Title

Following the rivers: historical reconstruction of California voles Microtus californicus (Rodentia: Cricetidae) in the deserts of eastern California

Journal

Biological Journal of the Linnean Society, 119(1)

ISSN

0024-4066

Authors

Conroy, Chris J
Patton, James L
Lim, Marisa CW
et al.

Publication Date

2016-09-01

DOI

10.1111/bij.12808

Peer reviewed

The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2016, **119**, 80–98. With 6 figures.

Following the rivers: historical reconstruction of California voles *Microtus californicus* (Rodentia: Cricetidae) in the deserts of eastern California

CHRIS J. CONROY[1]*, JAMES L. PATTON[1], MARISA C. W. LIM[2], MARK A. PHUONG[3], BRETT PARMENTER[4] and SEBASTIAN HÖHNA[5]

[1]*Museum of Vertebrate Zoology, University of California, 3101 Valley Life Sciences Building, Berkeley, CA, 94720, USA*
[2]*Department of Ecology & Evolution, Stony Brook University, 650 Life Sciences Building, Stony Brook, NY, 11794-5245, USA*
[3]*Department of Ecology & Evolutionary Biology, University of California, 612 Charles E. Young Dr., South Los Angeles, CA, 90095, USA*
[4]*Department of Biology, Warren Wilson College, P.O. Box 9000, Asheville, NC, 28815, USA*
[5]*Department of Integrative Biology, University of California, 1005 Valley Life Sciences Building #3140, Berkeley, CA, 94720, USA*

Received 20 November 2015; revised 1 February 2016; accepted for publication 2 February 2016

The California vole, *Microtus californicus*, restricted to habitat patches where water is available nearly year-round, is a remnant of the mesic history of the southern Great Basin and Mojave deserts of eastern California. The history of voles in this region is a model for species-edge population dynamics through periods of climatic change. We sampled voles from the eastern deserts of California and examined variation in the mitochondrial *cytb* gene, three nuclear intron regions, and across 12 nuclear microsatellite markers. Samples are allocated to two mitochondrial clades: one associated with southern California and the other with central and northern California. The limited mtDNA structure largely recovers the geographical distribution, replicated by both nuclear introns and microsatellites. The most remote population, *Microtus californicus scirpensis* at Tecopa near Death Valley, was the most distinct. This population shares microsatellite alleles with both mtDNA clades, and both its northern clade nuclear introns and southern clade mtDNA sequences support a hybrid origin for this endangered population. The overall patterns support two major invasions into the desert through an ancient system of riparian corridors along streams and lake margins during the latter part of the Pleistocene followed by local *in situ* divergence subsequent to late Pleistocene and Holocene drying events. Changes in current water resource use could easily remove California voles from parts of the desert landscape. *Biological Journal of the Linnean Society*, 2016, **119**, 80–98.

KEYWORDS: California – climate – microsatellite – *Microtus californicus* – mtDNA – Mojave desert – Owens Valley – phylogeography – vole.

INTRODUCTION

The California vole (*Microtus californicus* Peale; Cricetidae, subfamily Arvicolinae) is a small, herbivorous rodent widespread in Mediterranean habitats (warm, dry summers, with cool but wet winters) on the west coast of North American, extending from northern Baja California in Mexico throughout most of California west of the Sierra Nevada to west-central Oregon (Hall, 1981) (Fig. 1, inset). The species extends in habitat and elevation from the Pacific coast at sea level, where populations inhabit tidal marshes with halophytic vegetation, through varied grasslands, to montane moist meadows at elevations up to 8000 feet (2400 m). Niche reconstructions indicate that precipitation variables rather than temperature *per se* most strongly influence the species' range limits (McGuire & Davis, 2013). It is thus

*Corresponding author. E-mail: ondatra@berkeley.edu

perhaps surprising that *M. californicus* extends into the dry southern Great Basin and Mojave deserts in eastern California. Here, populations are fragmented, typically confined to moist seeps, springs, and narrow riparian corridors, which are surrounded by otherwise inhospitable desert scrub vegetation. Desert vole populations are found along the eastern base of the Sierra Nevada, through the Owens Valley and adjacent enclosed intermontane basins (Deep Springs and Saline valleys), and along the northern slopes of the Transverse Ranges (San Gabriel and San Bernardino mountains), across the Mojave Desert to the east as far as Tecopa south of Death Valley National Park (Fig. 1).

Rangewide, *M. californicus* is composed of northern and southern mtDNA clades, each displaying limited phylogeographical structure (Conroy & Neuwald, 2008; Conroy & Gupta, 2011). Representatives of both clades have been found in the few desert populations included in these earlier studies. In the present study, we examined the pattern of genetic variation through mitochondrial and nuclear markers to clarify the historical spread of voles through the desert of eastern California. Our samples cover the known range of the species through this region, except Saline Valley where voles have not been seen since the early 1900s. We estimate the divergence time of those vole populations inhabiting the deserts and test hypotheses of potential dispersal corridors from likely source areas in the context of the late Pleistocene climatic cycles and hydrological reconstructions. Not surprisingly, the recovered history is

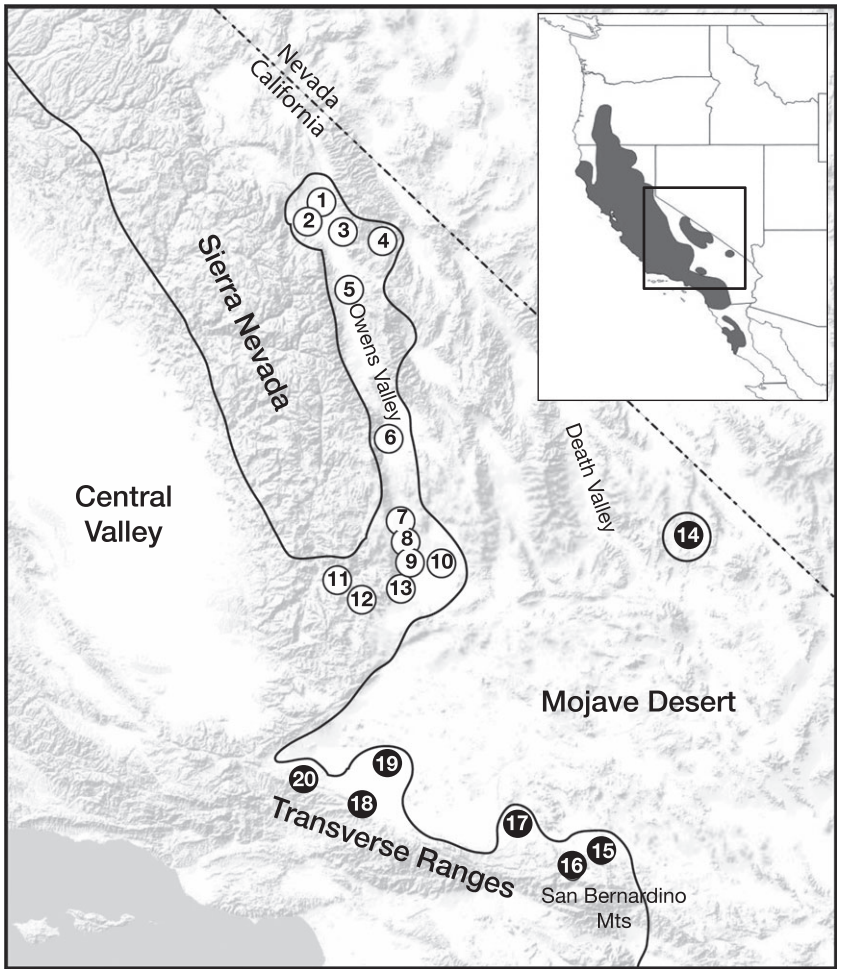

Figure 1. Map of general localities referred to in text and described in the Supporting information (Appendix S1). White circles identify populations belonging to the northern mtDNA clade, black circles represent southern mtDNA clade samples. 1, Fish Slough; 2, Lower Rock Creek; 3, Silver Canyon; 4, Deep Springs Valley; 5, Blackrock; 6, Cartago; 7, Little Lake; 8, Ninemile Canyon; 9, Sand Canyon; 10, Lark Seep; 11, Lake Isabella; 12, Piute Mts; 13, Soldier Wells; 14, Tecopa; 15, Cushenbury Springs; 16, Big Bear Lake; 17, Victorville; 18, Lake Palmdale; 19, Edwards Air Force Base; 20, Elizabeth Lake (see Supporting information, Appendix S1).

complex and confounded both by repeated wet–cool and warm–dry cycles over the past 0.15 Myr and the resultant episodes of vole expansion and retraction into localized refugia.

Taxonomy

Populations of *Microtus californicus* in the eastern California deserts are currently subdivided into three subspecies (Kellogg, 1918): *Microtus californicus vallicola* (type locality at Lone Pine, southern Owens Valley); *Microtus californicus mohavensis* (type locality at Victorville on the upper Mojave River at the northern base of the San Bernardino Mts); and *Microtus californicus scirpensis* (type locality at Shoshone on the Amargosa River east of Death Valley). In part because of the isolation of current populations, the subspecies *scirpensis* is Federally Listed as Endangered, whereas both *mohavensis* and *vallicola* are recognized as Species of Special Concern by the California Department of Fish and Wildlife. We treat populations irrespective of taxonomy, although we aim to integrate these molecular results with morphological analyses in a subsequent paper that addresses the systematic status of all known California vole populations in the eastern California deserts.

MATERIAL AND METHODS

Sample collection

Our samples are restricted to the desert region of eastern California, including Owens Valley to the east of the Sierra Nevada, to the mountainous area south of the Sierra Nevada, and to the Transverse Ranges, with the latter two the most likely historical source areas (Fig. 1; see also Supporting information, Appendix S1). For analysis, we grouped samples into 20 areas, and lumped smaller sampled areas if separate localities were geographically adjacent and part of a common hydrographic system; each group so defined is isolated from others by many square kilometers of inhospitable habitat.

Genetic data

We recovered genetic data from organ tissue of vouchered specimens or ear clips from released animals. We used mitochondrial cytochrome *b* (*cytb*) sequences for a deeper time, maternally inherited perspective and variation in nuclear DNA microsatellite markers for the finer scale necessary to test processes partitioning variation within and between populations. We also sequenced portions of the nuclear DBX gene intron 7, DBY gene intron 7, and AP5 introns to examine the origins of the Tecopa population (see Supporting information, Appendix S2).

