RESEARCH ARTICLE

AMERICAN JOURNAL OF BOTANY

Landscape genetics and population structure in Valley Oak (*Quercus lobata* Née)[1]

Mary V. Ashley[2,5], Saji T. Abraham[2,4], Janet R. Backs[2], and Walter D. Koenig[3]

PREMISE OF THE STUDY: Although long-distance pollen movement is common in wind-pollinated trees, barriers to gene flow may occur in species that have discontinuous ranges or are confined to certain habitat types. We investigated the genetic structure of *Quercus lobata* Née populations throughout much of their range in California. We assessed the connectivity of populations and determined if barriers to gene flow occurred, and if so, if they corresponded to landscape features.

METHODS: We collected leaf samples from 270 trees from 12 stands of *Quercus lobata* and genotyped these trees using eight polymorphic microsatellite loci. Genetic structure and clustering was evaluated using genetic distance methods, Bayesian clustering approaches, and network analysis of spatial genetic structure.

KEY RESULTS: The southernmost population of *Quercus lobata* sampled from the Santa Monica area comprised a separate genetic cluster from the rest of the species, suggesting that Transverse Ranges such as the San Gabriel Mountains limit gene flow. Population differentiation among the other sites was small but significant. Network analysis reflected higher connectivity among populations along the Central Coast range, with few connections spanning the dry, low Central Valley.

CONCLUSIONS: While long distance pollen movement has been shown to be common in oaks, on larger spatial scales, topographic features such as mountain ranges and the large, flat Central Valley of California limit gene flow. Such landscape features explain gene flow patterns much better than geographic distance alone.

KEY WORDS gene flow; landscape genetics; network analysis; oaks; pollen movement; population structure

Gene flow in plants occurs through the dispersal of pollen or seeds. Pollen-mediated gene flow is often extensive in plants that are pollinated by wind. Although long-distance (landscape-level) pollen movement is difficult to measure directly, paternity studies have shown that average pollination distances within stands of wind-pollinated trees is ten to hundreds of meters and pollen immigrations rates into study sites are about 50% (reviewed in Ashley, 2010). These levels of effective pollen movement suggest that wind-pollinated tree species may exhibit little population structure over large areas, especially if woodlands are relatively continuous and major landscape impediments to pollen movement do not occur. In tree species that have discontinuous ranges or those confined to certain habitat types, however, gene flow may be restricted through parts of a species range, and be reflected by significant population structure between some populations.

We studied population structure in valley oak, *Quercus lobata,* a wind-pollinated oak endemic to California. A winter-deciduous species, valley oak is the largest western North American oak, with trees 10-25 m tall and 0.5-0.7 m diameter at breast height (dbh; Munz, 1973). Valley oak is an ecologically important species in savannas and oak woodlands of California but has experienced extensive habitat loss as land within its range has been converted to farmland, vineyards, and other development (Allen-Diaz et al., 2007; Whipple et al., 2011). Ecological niche models for valley oak indicate that their current distribution has been relatively stable since the late Pleistocene (~120 kya; Gugger et al., 2013). This distribution forms a horseshoe around the Central Valley, found in the lower elevations of the Central and North Coast ranges and the Sierra Nevada (Fig. 1). The southernmost populations are close but

[1] Manuscript received 22 April 2015; revision accepted 5 November 2015.

[2] Department of Biological Sciences, University of Illinois at Chicago, 845 W. Taylor Street, Chicago, Illinois 60607 USA; and

[3] Lab of Ornithology and Department of Neurobiology and Behavior, Cornell University, Ithaca, New York 14850 USA

[4] Current address: Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan 48109 USA

[5] Author for correspondence (e-mail: ashley@uic.edu)

