ORIGINAL ARTICLE

Influence of climatic niche suitability and geographical overlap on hybridization patterns among southern Californian oaks

Joaquín Ortego[1,2,3,4]*, Paul F. Gugger[1], Erin C. Riordan[1] and Victoria L. Sork[1,5]

[1]*Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA,* [2]*Departamento de Ecología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), Madrid E-28006, Spain,* [3]*Grupo de Investigación de la Biodiversidad Genética y Cultural, Instituto de Investigación en Recursos Cinegéticos – IREC (CSIC, UCLM, JCCM), E-13005 Ciudad Real, Spain,* [4]*Conservation and Evolutionary Genetics Group, Department of Integrative Ecology, Estación Biológica de Doñana (EBD-CSIC), E-41092 Seville, Spain,* [5]*Institute of the Environment and Sustainability, University of California, Los Angeles, CA 90095, USA*

ABSTRACT

Aim Information on the association between geographical patterns of hybridization and the climatic niche requirements and co-occurrence of the species involved can provide important insights that further our understanding of the factors promoting the formation of hybrid zones. Here, we test whether climatic niche suitability explains patterns of hybridization beyond spatial overlap in distributions of the geographically restricted Engelmann oak (*Quercus engelmannii*) and the widespread scrub oak species complex (*Quercus* spp.).

Location Southern California, USA.

Methods We sampled Engelmann and scrub oaks across 31 localities and genotyped 343 individuals at nine microsatellite loci. We used climatic niche modelling to assess the suitability of habitats for each parental species and hybrids and to analyse whether climatic niche suitability in addition to co-occurrence are associated with observed spatial patterns of hybridization.

Results Our data indicated that hybrid zones are located in areas that are geographically constrained by the presence of scattered patches of Engelmann oak within a matrix broadly occupied by scrub oaks. However, the climatic niche of hybrids differed from that occupied by both Engelmann and scrub oaks, suggesting the importance of climate on their establishment and persistence beyond the simple role of geographical overlap of parental species distributions. We also found that even moderate levels of introgression contributed to increased genetic diversity in Engelmann oak but not in scrub oaks.

Main conclusions These results show that hybrid zones follow a mosaic pattern constrained by the presence of scattered patches occupied by Engelmann oak and climatically suitable areas that favour the establishment of hybrids. The low rates of ongoing hybridization and the bimodal distribution of parental genotypes found here suggest that hybridization is not threatening Engelmann oak and this phenomenon could instead be promoting the acquisition of favourable alleles that may increase the adaptive potential of this vulnerable species.

Keywords
Ecological niche modelling, extinction through hybridization, genetic diversity, hybridization, introgression, Maxent, microsatellite markers, *Quercus engelmannii*, peripheral populations, southern California.

Correspondence: Joaquín Ortego, Estación Biológica de Doñana, EBD-CSIC, Avda. Américo Vespucio s/n, E-41092 Seville, Spain. E-mail: joaquin.ortego@csic.es

INTRODUCTION

Hybridization and introgression are frequent phenomena across diverse taxa and are considered important evolutionary processes that contribute to shaping geographical patterns of biological diversity through species formation (Chapman & Burke, 2007) and extinction (Rhymer & Simberloff, 1996). For these reasons, the role of ecological factors in promoting or limiting interspecific gene exchange has long been discussed (e.g. Stebbins *et al.*, 1947; Anderson, 1948) and is

now receiving attention from researchers with new approaches that integrate genetic data and information provided by niche modelling (Chatfield *et al.*, 2010; Carling & Zuckerberg, 2011; Cullingham *et al.*, 2012).

Hybrid zones usually form when species separated by weak reproductive barriers come into geographical contact (Barton & Hewitt, 1985). These zones generally occur in ephemeral habitats or narrow ecotonal regions where two or more diverging lineages meet and interbreed after secondary contact (Barton & Hewitt, 1985; Swenson & Howard, 2005; Cullingham *et al.*, 2012). However, the width of hybrid zones is highly variable and some interbreeding taxa are co-distributed or share a large portion of their respective ranges (Barton & Hewitt, 1985). Geographical heterogeneity in hybridization rates has been often described in these widely co-distributed and interbreeding species, indicating that the location of hybridization hotspots cannot be predicted merely from overlapping distribution ranges and suggesting that factors other than simple geographical contact are probably involved (e.g. Lepais *et al.*, 2009; Zeng *et al.*, 2011). For example, hybridization rates or degree of admixture in hybrid individuals might also be a function of the suitability of the habitat for the hybrids or one or both of the parental species, irrespective of their co-occurrence (Anderson, 1948).

