Population Genetic Structure of California Hazelnut, An Important Food Source for People in Quiroste Valley in the Late Holocene

Paul V. A. Fine
Department of Integrative Biology, University of California, Berkeley, CA 94720 (paulfine@berkeley.edu)

Tracy M. Misiewicz
Department of Integrative Biology, University of California, Berkeley, CA 94720 (tmisiewicz@berkeley.edu)

Andreas S. Chavez
Department of Biology and Burke Museum, University of Washington, Seattle, WA 98195 (aschavez@u.washington.edu)

Rob Q. Cuthrell
Department of Anthropology, University of California, Berkeley, CA 94720 (rcuthrell@berkeley.edu)

Abstract California hazelnuts (*Corylus cornuta* var. *californica*) are abundant in the archaeological record of site CA-SMA-113 in Quiroste Valley Cultural Preserve, and hazel management on the Central Coast was recorded in late 18th century Spanish accounts. However, this species is currently absent from Quiroste Valley proper and rare in the watershed, though it is locally common elsewhere in the Santa Cruz Mountains. Because high California hazelnut abundance is associated with frequent fire regimes, we believe that its current low abundance could be due to fire suppression enforced in the region for the past two hundred years. We sequenced nuclear microsatellites from ten populations of California hazelnuts to test the hypothesis that this species has experienced demographic changes in response to changing climate and land management practices. We found that all populations exhibited high levels of genetic variation and negative population growth consistent with large population sizes in the past with some decline over time. We also found subtle patterns of geographic structure suggesting that Quiroste Valley and neighboring Butano may have been important refugia habitats during past climate warming events. These results provide an important foundation demonstrating that

California Archaeology, Volume 5, Number 2, December 2013, pp. 353–370.

population genetic approaches can be applied to eco-archaeological research on indigenous landscape management and set the stage for future work using genetics to reveal further details of the demographic history of Quiroste Valley hazelnut populations.

Resumen Las avellanas de California (*Corylus cornuta* var. *californica*) son abundantes en el registro arqueológico del sitio CA-SMA-113 en la Reserva Cultural del Valle Quiroste, y el manejo de los avellanos en la Costa Central fue mencionado en los registros españoles de fines del siglo XVIII. Sin embargo, esta especie está actualmente ausente del Valle Quiroste propiamente dicho, y es rara en la cuenca, aunque es común en otras partes de las Montañas Santa Cruz. Dado que la gran abundancia de avellanas está asociada con los regímenes de incendios frecuentes, creemos que la escasez actual podría deberse a la supresión del fuego ejecutada en la región en los últimos doscientos años. Hemos secuenciado microsatélites nucleares de diez poblaciones de avellanas de California para probar la hipótesis de que esta especie ha experimentado cambios demográficos en respuesta al cambio climático y a las prácticas de manejo de la tierra. Hemos descubierto que todas las poblaciones exhibían altos niveles de variación genética y un crecimiento negativo de la población, consistente con grandes poblaciones en el pasado con un cierto declive en el tiempo. También hemos descubierto sutiles patrones de estructura geográfica que sugieren que el Valle Quiroste y el cercano Butano pueden haber sido importantes hábitat de refugio durante los pasados eventos de calentamiento climático. Estos resultados proporcionan una base importante que demuestra que los enfoques de genética de poblaciones pueden ser aplicados a la investigación eco-arqueológica sobre el manejo indígena del paisaje y sentar las bases para el trabajo futuro, usando la genética para revelar más detalles de la historia demográfica de las poblaciones de avellana del Valle Quiroste.

From ethnographic accounts and historical documents, we know that California Indians actively managed the landscape to create stable food supplies, using fire as their primary means to modify vegetation at large scales (Anderson 2005). However, these records only go back to the eighteenth century, whereas Indians have inhabited California for at least 12,000 years (Erlandson et al. 2007). To what extent did these practices extend back across the many centuries of human settlement in California, and how large of an area was impacted by fire management during these times? The answers to these questions are relevant not only to reconstructing landscapes and cultures of the past, but also for

predicting how animal and plant communities may change in the future under scenarios of further global climate change. Moreover, our work with the Amah Mutsun Tribal Band has shown that many tribal members are eager to re-apply the traditional land management practices that sustained productive landscapes in California for centuries, and that have the potential to do so for future generations (Lightfoot and Lopez, this volume).

Traditionally, the investigation of human land management prior to historical documentation has remained in the realm of archaeology. Indeed, archaeological research can reveal a wealth of information about foodways, cultural practices, and landscape modifications in particular spatiotemporal contexts. However, we suggest the reconstruction of entire landscapes from such contexts can be improved with a broader approach including other methods of investigating the past, such as dendrochronology, phytolith analyses (Evett and Cuthrell, this volume), and surveys of pollen and charcoal in radiocarbon dated wetland sediment cores (Cowart and Byrne, this volume). In addition, investigating the genetic diversity of ancient DNA of biotic remains from archaeological sites can be used to test hypotheses of the strength of human impacts on food resources (Broughton et al. 2013).

The genetic diversity of a population of contemporary individuals also can serve as a window into past demographic events. Because genetic diversity in an individual derives from mutation rates and recombination during sexual reproduction of ancestors, measures of genetic diversity are highest in individuals belonging to populations that have been large and out-breeding over many generations and lowest in individuals belonging to populations that have experienced severe demographic declines and/or inbreeding over time (Soulé 1976). For example, sea otters have been found to have extremely low genetic diversity, a result of hunting pressure that reduced the Pacific Ocean population to just a few widely scattered populations representing less than 1 percent of their former numbers (Kenyon 1969). Today's Pacific otters descend from these survivors and bear the signature of low genetic diversity and high inbreeding (Larson et al. 2012). Conversely, Pacific whales were found to have much higher genetic diversity than expected, suggesting that whale populations before commercial whaling were much higher than previously believed (Roman and Palumbi 2003).

