

# Effects of landscape features on gene flow of valley oaks (*Quercus lobata*)

Maryam Gharehaghaji  · Emily S. Minor ·  
Mary V. Ashley · Saji T. Abraham · Walter D. Koenig

Received: 16 April 2016 / Accepted: 24 January 2017  
© Springer Science+Business Media Dordrecht 2017

**Abstract** Landscape features affect habitat connectivity and patterns of gene flow and hence influence genetic structure among populations. We studied valley oak (*Quercus lobata*), a threatened species of California (USA) savannas and oak woodlands, with a distribution forming a ring around the Central Valley grasslands. Our main goal was to determine the role of topography and land cover on patterns of gene flow and to test whether elevation or land cover forms stronger barriers to gene flow among valley oak populations. We sampled valley oaks in 12 populations across the range of this species, genotyped each tree at eight nuclear microsatellite loci, and created a series of resistance surfaces by assigning different

resistance values to land cover type and elevation. We also estimated recent migration rates and evaluated them with regard to landscape features. There was a significant but weak relationship between Euclidian distance and genetic distance. There was no relationship between genetic distances and land cover, but a significant relationship between genetic distances and elevation resistance. We conclude that gene flow is restricted by high elevations in the northern part of the valley oak range and by high elevations and the Central Valley further south. Migration rate analysis indicated some gene flow occurring east–west but we suggest that the high connectivity in the northern Central Valley is facilitating the formation of these links. We predict that southern populations may become more differentiated in the future through genetic isolation and local adaptation taking place in the face of climate change.

---

Communicated by Santiago C. Gonzalez-Martinez.

---

M. Gharehaghaji (✉) · E. S. Minor · M. V. Ashley  
Department of Biological Sciences, University of Illinois  
at Chicago, Chicago, IL, USA  
e-mail: mghare2@uic.edu

E. S. Minor  
Institute for Environmental Science and Policy,  
University of Illinois at Chicago, Chicago, IL, USA

S. T. Abraham  
Med-Nephrology Research lab, University of Michigan,  
Ann Arbor, MI, USA

W. D. Koenig  
Cornell Lab of Ornithology and Department of  
Neurobiology and Behavior, Cornell University, Ithaca,  
NY, USA

**Keywords** Landscape resistance · Landscape genetics · Circuit theory · Valley oak · Topography · Gene flow

## Introduction

Landscape features and topography affect habitat connectivity and gene flow and thus influence genetic structuring of populations at a regional level (Gomez et al. 2005). Reduced connectivity and gene flow can

decrease the resilience, adaptability, fitness, and fertility of populations, especially when populations are small (Frankham 2006). Despite the clear importance of habitat connectivity for population persistence, factors facilitating or inhibiting connectivity are largely unknown for most species.

Landscape genetic approaches offer powerful tools for explicitly quantifying the effects of landscape features on spatial genetic variation (Holderegger and Wagner 2008; Balkenhol et al. 2009), but landscape distance/resistance studies are less common in the study of plant systems compared to animals (Holderegger et al. 2010; Storfer et al. 2010). Examples include studies on Arkansas valley evening primrose (*Oenothera harringtonii*, Rhodes et al. 2014) and Engelmann oak (*Quercus engelmannii*, Ortego et al. 2012), both of which found that isolation by distance (IBD) and topographic features, especially elevation, have a significant effect on genetic patterns. Cushman et al. (2014), on the other hand, identified rivers and streams as landscape features that influence genetic structure and restrict gene flow of Fremont cottonwood (*Populus fremontii*), but IBD was not detected in their study.

The pronounced ecological gradients and complex climatic and geological history of California (CA) has produced one of the most geographically complex patterns of genetic diversity on Earth (Raven and Axelrod 1978; Calsbeek et al. 2003; Davis et al. 2008). As a result, the CA biota offers an excellent opportunity to study landscape features and their relation to observed patterns of genetic diversity. Oak woodlands and savannas make up nearly a quarter of California's forests and woodlands (Davis et al. 1998). California's oak landscapes have been a recurrent focus of conservation attention due to concerns about habitat conversion to farmland, vineyards, and other land uses, arrival and spread of invasive diseases such as sudden oak death syndrome, poor regeneration (Zavaleta et al. 2007), and climate change. Such factors can result in surviving populations being genetically structured and effectively isolated (Sork et al. 2010). Understanding how environmental and landscape-level features influence genetic variation and structure in these oaks will help us identify specific barriers to dispersal and corridors that facilitate gene flow, enabling improved prediction of the impact of future changes to the landscape (Schwartz et al. 2009).

Valley oak (*Quercus lobata*) is an ecologically important species endemic to CA savannas and oak woodlands. Due to its propensity to occupy locations that are desirable for farmlands, ranches, vineyards, and suburban developments, valley oak has lost more of its habitat than any other oak species in CA (Pluess et al. 2009; Whipple et al. 2011; Ashley et al. 2015). For example, it occupies 5% of its historical range in San Joaquin Valley (Kelly et al. 2005). Alterations in the structure and distribution of valley oak stands could affect main avian dispersers of the species, acorn woodpeckers (*Melanerpes formicivorus*) and California scrub-jays (*Aphelocoma californica*), a host of other species, both vertebrate and invertebrate, that are associated with CA oak woodlands, along with rates of soil and landform development (Pavlik et al. 1991; Howard 1992).

Many studies of genetic structure have revealed strong diverged lineages within species across geographic features (Rissler et al. 2006). In CA, one important feature is the dry, flat Central Valley. The ranges of dispersal-limited, habitat-specialized species and ones less adapted for the arid conditions are intersected by the Central Valley. Such species include Arthropoda such as *Greya politella* (Brown et al. 1997), mammals, particularly rodents (Maldonado et al. 2001), amphibians such as *Ensatina eschscholtzii* (Wake et al. 1986), *Batrachoseps* (Jockusch et al. 2001) reptiles including *Diadophis punctatus* (Feldman 2000), and several species of oaks, genus *Quercus* (Gugger et al. 2013).

Grivet et al. (2006) hypothesized that the geological history of CA is a primary driver of the current genetic structure of oak populations and that oak populations have been less strongly affected by past climatic changes such as the last glacial maximum (LGM). In addition, Lancaster and Kay (2013) have shown that California's topographic complexity and geographical location are critical for plant species persistence and diversification, whereas its temporary climatic conditions have been less important. On the other hand, Sork et al. (2010) and Gugger et al. (2013) showed that genetic variation in valley oak is significantly correlated with climatic variation and LGM climate.

