in ancient and modern times for each geographic region. It should be noted that the regional estimates of ancient haplogroup diversity/haplogroup frequencies represent composite measures based on population samples that may span many centuries. We are well aware of the difficulties inherent in using temporally dispersed samples from aDNA studies in population genetic analyses. We employ these methods here as a first approximation to evaluating stasis or change in genetic structure over time, but believe the results should be viewed more properly as qualitative rather than quantitative assessments. As such, they may serve as guidelines for future research in American population history using modern genomic methods.

*TABLE 1. Haplogroup frequencies and diversities (h) of ancient populations included in this review*

| | Site | Region | Date (YBP) | N | % A | % B | % C | % D | % X | *h* | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Dorset | Arctic/Subarctic | 2260–1216 | 2 | 0.000 | 0.000 | 0.000 | 1.000 | 0.000 | N.C.[a] | Hayes, 2002 |
| 2 | Sadlermiut | Arctic/Subarctic | 977–682 | 18 | 0.560 | 0.000 | 0.000 | 0.440 | 0.000 | 0.523 | Hayes, 2002 |
| 3 | Thule | Arctic/Subarctic | 1130–628 | 15 | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | Hayes, 2002 |
| 4 | Brooks River | Arctic/Subarctic | 938–1318 | 8 | 0.625 | 0.250 | 0.000 | 0.125 | 0.000 | 0.607 | Raff et al., 2010 |
| 5 | Hot Springs | Arctic/Subarctic | ~ 2070 | 3 | 0.333 | 0.000 | 0.000 | 0.666 | 0.000 | N.C. | Raff et al., 2010 |
| 6 | Mink Island | Arctic/Subarctic | 1029–1215 | 6 | 0.167 | 0.000 | 0.000 | 0.833 | 0.000 | 0.333 | Raff et al., 2010 |
| 7 | Tatshenshini-Alsek Glacier | Arctic/Subarctic | 550 | 1 | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 | N.C. | Monsalve et al., 2002 |
| 8 | Early Aleuts | Arctic/Subarctic | 3500–1200 | 11 | 0.727 | 0.000 | 0.000 | 0.273 | 0.000 | 0.436 | Smith et al., 2009 |
| 9 | Late Pre-Contact Aleuts | Arctic/Subarctic | 1000–400 | 52 | 0.231 | 0.000 | 0.000 | 0.769 | 0.000 | 0.362 | Smith et al., 2009 |
| 10 | On-Your-Knees Cave | Arctic/Subarctic | 10,300 | 1 | 0.000 | 0.000 | 0.000 | 1.000 | 0.000 | N.C. | Kemp et al., 2007 |
| 11 | La Caleta (Tainos) | Caribbean | 1330–320 | 24 | 0.000 | 0.000 | 0.750 | 0.250 | 0.000 | 0.391 | Lalueza-Fox et al., 2001 |
| 12 | Cuba (Ciboneys) | Caribbean | 4700–1620 | 15 | 0.067 | 0.000 | 0.600 | 0.333 | 0.000 | 0.562 | Lalueza-Fox et al., 2003 |
| 13 | Tlatelolco Post-Classic Aztec | Mesoamerica | 500–675 | 23 | 0.652 | 0.130 | 0.043 | 0.174 | 0.000 | 0.574 | Kemp et al., 2005 |
| 14 | Maya (Xcaret) | Mesoamerica | 480–1400 | 24 | 0.840 | 0.040 | 0.080 | 0.000 | 0.000 | 0.236 | González-Oliver et al., 2001 |
| 15 | Maya (Copan) | Mesoamerica | 750–1300 | 9 | 0.000 | 0.000 | 0.890 | 0.110 | 0.000 | 0.222 | Merriwether et al., 1997 |
| 16 | Great Western Park | Northeast/Midwest | 800 | 6 | 0.333 | 0.000 | 0.667 | 0.000 | 0.000 | 0.533 | Shook and Smith, 2008 |
| 17 | Glacial Kame | Northeast/Midwest | 2900 | 18 | 0.167 | 0.611 | 0.167 | 0.056 | 0.000 | 0.601 | Shook and Smith, 2008 |
| 18 | Morse | Northeast/Midwest | 2700 | 9 | 0.111 | 0.333 | 0.556 | 0.000 | 0.000 | 0.639 | Shook and Smith, 2008 |
| 19 | Orendorf | Northeast/Midwest | 800 | 11 | 0.455 | 0.000 | 0.273 | 0.273 | 0.000 | 0.709 | Shook and Smith, 2008 |
| 20 | Ohio Hopewell Mound Group | Northeast/Midwest | 1700 | 34 | 0.412 | 0.088 | 0.294 | 0.206 | 0.000 | 0.715 | Mills, 2003 |
| 21 | Pete Klunk Mound Group | Northeast/Midwest | 1825 | 39 | 0.231 | 0.128 | 0.487 | 0.128 | 0.026 | 0.694 | Bolnick and Smith, 2007 |
| 22 | Schild Mississippian | Northeast/Midwest | 900 | 47 | 0.383 | 0.128 | 0.234 | 0.085 | 0.170 | 0.762 | Raff, 2008 |
| 23 | Schild Late Woodland | Northeast/Midwest | 1200 | 19 | 0.263 | 0.053 | 0.211 | 0.053 | 0.421 | 0.743 | Raff, 2008 |
| 24 | Norris Farms | Northeast/Midwest | 700 | 108 | 0.315 | 0.120 | 0.426 | 0.083 | 0.056 | 0.702 | Stone and Stoneking, 1998 |
| 25 | Plateau Salish | Northwest | 200 | 11 | 0.273 | 0.545 | 0.091 | 0.091 | 0.000 | 0.673 | Malhi et al., 2004 |
| 26 | Plateau Sahaptian | Northwest | 200 | 8 | 0.000 | 0.500 | 0.250 | 0.250 | 0.000 | 0.714 | Malhi et al., 2004 |
| 27 | Wishram | Northwest | 200 | 33 | 0.212 | 0.515 | 0.000 | 0.273 | 0.000 | 0.635 | Malhi et al., 2004 |
| 28 | Vantage | Northwest | 500–1500 | 7 | 0.286 | 0.000 | 0.143 | 0.429 | 0.143 | 0.810 | Malhi et al., 2004 |
| 29 | Paisley 5 Mile Point Caves | Northwest | 14,000 | 2 | 0.500 | 0.500 | 0.000 | 0.000 | 0.000 | N.C. | Gilbert et al., 2008a |
| 30 | China Lake and Big Bar Lake | Northwest | 4975 | 3 | 0.333 | 0.000 | 0.000 | 0.000 | 0.000 | N.C. | Malhi et al., 2007[b] |
| 31 | Paracas (Peninsula) | South America | 2800–2200 | 10 | 0.000 | 0.000 | 0.300 | 0.700 | 0.000 | 0.467 | Fehren-Schmitz et al., 2010 |
| 32 | Paracas (Palpa) | South America | 2800–2200 | 28 | 0.071 | 0.000 | 0.143 | 0.786 | 0.000 | 0.370 | Fehren-Schmitz et al., 2010 |
| 33 | Nasca-Rural (Palpa) | South America | 2200–1400 | 37 | 0.027 | 0.108 | 0.216 | 0.649 | 0.000 | 0.535 | Fehren-Schmitz et al., 2010 |
| 34 | Nasca-Urban (Palpa) | South America | 2200–1400 | 28 | 0.000 | 0.179 | 0.429 | 0.393 | 0.000 | 0.653 | Fehren-Schmitz et al., 2010 |
| 35 | Middle Horizon (Palpa) | South America | 1400–1000 | 11 | 0.000 | 0.273 | 0.364 | 0.364 | 0.000 | 0.727 | Fehren-Schmitz et al., 2010 |
| 36 | Pacapaccari (Highlands) | South America | 820 | 16 | 0.000 | 0.688 | 0.313 | 0.000 | 0.000 | 0.458 | Fehren-Schmitz et al., 2010 |
| 37 | Yacotogia | South America | 1187 | 25 | 0.040 | 0.560 | 0.400 | 0.000 | 0.000 | 0.547 | Fehren-Schmitz et al., 2011 |
| 38 | Ocoro | South America | 1400–1000 | 5 | 0.200 | 0.600 | 0.200 | 0.000 | 0.000 | 0.700 | Fehren-Schmitz et al., 2011 |
| 39 | Botigiriayocc | South America | 1000–600 | 12 | 0.000 | 0.688 | 0.313 | 0.000 | 0.000 | 0.485 | Fehren-Schmitz et al., 2011 |
| 40 | Huayuncalla | South America | 978 | 5 | 0.200 | 0.400 | 0.400 | 0.000 | 0.000 | 0.800 | Fehren-Schmitz et al., 2011 |
| 41 | Layuni | South America | 1400–1000 | 9 | 0.000 | 0.556 | 0.333 | 0.111 | 0.000 | 0.639 | Fehren-Schmitz et al., 2011 |
| 42 | Conchapata | South America | 1400–1200 | 14 | 0.286 | 0.500 | 0.143 | 0.071 | 0.000 | 0.692 | Kemp et al., 2009 |
| 43 | Chen Chen | South America | 1215–1000 | 23 | 0.391 | 0.391 | 0.174 | 0.043 | 0.000 | 0.692 | Lewis et al., 2007 |
| 44 | Huari | South America | 900–600 | 18 | 0.167 | 0.222 | 0.555 | 0.056 | 0.000 | 0.647 | Kemp et al., 2009 |
| 45 | Pampa Grande | South America | 1600–1350 | 19 | 0.105 | 0.474 | 0.000 | 0.421 | 0.000 | 0.620 | Carnese et al., 2010 |
| 46 | Peruvian Highlanders | South America | 550–450 | 35 | 0.086 | 0.657 | 0.229 | 0.029 | 0.000 | 0.523 | Shinoda et al., 2006 |
| 47 | North Peruvian Coast | South America | 1000 | 21 | 0.286 | 0.333 | 0.048 | 0.333 | 0.000 | 0.729 | Shimada et al., 2004 |
| 48 | Cuzco | South America | 1020–830 | 1 | 0.000 | 0.000 | 0.000 | 1.000 | 0.000 | N.C. | Luciani et al., 2006 |
| 49 | North Chile Late Archaic | South America | 6000–3900 | 14 | 0.500 | 0.357 | 0.071 | 0.071 | 0.000 | 0.659 | Moraga et al., 2005 |