doi:10.3732/ajb.1500182

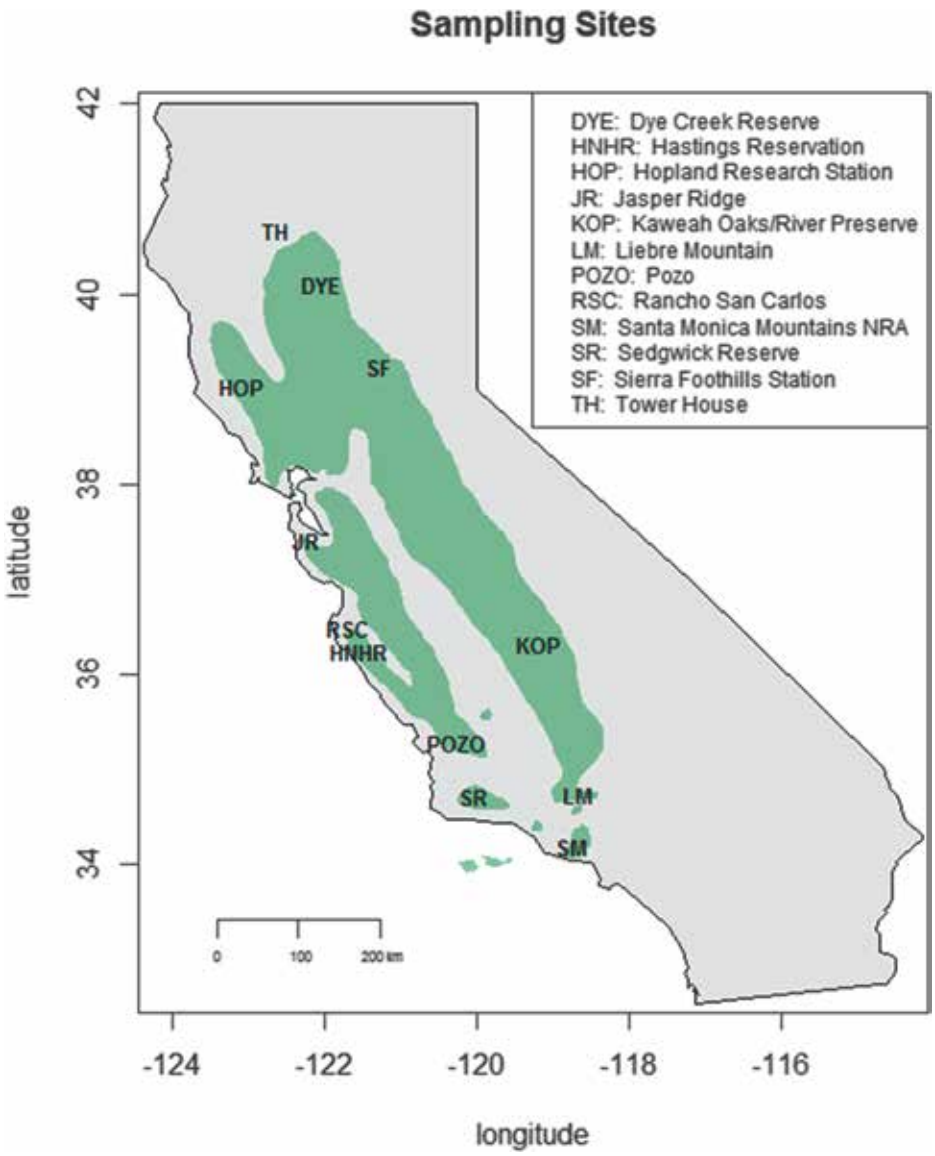

FIGURE 1 Map of California showing range of *Quercus lobata* (green shading) and locations of collection sites for this study.

disjunct from the main range of the species, separated by the Transverse Ranges of Southern California, which have an east-west orientation (in contrast to the north-south orientation of other major ranges in the state). The hot, dry Central Valley of California dissects the main part of the species' range further north. The Central Valley is known to restrict gene flow for several species with limited dispersal, including the classic ring species of salamanders, *Ensatina eschscholtzii* Gray (Wake et al., 1986; Moritz et al., 1992), but a recent study of *Quercus chrysolepis* Liebm. (canyon live oak) revealed little genetic differentiation between populations from the Coast ranges and the Sierra Nevada, on either side of the Central Valley (Ortego et al., 2015). Besides *Q. lobata* and *Q. chrysolepis*, several other California oaks, including *Q. kelloggii* Newb. (California black oak), *Q. douglasii* Hook. & Arn. (blue oak), *Q. wislizenii* A.DC. (interior live oak), and *Q. berberidifolia* Liebm. (scrub oak) share a similar distribution that circles the Central Valley, with disjunct populations at the southern end of their ranges. Given the importance of oaks as a key component of California woodlands, understanding how topographic and landscape features structure populations and gene flow in *Q. lobata* is of broader and more general importance for California oak woodlands.

Previous work on valley oak has examined local patterns of pollination (Sork et al., 2002; Pluess et al., 2009; Abraham et al., 2011) fine-scale genetic structure (Dutech et al., 2005), and hybridization (Craft et al., 2002; Abraham et al., 2011). There have also been studies of range-wide genetic patterns that have each come to somewhat different conclusions. Grivet et al. (2008) report strong genetic structure and restricted gene flow for chloroplast microsatellite markers, but not for nuclear microsatellites. Chloroplast markers will only reflect gene flow via acorn dispersal, which is likely to be much more limited than pollen dispersal in oaks. However, another analysis of chloroplast haplotypes indicated extensive historical connectivity and gene movement among most populations, including occasional long distance gene exchange across the Central Valley (Sork et al., 2010). Finally, a recent paper that reanalyzed the previously published data and added two nuclear gene sequences came to a third conclusion, identifying a strong split between coastal and Sierra Nevada populations for both nuclear and chloroplast markers (Gugger et al., 2013).

The objective of the current study was to characterize patterns of gene flow in valley oak and to evaluate these patterns in the context of distance, distribution, and landscape configuration. Our study differs from previous studies (Grivet et al., 2008; Sork et al., 2010; Gugger et al., 2013) because we sampled many more individuals from each population (but sampled fewer sites) and focused strictly on nuclear microsatellite markers. The highly variable nuclear microsatellites employed in this study allow for high resolution in inferring historical patterns of gene flow and detecting physical barriers to gene flow, should they exist.

METHODS

Sample collection and microsatellite genotyping—Leaf material was collected from 270 trees located in 12 populations that span the range of the species (Fig. 1), with 14-32 individuals sampled per population (Table 1). Genomic DNA was extracted from ~20-30 mg of ground leaf tissue using a Qiagen Plant Mini Kit (Qiagen, Valencia, California, USA). DNA quantities were measured using a Nanodrop (Nanodrop Technologies, Wilmington, Delaware, USA).

Samples were genotyped at eight microsatellite loci: (1) QpZag 9 and QpZAG1/5 were developed for *Quercus petraea* (Matt.) Liebl.

TABLE 1. Sampling sites, latitude and longitude, number of trees sampled (N), mean number of alleles per locus (N_A), mean effective number of alleles (N_E), observed heterozygosity (H_O), expected heterozygosity (H_E), and fixation index (F_{IS}) for *Quercus lobata* used in this study.