There are conflicting reports as to whether the arid flats of California's Great Central Valley are in fact a barrier to gene flow. Previous studies on valley oak (Grivet et al. 2008; Sork et al. 2010; Gugger et al. 2013) report genetic differentiation

among valley oak populations in the inland foothills of the Coastal Range and the western foothills of the Sierra Nevada, with higher gene flow north-to-south along climatically similar mountain corridors than east-to-west across the Central Valley. These findings suggest a role for environmentally mediated historical vicariance and that the Central Valley may be a barrier to gene flow in this species. Prior studies indicating that geographical barriers, possibly combined with climatological changes, may have affected the Californian fauna and flora include Calsbeek et al. (2003) and Lapointe and Rissler (2005). In contrast, Gugger et al. (2016) and Sork et al. (2016) failed to observe any clear east–west structure in valley oaks using candidate genes. Here we explicitly test if topographic features are a barrier to gene flow in valley oak by creating resistance surfaces.

Topography is but one of the factors likely to correlate with environmental variables important in the ecology of oak species (Thomassen et al. 2010). Land use change also threatens to reduce valley oak genetic diversity and evolutionary potential (Grivet et al. 2008). We hypothesize that different land cover influences genetic structure. In particular, woodland and mixed forest that are typical habitats for valley oaks, and to a lesser extent grassland and oak savannah, allow for pollen and seed dispersal by acorn dispersers, e.g., scrub-jays and, therefore, should facilitate gene flow, whereas shrubland, urban and non-vegetated land cover should inhibit it.

Here we extend a recent study by Ashley et al. (2015), who, along with others (Grivet et al. 2008; Sork et al. 2010), suggested that California's Central Valley and other topographical features including the Coast Ranges and the Sierra Nevada mountains are barriers to gene flow in valley oak. Our aim was to disentangle the effect of distance and landscape features such as elevation and land cover on the observed pattern of genetic variation of valley oak populations. We predicted that isolation by resistance (IBR) with different resistance values assigned to land cover types and elevation would better explain genetic variation than IBD. In addition, we measured recent migration rates among populations to test whether there are more north–south than east–west links and whether gene flow is matched to topographic and land cover barriers.

## Methods

### Study species

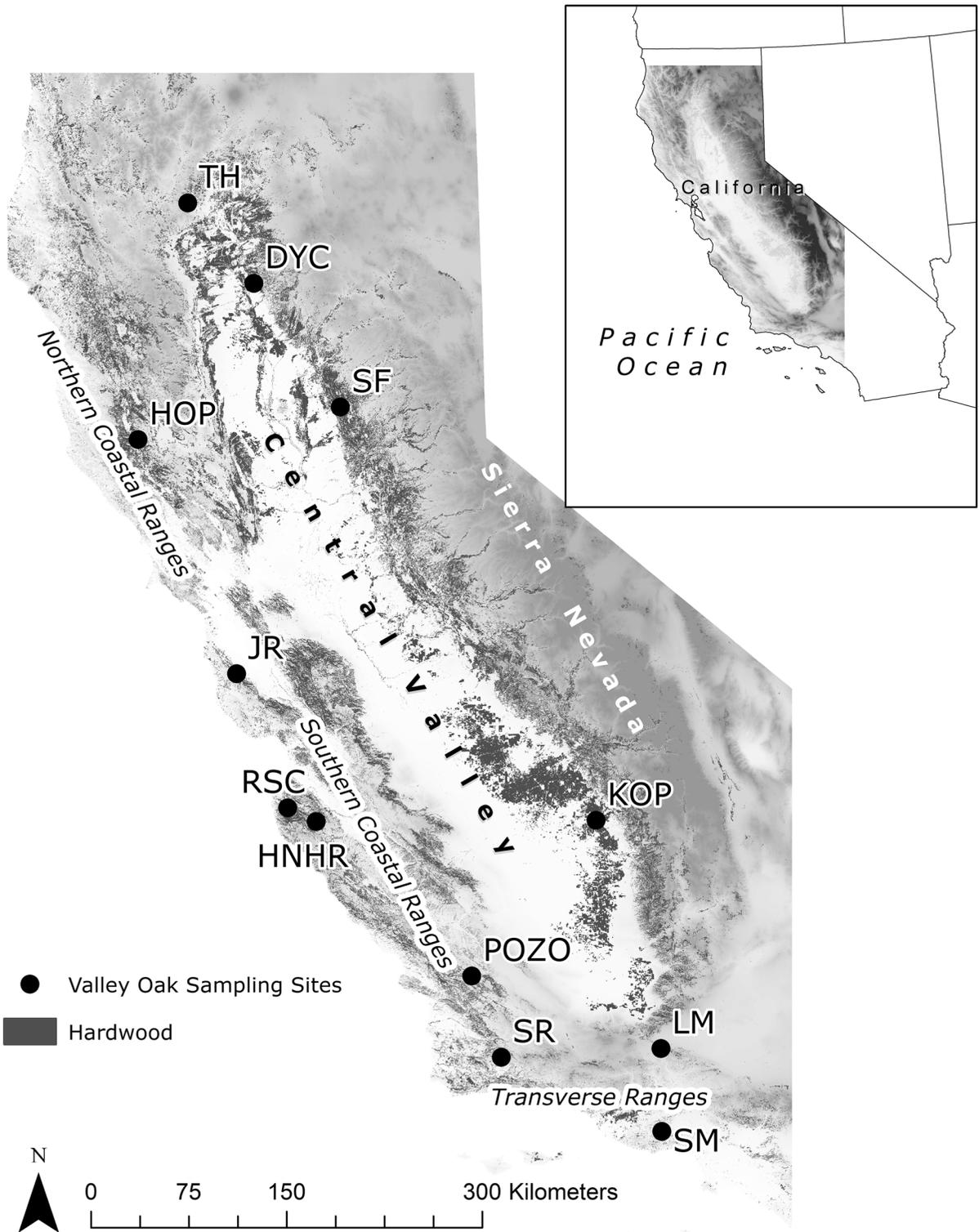
Valley oak is the largest North American oak, with trees standing 10–25 m tall and 0.5–0.7 m diameter at breast height (Munz 1973). It is wind pollinated and its seeds are dispersed by birds and rodents (Grivet et al. 2005; Pesendorfer et al. 2016). Studies of historical and contemporary gene flow indicate that the scale of pollen and seed dispersal in some populations is on the range of 100–300 m, which allows opportunity for adaptation to local environmental conditions (Pluess et al. 2009; Sork et al. 2010). Valley oak habitat is found in the lower elevations of the central and northern Coastal Ranges and the Sierra Nevada mountains and forms a ring around Central Valley grasslands (Fig. 1). The density of valley oaks varies widely, from closed-canopy forests in some regions to open savannahs in drier parts of their range.

California's Central Valley covers approximately 155,000 km<sup>2</sup> of land or nearly 40% of the total land area of the state. On average, it is about 725 km long and 80 km wide, is primarily agricultural, and is often in a state of moderate to severe drought (Whiteside 2007). It is also home to 588 rare and endangered species, more than any other region in the continental United States. The rapidly growing population of the Central Valley has caused the loss, degradation, and fragmentation of habitats through the development of agricultural and urban areas (Hosley et al. 2012). Scattered valley oaks occur in the Central Valley but are relatively uncommon due to agricultural development and loss of riparian habitat.