*TABLE 1. (Continued)*

| | Site | Region | Date (YBP) | N | % A | % B | % C | % D | % X | *h* | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 | North Chile Middle Horizon | South America | 1650–1000 | 19 | 0.316 | 0.421 | 0.263 | 0.000 | 0.000 | 0.690 | Moraga et al., 2005 |
| 51 | North Chile Late Intermediate | South America | 1000–500 | 15 | 0.200 | 0.533 | 0.200 | 0.067 | 0.000 | 0.676 | Moraga et al., 2005 |
| 52 | Pyramid Lake | Southwest/Great Basin/CA | 860–5905 | 18 | 0.111 | 0.333 | 0.000 | 0.556 | 0.000 | 0.601 | Kaestle and Smith, 2001 |
| 53 | Pyramid Lake (Wizards Beach) | Southwest/Great Basin/CA | 9200 | 1 | 0.000 | 0.000 | 1.000 | 0.000 | 0.000 | N.C. | Kaestle and Smith, 2001 |
| 54 | Stillwater Marsh | Southwest/Great Basin/CA | 290–3290 | 21 | 0.048 | 0.381 | 0.000 | 0.571 | 0.000 | 0.552 | Kaestle and Smith, 2001 |
| 55 | Tommy Site | Southwest/Great Basin/CA | 850–1150 | 36 | 0.028 | 0.694 | 0.139 | 0.139 | 0.000 | 0.492 | Snow et al., 2010 |
| 56 | Mine Canyon | Southwest/Great Basin/CA | 650–850 | 12 | 0.583 | 0.333 | 0.083 | 0.000 | 0.000 | 0.591 | Snow et al., 2010 |
| 57 | Fremont | Southwest/Great Basin/CA | 500–1500 | 30 | 0.000 | 0.800 | 0.133 | 0.067 | 0.000 | 0.349 | Parr et al., 1996; Carlyle et al., 2000 |
| 58 | Anasazi | Southwest/Great Basin/CA | 1010–2010 | 38 | 0.107 | 0.714 | 0.179 | 0.000 | 0.000 | 0.462 | Carlyle, 2003 |
| 59 | Western Basketmaker II | Southwest/Great Basin/CA | 2500–1300 | 23 | 0.130 | 0.783 | 0.043 | 0.043 | 0.000 | 0.383 | Carlyle, 2003; Leblanc et al., 2007 |
| 60 | Cecil | Southwest/Great Basin/CA | 3600–2860 | 16 | 0.000 | 0.063 | 0.563 | 0.375 | 0.000 | 0.575 | Eshleman, 2002 |
| 61 | Cook | Southwest/Great Basin/CA | 2000 | 23 | 0.043 | 0.087 | 0.435 | 0.435 | 0.000 | 0.640 | Eshleman, 2002 |
| 62 | Applegate | Southwest/Great Basin/CA | 1765–2055 | 6 | 0.000 | 0.333 | 0.667 | 0.000 | 0.000 | 0.533 | Eshleman, 2002 |
| 63 | Hourglass Cave | Southwest/Great Basin/CA | 8000 | 1 | 0.000 | 1.000 | 0.000 | 0.000 | 0.000 | N.C. | Stone and Stoneking, 1996 |

[a] N.C., Not Calculated.
[b] Also reported 0.667 frequency of haplogroup M.

## RESULTS

### The oldest samples (≥8,000 YBP)

Gilbert et al. (2008a) obtained DNA from 14 coprolites excavated from the Paisley 5 Mile Point Caves in south-central Oregon (see Fig. 1 for the locations of all sites in this study). These samples were dated to ≥14,000 YBP, making this site and the aDNA results coeval with the early site of Monte Verde in Chile, and predating the Clovis archaeological culture in North America. Although the coprolites were not excavated under sterile conditions and were contaminated with nonendogenous (primarily European) DNA, cloned and pyrosequenced PCR products from six samples yielded Native American SNPs diagnostic for mitochondrial haplotypes A2 and B2. These results appear to represent the oldest human DNA reported in the Americas and provide evidence of a pre-Clovis human presence in western North America, suggesting at least one pre-LGM founding event. They also suggest that mtDNA haplotypes A2 and B2 have the time depth expected for founding haplotypes (see Perego et al., 2009 for recent coalescence estimates).

Kemp et al. (2007) identified mtDNA haplotype D4h3a (Rickards et al. 1999; Perego et al., 2009) from a single sample dated to 10,300 YBP from the On-Your-Knees (O-Y-K) Cave site on Prince of Wales Island, Alaska. There are two clades within the D4h3 haplotype. D4h3a is commonly found in Pacific coastal populations of South America, and is represented in North America by the O-Y-K Cave specimen and one individual from the Pete Klunk mound group (1825 YBP) in Illinois (Bolnick and Smith, 2007). D4h3b is known from a single modern sample in China. The identification of D4h3a in the early O-Y-K Northwest coast individual was an important piece of evidence for Perego et al.'s (2009) dual migration colonization model for the Americas.

Two additional samples that predate 8,000 years ago have also been reported. Using HVS I sequencing, Stone and Stoneking (1998) reported an individual member of mtDNA haplogroup B from Hourglass Cave in Colorado dating to 8,000 YBP, while Kaestle (1998) and Kaestle and Smith (2001) identified the 9,200-year-old Wizard's Beach, Nevada individual as a member of mtDNA haplogroup C. Thus, in the small number of individuals from four archaeological sites dating to 8,000 years ago or earlier that have been characterized for mtDNA lineages, each carried a different representative of the currently recognized Native American mtDNA founding haplogroups.

