Impacts of human-induced environmental disturbances on hybridization between two ecologically differentiated Californian oak species

Joaquín Ortego[1], Paul F. Gugger[2] and Victoria L. Sork[3,4]

[1]Department of Integrative Ecology, Estación Biológica de Doñana, EBD-CSIC, Avda. Américo Vespucio s/n, E-41092 Seville, Spain; [2]University of Maryland Center for Environmental Science, Appalachian Laboratory, 301 Braddock Road, Frostburg, MD 21532, USA; [3]Department of Ecology and Evolutionary Biology, University of California, Box 957239, Los Angeles, CA 90095, USA; [4]Institute of the Environment and Sustainability, University of California, Box 951496, Los Angeles, CA 90095-1496, USA

Author for correspondence:
Joaquín Ortego
Estación Biológica de Doñana, EBD-CSIC,
Avda. Américo Vespucio s/n, E-41092 Seville, Spain
E-mail: joaquin.ortego@csic.es
Phone: +34 954 232 340

Running title: Ecological drivers of hybridization in oaks

Summary

- Natural hybridization, which can be involved in local adaptation and in speciation processes, has been linked to different sources of anthropogenic disturbance.
- Here, we use genotypic data to study range-wide patterns of genetic admixture between the serpentine-soil specialist leather oak (*Quercus durata*) and the widespread Californian scrub oak (*Q. berberidifolia*). First, we estimated hybridization rates and the direction of gene flow. Second, we tested the hypothesis that genetic admixture increases with different sources of environmental disturbance, namely anthropogenic destruction of natural habitats and wildfire frequency estimated from long-term records of fire occurrence.
- Our analyses indicate considerable rates of hybridization (>25%), asymmetric gene flow from *Q. durata* into *Q. berberidifolia*, and a higher occurrence of hybrids in areas where both species live in close parapatry. In accordance with the environmental disturbance hypothesis, we found that genetic admixture increases with wildfire frequency, but we did not find a significant effect of other sources of human-induced habitat alteration (urbanization, land clearing for agriculture) or a suite of ecological factors (climate, elevation, soil type).
- Our findings highlight that wildfires constitute an important source of environmental disturbance promoting hybridization between two ecologically well-differentiated native species.

Key words: environmental disturbance, California, genetic structure, genetic admixture, hybridization, introgression, *Quercus*, wildfire.

Introduction

Hybridization has attracted the attention of biologists for a long time (Linnaeus, 1735; Darwin, 1859; Anderson & Stebbins, 1954) and is widely recognized as an important evolutionary force involved in adaptation to novel environmental conditions (Lewontin & Birch, 1966; Morjan & Rieseberg, 2004; Baskett & Gomulkiewicz, 2011) and speciation (Anderson, 1948; Rieseberg, 1997; Abbott *et al*., 2013). Despite the potential benefits of hybridization to biodiversity, this phenomenon has also been linked to the disruption of local adaptation and species loss (Rhymer & Simberloff, 1996). This is of particular concern when exotic taxa are involved, as hybridization can potentially increase their invasiveness and lead to the extinction of native species (Ellstrand & Schierenbeck, 2000; Prentis *et al*., 2007). Beyond introductions of exotic species by humans, different forms of anthropogenic disturbance have also been hypothesized to increase hybridization rates among native, exotic and native-exotic taxa (Crispo *et al*., 2011; Guo, 2014). The disturbance hypothesis postulates that habitat alterations increase opportunities for hybridization via the breakdown of premating reproductive isolating mechanisms and/or creating environmental gradients with novel or intermediate niches where hybrids outperform parental species (Anderson & Stebbins, 1954; see also Anderson, 1948, 1953). Anthropogenic disturbances that have been linked with increased hybridization rates include land use alterations (Lamont *et al*., 2003; Thompson *et al*., 2010; Hoban *et al*., 2012; Guo, 2014), climate change (Muhlfeld *et al*., 2014) and, more counterintuitively, the disruption of natural disturbances that promote reproductive isolation and maintain species boundaries (e.g. suppression of natural wildfires; King *et al*., 2015; Stewart *et al*., 2015). Support for this hypothesis has been found across multiple taxa ("hybrid richness") at a continental scale (Guo, 2014)

and within pairs of interbreeding species at local (Hasselman *et al*., 2014), regional (Thompson *et al*., 2010; Muhlfeld *et al*., 2014) and range-wide scales (Hoban *et al*., 2012).

Quercus (oak) is a classic example of a genus with many highly hybridizing species that maintain their taxonomic and ecological identity in the presence of frequent interspecific gene flow (Muller, 1952; Whittemore & Schaal, 1991; Rushton, 1993). Species relative abundance and density (Lepais *et al*., 2009; Lagache *et al*., 2013), environment (e.g. Muller, 1952; Anderson & Stebbins, 1954; Williams *et al*., 2001; Ortego *et al*., 2014a) and population history (Zeng *et al*., 2011) have been found to be important explanatory factors of spatial patterns of hybridization in oaks, but less attention has been paid to the potential role of environmental disturbance. On the basis of morphological characters, Silliman & Leisner (1958) found evidence for higher hybridization rates in a stand showing signs of successive disturbances by fire and lumbering operations in comparison with a mixed oak forest established in a stable and undisturbed environment. At a local scale, Lagache *et al*. (2013) found that reduced conspecific density, likely resulting from environmental disturbance, increases hybridization rates through decreased intensity of pollen competition. However, despite considerable research on hybridization that has been performed on oaks, large-scale studies comparing rates of genetic admixture across multiple populations subjected to different sources of environmental disturbance are lacking (Rushton, 1993).

Here, we study the ecological drivers of range-wide patterns of genetic admixture between the widespread California scrub oak (*Quercus berberidifolia*) and the serpentine-soil specialist leather oak (*Quercus durata*), two Californian endemic sister taxa (Ortego *et al*., 2015a) with partly overlapping distributions and for which previous morphology- and molecular-based studies have reported a frequent occurrence

of interspecific hybrids (Forde & Faris, 1962; eFloras, 2015; Ortego *et al*., 2015a). A previous study on the Californian scrub white oak species complex showed that these two species probably diverged in peripatry or sympatry ~23-26 k years BP, the oldest split among the three pairs of sister taxa within the complex, and supported the monophyletic origin of *Q. durata* (Ortego *et al*., 2015a). In this study, we primarily aim to investigate the impacts of different sources of environmental disturbance on hybridization rates between these two shrub species (Anderson, 1948; Anderson & Stebbins, 1954), with particular emphasis on the potential role of human-induced landscape alterations (e.g. Thompson *et al*., 2010; Guo, 2014). California offers an excellent setting to address this question. It is a climatically and geologically complex region, a fact that has been linked with high rates of interspecific hybridization (Dodd & Afzal-Rafii, 2004; Guo, 2014; Ortego *et al*., 2014a), speciation (Calsbeek *et al*., 2003; Lancaster & Kay, 2013) and local adaptation processes across many organisms and spatiotemporal scales (Ortego *et al*., 2012; Langin *et al*., 2015). Much of this region is also highly impacted by urbanization and anthropogenic habitat degradation, which has altered population connectivity in many organisms (e.g. Riley *et al*., 2006; Vandergast *et al*., 2007) and increased risk of extinction in several species (Myers *et al*., 2000; Schwartz *et al*., 2006; Vandergast *et al*., 2008). Beyond urbanization and agriculture, wildfires are also an important source of environmental disturbance in Californian ecosystems and their occurrence and impact have steadily increased since European settlement associated with growing population densities, urbanization, and human-induced climate change (Westerling *et al*., 2006; Syphard *et al*., 2007; Moritz *et al*., 2014). Despite the fact that natural wildfire regimes are an important component of Californian ecosystems (Rundel, 1982; Keeley *et al*., 2012), their increased frequency associated with human activities has been linked with population declines in several

124 species already impacted by other sources of habitat loss and degradation (Syphard *et*
125 *al*., 2007; Barr *et al*., 2015). These different sources of environmental disturbance,
126 together with the high richness of native and exotic species in the region (Calsbeek *et*
127 *al*., 2003; Lancaster & Kay, 2013; Guo, 2014), have been suggested to underlie the high
128 proportion of hybrids found in California in comparison with most other states of the
129 country (Guo, 2014).

130 The overall goal of this study is to investigate the extent to which wildfire
131 frequency, human-induced habitat transformation, and local environmental factors
132 influence hybridization between *Q. berberidifolia* and *Q. durata*. Using genotypic data
133 for 58 stands (> 400 individuals) sampled across California we tested two specific
134 hypotheses. First, we estimated hybridization rates and direction of gene flow between
135 the two focal species and (i) tested whether the adaptation of *Q. durata* to serpentine
136 soils results in asymmetric gene flow from this species into *Q. berberidifolia*, which
137 would be expected if hybrids show a lower performance in serpentine than in non-
138 serpentine soils due to a higher niche breath of serpentine-adapted plants (i.e. more
139 tolerant to different soil types; Whittaker, 1954; Kruckeberg, 1984; Wright, 2007).
140 Then, we studied the potential role of different ecological factors on spatial patterns of
141 hybridization, primarily focusing on analyzing the impacts of human-induced
142 environmental disturbance (Anderson, 1948; Anderson & Stebbins, 1954; Arnold,
143 1997). In particular, we tested the hypothesis that (ii) hybridization increases with two
144 sources of environmental disturbance, namely anthropogenic large-scale destruction of
145 natural habitats (urbanization and agriculture) and wildfire frequency. We addressed
146 this main question controlling for other potential drivers of hybridization such as
147 species co-occurrence (range overlap) and environment (climate, elevation, soil type)
148 that have been previously reported to influence interspecific gene flow in oaks (e.g.

149 Williams *et al*., 2001; Dodd & Afzal-Rafii, 2004; Alberto *et al*., 2010; Ortego *et al*.,
150 2014a). We analyzed whether the contribution of these factors to explain population-
151 level genetic admixture varies with the spatial scale at which they are measured, an
152 issue that has been seldom addressed in the context of hybridization studies, despite the
153 possibility that it can have a non-negligible impact on the inferences obtained (Barton &
154 Hewitt, 1985; Harrison, 1986; Barton, 2001; Buggs, 2007), particularly in wind-
155 pollinated species with large scales of dispersal (e.g. Dow & Ashley, 1998; Buschbom
156 *et al*., 2011).
157
158 **Materials and Methods**
159
160 Population sampling
161
162 Between 2010 and 2014, we sampled 529 reproductive individuals (i.e. we did not
163 sample seedlings or saplings) from a total of 58 localities in California (Fig. 1;
164 Supporting Information Table S1). Plants were morphologically identified in the field
165 and tentatively assigned to parental taxa or hybrids according with available guides and
166 floras (Roberts, 1995; eFloras, 2015). We aimed to collect samples from populations
167 located across the entire distribution range of *Q. berberidifolia* Liebmann and *Q. durata*
168 Jepson and designed sampling using occurrence records available in the Calflora
169 database (http://www.calflora.org/). Some sampled individuals (n = 78) were
170 genetically identified as hybrids with other taxa within the scrub white oak species
171 complex (Ortego *et al*., 2015a) and were excluded from subsequent analyses. Thus, all
172 analyses presented in this study are based on 451 individuals of *Q. berberidifolia*, *Q.*
173 *durata* or hybrids between them (Table S1).

174

175 Microsatellite genotyping

176

177 We genotyped samples of *Q. berberidifolia* and *Q. durata* using 16 nuclear
178 microsatellite markers previously developed for other species (Table S2). DNA
179 extraction and microsatellite amplification and genotyping were performed as described
180 in Ortego *et al*. (2015a). Microsatellite data are available in the DRYAD Digital
181 Repository (doi: 10.5061/dryad.52504).

182

183 Genetic structure, hybrid identification and admixture analyses

184

185 We identified hybrid and purebred individuals in our empirical dataset using the
186 Bayesian Markov chain Monte Carlo clustering analyses implemented in the programs
187 STRUCTURE 2.3.3 (Pritchard *et al*., 2000; Falush *et al*., 2003; Hubisz *et al*., 2009) and
188 NEWHYBRIDS 1.1 (Anderson & Thompson, 2002). In STRUCTURE, the posterior
189 probability (*q*) describes the proportion of an individual genotype originating from each
190 of *K* clusters. In NEWHYBRIDS, *q* describes the probability that an individual belongs to
191 each of six different genotypic groups that include two parental species and four hybrid
192 classes (F1, F2 and first generation backcrosses). The assignment to a specific hybrid
193 class is often uncertain in NEWHYBRIDS (Burgarella *et al*., 2009). Thus, as done in
194 previous studies, we summed *q* values over all hybrid genotype frequency classes (e.g.
195 Cullingham *et al*., 2011; Hasselman *et al*., 2014; Haines *et al*., 2016). Given that
196 NEWHYBRIDS can only accommodate two species/clusters and our analyses support
197 genetic substructure within *Q. berberidifolia* (see Results), we conservatively based
198 subsequent analyses on estimates of admixed ancestry obtained from STRUCTURE (e.g.

199 Haines *et al*., 2016). Details of STRUCTURE and NEWHYBRIDS settings are presented in
200 Notes S1.

201 Complementary to Bayesian clustering analyses, we performed an individual-
202 based principle components analysis (PCA) using the R 3.0.3 (R Core Team, 2013)
203 package ADEGENET (Jombart, 2008). This analysis does not rely on Hardy-Weinberg or
204 linkage equilibrium and it has often been shown to be useful to complement the results
205 of clustering analyses in studies of hybridization (e.g. Saarman & Pogson, 2015). Then,
206 we employed a MANOVA to compare the PCA scores obtained for the first two principal
207 component (PC) axes among the two parental species and hybrids identified on the basis
208 of STRUCTURE and NEWHYBRIDS analyses. *Post hoc* Tukey tests were used to examine
209 differences between parental species and between parental species and their hybrids.
210 MANOVA were performed using SPSS 22.0.

211 We calculated the level of genetic differentiation (F_{ST}) between species in FSTAT
212 2.9.3 (Goudet, 1995) considering pure individuals from either parental species identified
213 by STRUCTURE and NEWHYBRIDS analyses. Confidence intervals (95% CI) were
214 estimated by bootstrapping over loci (10 000 randomizations).