### Genetic data

Leaf tissue was collected from 270 trees in 12 populations spread across most of the species' range (Fig. 1). Samples were genotyped at eight nuclear microsatellite loci. The eight loci used were highly variable, with an average of 8.12 alleles per locus (Ashley et al. 2015). Sampling sites and microsatellite genotyping are described in more detail by Ashley et al. (2015).

Ashley et al. (2015) calculated genetic differentiation statistics including Nei's (1973) multiallelic estimate of  $F_{ST}$  and  $D_{JOST}$  (Jost 2008) using the R (R Development Core Team 2014) package *diveR*sity



◀ **Fig. 1** Valley oak (*Quercus lobata*) sampling sites across hardwood forest (USDA 1981): Tower House, Shasta (TH), Dye Creek Reserve, Tehama (DYC), Sierra Foothills Station, Yuba (SF), Kaweah Oaks/River Preserve, Tulare (KOP), Liebre Mountain, Los Angeles (LM), Santa Monica Mountains, Ventura (SM), Sedgwick Reserve, Santa Barbara (SR), Pozo, San Louis (POZO), Hastings Reservation, Monterey, (HNHR), Rancho San Carlos, Monterey (RSC), Jasper Ridge, San Mateo (JR), Hopland Research Station, Mendocino (HOP)

(Keenan et al. 2013) and these were used as measures of genetic distance in this study.  $D_{\text{JOST}}$  (Jost 2008) measures the fraction of allelic variation among populations and overcomes the problem that  $F_{\text{ST}}$  is limited to low values when heterozygosity is high. However,  $D_{\text{JOST}}$  can be biased upwards. We therefore used both following the recommendation of Whitlock (2011).

#### Landscape analysis

To measure IBD, we computed the pairwise straight-line (Euclidean) distances between valley oak stands in GenAIEx (Peakall and Smouse 2012). To measure IBR, we modeled landscape resistance as a function of land cover and elevation. We generated multiple resistance grids in ArcGIS 9.3 (Esri 2014) in which each grid cell was assigned a value based on the hypothesized resistance of gene flow to the land in that cell.

As a precursor to the resistance grids, we first created a land cover raster by combining two GIS datasets. The most recent CALVEG land cover dataset from 2011 (30 × 30 m resolution), originally created by USDA Forest Service (1981), includes detailed vegetation classification including valley oak habitat but does not cover the entire study region. Therefore, we filled in areas with no CALVEG coverage with the USGS GAP Land Cover data. By combining the two datasets, we created a 30 × 30 m land cover map with seven land cover types (Table 1). We used Shuttle Radar Topographic Mission (SRTM) data for elevation, which are available in 3 arc second intervals (~75 m resolution in our study region).

To identify the optimum resistance values for elevation and land cover, we used the model optimization approach introduced by Cushman et al. (2006) and refined by Shirk et al. (2010). We related each land cover type and elevation to landscape

**Table 1** Ranking of cover types based on expert opinion indicating relative degree of resistance to gene flow

Land cover type	Resistance ranking
Valley oak	1
Hardwood	2
Mixed forest	3
Grassland	4
Shrubland	5
Conifer	6
Non-vegetation (urban, water, barren)	7

Woodland, mixed forest, and savanna that are typical habitats for valley oaks facilitate gene flow, whereas shrubland, urban and non-vegetated land cover inhibit it

resistance with a mathematical function (see model functions below) and used each function to reclassify appropriate raster data into a resistance surface. We generated alternative parameter values favoring the direction that increased the correlation until we observed a unimodal peak of support (Shirk et al. 2010).

For land cover, we first ranked resistance of each land cover type based on expert (M. V. Ashley personal recommendation) opinion about valley oak biology and habitat relations (Table 1). Our ranking of habitat was based on resistance or facilitation of pollen flow and seed movement, and in accordance with valley oak favored habitat and occurrence. Second, we reclassified the resulting categorical rank raster according to the function from Shirk et al. (2010):

$$R_i = (\text{Rank}_i / V_{\text{max}})^x \times R_{\text{max}} \quad (1)$$

where  $R_i$  is the resistance value of land cover type  $i$ ,  $\text{Rank}_i$  is the resistance rank of land cover type  $i$  (from Table 1),  $V_{\text{max}}$  is the maximum resistance rank (equal to 7, in our case, which was the resistance rank for all non-vegetation cover),  $x$  is the shape parameter that informs relative differences in resistance among different cover types, and  $R_{\text{max}}$  is the maximum resistance value for any land cover type. We tested five values of  $R_{\text{max}}$  (5, 10, 25, 50, and 100) and five values of  $x$  (0, 0.1, 0.25, 0.5, and 1) and calculated a resistance value for each land cover type for each set of parameters for a total of 25 candidate models. Valley oak habitat was always assigned a resistance value of 0.01. By systematically varying  $x$  and  $R_{\text{max}}$  values, we aimed to find the optimal hypothesis about the effect of land cover on gene flow.

Valley oak grows at higher elevations in the southern part of its range and lower elevations in the northern part (Griffin and Critchfield 1972). Therefore, we divided valley oak habitat range into north and south to account for this apparent geographic division. We tested several different latitudes as the dividing line and settled on 35°52'00" latitude as the dividing line between 'north' and 'south' that resulted in the strongest relation between resistance as a function of elevation and genetic variation. No study, to our knowledge, has specifically identified a latitude leading to a disjunction of genetic structure. Grivet et al. (2008) and Ashley et al. (2015), however, found that the western and eastern populations of valley oak differ genetically at the northern and southern ends of its range.

We then reclassified the digital elevation model according to an inverse Gaussian function.

$$R_{\max} = R_{\max} + 1 - R_{\max} \times e^{-\frac{(\text{elevation} - E_{\text{opt}})^2}{2 \times E_{\text{SD}}^2}} \quad (2)$$

where  $R_{\max}$ ,  $E_{\text{opt}}$ , and  $E_{\text{SD}}$  represent the maximum resistance, optimal elevation, and the standard deviation about the optimal elevation, respectively. Thus, as elevation increases or decreases away from  $E_{\text{opt}}$ , resistance increases to  $R_{\max}$  at a rate governed by  $E_{\text{SD}}$ .

We evaluated three values of  $R_{\max}$  (5, 10, and 25), three standard deviations ( $E_{\text{SD}}$ ) of 500, 1000, and 1500, and five values of  $E_{\text{opt}}$  in the north (ranging from 0 to 800 m in 200 m increments) and five values of  $E_{\text{opt}}$  in the south (ranging from 900 to 1700 m, in 200 m increments) for a total of 45 candidate models of elevation resistance.