Smith et al. (2005) attempted to genetically characterize all of the earliest human remains in North America, identifying 20 individuals dating to older than 7,400 YBP. Of the 18 samples that had not been studied genetically before, 10 yielded insufficient DNA for analysis, of which eight dated to earlier than 8,000 YBP. An additional three yielded apparently adequate DNA but contained inhibitors that prevented amplification and sequencing. To our knowledge, mtDNA haplogroups have not been reported for any of the remaining samples.

The paucity of early aDNA studies that might contribute directly to our understanding of the initial colonization of the Western Hemisphere is the result of two factors. First, very few remains have been excavated from early sites that are available for study, and second, the earliest samples have proven recalcitrant to molecular analysis due to poor nucleic acid preservation or inhibitors in the DNA extracts. Nevertheless, the few truly an-


**Fig. 1.** Location of sites reviewed in this study. Site names corresponding to numbers are given in Table 1.

cient samples in the Americas for which genetic information is available have been important in documenting the geographic range of human populations early in American prehistory, as well as the extent of mitochondrial genetic variation present in the initial populations of the Americas.

## More recent ancient samples (6000-200 YBP)

Our analysis of 59 more recent ancient samples (6000-200 YBP) and 98 modern populations (Supporting Information Table 1) suggests that continent-wide haplogroup patterns have not changed substantially over the last several thousand years. Haplogroup frequencies for the entire American continent are fairly similar in ancient and modern times, and the linear regression of population haplogroup diversity on time (see Fig. 2) indicates no statistically significant changes in population haplogroup diversity levels over the last 4,000 years ($t$ = 1.135, $P$ = 0.258, $r^2$ = 0.002). Furthermore, when populations are grouped temporally (i.e., ancient vs. modern), an ANOVA shows that population haplogroup diversities do not differ between the two temporal groups (Table 3).

However, some changes in regional haplogroup patterns (Table 2) are evident over time. Comparing regional haplogroup diversities in ancient and modern times reveals an increase in haplogroup diversity in the Southwest over time ($P$ = 0.006), but a significant decrease in haplogroup diversity in the Northeast ($P$ = 0.007), on the Columbian Plateau ($P$ = 0.012), and in South America ($P$ < 0.001). Ancient and modern haplogroup diversities do not differ for other regions ($P$ > 0.050). The PCA of regional haplogroup frequencies also shows that these frequencies have changed from ancient

*TABLE 2. Regional haplogroup frequencies and diversities (h) in ancient and modern times*

| Region | Time Period | *N* | % A | % B | % C | % D | % X | *h* |
|---|---|---|---|---|---|---|---|---|
| Arctic/Sub-Arctic | Ancient | 113 | 0.453 | 0.017 | 0.000 | 0.530 | 0.000 | 0.518 |
| | Modern | 332 | 0.551 | 0.000 | 0.012 | 0.437 | 0.000 | 0.507 |
| Caribbean | Ancient | 39 | 0.026 | 0.000 | 0.692 | 0.282 | 0.000 | 0.452 |
| | Modern | 570 | 0.546 | 0.082 | 0.328 | 0.044 | 0.000 | 0.587 |
| Great Basin/CA | Ancient | 85 | 0.047 | 0.224 | 0.283 | 0.447 | 0.000 | 0.676 |
| | Modern | 302 | 0.073 | 0.454 | 0.179 | 0.288 | 0.007 | 0.676 |
| Mesoamerica | Ancient | 70 | 0.629 | 0.100 | 0.172 | 0.100 | 0.000 | 0.564 |
| | Modern | 1409 | 0.472 | 0.258 | 0.231 | 0.038 | 0.000 | 0.656 |
| Northeast/Midwest | Ancient | 291 | 0.313 | 0.144 | 0.361 | 0.103 | 0.079 | 0.737 |
| | Modern | 502 | 0.484 | 0.118 | 0.227 | 0.030 | 0.142 | 0.681 |
| Northwest | Ancient | 64 | 0.219 | 0.437 | 0.063 | 0.235 | 0.016 | 0.712 |
| | Modern | 322 | 0.319 | 0.262 | 0.113 | 0.279 | 0.028 | 0.741 |
| South America | Ancient | 365 | 0.134 | 0.370 | 0.246 | 0.249 | 0.000 | 0.724 |
| | Modern | 1278 | 0.190 | 0.481 | 0.194 | 0.133 | 0.000 | 0.677 |
| Southwest | Ancient | 139 | 0.108 | 0.705 | 0.129 | 0.057 | 0.000 | 0.475 |
| | Modern | 740 | 0.127 | 0.496 | 0.354 | 0.005 | 0.017 | 0.613 |

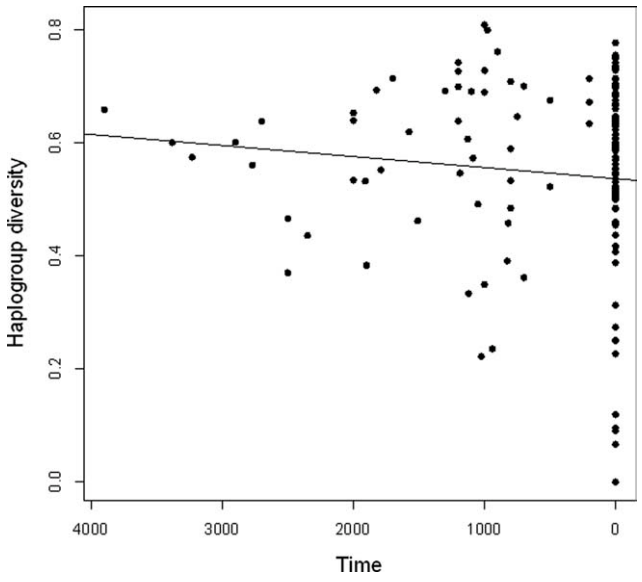


**Fig. 2.** Regression of population haplogroup diversity (h) on time.

*TABLE 3. Analysis of variance results*

| | df | SS | MS | *F* | *P* |
|---|---|---|---|---|---|
| Time Period | 1 | 0.053 | 0.053 | 2.343 | 0.128 |
| Region | 7 | 1.617 | 0.231 | 10.286 | <0.001 |
| Time Period*Region | 7 | 0.222 | 0.032 | 1.411 | 0.206 |
| Residuals | 136 | 3.054 | 0.022 | | |

to modern times (see Fig. 3). However, the magnitude and direction of change is not consistent across regions, and differences between regions generally persist over time. AMOVAs based on the regional haplogroup frequencies confirm this observation. Temporal groupings do not account for a statistically significant portion of the total genetic variance (−0.77%, $P = 0.415$), but regional groupings (i.e., groups composed of the ancient and modern haplogroup frequencies for the same region) do explain a significant amount of the observed haplogroup variation (9.42%, $P = 0.001$). Likewise, an ANOVA shows that population haplogroup diversities do vary significantly among geographic regions (Table 3).

These results demonstrate that regional differentiation and regional haplogroup patterns are not simply a modern phenomenon. They have existed for at least the last few thousand years, and recent historical events (such as those associated with European contact) do not seem to have drastically altered such patterns. Because regional differences have persisted over time, further description and analysis of the aDNA data is presented separately for each geographic region.

***Arctic/Sub-Arctic.*** Modern populations of the Arctic are characterized by high frequencies of haplotype A2, with Aleut populations exhibiting a high frequency of haplotype D2a (Crawford et al., 2010). Similarly, ancient populations from this region exhibit high frequencies of haplogroups A and D; other haplogroups are either absent or present at only extremely low frequencies. These similarities between the ancient and modern populations are confirmed by an AMOVA, which shows that ancient versus modern contrasts do not account for a significant portion of the total variance in population haplogroup frequencies (−20.36%, $P = 0.827$). A Mantel test also suggests that there is no significant correlation between genetic and temporal distances in this region ($r = 0.081$, $P = 0.191$). However, genetic changes over time are evident when populations from the Eastern Arctic and Aleutian Islands are considered separately.

Archaeological evidence suggests substantial and abrupt changes in population composition in both areas beginning around 1000 AD, consistent with dramatic genetic differences observed between the early Paleo-Eskimo Dorset and Neo-Eskimo Thule in the Eastern Canadian Arctic (Hayes, 2002; Hayes et al. 2003, 2005). These findings support the archaeological inference of a Thule migration circa-1000 AD as the foundation for the modern populations of the Eastern Arctic.

