

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

Joaquín Ortego<sup>1</sup>, Paul F. Gugger<sup>2</sup> and Victoria L. Sork<sup>3,4</sup>

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

Author for correspondence:

Joaquín Ortego

Tel: +34 954 232 340

Email: joaquin.ortego@csic.es

Received: 8 June 2016

Accepted: 4 August 2016

*New Phytologist* (2017) **213**: 942–955

doi: 10.1111/nph.14182

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

## 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 (*Quercus 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.

## 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.*, 2014b)

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 of 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, probably 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.*, 2015c) 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.*, 2015c). A previous study on the Californian scrub white oak species complex showed that these two species probably diverged in peripatry or sympatry *c.* 23–26 kyr before present (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.*, 2015c). 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.*, 2014b), 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

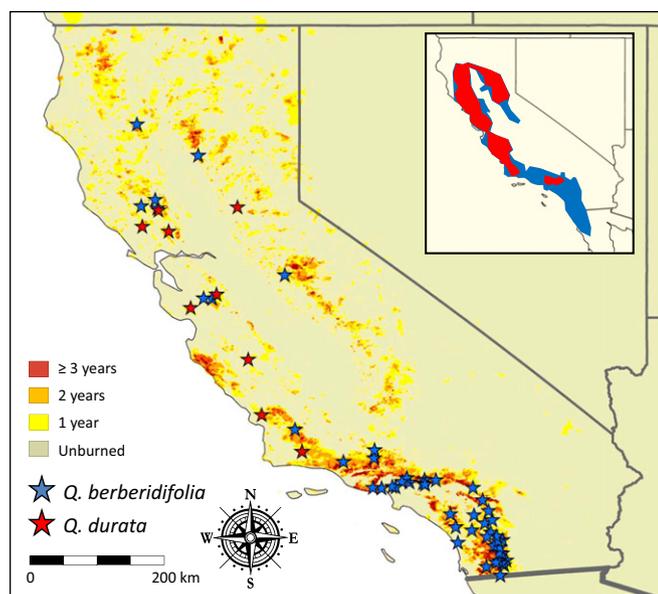
linked with population declines in several species already impacted by other sources of habitat loss and degradation (Syphard *et al.*, 2007; Barr *et al.*, 2015). These different sources of environmental disturbance, together with the high richness of native and exotic species in the region (Calsbeek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

The overall goal of this study is to investigate the extent to which wildfire frequency, human-induced habitat transformation, and local environmental factors influence hybridization between *Q. berberidifolia* and *Q. durata*. Using genotypic data for 58 stands (> 400 individuals) sampled across California we tested two specific hypotheses. First, we estimated hybridization rates and direction of gene flow between the two focal species and tested whether the adaptation of *Q. durata* to serpentine soils results in asymmetric gene flow from this species into *Q. berberidifolia*, which would be expected if hybrids show a lower performance in serpentine than in nonserpentine soils as a result of a higher niche breadth of serpentine-adapted plants (i.e. more tolerant to different soil types; Whittaker, 1954; Kruckeberg, 1984; Wright, 2007). Then, we studied the potential role of different ecological factors on spatial patterns of hybridization, primarily focusing on analyzing the impacts of human-induced environmental disturbance (Anderson, 1948; Anderson & Stebbins, 1954; Arnold, 1997). In particular, we tested the hypothesis that hybridization increases with two sources of environmental disturbance, namely anthropogenic large-scale destruction of natural habitats (urbanization and agriculture) and wildfire frequency. We addressed this main question controlling for other potential drivers of hybridization, such as species co-occurrence (range overlap) and environment (climate, elevation, soil type), that have previously been 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.*, 2014b). 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 seldom been addressed in the context of hybridization studies, despite the possibility that it can have a nonnegligible 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 to available guides and floras (Roberts, 1995; eFloras, 2015). We aimed to collect samples from populations located across the entire distribution range of *Quercus berberidifolia* Liebmann and



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

*Quercus 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.*, 2015c) 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).

### Microsatellite genotyping

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

### Genetic structure, hybrid identification and admixture analyses

We identified hybrid and purebred individuals in our empirical dataset using the Bayesian Markov chain Monte Carlo clustering analyses implemented in the programs STRUCTURE v.2.3.3 (Pritchard *et al.*, 2000; Falush *et al.*, 2003; Hubisz *et al.*, 2009) and NEWHYBRIDS v.1.1 (Anderson & Thompson, 2002). In STRUCTURE, the posterior probability ( $q$ ) describes the proportion of an individual genotype originating from each of  $K$  clusters. In NEWHYBRIDS,  $q$  describes the probability that an individual

belongs to each of six different genotypic groups that include two parental species and four hybrid classes ( $F_1$ ,  $F_2$  and first generation backcrosses). The assignment to a specific hybrid class is often uncertain in NEWHYBRIDS (Burgarella *et al.*, 2009). Thus, as done in previous studies, we summed  $q$ -values over all hybrid genotype frequency classes (e.g. Cullingham *et al.*, 2011; Hasselman *et al.*, 2014; Haines *et al.*, 2016). Given that NEWHYBRIDS can only accommodate two species/clusters and our analyses support genetic substructure within *Q. berberidifolia* (see the Results section), we conservatively based subsequent analyses on estimates of admixed ancestry obtained from STRUCTURE (e.g. Haines *et al.*, 2016). Details of STRUCTURE and NEWHYBRIDS settings are presented in Notes S1.