215

216 Hybrid simulation and genetic assignment

217

218 We used simulations to determine the accuracy, efficiency and overall performance
219 (Vähä & Primmer, 2006) of our set of markers for identifying hybrid and purebred
220 individuals on the basis of the probabilities of membership inferred from STRUCTURE
221 and NEWHYBRIDS analyses and the often used threshold of $Q \geq 0.90$ (e.g. Cullingham *et*
222 *al*., 2011; Hasselman *et al*., 2014). Further details of our approach are described in
223 Notes S2.

224

225 Bayesian comparison of gene flow models

226

227 We used MIGRATE-N 3.6.11 to test different scenarios of gene flow between our two
228 focal species (Beerli & Felsenstein, 2001; Beerli, 2006; Beerli & Palczewski, 2010).
229 This program estimates mutation-scaled effective population size ($\theta = 4N_e\mu$, where N_e =
230 effective population size and μ = mutation rate per generation) and migration rates (M =
231 m/μ, where m = migration rate) for multiple populations in a coalescent framework by
232 which alleles are traced back in time to a single ancestral copy (the most recent common
233 ancestor, MRCA) (Beerli, 2009). STRUCTURE analyses on our empirical database
234 revealed the presence of three genetic groups, one corresponding to *Q. durata* and the
235 other two corresponding to *Q. berberidifolia* (see Results section for more details).
236 Although the two genetic clusters identified within *Q. berberidifolia* have a high degree
237 of spatial genetic admixture, they roughly separate populations located in southern
238 California (south of the Transverse Ranges) and parapatric populations with *Q. durata*
239 in the north (see results section). Thus, we used MIGRATE-N to test six models that
240 considered different patterns of gene flow among three population groups defined *a*
241 *priori*, namely populations of *Q. durata* and the two clusters of *Q. berberidifolia*
242 identified by STRUCTURE analyses (Fig. S1a). The two population groups within *Q.*
243 *berberidifolia* were defined according to (i) the probabilities of genetic membership
244 inferred by STRUCTURE analyses (see Results section) and (ii) considering whether the
245 populations were located or not in areas overlapping with the distribution range of *Q.*
246 *durata* (i.e. north and south of the Transverse Ranges, respectively). Further, we tested
247 two different sets of models: one only including purebred individuals for each taxon and
248 cluster according with STRUCTURE analyses ($Q \geq 0.90$; see Results section for more

249 details) and another considering both purebred and hybrids individuals (e.g. Field *et al*.,
250 2011; Andrew *et al*., 2012; Starr *et al*., 2013). Hybrids were assigned to each of the
251 three groups according to majority population genetic assignment to each group ($Q >$
252 0.5) according to STRUCTURE analyses. Details of MIGRATE-N settings are presented in
253 Notes S3.
254
255 GIS analyses
256
257 We obtained information from wildfire frequency using the CALFIRE Fire Perimeters
258 Geodatabase version 13.2 (http://frap.cdf.ca.gov/). Briefly, we transformed the vector
259 layer containing the polygons delimiting wildfires perimeters for each year from 1900
260 to 2013 (114 years) into one raster layer per year in which burned areas (~50 m^2 pixels)
261 were given a pixel value equal to one. Then, we summed all year-based raster layers in
262 order to generate a new raster containing information on the number of years that each
263 pixel has been affected by wildfires. Finally, we calculated average wildfire frequency
264 within a circular area of 10, 100 and 1000 km^2 around each sampling locality, which
265 allowed us to assess the potential impact of spatial scale on our results (see next
266 section).
267 We estimated the proportion of habitats disturbed by agriculture and human
268 development from the Conterminous United States Land Cover 200 m resolution layer
269 (http://nationalmap.gov/small_scale/atlasftp.html). We considered as disturbed areas
270 those devoted to different forms of agriculture (categories 61, 71, 81, 82, 83 and 84),
271 urbanization (categories 21, 22, 23, and 85) and mining (category 32) (see Land Cover
272 layer legend for the description of the different categories). The proportion of disturbed
273 areas was calculated for the three same spatial scales considered for wildfire frequency.

Climate and elevation data were obtained from the WorldClim 1.4 dataset (http://www.worldclim.org/) (Hijmans *et al*., 2005). We downloaded the 19 climatic and elevation layers at a 30-arcsec resolution (c. 1-km) and extracted average values for each of them at the same spatial scales considered for wildfire frequency. We performed a PCA on the 19 climatic variables and retained for subsequent analyses the first principal component (PC1), which explained a large proportion of the variance at all the spatial scales considered (>93 % in all cases).

We obtained soil data for each sampling locality from the SSURGO datasets available at the Web Soil Survey from the United States Department of Agriculture (USDA) (http://websoilsurvey.sc.egov.usda.gov/; Staff, 2012). In this case, we only obtained soil order data (based on USDA Soil Taxonomy categories) for each sampling locality (i.e. this parameter was not estimated at multiple spatial scales) given that the successful establishment of seedlings is only expected to be influenced by soil properties a few meters around the germination site (e.g. Wright, 2007; Langhans *et al*., 2009). All GIS calculations were performed in ARCMAP 10.2.1 (ESRI, Redlands, CA, USA).

Analyses of genetic admixture

We estimated the degree of genetic admixture of the studied populations using the 'genetic admixture index' (G_{Admix}), calculated as described in Ortego *et al*. (2015b). G_{Admix} ranges from 0 to 1, with values equal to 0 indicating no admixture (i.e. genetically pure populations assigned to a single genetic cluster) and values equal to 1 indicating maximum admixture (i.e. genetically admixed populations with an equal probability of membership to each inferred genetic cluster). Thus, this summary statistic

299 provides information on within-population genetic admixture that can be directly
300 compared with different population characteristics (Ortego *et al*., 2015b). We used an
301 information-theoretic model selection approach to analyse which variables contribute to
302 explain patterns of G_{Admix} in the studied populations. We considered four covariates
303 (climatic conditions, PC1; elevation; proportion of habitats disturbed by agriculture and
304 human development; wildfire frequency) and two fixed factors (species range overlap;
305 soil type). Species range overlap was defined as a categorical variable, which
306 considered whether the studied populations were located (=1) or not (=0) in areas where
307 the distribution ranges of *Q. durata* and *Q. berberidifolia* overlap. Ranges for both
308 species were defined according with known records of the species obtained from
309 Calflora database (http://www.calflora.org/) and verified with distribution maps from
310 eFloras (2015). Given that the distribution of *Q. durata* in entirely embedded within the
311 range of *Q. berberidifolia* (Fig. 1), we just considered areas of range overlap as those
312 defined by the distribution of the former (eFloras, 2015). It should be noted that our
313 study does not aim to link contemporary hybridization with specific disturbance events,
314 as this approach would require genetic information from cohorts (e.g. seedlings or
315 saplings) established after the disturbance took place and, ideally, comparisons with
316 individuals collected from nearby non-disturbed areas (e.g. Stewart *et al*., 2015).
317 Instead, our large-scale study covering entire species ranges aims to retrieve information
318 from populations located in areas experiencing contrasting environmental disturbances
319 (e.g. with different wildfire regimes; Fig. 1) in order to link this information with their
320 past hybridization history reflected in their overall degree of genetic admixture (e.g.
321 Ortego *et al*., 2014a).

322 We analysed the data using General Linear Models (GLM) with a Gaussian error
323 structure and identity link function as implemented in the R 3.0.0 package LME4 (R

Core Team, 2012). The precision of G_{Admix} estimates may differ among populations due to differences in sample sizes and we took this into account using a weighted least square method, where weight equals the sample size for each studied population (Table S1). Model selection and averaging were performed using the R package AICCMODAVG (R Core Team, 2012) as detailed in Ortego *et al*. (2015c). We ran three subsets of models considering in each one the same variables but measured at the three different spatial scales described in the previous section. Complementarily, we also built three similar models in which we replaced species range overlap with latitude and longitude (fitted as covariates) to evaluate the potential impact of the spatial location of the study populations on our results. Note that species range overlap is highly associated with latitude ($F_{1, 56} = 50.53$, $P < 0.001$) and longitude ($F_{1, 56} = 89.50$, $P < 0.001$), so these three variables were not simultaneously fitted in the same models in order to avoid strong multicollinearity problems.

Results

Hybrid identification and genetic structure

Considering the dataset simulated on the basis of purebred individuals identified by STRUCTURE analyses, the assignment of purebred and hybrid individuals to their correct class was 93% for NEWHYBRIDS and 96% for STRUCTURE (Fig. S2a, b and Table S3a). Similarly, considering the dataset simulated on the basis of purebred individuals identified by NEWHYBRIDS analyses, the assignment of purebred and hybrid individuals to their correct class was 97% for NEWHYBRIDS and 98% for STRUCTURE (Fig. S2c, d and Table S3b). As found in previous studies, NEWHYBRIDS tended to detect a higher

number of hybrids than STRUCTURE (e.g. Haines *et al*., 2016) (Fig. S2). Overall, the performance of our set of markers to identify hybrids is similar to that reported in other microsatellite-based studies on trees (e.g. Lepais *et al*., 2009; Cullingham *et al*., 2012) and our simulations support that a threshold value of $Q = 0.9$ allows differentiating correctly purebred individuals from hybrids with a high confidence (Table S3). Thus, individuals with $Q \geq 0.90$ for either parental species were considered as purebred genotypes and all other individuals were classified as hybrids or introgressed genotypes (e.g. Vähä & Primmer, 2006; Lepais *et al*., 2009; Cullingham *et al*., 2011, 2012).

Log probabilities [Pr(X|*K*)] of STRUCTURE analyses on our empirical data sharply increased from $K = 1$ to $K = 2$ and reached a plateau at $K = 3$ (Fig. 2c). The statistic ΔK indicated an "optimal" value of $K = 2$ (Fig. 2c), which roughly grouped *Q. berberidifolia* and *Q. durata* in different genetic clusters (Fig. 2a). Log probabilities were significantly higher for any $K > 1$ than for $K = 1$ (Wilcoxon rank-sum tests, $P < 0.001$), rejecting the possibility of a single panmictic population (Fig. 2c). Analyses for $K = 2$ showed a very high degree of genetic admixture between both species (Fig. 2a). As a result, 210 individuals (47%) were classified as hybrids ($Q < 0.90$) and many of them belonged to populations from areas where the distribution ranges of both species do not overlap (60 hybrids, 39% of individuals from allopatric areas). Analyses for $K = 3$ showed a clearer separation between both species, with one cluster including individuals of *Q. durata* and the two other clusters reflecting a latitudinal cline of genetic differentiation within *Q. berberidifolia* (Ortego *et al*., 2015a). A much lower number of individuals were identified as hybrids for $K = 3$ ($n = 118$, 26%) and only a few were collected in areas of allopatry ($n = 18$; 12 % of individuals from allopatric areas). Thus, the number of hybrid individuals was much higher for $K = 2$ than for $K = 3$ and these differences were highly significant considering all populations ($\chi^2 = 40.55$; P

374 < 0.001), populations located in areas where the range of both species overlap (χ^2 =
375 17.31; $P < 0.001$) or populations from allopatric areas ($\chi^2 = 30.22$; $P < 0.001$). For K =
376 2, a considerable number of individuals morphologically identified in the field as *Q.*
377 *berberidifolia* were assigned to purebred *Q. durata* ($Q \geq 0.90$; $n = 19$) or had a much
378 higher probability of membership to *Q. durata* cluster than to *Q. berberidifolia* cluster
379 ($Q \geq 0.8$; $n = 41$). For $K = 3$, the number of misclassified individuals was much lower:
380 one individual identified in the field as *Q. berberidifolia* was assigned to purebred *Q.*
381 *durata* and two individuals morphological identified as *Q. durata* were genetically
382 assigned to purebred *Q. berberidifolia*. Accordingly, only a few individuals
383 morphologically identified as *Q. durata* ($n = 3$) and *Q. berberidifolia* ($n = 4$) had a high
384 probability of genetic membership ($Q \geq 0.8$) to the other species. Thus, the number of
385 misclassified individuals was much higher for $K = 2$ than for $K = 3$, either considering Q
386 ≥ 0.90 ($\chi^2 = 11.93$; $P < 0.001$) or $Q \geq 0.8$ ($\chi^2 = 25.43$; $P < 0.001$). For these reasons, K =
387 3 was regarded as a biologically more meaningful clustering solution than $K = 2$ and
388 considered for subsequent analyses.

389 In accordance with STRUCTURE analyses for $K = 2$, NEWHYBRIDS identified 188
390 hybrids (42%) and many of them belonged to areas where the distribution ranges of
391 both species do not overlap (29 hybrids, 19% of individuals from allopatric areas) (Fig.
392 2b). The number of hybrid individuals identified by NEWHYBRIDS was higher than those
393 identified by STRUCTURE analyses for $K = 3$ considering all populations ($\chi2 = 24.23$; P
394 < 0.001), populations located in areas where the range of both species overlap ($\chi2$ =
395 23.89; $P < 0.001$) or populations from allopatric areas ($\chi2 = 3.03$; $P = 0.08$). It is
396 remarkable the much lower number of purebred individuals of *Q. durata* identified by
397 NEWHYBRIDS ($n = 18$) in comparison with STRUCTURE analyses for both $K = 2$ ($n = 100$)
398 ($\chi2 = 28.26$; $P < 0.001$) and $K = 3$ ($n = 57$) ($\chi2 = 74.26$; $P < 0.001$).

Principal components analyses also supported the separation between the two parental species and their hybrids identified by STRUCTURE (MANOVA, PC1: $F_{2, 448}$ = 434.33, $P < 0.001$; PC2: $F_{2, 448} = 7.18$, $P = 0.001$; Fig. 3a) and NEWHYBRIDS analyses (MANOVA, PC1: $F_{2, 448} = 293.22$, $P < 0.001$; PC1: $F_{2, 448} = 8.69$, $P < 0.001$; Fig. 3b). There was no overlap along PC1 among purebred individuals of the two parental species identified by either STRUCTURE or NEWHYBRIDS analyses (Fig 3a, b). *Post hoc* Tukey tests showed that the only non-significant pairwise comparisons were those between *Q. berberidifolia* and *Q. durata* and between *Q. durata* and hybrid individuals for PC2 ($P > 0.05$). The two genetic clusters within *Q. berberidifolia* identified by STRUCTURE analyses for $K = 3$ were also well separated along PC1 of a PCA only including purebred individuals ($Q \geq 0.9$) of this species (MANOVA, PC1: $F_{2, 274} = 106.04$, $P < 0.001$; PC2: $F_{2, 274} = 1.96$, $P = 0.142$; Fig. 3c). *Post hoc* Tukey tests showed that all pair-wise comparisons for PC1 were highly significant (all Ps < 0.001), supporting the separation among individuals with a high probability of assignment to any of the two genetic clusters within *Q. berberidifolia* ($Q \geq 0.9$) and those showing admixed ancestry ($Q < 0.9$).