Haplogroup frequencies (determined by RFLP analysis) were also reported for samples from the Eastern Aleutian Islands by Hayes (2002) and later expanded by Smith et al. (2009). Here, too, genetic evidence suggests a population transition around 1000 AD. Samples predating this time exhibit high frequencies of haplogroup A (72%), while samples dating after AD 1000 are characterized by high frequencies of haplogroup D (77%), as are modern Unangan (Aleut) populations. Recently, Raff et al. (2010) reported HVS I sequence results for three

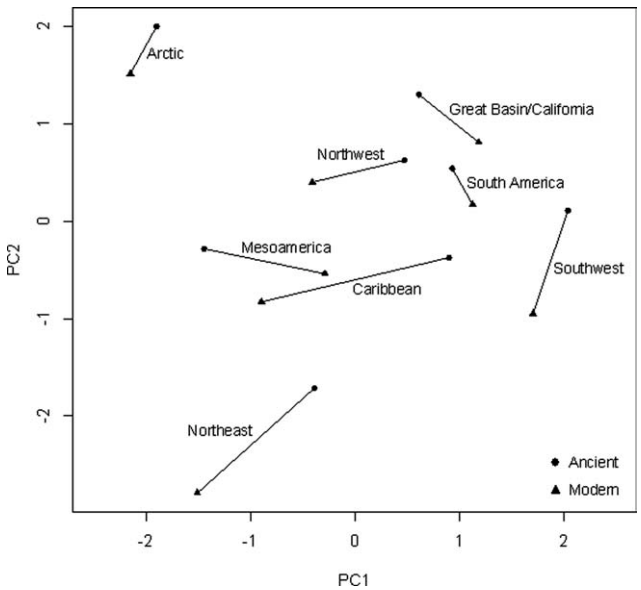


**Fig. 3.** Principle components analysis of regional haplogroup frequencies in ancient and modern times.

sites on (Hot Springs and Brooks River) or adjacent to (Mink Island) the Alaska Peninsula. Their results support Smith et al.'s (2009) finding of a population transition, perhaps a movement of people from the Peninsula to the Aleutian archipelago, around 1000 AD, and they also identify the northernmost examples of mtDNA haplotype B2 in an ancient American group (Brooks River).

Finally, two landmark studies have been conducted using aDNA from the Arctic. Gilbert et al. (2008b) sequenced the entire mitochondrial genome from a Greenland Paleo-Eskimo (Saqqaq) individual dated to about 4,170–3,600 $^{14}$C YBP, followed shortly by 79% of the nuclear genome (Rasmussen et al., 2010). The Saqqaq individual carried mtDNA haplotype D2a1, a haplotype found in modern Aleuts, Siberian Sireniki Yuits and American Na-Dene, but not in descendents of the Neo-Eskimo (Thule). The estimated coalescent for this haplotype was also older, as expected, than known Neo-Eskimo haplotypes (A2a, A2b, D3). SNP analysis of the nuclear genome consistently grouped the Saqqaq individual with the Koryaks and Chukchis of Chukotka and northern Kamchatka, and showed a more distant relationship to Native American populations. These three observations suggest that this Paleo-Eskimo originated from a Beringian source population, as did all other Native Americans, and therefore shares common ancestry with them, but at a time when the genetic constitution of the population(s) was different than that of either the earlier migrations into the Americas or the much more recent Thule migrations.

***Northwest.*** The Pacific Northwest exhibits high frequencies of haplogroups A, B, and D (as well as lower frequencies of C and X) in both ancient and modern times. When all populations from this region are considered, an AMOVA shows that temporal groupings (1,500–500 YBP, 200 YBP, and modern) do not explain the variance in population haplogroup frequencies (−4.30%, $P = 0.646$). Similarly, a Mantel test indicates no significant correlation between genetic and temporal distances in the Pacific Northwest ($r = -0.021$, $P = 0.555$).

However, clear geographic structuring exists in the mtDNA haplogroup frequency distributions for modern populations in this region, with the frequency of haplogroup A increasing with distance from the coast. The Columbian Plateau populations have high frequencies of haplogroups B and D, while the coastal populations are characterized by a high frequency of haplogroup A and only moderate frequencies of haplogroups B, C, and D. Since all published aDNA data from the Pacific Northwest are from the Columbian Plateau, we investigated whether genetic changes have occurred over time on the Plateau, separately. The available data suggest a decrease in the frequency of haplogroup A and an increase in the frequency of haplogroup B over time in this region. While an AMOVA found that temporal groupings (1500–500 YBP, 200 YBP, and modern) do not account for the variance in population haplogroup frequencies on the Columbian Plateau (1.65%, $P = 0.432$), a Mantel test does indicate a significant correlation between genetic and temporal distances ($r = 0.791$, $P = 0.030$). In addition, overall haplogroup diversity seems to have declined on the Columbian Plateau from ancient to modern times ($P = 0.012$).

These results indicate that the genetic makeup of populations on the Columbian Plateau has been altered over the last few thousand years. Malhi et al. (2004) identified lineage sharing as well as similar haplogroup frequencies between an ancient Northern Plateau group (the 500–1500 YBP Vantage site) and modern populations of the Northwest Coast, which they cited as evidence for a pre-1500 YBP population incursion from the Coastal/Subarctic region, possibly associated with the hypothesized expansion of Salishan-speaking peoples into the Northern Plateau.

It should also be noted that the only evidence of mtDNA haplogroup M in indigenous Americans comes from studies of aDNA from the Pacific Northwest. Malhi et al. (2007, 2010) used RFLP analysis and direct sequencing to identify two individuals from the China Lake site (4,950 $^{14}$C YBP) in British Columbia that belong to haplogroup M, an Asian haplogroup not previously found in Native Americans. This haplogroup has yet to be confirmed in other ancient or modern samples from the Americas, but continues to spur interest in identifying potential founder lineages that may be absent from modern populations.

***Northeast and Midwest.*** Following Stone and Stoneking's (1998) classic study of the 700 YBP Norris Farms site, there have been several investigations of mtDNA variation in ancient populations from the Northeast and Midwest. Mills (2003) analyzed individuals of the Hopewell Mound Group (1,700 YBP) in Ohio, while Bolnick and Smith (2007) studied the Middle Woodland population from the Pete Klunk Mound Group (1,825 YBP) in West-Central Illinois. Raff (2008) reported haplogroup frequencies for the Late Woodland (1,200 YBP) and Mississippian (900 YBP) components of the Schild site, also in West-Central Illinois. Shook and Smith (2008) analyzed mtDNA from four sites: two in the Central Illinois Valley (Morse and Orendorf, respectively dating to 2,700 YBP and 800 YBP), and two in southwestern Ontario (Great Western Park and Glacial Kame, respectively dating to 800 YPB and 2,900 YBP).

The haplogroup frequency distributions for many of these ancient populations are similar and fit with the overall pattern for modern populations in the Northeast

and Midwest (high frequencies of haplogroups A and C, and moderate frequencies of haplogroups B, D, and X). However, the two earliest populations (Glacial Kame and Morse) are distinct, exhibiting higher frequencies of haplogroup B and lower frequencies of A (Shook and Smith, 2008). The Late Woodland and Mississippian components of Schild also differ from the other ancient populations in exhibiting higher frequencies of haplogroup X (Raff, 2008). Nevertheless, these data suggest that haplogroups B and C decreased in frequency, whereas haplogroups A and X increased in frequency over time in this region. An AMOVA shows that a significant amount of variation in haplogroup frequencies occurred between four designated time periods (3.02%, $P = 0.024$), and a Mantel test indicates a significant correlation between genetic and temporal distances ($r = 0.372$, $P = 0.005$). Regional haplogroup diversity has also decreased over time ($P = 0.007$), as noted earlier. Altogether, these results suggest that haplogroup patterns changed gradually over time, but the general regional pattern seems to have been established by $\sim$ 2,000 years ago.

