

J O H N R. J O H N S O N A N D

J O S E P H G. L O R E N Z

GENETICS AND THE *CASTAS* OF COLONIAL CALIFORNIA

Cultural anthropologists have long been interested in the ways in which social groups maintain or lose their identities as they interact with other groups. In recent decades, they have increasingly come to believe that cultural identities are *socially constructed*. In other words, how a group's membership identifies itself and how others distinguish it is more important than a common biological inheritance from a geographically specific population.¹ Colonial society in Alta California had its origins in the multiracial society that had been forged in the northwest frontier of New Spain. While a variety of *castas* (classifications based on degree of miscegenation) were commonly used in colonial Mexico, these terms were often fluid, depending on how individuals identified themselves or how others in society perceived them. Mission sacramental registers of baptisms, marriages, and burials reflect the important distinction that developed between indigenous groups (*indios*) and Spanish Californians, called *gente de razón* no matter what biological ancestry they possessed. Thus, in colonial California, as long as a person could achieve the cultural status of being *de razón*, biological inheritance became unimportant in defining social identity.² While we recognize that the social construction of identity was paramount in Spanish California, we nonetheless find it of considerable anthropological and historical interest to understand how far along the process of *mestizaje*, the genetic mixing that occurred during colonization, had proceeded prior to the emigration of the Californio population from northwest Mexico. To pursue this line of inquiry, this study uses emerging techniques of genetic genealogy that have led scholars to return to older foundational debates about the origins of social, cultural, and racial identities in Spanish California.

The subject of genetic genealogy has recently engaged the public's imagination. Bestselling books such as *The Seven Daughters of Eve*, *Out*

of Eden, and *Deep Ancestry* speak to those eager to understand their place in the human family tree.³ The PBS documentary series *African American Lives*, hosted by Harvard professor Henry Louis Gates Jr., demonstrates how genetics has been used to trace the ancestry of a number of prominent African Americans, including Oprah Winfrey, Whoopi Goldberg, and Chris Tucker.⁴ Commercial labs in Britain and the United States vie to capitalize on the desire to determine one's own genetic origins. All of these developments are based on a relatively new field of molecular anthropology that goes beyond studies of comparative craniometrics and geographical patterns of blood groups to explore the ancestry of world populations based on detailed studies of mitochondrial, Y-chromosome, and autosomal DNA.

The interests of anthropologists, historians, and members of the general public converge in discovering what this emerging science can contribute to an understanding of the past of peoples whose ancestry remains clouded by virtue of myth and a partial documentary record. In this study, we take the first steps toward understanding how mitochondrial DNA evidence, used in combination with mission record research, clarifies the origins of Spanish-Mexican families who came to the region before 1790, the year Spanish officials undertook the first census of Alta California. A number of descendants of early California Spanish families participated in this study, as did other Mexican Americans. These two subgroups, the former with relatively well-documented ancestry and the second with ancestry that reflects their more recent immigration, can be compared and contrasted to produce network diagrams illustrating phylogenetic (ancestral) relationships. Our eventual goal is to place the mitochondrial DNA lineages that are of indigenous origin into an ethno-historical and geographical context, shedding light on peoples of north-west Mexico, especially Sonora and Sinaloa, whose population histories are poorly known.

The Castas of Colonial California

In 1813, Spanish officials circulated a questionnaire among the missionaries of Alta California, inquiring about the customs of the native inhabitants of the region. Although the majority of the questions pertained solely to the characteristics of California Indian peoples, the first question in particular elicited responses from many missionaries as to the colonial population of settlers and soldiers that had become established in the new territory since the first expedition of 1769 led by Gaspar de Portolá:

Question 1. Let them state into how many castes the population is divided: for example, whether they are Americans, Europeans, Indians, Mulattos, Negros, etc., omitting no group whatsoever.⁵

The response from Mission Santa Barbara is illustrative of the nature of cultural identity in the region:

It cannot be known with certainty into how many castes the inhabitants of the presidio adjoining this mission are divided. Although it is very well known that not all are genuine Spaniards either of European or American origin, yet at least they regard themselves as such. Furthermore if they were told to the contrary they would consider it an affront, so we are unable to state anything further on this particular matter.⁶

The missionaries at San Gabriel made a similar statement about the *pobladores* of Los Angeles and other Spanish-Mexican families under their ecclesiastical jurisdiction:

The general and distinctive population of this mission are only Indians. However . . . we have to attend to a town of whites who are known as *gente de razón*. In civil matters it is the headquarters of four ranchos composed of the same class of people. This mission also cares for another rancho belonging to the civil jurisdiction of Presidio San Diego, as well as for the families of the six soldiers who guard this mission. All the individuals who live in these places amount to 526. They are of various castes. How many these castes are and precisely which castes they are we do not know because, as we have said, they are all known as *gente de razón*.⁷

In response to a follow-up question, the San Gabriel missionaries articulated the prevailing myth regarding the homeland of the native Mexican peoples as being in the north and provided further details regarding the *gente de razón* families of their area:

In regard to the origin of our [California] Indian people we think that during the migration of the ancient Mexicans, various families in the course of their passage remained behind, . . . With regard to the people of the other class [gente de razón], it seems to us they have the same origin as those of New Spain, for the first settlers came from the provinces of Sonora, Sinaloa and Nueva Vizcaya. Some of them, however, both Spaniards and men of other castes, have consorted with Indian women of this peninsula.⁸

These responses make it clear that Alta California's colonial population ultimately derived from people of mixed ancestry who came largely from northwest Mexico. The comments of the San Gabriel missionaries in particular indicate that many of the settlers and soldiers were regarded as having considerable indigenous genetic inheritance.

Nearly a quarter of a century before the questionnaire, Spanish officials had undertaken a regionwide census of gente de razón families that recorded the caste designations in use among Alta California's pioneer families. The census of 1790 tabulated each individual and gave his or her occupation, approximate age, and caste designation, as well as the town of origin for the head of each household.⁹ This census, combined with the clues found in the various mission registers of baptisms, confirmations, marriages, and burials, provides an excellent beginning for studying the genetic origins of the Alta California colonists and comparing their public identities with what is revealed through DNA analysis.

Some Comments on DNA Analysis

Although DNA analysis has much to contribute to the study of population origins, genetics do not necessarily coincide with the ways in which people identify themselves. For instance, people who speak different languages often do reflect separate population histories, but a group's linguistic affiliation does not necessarily correspond to its genetic origins. Recent studies show that certain DNA lineages can crosscut language groupings in the Pre-Columbian Americas because of prehistoric patterns of intergroup marriage and cultural dominance of one group over another. Of course, examples of these processes can be seen following European contact, when miscegenation certainly occurred and the languages of Spanish, English, French, and Portuguese replaced many of the indigenous ones, even though descendants of the original populations

remained. Genetic patterns must be carefully compared to ethnolinguistic groupings in order to sort out what kinds of meanings underlie correlations between genes and languages.

Among the several types of DNA studies of Native American descendants, those based on mitochondrial DNA (mtDNA) have elucidated patterns among indigenous groups in the greater Southwest region, including California and northwest Mexico. Mitochondrial DNA is especially useful because it is extra-nuclear in origin and inherited only from one's mother. Thus it is unaffected by the mixing of chromosomes that results from mestizaje.¹⁰ Mitochondrial DNA exists in abundance, with hundreds of copies present in each cell. Its genome is relatively short, consisting of 16,569 nucleotide base pairs, composed of adenine, guanine, cytosine, and thymine. Each mtDNA molecule contains genes that provide instructions for the energy-producing functions of the cells. Also present are segments that do not code for metabolic activities; these can accumulate occasional random mutations in female lineages, rendering them invaluable for determining phylogenetic relationships among the world's populations. Because of its abundance, its short molecular length, and the variability of certain non-coding segments, mtDNA is relatively easy to extract and amplify, which increases its usefulness for genetic analysis.¹¹

Mutations that occurred in the distant past have given rise to broad groups (called *haplogroups* or *clades*) of mtDNA lineages that allow molecular anthropologists to reconstruct the interrelationships and migrations of the world's populations. Haplogroups are divided into daughter groups, herein called *lineages*, that reflect more recent mutations; any specific mtDNA sequence within a lineage is called a *haplotype*. Mitochondrial lineages indigenous to the Americas belong to one of five ancestral haplogroups: A, B, C, D, and X; these are also found in Asia.¹² Haplogroup X is almost entirely absent in the Californias and northwest Mexico and is not considered in this study.

Mitochondrial DNA can be analyzed in several different ways: (1) RFLP (restriction fragment length polymorphism) analysis determines which haplogroup a sample belongs to, (2) nucleotide-sequence analysis determines specific mutations that occur within one or two lengthy hypervariable segments in the mitochondrial DNA molecule's *control region*, and (3) complete genome sequencing detects additional mutations. Most of the analysis of this study uses the first two approaches because complete genome sequencing of mitochondrial DNA is only recently becoming widely reported for Native Americans.