The neutral genetic variation (i.e., variation which does not affect organism function) within populations often changes over spatial distance because individuals within populations usually mate with other individuals from the same population. Thus, the genetic variation found among populations of a species can exhibit "geographic structure", usually correlating with spatial distance or

topographic features that limit organism mobility (Avise 2000). Because species (and populations) shift their geographic distributions through time, it is sometimes possible to track such movement using genetic markers. For example, studies on North American plants have used these methods to infer migration rates following glacial retreat as well as to identify the putative refugia where tree populations survived during the last glacial maximum (McLachlan et al. 2005).

At Quiroste Valley, archaeological studies of plant and animal remains found in contexts dated to ca. AD 1,000–1,300 have included >19,000 archaeofaunal and >40,000 archaeobotanical specimens from more than 49 animal and 50 plant taxa (Cuthrell, this volume; Gifford-Gonzalez et al., this volume). Unlike most contemporary sites in Interior Central California, hazelnut is by far the most common nut food taxon at SMA-113, comprising over 85 percent of archaeological nutshell by weight (Cuthrell, this volume). It is possible that these hazelnuts were transported from elsewhere in the Santa Cruz Mountains, but most likely that they were harvested locally, within the foraging radius of the village (Kelly 1995). Historical accounts also mention hazelnut management in the vicinity. Father Juan Crespi, a member of the 1769 Portola expedition, reported stands of burnt hazelnut between the present-day locations of Watsonville and Santa Cruz, about 50 km ESE of Quiroste Valley. He wrote, "[there are] a great many knoll ranges wooded with hazelnut groves yielding very good-sized nuts ... the heathen, by what we could see burn them off" (Brown 2001:561).

These lines of evidence suggest that hazelnuts may have been an important food source for Quiroste Valley inhabitants from at least ca. 1,000 years ago until Spanish colonization. Today, however, hazelnut is almost entirely absent from the Whitehouse Creek watershed (containing Quiroste Valley), as extensive searches have uncovered only one population comprising two stands of less than 200 individuals in total, located approximately 2 km upstream of SMA-113 in a shady streamside habitat under redwood forest. The annual production of this stand would not yield large quantities of nuts (see Cuthrell, this volume, for more information on California hazelnut yields). Given the expected irregularity and sparsity of natural hazelnut production and the low density of the taxon in the area today, the abundance of hazelnuts in the SMA-113 assemblage would be difficult to explain if hazelnut had the same distribution and abundance during site occupation as it does today. Thus, it is plausible that the type of hazelnut management described by Crespi had a long history in this area. If hazelnut burning increased nut as well as sprout production, the abundance of hazel in the SMA-113 macrobotanical assemblage could be a

strong indicator of hazelnut management during the early part of the Late Holocene.

We hypothesize that the current rarity of hazelnut in Quiroste Valley may be due to the pervasive fire suppression that has occurred in the vicinity over the past 200 years. Besides Crespi's account, there are many studies that document Native Americans intentionally setting fires to promote hazelnut (LaLande and Pullen 1999; Levy 2005; Turner 1999). For example, the Coquille tribe of Oregon burns hazelnut fields approximately every five years to boost nut production and promote new clones (LaLande and Pullen 1999). Areas that are known to have experienced frequent fire returns have the highest densities of hazelnut, reaching very large densities (Halpern 1989). To be able to test our hypothesis, we need to know if Quiroste Valley hazelnuts represent a unique, genetically recognizable population. Furthermore, given the uncertainty of tracking genetic change with respect to absolute chronology, we must place the genetic heritage of Quiroste Valley hazelnuts within the context of the larger demographic history of hazelnuts over the past several millennia. Besides recent fire suppression in the area, there have been many other changes that could influence hazelnut population size and distribution over longer time scales, such as climate change over the last glacial cycles.

We used nuclear microsatellites to characterize genetic diversity and population structure of California hazelnut from California's Central Coast to pose the following three questions: (1) What is the genetic diversity within hazelnut populations of Central Coastal California hazelnut? (2) Do hazelnut populations exhibit geographic structure? (3) Is there evidence that hazelnut population size has changed through time on the Central Coast?

Methods

Study Organism and Field Collection

California hazelnut (*Corylus cornuta* var. *californica*) is a small, multi-stemmed tree that grows from Monterey County in California north to British Columbia, Canada (Figure 1). In California, it occurs in mixed evergreen forest and riparian areas in coastal areas, and in the Sierra Nevada it is present at mid-elevations south to Sequoia National Park. California hazelnuts reproduce through seed as well as vegetatively, with new stems growing from the root crown. Individual stems can reach over 20 years in age, but the clones themselves (i.e., genetically identical organisms) have much longer lifespans. California hazelnut will