Ancient DNA from eastern North America was also used in the first study of autosomal variation in ancient populations from the Americas. Halverson and Bolnick (2008) determined ABO genotypes for 15 ancient remains from Illinois (1,825 $\pm$ 75 YBP) and Kentucky (1,950–1,790 YBP). The ancient ABO frequencies were compared with frequencies from 45 extant Siberian populations and 15 extant Native American populations to investigate whether the high frequency of blood type O in most extant Native Americans (excluding the Na-Dene and Aleut-Eskimo) was the result of an initial founder effect during the peopling of the Americas or historical events associated with European contact. Halverson and Bolnick (2008) found no significant differences between the allele frequencies of the ancient, pre-contact North Americans (.033 A, 0.000 B, 0.967 O) and those of modern Eastern North Americans (.120 A, 0.014 B, 0.865 O), although both differed substantially from modern Siberians (0.190 A, 0.188 B, 0.623 O). This aDNA result suggests that European contact was not the main determinant of ABO frequency distributions in the Americas (Halverson and Bolnick, 2008), a result in agreement with O'Rourke et al. (2000b) and this study, both of which document a strong concordance between the geographic distribution of mtDNA haplogroups in both ancient and modern native North American populations.

***Southwest, Great Basin, and California.*** Today, haplogroups B, C, and D are the most prevalent haplogroups in the Southwest, Great Basin, and California, with haplogroup D more common in the Great Basin and California than in the Southwest. Ancient populations from this region generally show the same patterns, and AMOVAs suggest that temporal groupings (ancient vs. modern) do not account for a significant percentage of the observed variation in population haplogroup frequencies (Southwest: 2.99%, $P = 0.272$; Great Basin/California: 2.76%, $P = 0.109$). However, overall haplogroup diversity in the Southwest has increased from ancient to modern times ($P = 0.006$), perhaps due to migrations or other population changes in the region.

A number of aDNA papers have focused on questions of population movement in the Southwest, Great Basin, and California. First, the archaeological hypothesis of a migration of Numic-speaking peoples into the Great Basin (replacing the older, Fremont culture) was tested by Kaestle and Smith (2001). A statistically significant difference was found between the ancient and modern populations of the region, supporting the Numic Expansion hypothesis and suggesting that the ancient Fremont peoples are instead ancestral to the modern California Penutians (2008) (although Cabana et al.'s (2008) computer simulations indicate that low gene flow and genetic drift could also account for the observed differences between ancient and modern populations in the Great Basin). In a second study, Carlyle et al. (2000) examined mtDNA haplogroup diversity in Anasazi populations, adding to previously reported data from eastern Fremont samples excavated from the margins of the Great Salt Lake (Parr et al., 1996; Parr, 1998), and found the GSL Fremont resembled the ancient Ancestral Puebloan samples (Anasazi) more closely than they did the western Fremont samples reported by Kaestle and Smith (2001).

Eshleman (2002) also used aDNA to test the Hokan-Penutian linguistic hypothesis, which suggests that one or more migrations of Penutian speakers into California's Central Valley replaced the earlier Hokan speakers in that area. He found no statistically significant difference in haplogroup frequency distributions between three ancient populations of the region, representing two cultures (Windmiller 2800–3600 YBP, and Middle Horizon $\sim$ 1700–2000 YBP), and hypothesized to represent descendents of these immigrants. However, all three ancient populations were significantly different from all modern Penutian and Hokan speakers; instead, they were most similar to the Uto-Aztecan-speaking Takics from Southern California.

Finally, Snow et al. (2010) obtained aDNA from 48 ancient Puebloan individuals from the Tommy (850–1150 YBP) and Mine Canyon (650–850 YBP) sites near Farmington, New Mexico. The haplogroup frequency differences between these sites were statistically significant. The population of the Tommy site, with a high frequency of haplogroup B, most closely resembled the ancient Anasazi and Fremont (Carlyle et al., 2000; Leblanc et al., 2007), as well as the modern populations of Jemez and Zuni. The population from the Mine Canyon site resembled ancient/extant Mesoamerican populations (Mixtec, Nahua, Maya) and southern Athapascans (Apache and Navaho) in having a high frequency of haplogroup A. However, the HVS I sequence data reveal haplotypes at both the Tommy and Mine Canyon sites which are common throughout modern populations of the Southwest (but not in Mesoamerica), suggesting subsequent population continuity.

***Mesoamerica.*** Today, Mesoamerica is characterized by higher frequencies of haplogroups A, B, and C and only a low frequency of haplogroup D. Three ancient populations from this region have been studied: a 500–675 YBP sample of Post-Classic Aztec from Tlatelolco, Mexico (Kemp et al., 2005), a 480–1400 YBP Maya sample from Xcaret, Mexico (González-Oliver et al., 2001), and a 750–1300 YBP Maya sample from Copán, Honduras (Merriwether et al., 1997). The two ancient Mexican populations exhibit a high frequency of haplogroup A, like many modern Mesoamerican populations, but the ancient Maya population from Copán belongs primarily to haplogroup C. An AMOVA suggests that temporal groupings (ancient vs. modern) do not account for a significant portion of the variance in population haplogroup frequencies ($-0.91\%$, $P = 0.347$).

Ancient DNA data from this region has been used to test hypotheses about prehistoric interactions and population movements between Mesoamerica and the American Southwest, perhaps associated with the spread of Uto-Aztecan languages. Although linguistic and archaeological evidence suggest cultural ties, modern populations from the two regions exhibit significantly different mtDNA haplogroup frequencies. To investigate whether the same patterns were present before European contact, Kemp et al. (2005) determined the haplogroups (via RFLP) of 23 Aztec individuals from Tlatelolco. This Aztec population was distinct from Southwestern groups but similar to modern Mesoamerican populations (even those from different geographic areas and/or language families), suggesting that the pattern of haplogroup variation in Mesoamerica is quite ancient (Kemp et al., 2005). These aDNA data, like the modern mtDNA data, provide no evidence for a large-scale prehistoric migration from Mesoamerica to the Southwest.

***Caribbean.*** Unlike many other regions of the Americas, the Caribbean shows clear and dramatic differences between the ancient and modern populations that have been sampled. Two pre-Columbian populations have been analyzed—Tainos from the La Caleta site in the Dominican Republic (Lalueza-Fox et al., 2001) and Ciboneys from three sites in Cuba (Lalueza-Fox et al., 2003)—and both are characterized by high frequencies of haplogroups C and D. In contrast, modern Caribbean populations exhibit higher frequencies of haplogroups A and C, as well as many non-Native American haplogroups. An AMOVA (based just on the frequencies of the founding Native American haplogroups) shows that temporal groupings (ancient vs. modern) account for a significant percentage of the variation in population haplogroup frequencies (26.03%, $P < 0.001$). However, a Mantel test does not indicate any significant correlations between genetic and temporal distances ($r = -0.038$, $P = 0.600$). These results suggest that changes in haplogroup frequencies may have occurred fairly abruptly and recently in this region, possibly in association with the decimation of indigenous Caribbean populations following European contact.

***South America.*** Modern indigenous populations of South America have been well characterized molecularly, and have contributed disproportionately to our understanding of the genetic patterns and diversity in Native American populations. It is somewhat surprising, then, that comparatively few aDNA studies have been conducted on prehistoric samples from the continent. Over the past decade, ancient populations from South America that have been studied come from Peru, Argentina, and Chile (Demarchi et al., 2001; Shimada et al., 2004; Moraga et al., 2005; Luciani et al., 2006; Shinoda et al., 2006; Lewis et al., 2007; Kemp et al., 2009; Carnese et al., 2010; Fehren-Schmitz et al., 2010; Fehren-Schmitz et al., 2011). Our comparisons with modern South American populations, therefore, include only modern populations from the same regions of the continent.