The degree of genetic differentiation between the two parental species estimated on the basis of purebred individuals ($Q \geq 0.9$) identified by STRUCTURE ($F_{ST} = 0.041$, 95% CI: 0.030-0.053) and NEWHYBRIDS ($F_{ST} = 0.062$, 95% CI: 0.040-0.084) analyses was similar to that reported among other species within the scrub white oak species complex (Ortego *et al*., 2015a) and within the same order of magnitude previously reported for other hybridizing species (Cullingham *et al*., 2011; Haines *et al*., 2016).

Bayesian comparison of gene flow models

The scenario considering unidirectional gene flow from *Q. durata* into both parapatric and allopatric populations of *Q. berberidifolia* was the most supported in analyses both including and excluding hybrids individuals from the dataset (Table 1; Fig. S1b). The second best ranked model was the same but exclusively considering gene flow from *Q. durata* into populations of *Q. berberidifolia* located in areas where the distribution range of both species overlap (Table 1). MIGRATE-N analyses failed to converge when hybrids individuals were included, but model choice was consistent across replicated runs (data not shown). This indicates that we can be confident in model choice but parameter estimates for models including hybrids must be interpreted with caution (Fig. S1b). Convergence issues in models including hybrid individuals may be related with the fact that different runs provided a good fit of our data for contrasting combinations of estimates of effective population sizes and migration rates (see Beerli, 2006, 2009).

Factors associated with genetic admixture

Only wildfire frequency and range overlap were consistently included in models of genetic admixture (G_{Admix}) for all spatial scales (Table 2 and Table S4). Genetic admixture tended to be higher in populations located in areas where the distributions ranges of both species overlap, but these differences only were marginally significant (i.e. unconditional CIs crossed zero; Table 3). Accordingly, the frequency of hybrids ($Q < 0.90$) was significantly lower in areas of allopatry than in the vast region where the ranges of both species overlap ($\chi^2 = 25.88$, $P < 0.001$). Genetic admixture was positively associated with wildfire frequency at all spatial scales (Table 3 and Table S5; Fig. 4). Although wildfire frequency always had a significant effect, its effect size was higher at the two largest spatial scales (Table 3 and Table S5). All other variables were

449 not included in any model (climate) or were included in some of them but showed no
450 significant effects (elevation, proportion of habitats disturbed by agriculture and human
451 development, soil category, latitude, longitude) (Tables 2-3 and Supporting Information
452 Tables S4-5). Analyses based on soil suborder category (rather than order category)
453 provided analogous results as well as analyses exclusively focused on populations
454 located in areas where the distribution range of both species overlap (data not shown).
455
456 **Discussion**
457
458 Our STRUCTURE and NEWHYBRIDS analyses indicate considerable rates of hybridization
459 between *Q. durata* and *Q. berberidifolia* across their distribution ranges, which is
460 consistent with patterns found at local scales in a previous morphology-based study
461 (Forde & Faris, 1962). Although the two species are expected to share some alleles due
462 to common ancestry, several lines of evidence support that the observed patterns of
463 genetic admixture have resulted from genuine interspecific hybridization and cannot be
464 merely explained by incomplete lineage sorting (Muir & Schlotterer, 2005; Lexer *et al*.,
465 2006): i) Bayesian and PCA analyses identified two genetic clusters that are in good
466 agreement with the two morphological species and (ii) simulations demonstrated a high
467 performance of our set of markers to correctly identify hybrids and purebred
468 individuals; iii) We found strong differences in the rates of hybridization between
469 populations from parapatric and sympatric areas (Fig. 2), a spatial pattern that is
470 incompatible with ancestral polymorphism; iv) The presence within the same locality of
471 individuals with very different levels of admixed ancestry indicates that genetic
472 admixture is consequence of hybridization (i.e. the co-existence of purebred individuals,
473 first generation hybrids and backcrosses), as incomplete lineage sorting would have

474 resulted in a nearly identical background level of admixed ancestry across all
475 individuals from the same population (e.g. Fig. 2 in Tsuda *et al*., 2015).
476 STRUCTURE supports the presence of two main clusters corresponding with the
477 two parental species, but these analyses also point to a south to north cline of genetic
478 subdivision within *Q. berberidifolia* that seems to be the biologically most meaningful
479 clustering solution to explain spatial patterns of genetic admixture and species
480 boundaries (Fig. 2a; Ortego *et al*., 2015a). The two clusters of *Q. berberidifolia* roughly
481 separate populations north and south of the Transverse Ranges, a geographic barrier that
482 has been frequently identified to be associated with phylogeographic breaks in many
483 other Californian taxa (Calsbeek *et al*., 2003; Chatzimanolis & Caterino, 2007; Davis *et*
484 *al*., 2008). According to STRUCTURE analyses for K=3, twenty-six percent of the
485 analysed individuals were identified as hybrids according with the $Q \geq 0.90$ threshold
486 criterion (Vähä & Primmer, 2006; Lepais *et al*., 2009; Cullingham *et al*., 2012) (Fig.
487 2a), and this figure increased to 34 % when only populations located in areas of
488 geographic range overlap were considered. NEWHYBRIDS analyses estimated a higher
489 frequency of hybrids than STRUCTURE for both the entire study area (42%) and the
490 region of parapatry (54%), which may be explained by the presence of genetic
491 substructure within *Q. berberidifolia* and the fact that NEWHYBRIDS can only
492 accommodate two genetic clusters (Anderson & Thompson, 2002). These hybridization
493 rates are similar to those reported for two other hybridizing Californian oaks (Ortego *et*
494 *al*., 2014a) but markedly higher than those found among most interbreeding taxa within
495 the genus (e.g. Craft *et al*., 2002; Curtu *et al*., 2007; Cavender-Bares & Pahlich, 2009).
496 The higher occurrence of hybrids in areas where the geographical ranges of both
497 parental taxa overlap confirms previous studies indicating that spatial proximity is an
498 important factor determining hybridization rates (e.g. Dodd & Afzal-Rafii, 2004; Ortego

499 *et al*., 2014a). However, we found evidence of introgression of *Q. durata* in allopatric
500 populations of *Q. berberidifolia* located >190 km away from the closest populations of
501 the former (Fig. 2a). This could have resulted from long-distance pollen dispersal (e.g.
502 Dodd & Afzal-Rafii, 2004) or historic hybridization events followed by local extinction
503 of one parental species (Ortego *et al*., 2014a), being this last hypothesis very unlikely to
504 explain our results due to the lack of serpentine soils in southern Californian
505 (Kruckeberg, 1984). The high admixture proportions found in some individuals from
506 populations of *Q. berberidifolia* located far away from the distribution limit of *Q.*
507 *durata* (e.g. IGN and ELS; Fig. 2a) suggests the presence of first generation
508 hybrids/backcrosses and points to long-distance pollen dispersal as the most likely
509 proximate mechanism explaining the occurrence of a few introgressed genotypes in
510 allopatric populations (Dodd & Afzal-Rafii, 2004). These results are in agreement with
511 paternity-based studies on oaks showing that although pollen dispersal quickly decays
512 with distance from paternal trees (e.g. Streiff *et al*., 1999; Pluess *et al*., 2009), sporadic
513 long-distance pollination events can still have certain impact on the genetic structure
514 and diversity of faraway populations (e.g. Buschbom *et al*., 2011; Hampe *et al*., 2013).
515 Our data suggest that both species maintain their genetic and ecological identity
516 in the presence of frequent interspecific gene flow, a typical outcome for highly
517 interbreeding oaks in which the hybrid state often constitutes a transitory phase
518 followed by parental species “resurrection” in a few generations via recurrent
519 backcrossing and asymmetrical gene flow (Bacilieri *et al*., 1996; Petit *et al*., 2004;
520 Lepais & Gerber, 2011). The adaptation of *Q. durata* to serpentine soils in which *Q.*
521 *berberidifolia* is unable to form stable populations is likely to have resulted in disruptive
522 selection linked to microhabitat specialization (Whittaker, 1954; Brady *et al*., 2005;
523 Wright, 2007). Despite the fact that both taxa often grow in very close geographical

524 proximity and have high potential for interspecific pollen flow (Forde & Faris, 1962),
525 the formation of hybrid swarms may be prevented by strong selection against
526 introgressed individuals in microhabitats that are mostly optimal for either parental
527 species (Barton & Hewitt, 1985). However, coalescent-based estimates of migration
528 indicated that the model best supporting the data was the one considering asymmetric
529 gene flow from *Q. durata* into both allopatric and parapatric populations of *Q.*
530 *berberidifolia*, which suggests that hybrids may be less competitive in serpentine sites
531 because of their lower tolerance to these soils (Brooks, 1987; Harrison, 1999; Brady *et*
532 *al*., 2005). The fact that MIGRATE-N analyses based on purebred individuals identified
533 using the $Q = 0.90$ threshold criterion have also inferred the presence of asymmetric
534 gene flow (Table 1) indicates that the small degree of admixed ancestry (<10%) present
535 in putative "purebred" individuals contain a detectable genetic signal of asymmetric
536 introgression, highlighting the impacts of this phenomenon beyond first-generation
537 hybrids (see also Starr *et al*., 2013). Phenology mismatches have been also suggested as
538 an important isolation mechanism both within and across different oak species (e.g.
539 Cavender-Bares & Pahlich, 2009). Accordingly, Forde and Faris (1962) found evidence
540 for differences in flowering time between nearby populations of *Q. durata* and *Q.*
541 *berberidifolia*, which may be an important premating barrier reducing hybridization
542 rates even when both taxa live at a dispersal distance from each other. Thus,
543 environment-mediated selection and/or assortative mating are likely to explain why
544 populations located in areas where the ranges of both species overlap do not generally
545 converge into hybrid swarms (Fig. 2).
546
547 Environmental disturbance
548

549 In accordance with the environmental disturbance hypothesis, we found that genetic
550 admixture increases with wildfire frequency, an effect that was significant across a wide
551 range of spatial scales and also exclusively considering parapatric populations from
552 both species. Different mechanisms could explain the positive effects of wildfire
553 frequency on local patterns of hybridization. Frequent wildfires are expected to reduce
554 intra- and interspecific competition, which may relax selection against hybrids and
555 increase their rates of successful establishment even if they still perform worse than
556 either parental species (Brooks, 1987; Harrison, 1999; Brady *et al*., 2005). Although
557 scrub oaks are good resprouters (Keeley, 1992), the re-establishment of parental species
558 through vegetative regeneration is expected to be hampered by a high wildfire
559 frequency and intensity (Pausas & Keeley, 2014; Pausas *et al*., 2016), which may
560 ultimately increase the rates of seedling recruitment and the establishment of hybrids.
561 Wildfires also open habitat patches and reduce local population densities and sizes,
562 which is expected to decrease conspecific pollen availability (Breed *et al*., 2012; Ortego
563 *et al*., 2014b), allow immigrant pollen from other species to compete with local
564 conspecific pollen (Lagache *et al*., 2013), and favour more effective pollen and seed
565 dispersal from neighbouring population across barren landscapes (Bacles *et al*., 2006;
566 Bacles & Ennos, 2008). Finally, wildfires can potentially open new niches where
567 hybrids may outperform or perform similarly well than parental taxa (Lewontin &
568 Birch, 1966) due to biotic and abiotic changes in soil composition (Cerdà & Robichaud,
569 2009) or altered interactions with species involved in earlier stages of secondary
570 ecological succession (Horn, 1974). Thus, different factors can result in burned habitat
571 patches being occupied more frequently by genetically admixed individuals and their
572 persistence may be linked with the frequency of the disturbance, as has been proposed
573 to explain geographical patterns of genetic admixture and diversity for other plants in

574 relation with the degree of stability of suitable habitats across glacial and interglacial
575 cycles (i.e. higher admixture/hybridization in populations from climatically less stable
576 areas; e.g. Ortego *et al*., 2015b; Guo, 2014).

577 Despite support for an important role of wildfire frequency on patterns of
578 hybridization, habitat disturbance by urbanization and land clearing for agriculture was
579 only included in models for two spatial scales with a positive, but non-significant effect
580 on the degree of genetic admixture. Some previous studies on trees have found that
581 hybrids between native and introduced species are more frequently established in
582 human-altered landscapes such as roadsides and urbanized areas (Thompson *et al*.,
583 2010; Hoban *et al*., 2012), whereas others have reported widespread hybridization
584 across a variety of natural and disturbed areas (Zalapa *et al*., 2009). The fact that these
585 studies are focused on native-exotic species hybridization may be confounding the
586 effects purely due to environmental disturbance with those resulting from introduction
587 history, which is expected to be biased towards human-populated areas that, in turn, are
588 subjected to many different forms of anthropogenic alterations (Guo, 2014). On the
589 other hand, large-scale human alterations may have weak effects on hybridization rates
590 as a consequence of complete or nearly complete habitat depletion, i.e. environmental
591 disturbance does not result in a new niche that is suitable for either hybrids or any
592 parental species (Anderson, 1948; Anderson & Stebbins, 1954; Arnold, 1997).