Distributional analyses of mtDNA lineages found among Native American populations began with the work of Douglas Wallace and his associates at Emory University. Their studies demonstrated relationships between mtDNA types found among American Indian peoples and those of northeast Asia, and it determined that a non-random distribution of the four predominant mtDNA haplogroups (A, B, C, and D) existed throughout the Americas.¹³ Further studies have built on these findings, showing that specific mtDNA lineages arose among geographically separated tribal groups following the migrations that initially peopled the two continents of the Americas. Those mtDNA lineages that were determined to be ancestral to more specific haplotypes within each haplogroup were more widespread. Among the most ancient haplotypes were those considered to be founding lineages among the original migrants who spread southward from the Bering land bridge around 14,000 years ago or earlier, once pathways became available through previously glaciated areas of Alaska and Canada.¹⁴ The latest evidence suggests that a coastal migration route would have been accessible somewhat earlier than an ice-free interior passage east of the Rocky Mountains.¹⁵

Some individuals express the hope that DNA will shed light on their genetic heritage so that they can determine with certainty the tribal affiliations of their Native American ancestors. Others worry that DNA evidence will be used to exclude people from membership in particular federally recognized tribes if their DNA sequences do not align with others who are known tribal members. Yet, those who harbor these hopes and concerns attribute more to molecular evidence than it can presently deliver. While particular genetic lineages do appear to correlate with ethnolinguistic groupings to some degree, the reverse is not true. Speakers of a given language are typically not confined to specific genetic lineages. This is because the older, non-distinctive haplotypes within each haplogroup can be found among many widely separated populations, and because prehistoric patterns of migration and intermarriage have resulted in the sharing of mtDNA haplotypes among populations that come from different ethnolinguistic backgrounds.

Genetic Research in Northwest New Spain

Bearing in mind the potential contributions and limitations of genetic research, we can turn to the topic of mtDNA variation among indigenous populations that inhabited the greater American Southwest during the colonial period. Studies of indigenous groups in the Spanish Borderlands have revealed that populations characterized by Haplogroup B and

C lineages dominated much of the region. To the south, in central and southern Mexico, Haplogroup A lineages predominate. The indigenous peoples of the Central Valley of California and portions of the adjacent Great Basin include a considerable component descended from one of the founding types of Haplogroup D, as well as Haplogroup B. The Chumash peoples of south central California are characterized largely by Haplogroup A and a different founding haplotype from Haplogroup D. The Apache and Navajo, who speak an Athabaskan language, demonstrate their northern origins by the widespread presence of Haplogroup A lineages clearly related to the Na-Dene peoples of Alaska and Canada.¹⁶

These differences in the mtDNA haplogroup distributions among indigenous peoples of northwestern New Spain are derived from separate population histories. Isolated ethnolinguistic groups, such as the Chumash peoples of the Santa Barbara Channel region, the Zuni of New Mexico, and the Seri of the Sonoran coast, harbor distinctive mtDNA lineages that bespeak their longtime presence in those areas. The Yokuts and Miwok peoples of Alta California's Central Valley appear to have expanded across the Sierra Nevada from the Great Basin region many millennia ago, bringing with them genetic lineages characteristic of their homeland. With the exception of the Apache and Navajo, already mentioned, the remainder of the greater American Southwest is dominated by peoples speaking languages in the Uto-Aztecan and Yuman-Cochimí linguistic families.¹⁷ Some researchers have argued that the spread of these two language families within the Greater Southwest correlates with population growth associated with the adoption of agriculture and subsequent budding off of subgroups into adjacent areas.¹⁸ Brian Kemp, a molecular anthropologist who has considered the genetic implications of this hypothesis, discovered a distinctive genetic marker within a major mtDNA Haplogroup B subgroup that appears to be shared among many peoples in the Greater American Southwest. With this clue, he has dated the main spread of Uto-Aztecan and Yuman-Cochimí populations to sometime between 1,300 and 3,800 years ago, using a pedigree-based rate of mtDNA evolution.¹⁹

Based largely on the genetics of contemporary descendants, mtDNA research has thus revealed some important clues pertaining to the pre-history of population movements within the northwest region of the Spanish Borderlands following the conquest of Mexico in 1521. This picture is incomplete, however, because some groups have not been previously sampled and others are poorly known, having been displaced or absorbed into the expanding colonial state.²⁰ In northwest and west Mexico, only

certain populations have received attention from mtDNA investigative surveys: the aforementioned Seri, the Tarahumara of Sonora, the Cora and Huichol of Jalisco, and inhabitants of Ciudad Juárez and Ojinaga, who live along the Rio Grande.²¹ Uto-Aztecan peoples of Sonora and Sinaloa have not yet been studied genetically. Native peoples of certain portions of Sinaloa and Baja California are almost completely lacking with regard to evidence of linguistic affiliation, much less molecular genetics.²² Sonora and Sinaloa are precisely those regions from which the majority of the colonial families immigrated to the Californias during the seventeenth and eighteenth centuries.²³ Presuming that an underlying layer of indigenous Sonoran and Sinaloan mtDNA existed among these early settlers, a study of the Alta California colonial lineages has the potential to make a significant contribution to our understanding of the native peoples who inhabited the north frontier of New Spain. Our research represents the first effort to tackle this problem and will lay the groundwork for future studies in the region.

Sample Descriptions

Our sampling of Alta California's Spanish-Mexican population began entirely by accident, as a byproduct of research into the genetic lineages of California Indians.²⁴ One of the women whom we had believed to be descended from a female Native American ancestor baptized at Mission San Gabriel proved otherwise, once we had fully reconstructed her genealogy.²⁵ Instead, this individual could be traced back through her maternal line to a woman named María Rufina Hernández, who had been born in Sinaloa or Loreto, Baja California, according to different sources.²⁶ As our study continued, several other individuals who were uncertain as to whether they had California Indian or Spanish-Mexican ancestry provided us with samples. Further genealogical investigations demonstrated that some of these women could be traced back to women who were part of Alta California's colonial population. Soon we began expanding our study to include mtDNA samples from lineages descended from colonial California women.

In January 2005, we presented a lecture about mtDNA research to a gathering of Los Californianos, a group descended from the earliest colonists of Alta California. This resulted in additional volunteers for our study. A larger sample of descendants from northwestern Mexico was desirable, to compare with the accumulating colonial California mtDNA sequences, so we also sought Mexican American volunteers. Many of these were husbands or wives of participants in our California Indian

study. Others were acquaintances who agreed to be involved once they learned of our research. Finally, a Southern California organization called Familia Ancestral Research Association twice invited us to lecture about mtDNA research, and many of its members contributed mtDNA samples. This group actively sought information in Mexican archives in order to reconstruct its members' genealogies, thus assisting in the determination of the origin of many female lineages, sometimes tracing them back two centuries or more. Most Familia members were descended from northwestern Mexican immigrants from Sonora, Chihuahua, and Durango.

All who contributed samples were interviewed about their family genealogies. Most could provide the names and birthplaces of their maternal grandmothers; some were able to consult family records or other relatives to trace the lineages further. For about half of those forebears who could be demonstrated to have lived in Alta California prior to 1850, it was possible to use mission records to trace their earliest female ancestors who had immigrated during the colonial period. Most of these were listed in the census of 1790, so their caste descriptions are known, and most had their place of birth mentioned in contemporary documents or ecclesiastical records.²⁷ If two individuals had the same ancestor through their direct female lines, then only one sample was used in our analysis, to avoid duplication. Figure 1 illustrates the places of origin in Mexico, Central America, and the Spanish Borderlands of 115 female ancestors out of a total of 179 descendants who contributed mtDNA samples for analysis.²⁸ This total of 115 ancestors includes only those who were determined to possess Native American mtDNAs and for whom origins could be traced. The participants in our study included: (1) those who were "California Spanish" descendants, defined as those whose direct female ancestors were part of the old Californio population, (2) those who were from elsewhere in the Hispanic Southwest (that is, Arizona, New Mexico, or Texas), and (3) those of more recent Mexican or Central American immigrant backgrounds.

Mitochondrial Haplogroup Distributions

Table 1 presents the general categories for 179 samples based on ethnic attributions or regions of origin and provides the numbers for each that were assignable to the four principal mtDNA haplogroups found among Native Americans, as well as those that were European, African, or Asian in origin. Figure 2 illustrates these results in terms of percentages. More than 80 percent of the samples were from mitochondrial haplogroups



Figure 1. mtDNA samples determined to be of Native American derivation traceable to particular regions of the Hispanic Southwest, Mexico, and Central America. This map excludes thirty individuals who were of European or African mtDNA affiliation and thirty-four others whose ancestry in Mexico was unknown.

indigenous to the Americas, illustrating that the process of *mestizaje* has been predominantly characterized since the advent of the colonial period by men of European, African, or Asian origin fathering children with women of Mexican Indian ancestry. Women from the Old World did not emigrate in great numbers until years after an original *entrada* had taken place, so families produced by non-indigenous men and native wives were commonplace in frontier settings. This pattern is not unique to the Californias and has been documented in many other Latin American genetic studies. Thus, it is twenty times more likely that indigenous mitochondrial lineages will survive over indigenous Y-chromosome lineages in many areas of Hispanic America, especially in regions that were of low population density at the time of initial colonization. Given the small size of many indigenous groups and their lack of immunity to introduced European diseases, children of mixed marriages would be more likely to have genetic resistance to new pathogens, thus increasing their chances of survival. In this way an initial pattern of interracial mating skewed toward native women resulted in a much greater likelihood of continu-

Table 1. Distribution of Mitochondrial DNA Haplogroups according to Regional or Ethnic Affiliation

Regional or Ethnic Affiliation	Haplogroup A	Haplogroup B	Haplogroup C	Haplogroup D	Other Haplogroup	Total
Spanish Californian	16	10	7	9	11	53
Spanish Californian or Mexican American	5	1	2	0	0	8
Mexican American	19	17	25	5	18	84
Hispanic Southwest	9	4	8	1	1	23
Mexican Indian						
Cocopa	1	0	0	0	0	1
Huichol	0	0	1	0	0	1
Nahua	1	0	0	0	0	1
Opata	0	1	0	0	0	1
Tarahumara	0	0	1	0	0	1
Yaqui	2	1	2	0	0	5
Central America	1	0	0	0	0	1
Total	54	34	46	15	30	179

ation of indigenous mitochondrial lineages as compared to indigenous Y-chromosome inheritance.