When the overall haplogroup frequencies for South America (Table 2) are considered, we see higher frequencies of haplogroup B and more moderate frequencies of haplogroups A, C, and D in both ancient and modern times. An AMOVA indicates that these two groupings do not explain a significant portion of the total genetic variance (1.00%, $P = 0.146$). However, when the South American populations are divided into five temporal groupings (6,000–2,200 YBP, 2,200–1,400 YBP, 1,400–1,000 YBP, 1,000–500 YBP, and modern) to better represent the temporal variation present in the sampled populations, we do see evidence of significant genetic changes over time. The earliest populations generally exhibit higher frequencies of haplogroup D and lower frequencies of haplogroup B compared to population samples from later time periods. An AMOVA shows that these five temporal groupings do account for a significant component of the variation in population haplogroup frequencies (6.05%, $P = 0.016$). A Mantel test also indicates a significant correlation between genetic and temporal distances ($r = 0.436$, $P = 0.010$), suggesting that haplogroup frequencies have changed gradually in South America over the last 3,000–4,000 years. As noted earlier, overall haplogroup diversity appears to have declined in South America from ancient to modern times ($P < 0.001$).

Temporal changes in the genetic composition of South American populations have also been documented in studies of more localized regions of the continent. Demarchi et al. (2001), for example, found no evidence for the 9bp deletion (a discrete marker defining haplogroup B) in any samples from prehistoric Argentina. This result stands in sharp contrast to the high frequency of B found in modern Argentineans, suggesting that evolutionary events (such as the gene flow and depopulation from the expansion of the Inca Empire and Spanish contact) had a substantial effect on the population composition of this region.

In coastal southern Peru, Fehren-Schmitz et al. (2010) tested hypotheses of genetic continuity by analyzing DNA of 217 individuals from the Paracas (2,800-2,200 YBP) and Nasca cultures (2,200-1,400 YBP) of the Palpa region and several sites from the surrounding area (Peninsula, Highlands, Ancash, Arequipa, San Martin, Tayacaja). The authors combined direct sequencing of a 388bp region of HVS I ($N = 104$) with detection of coding region SNPs ($N = 130$) diagnostic for four American haplogroups (A-D). The Paracas populations from the Palpa and Peninsula regions were not statistically different in haplogroup composition from each other, having high frequencies (70–79%) of haplogroup D, moderate frequencies of haplogroup C (14–30%), and no haplogroup A or B. Among the rural and urban Nasca populations (both from the Palpa region), haplogroups A and B appear at low frequencies, D is decreased, and C is correspondingly increased. These haplogroup frequency changes, however, are not substantial enough to make the Nasca significantly different from the Paracas populations, or even the subsequent Middle Horizon population, which the authors interpret as evidence for genetic continuity between the successive cultures (although cautioning that sample sizes are too small to draw any final conclusions). These interpretations are borne out by genetic distance calculations, which are quite low between all of the Paracas period and Nasca period populations. However, the ancient coastal populations they characterized differed significantly in haplogroup composition from the ancient Andean highland populations, suggesting some degree of isolation (although the authors argue against genetic drift on the grounds that both populations were relatively large). Intriguingly, although significantly different in haplogroup frequency composition from modern Peruvian populations, the ancient Palpa populations are genetically similar to modern middle and

southern Chilean populations; a pattern which the authors suggest may date from the earliest peopling of the South American continent.

## DISCUSSION

Although genetic drift, population movements (migration), and other demographic processes have altered the genetic composition of some geographic regions of the Americas over the last few thousand years, the overall geographic structure and diversity patterns that so prominently characterize mtDNA variation in modern Native American populations appear to have been established at least several thousand years ago (O'Rourke et al., 2000b; this study). The time required to establish such continent wide geographic structure has yet to be established under the diverse demographic histories that have been proposed for early American colonizing populations. Effective demographic modeling based on a combination of modern and aDNA results can restrict plausible demographic models that warrant further, detailed examination (e.g., Marchani et al., 2007).

For example, many genetic studies are consistent with the inference that the initial colonization of the Americas was characterized by a dramatic bottleneck resulting in a small number of founders. If this bottleneck occurred before migration and dispersal into the American continents, then the founding haplotypes present in modern populations should be essentially the same as those in the founding populations and in any aDNA samples. However, if the bottleneck occurred as populations were beginning their dispersal throughout the continent, and/or if small population sizes persisted (e.g., serial founder effects, Wang et al., 2007), then the effects of drift over time in uniparentally inherited genomes like mitochondria should have reduced genetic variation (i.e., number of mtDNA lineages) post-colonization. The rate of such a reduction in lineage diversity would depend on population size, gene flow rates between colonizing demes, population growth rates, and the starting frequency of individual haplotypes. Nevertheless, under such a demographic scenario we would expect that a number of lineages present at founding would be lost in subsequent generations because of drift in small, dispersing groups. The aDNA data reviewed here provide no unequivocal evidence for the existence of founding haplotypes that no longer occur in the Americas, though as noted earlier the majority of aDNA data is recent (<6,000 YBP). Among the early aDNA samples studied that predate 8,000 YBP, it is noteworthy that five founding haplotypes (or "haplogroups" in older studies that typed only a limited set of markers) of the 15 currently recognized have been observed, and many others are present in samples dated only slightly later.

The joint analysis of aDNA data and other types of biological variation are also proving illustrative in clarifying evolutionary trends. For example, Perez et al. (2009) used craniometric and ancient DNA data (Lalueza et al., 1997; Garcia-Bour et al., 2004; Figueiro et al., 2007) from central Argentina to investigate the "two components hypothesis" for the peopling of the Americas. This hypothesis posits that modern morphometric variation in native populations is best explained by two successive migrations into the Americas. Perez et al. (2009) determined that while craniofacial metrics resulted in two morphometric clusters, genetically they all shared the same mtDNA haplotypes. This lack of concordance between the morphological discontinuity and genetic continuity in the data was also observed in samples from Tierra del Fuego (Perez et al., 2009). Several explanations for these observations are possible: drift, selection, developmental plasticity, and the uncoupling of shared morphological traits from shared ancestry as a result of environmental lability of morphological traits (Perez et al. 2009).

We caution, however, that reliance on lineage frequency estimates from aDNA studies is fraught with difficulty. Independent of the issues of authenticity that always accompany aDNA analyses, concerns with adequate sample sizes are also of considerable import (Smith et al., 2009). It is clear that our confidence in population history inferences from ancient DNA data require larger sample sizes per archaeological site or the analysis of many more loci than has typically been the case in past studies. In addition, aDNA contributions to our understanding of demographic history will require enhanced geographic as well as temporal sampling. This includes the need to directly date all aDNA samples studied to guard against biased temporal sample selection (Smith et al., 2009), as well as efforts to assure more effective links between genetic results and the archaeological and paleoecological contexts from whence ancient samples come.

As data acquisition strategies improve, we can more effectively exploit the greater resolution of ancient population structure afforded by newer genomic methods (e.g., whole mtDNA or nuclear genome sequencing, large-scale SNP screens, etc.). As the newer genomic methods become more routinely available, and as individual samples, especially very early samples, are more precisely dated, the promise of using aDNA as calibration points for coalescent estimates of lineages identified from modern population studies will be possible (Kemp et al., 2007; Endicott and Ho, 2008; Henn et al., 2008; Ho and Endicott, 2008).

## LITERATURE CITED

Bolnick DA, Smith DG. 2007. Migration and social structure among the Hopewell: evidence from ancient DNA. Am Antiq 72:627–644.

Cabana GS, Hunley K, Kaestle FA. 2008. Population continuity or replacement? A novel computer simulation approach and its application to the Numic expansion (western Great Basin, USA). Am J Phys Anthropol 135:438–447.

Carlyle SW, Parr RL, Hayes MG, O'Rourke DH. 2000. The context of maternal lineages in the Greater Southwest. Am J Phys Anthropol 113:85–101.

Carnese FR, Mendisco F, Keyser C, Dejean CB, Dugooujon J-M, Bravi CM, Ludes B, Crubézy E. 2010. Paleogenetical study of pre-Columbian samples from Pampa Grande (Salta Argentina). Am J Phys Anthropol 141:452–462.

Crawford MH, Rubicz RC, Zlojutro M. 2010. Origins of Aleuts and the genetic structure of populations of the archipelago: molecular and archaeological perspectives. Hum Biol 82:695–718.