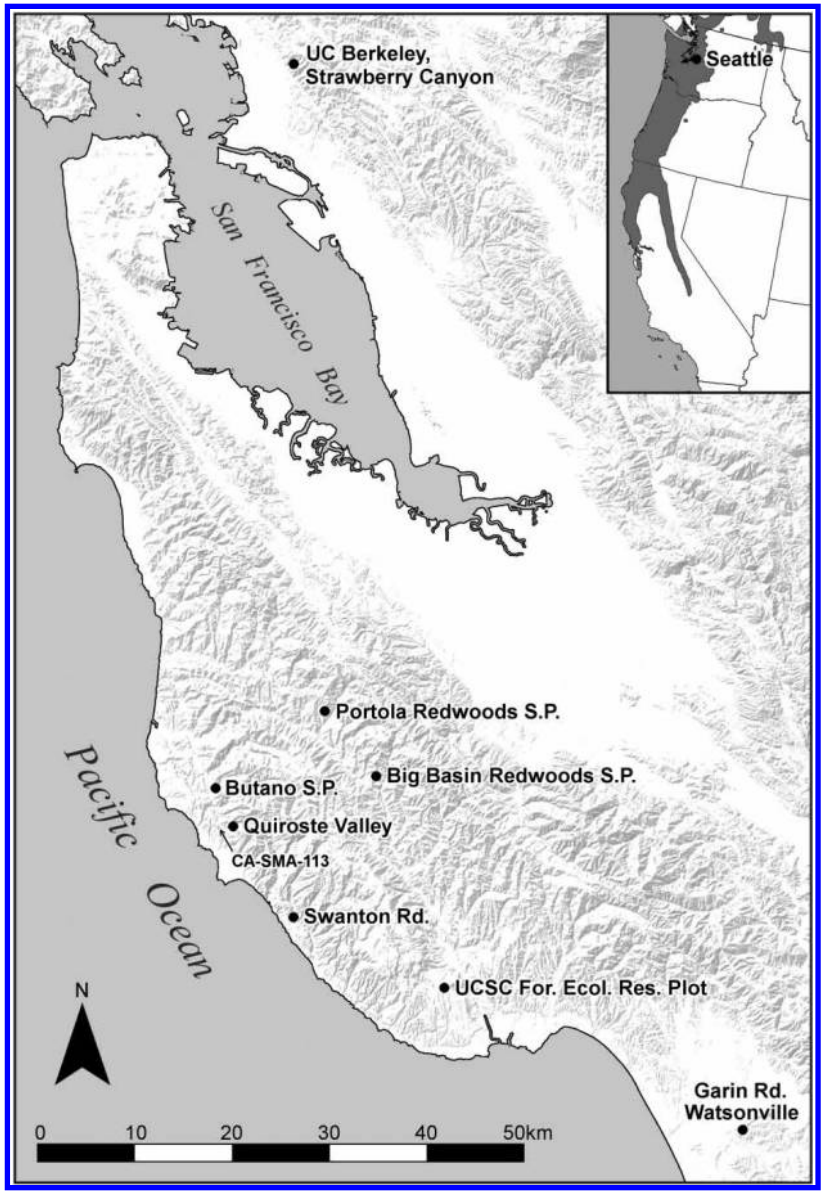

Figure 1. Sampling locations of California hazelnut populations on the central coast of California and location of CA-SMA-113. Inset: Contemporary distribution of California hazelnut (U.S. Geological Survey 2013) and sampling location at Seattle, Washington.

resprout from the root crown in response to top killing by fire. It is a minor understory component of most forests in which it occurs, with the exception of recently burned areas, where it can become quite dominant (Halpern 1989).

We collected California hazelnut specimens from eight sites in California and one site in Seattle, Washington (Figure 1). For each site besides Quiroste Valley, care was taken to sample individual plants more than 5 m from one another. At Quiroste Valley, all individuals were sampled.[1] Leaves were removed from each plant and placed in paper envelopes in a large plastic bag with silica gel and stored in a −20°C freezer at UC Berkeley.

DNA Extraction and PCR Amplification

DNA was extracted from the leaf tissue of 189 *Corylus californica* var. *cornuta* individuals collected from eight populations along the Central Coast of California and one near Seattle, Washington (Figure 1). Genomic DNA was extracted using a Qiagen DNEasy Plant Mini Kit (Valencia, California, USA). A 1:10 dilution of DNA and water was used for all amplifications.

Thirty-three primer pairs previously developed for the European hazelnut (*Corylus avellana*) and demonstrated to be transferable across *Corylus* species, including the California hazelnut, were assessed for use in this study (Bassil et al. 2005; Boccacci et al. 2005). Polymerase chain reactions (PCR) were performed in a total reaction volume of 25 μl containing 12.5 μl Promega GoTaq Green Master Mix (Madison, Wisconsin), 7.5 μl H_2O, 1.5 μl 10 μM forward primer, 1.5 μl 10 μM reverse primer, and 2 μl of diluted DNA template. All amplifications were conducted using a touchdown PCR protocol beginning with an initial temperature of 95°C for 150 seconds followed by 30 cycles at 95°C for 30 seconds, annealing at a temperature of 60°C decreased by 0.5°C every 30 seconds until 45°C, 72°C for 60 seconds, and ending with a final extension at 72°C for 20 minutes.

Characterization of Polymorphic Markers

Seventeen primers that were amplified successfully across populations were labeled with the florescent marker 6-FAM or HEX for visualization. Amplification products were co-loaded on an ABI 3730 automated sequencer (Applied Biosystems) with 0.3 μl GS-500 LIZ (Applied Biosystems) size standard to allow allele length sizing. Electrophoretic results were initially scored using Peak Scanner v. 1.0 (Applied Biosystems) followed by visual confirmation.