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

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

### Hybrid simulation and genetic assignment

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

### Bayesian comparison of gene flow models

We used MIGRATE-N v.3.6.11 to test different scenarios of gene flow between our two focal species (Beerli & Felsenstein, 2001; Beerli, 2006; Beerli & Palczewski, 2010). This program estimates mutation-scaled effective population size ( $\theta = 4N_e\mu$ , where  $N_e$  = effective population size and  $\mu$  = mutation rate per generation) and migration rates ( $M = m/\mu$ , where  $m$  = migration rate) for multiple populations in a coalescent framework by which alleles are traced back in time to a single ancestral copy (the most recent common ancestor, MRCA) (Beerli, 2009). STRUCTURE analyses on our empirical database revealed the presence of three genetic groups, one corresponding to *Q. durata* and the other

two corresponding to *Q. berberidifolia* (see the Results section for more details). Although the two genetic clusters identified within *Q. berberidifolia* have a high degree of spatial genetic admixture, they roughly separate populations located in southern California (south of the Transverse Ranges) and parapatric populations with *Q. durata* in the north (see the Results section). Thus, we used MIGRATE-N to test six models that considered different patterns of gene flow among three population groups defined *a priori*, namely populations of *Q. durata* and the two clusters of *Q. berberidifolia* identified by STRUCTURE analyses (Fig. S1a). The two population groups within *Q. berberidifolia* were defined according to the probabilities of genetic membership inferred by STRUCTURE analyses (see the Results section) and considering whether the populations were located or not in areas overlapping with the distribution range of *Q. durata* (i.e. north and south of the Transverse Ranges, respectively). Further, we tested two different sets of models: one only including purebred individuals for each taxon and cluster according to STRUCTURE analyses ( $Q \geq 0.90$ ; see the Results section for more details) and another considering both purebred and hybrid individuals (e.g. Field *et al.*, 2011; Andrew *et al.*, 2012; Starr *et al.*, 2013). Hybrids were assigned to each of the three groups according to majority population genetic assignment to each group ( $Q > 0.5$ ) according to STRUCTURE analyses. Details of MIGRATE-N settings are presented in Notes S3.

### GIS analyses

We obtained information from wildfire frequency using the CALFIRE Fire Perimeters Geodatabase v.13.2 (<http://frap.cdf.ca.gov/>). Briefly, we transformed the vector layer containing the polygons delimiting wildfire perimeters for each year from 1900 to 2013 (114 yr) into one raster layer  $\text{yr}^{-1}$  in which burned areas (c. 50 m<sup>2</sup> pixels) were given a pixel value equal to one. Then, we summed all year-based raster layers in order to generate a new raster containing information on the number of yr that each pixel has been affected by wildfires. Finally, we calculated average wildfire frequency within a circular area of 10, 100 and 1000 km<sup>2</sup> around each sampling locality, which allowed us to assess the potential impact of spatial scale on our results (see the 'Analyses of genetic admixture' subsection).

We estimated the proportion of habitats disturbed by agriculture and human development from the Conterminous United States Land Cover 200 m resolution layer ([http://nationalmap.gov/small\\_scale/atlasftp.html](http://nationalmap.gov/small_scale/atlasftp.html)). We considered as disturbed areas those devoted to different forms of agriculture (categories 61, 71, 81, 82, 83 and 84), urbanization (categories 21, 22, 23, and 85) and mining (category 32) (see land cover layer legend for the description of the different categories). The proportion of disturbed 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 v.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' ( $G_{\text{Admix}}$ ), calculated as described in Ortego *et al.* (2015b).  $G_{\text{Admix}}$  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 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  $G_{\text{Admix}}$  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 to 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* was 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 nondisturbed 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.*, 2014b).

We analyzed the data using general linear models (GLMs) with a Gaussian error structure and identity link function as implemented in the R v.3.0.0 package LME4 (R Core Team, 2015). The precision of  $G_{\text{Admix}}$  estimates may differ among populations as a result of differences in sample sizes and we took this into account using a weighted least-squares 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, 2015) as detailed in Ortego *et al.* (2015a). 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 subsection. 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

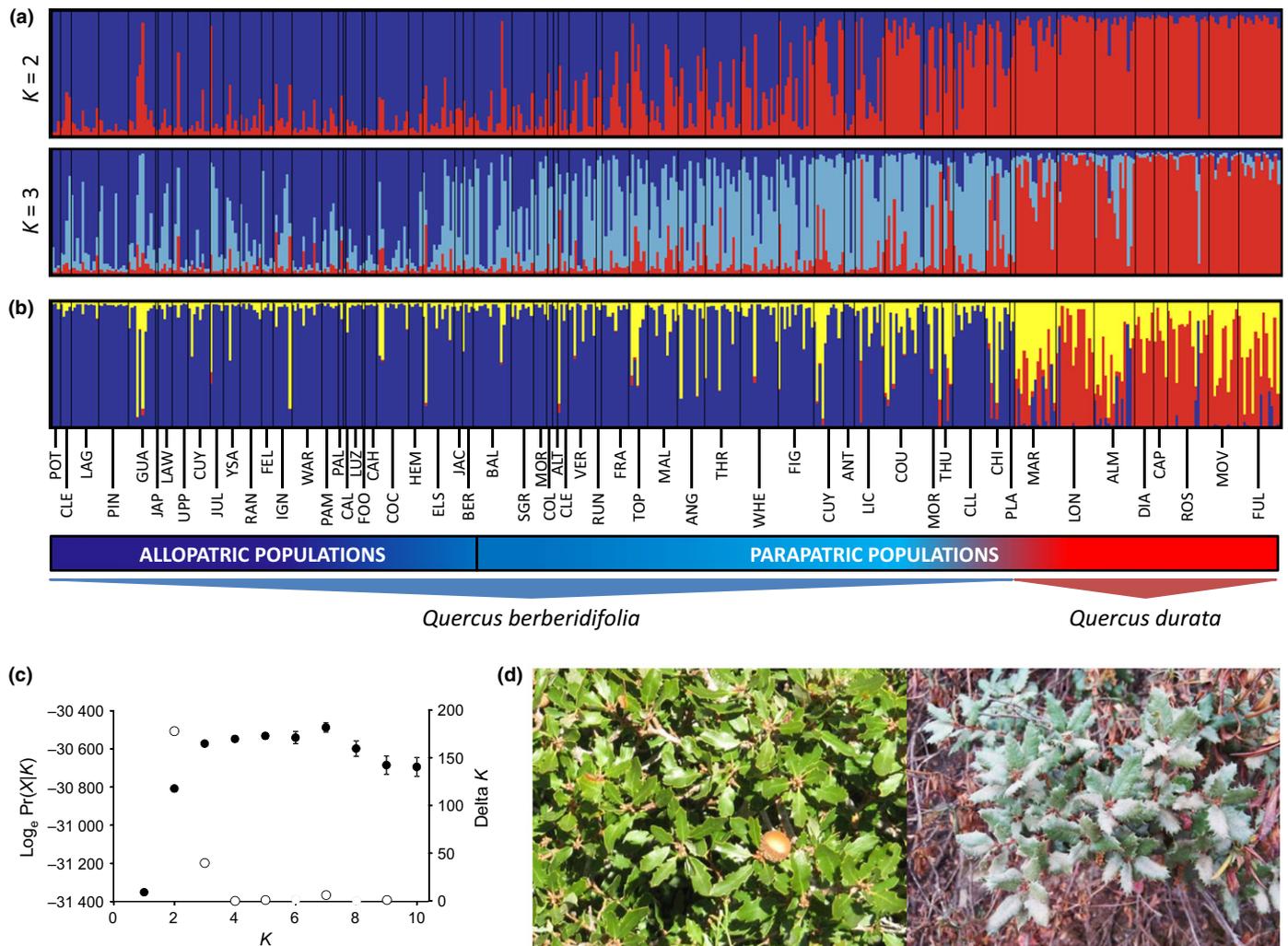
Considering the dataset simulated on the basis of purebred individuals identified by STRUCTURE analyses, the assignment of purebred and hybrid individuals to their correct class was 93% for NEWHYBRIDS and 96% for STRUCTURE (Fig. S2a,b; Table S3a). Similarly, considering the dataset simulated on the basis 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; 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 the idea that a threshold value of  $Q = 0.9$  allows purebred individuals to be differentiated correctly 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 ( $\text{Pr}(X|K)$ ) 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  $\Delta K$  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.*, 2015c). 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 ( $\chi^2 = 40.55$ ;  $P < 0.001$ ), populations located in areas where the range of both species overlap ( $\chi^2 = 17.31$ ;  $P < 0.001$ ) or populations from allopatric areas ( $\chi^2 = 30.22$ ;  $P < 0.001$ ). For  $K = 2$ , a considerable number of individuals morphologically identified in the field as *Q. berberidifolia* were assigned to purebred *Q. durata* ( $Q \geq 0.90$ ;  $n = 19$ ) or had a much higher probability of membership to *Q. durata* cluster than to *Q. berberidifolia* cluster ( $Q \geq 0.8$ ;  $n = 41$ ). For  $K = 3$ , the number of misclassified individuals was much lower: one individual identified in the field as *Q. berberidifolia* was assigned to purebred *Q. durata* and two individuals morphologically identified as *Q. durata* were genetically assigned to purebred *Q. berberidifolia*. Accordingly, only a few individuals morphologically identified as *Q. durata* ( $n = 3$ ) and *Q. berberidifolia* ( $n = 4$ ) had a high probability of genetic membership ( $Q \geq 0.8$ ) of the other species. Thus, the number of misclassified individuals was much higher for  $K = 2$  than for  $K = 3$ , considering either  $Q \geq 0.90$  ( $\chi^2 = 11.93$ ;  $P < 0.001$ ) or  $Q \geq 0.8$  ( $\chi^2 = 25.43$ ;  $P < 0.001$ ). For these reasons,  $K = 3$  was regarded as a biologically more meaningful clustering solution than  $K = 2$  and was considered for subsequent analyses.