Site (Abbreviation) [latitude, longitude]	N	N_A	N_E	H_O	H_E	F_{IS}
Dye Creek Reserve Tehama Co. (DYE) [40.103733, −122.04572]	20	7.50	4.37	0.766	0.696	−0.117
Hastings Reservation Monterey Co. (HNHR) [36.374996, −121.56021]	21	9.00	4.64	0.912	0.730	−0.117
Hopland Research Station Mendocino Co. (HOP) [39.027008, −123.08562]	21	8.50	4.29	0.900	0.696	−0.341
Jasper Ridge San Mateo Co. (JR) [37.404194, −122.23093]	26	10.50	4.92	0.792	0.748	−0.050
Kaweah Oaks/River Preserve Tulare Co. (KOP) [36.325928, −119.17060]	23	6.63	3.76	0.738	0.670	−0.128
Liebre Mountain Los Angeles Co. (LM) [34.739166, −118.66033]	14	8.75	4.62	0.819	0.766	−0.071
Pozo San Louis Obispo Co. (POZO) [35.283916, −120.26778]	23	9.25	5.30	0.861	0.786	−0.104
Rancho San Carlos Monterey Co. (RSC) [36.473100, −121.80199]	24	7.13	3.84	0.878	0.609	−0.909
Santa Monica Mountains National Recreational Area Ventura Co. (SM) [34.198925, −118.74638]	32	7.50	3.33	0.661	0.660	−0.006
Sedgwick Reserve Santa Barbara Co. (SR) [34.715590, −120.03887]	21	8.25	3.44	0.670	0.680	0.001
Sierra Foothills Station Yuba Co. (SF) [39.238552, −121.28862]	21	7.13	3.78	0.788	0.721	−0.095
Tower House Shasta Co. (TH) [40.663567, −122.63725]	24	7.25	3.57	0.842	0.690	−0.233
Total	270	8.12	4.16	0.802	0.703	−0.196

(Steinkellner et al., 1997); (2) MSQ4 and MSQ13 were developed for *Q. macrocarpa* Michx. (Dow et al., 1995; Dow and Ashley, 1996); (3) QpZAG15 and QpZAG11 were developed for *Q. robur* L. (Kampfer et al., 1998); and (4) QM69-2M1 and QM 57-3M were developed for *Q. myrsinifolia* Blume (Isagi and Suhandono, 1997). Polymerase Chain Reaction (PCR) was carried out using 50–100 ng genomic DNA in 10 µl PCR mix with the following reagents: (1) 0.5mM of 10mM dNTP mix (Denville Scientific, South Plainfield, New Jersey, USA); (2) 0.04 µM of the forward primer with the fluorescent-labeled M13 (−21) universal primer; (3) 0.6 – 0.8µM reverse primer; (4) 1.0 µg/µl bovine serum albumin; and (5) 0.25 U Taq Polymerase (Biotherm Taq, eEnzyme, Gaithersburg, Maryland, USA) with Biotherm Buffer (5×). PCR conditions are described in Craft and Ashley (2007). PCR products (1.5 µL) were analyzed in a capillary DNA sequencing machine (Applied Biosystem 3730) using a LIZ500 ladder (Applied Biosystems, Foster City, California, USA). All microsatellite genotypes were scored by analyzing the raw data using Applied Biosystems GeneMapper software (version 3.7).

Summary Statistics—Genetic diversity was estimated by the mean number of alleles per locus (N_A), mean effective number of alleles (N_E), observed (H_O), and expected heterozygosity (H_E) using GenAlEx version 6.5.1 (Table 1; Peakall and Smouse, 2006, 2012). Fixation indices (F_{IS}) were also calculated using GenAlEx version 6.5.1 (Table 1). Departures from Hardy-Weinberg expected genotype frequencies were tested using the Markov chain Monte Carlo (MCMC) method (dememorization 1000, batches 100, iterations per batch 1000) using GENEpop on the Web (Raymond and Rousset, 1995; Rousset, 2008).

Population Differentiation—We calculated genetic differentiation statistics including the traditional *F*-statistics analogues $\boldsymbol{\theta}$ (Weir and Cockerham, 1984) and G_{ST} (Nei and Chesser, 1983), as well as two newer differentiation statistics, G'_{ST} (Hedrick, 2005) and D_{JOST} (Jost, 2008), using the R (R Development Core Team, 2014) package diveRsity (Keenan et al., 2013). F_{ST} and its analogues have been criticized for use in highly variable markers such as microsatellites because their maximum possible value is dependent on within-population diversity (Hedrick, 1999; Jost, 2008); the other two statistics we calculated have been developed to overcome this drawback. We calculated 95% confidence intervals for these statistics using 1000 bootstraps as implemented in diveRsity.

To determine whether our microsatellite loci are useful in inferring demographic processes such as migration rates, we plotted estimates of $\boldsymbol{\theta}$, G_{ST}, G'_{ST} and D_{JOST} against the number of alleles using the corPlot function in diveRsity. If demographic processes have similar effects across neutral loci with differing mutation rates (resulting in differing allele numbers), F_{ST} estimates should be more or less equal across loci. These relationships also provide an assessment of if and how (positively or negatively) the different estimators are correlated with allele number and heterozygosity (Keenan et al., 2013). Isolation-by-distance (IBD) was tested in GenAlEx by using Mantel tests based on distance matrices of pairwise linearized F_{ST} and Euclidian distance (km) between sample locations with 9999 permutations.

Genetic Clustering, Barrier, and Network Analyses—We used three different approaches to identify genetic clusters and barriers to gene flow, Bayesian clustering algorithms, edge detection, and network analysis methods. Bayesian clustering methods delineate clusters of individuals based on the analysis of individual multilocus genotypes and assign individuals to the cluster where their posterior probability is highest. We used two Bayesian clustering approaches, STRUCTURE version 2.3.4 software (Pritchard et al., 2000), a nonspatial model that ignores geographic proximity, and GENELAND (Guillot et al., 2005) which uses geo-referenced individuals or populations. We chose GENELAND over other spatial Bayesian methods because it has been shown to have the highest power when analyzing simulated data (Safner et al., 2011; Blair et al., 2012). Edge detection approaches identify areas where changes in variables occur, in this case allele frequencies (Manni et al., 2004; Safner et al., 2011). We used Monmonier's algorithm to detect boundaries (edges) between populations. Finally, network analysis aims to connect populations in a network graph with connection (edges) that are weighted according to pairwise genetic distances. These different analyses are quite distinct but complimentary, and can provide confidence to conclusions that are supported by multiple approaches.

