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

4

5 Joaquín Ortego¹, Paul F. Gugger² and Victoria L. Sork^{3,4}

6

7 ¹Department of Integrative Ecology, Estación Biológica de Doñana, EBD-CSIC, Avda.
8 Américo Vespucio s/n, E-41092 Seville, Spain; ²University of Maryland Center for
9 Environmental Science, Appalachian Laboratory, 301 Braddock Road, Frostburg, MD
10 21532, USA; ³Department of Ecology and Evolutionary Biology, University of
11 California, Box 957239, Los Angeles, CA 90095, USA; ⁴Institute of the Environment
12 and Sustainability, University of California, Box 951496, Los Angeles, CA 90095-
13 1496, USA

14

15

16 Author for correspondence:

17 Joaquín Ortego

18 Estación Biológica de Doñana, EBD-CSIC,

19 Avda. Américo Vespucio s/n, E-41092 Seville, Spain

20 E-mail: joaquin.ortego@csic.es

21 Phone: +34 954 232 340

22

23 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.

49 **Introduction**

50

51 Hybridization has attracted the attention of biologists for a long time (Linnaeus, 1735;
52 Darwin, 1859; Anderson & Stebbins, 1954) and is widely recognized as an important
53 evolutionary force involved in adaptation to novel environmental conditions (Lewontin
54 & Birch, 1966; Morjan & Rieseberg, 2004; Baskett & Gomulkiewicz, 2011) and
55 speciation (Anderson, 1948; Rieseberg, 1997; Abbott *et al.*, 2013). Despite the potential
56 benefits of hybridization to biodiversity, this phenomenon has also been linked to the
57 disruption of local adaptation and species loss (Rhymer & Simberloff, 1996). This is of
58 particular concern when exotic taxa are involved, as hybridization can potentially
59 increase their invasiveness and lead to the extinction of native species (Ellstrand &
60 Schierenbeck, 2000; Prentis *et al.*, 2007). Beyond introductions of exotic species by
61 humans, different forms of anthropogenic disturbance have also been hypothesized to
62 increase hybridization rates among native, exotic and native-exotic taxa (Crispo *et al.*,
63 2011; Guo, 2014). The disturbance hypothesis postulates that habitat alterations
64 increase opportunities for hybridization via the breakdown of premating reproductive
65 isolating mechanisms and/or creating environmental gradients with novel or
66 intermediate niches where hybrids outperform parental species (Anderson & Stebbins,
67 1954; see also Anderson, 1948, 1953). Anthropogenic disturbances that have been
68 linked with increased hybridization rates include land use alterations (Lamont *et al.*,
69 2003; Thompson *et al.*, 2010; Hoban *et al.*, 2012; Guo, 2014), climate change (Muhlfeld
70 *et al.*, 2014) and, more counterintuitively, the disruption of natural disturbances that
71 promote reproductive isolation and maintain species boundaries (e.g. suppression of
72 natural wildfires; King *et al.*, 2015; Stewart *et al.*, 2015). Support for this hypothesis has
73 been found across multiple taxa (“hybrid richness”) at a continental scale (Guo, 2014)

74 and within pairs of interbreeding species at local (Hasselman *et al.*, 2014), regional
75 (Thompson *et al.*, 2010; Muhlfeld *et al.*, 2014) and range-wide scales (Hoban *et al.*,
76 2012).