Non-variable loci and loci with severe stutter were discarded from the study resulting in a total of 10 useful polymorphic loci. Deviation from Hardy

Weinberg Equilibrium was tested using GenAlEx 6.0 (Peakall and Smouse 2006) and raw data was checked for errors due to large allele dropout, stutter, and null alleles using MICROCHECKER (van Oosterhout et al. 2004). Three of the remaining 10 loci were discarded from the analysis after they were shown to significantly deviate from HWE with lower heterozygosity than expected across the majority of populations and tested positive for null alleles. The seven remaining loci used for all further analyses did not significantly deviate from HWE for most populations after sequential Bonferonni corrections (Rice 1989).

Genetic Diversity

Basic genetic diversity estimates including number of alleles (Na), number of effective alleles (Ne), observed heterozygosity (Ho), expected and unbiased expected heterozygosity (He and Uhe), information index (I), and the fixation index (F) were calculated using GenAlEx 6.0 (Peakall and Smouse 2006). Linkage disequilibrium was calculated using GenePop 4.0 (Raymond and Rousset 1995; Rousset 2008).

Population Differentiation

Pairwise F_{ST} values between all populations and across all loci as well as an analysis of molecular variance (AMOVA, Excoffier et al. 1992) were calculated using GenAlEx 6.0. AMOVA calculated variance among populations, among individuals, and within individuals using R_{ST} statistics with 999 permutations. Further evidence of population structure was assessed using population structure analysis implemented in Structure 2.3.4 (Pritchard et al. 2000). This analysis assessed the most probable number of genetic clusters that characterizes our samples and assigned individuals to populations. STRUCTURE is a model-based algorithm method that uses Bayesian statistics and Markov chain Monte Carlo (MCMC) to make cluster assignments using genetic data. We used the "admixture" model with correlated allele frequencies among populations and sampling locality as a prior. The use of sampling locality as a prior does not overestimate population structure when none is present, but can improve the ability of the clustering algorithm to detect structure when data are limited (Hubisz et al. 2009). We performed three replicate runs for each value of K ranging from 1 to 9 with a burn-in of 10,000 followed by 20,000 repetitions. Each run estimated the "log probability of data" (L(K)). We recorded the ln Prob(Data) for each run and averaged the ln Prob(Data) across runs for each value of K. Finally, genetic exchange was calculated using Slatkin and Barton's (1985) private allele method implemented in Genepop 4.0. The private allele method is an indirect estimate

of gene flow based on the theory that private alleles are unlikely to reach high frequencies if migration rates are high (Slatkin and Barton 1985).

Bottleneck Detection

A genetic bottleneck is experienced when a population undergoes a severe demographic decline followed by a recovery. Changes in genetic variation created by bottlenecks occur both within and between populations and have been well studied. When no data prior to the putative bottleneck is available, measurements including allelic diversity, average heterozygosity, and measurements of heterozygosity excess are commonly used for detection (Bonnell and Selander 1974; Bouzat et al. 1998; Cornuet and Luikart 1996; Nei et al. 1975). Attempts to uncover genetic signatures of a recent bottleneck were carried out using BOTTLENECK version 1.2.02 (Piry et al. 1999). Tests implemented in BOTTLENECK operate under the theory that once a population undergoes a bottleneck there will be a correlated reduction in both allele numbers and levels of heterozygosity. However, since allelic diversity will be lost faster then heterozygosity, a signature of heterozygosity excess (not heterozygote excess) is expected (Cornuet and Luikart 1996).

Tests were conducted across loci for each population using the sign test and Wilcoxon sign-rank test assuming a stepwise mutation model (SMM), an infinite alleles model (IMM), and a two-phase model (TPM), which incorporates aspects of the two extreme models. The standardized differences test was omitted due to lack of loci (a minimum of 20 loci are recommended). Because dinucleotide microsatellites tend to adhere more closely to an IAM model, the TPM was set to 10 percent SMM and 90 percent IAM. A description of the allele frequency distribution, referred to as the "mode-shift" indicator is also used to distinguish populations that have experienced recent bottlenecks from stable ones. A bottleneck is detected when the shape of the graph of allele frequencies deviates from the normal L-shaped distribution expected under mutation-drift equilibrium (Piry et al. 1999). One hazelnut population (Portola Redwood location) was not included in the calculations due to low sample size.

Population Demography

We assessed evidence of historical fluctuations in population size using data from 11 microsatellite loci with LAMARC 2.1.8 (Kuhner 2006). We investigated population fluctuation in four geographic locations that well represented the geographic structure in our sample: Quiroste Valley, Watsonville, Swanton Road, and Seattle. LAMARC estimates a population-growth parameter, g,

by using the coalescent and a Markov chain Monte Carlo genealogy sampler. Positive values of *g* indicate a growing population and negative values represent shrinking populations. A *g* of zero represents constant size. We used the 99 percent confidence intervals for *g* to test for significant difference from zero. For each maximum likelihood run, we used 20 short chains of 10,000 steps and two long chains with 200,000 steps. The number of short chains was based on when parameter estimates appeared to be stabilized. We used a Brownian-motion model for the mutation model, which is a statistical approximation of the stepwise model. Each run was replicated three times with different starting seeds and with three simultaneously heating searches.