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

Principal component 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 nonsignificant



**Fig. 2** Results of genetic assignments based on the programs (a) STRUCTURE ( $K = 2$  and  $K = 3$ ); (b) NEWHYBRIDS. Each individual is represented by a vertical bar, which is partitioned into  $K$  colored segments showing the individual's probability of assignment to *Quercus berberidifolia* (blue), *Quercus 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 to records of the two species available in the Calflora database (<http://www.calflora.org/>) and verified using distribution maps from eFloras (2015). (c) Mean ( $\pm$  SD) log probability of the data ( $\log_{10} \text{Pr}(X|K)$ ) over 10 runs of STRUCTURE (left axis, black dots and error bars) for each value of  $K$  and the magnitude of  $\Delta K$  (right axis, open dots). (d) Detail of the leaves of the two study species, *Q. berberidifolia* (left) and *Q. durata* (right) (photographs by Joaquín Ortego).

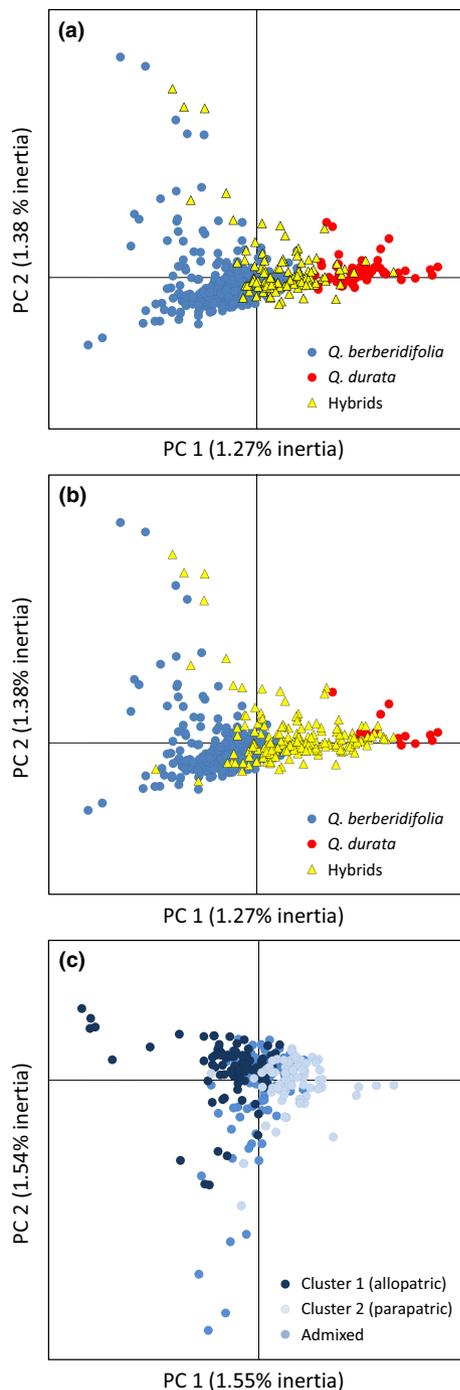
pairwise comparisons were those between *Q. berberidifolia* and *Q. durata* and between *Q. durata* and hybrid individuals for PC2 ( $P > 0.05$ ). The two genetic clusters within *Q. berberidifolia* identified by STRUCTURE analyses for  $K = 3$  were also well separated along PC1 of a PCA only including 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 pairwise comparisons for PC1 were highly significant (all  $P < 0.001$ ), 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.*, 2015c) 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



**Fig. 3** Principal component analyses (PCA) of genetic data of *Quercus berberidifolia*, *Quercus durata* and their hybrids. (a, b) PCA including all individuals and indicating their assignment to either parental species (red and blue circles) or their hybrids (yellow triangles) according to: (a) STRUCTURE analyses for  $K = 3$ ; (b) NEWHYBRIDS (considering a threshold of  $Q \geq 0.9$ ). (c) PCA only including purebred ( $Q \geq 0.9$ ) individuals of *Q. berberidifolia* and indicating their assignment to the two clusters identified within this species according to STRUCTURE analyses for  $K = 3$ .