Among sixty-one probable and suspected California Spanish mtDNAs shown in table 1, Haplogroup A represents about a third of the total, with the remainder split fairly evenly among Haplogroups B, C, D, and Other. It is interesting to note that a higher percentage of samples for colonial California appear to be derived from Haplogroup D than for other regions, a point that we discuss in greater detail below.

Table 2 breaks down the sample according to geographic origin documented for traceable female ancestors, pooling the results of the genealogical investigations of Alta California's colonial population with those derived from more recent Mexican American immigrants. As the data reported simply as "Alta California" show, we thus far have been unable to trace family trees of twenty-six individuals with indigenous Native American mtDNAs who appear to have descended from the early colonial population. Sometimes a female ancestor's birthplace in Mexico could not be determined because of inadequate genealogical clues in surviving documents. For the seventy-two samples traceable to northwestern Mexico, Haplogroups A, B, and C are all present in relatively equal proportions, whereas samples belonging to Haplogroup D are comparatively rare. Even though Haplogroup D appears to be less present in Sonora and Sinaloa than other mitochondrial haplogroups, its almost

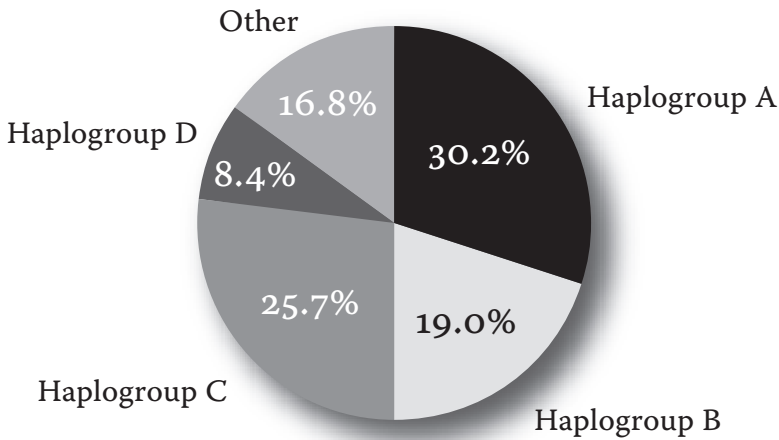


Figure 2. California Spanish and Mexican American Sample (N=179)

total absence elsewhere, except for among Alta California's colonial population, is perhaps significant. The surviving Haplogroup D lineages from northwest Mexico may be informing us about the genetic ancestry of otherwise poorly documented indigenous groups from this region.

Table 3 presents the results of the haplogroup assignments of thirty-five colonial California women listed in the 1790 census or otherwise determined to have resided in Alta California during the late eighteenth century, including five native Californian women who had married soldiers.²⁹ Table 3 does not include all samples that were classed as "California Spanish" because of problems encountered in tracing some genealogies (as mentioned previously). Nor does inclusion in this table necessarily mean that we have an associated mtDNA sequence reported in tables 4–7 (for reasons cited at the bottom of table 3). Of the thirty Spanish-Mexican women who had immigrated to Alta California, all but four derived from mtDNA lineages indigenous to the Americas. Of the four that came from the Old World, three were European (Haplogroups H, U5, and V) and one was African (Haplogroup L2). The overall predominance of Native American mtDNAs among Alta California's *gente de razón* population is quite at odds with the proportion of women in table 3 whose castas were reported in the 1790 census as *española* (40 percent), as opposed to such categories as *india* (nine individuals), *coyota* (three), *mestiza* (six), or *mulata* (one).³⁰

The overall distribution of mitochondrial haplogroups, as presented in tables 1–3, suggests that Sonora and Sinaloa were transitional areas

Table 2. Geographic Origins of California Spanish and Mexican American Mitochondrial DNA Samples^a

Origin	Haplogroup A	Haplogroup B	Haplogroup C	Haplogroup D	Total
<i>Spanish Borderlands (Origin in Mexico not established)</i>					
Alta California	11	5	4	6	26
Arizona	3	0	0	0	3
New Mexico	3	4	4	0	11
Texas	3	0	4	1	8
Borderland Totals	20	9	12	7	48
<i>Northwest Mexico</i>					
Sonora	9	7	8	3	27
Sinaloa	6	3	3	2	14
Chihuahua	5	5	6	3	19
Durango	0	4	1	0	5
Baja California Sur	1	1	2	0	4
Baja California (Norte)	2	0	1	0	3
Northwest Mexico Totals	23	20	21	8	72
<i>Other, Mostly of Mexican Origin</i>					
Coahuila	1	0	1	0	2
Zacatecas	2	2	3	0	7
Guanajuato	2	0	3	0	5
Jalisco	0	0	2	0	2
Michoacán	0	1	1	0	2
Other ^b and Unknown	6	2	3	0	11
Other Totals	11	5	13	0	29
<i>Total Mexican American Sample of Indigenous Origin</i>					
	54	34	46	15	149

^a This table includes only those samples for which sequences were obtained that belonged to mitochondrial haplogroups indigenous to the Americas, as well as one Polynesian sample belonging to Haplogroup B. Thirty samples listed in Table 1 that were determined to be from European or African mitochondrial haplogroups do not appear here.

^b Among the samples included as "Other" is one from Guatemala and one originally Hawaiian in origin, provided by an individual of partial Hawaiian descent, whose female ancestor married a California Spanish man during the colonial period.

between the central Mexican pattern of populations characterized by high percentages of Haplogroup A and populations of the American Southwest where Haplogroups B and C predominated. But to what degree were some of these mitochondrial patterns the result of the movement of colonial families of partial indigenous ancestry from central Mexico following the conquest of indigenous tribes in northwestern New Spain?

Certainly there was a dramatic population decline of tribes living in the region due to the depredations of Nuño de Guzmán in 1530–31 and continual, successive epidemics that spread through the province for many decades afterward.³¹ Later, the viceroy of New Spain encouraged colonies of Tlaxcaltecs from central Mexico to move into the area to serve as civilized examples for newly conquered peoples.³² In order to ascertain which mitochondrial DNA lineages might be the result of incoming settlers as opposed to original inhabitants, it is necessary to look deeper than haplogroup affiliations and determine original haplotypes. Those haplotypes present within our sample can be compared to those identified in a larger sample of Mexican and American Southwest mtDNA lineages in order to differentiate which may have been indigenous to Sonora and Sinaloa and which likely came from elsewhere.³³

Explorations of Deep Ancestry

Within each mitochondrial haplogroup that was indigenous to the Americas, it is useful to depict ancestral relationships among specific haplotypes by means of network diagrams. Given a large enough sample, such diagrams typically exhibit a star-like pattern, with the haplotypes of the American Indian founding mitochondrial lineages in the center of the star. Subsequent mutations define daughter lineages, which in turn spawned additional daughter lineages. If a branching chain of daughter lineages is specific to a particular ethnolinguistic group, an ancient geographic presence of that group in a particular region is implied because it takes time for haplotypes to become fixed within a population. An example of the branching-chain pattern occurs among the Chumash Indians of the Santa Barbara Channel region, who exhibit five haplotypes within Haplogroup A all characterized by a distinctive mutation at nucleotide position (np) 16093 in the mtDNA molecule.³⁴ A population must be established in a region over many millennia to “age” in this manner.

Detection of related haplotypes among disparate groups can also provide clues regarding ancient connections that have been disrupted by subsequent population migrations. An example of this pattern is observable in a rare haplotype of Haplogroup A—characterized by distinctive mutations at np 16257 and np 16263—that is present here and there among the Zuni and certain prehistoric samples from the American Southwest.³⁵ Despite being separated by hundreds of miles of desert, another haplotype containing these same two markers existed among the Ventureño Chumash.³⁶ Because the Chumash language family and the Zuni language each represent linguistic isolates with ancient presence in

Table 3. *Castas* and Mitochondrial Haplogroups of Early Colonial California Women^a