Results

Genetic Diversity

All loci exhibited moderate levels of polymorphism. The total number of alleles per locus ranged from 5 to 21. Private alleles were observed at most loci resulting in a lower number of effective alleles relative to the number of total alleles across loci (Table 1). Genetic variation within populations was moderate with observed heterozygosities ranging from 0.543–0.688 (Table 1). With the exception of locus CAC-B020 in Swanton Road and Big Basin populations and locus CAC-B109 in the Quiroste Valley population, no loci significantly differed from HWE after sequential Bonferonni corrections. The majority of loci did not show patterns of linkage disequilibrium after sequential Bonferonni corrections, the exceptions being CAC-A102 and CAC-B109 in the Watsonville population, CAC-B029b and CAC-B101 in the Butano population, and CAC-B010 and CAT-B107 in the Big Basin populations. Interestingly, the Swanton Road population showed high levels of disequilibrium between loci with CAC-A102 and CAC-B109, CAC029b and CACB109, CAC-B101 and CAC-B109, CAC-B029b and CAC-B101, and CAC-B020 and CAC-B109 showing significant signs of linkage (5 of the 21 pairwise comparisons).

Population Differentiation

Pairwise F_{ST} values were low between California populations ranging from 0.012–0.054. F_{ST} values between California populations did not show any trend relative to geographic distance, however F_{ST} values were the highest between the Seattle population and all other California populations, suggesting increased isolation by distance (Table 2). An indirect estimate of 7.75 migrants

Table 1. Genetic Diversity Statistics for All Populations Across All Loci.

		N	Na	Ne	I	Ho	He	UHe	F
UC Santa Cruz	*Mean*	25	7.43	3.44	1.44	0.68	0.65	0.67	−0.06
	SE		1.02	0.60	0.18	0.06	0.06	0.06	0.05
UC Berkeley	*Mean*	27	7.23	3.19	1.33	0.62	0.61	0.62	−0.01
	SE		1.48	0.65	0.22	0.08	0.07	0.07	0.05
Watsonville	*Mean*	25	7.00	2.98	1.25	0.58	0.58	0.59	0.02
	SE		1.20	0.71	0.21	0.07	0.09	0.07	0.08
Seattle	*Mean*	10	5.23	3.39	1.39	0.69	0.70	0.75	0.02
	SE		0.12	0.12	0.06	0.06	0.01	0.01	0.09
Swanton Road	*Mean*	25	7.00	3.47	1.41	0.62	0.65	0.66	0.04
	SE		0.93	0.52	0.18	0.08	0.71	0.07	0.05
Butano	*Mean*	25	6.00	3.21	1.23	0.62	0.60	0.61	−0.00
	SE		1.38	0.79	0.23	0.08	0.08	0.08	0.07
Big Basin	*Mean*	22	7.57	3.42	1.40	0.60	0.64	0.65	0.08
	SE		1.19	0.63	0.20	0.08	0.07	0.07	0.05
Quiroste Valley	*Mean*	25	7.43	3.99	1.42	0.62	0.64	0.66	0.05
	SE		1.46	1.02	0.25	0.09	0.08	0.08	0.06
Portola Redwoods	*Mean*	5	3.57	2.32	0.93	0.54	0.50	0.55	−0.12
	SE		0.53	0.39	0.17	0.08	0.08	0.09	0.07

Notes: N, sample size; Na, number of alleles; Ne, effective number of alleles; I, polymorphism information index; Ho, observed heterozygosity; He, expected heterozygosity; UHe, unbiased expected heterozygosity; and F, fixation index.

Table 2. Pairwise FST Values for All Population Pairs.

	UCSC	UCB	WAT	SEA	SWR	BUT	BGB	QV	PR
UC Santa Cruz	0.000								
UC Berkeley	0.023	0.000							
Watsonville	0.028	0.013	0.000						
Seattle	0.058	0.073	0.082	0.000					
Swanton Road	0.015	0.018	0.024	0.062	0.000				
Butano	0.047	0.029	0.029	0.085	0.043	0.000			
Big Basin	0.012	0.025	0.023	0.060	0.017	0.044	0.000		
Quiroste Valley	0.027	0.025	0.026	0.068	0.027	0.054	0.024	0.000	
Portola Redwood	0.033	0.023	0.029	0.105	0.035	0.046	0.037	0.042	0.000

Notes: UCSC, UC Santa Cruz; UCB, UC Berkeley; WAT, Watsonville; SEA, Seattle; SWR, Swanton Road; BUT, Butano; BGB, Big Basin; QV, Quiroste Valley; PR, Portola Redwood.

per generation (mean frequency of private alleles = 0.026) was calculated using Slatkin and Barton's (1985) private allele method.

The STRUCTURE analysis using allele profiles from seven microsatellite loci found that K = 5 was the most likely number of clusters of nuclear variation in the dataset (mean ln Prob(Data) = −3971.1). One of the five genetic clusters was geographically restricted to Seattle, Washington (Figure 2). The remaining four clusters were found at multiple sampling sites among the California populations. All samples at the Berkeley, Watsonville, and Swanton Road sites were mostly assigned to one of the genetic clusters (Figure 2). The remaining sites in California contained samples that had multiple populations-of-origin.

Bottleneck Detection

The IAM, SMM, and TPM were applied to test if the microsatellite loci showed departure from mutation-drift equilibrium using BOTTLENECK. Under the IAM, results indicated that there was no deviation from mutation-drift equilibrium using the sign test ($P > 0.05$) or the Wilcoxon rank sign test ($P > 0.1$, two tailed). Under the SMM, Seattle did not deviate from mutation drift equilibrium for both tests and Quiroste Valley did not deviate from mutation drift equilibrium using the sign test. All other results were significant under the SMM, indicating a deficiency in heterozygosity using both the sign test ($P < 0.05$) and Wilcoxon rank sign test ($P < 0.1$, two tailed). Results under TPM were consistent with IAM results (Table 3). Analysis of allele frequency patterns revealed normal L-shaped distributions, which are expected in populations that have not experienced a recent severe bottleneck.