Data analysis: mtDNA

We identified unique *cytb* haplotypes with TCS, version 1.21 (Clement, Posada & Crandall, 2000), with a 90% connection limit, and determined the clade membership of each in PAUP*, version 4.0b10 (Swofford, 2003) by sequence similarity. As reported, the clades differ by approximately 4.5% uncorrected *p*, and sequences within clades differ by < 1%. To establish clade monophyly, we estimated an unrooted topology using MrBayes, version 3.2.2 (Ronquist *et al.*, 2012). We chose a GTR + gamma substitution model and left all other settings at their default values. We found posterior probabilities of 1.0 supporting the monophyly of both the northern and southern clades based on the *cytb* sequences (see Supporting information, Appendix S3). Because all haplotypes of each clade are closely related, we depict population relationships by an unrooted Neighbour-joining (NJ) tree (Saitou & Nei, 1987) based on pairwise Slatkin's (1995) linearized F_{ST} calculated in ARLEQUIN, version 3.5 (Excoffier & Lischer, 2010). We used a Mantel test (IBDWS, version 3.23; Jensen, Bohonak & Kelley, 2005) to look for an association between genetic and geographical distances, with both linear and log distance between sites. This analysis compares F_{ST} matrices based on pairwise differences with 100 permutations and the geographical distances among localities on the surface of the WGS84 ellipsoid with a spreadsheet function based on Vincenty Formulas (Vincenty, 1975). To determine whether clusters of sample groups, delineated by microsatellite STRUCTURE analyses (see below), exhibited evidence of population stability or expansion, we used the mismatch distribution analysis in ARLEQUIN based on simple pairwise differences and 500 bootstrap replicates. We also used ARLEQUIN to estimate nucleotide diversity, as well as two statistics of selective neutrality, Fu's F_S (Fu, 1997) and Tajima's *D* (Tajima, 1989), for each population sample, with 1000 simulated samples.

Data analysis: nuclear DNA introns

Sequence data were input into PAUP* and compared with sequences from voles from other parts of California for indels and substitutions. Given very few differences, no phylogenetic or population genetic analyses were conducted.

Data analysis: nuclear DNA microsatellites

Data were first analyzed for anomalies such as null alleles (FREENA; Chapuis & Estoup, 2007; MICROCHECKER, Van Oosterhout *et al.*, 2004), linkage disequilibrium, and Hardy–Weinberg (HW) deviation

(ARLEQUIN, version 3.5). Linkage disequilibrium analysis was run with 10 000 permutations. For the HW analysis, we used 1000 steps in a Markov chain and 100 000 steps in a dememorization step. In an analysis of molecular variance (AMOVA), we computed population specific F_{IS} (an inbreeding coefficient). We used ARLEQUIN to generate basic diversity statistics and pairwise F_{ST} distances based on the number of different alleles. Finally, we tested for isolation-by-distance (IBD) using the same methods and geographical distances described for the mtDNA analysis. Analyses were run with 10 000 randomizations.

To test the hypothesis that voles lost diversity along hypothetical dispersal pathways, we contrasted two opposing scenarios: one where allelic richness is lost in successive populations with respect to distance from a source area and the second where all populations were initially equally diverse but lost diversity independently at approximately the same time following a retraction event. Probable source areas for desert populations of the southern mtDNA clade are the Transverse Ranges (localities 16, 18, and 20) (Fig. 1) and the southern Sierra Nevada (localities 11 and 12) for the northern mtDNA clade desert floor samples based on their geographical proximity. We used rarefied allelic richness estimates (A_R) (HP-RARE; Kalinowski, 2005), which accounts for unequal sample size, and tested for deviation from a negative relationship between richness and distance by regression analysis.

We used STRUCTURE, version 2.3 (Pritchard, Stephens & Donnelly, 2000) to determine microsatellite apportionment across sample sites. An initial analysis used the entire dataset of 407 voles and 20 population samples with $K = 1$–6, eight iterations per run, and a burn-in of 50 000 and a longer run of 500 000. We then analyzed each mtDNA clade separately [13 populations in the northern clade, six in the southern clade, excluding Tecopa (locality 14)], predefining each population and selecting the model for no admixture, with $K = 1$–13 for northern and $K = 1$–6 for southern mtDNA clade samples, each with 10 replicates, a burn-in of 50 000, and an additional 200 000 iterations. We used the Evanno, Regnaut & Goudet (2005) method implemented in STRUCTURE Harvester (Earl & von Holdt, 2012) to determine the most likely K given the rate of change between successive estimates. We used CLUMPP (Jakobsson & Rosenberg, 2007) and DISTRUCT (Rosenberg, 2004) with the Greedy option to summarize the results.

Timing of coalescent events

We are interested in estimating the timing of the mtDNA clades' origins in the region. Accordingly, we need a calibrated substitution rate for *cytb*. Unfortunately, there are no useful fossil or ancient DNA data for *M. californicus* to independently calibrate a rate. There are, however, numerous *cytb* substitution rates postulated for arvicoline rodents. Some span $0.3–0.5 \times 10^{-7}$ substitutions per site per year based on various fossil estimates (Brunhoff *et al.*, 2003). More recently, much faster estimates have been made, based on recent geological events, or combined with ancient DNA data. For example, Martínková *et al.* (2013) derived a substitution rate of 3.27×10^{-7} mutations per site per year for the invasion of the Orkney Islands by *M. arvalis* and Herman *et al.* (2014) derived a substitution rate of 4.572×10^{-7} mutations per site per year for the invasion of Scandinavia by *M. agrestis*.

Based on this rate range, we made two estimates of the time of the most recent common ancestor (tMRCA) of each mtDNA clade in the desert *M. californicus* using BEAST, version 1.8.0 (Drummond *et al.*, 2012). The first estimate is modified from Weksler, Lanier & Olson (2010), which is based on the range of previously published rates from arvicoline rodents, and implemented a Γ distribution with shape of 10 and a scale of 0.0066 that covers the typical range of variation in substitution rates for mammals based on deeper fossil record estimates. This analysis was run with all 56 unique mtDNA haplotypes, assuming a mean mtDNA substitution rate of 0.066 substitutions per site per Myr (or 13.1% divergence per Myr; Weksler *et al.*, 2010). We estimated parameters under a GTR + Γ substitution model, a constant size coalescent model as the tree prior, and an unweighted pair group method with arithmetic mean initial start tree. We assumed an uncorrelated lognormal (UCLN) relaxed clock with a Γ distribution with a shape parameter of 10 and a scale parameter of 0.006 as the hyperprior distribution on the mean parameter (ucld.mean) of the lognormal distribution. We used a chain length of 10 000 and burn-in of 1000. All other values were default values.

For a faster rate, we replaced the Γ hyperprior distribution on the ucld.mean by a Γ distribution with a shape of 20 and a scale of 0.02. This Γ distribution induces a mean substitution rate of 0.4 (equivalent to 80% divergence per Myr) and assigns 95% prior probability to the range in substitution rates from more recent European estimates based on ancient DNA and recent geological events (Herman & Searle, 2011; Martínková *et al.*, 2013; Herman *et al.*, 2014). We used a chain length of 50 000. We conducted both fast and slow analyses four times to look for anomalous results (nonconvergence). We examined all runs in TRACER, version 1.5 (Rambaut & Drummond, 2007) and confirmed that each parameter had an effective sample size greater than 200. We also

combined the results of the four runs. We repeated the two analyses assuming a strict molecular clock and an uncorrelated exponential clock with identical prior distributions on the mean clock rate to assess the impact of the molecular clock model on the divergence time estimates. Additionally, we estimated marginal likelihoods using stepping stone sampling to select between the different molecular clock models (Baele *et al.*, 2012; Xie *et al.*, 2011).

Historic gene flow

We used MIGRATE-N, version 3.6 (Beerli, 2006) to estimate Θ ($N_f\mu$ for mitochondrial genes and $4N_e\mu$ for nuclear genes) and the immigration parameter M (expressed as m/μ, where m is the immigration rate per generation and μ is the neutral mutation rate per site per generation) for both mtDNA and microsatellite datasets. We confined analyses to comparisons among geographically adjacent samples (Fig. 1), and included only populations with $N > 10$ individuals. We then estimated the strength of the recovered migration rate by multiplying M by the recipient population's Θ value to give $2Nm$ (for haploid data) or $4Nm$ (for diploid data). Values for either $2Nm$ or $4Nm$ above 1.0 suggest that migration is adequate to minimize differentiation by drift alone. Each analysis was run assuming a Brownian data model with default values, using a likelihood analysis of 10 initial chains sampling 500 genealogies, a sampling increment of 20, and 1000 trees discarded as burn-in. This sequence was followed by two chains, 100 000 trees, a sampling increment of 20, and 1000 trees discarded as burn-in.

RESULTS

Mitochondrial phylogeography

Across 414 sequences, we found 56 unique haplotypes: 25 from the northern clade and 31 from the southern clade (see Supporting information, Table S1, which lists unique haplotypes by sample with corresponding GenBank accession number). Several sequences were previously reported in Conroy & Neuwald (2008) as Genbank numbers EF506072, EF506066, EF506067, EF506081, EF506070, and EF506068 (see Supporting information, Table S1). Incomplete sequences were subsumed within the closest haplotype possible. Numbers of voles with a given haplotype ranged from 1 to 160 in the southern clade and 1 to 25 in the northern clade. Most haplotypes were unique to single population samples [42 of 56 (75%); 18 of 25 haplotypes in northern clade, 24 of 31 haplotypes in the southern clade], although the Tecopa sample [14] was the only one with completely unique haplotypes (five in this case). The number of haplotypes within populations ranged from two (Sand Canyon [9], Lark Seep [10]) to nine (Edwards AFB [19]). Gene diversity (see Supporting information, Table S2) ranged widely in the northern clade (0.111 in Sand Canyon [9] to 0.900 at Rock Creek [2]) with a relatively low mean (0.476); this measure was less variable (0.541 at Cushenbury Springs [15] to 0.857 Victorville [17]) and higher (mean = 0.705) in southern clade samples. The mean mismatch within a mtDNA clade was 2.914 in the northern clade and 5.835 in the southern clade. With several notable exceptions (e.g. Cushenbury Springs [15]), sample areas that are the smallest in extent and/or the most isolated (e.g. Antelope Springs in Deep Springs Valley [4], Cartago [6], Fish Slough [1], and Soldier Wells [13]) generally possess the lowest gene diversity levels. Slatkin's linearized F_{ST} ranged from 0 to 78.181 with a mean of 7.063 (SD = 10.013).