Name	<i>Casta</i> (1790 census)	Place of Birth	mtDNA Haplogroup
Alvarado, María Ignacia	Española	Loreto, Baja California	H
Beltrán, María Nicolasa	Española	Horcasitas, Sonora	B
Bojórquez, María Antonia	Española	Horcasitas, Sonora	L2
Carrillo, María Ignacia	Española	Loreto, Baja California	U5
Carrillo, María Matilde	Española	Loreto, Baja California	V
Domínguez, Ursula	India	Mission Santa Gertrudis, Baja California	C
Espinosa, Gertrudis Gregoria	Española	Villa Sinaloa, Sinaloa	B
Félix, Marcelina	Española	Cosalá, Sinaloa	C
Gonzales, Ana María	India	Villa Sinaloa, Sinaloa	B
Gutiérrez, María de los Santos	India	Culiacán, Sinaloa	A
Hernández, María Rufina	N/A	Loreto, Baja California	C
López, María Rosa (aka Monreal)	Española	Álamos, Sonora ^b	D
Lugo, María Pascuala	Mestiza	Villa Sinaloa, Sinaloa	C
Lugo, Rosa María	Española	Villa Sinaloa, Sinaloa	A
Martínez, Juana María Rita	N/A	Villa Sinaloa, Sinaloa	A
Noriega, María Ramona	Española	Loreto, Baja California	A
Ochoa, María Manuela	Mestiza	Villa Sinaloa, Sinaloa	B
Parra, Juana Paula	Mestiza	Santa Cruz de Mayo, Sonora ^b	A
Perez, María Guadalupe	Coyota	Rosario, Sinaloa	B
Piñuelas, María Josefa	Coyota	Villa Sinaloa, Sinaloa	C
Ramírez, Potenciana	Coyota	Villa Sinaloa, Sinaloa	A
Redondo, María Antonia	Española	Villa Sinaloa, Sinaloa	A
Rivera, María Ignacia (aka Velarde)	Mestiza	Maquipa, Santa Cruz de Mayo, Sonora ^b	A
Romero, María Francisca	Mestiza	Villa Sinaloa, Sinaloa	D
Rubio, Petra	Mulata	Álamos, Sonora ^b	A
Salgado, Lugarda	Mestiza	Loreto, Baja California	C
Sánchez, María Gertrudis (aka Límon)	Española	Villa Sinaloa, Sinaloa	B
Sotelo, Micaela	Española	Villa Sinaloa, Sinaloa	A
Tapia, María Antonia	India	Culiacán, Sinaloa	C
Valencia, María Dolores	Española	Horcasitas, Sonora	D
Irene, wife of Julián Ríos	India	Tipu (Obispeño Chumash)	D
Juana María, wife of Hilario Jiménez	India	Siujtun (Barbareño Chumash)	D
Margarita, wife of Juan María Ruiz	India	Lamaca (Antoniano Salinan)	A
María Ildelfonsa Bergas	India	Lima (Antoniano Salinan)	D
Regina Josefa Toypurina, wife of Manuel Montero	India	Japchivit (Gabrielino)	C

^aThis table includes haplogroup assignments for four samples typed by RFLP analysis that have not been reported in other tables or figures because precise mtDNA sequence information is still pending, as well as two additional samples traced to Alta California ancestors that were typed by Joseph Donohoe, a Californio descendant who has conducted genetic genealogical research.

^bÁlamos and Santa Cruz de Mayo, while currently within the state of Sonora, were formerly included in the province of Sinaloa at the time of the recruitment of settlers for Alta California. Women from these towns have been included in the totals for Sonora in Figure 1 and Table 2.

their respective regions, the occurrence of closely related haplotypes may well represent prehistoric contact predating the development of agriculture and subsequent expansion of mitochondrial lineages in Haplogroups B and C in the American Southwest sometime around 3,000 years ago.

To search for such meaningful patterns of relationships, we produced network diagrams for each mitochondrial haplogroup within 149 samples of indigenous origin analyzed for this study (see figures 3–6). Tables 4–7 present the mtDNA sequence data that were used to generate these diagrams. The designations next to the circles represent the arbitrarily numbered haplotypes listed in tables 4–7. The size of the circles within each network diagram varies according to the number of samples belonging to each haplotype, and the numbers between two circles refer to the molecular markers that distinguish between them. Those samples that were categorized as Spanish Californian are indicated in solid portions of the circles in these diagrams. To explore the indigenous origin of the specific mtDNA haplotypes occurring among the Spanish Californians, we undertook a close investigation for each of those women listed in the 1790 census whose origin in Sonora or Sinaloa was documented.

Haplogroup A Lineages

The Haplogroup A diagram (figure 3) represents the typical Native American pattern of daughter lineages radiating from a central haplotype (here designated as A₃₀) that represents the principal founding lineage for this mtDNA haplogroup. Six additional mutations are present among one particular daughter lineage (A₀₆) of this ancestral haplotype, but most haplotypes are distinguished by only one or two mutations. The network diagram of our Haplogroup A samples contains some reticulation (indicating multiple pathways that may have produced a haplotype) because of the tendency for certain nucleotide positions to be susceptible to frequent mutations. Two of these positions, np 16111 and 16325, are recognized “hot spots” in the mtDNA genome and have mutated more than once within Haplogroup A, accounting for all instances of reticulation apparent in figure 3.³⁷

The fifty-four mtDNA samples belonging to Haplogroup A within our study were sorted into thirty-six haplotypes. The two most common were the founding haplotype (A₃₀), indicated by the central node in figure 3, and Haplotype A₂₆ (six instances), which was distinguished from the founding haplotype by having a mutation at np 16092. Haplotype A₂₆ occurred commonly among our Spanish Californian samples and one individual who had descended from a Yaqui woman. One of the Spanish Californian samples was traceable back to Petra Rubio, who was listed

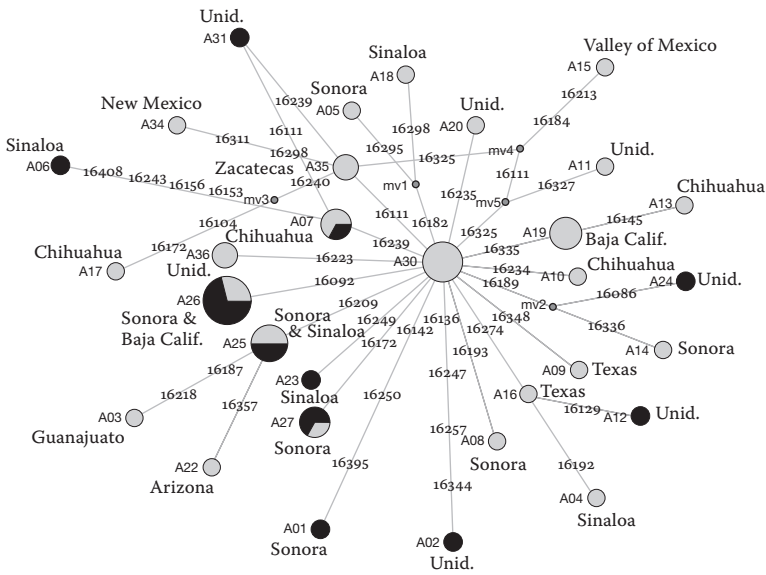


Figure 3. Haplogroup A Network Diagram Showing Ancestral Relationships among Spanish Californians and Mexican Americans

as a mulata from Álamos in the 1790 census. This area, now in the state of Sonora, was inhabited originally by speakers of Cahita, who occupied a considerable portion of the coastal plain and river valleys from the Yaqui River south to the Sinaloa River. Because of its common presence among the Spanish Californian descendants, as well as its occurrence among the Yaqui, this haplotype is hypothesized to have been indigenous to the Cahitan area. Comparisons with Kemp's larger study of mtDNAs show variations on this haplotype to be quite common among the linguistically related Cora Indians further south in Jalisco, but its virtual absence elsewhere supports the hypothesis that this haplotype derives from the original native populations of northwestern and western Mexico.

Two of the Spanish California mestiza women belonging to Haplogroup A were listed as being from Santa Cruz del Mayo, a pueblo formed around a former Jesuit mission in the lower Mayo Valley (see table 3). Indeed, these two women appear to have been the only two listed in the entire 1790 census as having been born at or near this pueblo.³⁸ One woman was specified as being from Maquipa in the jurisdiction of Santa Cruz del Mayo at the time she was married; but she was associated with the "Pueblo de Guasabe" in an entry in the Santa Barbara presidio's baptismal register.³⁹ If the identification of these two women as mestizas

is accurate, then they were of half-Indian ancestry, which further suggests that their mothers were indigenous to the area. One would have presumed that both women were Mayo, because Santa Cruz is located along the Mayo River; however, the direct statement that one woman, María Ignacia Rivera, was born in the Guasabe pueblo indicates that her maternal origin may have come from the Guasave Indians, a fishing people of undetermined linguistic affiliation who inhabited the coastline to the south of the Mayo.⁴⁰ This identification acquires significance because this woman's mitochondrial DNA markers are so distinctive, three mutations removed from the central node of the Haplogroup A diagram (Haplotype A01 in figure 3). Not a single one of these three mutations was found among 100 haplotypes belonging to Haplogroup A in Kemp's larger study of Mexican and American Southwest Indian mitochondrial lineages.⁴¹ Although a single sample cannot be considered representative, it is noteworthy nonetheless that the only Guasave mtDNA sample is so atypical, implying that this linguistically distinct, non-agricultural, maritime-oriented group was of a different origin from their neighbors who spoke Uto-Aztecan languages.

The second woman from Santa Cruz del Mayo, Juana Paula Parra, shared a haplotype (A27) with two other individuals in our sample. One of these was also of California Spanish descent and conceivably could be from the same female ancestor, since full genealogical investigation was not accomplished in this case. The other was from an individual whose mother was born in Texas and therefore had been grouped with the Hispanic Southwest category in our sample; however, the family's oral tradition was that the donor's maternal grandmother might have been of Yaqui ancestry. Kemp's study of Mexican and Southwestern Indian mtDNAs included one Tarahumara sample that matched this haplotype.⁴² These data, although too scanty for certain determination, are nonetheless consistent in attributing an indigenous Sonoran ancestry to this mtDNA haplotype, probably associated with groups speaking Uto-Aztecan languages (including the Mayo, Yaqui, and Tarahumara).