593

594 Environmental conditions

595

596 The studied environmental factors (elevation, climate, and soil type) unrelated with
597 habitat disturbance had no significant effect on estimates of mixed gene pools. This
598 result contrasts with previous studies showing that the occurrence of hybrids is

599 constrained by the presence of patches with particular environments suitable for
600 seedling establishment and survival (e.g. Williams *et al*., 2001; Dodd & Afzal-Rafii,
601 2004; Cullingham *et al*., 2012; Ortego *et al*., 2014a). In our study system, the random
602 distribution of hybrids with respect to climate and elevation may be explained by the
603 fact that the most important selective force controlling the distribution of the two focal
604 species is probably the presence of serpentine soils (Forde & Faris, 1962). These soils
605 are scattered in the landscape and often vary in terms of occurrence and chemical
606 composition at scales of a few tens of meters, a fine-grain heterogeneity not covered by
607 available soil maps including information on soil taxa order and suborder (Whittaker,
608 1954; Kruckeberg, 1984; Brady *et al*., 2005).
609
610 Conclusions and future directions
611
612 Overall, our study shows for the first time that wildfires are an important source of
613 environmental disturbance promoting genetic admixture between two ecologically well-
614 differentiated species. Our results suggest that the future predictions for increased
615 wildfire frequency in California and many other parts of the world linked with human-
616 induced climate and land use changes (Westerling *et al*., 2006; Westerling & Bryant,
617 2008) will have an important impact not only on species distribution and demographic
618 dynamics but also on interspecific patterns of gene flow (Barr *et al*., 2015).
619 Longitudinal studies analysing temporal changes in the proportion and genetic
620 composition of hybrids after specific disturbances (e.g. Muhlfeld *et al*., 2014) and
621 experimental/quasi-experimental approaches considering different spatial replicates and
622 intensities of environmental disturbances (e.g. Stewart *et al*., 2015) will provide
623 valuable information to understand more precisely the impacts of human activities on

processes of inter-specific hybridization. Future studies considering detailed analyses of soil chemistry (Forde & Faris, 1962; Yost *et al*., 2012), species phenology (Lamont *et al*., 2003), interspecific fecundity (Williams *et al*., 2001) and contemporary patterns of pollen flow (Lepais *et al*., 2009; Lagache *et al*., 2013) will also help to determine more accurately the factors influencing hybridization and how they interact with environmental heterogeneity and habitat disturbance. Finally, genome scans aimed to identify regions under selection and containing gene variants associated with local adaptation processes can provide important insights to understand the ecological and evolutionary mechanisms underlying asymmetric patterns of introgression in response to human-induced environmental perturbations and climate change (Gailing *et al*., 2004; Renaud *et al*., 2013).

Acknowledgements

We wish to thank to Isabel Afán (LAST-EBD) for her valuable help in GIS analyses. Four anonymous referees provided valuable comments on an earlier draft of this manuscript. JO was supported by a Ramón y Cajal (RYC-2013-12501) and Severo Ochoa (SEV-2012-0262) research fellowships. This work was supported by an internal EBD "Microproyectos" grant to JO, financed by the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Program for Centres of Excellence in R+D+I (SEV-2012-0262). Logistical support was provided by the Laboratorio de Ecología Molecular, Estación Biológica de Doñana, CSIC (LEM-EBD). Support for fieldwork was provided by UCLA research funds of VLS.

Author contributions

J.O. and V.L.S conceived and designed the study. J.O., P.F.G. and V.L.S collected the samples. J.O. performed the genetic analyses, analysed the data and wrote the manuscript.

References

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman JW, Brelsford A, Buerkle CA, Buggs R, *et al*. 2013. Hybridization and speciation. *Journal of Evolutionary Biology* **26**(2): 229-246.

Alberto F, Niort J, Derory J, Lepais O, Vitalis R, Galop D, Kremer A. 2010. Population differentiation of sessile oak at the altitudinal front of migration in the French Pyrenees. *Molecular Ecology* **19**(13): 2626-2639.

Anderson E. 1948. Hybridization of the habitat. *Evolution* **2**(1): 1-9.

Anderson E. 1953. Introgressive hybridization. *Biological Reviews of the Cambridge Philosophical Society* **28**(3): 280-307.

Anderson E, Stebbins GL. 1954. Hybridization as an evolutionary stimulus. *Evolution* **8**(4): 378-388.

Anderson E, Thompson EA. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**(3): 1217-1229.

Andrew RL, Ostevik KL, Ebert DP, Rieseberg LH. 2012. Adaptation with gene flow across the landscape in a dune sunflower. *Molecular Ecology* **21**(9): 2078-2091.

Arnold ML. 1997. *Natural Hybridization and Evolution*. Oxford, UK: Oxford University Press.

Bacilieri R, Ducousso A, Petit RJ, Kremer A. 1996. Mating system and asymmetric hybridization in a mixed stand of European oaks. *Evolution* **50**(2): 900-908.

Bacles CFE, Ennos RA. 2008. Paternity analysis of pollen-mediated gene flow for *Fraxinus excelsior* L. in a chronically fragmented landscape. *Heredity* **101**(4): 368-380.

Bacles CFE, Lowe AJ, Ennos RA. 2006. Effective seed dispersal across a fragmented landscape. *Science* **311**(5761): 628-628.

Barr KR, Kus BE, Preston KL, Howell S, Perkins E, Vandergast AG. 2015. Habitat fragmentation in coastal southern California disrupts genetic connectivity in the cactus wren (*Campylorhynchus brunneicapillus*). *Molecular Ecology* **24**(10): 2349-2363.

Barton NH. 2001. The role of hybridization in evolution. *Molecular Ecology* **10**(3): 551-568.

Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**: 113-148.

Baskett ML, Gomulkiewicz R. 2011. Introgressive hybridization as a mechanism for species rescue. *Theoretical Ecology* **4**(2): 223-239.

Beerli P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* **22**(3): 341-345.

Beerli P. 2009. How to Use MIGRATE or Why are Marko Chain Monte Carlo Programs Difficult to Use? New York, NY, USA: Cambridge University Press.

Beerli P, Felsenstein J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America* **98**(8): 4563-4568.

Beerli P, Palczewski M. 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* **185**(1): 313-326.

Brady KU, Kruckeberg AR, Bradshaw HD. 2005. Evolutionary ecology of plant adaptation to serpentine soils. *Annual Review of Ecology Evolution and Systematics* **36**: 243-266.

Breed MF, Gardner MG, Ottewell KM, Navarro CM, Lowe AJ. 2012. Shifts in reproductive assurance strategies and inbreeding costs associated with habitat fragmentation in Central American mahogany. *Ecology Letters* **15**: 444-452.

Brooks RR. 1987. *Serpentine and its Vegetation: A Multidisciplinary Approach*. Portland, OR, USA: Dioscorides.

Buggs RJA. 2007. Empirical study of hybrid zone movement. *Heredity* **99**(3): 301-312.

Burgarella C, Lorenzo Z, Jabbour-Zahad R *et al*. 2009. Detection of hybrids in nature: application to oaks (*Quercus suber* and *Q. ilex*). *Heredity* **102**(5), 442-452.

Buschbom J, Yanbaev Y, Degen B. 2011. Efficient long-distance gene flow into an isolated relict oak stand. *Journal of Heredity* **102**(4): 464-472.

Calsbeek R, Thompson JN, Richardson JE. 2003. Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. *Molecular Ecology* **12**(4): 1021-1029.

Cavender-Bares J, Pahlich A. 2009. Molecular, morphological and ecological niche differentiation of sympatric sister oak species, *Quercus virginiana* and *Q. geminata* (Fagaceae). *American Journal of Botany* **96**(9): 1690-1702.

Cerdà A, Robichaud P. 2009. *Fire Effects on Soils and Restoration Strategies*. Enfield, NH, USA: Science Publishers.

Craft KJ, Ashley MV, Koenig WD. 2002. Limited hybridization between *Quercus lobata* and *Quercus douglasii* (Fagaceae) in a mixed stand in central coastal California. *American Journal of Botany* **89**(11): 1792-1798.

724 **Crispo E, Moore JS, Lee-Yaw JA, Gray SM, Haller BC. 2011.** Broken barriers:
725 Human-induced changes to gene flow and introgression in animals. *Bioessays*
726 **33**(7): 508-518.
727 **Cullingham CI, James PMA, Cooke JEK, Coltman DW. 2012.** Characterizing the
728 physical and genetic structure of the lodgepole pine jack pine hybrid zone:
729 mosaic structure and differential introgression. *Evolutionary Applications* **5**(8):
730 879-891.
731 **Cullingham CI, Cooke JEK, Dang S, Davis CS, Cooke BJ, Coltman DW. 2011.**
732 Mountain pine beetle host-range expansion threatens the boreal forest.
733 *Molecular Ecology*, **20**(10):2157–2171.
734 **Curtu AL, Gailing O, Finkeldey R. 2007.** Evidence for hybridization and
735 introgression within a species-rich oak (*Quercus* spp.) community. *BMC*
736 *Evolutionary Biology* **7**: 218.
737 **Chatzimanolis S, Caterino MS. 2007.** Toward a better understanding of the
738 "Transverse Range Break": Lineage diversification in southern California.
739 *Evolution* **61**(9): 2127-2141.
740 **Darwin CR. 1859.** *On the Origin of Species by Means of Natural Selection*. London,
741 UK: John Murray.
742 **Davis EB, Koo MS, Conroy C, Patton JL, Moritz C. 2008.** The California Hotspots
743 Project: identifying regions of rapid diversification of mammals. *Molecular*
744 *Ecology* **17**(1): 120-138.
745 **Dodd RS, Afzal-Rafii Z. 2004.** Selection and dispersal in a multispecies oak hybrid
746 zone. *Evolution* **58**(2): 261-269.
747 **Dow BD, Ashley MV. 1998.** High levels of gene flow in bur oak revealed by paternity
748 analysis using microsatellites. *Journal of Heredity* **89**(1): 62-70.

749 **Doyle J, Doyle J. 1990.** Isolation of plant DNA from fresh tissue. *Focus*, **12**: 13-15.
750 **eFloras. 2015.** *Flora of North America*. URL http://www.efloras.org [accessed on 20
751 December 2015]. St. Louis, MO, USA & Cambridge, MA, USA: Missouri
752 Botanic Garden & Harvard University Herbaria.
753 **Ellstrand NC, Schierenbeck KA. 2000.** Hybridization as a stimulus for the evolution
754 of invasiveness in plants? *Proceedings of the National Academy of Sciences of*
755 *the United States of America* **97**(13): 7043-7050.
756 **Falush D, Stephens M, Pritchard JK. 2003.** Inference of population structure using
757 multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*
758 **164**(4): 1567-1587.
759 **Field DL, Ayre DJ, Whelan RJ, Young AG. 2011.** Patterns of hybridization and
760 asymmetrical gene flow in hybrid zones of the rare *Eucalyptus aggregata* and
761 common *E. rubida*. *Heredity* **106**(5): 841-853.
762 **Forde MB, Faris DG. 1962.** Effect of introgression on serpentine endemism of
763 *Quercus durata*. *Evolution* **16**(3): 338-347.
764 **Gailing O, Macnair MR, Bachmann K. 2004.** QTL mapping for a trade-off between
765 leaf and bud production in a recombinant inbred population of *Microseris*
766 *douglasii* and *M. bigelovii* (Asteraceae, Lactuceae): A potential preadaptation
767 for the colonization of serpentine soils. *Plant Biology* **6**(4): 440-446.
768 **Goudet J. 1995.** FSTAT (Version 1.2): A computer program to calculate F-statistics.
769 *Journal of Heredity* **86**(6): 485-486.
770 **Guo QF. 2014.** Plant hybridization: the role of human disturbance and biological
771 invasion. *Diversity and Distributions* **20**(11): 1345-1354.
772 **Haines ML, Melville J, Sumner J, Clemann N, Chapple DG, Stuart-Fox D. 2016.**
773 Geographic variation in hybridization and ecological differentiation between

774 three syntopic, morphologically similar species of montane lizards. *Molecular*
775 *Ecology*, **25**(12): 2887-2903.
776 **Hampe A, Pemonge MH, Petit RJ. 2013.** Efficient mitigation of founder effects
777 during the establishment of a leading-edge oak population. *Proceedings of the*
778 *Royal Society B-Biological Sciences* **280**: 20131070.
779 **Harrison RG. 1986.** Pattern and process in a narrow hybrid zone. *Heredity* **56**: 337-
780 349.
781 **Harrison S. 1999.** Local and regional diversity in a patchy landscape: Native, alien, and
782 endemic herbs on serpentine. *Ecology* **80**(1): 70-80.
783 **Hasselman DJ, Argo EE, McBride MC, Bentzen P, Schultz TF, Perez-Umphrey**
784 **AA, Palkovacs EP. 2014.** Human disturbance causes the formation of a hybrid
785 swarm between two naturally sympatric fish species. *Molecular Ecology* **23**(5):
786 1137-1152.
787 **Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005.** Very high resolution
788 interpolated climate surfaces for global land areas. *International Journal of*
789 *Climatology* **25**(15): 1965-1978.
790 **Hoban SM, McCleary TS, Schlarbaum SE, Anagnostakis SL, Romero-Severson J.**
791 **2012.** Human-impacted landscapes facilitate hybridization between a native and
792 an introduced tree. *Evolutionary Applications* **5**(7): 720-731.
793 **Horn HS. 1974.** The ecology of secondary succession. *Annual Review of Ecology and*
794 *Systematics* **5**: 25-37.
795 **Hubisz MJ, Falush D, Stephens M, Pritchard JK. 2009.** Inferring weak population
796 structure with the assistance of sample group information. *Molecular Ecology*
797 *Resources* **9**(5): 1322-1332.

798 **Jombart T. 2008.** ADEGENET: a R package for the multivariate analysis of genetic
799 markers. *Bioinformatics* **24**(11): 1403-1405.
800 **Keeley JE, Bond WJ, Bradstock RA, Pausas JG, Rundel PW. 2012.** *Fire in*
801 *Mediterranean Ecosystems: Ecology, Evolution and Management*. Cambridge,
802 UK: Cambridge University Press.
803 **Keeley JE. 1992.** Recruitment of seedlings and vegetative sprouts in unburned
804 chaparral. *Ecology* **73**(4): 1194-1208.
805 **King RS, Boysen JR, Brenneman JM, Cong RM, Hunter TS. 2015.** Fire created
806 habitat partitioning and isolation between hybridizing warblers. *Ecosphere* **6**(3):
807 34.
808 **Kruckeberg AR. 1984.** *California Serpentines: Flora, Vegetation, Geology, Soils, and*
809 *Management Problems*. Los Angeles, CA, USA: University of California
810 Publications in Botany.
811 **Lagache L, Klein EK, Guichoux E, Petit RJ. 2013.** Fine-scale environmental control
812 of hybridization in oaks. *Molecular Ecology* **22**(2): 423-436.
813 **Lamont BB, He T, Enright NJ, Krauss SL, Miller BP. 2003.** Anthropogenic
814 disturbance promotes hybridization between *Banksia* species by altering their
815 biology. *Journal of Evolutionary Biology* **16**(4): 551-557.
816 **Lancaster LT, Kay KM. 2013.** Origin and diversification of the California flora: re-
817 examining classic hypotheses with molecular phylogenies. *Evolution* **67**(4):
818 1041-1054.
819 **Langhans TM, Storm C, Schwabe A. 2009.** Biological soil crusts and their
820 microenvironment: Impact on emergence, survival and establishment of
821 seedlings. *Flora* **204**(2): 157-168.