Demarchi DA, Panzetta-Dutari GM, Colantonio SE, Marcellino AJ. 2001. Absence of the 9-bp deletion of mitochondrial DNA in pre-Hispanic inhabitants of Argentina. Hum Biol 73:575–582.

Endicott P, Ho SY. 2008. A Bayesian evaluation of human mitochondrial substitution rates. Am J Hum Genet 82:895–902.

Eshleman JA. 2002. Mitochondrial DNA and prehistoric population movements in Western North America. Ph.D. dissertation, University of California, Davis.

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

Fagundes NJR, Kkanitz R, Eckert R, Valls ACS, Bogo MR, Salzano FM, Smith DG, Silva WA, Zago MA, Ribeiro-dos-Santos AK, Santos SEB, Petzl-Erler ML, Bonatto SL. 2008. Mitochondrial population genomics supports a single Pre-Clovis origin with a coastal route for the peopling of the Americas. Am J Hum Gen 82:583–592.

Fehren-Schmitz L, Reindel M, Cagigao ET, Hummel S, Herrmann B. 2010. Pre-Columbian population dynamics in coastal southern Peru: a diachronic investigation of mtDNA patterns in the Palpa region by ancient DNA analysis. Am J Phys Anthropol 141:208–221.

Fehren-Schmitz L, Warnberg O, Reindel M, Seidenberg V, Tomasto-Cagigao E, Isla-Cuadrado J, Hummel S, Herrmann B. 2011. Diachronic investigations of mitochondrial and Y-chromosomal genetic markers in pre-Columbian Andean Highlanders from South Peru. Ann Hum Genet 75:266–283.

Figueiro G, Sans M. 2007. Primeros resultados del análisis de ADN mitochondrial del sitio Arroyo Seco 2. Provincia de Buenos Aires, Argentina. Rev Arg Antrop Biol 9:78.

Garcia-Bour J, Perez-Perez A, Alvarez S, Fernandez A, Lopez-Parra AM, Arroyo-Pardo E, Turbó Daniel. 2004. Early population differentiation in extinct aborigines from Tierra del Fuego-Patagonia: ancient mtDNA sequences and Y-chromosome STR characterization. Am J Phys Anthropol 123:361–370.

Gilbert MTP, Jenkins DL, Götherstrom A, Naveran N, Sanchez JJ, Hofreiter M, Thomsen PF, Binladen J, Higham TFG, Yohe II RM, Parr R, Cummings LS, Willerslev E. 2008a. DNA from pre-Clovis human coprolites in Oregon. North America. Science 320:786.

Gilbert MTP, Kivisild T, Gronnow B, Andersen PK, Metspalu E. Axelsson E, Götherström A, Campos PF, Rasmussen M, Metspalu M, Higham TFG, Schwenninger J-L, Nathan R, De Hoog C-J, Koch A, Møller LN, Andreasen C, Meldgaard M, Villems R, Bendixen C, Willerslev E. 2008b. Paleo-Eskimo mtDNA genome reveals matrilineal discontinuity in Greenland. Science 320:1787–1789.

Goebel T, Waters MR, O'Rourke DH. 2008. The Late Pleistocene dispersal of modern humans in the Americas. Science 319:1497–1502.

González-José R, Bortolini MC, Santos FR, Bonatto SL. 2008. The peopling of America: craniofacial shape variation on a continental scale and its interpretation from an interdisciplinary view. Am J Phys Anthropol 137:175–187.

González-Oliver A, Márquez-Morfín L, Jiménez JC, Torre-Blanco A. 2001. Founding Amerindian mitochondrial DNA lineages in ancient Maya from Xcaret. Quintana Roo. Am J Phys Anthropol 116:230–235.

Halverson MS, Bolnick DA. 2008. An ancient DNA test of a founder effect in Native American ABO blood group frequencies. Am J Phys Anthropol 137:342–247.

Hayes MG. 2002. Paleogenetic assessments of human migration and population replacement in North American arctic prehistory. Ph.D. Dissertation, Salt Lake City: University of Utah.

Hayes MG, Coltrain JB, O'Rourke DH. 2003. Mitochondrial analyses of Dorset, Thule,Sadlermiut, and Aleut skeletal samples from the prehistoric North American Arctic. In: Lynnerup N, Andreasen C, Berglund J, editors. Mummies in a new millennium: Proceedings of the 4th World Congress on Mummy Studies. Copenhagen: Danish Polar Center. p 125–128.

Hayes MG, Coltrain JB, O'Rourke DH. 2005. Molecular archaeology of the Dorset, Thule, and Sadlermiut: ancestor-descendant relationships in eastern North American arctic prehistory. In: Sutherland P, editor. Contributions to the Study of the Dorset Palaeo-Eskimos, Mercury Series, Archaeological Paper 167. Hull, Quebec: Canadian Museum of Civilization.

Henn BM, Gignoux CR, Feldman MW, Mountain JL. 2008. Characterizing the time dependency of human mitochondrial DNA mutation rate estimates. Mol Biol Evol 26:217–230.

Ho SY, Endicott P. 2008. The crucial role of calibration in molecular date estimates for the peopling of the Americas. Am J Hum Genet 83:142–146; author reply 146–147.

Kaestle FA. 1998. Molecular evidence for prehistoric Native American population movement: the Numic expansion. Ph.D. dissertation, Department of Anthropology, University of California, Davis.

Kaestle PA, Horsburgh KA. 2002. Ancient DNA in anthropology: methods, applications, and ethics. Yrbk Phys Anthropol 45:92–130.

Kaestle F, Smith DG. 2001. Ancient Native American DNA from western Nevada: implications for the Numic expansion hypothesis. Am J Phys Anthropol 115:1–12.

Kemp BM, Reséndez A, Berrelleza JAR, Malhi RS, Smith DG. 2005. An analysis of ancient Aztec mtDNA from Tlatelolco: pre-Columbian relations and the spread of Uto-Aztecan. In: Reed DM, editor. Biomolecular archaeology: genetic approaches to the past. Occasional Paper No. 32. Southern Illinois University, Carbondale: Center for Archaeological Investigations. p 22–46.

Kemp BM, Malhi RS, McDonough J, Bolnick DA, Eshleman JA, Rickards O, Martinez-Labarga C, Johnson JR, Lorenz JG, Dixon J, Fifield TE, Heaton TH, Worl R, Smith DG. 2007. Genetic analysis of Early Holocene skeletal remains from Alaska and its implications for the settlement of the Americas. Am J Phys Anthropol 132:605–621.

Kemp BM, Schurr TG. 2010. Ancient and modern genetic variation in the Americas. In: Auerbach BM, editor. Human variation in the Americas. Center for Archaeological Investigations, Occasional Paper No. 38. Southern Illinois University Carbondale. pp. 12–50.

Kemp BM, Tung TA, Summar ML. 2009. Genetic continuity after the collapse of the Wari empire: Mitochondrial DNA profiles from Wari and post-Wari populations in the ancient Andes. Am J Phys Anthropol 140:80–91.

Lalueza C, Perez-Perez A, Prats E, Cornudella L, Turbon D. 1997. Lack of founding Amerindian mitochondrial DNA lineages in extinct aborigines from Tierra de Fuego-Patagonia. Hum Mol Genet 6:41–46.

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

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

LeBlanc SA, Kreisman LSC, Kemp BM, Smiley FE, Carlyle SW, Dhody AN, Benjamin T. 2007. Quids and aprons: ancient DNA from artifacts from the American southwest. J Field Archeol 32:161–176.

Lewis CM, Buikstra JE, Stone AC. 2007. Ancient DNA and genetic continuity in the South Central Andes. Latin Am Antiq 18:145–160.

Luciani S, Fornaciari G, Rickards O, Labarga CM, Rollo F. 2006. Molecular characterization of a pre-Columbian mummy and in situ coprolite. Am J Phys Anthropol 129:620–629.

Malhi RS, Breece KE, Shook BAS, Kaestle FA, Chatters JC, Hackenberger S, Smith DG. 2004. Patterns of mtDNA diversity in northwestern North America. Hum Biol 76:33–54.