We used Circuitscape 4.0.5 (McRae et al. 2013) to measure pairwise landscape resistance among all populations for each landscape resistance scenario. Circuitscape proved to be faster and more efficient compared with the least cost path (LCP) procedure in ArcGIS due to the smaller number of resistance surfaces made compared to LCP. In the LCP method, least cost distance is measured from a given population to each of the other populations in each run, whereas in CIRCUITSCAPE all the populations can be given as the input, and thus resistance distance is measured from all population to every other population in a single run. We used a four-neighbor connection scheme to calculate resistance between point locations. Because our land cover map exceeded the maximum number of grid cells that Circuitscape

can handle, we aggregated groups of 25 pixels in each land cover resistance surface and groups of four pixels in each elevation resistance surface to create new resistance surfaces with  $150 \times 150$  m pixels whose values reflected the average resistance of the aggregated cells.

We used causal modeling (Legendre and Troussellier 1988) to find the relationships between genetic distance among populations and compared multiple alternative hypotheses of landscape resistance to evaluate the factors that limit gene flow (Cushman et al. 2006). After calculating resistance between each pair of populations for each resistance surface, we used Mantel tests (Mantel 1967) and partial Mantel tests (Smouse et al. 1986) to test the relationship between genetic distance and landscape resistance when removing the effects of Euclidean distance. In all Mantel tests, correlation coefficients for genetic distances were consistently higher for  $F_{\text{ST}}$  than  $D_{\text{JOST}}$ . Therefore, we only report results from  $F_{\text{ST}}$ , which is a basic descriptor of population structure (Neigel 2002). We identified the most supported model as the one with the highest significant correlation.

Mantel tests have been criticized for their lower power compared to traditional linear models leading to underestimation of the true magnitude of a relationship (Legendre and Fortin 2010). Therefore, we also conducted general linear mixed models (GLMMs), which account for dependency between pairwise observations in a distance matrix (Yang 2004) and are a more robust approach for landscape genetic inference. Before performing a GLMM, we checked collinearity between the explanatory variables. We fitted maximum likelihood population effects (MLPE) models (Clarke et al. 2002; Van Strien et al. 2012) with residual maximum likelihood (REML) estimation using the "lmer" function in the package "lme4" (Bates et al. 2011) for R (R Development Core Team 2014). Pairwise genetic distance ( $F_{\text{ST}}$ ) was used as the dependent variable, while elevation and land cover resistance between populations were the independent fixed variable. We used ANOVA to select the best model.

We explored recent migration rates among the twelve populations using BayesAss 3.0.3 (Wilson and Rannala 2003). BayesAss uses a Bayesian MCMC approach to estimate asymmetric migration over the last two to three generations (Wilson and Rannala 2003). We assessed convergence by using long ( $10^7$ )

iteration runs with large sampling frequencies (2000), and long ( $10^6$ ) burn-in periods and by comparing migration rates across 10 replicate runs using different starting seeds, as recommended by the program's authors (Wilson and Rannala 2003). We examined the results for evidence of convergence by comparing the posterior mean (the matrix parameter of migrational rates) for consistent results (Clark et al. 2013). We set delta values of allele frequency, migration rate, and inbreeding to 0.3. Delta is the parameter that defines the size of the proposed change to the parameter values at each iteration (Wilson and Rannala 2003).

## Results

Our landscape modeling results revealed a significant but weak relationship between genetic distance and Euclidean distance ( $r = 0.16$ ,  $P = 0.001$ ). The optimized IBR models provided better support, as both land cover ( $r = 0.25$ ,  $P = 0.041$ ) and elevation ( $r = 0.64$ ,  $P = 0.001$ ) were significantly related to genetic distance (Table 2). However, after removing the effect of Euclidean distance with a partial Mantel test, land cover resistance was no longer significantly related to genetic distance (Mantel  $r = 0.23$ ,  $P = 0.76$ ). Furthermore, there was no strong unimodal peak for any of the land cover resistance models (Fig. 2). In contrast, the Mantel correlation between resistance as a function of elevation and genetic distance remained high after removing the effect of Euclidean distance ( $r = 0.63$ ,  $P = 0.001$ ). The optimum elevation resistance model included a maximum resistance ( $R_{\max}$ ) of 25, an optimal elevation ( $E_{\text{opt}}$ ) of 200 m in the north and 1100 m in the south and a standard deviation about the optimal elevation ( $E_{\text{SD}}$ ) of 500 (Fig. 2).

The correlation between elevation resistance and land cover resistance was not significant ( $r = 0.32$ ,  $P = 0.09$ ). The full model showed that elevation resistance was significant ( $P < 0.001$ ) in determining genetic variation, whereas land cover resistance was not ( $P = 0.76$ ). In the reduced model with land cover resistance as the only fixed variable, land cover resistance was significant ( $P = 0.02$ ). The reduced model with elevation resistance as the fixed variable was selected as the best model to describe genetic variation (Table 3).

The Bayesian estimations of migration rates were low for many of the populations (Table 4). Nevertheless, they revealed significant gene flow occurring for the eastern and western populations. Gene flow was asymmetric for most site pairs (Fig. 3).

## Discussion

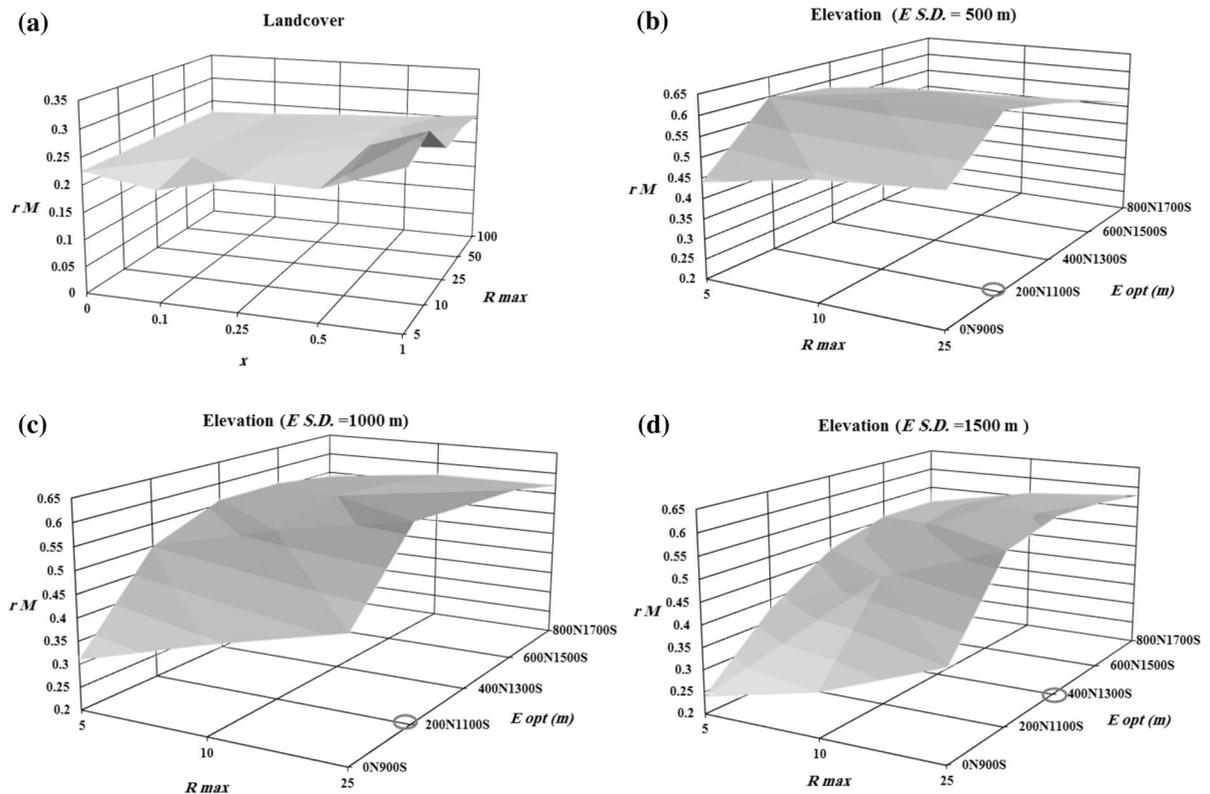
Valley oak demonstrates modest genetic structure across the species' range (Grivet et al. 2008; Ashley et al. 2015). Our results revealed a significant but weak relation between Euclidian distance and genetic distance across valley oak's range. Furthermore, we did not find land cover to be a significant predictor of genetic differentiation among populations. In contrast, elevation provided an important factor for explaining valley oak genetic structure. This result is in accordance with other studies that have found stable habitat conditions and topographic features to be the most relevant factors determining population differentiation within the CA region (Calsbeek et al. 2003; Davis et al. 2008; Lancaster and Kay 2013).