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

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

99 of interspecific hybrids (Forde & Faris, 1962; eFloras, 2015; Ortego *et al.*, 2015a). A
100 previous study on the Californian scrub white oak species complex showed that these
101 two species probably diverged in peripatry or sympatry ~23-26 k years BP, the oldest
102 split among the three pairs of sister taxa within the complex, and supported the
103 monophyletic origin of *Q. durata* (Ortego *et al.*, 2015a). In this study, we primarily aim
104 to investigate the impacts of different sources of environmental disturbance on
105 hybridization rates between these two shrub species (Anderson, 1948; Anderson &
106 Stebbins, 1954), with particular emphasis on the potential role of human-induced
107 landscape alterations (e.g. Thompson *et al.*, 2010; Guo, 2014). California offers an
108 excellent setting to address this question. It is a climatically and geologically complex
109 region, a fact that has been linked with high rates of interspecific hybridization (Dodd &
110 Afzal-Rafii, 2004; Guo, 2014; Ortego *et al.*, 2014a), speciation (Calsbeek *et al.*, 2003;
111 Lancaster & Kay, 2013) and local adaptation processes across many organisms and
112 spatiotemporal scales (Ortego *et al.*, 2012; Langin *et al.*, 2015). Much of this region is
113 also highly impacted by urbanization and anthropogenic habitat degradation, which has
114 altered population connectivity in many organisms (e.g. Riley *et al.*, 2006; Vandergast
115 *et al.*, 2007) and increased risk of extinction in several species (Myers *et al.*, 2000;
116 Schwartz *et al.*, 2006; Vandergast *et al.*, 2008). Beyond urbanization and agriculture,
117 wildfires are also an important source of environmental disturbance in Californian
118 ecosystems and their occurrence and impact have steadily increased since European
119 settlement associated with growing population densities, urbanization, and human-
120 induced climate change (Westerling *et al.*, 2006; Syphard *et al.*, 2007; Moritz *et al.*,
121 2014). Despite the fact that natural wildfire regimes are an important component of
122 Californian ecosystems (Rundel, 1982; Keeley *et al.*, 2012), their increased frequency
123 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 breadth 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.

Williams *et al.*, 2001; Dodd & Afzal-Rafii, 2004; Alberto *et al.*, 2010; Ortego *et al.*, 2014a). We analyzed whether the contribution of these factors to explain population-level genetic admixture varies with the spatial scale at which they are measured, an issue that has been seldom addressed in the context of hybridization studies, despite the possibility that it can have a non-negligible impact on the inferences obtained (Barton & Hewitt, 1985; Harrison, 1986; Barton, 2001; Buggs, 2007), particularly in wind-pollinated species with large scales of dispersal (e.g. Dow & Ashley, 1998; Buschbom *et al.*, 2011).

Materials and Methods

Population sampling

Between 2010 and 2014, we sampled 529 reproductive individuals (i.e. we did not sample seedlings or saplings) from a total of 58 localities in California (Fig. 1; Supporting Information Table S1). Plants were morphologically identified in the field and tentatively assigned to parental taxa or hybrids according with available guides and floras (Roberts, 1995; eFloras, 2015). We aimed to collect samples from populations located across the entire distribution range of *Q. berberidifolia* Liebm. and *Q. durata* Jepson and designed sampling using occurrence records available in the Calflora database (<http://www.calflora.org/>). Some sampled individuals ($n = 78$) were genetically identified as hybrids with other taxa within the scrub white oak species complex (Ortego *et al.*, 2015a) and were excluded from subsequent analyses. Thus, all analyses presented in this study are based on 451 individuals of *Q. berberidifolia*, *Q. 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 Cunningham *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 the235 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 ($\sim 50 \text{ m}^2$ 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.

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

281 We obtained soil data for each sampling locality from the SSURGO datasets
282 available at the Web Soil Survey from the United States Department of Agriculture
283 (USDA) (<http://websoilsurvey.sc.egov.usda.gov/>; Staff, 2012). In this case, we only
284 obtained soil order data (based on USDA Soil Taxonomy categories) for each sampling
285 locality (i.e. this parameter was not estimated at multiple spatial scales) given that the
286 successful establishment of seedlings is only expected to be influenced by soil
287 properties a few meters around the germination site (e.g. Wright, 2007; Langhans *et al.*,
288 2009). All GIS calculations were performed in ARCMAP 10.2.1 (ESRI, Redlands, CA,
289 USA).