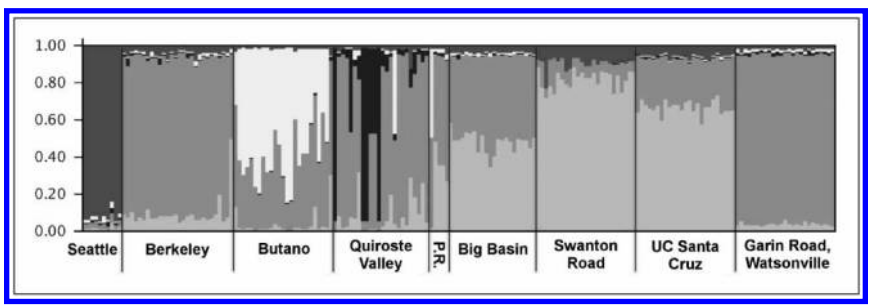

Figure 2. Structure results for K = 5 clusters across all populations. Each vertical bar represents an individual from the populations indicated on the x-axis, oriented from most northerly population towards the south, left to right. The different shades of gray correspond to ancestral population of origin. The y axis represents probability of assignment to one of these five clusters. Location "P.R." is Portola Redwoods State Park.

Table 3. Results for Tests of Heterozygosity Excess and Mode Shift Using BOTTLENECK.

	Sign Test			Wilcoxon Test			Mode Shift
	IAM	SMM	TPM	IAM	SMM	TPM	
UC Santa Cruz	0.403	0.002	0.312	0.938	0.008	0.578	No
UC Berkeley	0.328	0.217	0.321	0.813	0.234	0.375	No
Watsonville	0.117	0.002	0.116	0.297	0.008	0.078	No
Seattle	0.392	0.573	0.407	0.375	0.813	0.375	No
Swanton Road	0.600	0.001	0.591	0.578	0.008	0.812	No
Butano	0.639	0.022	0.647	0.297	0.055	0.469	No
Big Basin	0.415	0.002	0.104	0.938	0.008	0.078	No
Quiroste Valley	0.118	0.129	0.343	0.297	0.039	0.937	No

Notes: P-values are presented for tests performed across all populations for both the Sign Test and Wilcoxon Test using the infinite alleles model (IAM), stepwise mutation model (SMM), and the two-phase model (TPM). Significant p-values under the SMM signify excess homozygosity. Two-tailed P-values were calculated for the Wilcoxon Test.

Historical Population Demography

All four hazelnut populations that we analyzed showed evidence of a slight decline in historical population size based on our LAMARC analyses. The maximum likelihood estimates (MLE) of g were small and negative for Quiroste (MLE = −0.224), Watsonville (MLE = −0.117), Swanton (MLE = −0.113), and Seattle (MLE = −0.070). All estimates were shown with 99 percent certainty to be nonzero (Quiroste: range −0.117 to −0.368; Watsonville: range −0.227 to −0.035; Swanton: range −0.208 to −0.044; Seattle: range: −0.138 to −0.022). All independent runs of LAMARC provided similar results suggesting that the settings for the analyses were appropriate.

Discussion

Quiroste Valley hazelnuts yielded relatively high values of genetic diversity and there was a signal of large amounts of gene flow to the other populations in the area, mirroring the signal from all of the other Central California populations. The large values for genetic diversity coupled with the negative values of overall population growth indicate that Holocene hazelnut populations could have been much larger than they are now. Today, hazelnut populations attain the highest densities in frequently burned landscapes (Halpern 1989). Therefore, although our inference is weak, our findings are nevertheless broadly

congruent with the hypothesis that anthropogenic burning of landscapes at Quiroste promoted a larger population of hazelnuts in the past. However, it is important to emphasize that in our genetic results, it is difficult to disentangle the effects that fire suppression in recent times may have caused from older, Pleistocene and Holocene climate-driven distributional shifts of this species. For example, a recent genetic study of Chilean conifers concluded that the phylogeographic patterns they found were driven by both anthropogenic fire and global climate change (Souto et al. 2012).

The high degree of connectivity among all populations that we found is not surprising, given that hazelnuts are dioecious, wind pollinated, animal dispersed, and relatively long-lived woody plants, a combination of traits considered to result in low population structure over large geographic ranges (Loveless and Hamrick 1984). Nevertheless, we were able to recover some geographic structure for some Californian populations, especially at Quiroste Valley and Butano (Figure 2). The unique alleles found in each of these two populations were in contrast to the similar ancestral population signals shared by the rest of the California populations.

These results are notable for two reasons. First, since nuclear microsatellites in plants generally reflect pollen rather than seed dispersal, the fact that population structure is apparent in our results suggests that chloroplast alleles or microsatellites may yield even stronger signals of geographic structure. This is because hazelnut seeds almost certainly disperse smaller distances per generation than hazelnut pollen. Based on our results, we believe that future analysis of hazelnut populations with chloroplast DNA is warranted and is likely to give stronger within-population signals of demographic expansion or bottlenecks. Second, rare alleles like those found in the Butano and Quiroste Valley populations are often an indication of spatial isolation. Because the Santa Cruz mountains are so close to the current distributional limit of hazelnut, it is likely that during the warmer and drier climates of the mid Holocene, hazelnut population range contracted further north towards San Francisco and then re-expanded south within the past several thousand years. The protected, north-facing slopes within the Quiroste Valley and Butano watersheds may have harbored refugial populations of hazelnuts during these warmer and drier periods. Given that we are likely to again experience warmer climates in the near future, the Quiroste Valley and Butano areas may represent especially important habitat for this species.