Langin KM, Sillett TS, Funk WC, Morrison SA, Desrosiers MA, Ghalambor CK. 2015. Islands within an island: Repeated adaptive divergence in a single population. *Evolution* **69**(3): 653-665.

Lepais O, Gerber S. 2011. Reproductive patterns shape introgression dynamics and species succession within the European white oak species complex. *Evolution* **65**(1): 156-170.

Lepais O, Petit RJ, Guichoux E, Lavabre JE, Alberto F, Kremer A, Gerber S. 2009. Species relative abundance and direction of introgression in oaks. *Molecular Ecology* **18**(10): 2228-2242.

Lewontin RC, Birch LC. 1966. Hybridization as a sournce of variation for adaptation to new environments. *Evolution* **20**(3): 315-336.

Lexer C, Kremer A, Petit RJ. 2006. Shared alleles in sympatric oaks: recurrent gene flow is a more parsimonious explanation than ancestral polymorphism. *Molecular Ecology* **15**(7): 2007-2012.

Linnaeus C. 1735. *Systema Naturae 1st edition*. Leiden, the Netherlands: Theodorum Haak.

Moritz MA, Batllori E, Bradstock RA, Gill AM, Handmer J, Hessburg PF, Leonard J, McCaffrey S, Odion DC, Schoennagel T, *et al.* 2014. Learning to coexist with wildfire. *Nature* **515**(7525): 58-66.

Morjan CL, Rieseberg LH. 2004. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology* **13**(6): 1341-1356.

Muhlfeld CC, Kovach RP, Jones LA, Al-Chokhachy R, Boyer MC, Leary RF, Lowe WH, Luikart G, Allendorf FW. 2014. Invasive hybridization in a

threatened species is accelerated by climate change. *Nature Climate Change* **4**(7): 620-624.

Muir G, Schlotterer C. 2005. Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). *Molecular Ecology* **14**(2): 549-561.

Muller CH. 1952. Ecological control of hybridization in *Quercus* - A factor in the mechanism of evolution. *Evolution* **6**(2): 147-161.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* **403**(6772): 853-858.

Ortego J, Riordan EC, Gugger PF, Sork VL. 2012. Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. *Molecular Ecology* **21**(13): 3210-3223.

Ortego J, Gugger PF, Riordan EC, Sork VL. 2014a. Influence of climatic niche suitability and geographical overlap on hybridization patterns among southern Californian oaks. *Journal of Biogeography* **41**(10): 1895-1908.

Ortego J, Bonal R, Munoz A, Aparicio JM. 2014b. Extensive pollen immigration and no evidence of disrupted mating patterns or reproduction in a highly fragmented holm oak stand. *Journal of Plant Ecology* **7**(4): 384-395.

Ortego J, Noguerales V, Gugger PF, Sork VL. 2015a. Evolutionary and demographic history of the Californian scrub white oak species complex: An integrative approach. *Molecular Ecology* **24**(24): 6188-6208.

Ortego J, Gugger PF, Sork VL. 2015b. Climatically stable landscapes predict patterns of genetic structure and admixture in the Californian canyon live oak. *Journal of Biogeography* **42**(2): 328-338.

Ortego J, Aguirre MP, Noguerales V, Cordero PJ. 2015c. Consequences of extensive habitat fragmentation in landscape-level patterns of genetic diversity and structure in the Mediterranean esparto grasshopper. *Evolutionary Applications* **8**(6): 621-632.

Pausas JG, Keeley JE. 2014. Evolutionary ecology of resprouting and seeding in fire-prone ecosystems. *New Phytologist*, **204**(1): 55-65.

Pausas JG, Pratt RB, Keeley JE, Jacobsen AL, Ramirez AR, Vilagrosa A, Paula S, Kaneakua-Pia IN, Davis SD. 2016. Towards understanding resprouting at the global scale. *New Phytologist*, **209**(3): 945-954.

Petit RJ, Bodenes C, Ducousso A, Roussel G, Kremer A. 2004. Hybridization as a mechanism of invasion in oaks. *New Phytologist* **161**(1): 151-164.

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 Ecology and Management* **258**(5): 735-744.

Prentis PJ, White EM, Radford IJ, Lowe AJ, Clarke AR. 2007. Can hybridization cause local extinction: a case for demographic swamping of the Australian native *Senecio pinnatifolius* by the invasive *Senecio madagascariensis*? *New Phytologist* **176**(4): 902-912.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**(2): 945-959.

R Core Team. 2015. *R: A Language and Environment for Statistical Computing. Version 3.0.2*. URL http://www.rproject.org. Vienna, Austria: R Foundation for Statistical Computing.

Renaut S, Grassa CJ, Yeaman S, Moyers BT, Lai Z, Kane NC, Bowers JE, Burke JM, Rieseberg LH. 2013. Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nature Communications* **4**: 1827.

Rhymer JM, Simberloff D. 1996. Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics* **27**: 83-109.

Rieseberg LH. 1997. Hybrid origins of plant species. *Annual Review of Ecology and Systematics* **28**: 359-389.

Riley SPD, Pollinger JP, Sauvajot RM, York EC, Bromley C, Fuller TK, Wayne RK. 2006. A southern California freeway is a physical and social barrier to gene flow in carnivores. *Molecular Ecology* **15**(7): 1733-1741.

Roberts FM. 1995. *Illustrated guide to the oaks of the Southern California Floristic Province*. Encinitas, CA, USA: F.M. Roberts Publications.

Rundel PW. 1982. Fire as an ecological factor. In: Lange OL, Nobel PS, Osmond CB, Ziegler H, eds. *Physiological Plant Ecology I*. New York, USA: Springer-Verlag, 501-538.

Rushton BS. 1993. Natural hybridization within the genus *Quercus* 1. *Annales des Sciences Forestieres, Vol 50, Supplement 1, 1993 - Genetics of Oaks*: S73-S90.

Saarman NP, Pogson GH. 2015. Introgression between invasive and native blue mussels (genus *Mytilus*) in the central California hybrid zone. *Molecular Ecology* **24**(18): 4723-4738.

Schwartz MW, Thorne JH, Viers JH. 2006. Biotic homogenization of the California flora in urban and urbanizing regions. *Biological Conservation* **127**(3): 282-291.

Silliman FE, Leisner RS. 1958. An analysis of a colony of hybrid oaks. *American Journal of Botany* **45**(10): 730-736.

Staff SS. 2012. *Soil Survey Geographic (SSURGO) database for California*. URL http://websoilsurvey.sc.egov.usda.gov [accessed on 20 December 2015].

Starr TN, Gadek KE, Yoder JB, Flatz R, Smith CI. 2013. Asymmetric hybridization and gene flow between Joshua trees (Agavaceae: Yucca) reflect differences in pollinator host specificity. *Molecular Ecology* **22**(2): 437-449.

Stewart JF, Will RE, Robertson KM, Nelson CD. 2015. Frequent fire protects shortleaf pine (*Pinus echinata*) from introgression by loblolly pine (*P. taeda*). *Conservation Genetics* **16**(2): 491-495.

Streiff R, Ducousso A, Lexer C, Steinkellner H, Gloessl J, Kremer A. 1999. Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L- and *Q. petraea* (Matt.) Liebl. *Molecular Ecology* **8**(5): 831-841.

Syphard AD, Radeloff VC, Keeley JE, Hawbaker TJ, Clayton MK, Stewart SI, Hammer RB. 2007. Human influence on California fire regimes. *Ecological Applications* **17**(5): 1388-1402.

Thompson SL, Lamothe M, Meirmans PG, Perinet P, Isabel N. 2010. Repeated unidirectional introgression towards *Populus balsamifera* in contact zones of exotic and native poplars. *Molecular Ecology* **19**(1): 132-145.

Tsuda Y, Nakao K, Ide Y, Tsumura Y. 2015. The population demography of *Betula maximowicziana*, a cool-temperate tree species in Japan, in relation to the last glacial period: its admixture-like genetic structure is the result of simple population splitting not admixing. *Molecular Ecology* **24**(7): 1403-1418.

Vähä JP, Primmer CR. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* **15**(1): 63-72.

941 **Vandergast AG, Bohonak AJ, Hathaway SA, Boys J, Fisher RN. 2008.** Are hotspots
942 of evolutionary potential adequately protected in southern California? *Biological*
943 *Conservation* **141**(6): 1648-1664.
944 **Vandergast AG, Bohonak AJ, Weissman DB, Fisher RN. 2007.** Understanding the
945 genetic effects of recent habitat fragmentation in the context of evolutionary
946 history: phylogeography and landscape genetics of a southern California
947 endemic Jerusalem cricket (Orthoptera: Stenopelmatidae: Stenopelmatus).
948 *Molecular Ecology* **16**(5): 977-992.
949 **Westerling AL, Bryant BP. 2008.** Climate change and wildfire in California. *Climatic*
950 *Change* **87**: S231-S249.
951 **Westerling AL, Hidalgo HG, Cayan DR, Swetnam TW. 2006.** Warming and earlier
952 spring increase western US forest wildfire activity. *Science* **313**(5789): 940-943.
953 **Whittaker RH. 1954.** The ecology of serpentine soils. *Ecology* **35**(2): 258-259.
954 **Whittemore AT, Schaal BA. 1991.** Interspecific gene flow in sympatric oaks.
955 *Proceedings of the National Academy of Sciences of the United States of*
956 *America* **88**(6): 2540-2544.
957 **Williams JH, Boecklen WJ, Howard DJ. 2001.** Reproductive processes in two oak
958 (*Quercus*) contact zones with different levels of hybridization. *Heredity* **87**: 680-
959 690.
960 **Wright JW. 2007.** Local adaptation to serpentine soils in *Pinus ponderosa*. *Plant and*
961 *Soil* **293**(1-2): 209-217.
962 **Yost JM, Barry T, Kay KM, Rajakaruna N. 2012.** Edaphic adaptation maintains the
963 coexistence of two cryptic species on serpentine soils. *American Journal of*
964 *Botany* **99**(5): 890-897.

Zalapa JE, Brunet J, Guries RP. 2009. Patterns of hybridization and introgression between invasive *Ulmus pumila* (Ulmaceae) and native *U. rubra*. *American Journal of Botany* **96**(6): 1116-1128.

Zeng YF, Liao WJ, Petit RJ, Zhang DY. 2011. Geographic variation in the structure of oak hybrid zones provides insights into the dynamics of speciation. *Molecular Ecology* **20**(23): 4995-5011.

Supporting Information

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

Table S1 Geographical location of *Quercus berberidifolia* and *Q. durata* sampling sites in California.

Table S2 Microsatellite loci used to genotype *Quercus berberidifolia* and *Q. durata*.

Table S3 Accuracy, efficiency and overall performance of assignment of simulated genotypes of *Quercus berberidifolia*, *Q. durata* and their hybrids.

Table S4 Model selection for analyses of genetic admixture (G_{Admix}).

Table S5 General linear model summaries for analyses of genetic admixture (G_{Admix}).

Fig. S1 Models tested using MIGRATE-N and Bayesian estimates of mutation-scaled effective population sizes and migration rates.

991

992 **Fig. S2** Ancestry plots for simulated genotypes of *Q. berberidifolia*, *Q. durata* and four

993 hybrid classes analyzed with STRUCTURE and NEWHYBRIDS.

994

995 **Notes S1** Settings of Structure and Newhybrids analyses.

996

997 **Notes S2** Hybrid simulation and genetic assignment.

998

999 **Notes S3** Settings of MIGRATE-N analyses.

Table 1 Model description and results from model comparison in MIGRATE-N for analyses excluding and including interspecific hybrids between *Quercus berberidifolia* and *Q. durata* identified by STRUCTURE analyses ($Q < 0.90$) (see also Supporting Information Fig. S1a). Bézier approximation scores of log marginal likelihoods, log Bayes factors (LBF) and model probabilities are shown. Best supported model is indicated in bold.

	Excluding hybrids			Including hybrids		
Model description	Bézier	LBF	Probability	Bézier	LBF	Probability
(I) Full migration model	-270658	-525823	<0.0001	-415896	-691754	<0.0001
(II) Bidirectional interspecific gene flow restricted to parapatric populations	-140090	-264686	<0.0001	-255323	-370607	<0.0001
(III) Unidirectional gene flow from *Q. berberidifolia* to *Q. durata*	-113727	-211961	<0.0001	-292242	-444446	<0.0001
(IV) Unidirectional gene flow from *Q. durata* to *Q. berberidifolia*	**-7747**	0	**1.0000**	**-70019**	0	**1.0000**
(V) Unidirectional gene flow from *Q. berberidifolia* to *Q. durata* restricted to parapatric populations	-81256	-147017	<0.0001	-295109	-450180	<0.0001
(VI) Unidirectional gene flow from *Q. durata* to *Q. berberidifolia* restricted to parapatric populations	-16421	-17347	<0.0001	-115182	-90325	<0.0001

Table 2 Model selection to assess the relationship between genetic admixture (G_{Admix}) of *Quercus berberidifolia* and *Q. durata* and [A] species range overlap, [B] climate, [C] elevation, [D] soil type, [E] proportion of habitats disturbed by agriculture and human development, and [F] wildfire frequency. We run three subsets of models considering the same parameters but with some of them (B, C, E and F) measured at three different spatial scales [(a) 10 km^2, (b) 100 km^2, and (c) 1000 km^2 around sampling localities]. Only best ranked equivalent models ($\Delta AICc \leq 2$) are shown.