290

291 Analyses of genetic admixture

292

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

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

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

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

337

338 Results

339

340 Hybrid identification and genetic structure

341

342 Considering the dataset simulated on the basis of purebred individuals identified by
343 STRUCTURE analyses, the assignment of purebred and hybrid individuals to their correct
344 class was 93% for NEWHYBRIDS and 96% for STRUCTURE (Fig. S2a, b and Table S3a).
345 Similarly, considering the dataset simulated on the basis of purebred individuals
346 identified by NEWHYBRIDS analyses, the assignment of purebred and hybrid individuals
347 to their correct class was 97% for NEWHYBRIDS and 98% for STRUCTURE (Fig. S2c, d
348 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 ($\chi^2 =$
 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 ($\chi^2 = 24.23$; P
 394 < 0.001), populations located in areas where the range of both species overlap ($\chi^2 =$
 395 23.89; $P < 0.001$) or populations from allopatric areas ($\chi^2 = 3.03$; $P = 0.08$). 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 ($\chi^2 = 28.26$; $P < 0.001$) and $K = 3$ ($n = 57$) ($\chi^2 = 74.26$; $P < 0.001$).

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 $P_s < 0.001$), 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).

421

Bayesian comparison of gene flow models

423

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

436

437 Factors associated with genetic admixture

438

439 Only wildfire frequency and range overlap were consistently included in models of
440 genetic admixture (G_{Admix}) for all spatial scales (Table 2 and Table S4). Genetic
441 admixture tended to be higher in populations located in areas where the distributions
442 ranges of both species overlap, but these differences only were marginally significant
443 (i.e. unconditional CIs crossed zero; Table 3). Accordingly, the frequency of hybrids (Q
444 < 0.90) was significantly lower in areas of allopatry than in the vast region where the
445 ranges of both species overlap ($\chi^2 = 25.88$, $P < 0.001$). Genetic admixture was
446 positively associated with wildfire frequency at all spatial scales (Table 3 and Table S5;
447 Fig. 4). Although wildfire frequency always had a significant effect, its effect size was
448 higher at the two largest spatial scales (Table 3 and Table S5). All other variables were

not included in any model (climate) or were included in some of them but showed no significant effects (elevation, proportion of habitats disturbed by agriculture and human development, soil category, latitude, longitude) (Tables 2-3 and Supporting Information Tables S4-5). Analyses based on soil suborder category (rather than order category) provided analogous results as well as analyses exclusively focused on populations located in areas where the distribution range of both species overlap (data not shown).

Discussion

Our STRUCTURE and NEWHYBRIDS analyses indicate considerable rates of hybridization between *Q. durata* and *Q. berberidifolia* across their distribution ranges, which is consistent with patterns found at local scales in a previous morphology-based study (Forde & Faris, 1962). Although the two species are expected to share some alleles due to common ancestry, several lines of evidence support that the observed patterns of genetic admixture have resulted from genuine interspecific hybridization and cannot be merely explained by incomplete lineage sorting (Muir & Schlotterer, 2005; Lexer *et al.*, 2006): i) Bayesian and PCA analyses identified two genetic clusters that are in good agreement with the two morphological species and (ii) simulations demonstrated a high performance of our set of markers to correctly identify hybrids and purebred individuals; iii) We found strong differences in the rates of hybridization between populations from parapatric and sympatric areas (Fig. 2), a spatial pattern that is incompatible with ancestral polymorphism; iv) The presence within the same locality of individuals with very different levels of admixed ancestry indicates that genetic admixture is consequence of hybridization (i.e. the co-existence of purebred individuals, 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 al.*, 2008). According to STRUCTURE analyses for $K=3$, twenty-six percent of the
484 analysed individuals were identified as hybrids according with the $Q \geq 0.90$ threshold
485 criterion (Vähä & Primmer, 2006; Lepais *et al.*, 2009; Cullingham *et al.*, 2012) (Fig.
486 2a), and this figure increased to 34 % when only populations located in areas of
487 geographic range overlap were considered. NEWHYBRIDS analyses estimated a higher
488 frequency of hybrids than STRUCTURE for both the entire study area (42%) and the
489 region of parapatry (54%), which may be explained by the presence of genetic
490 substructure within *Q. berberidifolia* and the fact that NEWHYBRIDS can only
491 accommodate two genetic clusters (Anderson & Thompson, 2002). These hybridization
492 rates are similar to those reported for two other hybridizing Californian oaks (Ortego *et al.*, 2014a) but markedly higher than those found among most interbreeding taxa within
493 the genus (e.g. Craft *et al.*, 2002; Curtu *et al.*, 2007; Cavender-Bares & Pahlich, 2009).
494 The higher occurrence of hybrids in areas where the geographical ranges of both
495 parental taxa overlap confirms previous studies indicating that spatial proximity is an
496 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