Among the California Spanish women who belonged to Haplogroup A is a woman named María de los Santos Gutiérrez from Culiacán, Sinaloa. The 1790 census identifies her *casta* as *india*. The original language of Culiacán was Tahue, probably closely related to Cahita, the language of the Mayos and Yaquis, but the dramatic population decline in the sixteenth and seventeenth centuries greatly reduced the original native population. By the 1700s coalescence with immigrants from other Mexican tribes meant that someone described as an Indian from Culi-

acán could have descended from any number of indigenous groups.⁴³ The mtDNA haplotype (A23) of Gutiérrez is one mutation removed from the central node of the Haplogroup A diagram. This mutation at np 16249 was also present among a Zapotec sample analyzed by Kemp, but other markers in that sample indicate that parallel mutations have occurred at the same molecular location, rather than the two being related phylogenetically.⁴⁴ Thus, it would appear that the haplotype of Gutiérrez represents an indigenous northwestern Mexican mtDNA lineage, although not enough evidence exists for further determination of ethnic affiliation.

The remaining three Haplogroup A samples for California Spanish women listed in table 3 all descend from women from Villa Sinaloa, a small town that contributed more colonists to Alta California than any other town in northwestern Mexico.⁴⁵ Two of these three women were listed as *española*, and the third was a *coyota* (table 3). Two samples from these women from Villa Sinaloa represented haplotypes that had no exact matches in our sample nor in Kemp's corpus of 100 Mexican and Southwest Indian mtDNA haplotypes belonging to Haplogroup A. One of these unique haplotypes (A32) belonged to Potenciana Ramírez, the woman described as *coyota*—that is, a person of three-quarters Indian ancestry. Her haplotype contained genetic markers that were so unusual that further study was deemed necessary. It is possible, however, that the presence of two isolate language groups in the Sinaloa River watershed, known as Ocoroni and Nio, might have harbored distinctive mitochondrial lineages, and women from these groups eventually joined with other mestizo and Indian families who settled in the nearby Villa Sinaloa. The Ocoroni and Nio, if indeed they represent linguistic isolates, are likely to have descended from more ancient groups that were subsequently overwhelmed and isolated after the spread of peoples speaking the Uto-Aztecan Cahita and Tahue languages.⁴⁶

The three *española* women from Villa Sinaloa who belonged to Haplogroup A may descend from an immigrant population from outside the local region. One of these, Micaela Sotelo, appears to belong to a haplotype (A06) that is derived from a mitochondrial lineage found among the Zapotec.⁴⁷ A second *española* from Villa Sinaloa, María Antonia Redondo, shared an mtDNA haplotype (A25) with two other participants in our study, one being a Spanish Californian descendant whose ancestry had not been successfully traced and the other coming from a great-grandmother from Sonora, who may have been of Mayo Indian descent, according to family tradition. The marker that characterized this haplotype at np 16209 was also found among Mixtec and Nahua samples

analyzed by Kemp, but the Mixtec and Nahua mtDNAs had additional markers that further differentiated them from this ancestral type.⁴⁸ In the samples analyzed for the present study, haplotypes A03 and A22, from Guanajuato and Arizona Mexican Americans respectively, were also derived from A25 (see figure 3). The fact that the two women who represented haplotypes possibly from elsewhere in Mexico were also those with enough European admixture to be considered españolas suggests that they could have descended from Indians or mestizos who came with the original Spanish settlers who began settling in the Sinaloa Valley in the late sixteenth century.⁴⁹

Haplogroup B Lineages

A total of twenty-two haplotypes occurred among thirty-four samples belonging to Haplogroup B. Eleven of these represented Haplotype B01, the founding lineage occurring as the central node in the Haplogroup B diagram (figure 4). The difficulty in determining the number of sequential cytosine molecules in the “poly-C” region between np 16180 and np 16193 in the mtDNA sequence introduces some uncertainty in the assignment of haplotypes, so we have eliminated positions np 16182 through 16183.2 from consideration in creating the network diagram in figure 4. This has resulted in combining haplotypes B01 and B06 and haplotypes B02, B08, and B09 in our figures, because the sequence variability reported at these locations appears not to be phylogenetically meaningful in all cases.

The vast majority of the samples belonging to Haplogroup B came from ancestors from Sonora, Chihuahua, and Durango, conforming to the predominance of this haplogroup among groups speaking Uto-Aztecan and Yuman languages in the arid, interior regions of northwestern Mexico and the American Southwest.⁵⁰ Although six California Spanish women in table 3 belonged to Haplogroup B, only four of these currently have sequences that can be identified with regard to haplotype. Three of these, an española born in Horcasitas, Sonora, an española from Villa Sinaloa, and a coyota from Rosario, Sinaloa, represent the principal founding lineage within Haplogroup B and therefore cannot be considered distinctive of particular ethnic groups. Similarly, a mestiza from Villa Sinaloa belonging to B08 is only one mutation removed from the founding haplotype.⁵¹ The resettlement of Ópata, Pima, and Tarahumara Indians from interior Sonora and Chihuahua to form enclaves on the Sinaloa River in the eighteenth century may also have introduced Haplogroup B lineages into an area where they may not have been common in precontact times.⁵²

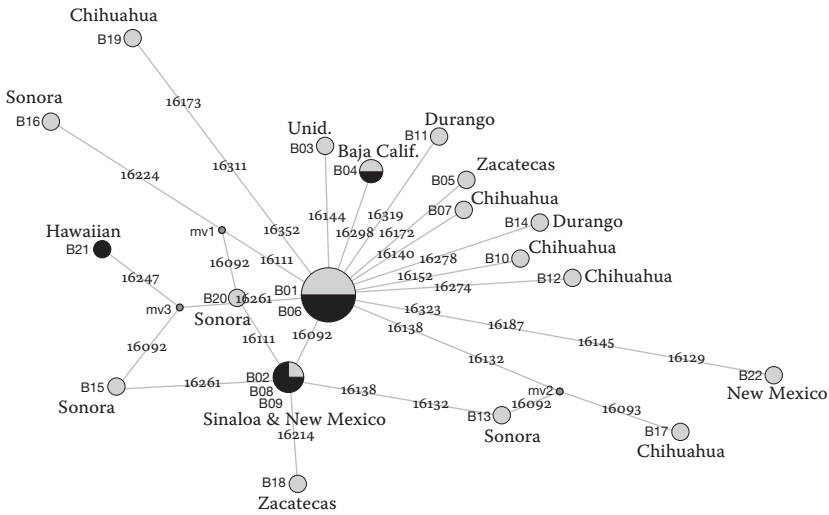


Figure 4. Haplogroup B Network Diagram Showing Ancestral Relationships among Spanish Californians and Mexican Americans

Haplogroup C Lineages

The forty-six samples belonging to Haplogroup C that were analyzed for this study included thirty-four haplotypes. As was the case with Haplogroups A and B, some reticulation is present, obscuring phylogenetic relationships among certain haplotypes because of parallel mutations and reverse mutations at certain molecular “hot spots” in the mitochondrial genome. The founding haplotype, represented by nine samples (here designated as C04), is clearly identifiable at the central node of the network. Although Haplogroup C is extremely common in the American Southwest and northwestern Mexico, it is also frequently encountered in other parts of Mexico, as is indicated in figure 5.

A total of five haplotypes of Haplogroup C have been identified among women listed in the 1790 census. Four of these California Spanish samples in three haplotypes (C31, C33, and C34) share a common marker at np 16295. This marker was present in the mtDNAs from an Indian woman from Culiacán, a coyota from Villa Sinaloa, and two españolas, one from Cosalá, Sinaloa, and the other from Loreto, Baja California (table 3). This particular mutation is also quite common among Akimel O’odham (Pima) and other American Southwest samples belonging to Haplogroup C, even though further mutations differentiate these from those haplotypes occurring among the California Spanish women. This marker also appears among haplotypes present among the

Huichol in Jalisco.⁵³ At least some of these Haplogroup C haplotypes carried to Alta California by settlers in the late eighteenth century would appear to be derived from people speaking Uto-Aztecan languages in Sonora and Sinaloa. The presence of the np 16295 marker among native groups in Arizona and Jalisco, to either side of Sonora and Sinaloa, suggests an ancestral association with the nearly continuous chain of indigenous Uto-Aztecan groups connecting these regions.

Haplogroup D Lineages

Haplogroup D is quite rare in Mesoamerica and the American Southwest. Kemp's study of 713 mtDNA samples from these regions produced only 22 that belonged to Haplogroup D, representing 12 haplotypes.⁵⁴ Another study yielded only one Haplogroup D individual out of 223 samples from towns on the Texas-Chihuahua international border.⁵⁵ The current study of 179 samples produced 15 that belonged to Haplogroup D, representing 12 haplotypes. Although our study supports the observation that Haplogroup D lineages are a minority, they would appear to be not quite as rare among people originating in coastal Sonora and Sinaloa as they are elsewhere in Mexico or the American Southwest. Thus, the California Spanish samples in particular acquire some significance in reconstructing the genetic mosaic of the original inhabitants of this sub-region of northwestern Mexico.