Model no.	Model	*K*	AIC_c	ΔAIC_c	ω_i
(a) G_{Admix} (10 km^2)					
1	F	3	11.33	0.00	0.13
2	A+C+F	5	12.46	1.13	0.07
3	A+F	4	12.68	1.35	0.07
4	C+F	4	13.01	1.67	0.06
5	E+F	4	13.06	1.73	0.05
(b) G_{Admix} (100 km^2)					
1	F	3	7.41	0.00	0.19
2	A+F	4	8.42	1.01	0.11
(c) G_{Admix} (1000 km^2)					
1	A+F	4	3.01	0.00	0.17
2	E+F	4	3.46	0.44	0.13
3	A+E+F	5	4.49	1.48	0.08
4	A+D+F	5	5.01	2.00	0.06

K, number of parameters in the model; AIC_c, corrected Akaike's information criterion (AIC) value; ΔAIC_c, difference in AIC_c value from that of the strongest model; ω_i, AIC_c weight

1022 **Table 3** General linear models (GLMs) for genetic admixture (G_{Admix}) of *Quercus*
1023 *berberidifolia* and *Q. durata*. Parameter estimates and unconditional standard errors
1024 (USE) were obtained performing model averaging of the best ranked equivalent models
1025 ($\Delta AICc \leq 2$) (see Table 1). Variables are sorted according with their relative importance
1026 based on the sum of Akaike weights ($\Sigma\ \omega_i$) of those models with $\Delta AICc \leq 2$ in which
1027 the variable was present. Bold type indicates significant variables, i.e. variables for
1028 which their unconditional 95 % confidence interval (CI) did not cross zero. The
1029 percentage of explained deviance for each model is indicated in parentheses. We run
1030 three subsets of models considering the same parameters but with some of them (fire
1031 frequency, climate, proportion of habitats disturbed by agriculture and human
1032 development and elevation) measured at three different spatial scales [(a) 10 km^2, (b)
1033 100 km^2, and (c) 1000 km^2 around sampling localities].
1034
1035

	Estimate ± USE	$\Sigma\ \omega_i$	Lower 95% CI	Upper 95% CI
(a) G_{Admix} (10 km^2) (% of explained deviance: 18.20)				
Intercept	0.331 ± 0.100			
Wildfire frequency	0.087 ± 0.033	0.38	**0.02**	**0.15**
Range overlap	0.096 ± 0.078	0.14	-0.06	0.25
Elevation	0.001 ± 0.001	0.13	-0.01	0.01
% of disturbed habitat	-0.001 ± 0.002	0.05	-0.01	0.01
(b) G_{Admix} (100 km^2) (% of explained deviance: 19.15)				
Intercept	0.313 ± 0.063			
Wildfire frequency	0.127 ± 0.037	0.30	**0.06**	**0.20**
Range overlap	0.069 ± 0.062	0.11	-0.05	0.19
(c) G_{Admix} (1000 km^2) (% of explained deviance: 28.03)				
Intercept	0.082 ± 0.124			
Wildfire frequency	0.259 ± 0.067	0.44	**0.13**	**0.39**
Range overlap	0.136 ± 0.072	0.31	-0.01	0.28
% of disturbed habitat	0.004 ± 0.003	0.21	-0.01	0.01
Soil type	0.012 ± 0.020	0.06	-0.02	0.05

Figure legends

Fig. 1 Map from California representing wildfire frequency estimated as the number of years that a given area was burned between 1900 and 2013 (source: CALFIRE Fire Perimeters Geodatabase version 13.2). The map also represents sampling localities for putative *Quercus berberidifolia* (blue stars) and *Q. durata* (red stars). Right top inset shows the range distributions of *Q. berberidifolia* (blue) and *Q. durata* (red) according with records of the two species available in Calflora database (http://www.calflora.org/).

Fig. 2 Results of genetic assignments based on the programs (a) STRUCTURE (K = 2 and K = 3) and (b) NEWHYBRIDS. Each individual is represented by a vertical bar, which is partitioned into K coloured segments showing the individual's probability of assignment to *Quercus berberidifolia* (blue), *Q. durata* (red) and hybrids (yellow, only for NEWHYBRIDS). Thin vertical black lines separate individuals from different sampling localities arranged according to their geographical location from southeast (left) to northwest (right) (population codes are described in Supporting Information Table S1). The horizontal bar indicates whether the different sampled populations are located or not in areas of geographical range overlap between the two focal studied species (allopatry *vs*. parapatry). Areas of parapatry and allopatry were defined according with records of the two species available in Calflora database (http://www.calflora.org/) and verified with distribution maps from eFloras (2015). Panel (c) shows the mean (± SD) log probability of the data (ln Pr(X|K)) over 10 runs of STRUCTURE (left axis, black dots and error bars) for each value of K and the magnitude of ΔK (right axis, open dots). (d) Detail of the leaves of the two study species, *Quercus berberidifolia* (left) and *Q. durata* (right) (photographs by Joaquín Ortego).

1061

1062 **Fig. 3** Principal component analyses (PCA) for genetic data of *Q. berberidifolia*, *Q.*
1063 *durata* and their hybrids. Panels a-b) show a PCA including all individuals and
1064 indicating their assignment to either parental species (red and blue circles) or their
1065 hybrids (yellow triangles) according to (a) STRUCTURE analyses for $K = 3$ and (b)
1066 NEWHYBRIDS (considering a threshold of $Q \geq 0.9$). Panel c) shows a PCA only
1067 including purebred ($Q \geq 0.9$) individuals of *Q. berberidifolia* and indicating their
1068 assignment to the two clusters identified within this species according to STRUCTURE
1069 analyses for $K = 3$.

1070

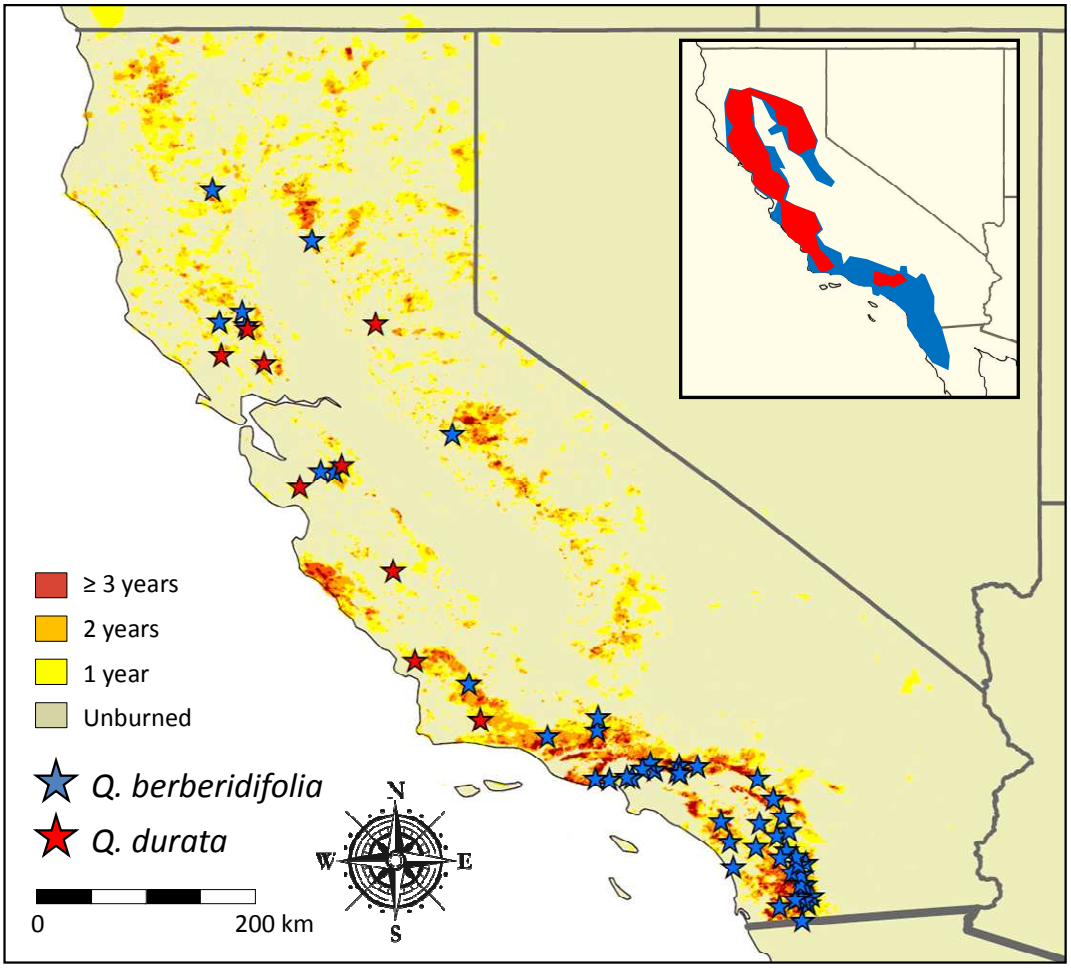
1071 **Fig. 4** Relationship between genetic admixture of *Quercus berberidifolia* and *Q. durata*
1072 (G_{Admix}) and average wildfire frequency estimated in an area of (a) 10 km^2, (b) 100 km^2,
1073 and (c) 1000 km^2 around sampling localities. Regression lines are indicated and dot size
1074 is proportional to sample size for each studied population.