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

Environmental disturbance

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

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

635

636 **Acknowledgements**

637

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

647

648 **Author contributions**

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

652

653 **References**

654

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

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

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

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

664 **Anderson E, Stebbins GL. 1954.** Hybridization as an evolutionary stimulus. *Evolution*
665 **8**(4): 378-388.

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

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

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

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

- 674 **Bacles CFE, Ennos RA. 2008.** Paternity analysis of pollen-mediated gene flow for
675 *Fraxinus excelsior* L. in a chronically fragmented landscape. *Heredity* **101**(4):
676 368-380.
- 677 **Bacles CFE, Lowe AJ, Ennos RA. 2006.** Effective seed dispersal across a fragmented
678 landscape. *Science* **311**(5761): 628-628.
- 679 **Barr KR, Kus BE, Preston KL, Howell S, Perkins E, Vandergast AG. 2015.** Habitat
680 fragmentation in coastal southern California disrupts genetic connectivity in the
681 cactus wren (*Campylorhynchus brunneicapillus*). *Molecular Ecology* **24**(10):
682 2349-2363.
- 683 **Barton NH. 2001.** The role of hybridization in evolution. *Molecular Ecology* **10**(3):
684 551-568.
- 685 **Barton NH, Hewitt GM. 1985.** Analysis of hybrid zones. *Annual Review of Ecology*
686 *and Systematics* **16**: 113-148.
- 687 **Baskett ML, Gomulkiewicz R. 2011.** Introgressive hybridization as a mechanism for
688 species rescue. *Theoretical Ecology* **4**(2): 223-239.
- 689 **Beerli P. 2006.** Comparison of Bayesian and maximum-likelihood inference of
690 population genetic parameters. *Bioinformatics* **22**(3): 341-345.
- 691 **Beerli P. 2009.** How to Use MIGRATE or Why are Marko Chain Monte Carlo Programs
692 Difficult to Use? New York, NY, USA: Cambridge University Press.
- 693 **Beerli P, Felsenstein J. 2001.** Maximum likelihood estimation of a migration matrix
694 and effective population sizes in n subpopulations by using a coalescent
695 approach. *Proceedings of the National Academy of Sciences of the United States*
696 *of America* **98**(8): 4563-4568.
- 697 **Beerli P, Palczewski M. 2010.** Unified framework to evaluate panmixia and migration
698 direction among multiple sampling locations. *Genetics* **185**(1): 313-326.