For nonspatial Bayesian clustering algorithm, implemented in STRUCTURE version 2.3.4 software (Pritchard et al., 2000), the number of assumed genetic clusters (K) was set from 1 to 12, and 10 runs with 100 000 MCMC iterations were performed for each K following a burn-in of 50 000 iterations. The admixture model was run with correlated allele frequencies, and included a priori sampling locations as prior information (LOCPRIOR) to detect weak population structure (Hubisz et al., 2009). The LOCPRIOR option is not biased toward detecting structure when it is not present (Hubisz et al., 2009). The optimal value of K was determined depending on ΔK value (Evanno et al., 2005) using the STRUCTURE HARVESTER version 0.6.93 application (Earl and vonHoldt 2012). Consensus analyses were performed using CLUMPP 1.1.2 version (Jakobsson and Rosenberg, 2007) on the average scores for the inferred K value. STRUCTURE PLOT (Ramasamy et al., 2014) was used to visualize the STRUCTURE output.

For GENELAND, the GUI interface (GENELAND R-package version 4.0.5) was used to run the correlated frequency, spatial model at 100K MCMC, thin rate 100, and burn-in 200 for 10 runs at K values from one to fifteen. Because a single set of spatial coordinates was used for each collection site, we set *uncertainty on coordinate* to 0.0005 (about 75 m) to prevent individuals sampled from each site being automatically assigned to the same inferred group by GENELAND. This procedure was repeated five times. For each repetition, the run with the highest posterior probability was chosen, and from these five results the best K was inferred from the modal value of K with the highest posterior probability.

We used the Monmonier's function in the R-package *adegenet* version 1.4-1 (Jombart, 2008) to describe paths of strongest genetic distances between populations. The initial connection network was built using coordinates for the sites. A Euclidean distance matrix was defined based on analogues $\boldsymbol{\theta}$ (Weir and Cockerham, 1984) for the 12 populations in the study. Random noise was eliminated using Principal Coordinate Analysis (dudi.pco). The algorithm then builds genetic boundaries based on maximum pairwise distances.

We applied graph-based network theory analysis to create a genetic network of populations or minimum spanning network (MSN), implemented in EDENetworks version 2.18 (Kivelä et al., 2015). The program plots all agents (individuals or populations) as nodes in a network graph with connections (edges or links) between nodes weighted by their pairwise genetic distance F_{ST}. EDENetworks illustrates the distribution of links among populations using percolation theory and without an a priori hypothesis based on population identity or sampling location. The percolation theory allows for the splitting of a fully connected network into discrete clusters, and the critical threshold distance is known as the percolation threshold. The layout of the MSN was recalculated 10 times to test for possible alternative network shapes.

RESULTS

The eight loci used were highly variable, with an average of 8.12 alleles per locus and a mean expected heterozygosity (H_E) of 0.703 per population (Table 1). Average observed heterozygosity (H_O) was larger than expected heterozygosities at all sites except Sedgwick Reserve. Indeed, several loci/site combinations showed significant heterozygote excess, but global tests of heterozygote deficiencies and excess were not significant. Excess heterozygosity has been reported in other studies of oaks (Dutech et al., 2005; Craft and Ashley, 2007, 2010) and may be due to the high levels of outcrossing in oak species.

Global estimates of genetic differentiation were: $\boldsymbol{\theta}$ = 0.0639 (Weir and Cockerham, 1984); G_{ST} (Nei and Chesser, 1983) = 0.0549; G'_{ST} (Hedrick, 2005) = 0.2086; and D_{JOST} (Jost, 2008) = 0.1279. The lower 95% confidence interval estimated by bootstrapping did not go below zero for any of these statistics, indicating significant population differentiation. As is expected, the relationship between $\boldsymbol{\theta}$ and the number of alleles and G_{ST} and the number of alleles were both slightly negative whereas the plots for G'_{ST} and D_{JOST} and the number of alleles showed a small positive relationship (plots not shown); however, none of the correlations was significant. This suggests that the mutation rates exhibited by our markers were not high enough to obscure past demographic processes and both classes of genetic differentiation statistics should be useful in inferring historical gene flow patterns.

Pairwise estimates of $\boldsymbol{\theta}$ and D_{JOST} are shown in Table 2. Except for the comparison between Liebre Mountain (LM) and Sedgwick Reserve (SR), the lower 95% confidence interval estimated by bootstrapping did not go below zero for any of pairwise statistics, again indicating significant population differentiation. The results of the Mantel test, using the total sample set, was Rxy = 0.158 (P = 0.0001), indicating a significant positive, but relatively weak, correlation between the genetic and geographic matrices and limited isolation by

TABLE 2. Pairwise estimates of genetic differentiation: below diagonal θ (Weir and Cockerham, 1984); above diagonal D_{JOST} (Jost, 2008). Sampling site abbreviations are given in Table 1.

	DYE	HNHR	HOP	JR	KOP	LM	POZO	RSC	SM	SR	SF	TH
DYE	—–	0.0744	0.0843	0.0605	0.0479	0.061	0.100	0.152	0.211	0.047	0.052	0.084
HNHR	0.0649	—–	0.0272	0.0242	0.0817	0.007	0.045	0.028	0.174	0.054	0.136	0.019
HOP	0.0547	0.019	—–	0.035	0.1231	0.040	0.109	0.083	0.250	0.056	0.140	0.052
JR	0.0279	0.021	0.016	—–	0.0456	0.002	0.031	0.143	0.115	0.042	0.083	0.038
KOP	0.0355	0.0821	0.0792	0.0351	—–	0.046	0.113	0.191	0.134	0.049	0.077	0.128
LM	0.036	0.028	0.038	0.007	0.038	—–	0.046	0.081	0.084	−0.007	0.022	0.033
POZO	0.047	0.023	0.039	0.016	0.062	0.019	—–	0.157	0.190	0.137	0.077	0.051
RSC	0.110	0.023	0.045	0.069	0.136	0.077	0.074	—–	0.177	0.121	0.183	0.090
SM	0.119	0.116	0.145	0.089	0.113	0.060	0.088	0.149	—–	0.092	0.194	0.193
SR	0.038	0.059	0.049	0.023	0.039	−0.005	0.056	0.103	0.102	—–	0.082	0.061
SF	0.036	0.065	0.061	0.036	0.054	0.020	0.040	0.112	0.111	0.053	—–	0.055
TH	0.071	0.015	0.031	0.027	0.093	0.034	0.041	0.068	0.131	0.059	0.059	—–

distance. For the sample set without the Santa Monica Mountains (SM) site, Rxy = 0.070 (P = 0.0004).