The reconstructed network diagram based on the twelve haplotypes that characterize Haplogroup D reveals a pattern quite different from what we have come to expect from the diagrams for Haplogroups A, B, and C (figures 3–5). Instead of a star-like configuration of daughter lineages arranged around a central node, Haplogroup D illustrates a double-star pattern (figure 6). Based on a discovery made by a team of researchers led by Olga Rickards, who studied the Cayapa Indians of coastal Ecuador, and supported by additional evidence gathered by Kemp and his colleagues, it is now recognized that there are two substantially different founding lineages of Haplogroup D that occur among Native Americans.⁵⁶ In figure 6, these two founding haplotypes are differentiated from each other by a minimum of three genetic markers at np 16342, 16325, and 16241, creating a “left” group of nine samples sorted into seven haplotypes and a “right” group of six samples sorted into five haplotypes. The subgroup on the left represents the central node (D01) and daughter branches of the most widespread of the founding haplotypes, which was recognized first. The subgroup on the right is arranged around the newly discovered founding haplotype, labeled D04 in table 7 and figure 6, but now officially designated as D4h3 in the molecular anthropological

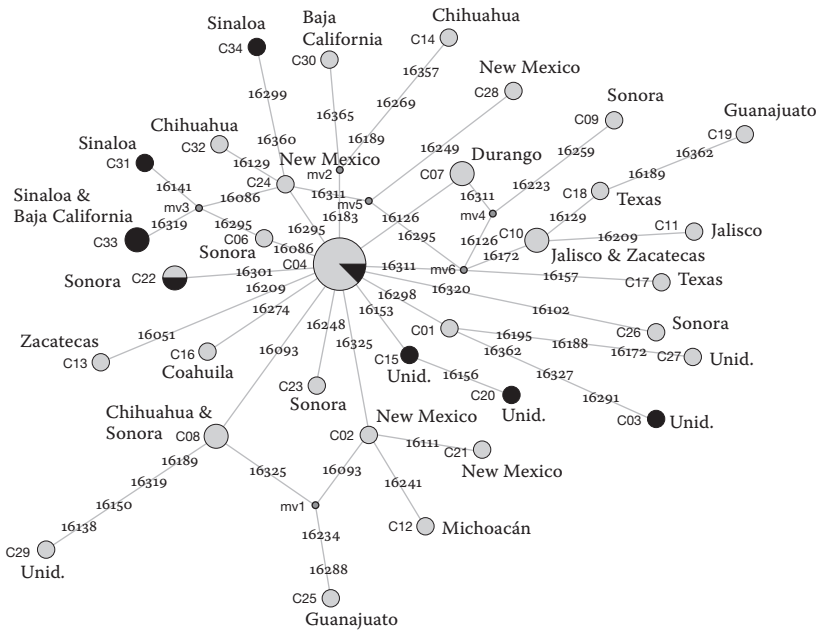


Figure 5. Haplogroup C Network Diagram Showing Ancestral Relationships among Spanish Californians and Mexican Americans

literature.⁵⁷ The recognized central node of this rarer subgroup is represented by one sample of presumed California Spanish origin (Sample JJ216), which exactly matches samples documented for the Chumash Indians of coastal California. Because this sample comes from an individual living in Santa Barbara in the central portion of territory inhabited by Chumash peoples and because the female lineage has not been traced further than the participant's grandmother, it cannot be ruled out that this lineage may have come from Chumash ancestors.

The newly discovered founding haplotype (D4h3) and its related haplotypes exhibit an interesting distribution of discontinuous occurrences, geographically separated from one other along the western margin of the American continents.⁵⁸ In the view of the authors of this study, this geographic pattern provides some of the best evidence to date in support of the hypothesis, long advanced by some archaeologists, that an ancient coastal migration may have taken place during the initial peopling of the Americas.⁵⁹ In view of the significance of this rare subgroup, it would be interesting to know if any of the six samples belonging to this subgroup could be traced to places of origin in coastal Sonora or Sinaloa. Among the haplotypes included in the Haplogroup D subgroup,

surrounding haplotype Do4 are three that descended from the Spanish Californian population (one containing two samples not obviously related to one another). Unfortunately only one of these has so far been traced to a female ancestor living in northwestern Mexico, María Dolores Valencia, who had been born in Horcasitas, Sonora (table 3). Her casta is listed as *española*, so it cannot be presumed that she was originally of native ancestors indigenous to the Pimería Bajo, where Horcasitas was founded. She might well have descended from a mestizo contingent of settlers who moved into the area following its pacification in the sixteenth century. None of the other haplotypes in this subgroup were traceable to more than general regions of the Spanish Borderlands (for example, Chihuahua and Texas); however, Kemp reported the presence of Haplotype Do4 (D4h3) among one of his Tarahumara samples, and Lance Green and colleagues discovered a closely related haplotype in their genetic study of residents from two Rio Grande border towns.⁶⁰ Even though the Tarahumara affiliation is of interest, most Tarahumara samples belonged to Haplogroups B and C, so its presence in the twenty-first century may be an accident of postcolonial intermarriage with people from other parts of Mexico. Clearly not enough evidence yet exists to associate the rare subgroup of Haplogroup D with any particular ethnolinguistic group or with a particular geographic area.

The other, more common subgroup of Haplogroup D included six individuals who were descended from the Spanish Californian colonial population (figure 6). Only two of these were successfully traced through their female ancestry to women listed in the 1790 census. One of these was María Rosa López (also known as Monreal), an *española* from Álamos, a prominent mining town, once part of Sinaloa but later annexed to Sonora, which drew many settlers from other regions of Mexico. The other individual was María Francisca Romero, a *mestiza* from Villa Sinaloa. Villa Sinaloa was founded in the area where Cahita was spoken, but there were other nearby local groups reported by early missionaries to be linguistically distinct.⁶¹ Also, enclaves of Ópata, Pima, and Tarahumara later settled on the Sinaloa River during the eighteenth century.⁶² So although it might be suspected that Romero's casta hints at an origin indigenous to the Sinaloa region, the ethnolinguistic affiliation of her ancestors cannot be reliably determined. It is interesting to note that her haplotype was shared with an individual whose mother was born in Texas to an immigrant woman from Chihuahua (figure 6). Also, Kemp's larger study of Mexican Indian mtDNA lineages reported three Tarahumara individuals who belonged to this haplotype.⁶³ These findings collectively support the conclusion that Romero's haplotype was indigenous

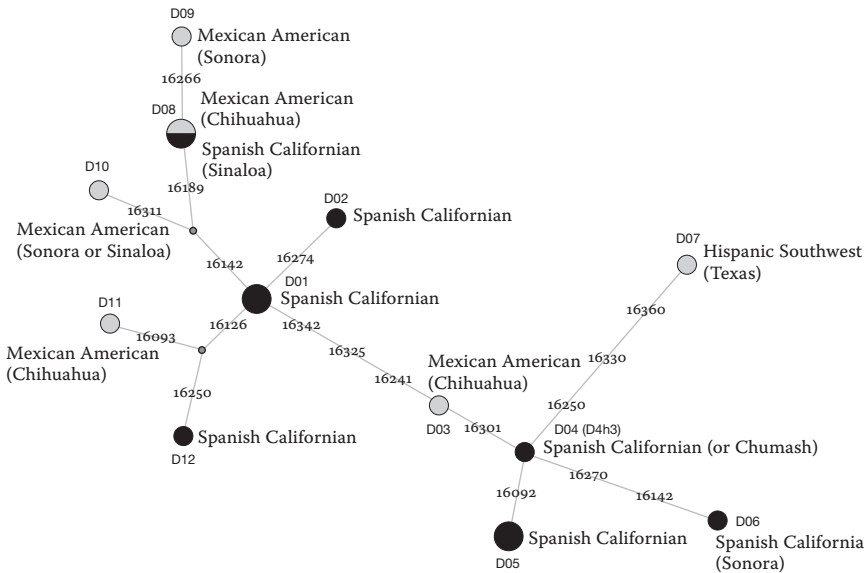


Figure 6. Haplogroup D Network Diagram Showing Ancestral Relationships among Spanish Californians and Mexican Americans

to northwestern Mexico, and we might further hypothesize that this mtDNA lineage type was found in a minority of lineages among several different tribes.

Conclusion

The study of the mitochondrial DNA lineages found among Alta California's colonial population has resulted in substantive new discoveries that supplement the findings based on historical documentary research. No matter what castas were indicated for Spanish California women of the colonial period, part of the descent of the greater majority of female settlers, exceeding 80 percent, was derived from indigenous groups of Mexico. The explanation for this apparent discrepancy has to do with the prevalent pattern of directional mating that occurred following the conquest of Mexico and subsequent spread of colonial control northward. The simple, traditional formula used during colonial times for determining castas illustrates how women with indigenous mitochondrial DNA could have great-granddaughters who were considered españolas:

1. De español é india, produce mestizo,
2. De español y mestiza, castiza.
3. De español y castiza, produce española.⁶⁴

Thus, the *casta* of a woman who was seven-eighths European would be considered *española* in colonial times, even though her mitochondrial DNA lineage was of indigenous origin. Genetic data demonstrate that social identity, as expressed by the *castas* reported in the 1790 census, did not necessarily reflect actual ancestry.

Our ultimate goal was to explore in greater detail whether these mitochondrial lineages might shed light on the original Native American groups that inhabited the region of Sonora and Sinaloa, where most of the Alta California female ancestors had been born. Some progress was made toward this objective, even though it was far from clear in most cases, given the limitations of surviving documentary evidence, which ethnolinguistic group might be the ancestral source of the lineage. Further hope is provided in that many mitochondrial sequences of Spanish Californians harbor unique genetic markers that can be the basis for research regarding their indigenous origin in Mexico.