Our population demographic analyses revealed no recent genetic bottleneck in hazelnuts, but instead large heterozygosity consistent with large, outbreeding populations. These results are in stark contrast to the current rarity of hazelnut

in Quiroste Valley and its low current abundance in the Central Coast of California. These contrasting observations might be reconciled by considering that the microsatellite data is likely giving us a signal of deeper time rather than the past 200–300 years. Even though individual hazelnut stems appear not to persist more than 10–20 years, they are clonal and could thus be extremely long-lived. The recent decimation of past large populations due to fire suppression may not be reflected in the genetic diversity of the Quiroste population if those individuals represent genetic entities that are more than a hundred years old and thus have not yet experienced any inbreeding.

Conclusions

Recently, the analysis of ancient DNA has allowed researchers to refine their understanding of populations of animals and plants useful to humans in deep time. For example, ancient DNA can help place archaeological samples into a genealogical framework, providing insight into the geographic and cultural origins of domesticated foods (Kistler and Shapiro 2011). Given the large number of hazelnut shells in the SMA-113 archaeobotanical assemblage (Cuthrell, this volume) we think that it is very likely that some of them will yield ancient DNA. This may allow us to place the ancient Quiroste Valley hazelnut population into the contemporary hazelnut population genetic structure that we have characterized with our present study and give more precise insight into the timing of Central Coast hazelnut demographic change and by extension, the extent of anthropogenic burning over time and space in this region.

We found that all hazelnut populations exhibited high levels of genetic variation and negative population growth consistent with large populations' sizes during the Holocene, with some degree of population decline over time. We also found subtle patterns of geographic structure suggesting that Quiroste Valley and neighboring Butano may have been important refugia during past climate warming events. These results provide an important foundation demonstrating that population genetic approaches can be successfully applied to research questions about indigenous landscape management. Results of this study also suggest several promising avenues for future work, including evaluation of hazelnut population genetic structure using chloroplast DNA and analysis of ancient DNA extracted from semi-carbonized archaeological hazelnut shells.

Acknowledgements

We would like to thank Valentin Lopez and the Amah Mutsun Tribal Band for their support of our research. We thank Chuck Striplen and Kent Lightfoot for their valuable suggestions. We would like to thank the Gordon and Betty Moore Foundation for funding the Berkeley Initiative for Global Change Biology, which funded this project. Finally, we would like to thank the following individuals and organizations for the necessary collection permits: California State Parks, Santa Cruz County Parks, Jim West, Greg Gilbert, the University of California Santa Cruz Natural Reserve System, and the Amah Mutsun Tribal Band.

Note

1. Hazelnut samples were collected only from Whitehouse Canyon in the upper portion of the Whitehouse Creek watershed. Quiroste Valley Cultural Preserve is located in the lower portion of the Whitehouse Creek watershed. In the text, hazelnut samples from this area are referred to as the "Quiroste Valley" hazelnut population, though Quiroste Valley Cultural Preserve does not currently contain any hazelnut.

References Cited

Anderson, M. Kat
2005 *Tending the Wild: Native American Knowledge and the Management of California's Natural Resources*. University of California Press, Berkeley, California.

Avise, John C.
2000 *Phylogeography: the History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Brown, Alan K. (editor)
2001 *A Description of Distant Roads: Original Journals of the First Expedition into California, 1769–1770, by Juan Crespi*. Edited and translated by Alan. K. Brown. San Diego State University Press, San Diego, California.

Bassil, Nahla V., Roberto Botta, and Shawn A. Mehlenbacher
2005 Microsatellite Markers in Hazelnut: Isolation, Characterization, and Cross-species Amplification. *Journal of American Horticultural Science* 130:543–549.

Boccacci, Paolo, Aziz Akkak, Nahla Bassil, Shawn A. Mehlenbacher, and Roberto Botta
2005 Characterization and Evaluation of Microsatellite Loci in European Hazelnut (*Corylus avellana* L.) and Their Transferability to Other *Corylus Species*. *Molecular Ecology Notes* 5:934–937.

Bonnell, Michael, and Robert K. Selander
1974 Elephant Seals: Genetic Variation and Near Extinction. *Science* 184:908–909.

Bouzat, Juan L., Harris A. Lewin, and Ken N. Paige
1998 The Ghost of Genetic Diversity Past: Historical DNA Analysis of the Greater Prairie Chicken. *The American Naturalist* 152:1–6.

Broughton, Jack M., R. Kelly Beck, Joan B. Coltrain, Dennis H. O'Rourke, and Alan Rogers
2013 A Late Holocene Population Bottleneck in California Tule Elk (*Cervus elaphus nannodes*): Provisional Support from Ancient DNA. *Journal of Archaeological Method and Theory* 20:495–524.

Cornuet, Jean-Marie, and Gordon Luikart
1996 Description and Power Analysis of Two Tests for Detecting Recent Population Bottlenecks from Allele Frequency Data. *Genetics* 144:2001–2014.