1075

1076

1077 Figure 1

1078

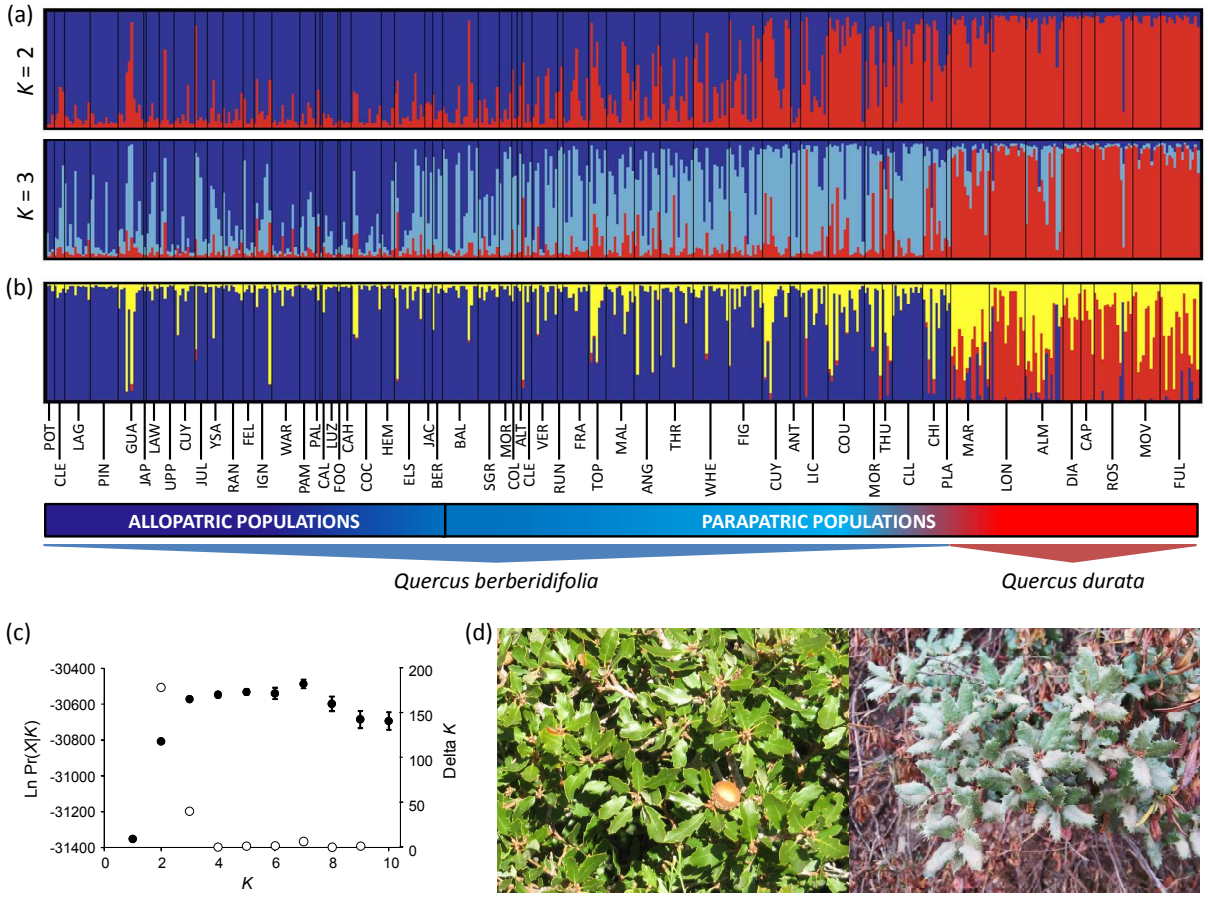


1079

1080

1081 Figure 2

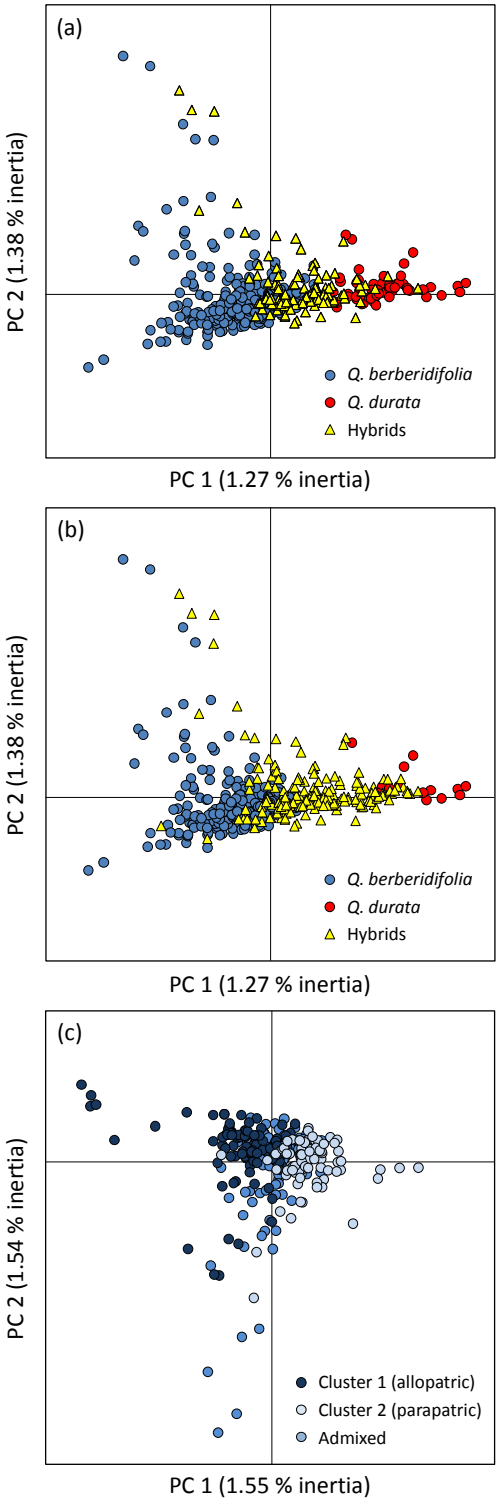
1082



1083

1084 Figure 3

1085

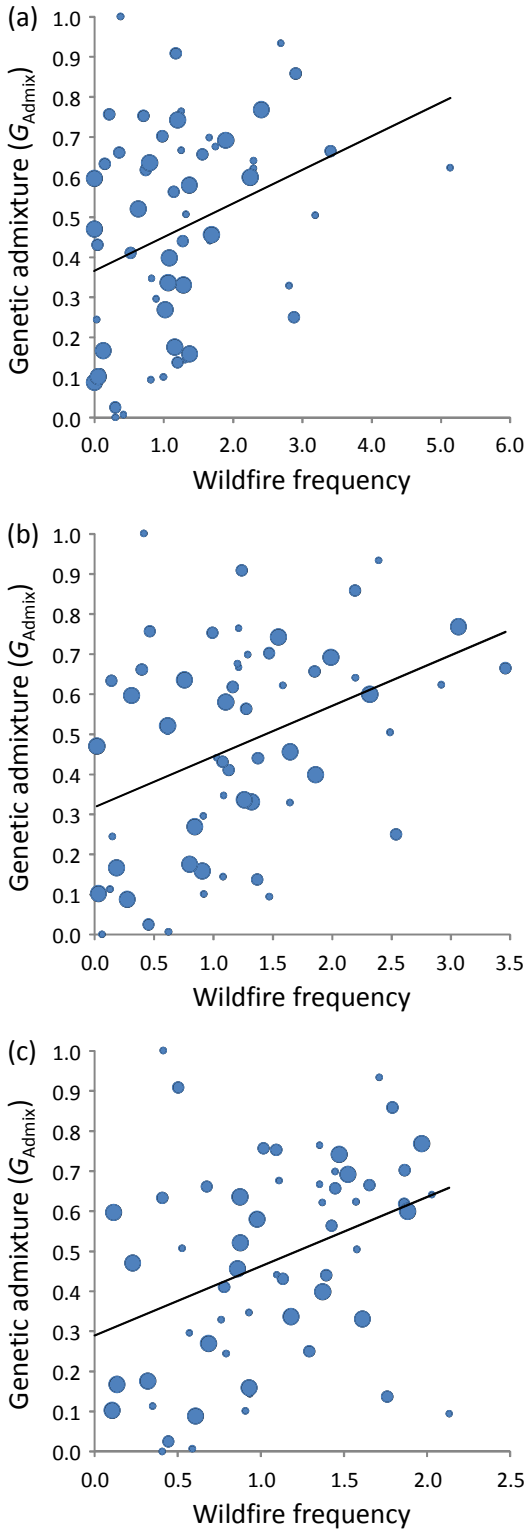


1086

1087

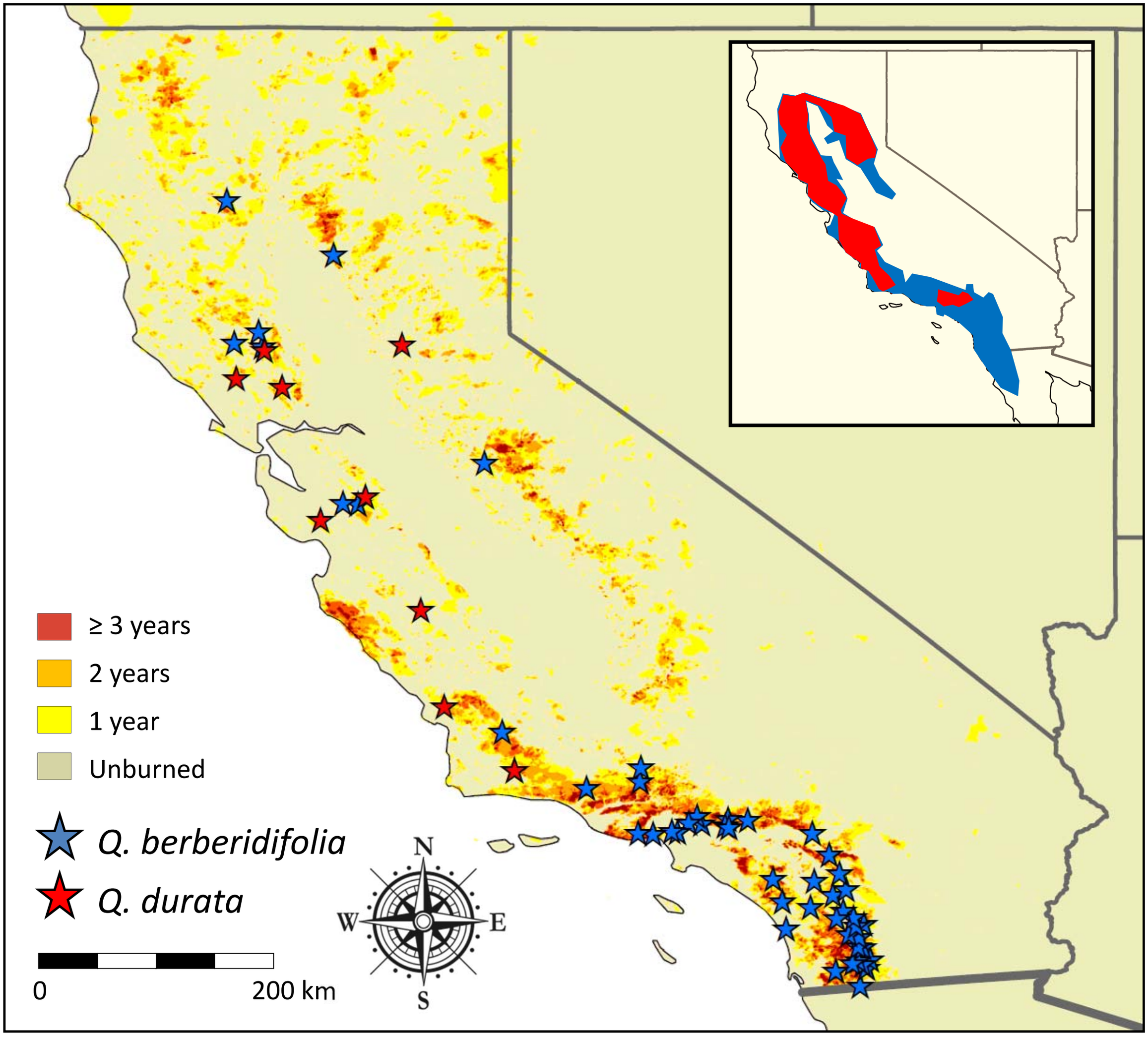
1088

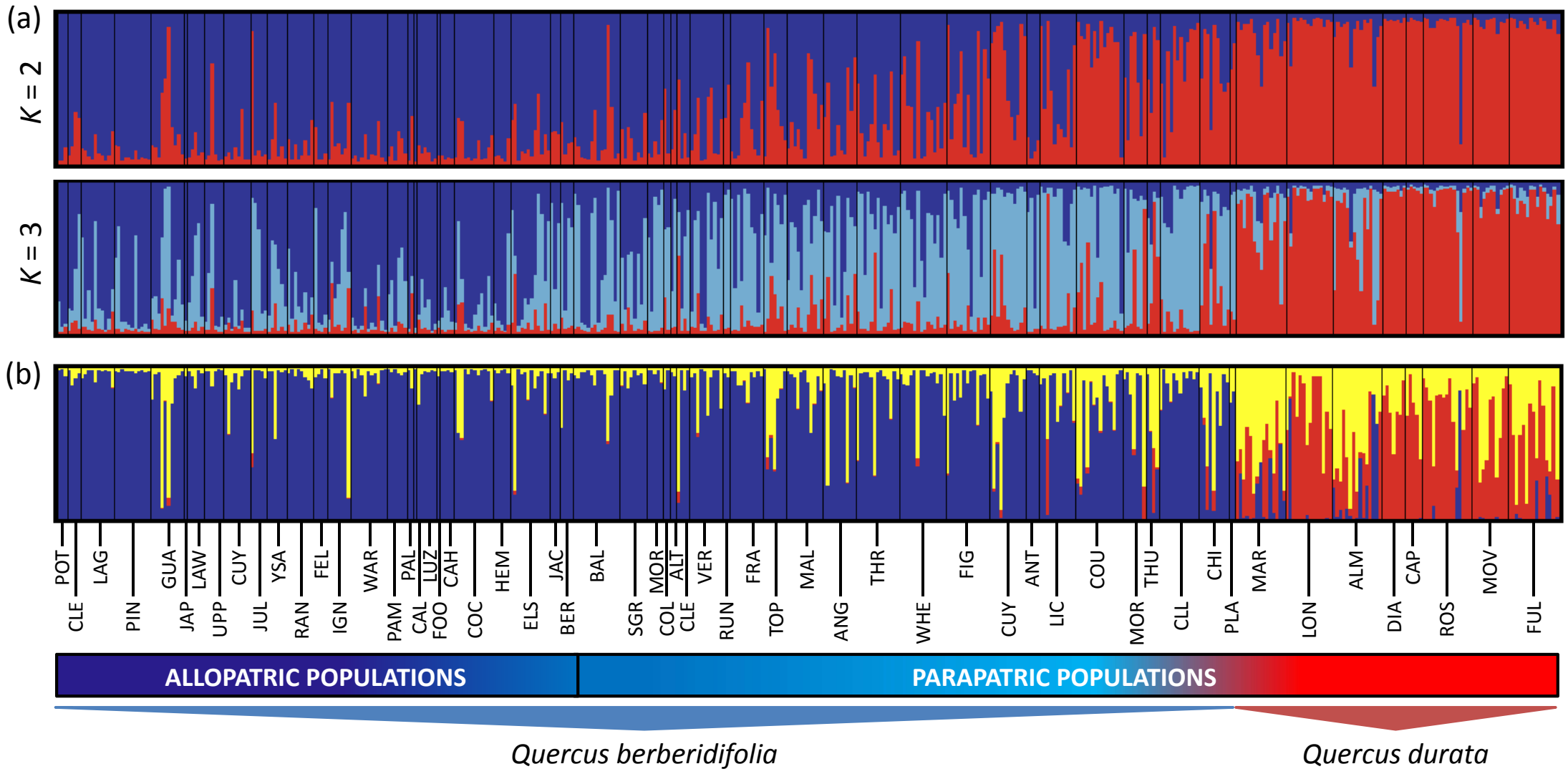
1089 Figure 4

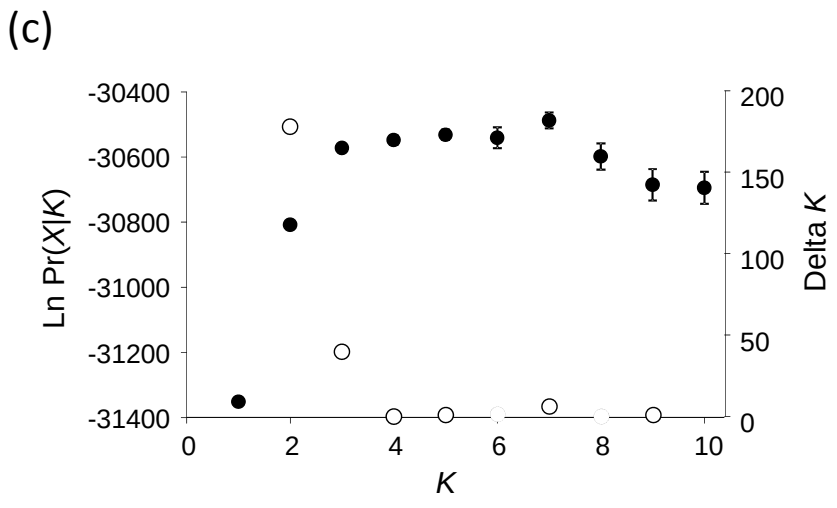


1090

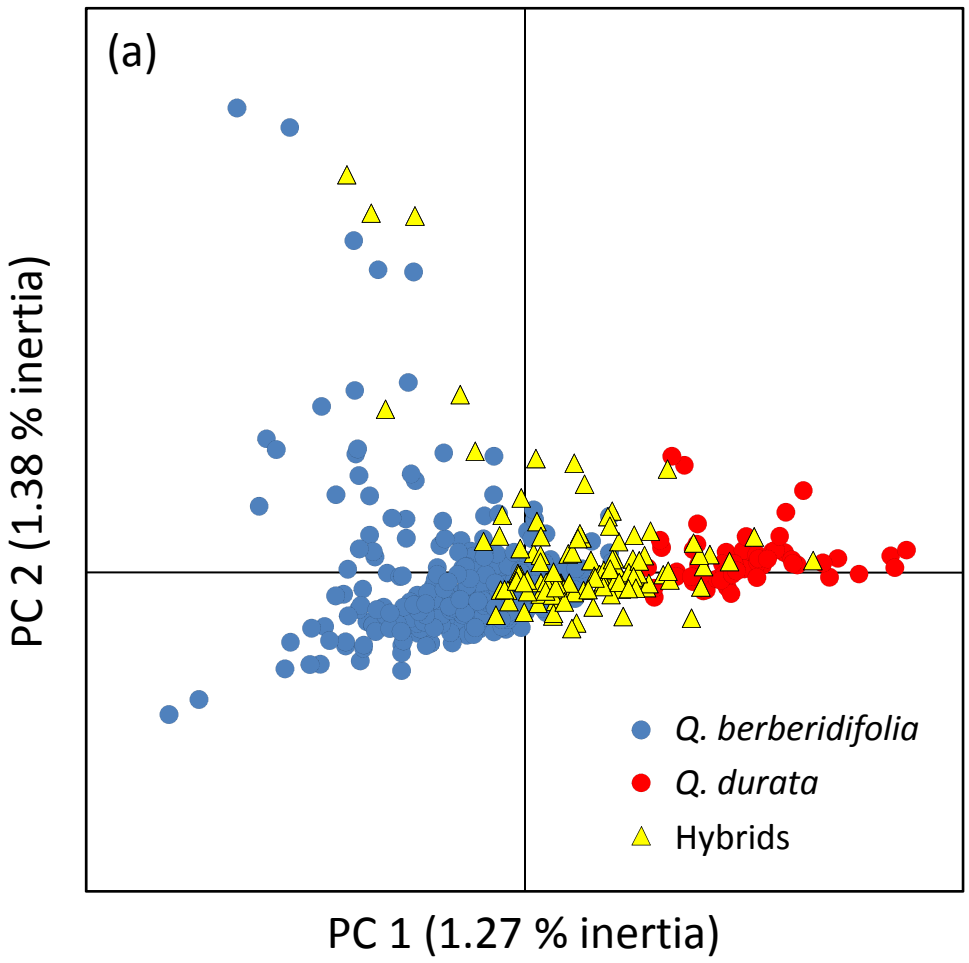
1091

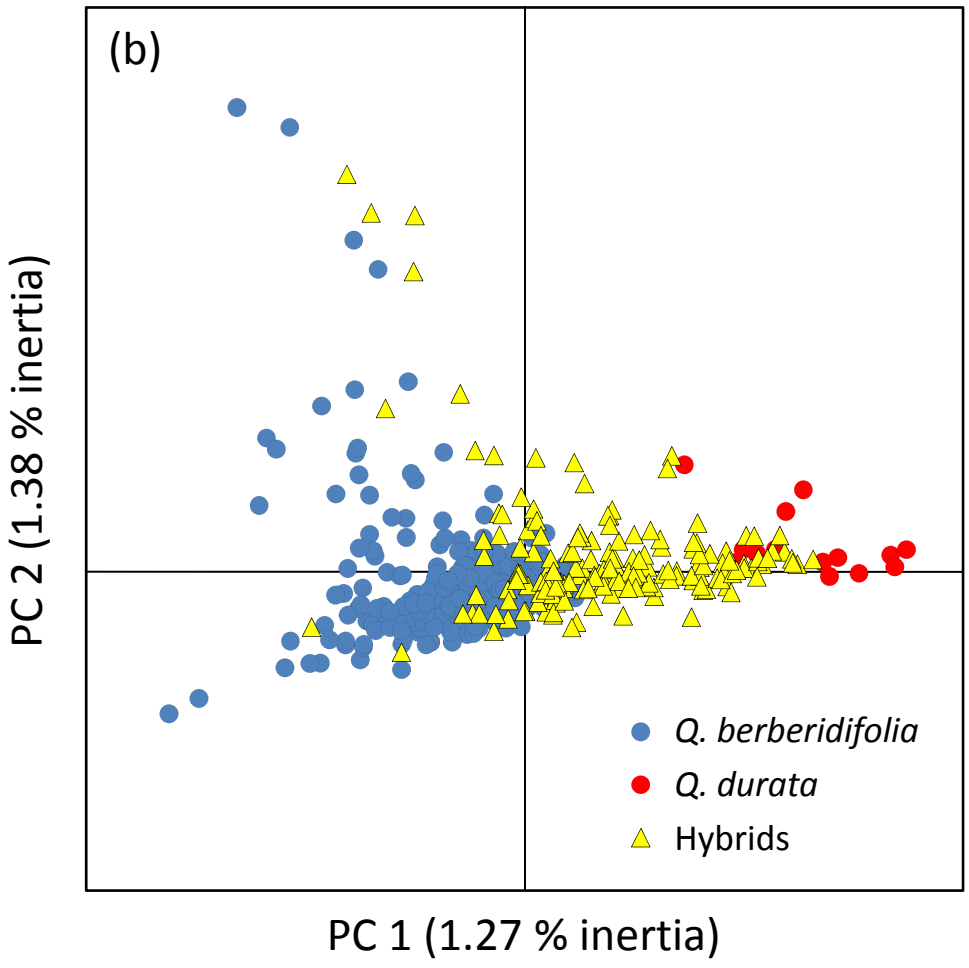

≥ 3 years
2 years
1 year
Unburned
Q. berberidifolia
Q. durata
0
200 km
N
W
E
S

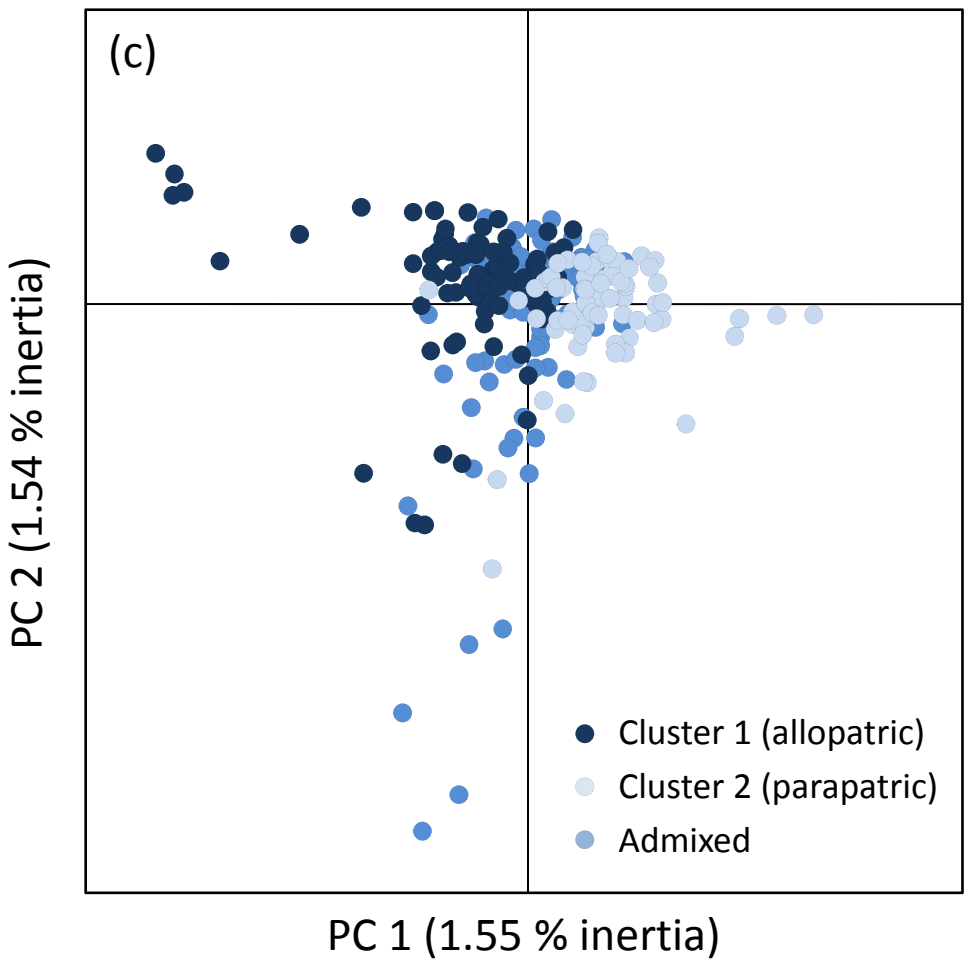

(a)
K = 2
K = 3
(b)
POT
CLE
LAG
PIN
GUA
JAP
LAW
UPP
CUY
JUL
YSA
RAN
FEL
IGN
WAR
PAM
PAL
CAL
LUZ
FOO
CAH
COC
HEM
ELS
JAC
BER
BAL
SGR
MOR
COL
ALT
CLE
VER
RUN
FRA
TOP
MAL
ANG
THR
WHE
FIG
CUY
ANT
LIC
COU
MOR
THU
CLL
CHI
PLA
MAR
LON
ALM
DIA
CAP
ROS
MOV
FUL
ALLOPATRIC POPULATIONS
PARAPATRIC POPULATIONS