STRUCTURE analysis revealed the peak distribution of ΔK (Evanno et al., 2005) occurred at K = 2 for 10 simulations at K values from 1 to 12 (Fig. 2). One of the two clusters represented the oaks collected from the Santa Monica Mountains and the other was comprised of all other populations. The consensus membership coefficients, implemented in CLUMPP version 1.1.2 (Jakobsson and Rosenberg, 2007) are shown in Fig. 3. Most trees had high membership coefficients in their respective cluster, with a mean of 0.929 for trees in the larger cluster and 0.948 for trees from the Santa Monica Mountains cluster (Fig. 3). GENELAND identified 11 genetic clusters, each corresponding to a sampling site with the exception of Sedgwick Reserve (SR) and Liebre Mountain (LM), which together comprised a single genetic cluster. The Monmonier's function implemented in *adegenet* identified a single barrier separating the Santa Monica population from all populations north of the San Gabriel Mountains (Fig. 4).

The MSN produced by EDENetworks based on populations is shown in Fig. 4. Several distinct topological features are revealed.

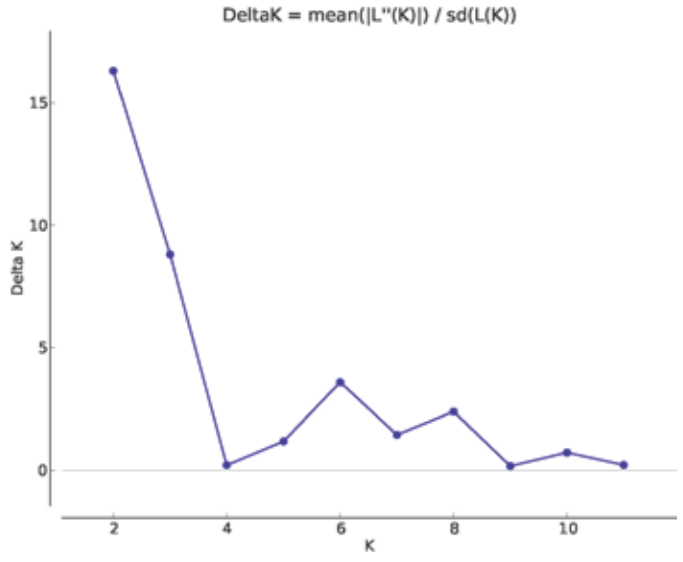

FIGURE 2 Posterior probability of the data (Ln[$P(D|K)$]) and values of ΔK (Evanno et al., 2005) as a function of K (number of clusters), associated with the results shown in Fig. 3.

For example, despite the close physical distance between SM and other populations (LM and SR), there is only a very weak link connecting them. This confirms our other results indicating that the *Quercus lobata* population in the Santa Monica Mountains is genetically distinct from the rest of the range and experiences very limited gene flow from the north. The strong link between LM and SR also supports the GENELAND clustering results which places these two populations into a single genetic cluster. Only one link spanned the Central Valley (between Jasper Ridge (JR) and Kaweah Oak/River Preserve (KOP), indicating that this large landscape feature does act to limit gene flow. Populations along the Coast Range sites were well connected, with the JR population being an especially important node in facilitating or allowing connectivity of populations throughout much of the species range.

DISCUSSION

Oak woodlands and savannas are one of the most important and valued ecosystems of California, occupying about four million hectares west of the higher elevations of the Sierra Nevada mountains (Allen-Diaz et al., 2007). While widespread, oak woodlands are confined to a relatively narrow elevational corridor, bounded by grasslands at lower elevations and conifer forests at higher elevations (Barbour and Minnich, 2000). Their distribution is influenced by the Coast Mountain Range in the western part of the state, the Sierra Nevada mountains in the east, the Transverse Ranges in the south, and the Central Valley in the middle of the state. Thus the oak species that define this ecosystem are excellent candidates for studying how landscape and topographical features influence gene flow and population structure. The twelve sites where we sampled *Quercus lobata* span the native range of the species, from Ventura and Los Angeles Counties in the south to Shasta County in the north, and with sites on both sides of the Central Valley (Fig. 1, Table 1).

Quercus lobata, the largest of the oak species within its range, was widespread before European colonization but is now more restricted as native valley oak woodland habitat has been lost. Valley oak woodlands now exist as patches or fragments embedded in a matrix of agricultural and urban/suburban land, with conversion to irrigated agricultural land responsible for the largest acreage decline. Such forest fragmentation might be expected to bring about reproductive isolation of remnants and genetic declines associated with genetic drift or inbreeding (Grivet et al., 2008; Kramer et al., 2008).