One of the interesting observations about the native groups that inhabited the coastal plains and foothills of Sonora and Sinaloa (as well as southern Baja California) is that they had a fair amount of linguistic diversity, according to missionary testimony. Unfortunately, little evidence in the form of word lists or texts from these linguistically distinct local groups has survived. In the absence of sufficient linguistic data, some scholars have been inclined to attribute Uto-Aztecan affiliation to some of these groups, but this is clearly unwarranted in some cases.⁶⁵ Groups like the Guasave, Huite, Nio, and Ocoroni of Sinaloa, the Seri of Sonora, and the Guaicura and Pericú of Baja California del Sur likely represent survivals of earlier hunter-gatherers and fishermen who inhabited pericoastal areas prior to the arrival of agriculturalists speaking Uto-Aztecan and Yuman-Cochimí languages. Some of these locally distinct peoples of Sinaloa had adopted agriculture by the time of contact, while others, like the Guasave, had not. We can hypothesize that in the prehistoric past, some local Sinaloan groups had shifted to speaking Uto-Aztecan languages, especially Cahita and Tahue, through a process of cultural dominance. These languages spread at the expense of others during a period of demographic expansion resulting from the dependability of food supplies afforded by cultivation. As has been demonstrated in both Europe and North America, the spread of agriculture often did accompany the arrival of a new genetic component; however, older mtDNA lineages from the pre-agricultural population survived as well (or even predominated, such as in north-western Europe). As Kemp's research has shown, there is a substantive disconnect between mtDNA lineages among the Uto-Aztecan peoples of

the American Southwest, where Haplogroups B and C predominate, and those of Central Mexico, where Haplogroup A prevails.⁶⁶ Thus, genes and languages do not necessarily coincide. The coastal regions of Sonora and Sinaloa lie directly between these two regions and thus have the potential to inform us regarding the processes of demographic and cultural change that left a genetic signature in native populations.

The preliminary data derived from our analysis of Alta California's Spanish colonial population help answer some of the questions posed regarding the genetic affinities of the original peoples of Sonora and Sinaloa. It appears clear that the predominance of Haplogroups B and C is not likely to extend to the Cahitan peoples (Yaqui, Mayo, and related groups) of Sonora and Sinaloa. Enough information has now been accumulated to suggest that Haplogroup A lineages were likely to have been common in this area, including a distinctive haplotype that may have been present among the Guasave, and Haplogroup D appears to have been more prevalent there than elsewhere in Mexico. The presence of the rare subgroup of Haplogroup D in northwestern Mexico is surely significant, because of its hypothesized ancestral connection with an ancient coastal migration; however, its association with any particular ethnolinguistic group is conjectural at present. At this stage, it is not inappropriate to draw an analogy between the native peoples of Sinaloa and the Chumash population of coastal California, both of whom contrasted with their neighbors by harboring distinctive lineages belonging to Haplogroups A and D. We wonder if both populations might be descended from an ancient maritime-oriented group that established itself early in the settlement of the Americas during a fairly rapid Late Pleistocene coastal migration. To further test this hypothesis, of course, more research is necessary. To gain confidence in the patterns presented here, more sampling is needed among the surviving Yaqui and Mayo peoples, and further genealogical sleuthing should be undertaken in archives containing Sinaloan and Sonoran colonial records.

Certainly many Spanish Californians must have been aware of their partial Indian ancestry; in the new frontier setting in which they found themselves, however, this genetic inheritance became less important. In order to differentiate themselves from the populations indigenous to Alta California, the newly arrived colonists from Sonora and Sinaloa constructed a *gente de razón* social identity that eclipsed the old *casta* distinctions of their homeland.

Table 7. Haplogroup D mtDNA Sequence Data for Spanish Californians and Mexican Americans

Sample	Ethnicity	Traceable Origin	Haplotype	16092	16093	16126	16142	16183	16183.1	16189	16223	16241	16250	16266	16270	16274	16301	16311	16325	16330	16342	16360	16362
JJ255	Spanish Californian	Sonora or Sinaloa	D01	T	C	.	.	.	C
JJ069	Spanish Californian	Unidentified	D01	T	C	.	.	.	C
JJ190	Spanish Californian?	Unidentified	D02	T	A	.	.	C	.	.	.	C
JJ284	Mexican American	Chihuahua	D03	T	G	C	.	.	.	C
JJ216	Spanish Californian (or Chumash)	Unidentified	D04	T	G	T	.	C	.	.	.	C
JJ297	Spanish Californian	Unidentified	D05	C	T	G	T	.	C	.	.	.	C
JJ403	Spanish Californian?	Unidentified	D05	C	T	G	T	.	C	.	.	.	C
JJ280	Spanish Californian	Sonora	D06	.	.	.	T	.	.	.	T	G	T	.	C	.	.	.	C
JJ142	Hispanic Southwest	Texas	D07	T	G	T	.	.	.	T	.	C	.	A	G	C
JJ128	Spanish Californian	Sinaloa	D08	.	.	.	T	.	.	.	C	T	C	.	.	.	C
JJ198	Mexican American	Chihuahua	D08	.	.	.	T	.	.	.	C	T	C	.	.	x	C
JJ345	Mexican American	Sonora	D09	.	.	.	T	C	C	C	T	C	.	.	.	C
JJ349	Mexican American	Sonora or Sinaloa	D10	.	.	.	T	.	.	.	T	C	C	.	.	C
JJ285	Mexican American	Chihuahua	D11	.	C	C	T	C	.	.	.	C
JJ144	Spanish Californian?	Unidentified	D12	.	.	C	T	.	T	C	.	.	.	C

NOTES

- 1 For theoretical discussions of social identity, see Fredrik Barth, "Introduction," in *Ethnic Groups and Boundaries: The Social Organization of Cultural Difference*, ed. Fredrik Barth (Boston: Little, Brown, 1969), 9–38; Thomas Hylland Eriksen, *Ethnicity and Nationalism: Anthropological Perspectives* (Boulder, Colo.: Pluto Press, 1993).
- 2 A number of historians and anthropologists have examined changing identities in Alta California. See especially William Marvin Mason, *The Census of 1790: A Demographic History of Colonial California* (Menlo Park, Calif., Ballena Press, 1998); Lisbeth Haas, *Conquests and Historical Identities in California, 1769–1836* (Berkeley: University of California Press, 1995); Brian D. Haley and Larry R. Wilcoxon, "How Spaniards Became Chumash and Other Tales of Ethnogenesis," *American Anthropologist* 107 (2005): 432–45; and John R. Johnson and William M. Williams, "Toypurina's Descendants: Three Generations of an Alta California Family," *Boletín: The Journal of the California Mission Studies Association* 24, no. 2 (2007): 30–55.
- 3 Bryan Sykes, *The Seven Daughters of Eve* (London and New York: Bantam Press, 2001); Stephen Oppenheimer, *Out of Eden: The Peopling of the World* (London: Constable and Robinson, 2003); Spencer Wells, *Deep Ancestry: Inside the Genographic Project* (Washington, D.C.: National Geographic, 2006).
- 4 Henry Louis Gates Jr., series host and executive producer, *African American Lives*, DVD (PBS Paramount, 2006).
- 5 Maynard Geiger and Clement W. Meighan, *As the Padres Saw Them: California Indian Life and Customs as Reported by the Franciscan Missionaries, 1813–1815* (Santa Barbara, Calif.: Santa Barbara Mission Archive-Library, 1976), 11.
- 6 *Ibid.*, 12.
- 7 *Ibid.*, 11.
- 8 *Ibid.*, 15.
- 9 Mason, *Census of 1790*.
- 10 Claudio Esteva-Fabregat, *Mestizaje in Ibero-America*, trans. John Wheat (Tucson, Ariz.: University of Arizona Press, 1995).
- 11 D. Andrew Merriwether, "Mitochondrial DNA," in *Handbook of North American Indians*, vol. 3, *Environment, Origins, and Population*, ed. Dennis Stanford et al. (Washington, D.C.: Smithsonian, 2006), 817–30.
- 12 Antonio Torroni, "Mitochondrial DNA and the Origin of Native Americans," in *America Past, America Present: Genes and Languages in the Americas and Beyond*, ed. Colin Renfrew (Cambridge: McDonald Institute for Archaeological Research, 2000), 77–87.
- 13 Ann Gibbons, "Geneticists Trace the DNA Trail of the First Americans," *Science* 259 (1993): 312–13; Antonio Torroni et al., "Asian Affinities and Continental Radiation of the Four Founding Native American mtDNAs," *American Journal of Human Genetics* 53 (1993): 563–90; Douglas C. Wallace and Antonio Torroni, "American Indian Prehistory as Written in the Mitochondrial DNA: A Review," *Human Biology* 64 (1992): 403–16.