Malhi RS, Kemp BM, Eshleman JA, Cybulski J, Smith DG, Cousins S, Harry H. 2007. Mitochondrial haplogroup M discovered in prehistoric North Americans. J Archeol Sci 34:642–648.

Malhi RS, Cybulski JS, Tito RY, Johnson J, Harry H, Dan C. 2010. Brief communication: Mitochondrial haplotype C4c confirmed as a founding genome in the Americas. Am J Phys Anthropol 141:494–497.

Marchani EE, Rogers AR, O'Rourke DH. 2007. The Thule migration: rejecting population histories using computer simulation. Am J Phys Anthropol 131:281–284.

Merriwether DA, Reed DM, Ferrell RE. 1997. Ancient and contemporary mitochondrial DNA variation in the Maya. In: Whittington, SL, Reed DM, editors. Bones of the Maya: studies of ancient skeletons. Washington, DC: Smithsonian Institute Press. p 208–217.

Mills L. 2003. Mitochondrial DNA analysis of the Ohio Hopewell of the Hopewell Mound Group. Ph.D. dissertation. The Ohio State University, Columbus.

Monsalve MV, Stone AC, Lewis CM, Rempel A, Richards M, Straathof D, Devine DV. Molecular Analysis of the Kwäday Dän Ts'ínchi Ancient Remains Found in a Glacier in Canada. Am J Phys Anthropol 119:288–291.

Moraga M, Santoro CM, Standen VG, Carvallo P, Rothhammer F. 2005. Microevolution in prehistoric Andean populations: chronologic mtDNA variation in the desert valleys of northern Chile. Am J Phys Anthropol 127:170–181.

Mulligan CJ, Kitchen A, Miyamoto MM. 2008. Updated three-stage model for the peopling of the Americas. PLoS One.3:e3199.

O'Rourke DH, Hayes MG, Carlyle SW. 2000a. Ancient DNA studies in physical anthropology. Ann Rev Anthropol 29:217–242.

O'Rourke DH, Hayes MG, Carlyle SW. 2000b. Spatial and temporal stability of mtDNA haplogroup frequencies in native North America. Hum Biol 72:15–34.

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

Parr RL, Carlyle SW, O'Rourke DH. 1996. Ancient DNA analysis of Fremont Amerindians of the Great Salt Lake wetlands. Am J Phys Anthropol 99:507–518.

Parr RL. 1998. Molecular genetic analysis of the Great Salt Lake wetlands Fremont. Unpublished Ph.D. dissertation. Department of Anthropology, Salt Lake City: University of Utah.

Perego UA, Achilli A, Angerhofer N, Accetturo M, Pala M, Olivieri A, Kashani BH, Ritchie KH, Scozzari R, Kong Q-P, Myres NM, Salas A, Semino O, Bandelt H-J, Woodward SR, Torroni A. 2009. Distinctive but concomitant Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. Curr Biol 19:1–8.

Perego UA, Angerhofer N, Pala M, Olivieri A, Lancioni H, Kashani BH, Carossa V, Ekins JE, Gómez-Carballa A, Huber G, Zimmermann B, Corach D, Babudri N, Panara F, Myres NM, Parson W, Semino O, Salas A, Woodward SR, Achilli A, Torroni A. 2010. The initial peopling of the Americas: a growing number of founding mitochondrial genomes from Beringia. Genome Res 20:1174–1179.

Perez SI, Bernal V, Gonzalez PN, Sardi M and Politis GG. 2009. Discrepancy between cranial and DNA data of early Americans: implications for American peopling. PLoS ONE 4(5):e5746. doi:10.1371/journal.pone.0005746.

Raff JA. 2008. An ancient DNA perspective on the prehistory of the Lower Illinois Valley. Ph.D. dissertation. Indiana University, Bloomington.

Raff JA, Tackney J, O'Rourke DH. 2010. South from Alaska: a pilot aDNA study of genetic history on the Alaska Peninsula and the Eastern Aleutians. Hum Biol 82:677–693.

Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A. Moltke I, Metspalu M, Metspalu E, Kivisild T, Gupta R, Bertalan M, Nielsen K, Gilbert MTP, Wang Y, Raghavan M, Campos PF, Kamp HM, Wilson AS, Gledhill A, Tridico S, Bunce M, Lorenzen ED, Binladen J, Guo X, Zhao J, Zhang X, Zhang H, Li Z, Chen M, Orlando L, Kristiansen K, Bak M, Tommerup N, Bendixen C, Pierre TL, Grønnow B, Meldgaard M, Andreasen C, Fedorova SA, Osipova LP, Higham TFG, Ramsey CB, Hansen TVO, Nielsen FC, Crawford MH, Brunak S, Sicheritz-Pontén T, Villems R, Nielsen R, Krogh A, Wang J, Willerslev E. 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature 463:757–762.

Rickards O, Martínez-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.

Schroeder KB, Jakobsson M, Crawford MH, Schurr TG, Boca SM, Conrad DF, Tito RY, Osipova LP, Tarskaia LA, Zhadanov SI, Wall JD, Pritchard JK, Malhi RS, Smith DG, Rosenberg NA. 2009. Haplotypic background of a private allele at high frequency in the Americas. Mol Biol Evol 26:995–1016.

Shimada I, Shinoda K, Farnum J, Corruccini R, Watanabe H. 2004. An integrated analysis of pre-Hispanic mortuary practices. Curr Anthropol 45:369–390.

Shinoda K, Adachi N, Guillen S, Shimada I. 2006. Mitochondrial DNA analysis of ancient Peruvian highlanders. Am J Phys Anthropol 131:98–107.

Shook BAS, Smith DG. 2008. Using ancient mtDNA to reconstruct the population history of northeastern North America. Am J Phys Anthropol 137:14–29.

Smith S, Hayes MG, Cabana G, Huff C, Coltrain JB, O'Rourke DH. 2009. Inferring population continuity versus replacement with aDNA: a cautionary tale from the Aleutian Islands. Hum Biol 81:19–38.

Smith DG, Malhi RS, Eshleman JA, Kaestle FA, Kemp BM. 2005. Mitochondrial DNA Haplogroups of Paleoamericans in North America. In: Bonnichsen R, Lepper BT, Stanford D, Waters MR, editors. Paleoamerican origins: beyond Clovis. College Station, TX: Texas A&M University Press. p 243–254.

Snow MH, Durand KR, Smith DG. 2010. Ancestral Puebloan mtDNA in context of the greater southwest. J Archeol Sci 37:1635–1645.

Stone A, Stoneking M. 1998. mtDNA analysis of a prehistoric Oneota population: implications for the peopling of the New World. Am J Hum Gen 62:1153–1170.

Tamm E, Kivisild T, Reidla M, Metspalu M, Smith DG, Mulligan CJ, Bravis CM, Rickards O, Martinez-Labarga C, Khusnutdinova EK, Fedrova SA, Golubenko MV, Stepanov VA, Grubina MA, Zhadanov SI, Ossipova LP, Damba L, Voevoda MI, Dipierri JE, Villems R, Malhi RS. 2007. Beringian standstill and spread of Native American founders. PLoS ONE 2:e829.

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

Wang S, Lewis CM, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W, Parra MV, Molina JA, Gallo C, Mazzotti G, Poletti G, Hill K, Hurtado AM, Labuda D, Klitz W, Barrantes R, Bortolini MC, Salzano FM, Petzl-Erler ML, Tsuneto LT, Llop E, Rothhammer F, Excoffier L, Feldman MW, Rosenberg NA, Ruiz-Linares A. 2007. Genetic variation and population structure in Native Americans. PLoS Genet 3:e185.

Yang NN, Mazières S, Bravi C, Ray N, Wang S, Burley M-W, Bedoya G, Rojas W, Parra MV, Molina JA, Gallo C, Poletti G, Hill K, Hurtado AM, Petzl-Erler ML, Tsuneto LT, Klintz W, Barrantes R, Llop E, Rothhammer F, Labuda D, Salzano FM, Bortolini M-C, Excoffier L, Dugoujon JM, Ruiz-Linares A. 2010. Contrasting patterns of nuclear and mtDNA diversity in Native American populations. Ann Hum Genet 74:525–538.