Land cover was not a strong indicator of gene flow. This may be due to the fact that the adult trees that were sampled were in most cases established over

**Table 2** Mantel correlations between the most highly supported resistance models (as a function of elevation or land cover) and genetic distance ( $F_{\text{ST}}$ ) alone and after partialling out the effect of Euclidean distance (IBD)

Model	Mantel $r$	$P$ value	Optimized parameter value
$F_{\text{ST}} \sim \text{IBD}$	0.16	<b>0.001</b>	
$F_{\text{ST}} \sim \text{elevation}$	0.64	<b>0.001</b>	$E_{\text{opt}} = 200$ m (N), 1100 m (S), $R_{\max} = 25$
$F_{\text{ST}} \sim \text{land cover}$	0.25	<b>0.041</b>	$X = 0.25$ , $R_{\max} = 25$
$F_{\text{ST}} \sim \text{elevation} \mid \text{IBD}$	0.63	<b>0.001</b>	$E_{\text{opt}} = 200$ m (N), 1100 m (S), $R_{\max} = 25$
$F_{\text{ST}} \sim \text{land cover} \mid \text{IBD}$	0.23	0.760	$X = 0.25$ , $R_{\max} = 25$

$N$  and  $S$  stand for North and South range of valley oak. Significant values are indicated in bold



**Fig. 2** Different scenarios representing Mantel correlation ( $rM$ ) between genetic distance ( $F_{ST}$ ) and resistance as a function of Elevation and land cover.  $R_{max}$  is the maximum resistance ( $x$ -axis) and  $E_{opt}$  is optimum elevation ( $z$ -axis) in the elevation

scenarios.  $X$  is the response shape in the land cover scenarios. The resistance model with peak of support for highest correlation to genetic distance is shown with a *gray oval*

**Table 3** Parameter estimates for the linear mixed effects models

Factor	Parameter	$\beta$	S.E.	$T$ -value	$P$ value	AIC
Elevation and land cover	Intercept	-0.0050	0.0129	-0.46	0.65	-279.7
	Elevation	0.0019	0.0003	5.56	<0.001	
	Landcover	0.0003	0.0010	0.31	0.76	
<b>Elevation</b>	Intercept	-0.0030	0.0109	-0.35	0.73	<b>-281.6</b>
	Elevation	0.0019	0.0003	6.29	<0.001	
Land cover	Intercept	0.0256	0.0144	1.78	0.08	-256.5
	Land cover	0.0029	0.0011	2.43	0.02	
<b>Euclidian distance</b>	Intercept	0.0472	0.0092	4.75	<0.001	<b>-252.7</b>
	Euclidian distance	0.0002	0.0002	1.36		

Elevation and land cover resistance surface are the explanatory fixed variables and  $F_{ST}$  is the response variable. Effect of Euclidian distance was measured in the univariate model. Beta, standard error,  $t$ -value,  $P$  value estimates are from the mixed effect model fit by restricted maximum likelihood using lme4. The best fit model is in bold

100 years ago, before most of the anthropogenic landscape changes in this region occurred. Our measure of landscape resistance reflects the current

landscape, whereas genetic distance reflects historic gene flow. Although land cover changes are likely to eventually have strong population effects through loss

**Table 4** Bayesian estimates of recent migration rates in BayesAss among 12 valley oak populations

From	To											
	TH	DYC	SF	KOP	LM	SM	SR	POZ	HNR	RSC	JR	HOP
TH	<u>0.679</u>	0.004	0.001	0.004	0.005	0.001	0.004	0.004	0.002	0.004	0.003	0.004
DYC	0.003	<u>0.682</u>	0.001	0.004	0.005	0.001	0.004	0.004	0.002	0.004	0.004	0.004
SF	0.010	<b>0.266</b>	<u>0.982</u>	<b>0.265</b>	0.011	0.001	0.030	<b>0.139</b>	0.013	0.004	<b>0.112</b>	0.030
KOP	0.003	0.004	0.001	<u>0.679</u>	0.005	0.001	0.004	0.004	0.002	0.004	0.004	0.004
LM	0.003	0.004	0.001	0.004	<u>0.687</u>	0.001	0.004	0.004	0.002	0.004	0.004	0.004
SM	0.003	0.004	0.001	0.004	0.007	<u>0.987</u>	0.004	0.005	0.003	0.004	0.005	0.004
SR	0.003	0.004	0.001	0.004	0.005	0.001	<u>0.681</u>	0.004	0.002	0.004	0.004	0.004
POZ	0.004	0.004	0.001	0.004	0.006	0.001	0.011	<u>0.706</u>	0.003	0.004	0.004	0.004
HNR	<b>0.245</b>	0.007	0.002	0.008	0.041	0.001	0.025	<b>0.102</b>	<u>0.955</u>	<b>0.279</b>	<b>0.111</b>	<b>0.251</b>
RSC	0.003	0.004	0.001	0.004	0.005	0.001	0.004	0.004	0.002	<u>0.679</u>	0.004	0.004
JR	0.034	0.006	0.001	0.013	<b>0.210</b>	0.001	<b>0.229</b>	0.015	0.005	0.004	<u>0.699</u>	0.010
HOP	0.003	0.004	0.001	0.004	0.005	0.001	0.004	0.004	0.002	0.004	0.004	<u>0.681</u>

Underlined values on the diagonal indicate the proportion of individuals in each generation that are not migrants (resident populations). Simulations show that in instances where there is no information in the data, migration rate mean and 95% confidence interval for datasets with 12 populations are 0.0150 (0.000, 0.100). Bold values show informative (>0.100) migration rates

of corridors and stepping stones, there is likely to be a considerable time lag in response, especially in long-lived species (Wright 1943; Waples 1998). This mismatch can be an issue when studying landscape genetics in landscapes that are changing rapidly. On the other hand, land cover may offer little resistance to pollen movement, as has been shown in valley oak (Abraham et al. 2011) and other oak species (Ashley 2010; Craft and Ashley 2010).