- 14 Peter Forster et al., "Origin and Evolution of Native American mtDNA Variation: A Reappraisal," *American Journal of Human Genetics* 59 (1996): 935–45; Joseph G. Lorenz and David Glenn Smith, "Distribution of Sequence Variation in the mtDNA Control Region of Native North Americans," *Human Biology* 69 (1997): 749–76; Ripan S. Malhi et al., "The Structure of Diversity within New World Mitochondrial DNA Haplogroups: Implications for the Prehistory of North America," *American Journal of Human Genetics* 70 (2002): 905–19.
- 15 Carol H. Mandryk et al., "Late Quaternary Paleoenvironments of Northwestern North America: Implications for Inland versus Coastal Migration Routes," *Quaternary Science Reviews* 20 (2001): 301–14; John J. Clague et al., "Environments of Northwestern North America before the Last Glacial Maximum," in *Entering America: Northeast Asia and Beringia before the Last Glacial Maximum*, ed. D. B. Madsen (Salt Lake City: University of Utah Press, 2004), 92.
- 16 John R. Johnson and Joseph G. Lorenz, "Genetics, Linguistics, and Prehistoric Migrations: An Analysis of California Indian Mitochondrial DNA Lineages," *Journal of California and Great Basin Anthropology* 26 (2006): 33–64; Brian M. Kemp, "Mesoamerica and Southwest Prehistory and the Entrance of Humans into the Americas: Mitochondrial DNA Evidence" (PhD diss., University of California, Davis, 2006); Ripan S. Malhi et al., "Native American mtDNA Prehistory in the American Southwest," *American Journal of Physical Anthropology* 120 (2003): 108–24.
- 17 Ives Goddard, "Native Languages and Language Families of North America," map accompanying *Handbook of North American Indians*, vol. 17, *Languages*, ed. Ives Goddard (Washington, D.C.: Smithsonian, 1996).
- 18 Jared Diamond and Peter Bellwood, "Farmers and Their Languages: The First Expansions," *Science* 300 (2003): 597–603; Peter Bellwood and Colin Renfrew, eds., *Examining the Farming/Language Dispersal Hypothesis* (Cambridge: McDonald Institute for Archaeological Research, 2002), 331–40.
- 19 This date range is based upon a 99.5 percent confidence interval (Kemp, "Mesoamerica and Southwest Prehistory," 38–39).
- 20 Compare Andrés Reséndez and Brian M. Kemp, "Genetics and the History of Latin America," *Hispanic American Historical Review* 85 (2005): 283–98.
- 21 Lance D. Green, James N. Derr, and Alec Knight, "mtDNA Affinities of the Peoples of North-Central Mexico," *American Journal of Human Genetics* 66 (2000): 989–98; Kemp, "Mesoamerica and Southwest Prehistory"; Malhi et al., "mtDNA Prehistory in the American Southwest"; Karla Sandoval-Mendoza, Leonor Buentello-Malo, and David Comas, "Insights into the Genetic Diversity of Indigenous Mexican Populations" (poster presentation, Languages and Genes: An Interdisciplinary Conference, University of California, Santa Barbara, September 8–10, 2006).
- 22 Goddard, "Native Languages"; Carl Sauer, *The Distribution of Aboriginal Tribes and Languages in Northwestern Mexico*, Ibero-Americana 5 (Berkeley: University of California, 1934).
- 23 Mason, *Census of 1790*, 65–72.
- 24 Johnson and Lorenz, "Genetics, Linguistics, and Prehistoric Migrations."

- 25 Genealogical research conducted for this study utilized copies of the mission registers available at the Santa Barbara Mission Archive-Library.
- 26 Marie E. Northrop, *Spanish-Mexican Families of Early California, 1769–1850*, 3 vols. (Burbank, Calif.: Southern California Genealogical Society, 1984–2004), 1:24; Dorothy Mutnick, *Some Alta California Pioneers and Descendants*, 5 vols. (Pleasant Hills, Calif.: Contra Costa County Historical Society), 1:65.
- 27 Several compilations of genealogical information were consulted for this project, in addition to primary records of baptisms, marriages, burials, and censuses. The most useful were Mason, *Census of 1790*; Mutnick, *Some Early California Pioneers*, vols. 1–5; and Northrop, *Spanish-Mexican Families of Early California*, vols. 1–3. The Huntington Library's Early California Population Project, <http://missions.huntington.org>, has also facilitated genealogical research.
- 28 Not included in figure 1 are: (1) thirty samples from mtDNA haplogroups indigenous to Europe and Africa, (2) twenty-five "California Spanish" descendants whose female lineages were not successfully traced to an origin in eighteenth-century Mexico, (3) nine samples from people of more recent Mexican immigrant ancestry who were uncertain of the region in Mexico from where their grandmothers came, and (4) one Polynesian sample descending through the female line from a Native Hawaiian woman who had married a Californio during the Mexican period.
- 29 This table expands on the information reported previously for twenty women in John R. Johnson, "Identifying the Ancestors of Los Californianos," *Noticias para Los Californianos* 37, no. 4 (2005): 41–43. The samples from four of the Alta California Indian women and one Baja California Indian woman are further discussed in Johnson and Lorenz, "Genetics, Linguistics, and Prehistoric Migrations." The fifth sample, contributed by a descendant from the direct female line from Regina Josefa Toypurina, has only recently been sequenced (see Johnson and Williams, "Toypurina's Descendants," for further information on this family's genealogical background).
- 30 The proportion of Native American mtDNAs among Spanish Californians (including probable Spanish Californians) is 82.7 percent or 88.6 percent, as calculated from tables 1 and 3 respectively.
- 31 Daniel T. Reff, *Disease, Depopulation, and Culture Change in Northwestern New Spain* (Salt Lake City: University of Utah Press, 1991); Carl Sauer, *The Aboriginal Population of Northwestern Mexico*, Ibero-Americana 10 (Berkeley: University of California, 1935).
- 32 Peter Gerhard, *The North Frontier of New Spain*, rev. ed. (Norman, Okla.: University of Oklahoma Press, 1993), 19.
- 33 Comparisons were made with 330 haplotypes reported by Kemp, "Mesoamerica and Southwest Prehistory"; however, exact matches could not be made with certainty because Kemp's haplotypes were based in part on genetic markers found in the second hypervariable region, which was not sequenced for the present study.
- 34 Johnson and Lorenz, "Genetics, Linguistics, and Prehistoric Migrations," fig. 5.

- 35 Kemp, "Mesoamerica and Southwest Prehistory," 38; Meredith Snow, Kathryn Durand, and David Glenn Smith, "Analysis of the Tommy and Mine Canyon Sites" (paper presented at the Society for American Archaeology annual meeting, Austin, Texas, April 28, 2007).
- 36 Johnson and Lorenz, "Genetics, Linguistics, and Prehistoric Migrations," fig. 5.
- 37 Certain specific haplotypes from table 4 have not been depicted in figure 3 because "hot spots" in the mtDNA genome have obscured phylogenetic pathways.
- 38 Mason, *Census of 1790*, 68, 89.
- 39 Mutnick, *Some Alta California Pioneers*, 1:345; Northrop, *Spanish-Mexican Families*, 2:63.
- 40 Sauer, *Distribution of Tribes and Languages*, 28–30.
- 41 Kemp, "Mesoamerica and Southwest Prehistory," 79–87.
- 42 Ibid., table C-1, HapA075.
- 43 Gerhard, *North Frontier of New Spain*, 257–61.
- 44 Kemp, "Mesoamerica and Southwest Prehistory," table C-1, HapA083.
- 45 Mason, *Census of 1790*, 65, 68, 71.
- 46 Sauer, "Distribution of Tribes and Languages," 25–26.
- 47 Kemp, "Mesoamerica and Southwest Prehistory," table C-1, HapA080.
- 48 Ibid., table C-1, HapA028, HapA051.
- 49 Gerhard, *North Frontier of New Spain*, 277.
- 50 See also Green et al., "mtDNA Affinities"; Kemp, "Mesoamerica and Southwest Prehistory."
- 51 Combined with Bo2 in figure 4.
- 52 Gerhard, *North Frontier of New Spain*, 274.
- 53 Kemp, "Mesoamerica and Southwest Prehistory," 92–95.
- 54 Ibid., 96.
- 55 Green et al., "mtDNA Affinities," 994.
- 56 Olga Rickards et al., "mtDNA History of the Cayapa Amerinds of Ecuador: Detection of Additional Founding Lineages for Native American Populations," *American Journal of Human Genetics* 65 (1999): 519–30; Brian Kemp et al., "Genetic Analysis of Early Holocene Skeletal Remains from Alaska and its Implications for the Settlement of the Americas," *American Journal of Physical Anthropology* 132 (2007): 605–21.
- 57 Erika Tamm et al., "Beringian Standstill and Spread of Native American Founders," *PLoS ONE* 2, no. 9 (2007): e829 (doi:10.1371/journal.pone.0000829); Ugo Parego et al., "Distinctive Paleo-Indian Migration Routes from Beringia Marked by Two Rare mtDNA Haplogroups," *Human Biology* 19 (2008): 1–8.
- 58 Kemp et al., "Genetic Analysis," 615.
- 59 E. James Dixon, *Bone, Boats, & Bison: Archeology and the First Colonization of Western North America* (Albuquerque: University of New Mexico, 1999); K. R. Fladmark, "Routes: Alternate Migration Corridors for Early Man in North America," *American Antiquity* 44 (1979): 55–69; Johnson and Lorenz,

- “Genetics, Linguistics, and Prehistoric Migrations,” 56; Parego et al, “Distinctive Paleoindian Migration Routes.”
- 60 Green et al., “mtDNA Affinities,” 994; Kemp et al., “Genetic Analysis,” 214.
- 61 Wick R. Miller, “Uto-Aztecan Languages,” in *Handbook of North American Indians*, vol. 10, *Southwest*, ed. Alfonso Ortiz (Washington, D.C.: Smithsonian, 1983), 122; Sauer, *Distribution of Tribes and Languages*, 25–26, 28–30, 37.
- 62 Gerhard, *North Frontier of New Spain*, 274.
- 63 Kemp, “Mesoamerica and Southwest Prehistory,” table C-4, HapDoo9.
- 64 “1. From Spaniard and Indian, a *mestizo* is born. 2. From Spaniard and *mestiza*, *castiza*. 3. From Spaniard and *castiza*, a Spaniard is born” (Ilona Katzew, *Casta Painting: Images of Race in Eighteenth-Century Mexico* [New Haven, Conn.: Yale University Press, 2004], figs. 163–65 [pp. 130–31]).
- 65 N. Ross Crumrine, “Mayo,” in *Handbook of North American Indians*, 10:264; Miller, “Uto-Aztecan Languages,” 122; Sauer, *Distribution of Tribes and Languages*, 15–37.
- 66 Kemp, “Mesoamerica and Southwestern Prehistory,” chap. 2.