- 699 **Brady KU, Kruckeberg AR, Bradshaw HD. 2005.** Evolutionary ecology of plant
700 adaptation to serpentine soils. *Annual Review of Ecology Evolution and*
701 *Systematics* **36**: 243-266.
- 702 **Breed MF, Gardner MG, Ottewell KM, Navarro CM, Lowe AJ. 2012.** Shifts in
703 reproductive assurance strategies and inbreeding costs associated with habitat
704 fragmentation in Central American mahogany. *Ecology Letters* **15**: 444-452.
- 705 **Brooks RR. 1987.** *Serpentine and its Vegetation: A Multidisciplinary Approach.*
706 Portland, OR, USA: Dioscorides.
- 707 **Buggs RJA. 2007.** Empirical study of hybrid zone movement. *Heredity* **99**(3): 301-312.
- 708 **Burgarella C, Lorenzo Z, Jabbour-Zahad R et al. 2009.** Detection of hybrids in
709 nature: application to oaks (*Quercus suber* and *Q. ilex*). *Heredity* **102**(5), 442-
710 452.
- 711 **Buschbom J, Yanbaev Y, Degen B. 2011.** Efficient long-distance gene flow into an
712 isolated relict oak stand. *Journal of Heredity* **102**(4): 464-472.
- 713 **Calsbeek R, Thompson JN, Richardson JE. 2003.** Patterns of molecular evolution
714 and diversification in a biodiversity hotspot: the California Floristic Province.
715 *Molecular Ecology* **12**(4): 1021-1029.
- 716 **Cavender-Bares J, Pahlich A. 2009.** Molecular, morphological and ecological niche
717 differentiation of sympatric sister oak species, *Quercus virginiana* and *Q.*
718 *geminata* (Fagaceae). *American Journal of Botany* **96**(9): 1690-1702.
- 719 **Cerdà A, Robichaud P. 2009.** *Fire Effects on Soils and Restoration Strategies.* Enfield,
720 NH, USA: Science Publishers.
- 721 **Craft KJ, Ashley MV, Koenig WD. 2002.** Limited hybridization between *Quercus*
722 *lobata* and *Quercus douglasii* (Fagaceae) in a mixed stand in central coastal
723 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 Serpentes: 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.

- 822 **Langin KM, Sillett TS, Funk WC, Morrison SA, Desrosiers MA, Ghalambor CK.**
823 **2015.** Islands within an island: Repeated adaptive divergence in a single
824 population. *Evolution* **69**(3): 653-665.
- 825 **Lepais O, Gerber S. 2011.** Reproductive patterns shape introgression dynamics and
826 species succession within the European white oak species complex. *Evolution*
827 **65**(1): 156-170.
- 828 **Lepais O, Petit RJ, Guichoux E, Lavabre JE, Alberto F, Kremer A, Gerber S.**
829 **2009.** Species relative abundance and direction of introgression in oaks.
830 *Molecular Ecology* **18**(10): 2228-2242.
- 831 **Lewontin RC, Birch LC. 1966.** Hybridization as a source of variation for adaptation
832 to new environments. *Evolution* **20**(3): 315-336.
- 833 **Lexer C, Kremer A, Petit RJ. 2006.** Shared alleles in sympatric oaks: recurrent gene
834 flow is a more parsimonious explanation than ancestral polymorphism.
835 *Molecular Ecology* **15**(7): 2007-2012.
- 836 **Linnaeus C. 1735.** *Systema Naturae 1st edition*. Leiden, the Netherlands: Theodorum
837 Haak.
- 838 **Moritz MA, Batllori E, Bradstock RA, Gill AM, Handmer J, Hessburg PF,**
839 **Leonard J, McCaffrey S, Odion DC, Schoennagel T, et al. 2014.** Learning to
840 coexist with wildfire. *Nature* **515**(7525): 58-66.
- 841 **Morjan CL, Rieseberg LH. 2004.** How species evolve collectively: implications of
842 gene flow and selection for the spread of advantageous alleles. *Molecular*
843 *Ecology* **13**(6): 1341-1356.
- 844 **Muhlfeld CC, Kovach RP, Jones LA, Al-Chokhachy R, Boyer MC, Leary RF,**
845 **Lowe WH, Luikart G, Allendorf FW. 2014.** Invasive hybridization in a