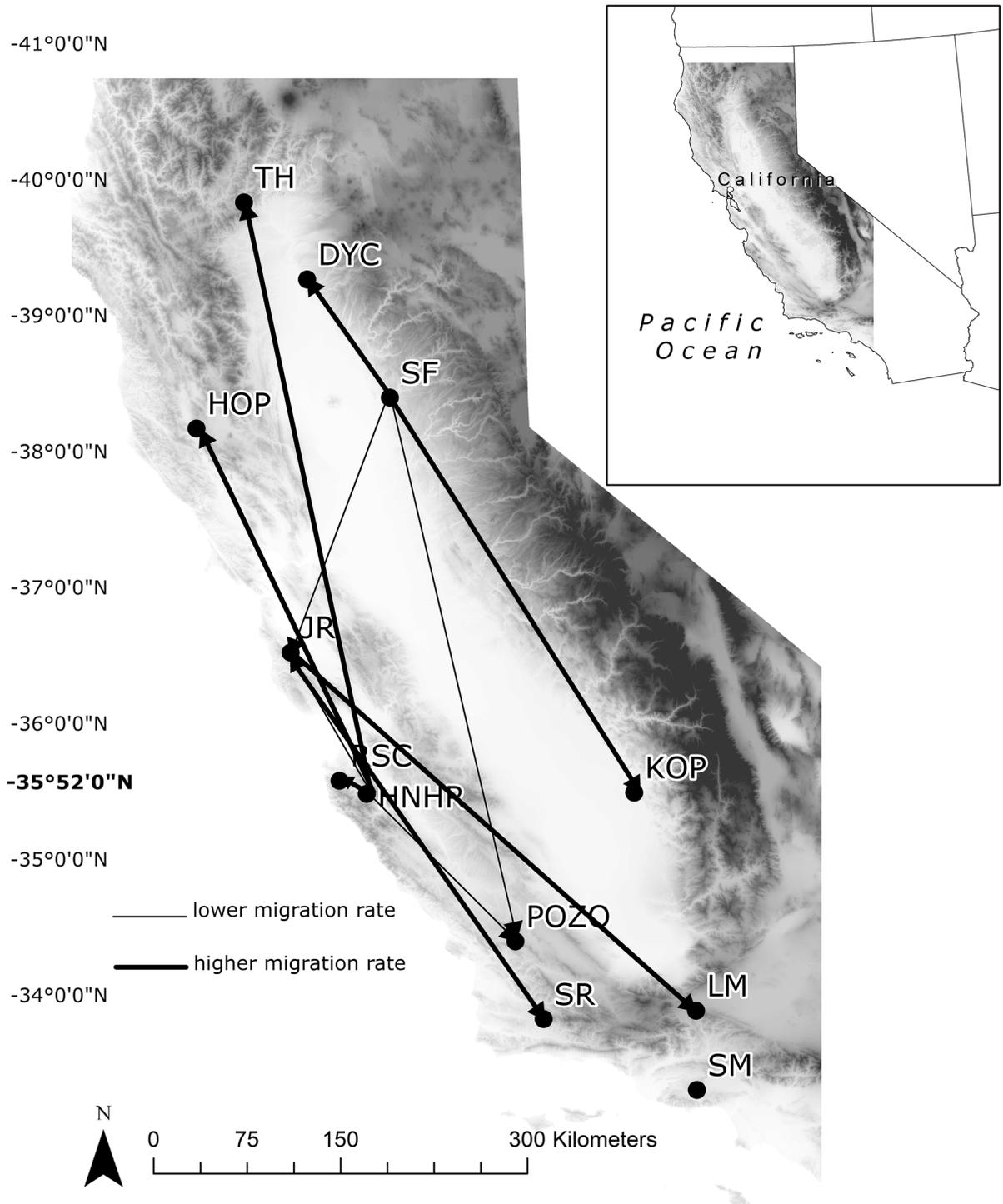
After dividing the distribution of valley oak latitudinally, we deduced that gene flow was restricted by high elevations in the northern part of its range and by high elevations and the low elevation of Central Valley further south. Our results provide support for the hypothesis put forth by previous studies that topographic features such as mountain ranges and the large, flat Central Valley impact patterns of gene flow in this species (Grivet et al. 2006; Sork et al. 2010; Ashley et al. 2015).

We found a greater resistance effect of elevation in the southern range of valley oak. In other words, the Central Valley is a smaller barrier to gene flow as valley oaks grow at lower elevation in the more northern part of its range, whereas in the southern part of its range where they inhabit higher elevations, the Central Valley is far below their optimum elevation. This supports the results of Grivet et al. (2008), who found shared chloroplast haplotypes between eastern

and western populations north of the Monterey Bay area indicating connectivity in the north part of Central Valley. Similarly, Sork et al. (2010) found that gene flow in valley oak is less restricted in the east–west direction in the northern part of its range, probably due to a strong riparian network in the north. These results all support the conclusion that topographic features have a greater impact on southern-most populations of this species.

In CA, an increase in temperature associated with climate change could increase vegetation productivity given adequate moisture availability, especially in cooler regions of the state along the north-central coast and at high elevations, and could also promote advancement of grassland and reduction of forest particularly in the southern end of the Central Valley (Lenihan et al. 2003). We therefore predict higher connectivity of northern populations compared to the south because of more contiguous forest and riverine ecosystems. We also predict that the southern end of Central Valley will become a greater barrier in the future. Higher connectivity of northern populations may facilitate resilience to climate change through dispersal of alleles and genotypes better adapted to changing conditions.

Our estimates of recent migration rates show higher gene flow north–south than east–west across the Central Valley (Fig. 3), as previously reported by



**Fig. 3** Recent migration pattern for the 12 valley oak sampling sites using BayesAss. The informative ( $>0.100$ ) values of migration rate are shown. *Thicker lines* show migration

rate  $>0.2$  and *thinner lines* show migration rate  $<0.2$ . The latitude in which we used to divide valley oak habitat range into north and south is shown in **bold**

Grivet et al. (2008) and Sork et al. (2010). We did, however, detect evidence of some east–west gene flow. We also saw more connections in the west, which is in accordance with Grivet et al. (2008). In the west, JR and HNHR were the most important gene donors, while in the east SF was an important gene donor. These nodes appear to facilitate gene flow by connecting distant populations through a series of migrational routes that mostly occur in the northern Central Valley and should be prioritized for conservation efforts.