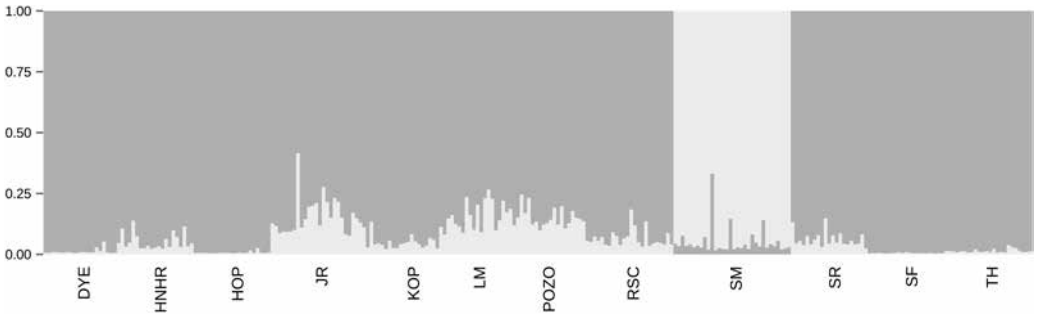

FIGURE 3 The proportion of the membership coefficient for each individual in the 12 *Quercus lobata* populations for the inferred clusters when $K = 2$ according to STRUCTURE analysis. Population abbreviations are as given in Table1 and Fig. 1.

Several studies have tested whether isolated or relict populations of oaks are reproductively isolated, genetically differentiated, or genetically depauperate. These studies have shown that even very remote stands of oaks are genetically diverse and reproductively connected through pollen flow. For example, peripheral, remnant populations of *Quercus petraea* (Matt.) Liebl. in Ireland had high levels of genetic diversity and were not differentiated from populations in Spain and southwestern France (Muir et al., 2004). Isolated remnant stands of *Q. macrocarpa* received more than half their pollen from outside the stand (Craft and Ashley, 2010), and small remnants showed little or no population structure despite a fragmented distribution lasting thousands of years (Craft and Ashley, 2007). Two tiny remnant stands of *Q. robur* isolated more than 80 km from other oaks show high genetic diversity and received substantial pollen immigration (Buschbom et al., 2011). Extensive pollen immigration was reported in a stand of *Q. ilex* L. that was highly fragmented with very low tree density, and seed set and progeny genetic diversity were not reduced (Ortego et al., 2014). These studies suggest that long distance pollination is the norm in oaks and distance, low densities, or landscape features do not notably disrupt pollen dispersal and reproduction.

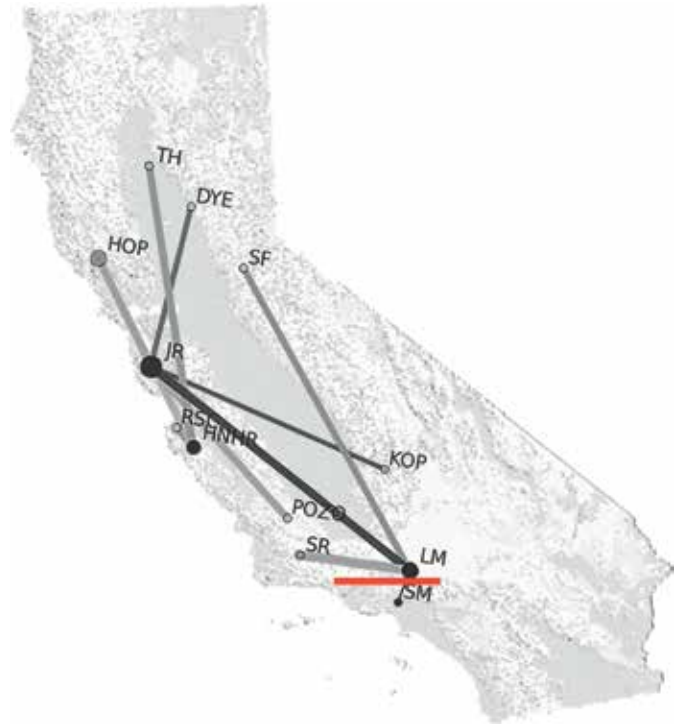

FIGURE 4 Topographic map of California with sampling sites overlaid. Blue and green lines show network linkages identified by EDENetworks (Kivelä et al., 2015) between nodes (sampling sites). Line thickness is proportional to linkage strength and node size is proportional to the number of linkages for each node. The red line indicates the position of the single barrier to gene flow identified by the Monmonier's function implemented in *adegenet* (Jombart, 2008).

Two previous studies examined local pollination patterns for *Quercus lobata* at two of the sites included here. At the Hastings Reservation (HNHR) site, paternity assignment of acorns was used to measure pollination distances (Abraham et al., 2011). For most acorns sampled (76/108, 70%) the pollen donor was beyond 200 m of the maternal tree, indicating that the majority of effective pollen travels more than this distance. Work at the Sedgwick (SR) site using paternity assignment suggested more limited pollen immigration, 18.5%, from beyond 250 m of maternal trees (Pluess et al., 2009). These two values span the range of estimates for all oaks (Ashley, 2010), and may reflect either methodological differences or actual intraspecific variation in pollination patterns due to habitat, density or other factors that differ among sites. But even the lower values of pollen immigration would be expected to homogenize allele frequencies across populations and prevent differentiation and genetic structure to develop, as in the other studies of oaks described above in the previous paragraph.

Thus, in light of previous studies on gene flow in oaks, the results we report here indicating genetic discontinuities in *Quercus lobata*, are surprising. We identified features of the California landscape that act as barriers to gene flow and create population structure for *Q. lobata*. The most striking discontinuity detected was between the trees sampled at our southernmost Santa Monica Mountains site in Ventura County and all sites to the north, a result supported by all of our analyses. Indeed, STRUCTURE found only two genetic clusters in the data, one comprised of the eleven populations north of the San Gabriel Mountains and the other comprised of the single population south of this range (Fig. 3). The EDENetwork analysis found a single small linkage between SM and the closest sampling site, LM, which is only about 60 km away but separated by the San Gabriel Mountains. These findings indicate that there has been little gene flow across the transverse ranges for *Q. lobata* and raise the possibility that gene flow barriers might exist for other oaks with distributions that span these mountains. Another possibility is that divergence in flowering phenology, rather than a topographical barrier, isolates populations at the southern end of their distribution, as has been suggested for peripheral populations of *Q. engelmannii* Greene in California (Ortego et al., 2012). From a conservation standpoint, the few remaining stands of *Q. lobata* in the southern part of their range are genetically distinct and should be a focus of conservation efforts, as previously suggested by Grivet et al. (2008).