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 hybrid 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 that parameter estimates for models including hybrids must be interpreted with caution (Fig. S1b). Convergence issues in models including hybrid individuals may be related to 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 (Tables 2, S4). Genetic admixture tended to be higher in populations located in areas where the distribution ranges of both species overlap, but these differences were only 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 (Tables 3, S5; Fig. 4). Although wildfire frequency always had a significant effect, its effect size was higher at the two largest spatial scales (Tables 3, 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, S4, S5). 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 as a result of common ancestry, several lines of evidence support that the observed patterns of genetic admixture have resulted from genuine interspecific hybridization and cannot merely be explained by incomplete lineage sorting (Muir & Schlotterer, 2005; Lexer *et al.*, 2006): Bayesian and PCA analyses identified two genetic clusters that are in good agreement with the two morphological species, and simulations demonstrated a high performance of our set of markers to correctly identify hybrids and purebred individuals; 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; the presence within the same locality of individuals with very different levels of admixed ancestry indicates that genetic admixture is a consequence of hybridization (i.e. the coexistence of purebred individuals, first-generation hybrids and backcrosses), as incomplete lineage sorting would

**Table 1** Model description and results from model comparison in MIGRATE-N for analyses excluding and including interspecific hybrids between *Quercus berberidifolia* and *Quercus durata* identified by STRUCTURE analyses ( $Q < 0.90$ ) (see also Supporting Information Fig. S1a)

Model description	Excluding hybrids			Including hybrids		
	Bézier	LBF	Probability	Bézier	LBF	Probability
(1) Full migration model	-270 658	-525 823	<0.0001	-415 896	-691 754	<0.0001
(2) Bidirectional interspecific gene flow restricted to parapatric populations	-140 090	-264 686	<0.0001	-255 323	-370 607	<0.0001
(3) Unidirectional gene flow from <i>Q. berberidifolia</i> to <i>Q. durata</i>	-113 727	-211 961	<0.0001	-292 242	-444 446	<0.0001
(4) Unidirectional gene flow from <i>Q. durata</i> to <i>Q. berberidifolia</i>	<b>-7747</b>	<b>0</b>	<b>1.0000</b>	<b>-70 019</b>	<b>0</b>	<b>1.0000</b>
(5) Unidirectional gene flow from <i>Q. berberidifolia</i> to <i>Q. durata</i> restricted to parapatric populations	-81 256	-147 017	<0.0001	-295 109	-450 180	<0.0001
(6) Unidirectional gene flow from <i>Q. durata</i> to <i>Q. berberidifolia</i> restricted to parapatric populations	-16 421	-17 347	<0.0001	-115 182	-90 325	<0.0001

Bézier approximation scores of log marginal likelihoods, log Bayes factors (LBFs) and model probabilities are shown. The best supported model is indicated in bold.

have resulted in a nearly identical background level of admixed ancestry across all individuals from the same population (e.g. Fig. 2 in Tsuda *et al.*, 2015).

STRUCTURE supports the presence of two main clusters corresponding with the two parental species, but these analyses also point to a south to north cline of genetic subdivision within *Q. berberidifolia* that seems to be the biologically most meaningful clustering solution to explain spatial patterns of genetic admixture and species boundaries (Fig. 2a; Ortego *et al.*, 2015c). The two clusters of *Q. berberidifolia* roughly separate populations north and south of the Transverse Ranges, a geographic barrier that has been frequently identified to be associated with phylogeographic breaks in many other Californian taxa (Calsbeek

**Table 2** Model selection to assess the relationship between genetic admixture ( $G_{Admix}$ ) of *Quercus berberidifolia* and *Quercus durata* and species range overlap (A), climate (B), elevation (C), soil type (D), proportion of habitats disturbed by agriculture and human development (E), and wildfire frequency (F)

Model no.	Model	K	AIC <sub>c</sub>	ΔAIC <sub>c</sub>	$\omega_i$
<b>(a) <math>G_{Admix}</math> (10 km<sup>2</sup>)</b>					
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>(b) <math>G_{Admix}</math> (100 km<sup>2</sup>)</b>					
1	F	3	7.41	0.00	0.19
2	A + F	4	8.42	1.01	0.11
<b>(c) <math>G_{Admix}</math> (1000 km<sup>2</sup>)</b>					
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

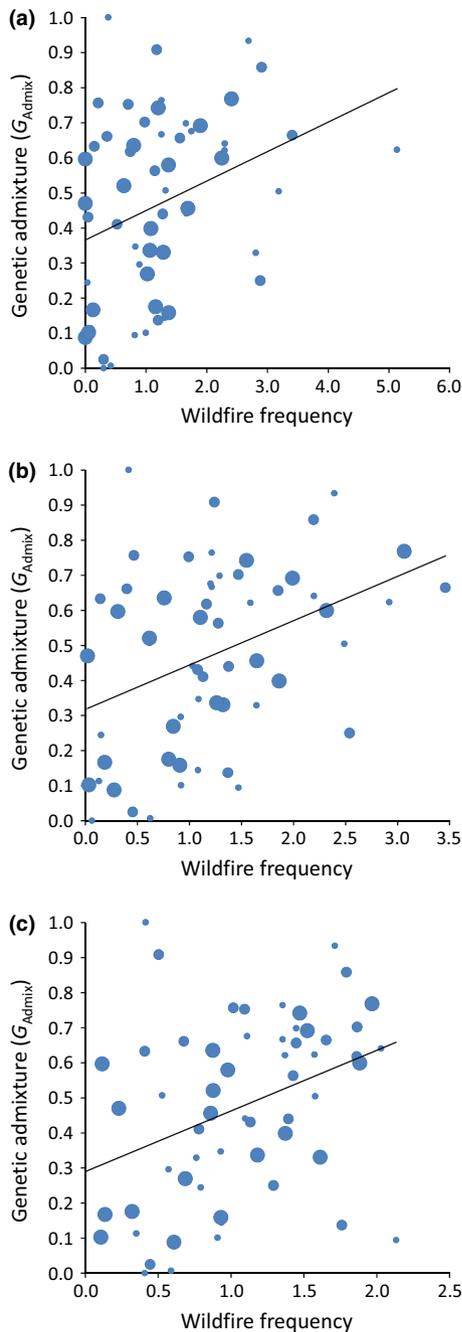
We ran three subsets of models considering the same parameters but with some of them (B, C, E and F) measured at three different spatial scales: (a) 10 km<sup>2</sup>, (b) 100 km<sup>2</sup>, and (c) 1000 km<sup>2</sup> around sampling localities. Only best-ranked equivalent models ( $\Delta AIC_c \leq 2$ ) are shown. K, number of parameters in the model; AIC<sub>c</sub>, corrected Akaike's information criterion (AIC) value; ΔAIC<sub>c</sub>, difference in AIC<sub>c</sub> value from that of the strongest model;  $\omega_i$ , AIC<sub>c</sub> weight.