Each geographical sample contains haplotypes belonging to only one of the two previously recognized mtDNA clades (Conroy & Neuwald, 2008). Localities with northern clade membership include the Deep Springs Valley (sample 4) (Fig. 1), all Owens Valley (samples 1–3, 5–6), and all southern Sierra Nevada samples (samples 7–9, 11–13), as well as that which extends onto the desert floor at Lark Seep (sample 10). Southern clade samples include all of those from the Transverse Ranges (samples 16, 18, and 20) and southern margins of the Mojave Desert (samples 15, 17, and 19), as well as those from Tecopa (sample 14). Not surprisingly, the AMOVA analysis apportioned most of this variation to that between the major mtDNA groups (85.36%), with more variation found within populations (8.44%) than among those within each mtDNA group (6.2%).

The geographical pattern is evident among northern *cytb* clade members. Samples in northern (Fish Slough [1], Rock Creek [2], Silver Canyon [3]) and central (Blackrock [5]) Owens Valley plus nearby Deep Springs Valley [4] cluster in the middle of the unrooted NJ tree of mean F_{ST} values (Fig. 2, Table 1). Those from the southern Owens Valley (Cartago [6]), eastern Sierra Nevada slope (Little Lake [7] and Soldier Wells [13]), and Lark Seep [10] on the desert floor group together, whereas those samples from the Kern River Plateau [11], Piute Mts [12], and two eastern slope samples (Ninemile Canyon [8], Sand Canyon [9]) form a third group. Connections among these groups in the NJ tree (Fig. 2) do not reflect a nearest-Neighbour geographical sequence, nor do the pooled northern clade samples exhibit IBD (Mantel tests of F_{ST} vs. either km or log-km matrices are nonsignificant: $r = -0.157$, $P = 0.094$, and $r = -0.11$, $P = 0.141$, respectively).

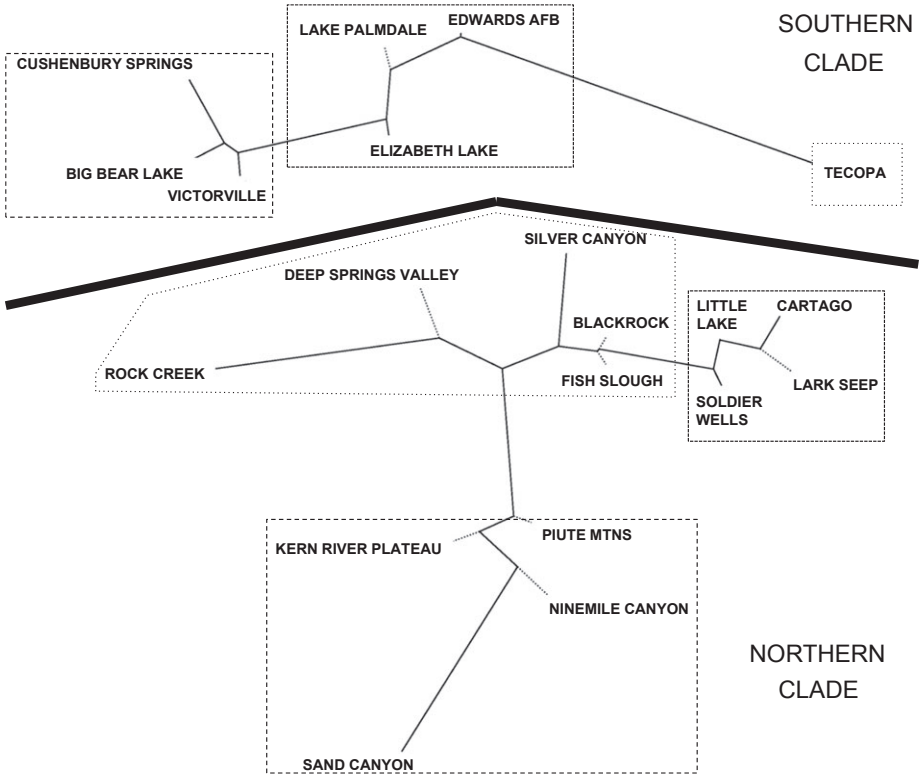

Figure 2. Unrooted Neighbour-joining trees for the southern and northern mtDNA clades, based on F_{ST} distances. Boxes with dashed lines indicate geographically cohesive groups.

The Tecopa sample [14] is very strongly differentiated from all other southern clade members in the unrooted Neighbour-joining tree based on F_{ST} distances (Fig. 2). The remaining samples fall into two groups, each forming a cluster of geographically, and hydrographically, connected populations: a western Lake Palmdale [18], Elizabeth Lake [20], and Edwards AFB [19] group vs. an eastern Big Bear Lake [16], Cushenbury Springs [15], and Victorville [17] group. The extreme divergence of the Tecopa sample and its closer connection to western, southern-clade samples (Elizabeth Lake, Lake Palmdale, and Edwards AFB) belies at least any direct, recent connection with populations to the south through the historic Mojave-Amargosa river nexus (see below). Cushenbury Springs and Big Bear Lake are geographically and, notably, genetically very close. Unlike the northern clade samples, those of the combined southern clade exhibit significant IBD (Mantel tests of F_{ST} vs. either km or log-km matrices are significant: $r = -0.810$, $P = 0.005$, and $r = -0.748$, $P = 0.006$, respectively).

In the *cytb* mismatch analysis, Elizabeth Lake, Lake Palmdale, Tecopa, Lake Isabella, and Piute Mts all deviated from the demographic expansion model ($P < 0.05$). Only Tecopa deviated from the spatial expansion with constant deme size ($P = 0.04$). However, all pooled samples defined by the best supported K levels from the STRUCTURE analysis (below) deviated from a model of spatial expansion, although two groups, the Lake Isabella–Piute Mts–Lark Seep and Lake Palmdale regions, also deviated from a demographic expansion model, suggesting stability in population size over time. Only a few populations exhibited significant deviations from neutrality, as indicated by Fu's F_S and Tajima's D tests (see Supporting information, Table S2). Tajima's D values were significant only in the Sand Canyon ($D = -2.034$, $P = 0.005$) and Silver Canyon samples ($D = -1.728$, $P = 0.02$), and only the Cartago sample was significant for Fu's F_S ($F_S = -2.98$, $P = 0.002$).

Within-sample microsatellite patterns

The mean number of microsatellite alleles within populations ranged from 2.857 at Tecopa to 9.818 at Edwards AFB (see Supporting information, Table S3). Observed heterozygosity ranged from 0.396 (Tecopa) to 0.823 (Elizabeth Lake). Gene diversity ranged from 0.236 (Tecopa) up to 0.747 (Elizabeth Lake). Deviations from HW equilibrium were

Table 1. Pairwise F_{ST} values for all 20 populations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 – Fish Slough	0.000	0.156	0.067	0.180	0.100	0.163	0.176	0.179	0.169	0.247	0.148	0.192	0.226	0.470	0.355	0.262	0.263	0.199	0.229	0.255
2 – Rock Creek	1.020	0.000	0.127	0.236	0.152	0.190	0.217	0.206	0.217	0.264	0.197	0.224	0.263	0.564	0.374	0.288	0.289	0.214	0.259	0.263
3 – Silver Canyon	0.157	1.469	0.000	0.173	0.062	0.130	0.166	0.157	0.173	0.229	0.145	0.173	0.218	0.481	0.349	0.265	0.254	0.189	0.224	0.246
4 – Deep Springs	0.010	0.322	0.181	0.000	0.138	0.198	0.197	0.221	0.234	0.287	0.218	0.235	0.282	0.500	0.374	0.312	0.305	0.267	0.285	0.314
5 – Blackrock	0.007	0.850	0.205	0.071	0.000	0.098	0.143	0.143	0.150	0.197	0.130	0.155	0.174	0.386	0.322	0.251	0.235	0.189	0.216	0.236
6 – Cartago	0.368	11.414	1.152	0.755	0.327	0.000	0.113	0.133	0.151	0.190	0.124	0.173	0.204	0.398	0.342	0.269	0.255	0.187	0.220	0.253
7 – Little Lake	0.190	4.449	0.274	0.437	0.227	0.057	0.000	0.086	0.112	0.167	0.100	0.160	0.159	0.424	0.326	0.235	0.229	0.181	0.213	0.232
8 – Ninemile Canyon	0.630	1.274	0.863	0.588	0.618	2.013	1.262	0.000	0.072	0.128	0.079	0.133	0.111	0.427	0.292	0.212	0.208	0.170	0.192	0.205
9 – Sand Canyon	3.529	5.464	5.410	2.153	2.168	17.790	8.421	0.293	0.000	0.134	0.069	0.112	0.129	0.441	0.312	0.223	0.233	0.176	0.199	0.215
10 – Lark Seep	0.171	6.275	0.679	0.409	0.181	0.000	0.033	1.166	12.916	0.000	0.159	0.199	0.206	0.504	0.356	0.277	0.301	0.222	0.258	0.257
11 – Lake Isabella	0.475	1.095	0.664	0.466	0.483	1.752	1.038	0.000	0.398	0.949	0.000	0.059	0.079	0.392	0.276	0.181	0.177	0.146	0.170	0.189
12 – Piute Mts	0.394	1.009	0.564	0.413	0.425	1.434	0.860	0.018	0.559	0.763	0.000	0.000	0.132	0.394	0.282	0.193	0.217	0.160	0.193	0.201
13 – Soldier Wells	0.173	4.547	0.589	0.406	0.188	0.134	0.105	1.102	9.481	0.073	0.892	0.721	0.000	0.490	0.320	0.237	0.229	0.195	0.224	0.227
14 – Tecopa	22.071	26.949	28.523	14.600	13.434	78.181	41.239	15.089	41.666	56.175	14.881	13.870	44.730	0.000	0.577	0.513	0.527	0.448	0.455	0.502
15 – Cushenbury Springs	9.512	8.510	9.867	8.263	8.621	24.495	15.133	8.366	13.179	14.399	7.875	7.590	13.446	3.589	0.000	0.093	0.160	0.175	0.167	0.170
16 – Big Bear Lake	9.772	8.833	10.106	8.464	8.795	24.294	15.238	8.574	13.394	14.595	8.093	7.821	13.670	3.736	0.151	0.000	0.106	0.087	0.090	0.078
17 – Victorville	9.428	8.017	9.812	7.938	8.424	30.496	16.950	8.088	14.406	16.426	7.481	7.149	14.852	3.780	0.209	0.047	0.000	0.093	0.097	0.125
18 – Lake Palmdale	5.639	4.629	5.492	5.501	6.016	13.796	8.712	5.701	7.628	7.641	5.222	5.083	7.335	0.813	0.451	0.427	0.280	0.000	0.062	0.048
19 – Edwards AFB	6.477	5.737	6.416	6.246	6.605	13.053	9.031	6.436	8.248	8.225	6.048	5.915	7.981	0.726	0.548	0.680	0.481	0.103	0.000	0.066
20 – Elizabeth Lake	7.873	6.562	7.973	6.973	7.475	23.533	13.568	7.199	11.739	12.585	6.616	6.353	11.639	1.597	0.469	0.358	0.237	0.013	0.222	0.000

Slatkin's linearized F_{ST} values below diagonal and F_{ST} based on number of different alleles above the diagonal.

observed in 51 cases, although these were scattered across the pool of 12 loci and 20 samples (data not shown). Presumptive null alleles were present in seven of the 12 loci, although null alleles were only found in more than one or two samples for locus Mar105 (12 of 20 samples in this case). Overall, we ignored this issue because a single locus is not expected to influence our general findings (Carlsson, 2008) and removing it (not shown) changed neither individual measures, nor the STRUCTURE analysis.