The only notable gene flow crossing the southern Central Valley was from SF to POZO, but we suggest that in this case pollen has not directly crossed the Central Valley but rather has occurred through chains of populations, with recent migration (Table 4) between SF and JR (0.112), JR and HNHR (0.111), and HNHR and POZO (0.102). Our results also indicate that gene movement occurs in long-distance events and that all populations of valley oak are connected to at least one other population either directly or by a chain of populations. The exception is the Santa Monica population, as previously reported by Ashley et al. (2015). Populations in the Santa Monica mountains in the Transverse ranges compose a distinct group and should be the focus of conservation efforts (Grivet et al. 2008; Ashley et al. 2015).

Further research into the role and influence of other landscape variables such as slope and aspect will improve our understanding of gene flow for this species. For example, the role of slope and aspect in structuring valley oak populations may be important for steep and opposing north and south slopes such as the transverse ranges if pollen is unable to reach the other side. Future studies would also benefit by including more study sites and marker loci to further elucidate ecological barriers to gene flow. Using more loci would improve resolution as the magnitude of correlations between landscape pattern and genetic structure of a population of organisms is highly dependent on the number of loci analyzed per individual (Landguth et al. 2012).

The more genetically isolated populations located south of the Transverse Ranges are at great risk of diversity loss and thus likely to be more constrained in their ability to tolerate rapid climate change. Although long-distance pollen dispersal is common in oaks (Ashley et al. 2015) and seed dispersal can go beyond local sites (Sezen et al. 2005), the diversity of the

pollen/seed pool of the original source is also an important factor for adaptation to climate change. A long-term connectivity plan to restore valley oaks throughout their historical distribution, especially in mid-elevation areas, would be particularly valuable. Conservation strategies should include restoring and conserving suitable habitat for valley oak based on optimum elevation and topographic structure to maintain functional connectivity across its habitat range, protecting the nodes of its gene flow network, maintenance of valley oak habitat especially in the northern Central Valley, and protection of riverine ecosystems that create migrational corridors and enhance gene flow.

**Acknowledgements** We thank Nina Savar, GIS instructor at UIC for her invaluable advice on spatial analyses and Marti Whittier for collecting samples for one site used in this study. Funding was provided by NSF grant DEB-0816691 to WDK. We are grateful to Matthew McCary and David Lowenstein, Elsa Anderson, Alexis Smith, and Megan Garfinkel of the Minor lab for feedback on the manuscript.

## References

- Abraham ST, Zaya DN, Koenig WD, Ashley MV (2011) Interspecific and intraspecific pollination patterns of valley oak, *Quercus lobata*, in a mixed stand in coastal central California. *Int J Plant Sci* 172:691–699
- Ashley MV (2010) Plant parentage, pollination, and dispersal: how DNA microsatellites have altered the landscape. *Crit Rev Plant Sci* 29:148–161
- Ashley MV, Abraham ST, Backs JR, Koenig WD (2015) Landscape genetics and population structure in valley oak (*Quercus lobata* Née). *Am J Bot* 102:2124–2131
- Balkenhol N, Waits LP, Dezzani RJ (2009) Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. *Ecography* 32:818–830
- Bates D, Maechler M, Bolker B (2011) lme4: linear mixed-effects models using S4 classes. R package version 0. p. 999375–999342
- Brown JM, Leebens-Mack JH, Thompson JN, Pellmyr O, Harrison RG (1997) Phylogeography and host association in a pollinating seed parasite *Greya politella* (Lepidoptera: Prodoxidae). *Mol Ecol* 6:215–224
- Calsbeek R, Thompson JN, Richardson JE (2003) Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. *Mol Ecol* 12:1021–1029
- Clarke RT, Rothery P, Raybould AF (2002) Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. *J Agric Biol Envir Stat* 7:361–372
- Clarke LJ, Whalen MA, Mackay DA (2013) Cutting grass on desert islands: genetic structure of disjunct coastal and

- central Australian populations of *Gahnia trifida* (Cyperaceae). *J Biogeogr* 40:1071–1081
- Craft KJ, Ashley MV (2010) Pollen-mediated gene flow in isolated and continuous stands of bur oak, *Quercus macrocarpa* (Fagaceae). *Am J Bot* 97:1999–2006
- Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *Am Nat* 168:486–499
- Cushman SA, Max TL, Whitham TG, Allan GJ (2014) River network connectivity and climate gradients drive genetic differentiation in a riparian foundation tree. *Ecol Appl* 24:1000–1014
- Davis MA, Wrage KJ, Reich PB (1998) Competition between tree seedlings and herbaceous vegetation: support for a theory of resource supply and demand. *J Ecol* 86:652–661
- Davis EB, Koo MS, Conroy C, Patton JL, Moritz C (2008) The California hotspots project: identifying regions of rapid diversification of mammals. *Mol Ecol* 17:120–138
- ESRI (2014) ArcGIS Desktop : Release 10.3. Environmental Systems Research Institute, Redlands
- Feldman CR (2000) Comparative phylogeography of three Californian reptiles: *Contia tenuis*, *Diadophis punctatus*, *Elgaria multicarinata*. *Biology*. San Francisco State University, San Francisco
- Frankham R (2006) Genetics and landscape connectivity. In: Crooks KR, Sanjayan M (eds) *Connectivity conservation*. Cambridge University Press, New York, pp 72–96
- Gomez A, Vendramin GG, Gonzalez-Martinez S, Alia R (2005) Genetic diversity and differentiation of two Mediterranean pines (*Pinus halepensis* Mill. and *Pinus pinaster* Ait.) along a latitudinal cline using chloroplast microsatellite markers. *Divers Distrib* 11:257–263
- Griffin JR, Critchfield WB (1972) The distribution of forest trees in California. USDA Forest Research Paper PSW-82/1972
- Grivet D, Smouse PE, Sork VL (2005) A novel approach to an old problem: tracking dispersed seeds. *Mol Ecol* 14:3585–3595
- Grivet DM, Deguilloux MF, Petit JR, Sork VL (2006) Contrasting patterns of historical colonization in white oaks (*Quercus* spp.) in California and Europe. *Mol Ecol* 15:4085–4093
- Grivet DM, Sork VL, Westfall RD, Davis FW (2008) Conserving the evolutionary potential of California valley oak (*Quercus lobata* Née): a multivariate genetic approach to conservation planning. *Mol Ecol* 17:139–156
- Gugger PF, Ikegami M, Sork VL (2013) Influence of late Quaternary climate change on present patterns of genetic variation in valley oak, *Quercus lobata* Née. *Mol Ecol* 22:3598–3612
- Gugger PF, Cokus SJ, Sork VL (2016) Association of transcriptome-wide sequence variation with climate gradients in valley oak (*Quercus lobata*). *Tree Genet Genomes* 12:15
- Holderegger R, Wagner HH (2008) Landscape genetics. *BioScience* 58:199–207
- Holderegger R, Buehler D, Gugerli F, Manel S (2010) Landscape genetics of plants. *Trends Plant Sci* 15:675–683
- Hosley D, Hoile L, Weikel B (2012) Assessing the region via indicators 2006–2011: The State of the Great Central Valley. <http://www.greatvalley.org/wp-content/uploads/2012/01/environmental-indicators-2011.pdf>. Accessed 7 Sep 2016
- Howard JL (1992) *Quercus lobata* in Fire Effects Information System. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory (Producer). <http://www.fs.fed.us/database/feis/>
- Jockusch EL, Yanev KP, Wake DB (2001) Molecular phylogenetic analysis of slender salamanders, genus *Batrachoseps* (Amphibia: Plethodontidae), from central coastal California with descriptions of four new species. *Herpetol Monogr* 15:54–99
- Jost L (2008)  $G_{ST}$  and its relatives do not measure differentiation. *Mol Ecol* 17:4015–4026
- Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA (2013) diveRsim: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol Evol* 4:782–788
- Kelly PA, Phillips SE, Williams DF (2005) Documenting ecological change in time and space: the San Joaquin Valley of California. In: Patton JL, Lacey EA, Myers P (eds) *Mammalian diversification: from chromosomes to phylogeography*. University of California Press, Berkeley, pp 57–78
- Lancaster LT, Kay KM (2013) Origin and diversification of the California flora: re-examining classic hypotheses with molecular phylogenies. *Evolution* 67:1041–1054
- Landguth EL, Fedy BC, Oyler-McCance SJ, Garey AL, Emel SL, Mumma M et al (2012) Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Mol Ecol Resour* 12:276–284
- Lapointe FJ, Rissler LJ (2005) Congruence, consensus, and the comparative phylogeography of codistributed species in California. *Am Nat* 166:290–299
- Legendre P, Fortin MJ (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Mol Ecol Resour* 10:831–844
- Legendre P, Troussellier M (1988) Aquatic heterotrophic bacteria: modeling in the presence of spatial autocorrelation. *Limnol Oceanogr* 33:1055–1067
- Lenihan JM, Drapek R, Bachelet D, Neilson RP (2003) Climate change effects on vegetation distribution, carbon, and fire in California. *Ecol Appl* 13:1667–1681
- Maldonado JE, Vila C, Wayne RK (2001) Tripartite genetic subdivisions in the ornate shrew (*Sorex ornatus*). *Mol Ecol* 10:127–147
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27:209–220
- McRae BH, Shah V, Mohapatra T (2013) Circuitscape 4 user guide. The nature conservancy. <http://www.circuitscape.org>
- Munz PA (1973) *A California flora and supplement*. University of California Press, Berkeley
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Nat Acad Sci USA* 70:3321–3323
- Neigel JE (2002) Is  $F_{ST}$  obsolete? *Conserv Genet* 3:167–173
- Ortego J, Riordan EC, Gugger PF, Sork VL (2012) Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. *Mol Ecol* 21:3210–3223
- Pavlik BM, Muick P, Johnson S (1991) *Oaks of California*. Cachuma Press and California Oaks Foundation, Los Olivos

- Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539
- Pesendorfer MB, Sillett TS, Koenig WD, Morrison SA (2016) Scatter-hoarding corvids as seed dispersers for oaks and pines: a review of a widely distributed mutualism and its utility to habitat restoration. *Condor: Ornithological Applications* 118:215–237
- Pluess AR, Sork VL, Dolan B, Davis FW, Grivet D, Merg K, Papp J, Smouse PE (2009) Short distance pollen movement in a wind-pollinated tree, *Quercus lobata* (Fagaceae). *Forest Ecol Manag* 258:735–744
- R Development Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <http://www.R-project.org>
- Raven PH, Axelrod DI (1978) Origin and relationships of the California flora. *Univ Calif Publ Bot* 72:1–134
- Rhodes MK, Fant JB, Skogen KA (2014) Local topography shapes fine-scale spatial genetic structure in the Arkansas valley evening primrose, *Oenothera harringtonii* (Onagraceae). *J Heredity* 105:806–815
- Rissler LJ, Hijmans RJ, Graham CH, Moritz C, Wake DB (2006) Phylogeographic lineages and species comparisons in conservation analyses: a case study of California herpetofauna. *Am Nat* 167:655–666
- Schwartz MK, Copeland JP, Anderson NJ, Squires JR, Inman RM, McKelvey KS, Pilgrim KL, Waits LP, Cushman SA (2009) Wolverine gene flow across a narrow climatic niche. *Ecology* 90:3222–3232
- Sezen UU, Chazdon RL, Holsinger KE (2005) Genetic consequences of tropical second-growth forest regeneration. *Science* 307:891–894
- Shirk AJ, Wallin DO, Cushman SA, Rice CG, Warheit KI (2010) Inferring landscape effects on gene flow: a new multi-scale model selection framework. *Mol Ecol* 19:3603–3619
- Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Zool* 35:627–632
- Sork VL, Davis FW, Westfall R, Flint A, Ikegami M, Wang H, Grivet D (2010) Gene movement and genetic association with regional climate gradients in California valley oak (*Quercus lobata* Née) in the face of climate change. *Mol Ecol* 19:3806–3823
- Sork VL, Squire K, Gugger PF, Steele SE, Levy ED, Eckert AJ (2016) Landscape genomic analysis of candidate genes for climate adaptation in a California endemic oak, *Quercus lobata*. *Am J Bot* 103:33–46
- Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010) Landscape genetics: where are we now? *Mol Ecol* 19:3496–3514
- Thomassen HA, Cheviron ZA, Freedman AH, Harrigan RJ, Wayne RK, Smith TB (2010) Spatial modelling and landscape-level approaches for visualizing intra-specific variation. *Mol Ecol* 19:3532–3548
- USDA Forest Service (1981) CALVEG: a classification of California vegetation. Pacific Southwest Region, Regional Ecology Group, San Francisco, p 168
- Van Strien MJ, Keller D, Holderegger R (2012) A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Mol Ecol* 21:4010–4023
- Wake DB, Yanev KP, Brown CW (1986) Intraspecific sympatry in a “ring species,” the plethodontid salamander *Ensatina eschscholtzii*, in southern California. *Evolution* 40:866–868
- Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J Heredity* 89:438–450
- Whipple AA, Grossinger RM, Davis FW (2011) Shifting baselines in a California oak savanna: nineteenth century data to inform restoration scenarios. *Restor Ecol* 19:88–101
- Whiteside C (2007) Challenges and choices for California’s great Central Valley. *Berkeley J Afr Am Law Policy* 9:160–165
- Whitlock MC (2011)  $G'_{ST}$  and  $D$  do not replace  $F_{ST}$ . *Mol Ecol* 20:1083–1091
- Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163:1177–1191
- Wright S (1943) Isolation by distance. *Genetics* 28:114–138
- Yang RC (2004) A likelihood-based approach to estimating and testing for isolation by distance. *Evolution* 58:1839–1845
- Zavaleta ES, Hulvey KB, Fulfrost B (2007) Regional patterns of recruitment success and failure in two endemic California oaks. *Divers Distrib* 13:735–745