Quercus berberidifolia
Quercus durata

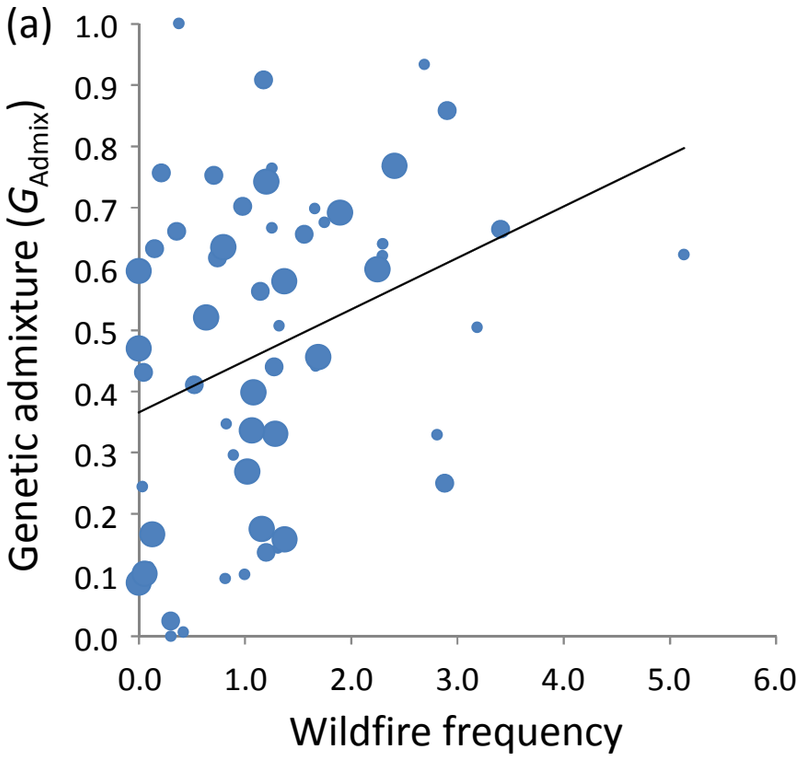

(c)
Ln Pr(X|K)
Delta K
K
-30400
-30600
-30800
-31000
-31200
-31400
200
150
100
50
0
0
2
4
6
8
10

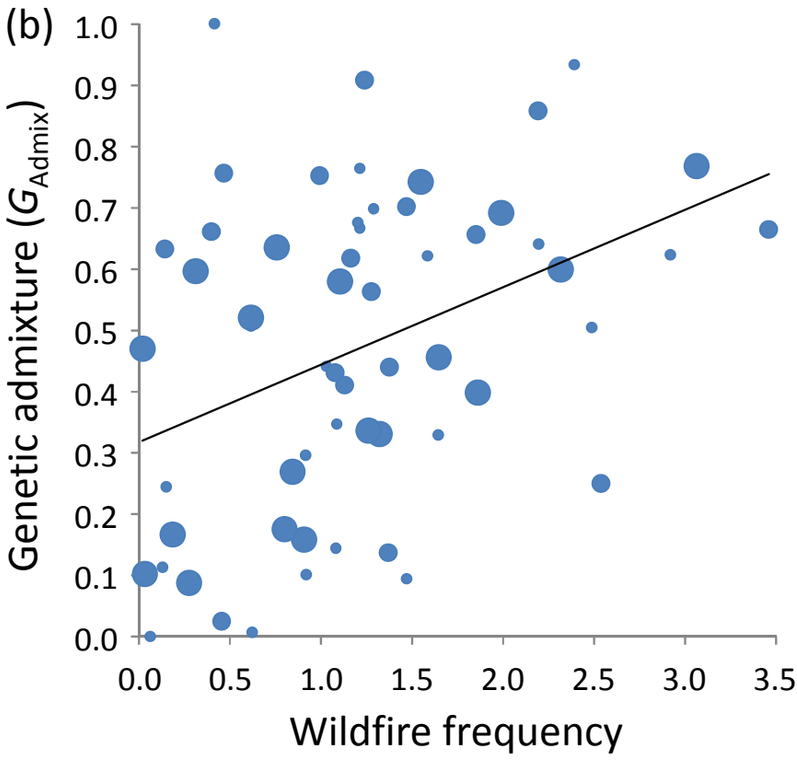

(d)

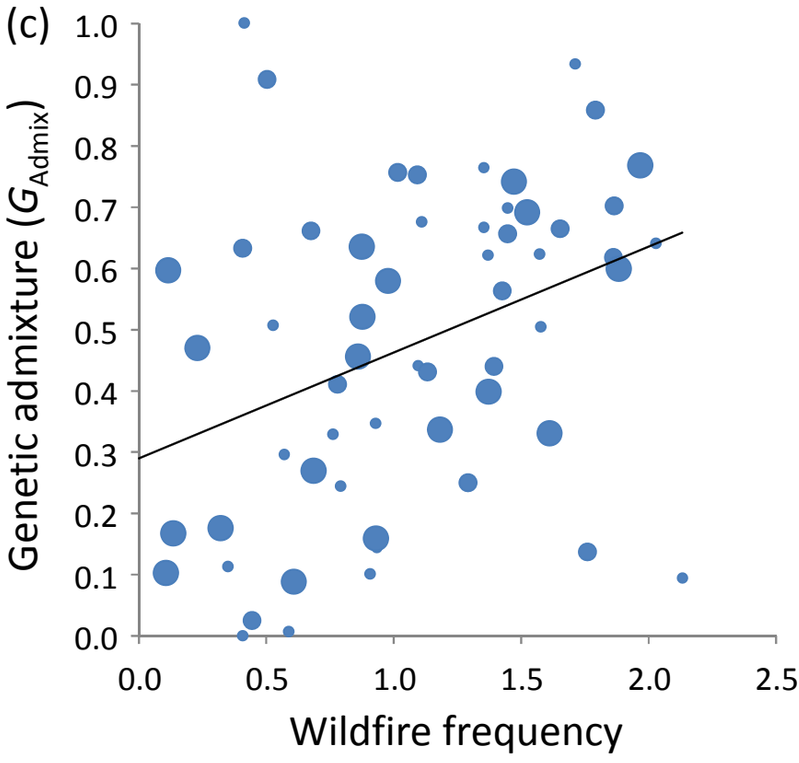

(a)
PC 2 (1.38 % inertia)
PC 1 (1.27 % inertia)
Q. berberidifolia
Q. durata
Hybrids


(b)
PC 2 (1.38 % inertia)
PC 1 (1.27 % inertia)
Q. berberidifolia
Q. durata
Hybrids


(c)
PC 2 (1.54 % inertia)
PC 1 (1.55 % inertia)
Cluster 1 (allopatric)
Cluster 2 (parapatric)
Admixed


(a)
Genetic admixture (G_{Admix})
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0
1.0
2.0
3.0
4.0
5.0
6.0
Wildfire frequency


(b)
Genetic admixture (G_{Admix})
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Wildfire frequency


(c)
Genetic admixture (G_{Admix})
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0
0.5
1.0
1.5
2.0
2.5
Wildfire frequency