*et al.*, 2003; Chatzimanolis & Caterino, 2007; Davis *et al.*, 2008). According to STRUCTURE analyses for  $K=3$ , 26% of the analyzed individuals were identified as hybrids according to the  $Q \geq 0.90$  threshold criterion (Vähä & Primmer, 2006; Lepais *et al.*, 2009; Cullingham *et al.*, 2012) (Fig. 2a), and this figure increased to 34% when only populations located in areas of geographic range overlap were considered. NEWHYBRIDS analyses estimated a higher frequency of hybrids than did STRUCTURE for

**Table 3** General linear models (GLMs) for genetic admixture ( $G_{Admix}$ ) of *Quercus berberidifolia* and *Quercus durata*

	Estimate ± USE	$\Sigma \omega_i$	Lower 95% CI	Upper 95% CI
<b>(a) <math>G_{Admix}</math> (10 km<sup>2</sup>) (% of explained deviance: 18.20)</b>				
Intercept	0.331 ± 0.100			
Wildfire frequency	0.087 ± 0.033	0.38	<b>0.02</b>	<b>0.15</b>
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>(b) <math>G_{Admix}</math> (100 km<sup>2</sup>) (% of explained deviance: 19.15)</b>				
Intercept	0.313 ± 0.063			
Wildfire frequency	0.127 ± 0.037	0.30	<b>0.06</b>	<b>0.20</b>
Range overlap	0.069 ± 0.062	0.11	-0.05	0.19
<b>(c) <math>G_{Admix}</math> (1000 km<sup>2</sup>) (% of explained deviance: 28.03)</b>				
Intercept	0.082 ± 0.124			
Wildfire frequency	0.259 ± 0.067	0.44	<b>0.13</b>	<b>0.39</b>
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

Parameter estimates and unconditional SE (USE) were obtained performing model averaging of the best ranked equivalent models ( $\Delta AIC_c \leq 2$ ) (see Table 1). Variables are sorted according to their relative importance based on the sum of Akaike weights ( $\Sigma \omega_i$ ) of those models with  $\Delta AIC_c \leq 2$  in which the variable was present. Bold type indicates significant variables, that is, 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 km<sup>2</sup>, (b) 100 km<sup>2</sup>, and (c) 1000 km<sup>2</sup> around sampling localities.



**Fig. 4** Relationship between genetic admixture of *Quercus berberidifolia* and *Quercus durata* ( $G_{Admix}$ ) and average wildfire frequency estimated in areas of 10 km<sup>2</sup> (a), 100 km<sup>2</sup> (b), and 1000 km<sup>2</sup> (c) around sampling localities. Regression lines are indicated and dot size is proportional to sample size for each studied population.

both the entire study area (42%) and the region of parapatry (54%), which may be explained by the presence of genetic substructure within *Q. berberidifolia* and the fact that NEWHYBRIDS can only accommodate two genetic clusters (Anderson & Thompson, 2002). These hybridization rates are similar to those reported for two other hybridizing Californian oaks (Ortego *et al.*, 2014b) but markedly higher than those found among most interbreeding taxa within the genus (e.g. Craft *et al.*, 2002; Curtu *et al.*, 2007; Cavender-Bares & Pahlich, 2009). The higher

occurrence of hybrids in areas where the geographical ranges of both parental taxa overlap confirms previous studies indicating that spatial proximity is an important factor determining hybridization rates (e.g. Dodd & Afzal-Rafii, 2004; Ortego *et al.*, 2014b). However, we found evidence of introgression of *Q. durata* in allopatric populations of *Q. berberidifolia* located >190 km away from the closest populations of the former (Fig. 2a). This could have resulted from long-distance pollen dispersal (e.g. Dodd & Afzal-Rafii, 2004) or historic hybridization events followed by local extinction of one parental species (Ortego *et al.*, 2014b), this last hypothesis being very unlikely to explain our results as a result of the lack of serpentine soils in southern Californian (Kruckeberg, 1984). The high admixture proportions found in some individuals from populations of *Q. berberidifolia* located far away from the distribution limit of *Q. durata* (e.g. IGN and ELS; Fig. 2a) suggests the presence of first-generation hybrids/backcrosses and points to long-distance pollen dispersal as the most likely proximate mechanism explaining the occurrence of a few introgressed genotypes in allopatric populations (Dodd & Afzal-Rafii, 2004). These results are in agreement with paternity-based studies on oaks showing that although pollen dispersal quickly decays with distance from paternal trees (e.g. Streiff *et al.*, 1999; Pluess *et al.*, 2009), sporadic long-distance pollination events can still have some impact on the genetic structure and diversity of faraway populations (e.g. Buschbom *et al.*, 2011; Hampe *et al.*, 2013).