AMOVA results partitioned most allelic variation in both clades to that among individuals (74% and 81%, respectively, for southern and northern clade samples), followed by among-populations (24% and 16%), and then among-individuals within populations (2% and 3%).

The inbreeding measure F_{IS} was significant only in the Lake Palmdale ($F_{IS} = 0.079$, $P = 0.0117$) and Victorville ($F_{IS} = 0.117$, $P = 0.012$) samples belonging to the southern mtDNA clade (see Supporting information, Table S4). Twice as many northern clade samples exhibited significant values, including Blackrock ($F_{IS} = 0.925$, $P = 0.001$), Silver Canyon ($F_{IS} = 0.132$, $P = 0.024$), Fish Slough ($F_{IS} = 0.123$, $P = 0.022$), and Lake Isabella ($F_{IS} = 0.084$, $P = 0.018$).

Nuclear intron patterns

We found that the Tecopa population consistently shows the northern genotypes at the DBX ($N = 19$), DBY ($N = 12$), and AP5 ($N = 19$) markers, in contrast to their consistently southern ($N = 33$) mtDNA (see Supporting information, Table S5). We also typed samples from northern mtDNA clade populations Antelope Springs, Carroll Creek, Cartago, Little Lake, Lubken Creek, Olancha Creek, Piute Mountains, Sand Canyon, Silver Canyon, Soldier Wells, and Wyman Canyon (in Deep Springs Valley). These populations shared northern nuclear genotypes (see Supporting information, Table S6). We typed samples from southern mtDNA clade populations Big Bear Lake, Edwards AFB, Lake Hughes, and Victorville, and all had southern genotypes.

Microsatellite geographical patterns

The northern clade samples (Table 2) possess slightly fewer alleles per locus than do those of the southern clade (mean of 12.25 vs. 14.83), and have approximately half as many unique, or private, alleles per clade (2.92 vs. 5.58). The number of alleles observed per locus ranged from five in Mar080 up to 30 in Mar049.

Mantel tests of the association of pairwise sample F_{ST} with geographical distance for each mitochondrial clade were significant (for the northern clade: $r = -0.729$, $P < 0.001$ for km distance, and $r = -0.664$, $P < 0.0001$ for log km distance; for the southern clade: $r = -0.819$, $P = 0.016$ for km distance, and $r = -0.670$, $P = 0.004$ for log km distance). Even if the Tecopa sample is removed from the southern clade group of populations, a significant negative relationship between genetic and geographical distance remains ($r = -0.474$, $P = 0.02$, and $r = -0.471$, $P = 0.006$).

The ΔK method of Evanno *et al.* (2005) supported $K = 4$ groups in our initial STRUCTURE analysis that combined all samples (Fig. 3). Delineated groups include (1) the Tecopa sample (locality 14) (Fig. 1); (2) all southern mtDNA clade populations in the

Table 2. Summary of all microsatellite alleles present in all southern and northern mtDNA clade samples, and those unique (private) to either the southern or northern clades

Marker	Number of alleles observed	Number of alleles observed in northern clade	Number of alleles observed in southern clade	Number of alleles shared between clades	Number of alleles unique to northern clade	Number of alleles unique to southern clade
MSMM3	14	10	13	9	1	4
MSMM5	15	9	14	8	1	6
MSMM6	11	2	10	1	1	9
Ma36	21	17	19	15	2	4
Moe7	23	21	18	16	5	2
Moe2	19	16	15	12	4	3
Mar080	5	4	3	2	2	1
MSMM2	22	15	20	13	2	7
Mar105	15	13	10	8	5	2
Mar049	30	19	22	11	7	11
Mar063	20	9	18	7	2	11
Ma54	19	12	16	9	3	7
Mean	17.83	12.25	14.83	9.25	2.92	5.58

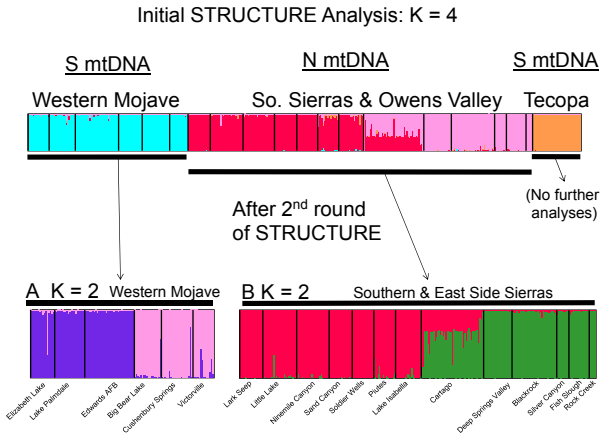

Figure 3. Results from STRUCTURE analysis, showing the initial run with all samples (above), then two subsequent runs with (A) southern (S) mtDNA group except Tecopa (bottom, left) and (B) northern (N) mtDNA group (bottom, right).

western Mojave desert (localities 15–20); (3) samples from Lake Isabella and the Piute Mts plus those from the adjacent eastern slopes and desert floor (localities 7–13); and (4) most of those in the Owens Valley plus Deep Springs Valley (localities 1–5). Samples from the southernmost Owens Valley (Cartago at Owens Lake and streams in the Lone Pine area to the immediate north; locality 6) include admixed individuals between groups 3 and 4, which are geographically adjacent to the immediate north and south, respectively.

Separate STRUCTURE analyses for each mtDNA clade, excluding Tecopa, indicated $K = 2$ as most likely. For the northern clade, the Cartago + Lone Pine samples (locality 6) are again comprised of individuals of mixed ancestry with respect to those north and south. Deep Springs Valley, which is now completely isolated from the Owens Valley, nonetheless is indistinguishable from adjacent samples in the northern Owens Valley. The southern mtDNA clade also comprised two subgroups: one composed of samples at the base of the western-most Transverse Ranges (Elizabeth Lake, Lake Palmdale, and Edwards AFB) and the other combining samples from the San Bernardino Mts with those to the north at Victorville and Cushenbury Springs. This pair of subgroups coincides with the drainage patterns in both areas.

There is no relationship between allelic richness (A_R) and geographical distance across all northern mtDNA clade samples ($r^2 = 0.1664$, $P = 0.1665$) or for each of the geographical subgroups identified by STRUCTURE (northern Owens Valley: $r = 0.556$, $P = 0.089$; southern Sierras and adjacent localities: $r^2 = 0.489$, $P = 0.056$). By contrast, all southern mtDNA clade samples taken together did exhibit a significant negative trend ($r^2 = 0.684$, $P = 0.022$), supporting a progressively eastward pattern of declining allelic diversity from Elizabeth Lake in the west to Tecopa in the east. Treating the San Bernardino Mts (Big Bear) as the point of origin for the Tecopa did not provide a similar pattern ($r^2 = 0.094$, $P = 0.503$).

Time of diversification and gene flow estimates among regions

We found that the point estimates for the UCLN and strict models were almost identical (see Supporting information, Appendix S4). Furthermore, we found that, among the UCLN, strict, and uncorrelated exponential clock models, the UCLN model was slightly better supported than the others based on

Table 3. Summary of BEAST runs to estimate time of the most recent common ancestor (tMRCA) of mtDNA haplotypes

Run	North tMRCA	HPD lower	HPD upper	ESS	South tMRCA	HPD lower	HPD upper	ESS
1	0.163	0.050	0.300	2757.7	0.138	0.047	0.253	2525.7
2	0.161	0.050	0.300	2627.5	0.137	0.049	0.252	2417.8
3	0.164	0.059	0.310	2391.5	0.140	0.047	0.254	2416.6
4	0.162	0.058	0.303	2853.1	0.137	0.048	0.250	2464.6
Combined	0.163	0.054	0.304	10 631.1	0.138	0.049	0.254	10 015.6
1	0.023	0.010	0.040	14 614.5	0.020	0.009	0.032	15 352.8
2	0.023	0.010	0.039	16 331.6	0.020	0.009	0.033	15 772.8
3	0.023	0.0098	0.039	16 423.2	0.020	0.009	0.033	16 243.5
4	0.023	0.010	0.040	15 844.2	0.0196	0.009	0.033	15 446.9
Combined	0.023	0.010	0.040	63 097	0.0196	0.009	0.033	62 842.5

Upper: slower rate. Lower: faster rate. All time estimates are in Myr.ESS, effective sample size; HPD, highest posterior density.

Bayes factors (see Supporting information, Table S7). Thus, we report only the results from the UCLN model. The two substitution rates that we used in the UCLN model gave substantially different estimates of the tMRCA for each mtDNA clade. The results from four runs in each time estimate varied very little. From the deeper, slower rate estimate (Table 3), sequences within the northern mtDNA clade coalesced at approximately 0.163 Myr [95% highest posterior density (HPD): 0.054–0.304 Myr], and the southern mtDNA clade coalesced at approximately 0.138 Myr (95% HPD: 0.049–0.254 Myr). These dates place the initial diversification in both clades near the second to last glacial interval (Illinoian) and well before the beginning of the Holocene (Fig. 4). These values contrast strongly with those derived from the faster rate (see chronograms in the Supporting information, Appendix S5). In the faster rate, the northern clade sequences tMRCA was estimated at 23 300 years BP (95% HPD: 10 000–40 000 years BP), and the southern clade tMRCA was estimated at 19 600 years BP (95% HPD: 9000–33 000 years BP). These ranges of dates coincide with the Last Glacial Maximum, 26 000 to 18 000 years BP (Garcia *et al.*, 2014) into the early Holocene.