- 846 threatened species is accelerated by climate change. *Nature Climate Change*
847 **4**(7): 620-624.
- 848 **Muir G, Schlotterer C. 2005.** Evidence for shared ancestral polymorphism rather than
849 recurrent gene flow at microsatellite loci differentiating two hybridizing oaks
850 (*Quercus* spp.). *Molecular Ecology* **14**(2): 549-561.
- 851 **Muller CH. 1952.** Ecological control of hybridization in *Quercus* - A factor in the
852 mechanism of evolution. *Evolution* **6**(2): 147-161.
- 853 **Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000.**
854 Biodiversity hotspots for conservation priorities. *Nature* **403**(6772): 853-858.
- 855 **Ortego J, Riordan EC, Gugger PF, Sork VL. 2012.** Influence of environmental
856 heterogeneity on genetic diversity and structure in an endemic southern
857 Californian oak. *Molecular Ecology* **21**(13): 3210-3223.
- 858 **Ortego J, Gugger PF, Riordan EC, Sork VL. 2014a.** Influence of climatic niche
859 suitability and geographical overlap on hybridization patterns among southern
860 Californian oaks. *Journal of Biogeography* **41**(10): 1895-1908.
- 861 **Ortego J, Bonal R, Munoz A, Aparicio JM. 2014b.** Extensive pollen immigration and
862 no evidence of disrupted mating patterns or reproduction in a highly fragmented
863 holm oak stand. *Journal of Plant Ecology* **7**(4): 384-395.
- 864 **Ortego J, Noguerales V, Gugger PF, Sork VL. 2015a.** Evolutionary and demographic
865 history of the Californian scrub white oak species complex: An integrative
866 approach. *Molecular Ecology* **24**(24): 6188-6208.
- 867 **Ortego J, Gugger PF, Sork VL. 2015b.** Climatically stable landscapes predict patterns
868 of genetic structure and admixture in the Californian canyon live oak. *Journal of*
869 *Biogeography* **42**(2): 328-338.

- 870 **Ortego J, Aguirre MP, Nogueras V, Cordero PJ. 2015c.** Consequences of
871 extensive habitat fragmentation in landscape-level patterns of genetic diversity
872 and structure in the Mediterranean esparto grasshopper. *Evolutionary*
873 *Applications* **8**(6): 621-632.
- 874 **Pausas JG, Keeley JE. 2014.** Evolutionary ecology of resprouting and seeding in fire-
875 prone ecosystems. *New Phytologist*, **204**(1): 55-65.
- 876 **Pausas JG, Pratt RB, Keeley JE, Jacobsen AL, Ramirez AR, Vilagrosa A, Paula S,**
877 **Kaneakua-Pia IN, Davis SD. 2016.** Towards understanding resprouting at the
878 global scale. *New Phytologist*, **209**(3): 945-954.
- 879 **Petit RJ, Bodenes C, Ducousso A, Roussel G, Kremer A. 2004.** Hybridization as a
880 mechanism of invasion in oaks. *New Phytologist* **161**(1): 151-164.
- 881 **Pluess AR, Sork VL, Dolan B, Davis FW, Grivet D, Merg K, Papp J, Smouse PE.**
882 **2009.** Short distance pollen movement in a wind-pollinated tree, *Quercus lobata*
883 (Fagaceae). *Forest Ecology and Management* **258**(5): 735-744.
- 884 **Prentis PJ, White EM, Radford IJ, Lowe AJ, Clarke AR. 2007.** Can hybridization
885 cause local extinction: a case for demographic swamping of the Australian
886 native *Senecio pinnatifolius* by the invasive *Senecio madagascariensis*? *New*
887 *Phytologist* **176**(4): 902-912.
- 888 **Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure using
889 multilocus genotype data. *Genetics* **155**(2): 945-959.
- 890 **R Core Team. 2015.** *R: A Language and Environment for Statistical Computing.*
891 *Version 3.0.2.* URL <http://www.rproject.org>. Vienna, Austria: R Foundation for
892 Statistical Computing.