Our data suggest that both species maintain their genetic and ecological identity in the presence of frequent interspecific gene flow, a typical outcome for highly interbreeding oaks in which the hybrid state often constitutes a transitory phase followed by parental species 'resurrection' in a few generations via recurrent backcrossing and asymmetrical gene flow (Bacilieri *et al.*, 1996; Petit *et al.*, 2004; Lepais & Gerber, 2011). The adaptation of *Q. durata* to serpentine soils in which *Q. berberidifolia* is unable to form stable populations is likely to have resulted in disruptive selection linked to microhabitat specialization (Whittaker, 1954; Brady *et al.*, 2005; 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 contains 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 & Pahlisch, 2009). Accordingly, Forde & Faris (1962) found evidence of differences in flowering time between nearby populations of *Q. durata* and *Q. berberidifolia*, which may be an important pre-mating 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

In accordance with the environmental disturbance hypothesis, we found that genetic admixture increases with wildfire frequency, an effect that was significant across a wide range of spatial scales and also exclusively considering parapatric populations from both species. Different mechanisms could explain the positive effects of wildfire frequency on local patterns of hybridization. Frequent wildfires are expected to reduce intra- and interspecific competition, which may relax selection against hybrids and increase their rates of successful establishment even if they still perform worse than either parental species (Brooks, 1987; Harrison, 1999; Brady *et al.*, 2005). Although scrub oaks are good resprouters (Keeley, 1992), the re-establishment of parental species through vegetative regeneration is expected to be hampered by a high wildfire frequency and intensity (Pausas & Keeley, 2014; Pausas *et al.*, 2016), which may ultimately increase the rates of seedling recruitment and the establishment of hybrids. Wildfires also open habitat patches and reduce local population densities and sizes, which is expected to decrease conspecific pollen availability (Breed *et al.*, 2012; Ortego *et al.*, 2014a), allow immigrant pollen from other species to compete with local conspecific pollen (Lagache *et al.*, 2013) and favor more effective pollen and seed dispersal from neighboring populations across barren landscapes (Bacles *et al.*, 2006; Bacles & Ennos, 2008). Finally, wildfires can potentially open new niches where hybrids may outperform or perform similarly well to parental taxa (Lewontin & Birch, 1966) as a result of biotic and abiotic changes in soil composition (Cerdà & Robichaud, 2009) or altered interactions with species involved in earlier stages of secondary ecological succession (Horn, 1974). Thus, different factors can result in burned habitat patches being occupied more frequently by genetically admixed individuals and their persistence may be linked with the frequency of the disturbance, as has been proposed to explain geographical patterns of genetic admixture and diversity for other plants in relation to the degree of stability of suitable habitats across glacial and interglacial cycles (i.e. higher admixture/hybridization in populations from climatically less stable areas; e.g. Ortego *et al.*, 2015b; Guo, 2014).

Despite support for an important role of wildfire frequency on patterns of hybridization, habitat disturbance by urbanization and land clearing for agriculture was only included in models for two spatial scales with a positive, but nonsignificant, effect on the degree of genetic admixture. Some previous studies on trees have

found that hybrids between native and introduced species are more frequently established in human-altered landscapes such as roadsides and urbanized areas (Thompson *et al.*, 2010; Hoban *et al.*, 2012), whereas others have reported widespread hybridization across a variety of natural and disturbed areas (Zalapa *et al.*, 2009). The fact that these studies are focused on native–exotic species hybridization may be confounding the effects purely as a result of environmental disturbance with those resulting from introduction history, which is expected to be biased towards human-populated areas that, in turn, are subjected to many different forms of anthropogenic alterations (Guo, 2014). On the other hand, large-scale human alterations may have weak effects on hybridization rates as a consequence of complete or nearly complete habitat depletion, that is, environmental disturbance does not result in a new niche that is suitable for either hybrids or any parental species (Anderson, 1948; Anderson & Stebbins, 1954; Arnold, 1997).

### Environmental conditions

The studied environmental factors (elevation, climate and soil type) unrelated to habitat disturbance had no significant effect on estimates of mixed gene pools. This result contrasts with previous studies showing that the occurrence of hybrids is constrained by the presence of patches with particular environments suitable for seedling establishment and survival (e.g. Williams *et al.*, 2001; Dodd & Afzal-Rafii, 2004; Cullingham *et al.*, 2012; Ortego *et al.*, 2014b). In our study system, the random distribution of hybrids with respect to climate and elevation may be explained by the fact that the most important selective force controlling the distribution of the two focal species is probably the presence of serpentine soils (Forde & Faris, 1962). These soils are scattered in the landscape and often vary in terms of occurrence and chemical composition at scales of a few tens of meters, a fine-grain heterogeneity not covered by available soil maps including information on soil taxa order and suborder (Whittaker, 1954; Kruckeberg, 1984; Brady *et al.*, 2005).

### Conclusions and future directions

Overall, our study shows for the first time that wildfires are an important source of environmental disturbance promoting genetic admixture between two ecologically well-differentiated species. Our results suggest that the future predictions for increased wildfire frequency in California and many other parts of the world linked with human-induced climate and land-use changes (Westerling *et al.*, 2006; Westerling & Bryant, 2008) will have an important impact not only on species distribution and demographic dynamics but also on interspecific patterns of gene flow (Barr *et al.*, 2015). Longitudinal studies analyzing temporal changes in the proportion and genetic composition of hybrids after specific disturbances (e.g. Muhlfeld *et al.*, 2014) and experimental/quasi-experimental approaches considering different spatial replicates and intensities of environmental disturbances (e.g. Stewart *et al.*, 2015) will provide valuable information to understand more precisely the impacts of human

activities on processes of interspecific 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 at identifying regions under selection and containing gene variants associated with local adaptation processes can provide important insights into the ecological and evolutionary mechanisms underlying asymmetric patterns of introgression in response to human-induced environmental perturbations and climate change (Gailing *et al.*, 2004; Renaut *et al.*, 2013).

## Acknowledgements

We wish to thank to Isabel Afán (LAST-EBD) for her valuable help with GIS analyses. Four anonymous referees provided valuable comments on an earlier draft of this manuscript. J.O. was supported by 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 J.O., 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 to V.L.S.