Migration rates among all pairs of geographically adjacent samples, depicted as 2*Nm* (for haploid *cytb* haplotypes) and 4*Nm* (for diploid microsatellites) derived from the MIGRATE analyses, are shown in Figure 5. Pairwise connections are drawn only where estimated rates are above 1.0, which is the theoretical threshold where gene flow is sufficient to counteract the effects of local drift, and the strength of emigration into and immigration from for each pair of localities is indicated by the size and direction of each arrow. For the *cytb* data (Fig 5, left), of the 70 total emigration/immigration possibilities (35 pairwise comparisons), 43 (61%) had *M*-values > 1.0, although only 16 (23%) were sufficiently strong to yield *2Nm* values also > 1.0. Migration connections, either one-way or both ways, are present among geographically localized samples that mirror the STRUCTURE results. Samples from the northern Owens Valley (including Deep Springs Valley) exhibit one- or two-way connections, albeit relatively weak ones, as do the three samples in the Kern River Plateau and vicinity (Lake Isabella, Piute Mts, and Soldier Wells). The cluster of samples from Cartago (in the southern Owens Valley) combined with those along the eastern flank of mountains to the south and Lark Seep at China Lake generally group more strongly. All three of these regional groups with northern mtDNA clade haplotypes are unconnected, always within 2*Nm* values < 1.0. Notably, although not surprisingly, these northern mtDNA clade samples also exhibit no connection to any southern clade locality. Among the latter, each triad of samples from the western Transverse Ranges (Elizabeth Lake, Lake Palmdale, and Edwards AFB) is interconnected,

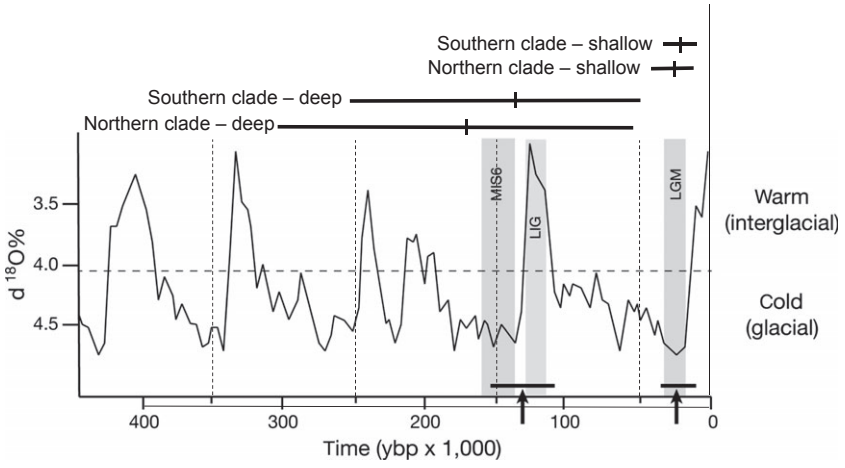

Figure 4. The benthic $d^{18}O$ record, which reflects changes in seawater temperature during climatic oscillations and thus depicts the major glacial-interglacial episodes over the past 450 000 years (Cohen & Gibbard, 2011): LGM, last glacial maximum (Wisconsin); LIG, last interglacial (Sangamonian); and MIS6, penultimate glacial maximum (Illinoian). Ranges in mean estimates of the time of the most recent common ancestor for mtDNA haplotypes for both the northern and southern clades in the eastern California deserts are depicted by the vertical bars spanning horizontal lines above the temperature profile (for the estimated range bounding each mean, see text). Time intervals of significance to the potential pathways across the desert floor to form the modern populations are indicated by the horizontal lines and arrows below (Fig. 6).

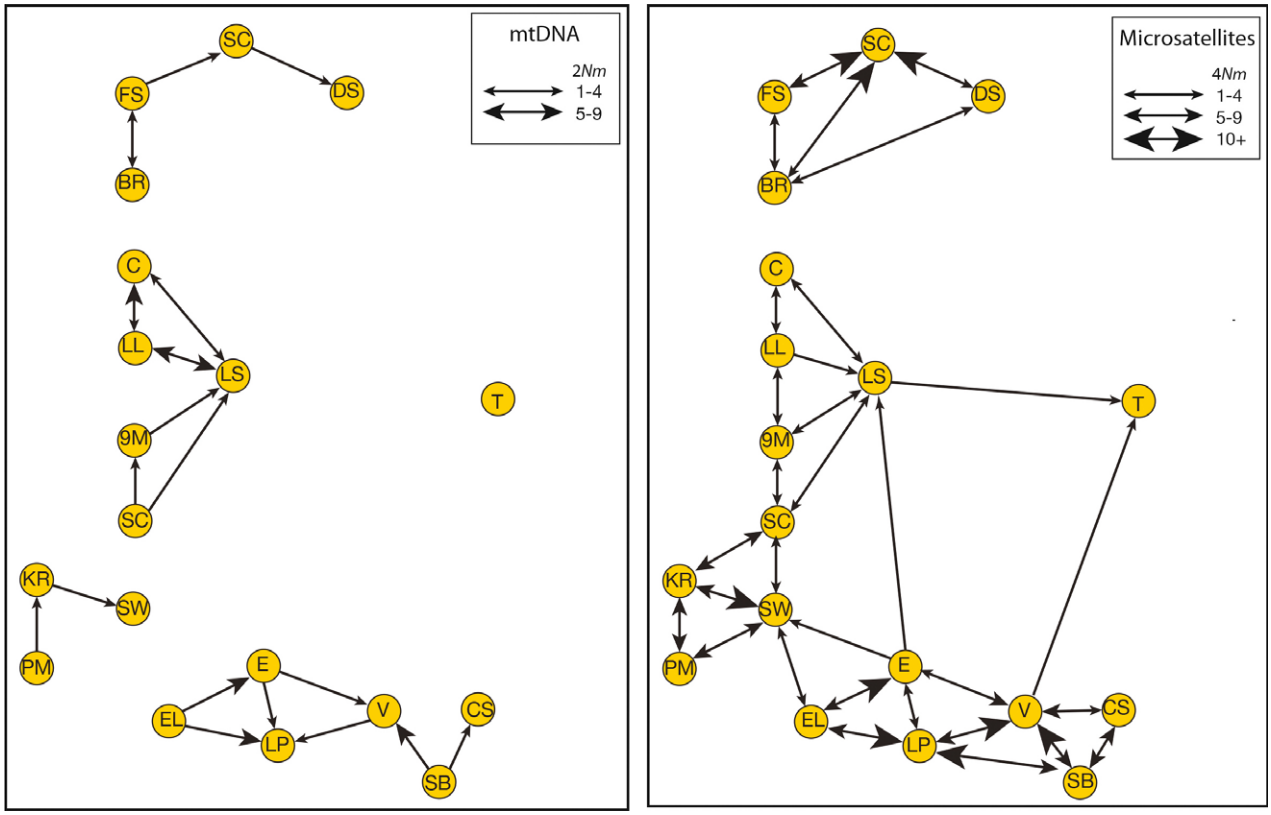

Figure 5. The pattern of gene flow between each pair of geographically adjacent samples where $N > 10$. Arrows indicate the direction of connectivity, and their size the strength of migration, given as $M \times \Theta$, or $2Nm$ for haploid mtDNA (left) and $4Nm$ for the 12 microsatellite loci (right); for further explanation, see text. Only values > 1.0, the theoretical limit above which differentiation by drift alone is minimized, are shown. Abbreviations refer to named populations in Figure 1.

although mostly in only single directions, as are those from the San Bernardino Mts and the adjacent desert floor (Big Bear Lake, Victorville, and Cushenbury Springs). There is also evidence of interchange between these western and eastern triad samples, between Edwards AFB unidirectionally into Victorville, and between Victorville and Lake Palmdale. Finally, and importantly, the Tecopa sample is unconnected to any other sample area, including those that also have southern *cytb* haplotypes.

The microsatellites indicate greater, and more bidirectional, population connections between geographically adjacent sample pairs (Fig. 5, right). *M*-values for 67 of the possible 70 comparisons are above 1.0 (96%), and 58 comparisons (83%) also exhibit $4Nm$ values > 1.0. Overall, the geographical pattern of emigration–immigration rates is similar to that for mtDNA (compare right and left in Fig. 5). Among the northern mtDNA clade samples, relatively high $4Nm$ values are clustered among those from the northern Owens Valley (including Deep Springs Valley) and the triad of samples in the Kern River Plateau. Samples from the southern Owens Valley + eastern mountain slopes + Lark Seep are also all interconnected but at lower $4Nm$ values. Among southern mtDNA clade samples, the three from the western Transverse Ranges exhibit especially strong interconnections, whereas those in the San Bernardino Mts and vicinity are somewhat less. Connections also exist between both the northern mtDNA clade Kern River Plateau triad and the adjacent southern clade group from the Transverse Range. Importantly, there is supportable gene flow input into the Tecopa population from both the northern mtDNA clade population at Lark Seep and the southern clade sample at Victorville, albeit relatively weakly so in each case.