Erlandson, Jon M., Torben C. Rick, Terry L. Jones, and Judith F. Porcasi
2007 One If by Land, Two If by Sea: Who Were the First Californians? In *California Prehistory: Colonization, Culture, and Complexity*, edited by Terry L. Jones, and Kathryn Klar, pp. 53–63. Altamira Press, Lanham, Maryland.

Excoffier, Laurent, Peter E. Smouse, and Joseph M. Quattro
1992 Analysis of Molecular Variance Inferred from Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics* 131:479–491.

Halpern, Charles B.
1989 Early Successional Patterns of Forest Species: Interactions of Life History Traits and Disturbance. *Ecology* 70:704–720.

Hubisz, Melissa J., Daniel Falush, Matthew Stephens, and Jonathan K. Pritchard
2009 Inferring Weak Population Structure with the Assistance of Sample Group Information. *Molecular Ecology Resources* 9:1322–1332.

Kelly, Robert L.
1995 *The Foraging Spectrum, Diversity in Hunter-Gatherer Lifeways*. Smithsonian Institution Press, Washington, DC.

Kenyon, Karl W.
1969 *The Sea Otter in the Eastern Pacific Ocean*. North American Fauna No. 68, U.S. Fish and Wildlife Service, Washington, DC.

Kistler, Logan, and Beth Shapiro
2011 Ancient DNA Confirms a Local Origin of Domesticated Chenopod in Eastern North America. *Journal of Archaeological Science* 38:3459–3554.

Kuhner, Mary K.
2006 LAMARC 2.0: Maximum Likelihood and Bayesian Estimation of Population Parameters. *Bioinformatics* 22:76–770.

LaLande, Jeff, and Reg Pullen
1999 Burning for a "Fine and Beautiful Open Country": Native Uses of Fire in Southwestern Oregon. In *Indians, Fire and the Land in the Pacific Northwest*, edited by Robert Boyd, pp. 255–276. Oregon State University Press, Corvallis, Oregon.

Larson, Shawn, Ronald Jameson, Michael Etnier, Terry Jones, and Roberta Hall
2012 Genetic Diversity and Population Parameters of Sea Otters, Enhydra lutris, before Fur Trade Extirpation from 1741–1911. *PLoS ONE* 7(3):e32205. [doi:10.1371/journal.pone.0032205].

Levy, Sharon
2005 Rekindling Native Fires. *Bioscience* 55:303–308.

Loveless, Marilyn D., and James L. Hamrick
1984 Ecological Determinants of Genetic Structure in Plant Populations. *Annual Reviews in Ecology and Systematics* 15:65–95.

McLachlan, Jason S., James S. Clark, and Paul S. Manos
2005 Molecular Indicators of Tree Migration Capacity Under Rapid Climate Change. *Ecology* 86 (8):2088–2098.
Nei, Masatoshi, Takeo Maruyama, and Ranajit Chakraborty
1975 The Bottleneck Effect and Genetic Variability in Populations. *Evolution* 29:1–10.
Peakall, Rod, and Peter E. Smouse
2006 GENALEX 6: Genetic Analysis in Excel. Population Software for Teaching and Research. *Molecular Ecology Notes* 6:288–295.
Piry, Sylvain, Gordon Luikart, and Jean-Marie Cornuet
1999 Bottleneck: A Computer Program for Detecting Recent Reductions in the Effective Size Using Allele Frequency Data. *Journal of Heredity* 90:502–503.
Pritchard, Jonathan K., Matthew Stephens, and Peter Donnelly
2000 Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155:945–959.
Raymond, Michel, and Francois Rousset
1995 GENEPOP (version 1.2): Population Genetics Software for Exact Tests and Ecunmenicism. *Journal of Heredity* 86:248–249.
Rice, William R.
1989 Analyzing Tables of Statistical Tests. *Evolution* 43:223–225.
Roman, Joe, and Stephen R. Palumbi
2003 Whales Before Whaling in the North Atlantic. *Science* 301:508–510.
Rousset, Francois
2008 Genepop'007: A Complete Reimplementation of the Genepop Software for Windows and Linux. *Molecular Ecology Resources* 8:103–106.
Slatkin, Monty, and Nicholas H. Barton.
1985 Rare Alleles as Indicators of Gene Flow. *Evolution* 39:53–65.
Soulé, Michael. E.
1976 Allozyme Variation, Its Determinants in Space and Time. In *Molecular Evolution*, edited by Francisco J. Ayala, pp. 60–77. Sinauer Associates, Sunderland, Massachusetts.
Souto, Cintia P., Karin Heinemann, Thomas Kitzberger, Adrian C. Newton, and Andrea C. Premoli
2012 Genetic Diversity and Structure in *Austrocedrus chilensis* Populations: Implications for Dryland Forest Restoration. *Restoration Ecology* 20:568–575.
Turner, Nancy J.
1999 "Time to Burn": Traditional Use of Fire to Enhance Resource Production by Aboriginal Peoples in British Columbia. In *Indians, Fire and the Land in the Pacific Northwest*, edited by Robert Boyd, pp. 185–218. Oregon State University Press, Corvallis, Oregon.
U.S. Geological Survey
2013 Digital Representations of Tree Species Range Maps from "Atlas of United States Trees" by Elbert L. Little, Jr. (and Other Publications). Electronic document, http://esp.cr.usgs.gov/data/little/, accessed May 2013.
Van Oosterhout, Cock, William F. Hutchinson, Derek P. M. Wills, and Peter Shipley
2004 Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. *Molecular Ecology Notes* 4:535–538.