- 893 **Renaut S, Grassa CJ, Yeaman S, Moyers BT, Lai Z, Kane NC, Bowers JE, Burke**
894 **JM, Rieseberg LH. 2013.** Genomic islands of divergence are not affected by
895 geography of speciation in sunflowers. *Nature Communications* **4**: 1827.
- 896 **Rhymer JM, Simberloff D. 1996.** Extinction by hybridization and introgression.
897 *Annual Review of Ecology and Systematics* **27**: 83-109.
- 898 **Rieseberg LH. 1997.** Hybrid origins of plant species. *Annual Review of Ecology and*
899 *Systematics* **28**: 359-389.
- 900 **Riley SPD, Pollinger JP, Sauvajot RM, York EC, Bromley C, Fuller TK, Wayne**
901 **RK. 2006.** A southern California freeway is a physical and social barrier to gene
902 flow in carnivores. *Molecular Ecology* **15**(7): 1733-1741.
- 903 **Roberts FM. 1995.** *Illustrated guide to the oaks of the Southern California Floristic*
904 *Province*. Encinitas, CA, USA: F.M. Roberts Publications.
- 905 **Rundel PW. 1982.** Fire as an ecological factor. In: Lange OL, Nobel PS, Osmond CB,
906 Ziegler H, eds. *Physiological Plant Ecology I*. New York, USA: Springer-
907 Verlag, 501-538.
- 908 **Rushton BS. 1993.** Natural hybridization within the genus *Quercus* 1. *Annales des*
909 *Sciences Forestieres, Vol 50, Supplement 1, 1993 - Genetics of Oaks*: S73-S90.
- 910 **Saarman NP, Pogson GH. 2015.** Introgression between invasive and native blue
911 mussels (genus *Mytilus*) in the central California hybrid zone. *Molecular*
912 *Ecology* **24**(18): 4723-4738.
- 913 **Schwartz MW, Thorne JH, Viers JH. 2006.** Biotic homogenization of the California
914 flora in urban and urbanizing regions. *Biological Conservation* **127**(3): 282-291.
- 915 **Silliman FE, Leisner RS. 1958.** An analysis of a colony of hybrid oaks. *American*
916 *Journal of Botany* **45**(10): 730-736.

- 917 **Staff SS. 2012.** *Soil Survey Geographic (SSURGO) database for California*. URL
918 <http://websoilsurvey.sc.egov.usda.gov> [accessed on 20 December 2015].
- 919 **Starr TN, Gadek KE, Yoder JB, Flatz R, Smith CI. 2013.** Asymmetric hybridization
920 and gene flow between Joshua trees (Agavaceae: *Yucca*) reflect differences in
921 pollinator host specificity. *Molecular Ecology* **22**(2): 437-449.
- 922 **Stewart JF, Will RE, Robertson KM, Nelson CD. 2015.** Frequent fire protects
923 shortleaf pine (*Pinus echinata*) from introgression by loblolly pine (*P. taeda*).
924 *Conservation Genetics* **16**(2): 491-495.
- 925 **Streiff R, Ducousso A, Lexer C, Steinkellner H, Gloessl J, Kremer A. 1999.** Pollen
926 dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur*
927 L- and *Q. petraea* (Matt.) Liebl. *Molecular Ecology* **8**(5): 831-841.
- 928 **Syphard AD, Radeloff VC, Keeley JE, Hawbaker TJ, Clayton MK, Stewart SI,**
929 **Hammer RB. 2007.** Human influence on California fire regimes. *Ecological*
930 *Applications* **17**(5): 1388-1402.
- 931 **Thompson SL, Lamothe M, Meirmans PG, Perinet P, Isabel N. 2010.** Repeated
932 unidirectional introgression towards *Populus balsamifera* in contact zones of
933 exotic and native poplars. *Molecular Ecology* **19**(1): 132-145.
- 934 **Tsuda Y, Nakao K, Ide Y, Tsumura Y. 2015.** The population demography of *Betula*
935 *maximowicziana*, a cool-temperate tree species in Japan, in relation to the last
936 glacial period: its admixture-like genetic structure is the result of simple
937 population splitting not admixing. *Molecular Ecology* **24**(7): 1403-1418.
- 938 **Vähä JP, Primmer CR. 2006.** Efficiency of model-based Bayesian methods for
939 detecting hybrid individuals under different hybridization scenarios and with
940 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.