## 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, analyzed 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: 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: 2626–2639.
- Anderson E. 1948. Hybridization of the habitat. *Evolution* 2: 1–9.
- Anderson E. 1953. Introgressive hybridization. *Biological Reviews of the Cambridge Philosophical Society* 28: 280–307.
- Anderson E, Stebbins GL. 1954. Hybridization as an evolutionary stimulus. *Evolution* 8: 378–388.
- Anderson E, Thompson EA. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160: 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: 2078–2091.
- Arnold ML. 1997. *Natural hybridization and evolution*. Oxford, UK: Oxford University Press.
- Bacilieri R, Ducouso A, Petit RJ, Kremer A. 1996. Mating system and asymmetric hybridization in a mixed stand of European oaks. *Evolution* 50: 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: 368–380.
- Bacles CFE, Lowe AJ, Ennos RA. 2006. Effective seed dispersal across a fragmented landscape. *Science* 311: 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: 2349–2363.
- Barton NH. 2001. The role of hybridization in evolution. *Molecular Ecology* 10: 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: 223–239.
- Beerli P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22: 341–345.
- Beerli P. 2009. *How to Use Migrate or why are Markov 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, USA* 98: 4563–4568.
- Beerli P, Palczewski M. 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* 185: 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: 301–312.
- Burgarella C, Lorenzo Z, Jabbar-Zahad R, Lumaret R, Guichoux E, Petit RJ, Soto A, Gil L. 2009. Detection of hybrids in nature: application to oaks (*Quercus suber* and *Q. ilex*). *Heredity* 102: 442–452.
- Buschbom J, Yanbaev Y, Degen B. 2011. Efficient long-distance gene flow into an isolated relict oak stand. *Journal of Heredity* 102: 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: 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: 1690–1702.
- Cerdà A, Robichaud P. 2009. *Fire effects on soils and restoration strategies*. Enfield, NH, USA: Science Publishers.
- Chatzimanolis S, Caterino MS. 2007. Toward a better understanding of the “Transverse Range Break”: lineage diversification in southern California. *Evolution* 61: 2127–2141.
- 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: 1792–1798.
- Crispo E, Moore JS, Lee-Yaw JA, Gray SM, Haller BC. 2011. Broken barriers: human-induced changes to gene flow and introgression in animals. *BioEssays* 33: 508–518.
- Cunningham CI, Cooke JEK, Dang S, Davis CS, Cooke BJ, Coltman DW. 2011. Mountain pine beetle host-range expansion threatens the boreal forest. *Molecular Ecology* 20: 2157–2171.
- Cunningham CI, James PMA, Cooke JEK, Coltman DW. 2012. Characterizing the physical and genetic structure of the lodgepole pine jack pine hybrid zone: mosaic structure and differential introgression. *Evolutionary Applications* 5: 879–891.
- Curtu AL, Gailing O, Finkeldey R. 2007. Evidence for hybridization and introgression within a species-rich oak (*Quercus* spp.) community. *BMC Evolutionary Biology* 7: 218.

- Darwin CR. 1859. *On the origin of species by means of natural selection*. London, UK: John Murray.
- Davis EB, Koo MS, Conroy C, Patton JL, Moritz C. 2008. The California Hotspots Project: identifying regions of rapid diversification of mammals. *Molecular Ecology* 17: 120–138.
- Dodd RS, Afzal-Rafii Z. 2004. Selection and dispersal in a multispecies oak hybrid zone. *Evolution* 58: 261–269.
- Dow BD, Ashley MV. 1998. High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. *Journal of Heredity* 89: 62–70.
- Ellstrand NC, Schierenbeck KA. 2000. Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences, USA* 97: 7043–7050.
- Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.
- Field DL, Ayre DJ, Whelan RJ, Young AG. 2011. Patterns of hybridization and asymmetrical gene flow in hybrid zones of the rare *Eucalyptus aggregata* and common *E. rubida*. *Heredity* 106: 841–853.
- eFloras. 2015. *Flora of North America*. St Louis, MO, USA & Cambridge, MA, USA: Missouri Botanic Garden & Harvard University Herbaria. [WWW document] URL <http://www.efloras.org> [accessed 20 December 2015].
- Forde MB, Farris DG. 1962. Effect of introgression on serpentine endemism of *Quercus durata*. *Evolution* 16: 338–347.
- Gailing O, Macnair MR, Bachmann K. 2004. QTL mapping for a trade-off between leaf and bud production in a recombinant inbred population of *Microseris douglasii* and *M. bigelovii* (Asteraceae, Lactuceae): a potential preadaptation for the colonization of serpentine soils. *Plant Biology* 6: 440–446.
- Goudet J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* 86: 485–486.
- Guo QF. 2014. Plant hybridization: the role of human disturbance and biological invasion. *Diversity and Distributions* 20: 1345–1354.
- Haines ML, Melville J, Sumner J, Clemann N, Chapple DG, Stuart-Fox D. 2016. Geographic variation in hybridization and ecological differentiation between three syntopic, morphologically similar species of montane lizards. *Molecular Ecology* 25: 2887–2903.
- Hampe A, Pemonge MH, Petit RJ. 2013. Efficient mitigation of founder effects during the establishment of a leading-edge oak population. *Proceedings of the Royal Society B-Biological Sciences* 280: 20131070.
- Harrison RG. 1986. Pattern and process in a narrow hybrid zone. *Heredity* 56: 337–349.
- Harrison S. 1999. Local and regional diversity in a patchy landscape: native, alien, and endemic herbs on serpentine. *Ecology* 80: 70–80.
- Hasselman DJ, Argo EE, McBride MC, Bentzen P, Schultz TF, Perez-Umphrey AA, Palkovacs EP. 2014. Human disturbance causes the formation of a hybrid swarm between two naturally sympatric fish species. *Molecular Ecology* 23: 1137–1152.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.
- Hoban SM, McCleary TS, Schlarbaum SE, Anagnostakis SL, Romero-Severson J. 2012. Human-impacted landscapes facilitate hybridization between a native and an introduced tree. *Evolutionary Applications* 5: 720–731.
- Horn HS. 1974. The ecology of secondary succession. *Annual Review of Ecology and Systematics* 5: 25–37.
- Hubisz MJ, Falush D, Stephens M, Pritchard JK. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 9: 1322–1332.
- Jombart T. 2008. ADEGENET: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405.
- Keeley JE. 1992. Recruitment of seedlings and vegetative sprouts in unburned chaparral. *Ecology* 73: 1194–1208.
- Keeley JE, Bond WJ, Bradstock RA, Pausas JG, Rundel PW. 2012. *Fire in Mediterranean ecosystems: ecology, evolution and management*. Cambridge, UK: Cambridge University Press.
- King RS, Boyesen JR, Brennehan JM, Cong RM, Hunter TS. 2015. Fire created habitat partitioning and isolation between hybridizing warblers. *Ecosphere* 6: 34.
- Kruckeberg AR. 1984. *California serpentine: flora, vegetation, geology, soils, and management problems*. Los Angeles, CA, USA: University of California Publications in Botany.
- Lagache L, Klein EK, Guichoux E, Petit RJ. 2013. Fine-scale environmental control of hybridization in oaks. *Molecular Ecology* 22: 423–436.
- Lamont BB, He T, Enright NJ, Krauss SL, Miller BP. 2003. Anthropogenic disturbance promotes hybridization between *Banksia* species by altering their biology. *Journal of Evolutionary Biology* 16: 551–557.
- Lancaster LT, Kay KM. 2013. Origin and diversification of the California flora: re-examining classic hypotheses with molecular phylogenies. *Evolution* 67: 1041–1054.
- Langhans TM, Storm C, Schwabe A. 2009. Biological soil crusts and their microenvironment: impact on emergence, survival and establishment of seedlings. *Flora* 204: 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: 653–665.
- Lepais O, Gerber S. 2011. Reproductive patterns shape introgression dynamics and species succession within the European white oak species complex. *Evolution* 65: 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: 2228–2242.
- Lewontin RC, Birch LC. 1966. Hybridization as a source of variation for adaptation to new environments. *Evolution* 20: 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: 2007–2012.
- Linnaeus C. 1735. *Systema Naturae*, 1st edn. Leiden, the Netherlands: Theodorum Haak.
- Moritz MA, Battlori 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: 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: 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: 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: 549–561.
- Muller CH. 1952. Ecological control of hybridization in *Quercus* – a factor in the mechanism of evolution. *Evolution* 6: 147–161.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
- Ortego J, Aguirre MP, Nogueras V, Cordero PJ. 2015a. Consequences of extensive habitat fragmentation in landscape-level patterns of genetic diversity and structure in the Mediterranean esparto grasshopper. *Evolutionary Applications* 8: 621–632.
- Ortego J, Bonal R, Munoz A, Aparicio JM. 2014a. Extensive pollen immigration and no evidence of disrupted mating patterns or reproduction in a highly fragmented holm oak stand. *Journal of Plant Ecology* 7: 384–395.
- Ortego J, Gugger PF, Riordan EC, Sork VL. 2014b. Influence of climatic niche suitability and geographical overlap on hybridization patterns among southern Californian oaks. *Journal of Biogeography* 41: 1895–1908.
- 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: 328–338.
- Ortego J, Nogueras V, Gugger PF, Sork VL. 2015c. Evolutionary and demographic history of the Californian scrub white oak species complex: an integrative approach. *Molecular Ecology* 24: 6188–6208.
- 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: 3210–3223.