In contrast to the isolation of the SM site, two sites to the north and west of the San Gabriel Mountains, LM and SR, are about 125 km apart and separated by Los Padres National Forest, but surprisingly, are genetically very similar. These were the only population pairs for which pairwise estimates of $\boldsymbol{\theta}$ and D_{JOST} were not significantly different. Furthermore, the GENELAND analysis placed these two populations in the same genetic cluster while assigning all other sites to different genetic clusters.

Throughout the rest of the species range, populations exhibited modest but significant levels of genetic differentiation, indicating that gene flow occurs but is not sufficient to keep populations of *Quercus lobata* panmictic throughout its range. Tests for IBD showed weak but significant correlations. The network analysis shown in Fig. 4 demonstrates why geographical distances alone poorly reflect population connectivity. Sites spanning the Central Coast Range are strongly linked, with the centrally located JR site being an important node linking several sites. Only one linkage connecting JR and KOP spanned the largely treeless Central Valley. Sites east of the Central Valley in the foothills of the Sierra Nevada showed less population connectivity and seemingly were not linked to the sites north of the Central Valley.

Our findings differ considerably from those of Gugger et al. (2013), who also examined genetic differentiation and population structure in *Quercus lobata*. These authors also identified two major genetic clusters within valley oaks, but one was comprised of a western cluster from the Coast Range and an eastern one found on the western Sierra Nevada foothills. Their analysis did not identify the strong discontinuity we found between the southernmost population and those located north of the Transverse Ranges. Instead they found the Transverse Range populations to be genetically connected to the Sierran populations. These differences may be because of the different sampling or analytical methods used.

In conclusion, our findings indicate that the striking topographical and climatic gradients that shape California oak woodland distributions also impact patterns of gene flow. For populations of *Quercus lobata*, distance alone is a poor predictor of population differentiation, and in particular, the San Gabriel Mountains appears to be a strong barrier to gene flow, isolating the southernmost population from the rest of the species' range.

ACKNOWLEDGEMENTS

The authors thank Marti Whittier for collecting samples at one site used in this study. Funding was provided by a grant from the Integrated Hardwood Range Management Program and NSF grant DEB-0816691 (to WDK). The authors also thank the anonymous reviewers and the Associate Editor for comments that improved the manuscript.

LITERATURE CITED

Abraham, S. T., D. N. Zaya, W. D. Koenig, and M. V. Ashley. 2011. Interspecific and intraspecific pollination patterns of valley oak, *Quercus lobata*, in a mixed stand in coastal central California. *International Journal of Plant Sciences* 172: 691–699.

Allen-Diaz, B., R. Standiford, and R. D. Jackson. 2007. Oak woodlands and forests. *In* M. Barbour, T. Keeler-Wolf, and A. A. Schoenherr [eds.], Terrestrial vegetation of California, 3rd ed., 313-338. University of California Press, Oakland, California, USA.

Ashley, M. V. 2010. Plant parentage, pollination, and dispersal: How DNA microsatellites have altered the landscape. *Critical Reviews in Plant Sciences* 29: 148–161.

Barbour, M. G., and R. A. Minnich. 2000. Californian upland forests and woodlands. *In* M. G. Barbour and W. D. Billings [eds.], North American terrestrial vegetation. 2nd ed., 161-202. Cambridge University Press, New York, New York, USA.

Blair, C., D. E. Weigel, M. Balazik, A. T. Keeley, F. M. Walker, E. Landguth, S. Cushman, et al. 2012. A simulation-based evaluation of methods for inferring linear barriers to gene flow. *Molecular Ecology Resources* 12: 822–833.

Buschbom, J., Y. Yanbaev, and B. Degen. 2011. Efficient long-distance gene flow into an isolated relict oak stand. *Journal of Heredity* 102: 464–472.

Craft, K. J., and M. V. Ashley. 2007. Landscape genetic structure of bur oak (*Quercus macrocarpa*) savannas in Illinois. *Forest Ecology and Management* 239: 13–20.

Craft, K. J., and M. V. Ashley. 2010. Pollen-mediated gene flow in isolated and continuous stands of bur oak, *Quercus macrocarpa* (Fagaceae). *American Journal of Botany* 97: 1999–2006.

Craft, K. J., M. V. Ashley, and W. D. Koenig. 2002. Limited hybridization between *Quercus lobata* and *Quercus douglasii* (Fagaceae) in a mixed stand in central coastal California. *American Journal of Botany* 89: 1792–1798.

Dow, B. D., and M. V. Ashley. 1996. Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa. Molecular Ecology* 5: 615–627.

Dow, B. D., M. V. Ashley, and H. F. Howe. 1995. Characterization of highly variable $(GA/CT)_{(N)}$ microsatellites in the bur oak, *Quercus macrocarpa. Theoretical and Applied Genetics* 91: 137–141.

Dutech, C., V. L. Sork, A. J. Irwin, P. E. Smouse, and F. W. Davis. 2005. Gene flow and fine-scale genetic structure in a wind-pollinated tree species *Quercus lobata* (Fagaceaee). *American Journal of Botany* 92: 252–261.

Earl, D. A., and B. M. vonHoldt. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.

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

Grivet, D., V. L. Sork, R. D. Westfall, and F. W. Davis. 2008. Conserving the evolutionary potential of California valley oak (*Quercus lobata* Née): A multivariate genetic approach to conservation planning. *Molecular Ecology* 17: 139–156.