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

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

971

972 **Supporting Information**

973

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

976

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

979

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

981

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

984

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

986

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

988

989 **Fig. S1** Models tested using MIGRATE-N and Bayesian estimates of mutation-scaled
990 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.

Model description	Excluding hybrids			Including hybrids		
	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 <i>Q. berberidifolia</i> to <i>Q. durata</i>	-113727	-211961	<0.0001	-292242	-444446	<0.0001
(IV) Unidirectional gene flow from <i>Q. durata</i> to <i>Q. berberidifolia</i>	-7747	0	1.0000	-70019	0	1.0000
(V) Unidirectional gene flow from <i>Q. berberidifolia</i> to <i>Q. durata</i> restricted to parapatric populations	-81256	-147017	<0.0001	-295109	-450180	<0.0001
(VI) Unidirectional gene flow from <i>Q. durata</i> to <i>Q. berberidifolia</i> 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 km², (b) 100 km², and (c) 1000 km² around sampling localities]. Only best ranked equivalent models ($\Delta\text{AIC}_c \leq 2$) 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

Table 3 General linear models (GLMs) for genetic admixture (G_{Admix}) of *Quercus berberidifolia* and *Q. durata*. Parameter estimates and unconditional standard errors (USE) were obtained performing model averaging of the best ranked equivalent models ($\Delta\text{AICc} \leq 2$) (see Table 1). Variables are sorted according with their relative importance based on the sum of Akaike weights ($\Sigma \omega_i$) of those models with $\Delta\text{AICc} \leq 2$ in which the variable was present. Bold type indicates significant variables, i.e. variables for which their unconditional 95 % confidence interval (CI) did not cross zero. The percentage of explained deviance for each model is indicated in parentheses. We run three subsets of models considering the same parameters but with some of them (fire frequency, climate, proportion of habitats disturbed by agriculture and human development and elevation) measured at three different spatial scales [(a) 10 km², (b) 100 km², and (c) 1000 km² around sampling localities].

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

1036 Figure legends

1037

1038 **Fig. 1** Map from California representing wildfire frequency estimated as the number of
 1039 years that a given area was burned between 1900 and 2013 (source: CALFIRE Fire
 1040 Perimeters Geodatabase version 13.2). The map also represents sampling localities for
 1041 putative *Quercus berberidifolia* (blue stars) and *Q. durata* (red stars). Right top inset
 1042 shows the range distributions of *Q. berberidifolia* (blue) and *Q. durata* (red) according
 1043 with records of the two species available in Calflora database (<http://www.calflora.org/>).
 1044

1045 **Fig. 2** Results of genetic assignments based on the programs (a) STRUCTURE ($K = 2$ and
 1046 $K = 3$) and (b) NEWHYBRIDS. Each individual is represented by a vertical bar, which is
 1047 partitioned into K coloured segments showing the individual's probability of assignment
 1048 to *Quercus berberidifolia* (blue), *Q. durata* (red) and hybrids (yellow, only for
 1049 NEWHYBRIDS). Thin vertical black lines separate individuals from different sampling
 1050 localities arranged according to their geographical location from southeast (left) to
 1051 northwest (right) (population codes are described in Supporting Information Table S1).
 1052 The horizontal bar indicates whether the different sampled populations are located or
 1053 not in areas of geographical range overlap between the two focal studied species
 1054 (allopatry vs. parapatry). Areas of parapatry and allopatry were defined according with
 1055 records of the two species available in Calflora database (<http://www.calflora.org/>) and
 1056 verified with distribution maps from eFloras (2015). Panel (c) shows the mean ($\pm$ SD)
 1057 log probability of the data ($\ln \Pr(X|K)$) over 10 runs of STRUCTURE (left axis, black dots
 1058 and error bars) for each value of K and the magnitude of ΔK (right axis, open dots). (d)
 1059 Detail of the leaves of the two study species, *Quercus berberidifolia* (left) and *Q. durata*
 1060 (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 km², (b) 100 km²,
1073 and (c) 1000 km² 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

Figure 2

Figure 3

1089 Figure 4

1090

1091