DISCUSSION

Origin, direction, and timing of vole colonization in the Mojave Desert

The mesic-adapted California voles ultimately derive from populations in coastal southern and central California, and likely spread through the desert during

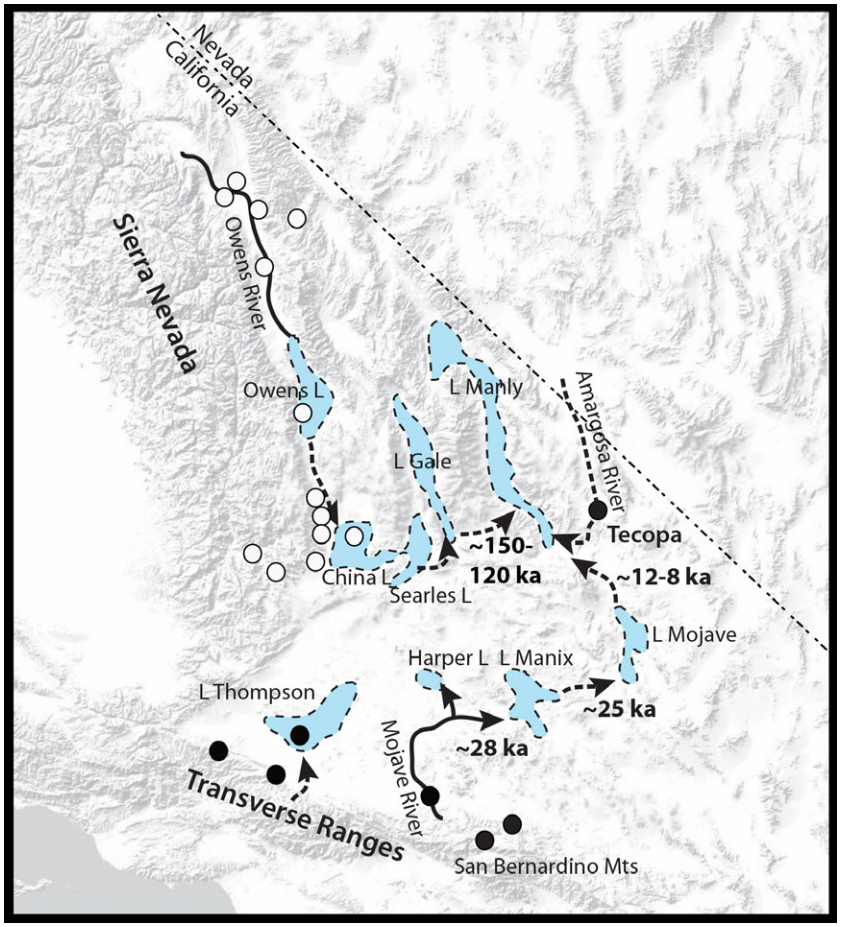

Figure 6. Palaeo-hydrological systems of the desert region of eastern California, including the Owens Valley east of the crest of the Sierra Nevada and the western Mojave Desert. The Owens, Mojave, and Amargosa rivers are identified, as are each lake basin, including Pleistocene Lake Gale in Panamint Valley, Pleistocene Lake Manly in Death Valley, Pleistocene Lake Thompson (current Rosamond, Buckhorn, and Rogers dry lakes), Pleistocene Lake Manix (Troy and Coyote dry lakes), and Pleistocene Lake Mojave (Soda dry lake). Dashed lines and arrows identify the direction of historical water flow connecting lake basins, based on palaeo-hydrological reconstructions. The approximate dates when lake highstand was sufficient for spillover from Owens Lake in the north-west through the successive intervening lake basins into Death Valley, and from both the Amargosa drainage and Mojave River into Death Valley from the north-east and south-west, respectively, are shown in bold. Open circles indicate sample sites for northern mtDNA clade voles; closed circles indicate southern mtDNA clade voles (Fig. 1). For sources of ages, see text.

cooler, moister periods along a network of rivers and streams. These mesic patches are now surrounded by desert scrub vegetation, an assemblage of species that did not form until early to mid Holocene, with the Mojave Desert creosote bush not arriving until 10 000 to 5000 years BP (Grayson, 2011).

Episodes of vole colonization of the desert floor likely occurred more than once, at different time intervals and through separate drainage systems, with the result that the history of earlier vole invasions likely has been over-written by later episodes. Although some hydrological systems are isolated today, some were interconnected at various times during past mesic periods, which themselves were of different temporal duration. Although, as described below, some of palaeo-hydrological connections are reasonably well understood, and thus provide the basis for hypothesized routes and times of vole dispersal across what is now desert landscape, other biogeographical events resulting in the current distribution pattern of California voles in the California desert are far from clear. For example, the route by which the isolated Deep Springs and Saline valleys of the White-Inyo range were colonized remains an enigma.

Mitochondrial clade structure (Fig. 2), microsatellite apportionment (Fig. 3), patterns of nearest-Neighbour migration rates (Fig. 5), and estimates of

the time of the tMRCA all support colonization of the eastern California deserts from a minimum of two separate source areas within the last 300 000 years BP. Our deeper time estimates suggest the northern mtDNA clade was likely derived from populations at the southern terminus of the Sierra Nevada at a mean tMRCA of approximately 163 000 years BP (95% HPD: 0.054–0.304 Myr); those of the southern mtDNA clade spread out from the flanking Transverse Ranges slightly later: approximately 138 000 years BP. These time estimates are coincidental with the Second to Last Glacial Maximum (Fig. 4). The more recent estimates, based on a much faster substitution rate, suggested a tMRCA of the northern samples at approximately 23 300 years BP (95% HPD: 10 000–40 000 years BP), and the southern samples at 19 500 years BP (95% HPD: 9000–33 000 years BP). This corresponds well to the Last Glacial Maximum (26 500–19 000 years BP) (Clark *et al.*, 2009). Both sets of estimates would suggest invasion during cooler periods. However, both cannot be correct.

Origin and spread of the northern mtDNA clade

Our a priori hypothesis is that the palaeo-Owens River drainage system (including Deep Springs Valley) was colonized from a source population probably located in the combined Kern River Plateau-Piute Mts region at the southern end of the Sierra Nevada. Invasion from the north is unlikely given the higher mountain passes (all > 3 100 m) that voles would have had to overcome. Because the vole populations that comprise the northern clade are arranged linearly from south to north, a one-way stepping-stone model from that source predicts a simple IBD pattern. This pattern characterizes nuclear DNA for this set of samples not only both with respect to IBD of F_{ST} on distance, but also for south-north trends in allelic richness, although not their mtDNA. Furthermore, interconnections between nearest-Neighbour sample pairs are neither uniform in extent or direction for either *cytb* sequences or microsatellites (Fig. 5). The three foci of localized but strongly geographical migration interconnections (northern Owens Valley, including neighbouring, but geologically isolated Deep Springs Valley; southern Owens Valley, slope, and desert floor at Lark Seep samples; and Kern River Plateau and Piute Mts), with limited connections between them, also belie a simple south-to-north historical spread of voles through this large area. The substantial time depth of the northern mtDNA clade and compounding cycles of population expansion and retraction coincidental with pluvial–interpluvial cycles, has likely obscured the early history of colonization. The patterns for both sets of molecular markers that remain likely reflect the more recent periods of exchange. One lasting and important signature of exchange that warrants special mention is the strong and progressive north-to-south series of connections following the length of the Owens Valley, down the foothills on the eastern slope and onto the floor of China Lake. The likely wetland connection between Owens Lake and China Lake down the eastern side of the Sierra Nevada, including its feeder streams, was not severed until sometime after 16 500 years BP. Today, no suitable habitat connects any of the populations of voles now present in this narrow region.

Origin and spread of the southern mtDNA clade

Our a priori hypothesis for this region is similar to that posited for the northern clade in that the Transverse Ranges would likely have served as the source for all southern clade populations on the floor of the Mojave Desert, including that at Tecopa. However, because of separate drainage systems, we supposed that western Mojave populations were colonized from the adjacent San Gabriel Mts, whereas those to the east ultimately were founded from the San Bernardino Mts. All data and analyses support this dual origin scenario, although remnants of some strong migration interconnections between the two areas remain. In part, both patterns reflect the more recent history of the wetlands connecting areas, and thus the spread of voles among them in comparison with northern clade samples.

The generally high migration rates between pairs of the western-most triad of populations (Elizabeth Lake-Lake Palmdale-Edwards AFB; Fig. 5) are consistent with estimates of long-term population stability for this group measured by the mismatch distribution. Importantly, the levels of migration connecting these populations are either unidirectional from Lake Elizabeth (mtDNA data; Fig. 5, left) or strongly skewed outward from Lake Elizabeth (microsatellite data; Fig. 5 right), with the latter completely concordant with declining allelic diversity away from this site. Although the more eastern of the two Mojave groups (Cushenbury Springs/Big Bear Lake/Victorville) did not deviate significantly from a demographic expansion model, a stable history was only barely rejected ($P = 0.056$). There is extensive habitat for voles in the San Bernardino Mountains and this is likely a stable region for the species. None of the reviewed geological reports for the palaeo-Mojave River drainage has examined interconnections with Lake Thompson to the west, although that connection is supported by the

migration rate patterns for both mitochondrial and nuclear markers (Fig. 5). The above observations agree with our null hypothesis of voles populating the southern Mojave Desert in a gradual stepping stone fashion along existing stream channels and the current populations not merely being remnants of an earlier, single panmictic population that decayed into the isolated patches of today.

As we summarize briefly below, a deeper understanding of the Quaternary hydrological history of this region is necessary to interpret our findings. The hydrological history aids us in identifying probable routes of colonization from source areas, as well as the time periods when voles most likely spread and then subsequently retracted into their currently isolated habitat patches.

The eastern California deserts experienced cyclical cool–wet and warm–dry periods over the Pleistocene coincident with glacial–interglacial periods (Fig. 4) (Phillips, 2008). Meltwater from alpine glaciers in the Sierra Nevada and regional precipitation provided a continual input of surface water into the Owens Valley to their east (Koehler & Anderson, 1994; Orme & Orme, 2008). Increased winter snowpack during glacial periods in the mountains south of the Sierra Nevada, the Transverse Ranges, and high desert ranges also served as an expanded source of water into the Mojave Desert (Miller *et al.*, 2010). These large, climate-driven fluctuations in the regional water balance resulted in repeated interconnection and disconnection of various basins that make up the palaeo-Owens and palaeo-Mojave drainage systems and the higher water tables and emergent ground water that supported extensive wetlands (Pigati *et al.*, 2011), which represents an ideal habitat for expanding vole populations.

Vole populations currently occupy parts of three palaeo-hydrological systems in eastern California, the separate palaeo-Owens and Mojave drainages, as well as Pleistocene Lake Thompson in the westernmost part of the Mojave Desert. Although these systems and areas were concordantly affected by the cyclical shifts in precipitation and temperature, interconnections between the separate basins within each differed substantially in temporal pattern and duration, as most likely did the timing of vole colonization.