- Pausas JG, Keeley JE. 2014. Evolutionary ecology of resprouting and seeding in fire-prone ecosystems. *New Phytologist* 204: 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: 945–954.
- Petit RJ, Bodenes C, Ducouso A, Roussel G, Kremer A. 2004. Hybridization as a mechanism of invasion in oaks. *New Phytologist* 161: 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: 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: 902–912.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- R Core Team. 2015. *R: a language and environment for statistical computing, v.3.0.2*. Vienna, Austria: R Foundation for Statistical Computing. [WWW document] URL <http://www.rproject.org>.
- 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: 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, NY, USA: Springer-Verlag, 501–538.
- Rushton BS. 1993. Natural hybridization within the genus *Quercus* L. *Annales des Sciences Forestières* 50: S73–S90.
- Saarman NP, Pogsoson GH. 2015. Introgression between invasive and native blue mussels (genus *Mytilus*) in the central California hybrid zone. *Molecular Ecology* 24: 4723–4738.
- Schwartz MW, Thorne JH, Viers JH. 2006. Biotic homogenization of the California flora in urban and urbanizing regions. *Biological Conservation* 127: 282–291.
- Silliman FE, Leisner RS. 1958. An analysis of a colony of hybrid oaks. *American Journal of Botany* 45: 730–736.
- Staff SS. 2012. *Soil Survey Geographic (SSURGO) database for California*. [WWW document] URL <http://websoilsurvey.sc.egov.usda.gov> [accessed 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: 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: 491–495.
- Streiff R, Ducouso 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: 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: 1388–1402.
- Thompson SL, Lamothe M, Meirans PG, Perinet P, Isabel N. 2010. Repeated unidirectional introgression towards *Populus balsamifera* in contact zones of exotic and native poplars. *Molecular Ecology* 19: 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: 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: 63–72.
- Vandergast AG, Bohonak AJ, Hathaway SA, Boys J, Fisher RN. 2008. Are hotspots of evolutionary potential adequately protected in southern California? *Biological Conservation* 141: 1648–1664.
- Vandergast AG, Bohonak AJ, Weissman DB, Fisher RN. 2007. Understanding the genetic effects of recent habitat fragmentation in the context of evolutionary history: phylogeography and landscape genetics of a southern California endemic Jerusalem cricket (Orthoptera: Stenopelmatidae: *Stenopelmatus*). *Molecular Ecology* 16: 977–992.
- Westerling AL, Bryant BP. 2008. Climate change and wildfire in California. *Climatic Change* 87: S231–S249.
- Westerling AL, Hidalgo HG, Cayan DR, Swetnam TW. 2006. Warming and earlier spring increase western US forest wildfire activity. *Science* 313: 940–943.
- Whittaker RH. 1954. The ecology of serpentine soils. *Ecology* 35: 258–259.
- Whittemore AT, Schaal BA. 1991. Interspecific gene flow in sympatric oaks. *Proceedings of the National Academy of Sciences, USA* 88: 2540–2544.
- Williams JH, Boecklen WJ, Howard DJ. 2001. Reproductive processes in two oak (*Quercus*) contact zones with different levels of hybridization. *Heredity* 87: 680–690.
- Wright JW. 2007. Local adaptation to serpentine soils in *Pinus ponderosa*. *Plant and Soil* 293: 209–217.
- Yost JM, Barry T, Kay KM, Rajakaruna N. 2012. Edaphic adaptation maintains the coexistence of two cryptic species on serpentine soils. *American Journal of Botany* 99: 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: 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: 4995–5011.

## Supporting Information

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

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

**Fig. S2** Ancestry plots for simulated genotypes of *Quercus berberidifolia*, *Q. durata* and four hybrid classes analyzed with STRUCTURE and NEWHYBRIDS.

**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 ( $G_{\text{Admix}}$ )

**Notes S1** Settings of STRUCTURE and NEWHYBRIDS analyses.

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

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

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.



## About *New Phytologist*

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <28 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)
- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**