Gugger, P. F., M. Ikegami, and V. L. Sork. 2013. Influence of late Quaternary climate change on present patterns of genetic variation in valley oak, *Quercus lobata* Née. *Molecular Ecology* 22: 3598–3612.

Guillot, G., F. Mortier, and A. Estoup. 2005. GENELAND: A computer package for landscape genetics. *Molecular Ecology Resources* 5: 712–715.

Hedrick, P. W. 1999. Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution* 53: 313–318.

Hedrick, P. W. 2005. A standardized genetic differentiation measure. *Evolution* 59: 1633–1638.

Hubisz, M. J., D. Falush, M. Stephens, and J. K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 9: 1322–1332.

Isagi, Y., and S. Suhandono. 1997. PCR primers amplifying microsatellite loci of *Quercus myrsinifolia* Blume and their conservation between oak species. *Molecular Ecology* 6: 897–899.

Jakobsson, M., and N. A. Rosenberg. 2007. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics (Oxford, England)* 23: 1801–1806.

Jombart, T. 2008. adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics (Oxford, England)* 24: 1403–1405.

Jost, L. 2008. G_{ST} and its relatives do not measure differentiation. *Molecular Ecology* 17: 4015–4026.

Kampfer, S., C. Lexer, J. Glössl, and H. Steinkellner. 1998. Characterization of $(GA)_n$ microsatellite loci from *Quercus robur. Hereditas* 129: 183–186.

Keenan, K., P. McGinnity, T. F. Cross, W. W. Crozier, and P. A. Prodöhl. 2013. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* 4: 782–788.

Kivelä, M., S. Arnaud-Haond, and J. Saramäki. 2015. EDENetworks: A user-friendly software to build and analyse networks in biogeography, ecology and population genetics. *Molecular Ecology Resources* 15: 117–122.

Kramer, A. T., J. L. Ison, M. V. Ashley, and H. F. Howe. 2008. The paradox of forest fragmentation genetics. *Conservation Biology* 22: 878–885.

Manni, F., E. Guerard, and E. Heyer. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. *Human Biology* 76: 173–190.

Moritz, C., C. J. Schneider, and D. B. Wake. 1992. Evolutionary relationships within the *Ensatina eschscholtzii* complex confirm the ring species interpretation. *Systematic Biology* 41: 273–291.

Muir, G., A. J. Lowe, C. C. Fleming, and C. Vogl. 2004. High nuclear genetic diversity, high levels of outcrossing and low differentiation among remnant populations of *Quercus petraea* at the margin of its range in Ireland. *Annals of Botany* 93: 691–697.

Munz, P. A. 1973. A California flora and supplement. University of California Press, Berkeley, California, USA.

Nei, M., and R. K. Chesser. 1983. Estimation of fixation indices and gene diversities. *Annals of Human Genetics* 47: 253–259.

Ortego, J., R. Bonal, A. Muñoz, and J. M. Aparicio. 2014. Extensive pollen immigration and no evidence of disrupted mating patterns or reproduction in a highly fragmented holm oak stand. *Journal of Plant Ecology* 7: 384–395.

Ortego, J., P. F. Gugger, and V. L. Sork. 2015. Climatically stable landscapes predict patterns of genetic structure and admixture in the Californian canyon live oak. *Journal of Biogeography* 42: 328–338.

Ortego, J., E. C. Riordan, P. F. Gugger, and V. L. Sork. 2012. Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. *Molecular Ecology* 21: 3210–3223.

Peakall, R., and P. E. Smouse. 2006. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

Peakall, R., and P. E. Smouse. 2012. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics (Oxford, England)* 28: 2537–2539.

Pluess, A. R., V. L. Sork, B. Dolan, F. W. Davis, D. Grivet, K. Merg, J. Papp, and P. E. Smouse. 2009. Short distance pollen movement in a wind-pollinated tree, *Quercus lobata* (Fagaceae). *Forest Ecology and Management* 258: 735–744.

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

R Development Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Website: http://www.R-project.org.

Ramasamy, R. K., S. Ramasamy, B. B. Bindroo, and V. G. Naik. 2014. STRUCTURE PLOT: A program for drawing elegant STRUCTURE bar plots in user friendly interface. *SpringerPlus* 3: 431.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248–249.

Rousset, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.

Safner, T., M. P. Miller, B. H. McRae, M.-J. Fortin, and S. Manel. 2011. Comparison of Bayesian clustering and edge detection methods for inferring boundaries in landscape genetics. *International Journal of Molecular Sciences* 12: 865–889.

Sork, V. L., F. W. Davis, P. E. Smouse, V. J. Apsit, R. J. Dyer, J. F. Fernandez-M, and B. Kuhn. 2002. Pollen movement in declining populations of California Valley oak, *Quercus lobata*: Where have all the fathers gone? *Molecular Ecology* 11: 1657–1668.

Sork, V. L., F. W. Davis, R. Westfall, A. Flint, M. Ikegami, H. Wang, and D. Grivet. 2010. Gene movement and genetic association with regional climate gradients in California valley oak (*Quercus lobata* Née) in the face of climate change. *Molecular Ecology* 19: 3806–3823.

Steinkellner, H., S. Fluch, E. Turetschek, C. Lexer, R. Streiff, A. Kremer, K. Burg, and J. Glössl. 1997. Identification and characterization of $(GA/CT)_n$ - microsatellite loci from *Quercus petraea*. *Plant Molecular Biology* 33: 1093–1096.

Wake, D. B., K. P. Yanev, and C. W. Brown. 1986. Intraspecific sympatry in a "ring species," the plethodontid salamander *Ensatina eschscholtzii*, in southern California. *Evolution* 40: 866–868.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.

Whipple, A. A., R. M. Grossinger, and F. W. Davis. 2011. Shifting baselines in a California oak savanna: Nineteenth century data to inform restoration scenarios. *Restoration Ecology* 19: 88–101.