The current terminus of the Owens River is Owens Lake, which is now largely a dry bed (Fig. 6). Owens Lake, however, has overflowed to the south into China and Searles lakes periodically for the past 1.8 Mya and, for the last approximately 150 000 years BP, overflows have been tightly correlated with glacial advances in the Sierra Nevada (Phillips, 2008). Oxygen isotope data indicate that the period from 190 to 120 000 years BP was extremely wet with cool temperatures (Jayko *et al.*, 2008), resulting in a major influx of Owens River water via China–Searles lakes into Panamint Valley (Pleistocene Lake Gale), and the subsequent arrival of Owens River water into Death Valley (Pleistocene Lake Manly) over Wingate Pass and via Wingate Wash (Smith & Bischoff, 1997; Forester, Lowenstein & Spencer, 2005; Jayko *et al.*, 2008). Overflow connections between Owens Lake, China–Searles lakes, and Panamint Valley persisted periodically until approximately 23 500 years BP (95% HPD: 10 000–39 800 years BP), although none of these overspills produced sill elevations sufficient to connect to Death Valley (Jayko *et al.*, 2008). The Owens Valley–China and Searles lake connection was severed by uplift in the intervening Coso Range and subsidence in the Owens Valley before 16 500 years BP (Orme & Orme, 2008). These historical interbasin connections thus likely provided a dispersal corridor of suitable vegetation for northern mtDNA clade voles from the Owens Valley, and also probably the drainages stemming down the mountains south of the Sierra Nevada, through China Lake, and eastward as far as the Amargosa drainage, including Tecopa. The timing of these interconnections is completely consistent with the mean tMRCA of approximately 160 000 years BP (95% HPD: 0.054–0.304 Myr) estimated for the northern mtDNA clade with the slower substitution rate (Fig. 4). The faster substitution rate, with a tMRCA of 23 300 years BP (95% HPD: 10 000–39 800 years BP), would suggest a much faster colonization of the region, although this is still largely compatible with the geological evidence for colonization around the Owens Valley and China Lake region. Although earlier invasion was possible, this estimate would suggest that either earlier invasions might have been replaced by later ones, or the surviving haplotypes derive from only a small set of the earliest invaders. What is not compatible is the invasion of voles all the way to Tecopa when there was a hydrological connection sufficiently substantial to leave a record in the geological history. Voles may be able to persist and travel in moist habitats that did not leave a substantial geological record. Currently, voles persist in some very small patches of moist habitat that are not part of large river systems. A reassessment of the types of habitats necessary for the movement of voles, outside of large flowing rivers and lakes, will be needed to better test this hypothesis.

The history of the palaeo-Mojave hydrological system is similar to that of the palaeo-Owens system but differs in the timing of connections between the intervening basins (and their Pleistocene lakes) (Fig. 6). The Mojave River originated in the late

Pliocene-early Pleistocene with the tectonic uplift of the Transverse Ranges (approximately 1.94 Mya) (Jefferson, 2008), flowing downslope to fill the Victorville basin (Enzel, Wells & Lancaster, 2003). By 0.5 Myr, the river terminated in the Victorville or perhaps the Harper Lake basins, and periodically, between approximately 130 000 and 28 000 years BP, extended east to terminate in Pleistocene Lake Manix (Knott *et al.*, 2008). Highstands in Lake Manix were sufficient to cause overflow through Afton Canyon into the Soda Lake–Silver Lake basins, forming Lake Mojave between 50–35 000 years BP (Reheis & Redwine, 2008) and at approximately 25 000 years BP (Garcia *et al.*, 2014). Lake Manly in Death Valley included surface-water input from both the Amargosa and Mojave rivers at 18 000 years BP and 12 000 years BP (Anderson & Wells, 2003a,b), with the latter resulting from overflow of Lake Mojave between 13 700 and 11 400 years BP (Wells *et al.*, 2003) and as recent as 8000 years BP (Garcia *et al.*, 2014). This connection between the Mojave River through Pleistocene Lake Manix and Lake Mojave would have provided the corridor necessary for vole expansion to the present-day Amargosa drainage, as represented by our Tecopa sample. Our older tMRCA estimate of 0.138 Myr (95% HPD: 0.049–0.254 Myr) is easily compatible with this hypothesis, with an expansion of southern clade voles beginning over a 100 000 years BP, primed for expansion into the Amargosa drainage. If the faster rate employed in the BEAST analysis is more accurate, this would suggest a much shorter period of invasion, from 19 500 years BP, arriving in the Amargosa drainage by approximately 8000 years BP. However, there are large degrees of error with these date estimates.

Less is understood about the temporal depth of connections between the western Transverse Ranges and Rosamond, Buckhorn, and Rogers dry lakes of today that comprised Pleistocene Lake Thompson (Fig. 6). Orme (2008) documented a perennial lake at periods between 36 000 to at least 34 000 years BP, from before 26 000 to approximately 21 000 years BP, and between 16 200 and 12 600 years BP. The likely input source in each case would have been streams extending north from the slopes of the Transverse Range to the south (Fig. 6). Certainly, water flowed from this range via channels, such as Little Rock Creek, into the Rosamond lake basin during the Little Ice Age (AD 1650) of the Holocene (Miller *et al.*, 2010). These data suggest that voles extended onto the desert floor by the late Pleistocene and have had continual if sporadic access until recently. This hypothesis is supported by the low F_{ST} values between our Elizabeth Lake, Lake Palmdale, and Edwards AFB samples (Table 1) (depicted graphically in Fig. 2) and the relatively large migration rate measures (Fig. 5).

Origin of the Tecopa population

The Tecopa sample (locality 14) (Fig. 1) contains only unique mtDNA haplotypes that are nested within the southern clade (Fig. 2). Although this sample is sharply differentiated from both southern and northern samples in the overall apportionment of microsatellites (Fig. 3), it shares microsatellite alleles common to both mtDNA clades. Tecopa voles have two marker/allele combinations that are known only from southern mtDNA clade populations, seven combinations known only from the northern mtDNA clade, and two known only from Tecopa (see Supporting information, Table S6). Fourteen other marker/alleles combinations are present in both southern and northern clades. Although the Tecopa population formed its own group in the best supported STRUCTURE analysis (Fig. 3, with $K = 4$), Tecopa was linked to the northern mtDNA clade in a less well supported $K = 3$ analysis (data not shown). Furthermore, migration rates of $4Nm$ 1.0 for both mtDNA and microsatellites (Fig. 5) exhibit linkages between Tecopa and both northern (Lark Seep in China Lake) and southern mtDNA clade samples (Victorville). Additional nuclear intron data also supports a partial northern clade origin for the Tecopa voles. The intron data show a clear division agreeing with the mtDNA north–south division. Although we have not surveyed every population for these introns, it appears that Tecopa shares these genotypes as a result of a shared history.

Our working hypothesis is that voles entered the Amargosa drainage from two source areas: one from the northern mtDNA clade and one representing southern mtDNA clade members. The former occurred earlier, following the palaeo-Owens River connection through China Lake–Searles Lakes into Panamint Valley and hence into the Amargosa drainage via Lake Manly in Death Valley (Fig. 6). The southern, mitochondrial input was later, at the very end of the Pleistocene or early Holocene. Given the small effective population size of haploid genes, such as mtDNA, the differential timing of entry from separate source areas, and the severe retraction of vole habitat in the Amargosa drainage today, a fixation of southern mtDNA haplotypes but with retention of some evidence for the retention of northern mtDNA nuclear genes is not unexpected. One element that is incompatible with this hypothesis is the proposed timing using the more recent, faster substitution rate estimate in BEAST. This estimate suggested that the northern clade voles coalesced approximately 23 300 years BP (95% HPD: 10 000–39 800 years BP). However, the Wingate Pass con-

nection between Panamint Lake and Lake Manly likely did not exist after approximately 30 000 years BP (Anderson & Wells, 2003a). One explanation would be that the rate estimate needs adjustment. The mean is incompatible, although the 95% interval, 39 800–10 000 years BP would encompass possible dates for gene flow during spillover from Panamint Lake. Another explanation is that our understanding of the habitats during this period need improvement. The lack of fossil evidence as a result of water not flowing through Wingate Pass does not mean that the habitat was incompatible with vole movement. It is unclear how quickly the habitat changed from a flowing river to the inhospitable desert habitat it is today.

Comparison with other species

Few other studies have examined the phylogeographical structure for animals in the Mojave Desert. Wood *et al.* (2013) examined genetic variation across 12 taxonomically disparate species and found most of the divergences across the Mojave and Sonoran deserts dated prior to the Last Glacial Maximum. However, the species examined were primarily desert-adapted and thus do not serve as a proper comparison for mesic-associated species such as the California vole. Unfortunately, the only comparative studies for mesic-adapted taxa are restricted to fully aquatic species, such as pupfish (Echelle, 2008; Knott *et al.*, 2008; Spencer, Smith & Dowling, 2008) or freshwater amphipods (Witt, Threloff & Hebert, 2008). Although these organisms clearly experienced the same palaeo-hydrology, these studies are relatively uninformative in a comparison with vole phylogeography. Lineages for both groups are taxonomically much more diverse, are more broadly distributed throughout the intermontane west, have much deeper temporal histories, and their restriction to emergent water lessens the likelihood that late and post-Pleistocene habitat connections influenced vole distributions over this co-inhabited area. To date, there are no studies of truly parallel wetland and riparian taxa to compare with the patterns uncovered by the present study.

The present study sheds light on a mesic adapted species that colonized the Mojave Desert region and persists in small pockets of habitat. Its history of expansion into and contraction across this region reflects the complexity of the changing fluvial landscape of this desert ecosystem. The continued sampling of vole populations, as well as other mesic species, combined with approaches that utilize a diversity of morphological and molecular markers will lead to a more comprehensive understanding of the effects of Pleistocene climate change on the fauna of the Mojave Desert.

ACKNOWLEDGEMENTS

We thank Alisa Ellsworth and Tammy Branston of California Fish & Wildlife for help with access to samples from Fish Slough, Cartago, and Blackrock; Keith Dyas and Chris Herbst for access to Edwards Air Force Base; Shelley Ellis of the Bureau of Land Management Ridgecrest office for advice and access to land; Tom Campbell and Anna-Maria Easley at China Lake Naval Air Weapons Station; Tom Gui for access to personal and commercial property at Elizabeth Lake; Mitsubishi Cement Corporation for access to Cushenbury Springs; Robin Eliason of the USFS for access to lands in the San Bernardino National Forest; Palmdale Fin and Feather Club for access to Lake Palmdale; Link Olson at University of Alaska's Museum of the North for samples from Cushenbury Springs; and Janet Foley, her students, and USGS, California Department of Fish &Wildlife, and Bureau of Land Management biologists for access to samples from Tecopa. The US Fish & Wildlife Service provided financial support (grant #VENFWO-81440-1201) for part of this research. Jim McGuire, Rauri Bowie, Marcelo Weksler, and Pauline Kamath provided advice on the use of software for analysis of genetic data; John Wieczorek made the EXCEL spreadsheet function for estimating geographical distances from coordinate data. Michelle Koo gave advice on constructing maps. Garrett Harada conducted some of the microsatellite and DNA laboratory work. Finally, we would like to thank three anonymous reviewers for their helpful comments on the submitted manuscript.

REFERENCES

Anderson KC, Wells SG. 2003a. Pleistocene lake highstands in Death Valley, California. *GSA Special Papers* **368:** 115–128.

Anderson KC, Wells SG. 2003b. Latest Quaternary paleohydrology of Silurian Lake and Salt Spring basin, Silurian Valley, California. *GSA Special Papers* **368:** 29–141.

Baele G, Lemey P, Bedford T, Rambaut A, Suchard M, Alekseyenko A. 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution* **29:** 2157–2167.

Beerli P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* **22:** 341–345.

Brunhoff C, Ke G, Vb F, Ja C, Jaarola M. 2003. Holarctic phylogeography of the root vole (*Microtus oeconomus*): implications for late Quaternary biogeography of high latitudes. *Molecular Ecology* **12:** 957–968.

Carlsson J. 2008. Effects of microsatellite null alleles on assignment testing. *Journal of Heredity* **99:** 616–623.

Chapuis MP, Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* **24:** 621–631.

Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, Mitrovica JX, Hostetler SW, McCabe AM. 2009. The Last Glacial Maximum. *Science* **325:** 710–714.

Clement M, Posada D, Crandall K. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9:** 1657–1660.

Cohen KM, Gibbard P. 2011. *Global chronostratigraphic correlation table for the last 2.7 million years*. Cambridge: Subcommission on Quaternary Stratigraphy, International Commission on Stratigraphy. Available at: http://quaternary.stratigraphy.org/charts

Conroy CJ, Gupta A. 2011. Morphological diversity in California voles, *Microtus californicus*, across a hybrid zone. *Biological Journal of the Linnean Society* **104:** 264–283.

Conroy CJ, Neuwald JL. 2008. Phylogeographic study of the California vole, *Microtus californicus*. *Journal of Mammalogy* **89:** 755–767.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969–1973.

Earl DA, von Holdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4:** 359–361.

Echelle AA. 2008. The western North American pupfish clade (Cyprinodontidae: *Cyprinodon*): mitochondrial DNA divergence and drainage history. *GSA Special Papers* **439:** 27–38.

Enzel Y, Wells SG, Lancaster N. 2003. Pleistocene lakes along the Mojave River, southeast California. *GSA Special Papers* **368:** 61–77.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611–2620.

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

Forester RM, Lowenstein TK, Spencer RJ. 2005. An ostracode based paleolimnologic and paleohydrologic history of Death Valley: 200 to 0 ka. *GSA Bulletin* **117:** 1379–1386.

Fu Y-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and backgroud selection. *Genetics* **147:** 915–925.

Garcia AL, Knott JR, Mahan SA, Bright J. 2014. Geochronology and paleoenvironment of pluvial Harper Lake, Mojave Desert, California, USA. *Quarternary Research* **81:** 305–317.

Grayson D. 2011. *The Great Basin: a natural prehistory*. Berkeley, CA: University of California Press.

Hall ER. 1981. *The mammals of North America, 2nd edn*. New York, NY: John Wiley & Sons.

Herman JS, Searle JB. 2011. Post-glacial partitioning of mitochondrial genetic variation in the field vole. *Proceedings of the Royal Society of London Series B, Biological Sciences* **278:** 3601–3607.

Herman JS, McDevitt AD, Kawałko A, Jaarola M, Wójcik JM, Searle JB. 2014. Land-bridge calibration of molecular clocks and the post-glacial colonization of Scandinavia by the Eurasian field vole *Microtus agrestis*. *PLoS ONE* **9:** e103949.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23:** 1801–1806.

Jayko AS, Forester RM, Kaufman DS, Phillips FM, Yount JC, McGeehin J, Mahan SA. 2008. Late Pleistocene lakes and wetlands, Panamint Valley, Inyo County, California. *GSA Special Papers* **439:** 151–184.

Jefferson GT. 2008. Stratigraphy and paleontology of the middle to late Pleistocene Manix Formation and paleoenvironments of the central Mojave River, southern California. *GSA Special Papers* **368:** 43–60.

Jensen JL, Bohonak AJ, Kelley ST. 2005. Isolation by distance, web service. *BMC Genetics* **6:** 13.

Kalinowski ST. 2005. HP-Rare. A computer program for performing rarefaction on measures of allelic diversity. *Molecular Ecology Notes* **5:** 187–189.

Kellogg R. 1918. A revision of the *Microtus californicus* group of meadow mice. *University of California Publications in Zoology* **21:** 1–42.

Knott JR, Machette MN, Klinger RE, Sarna-Wojcicki AM, Liddicoat JC, Tinsley JC III, David BT, Ebbs VM. 2008. Reconstructing late Pliocene to middle Pleistocene Death Valley lakes and river systems as a test of pupfish (Cyprinodontidae) dispersal hypotheses. *GSA Special Papers* **439:** 1–26.

Koehler PA, Anderson RS. 1994. Full-glacial shoreline vegetation during the maximum highstand at Owens Lake, California. *Great Basin Naturalist* **54:** 142–149.

Martínková N, Barnett R, Cucchi T, Struchen R, Pascal M, Pascal M, Fischer MC, Higham T, Brace S, Ho SYW, Quéré J-P, O'Higgins P, Excoffier L, Heckel G, Hoelzel AR, Dobney KM, Searle JB. 2013. Divergent evolutionary processes associated with colonization of offshore islands. *Molecular Ecology* **22:** 5205–5220.

McGuire JL, Davis EB. 2013. Using the paleontological record of *Microtus* to test species distribution models and reveal responses to climate change. *Journal of Biogeography* **40:** 1490–1500.

Miller DM, Schmidt KM, Mahan SA, McGeehin JP, Owen LA, Barron JA, Lehmkuhl F, Löhrer R. 2010. Holocene landscape response to seasonality of storms in the Mojave Desert. *Quaternary International* **215:** 45–61.

Orme AR. 2008. Lake Thompson, Mojave Desert, California: the late Pleistocene lake system and its Holocene dessication. *GSA Special Papers* **439:** 261–278.

Orme AR, Orme AJ. 2008. Late Pleistocene shorelines of Owens Lake, California, and their hydroclimatic and tectonic implications. *GSA Special Papers* **439:** 207–225.

Phillips FM. 2008. Geological and hydrological history of the paleo-Owens River drainage since the late Miocene. *GSA Special Papers* **439:** 151–184.

Pigati JS, Miller DM, Bright JE, Mahan SA, Nekola JC, Paces JB. 2011. Chronology, sedimentology, and microfauna of groundwater discharge deposits in the central Mojave Desert, Valley Wells, California. *GSA Bulletin* **123:** 2224–2239.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

Rambaut A, Drummond AJ. 2007. Tracer, Version 1.5. Available at: http://beast.bio.ed.ac.uk/Tracer

Reheis MC, Redwine JL. 2008. Lake Manix shorelines and Afton Canyon terraces: implications for incision of Afton Canyon. *GSA Special Papers* **439:** 227–259.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61:** 539–542.

Rosenberg NA. 2004. Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* **4:** 137–138.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4:** 406–425.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139:** 457–462.

Smith GI, Bischoff JL, eds. 1997. An 800,000-year paleoclimatic record from Core OL-92, Owens Lake, southeast California. *GSA Special Papers* **317:** 1–165.

Spencer JE, Smith GR, Dowling TE. 2008. Middle to late Cenozoic geology, hydrography, and fish evolution in the American Southwest. *GSA Special Papers* **439:** 279–299.

Swofford DL. 2003. *PAUP*: phylogenetic analysis using parsimony (*and other methods), Version 4.0b10*. Sunderland, MA: Sinauer Associates Inc.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123:** 585–595.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. Microchecker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4:** 535–538.

Vincenty T. 1975. Direct and inverse solutions of geodesics 42 on the ellipsoid with application of nested equations. *Survey 43 Review* **22:** 88–93.

Weksler M, Lanier HC, Olson LE. 2010. Eastern Beringian biogeography: historical and spatial genetic structure of singing voles in Alaska. *Journal of Biogeography* **37:** 1414–1431.

Wells SG, Brown WJ, Enzel Y, Anderson RY, McFadden LD. 2003. Late Quaternary geology and paleohydrology of pluvial Lake Mojave, southern California. *GSA Special Papers* **368:** 79–114.

Witt JDS, Threloff DL, Hebert PDN. 2008. Genetic zoogeography of the *Hyalella azteca* species complex in the Great Basin: rapid rates of molecular diversification in desert springs. *GSA Special Papers* **439:** 103–114.

Wood DA, Vandergast AG, Barr KR, Inman RD, Esque TC, Nussear KE, Fisher RN. 2013. Comparative phylogeography reveals deep lineages and regional evolutionary hotspots in the Mojave and Sonoran Deserts. *Diversity and Distributions* **19:** 722–737.

Xie W, Lewis P, Fan Y, Kuo L, Chen M. 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology* **60:** 150–160.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1. Unique mtDNA haplotypes (N, northern clade; S, southern clade), samples in which each occurs, and GenBank accession numbers.

Table S2. Summary statistics for *cytb* data.

Table S3. Mean, SD, number of alleles, observed and expected heterozygosity, gene diversity, and F_{IS} across 12 microsatellite loci for each population sample of desert California voles. Highlighted values are significant ($P < 0.05$).

Table S4. Population specific F_{IS} indices (1023 permutations). Values shown in bold are significant at the 0.05 level. Value for P is the probability that a random $F_{IS} \geq$ observed F_{IS}.

Table S5. Breakdown of nuclear DNA introns DBX, DBY, and AP5 at several desert populations.

Table S6. Alleles present in Tecopa, presence across Southern and Northern mtDNA clade populations, or unique to Tecopa (i.e. absent from all other populations sampled). An X indicates detection. No reference to relative abundance is made.

Table S7. Table of log marginal likelihood estimates for three clock models.

Appendix S1. Divergence time estimation in BEAST under the three different clock models (strict, uncorrelated lognormal, uncorrelated exponential). A, northern clade sequences with the faster rate. B, southern clade sequences with the faster rate. C, northern clade sequences with the slower rate. D. southern clade sequences with the slower rate.

Appendix S2. Localities comprising each pooled group identified in Fig. 1 used in the analyses, including assigned subspecies, locality designation(s), latitude and longitude (WGS datum), sample size, voucher museum specimen or field numbers, and *cytb* haplotype designations.
Appendix S3. DNA extraction and amplification protocols.
Appendix S4. An unrooted topology using MrBayes, version 3.2.2 (Ronquist *et al.*, 2012), based on a GTR + gamma substitution model.
Appendix S5. Chronograms derived from BEAST analysis. Each is based on a combination of the four runs used to look for stationarity. LOGCOMBINER was used to combine those trees. We used TREEANNOTATER to build a single chronogram from those combined trees. Trees were then opened and edited in FIGTREE. Units on the *x*-axis are in millions of years.