Oaks (*Quercus*, Fagaceae) are a classic example of a taxonomic group that often forms mixed stands of two or more species that maintain their morphological, ecological and genetic identities in the presence of relatively frequent hybridization (e.g. Stebbins *et al.*, 1947; Van Valen, 1976; Howard *et al.*, 1997; González-Rodríguez & Oyama, 2005; Curtu *et al.*, 2007; Cavender-Bares & Pahlich, 2009; Ortego & Bonal, 2010). However, patterns of hybridization among oak species are heterogeneous, probably because of the different impacts of multiple factors, such as relative species abundance (Lepais *et al.*, 2009), interpopulation differences in reproductive barriers (Zeng *et al.*, 2011), and local ecological or environmental conditions determining the production and survival of hybrids (Anderson, 1948; Muller, 1952; Benson *et al.*, 1967; Williams *et al.*, 2001; Dodd & Afzal-Rafii, 2004; Alberto *et al.*, 2010). For these reasons, oaks are an excellent study group with which to analyse the factors associated with hybridization patterns across a wide range of environmental gradients and contrasting patterns of species relative abundance (Dodd & Afzal-Rafii, 2004; Lepais *et al.*, 2009; Zeng *et al.*, 2011). Information on the association between ecological niche requirements and geographical patterns of hybridization could provide further insights into this environmental role. Specifically, it will further our understanding of whether hybridization occurs simply in areas of geographical overlap among species or whether it is further influenced by how environmentally suitable a particular site is for each parental species and the hybrids. Climatic niche models provide an estimate of habitat suitability across the range of a species and sites within regions of species overlap can be compared in terms of habitat suitability for each parental species and for hybrids (Chatfield *et al.*, 2010). The climatic conditions at a particular site will be the same for co-occurring individuals but can vary in terms of suitability. Thus, a climatic niche modelling approach can be particularly advantageous when comparing species with overlapping distributions, offering a potential advantage to traditional approaches using raw climatic or environmental variables (e.g. Chatfield *et al.*, 2010; Swenson, 2006; see also Barton & Hewitt, 1985 and references therein).

In this study, we analyse patterns of hybridization between Engelmann oak (*Quercus engelmannii* Greene) and the scrub oak complex (*Quercus* spp.) (Nixon, 2002; Pearse & Hipp, 2009) (Fig. 1). Engelmann oak is a tree species endemic to the southern Californian Floristic Province whose entire distribution range consists of southern California (USA) and northern Baja California (Mexico) (Scott, 1990; Roberts, 1995). Species within the scrub oak complex, *Q. cornelius-mulleri* Nixon & Steele, *Q. berberidifolia* Liebmann, *Q. durata* Jepson and *Q. dumosa* Nuttall (Roberts, 1995; Nixon, 2002; eFloras, 2011), have never been definitively resolved phylogenetically and their taxonomic status has been debated (Tucker, 1952, 1953; Nixon & Steele, 1981; Nixon & Muller, 1994; eFloras, 2011). Scrub oaks have a much larger distribution range than the Engelmann oak, being present from northern California to Baja California and overlapping with the entire geographical range occupied by Engelmann oak (Roberts, 1995). Engelmann oak is included in the IUCN Red List of Threatened Species due to its small range size and extensive habitat destruction (category Vulnerable A1c; IUCN, 2010). The core of the species' distribution range is located in San Diego County (USA), where the species is generally distributed in scattered patches often consisting of a few individuals (Scott, 1991; Ortego *et al.*, 2012). Hybridization with the widespread scrub oaks has implications for the persistence of the species. Extensive hybridization and pollen swamping could dilute Engelmann oak's genome and further decimate the population (Rhymer & Simberloff, 1996; Allendorf *et al.*, 2001; Lepais *et al.*, 2009) or, alternatively, some hybridization could facilitate the exchange of adaptive alleles across species boundaries and thus bolster the population (Barrett & Schluter, 2008; Becker *et al.*, 2013).

We combine information from nuclear microsatellite markers and climatic niche models (CNM) to study the factors associated with hybridization between Engelmann oak and the scrub oak complex. We first quantify the extent of hybridization across the distribution range of Engelmann oak and use CNMs to create climatic suitability maps for Engelmann oak, scrub oaks and hybrids. We then (1) analyse whether the presence of hybrid individuals in a given locality is explained by the climatic suitability of that site for both parental species or simply by their co-occurrence, and (2) test the null hypothesis that hybrids and either parental species have equivalent climatic niches. If the geographical presence of hybrids is associated with co-occurrence of Engelmann and scrub oaks, but they are located in areas that differ from those occupied by both parental species, our findings would then suggest that
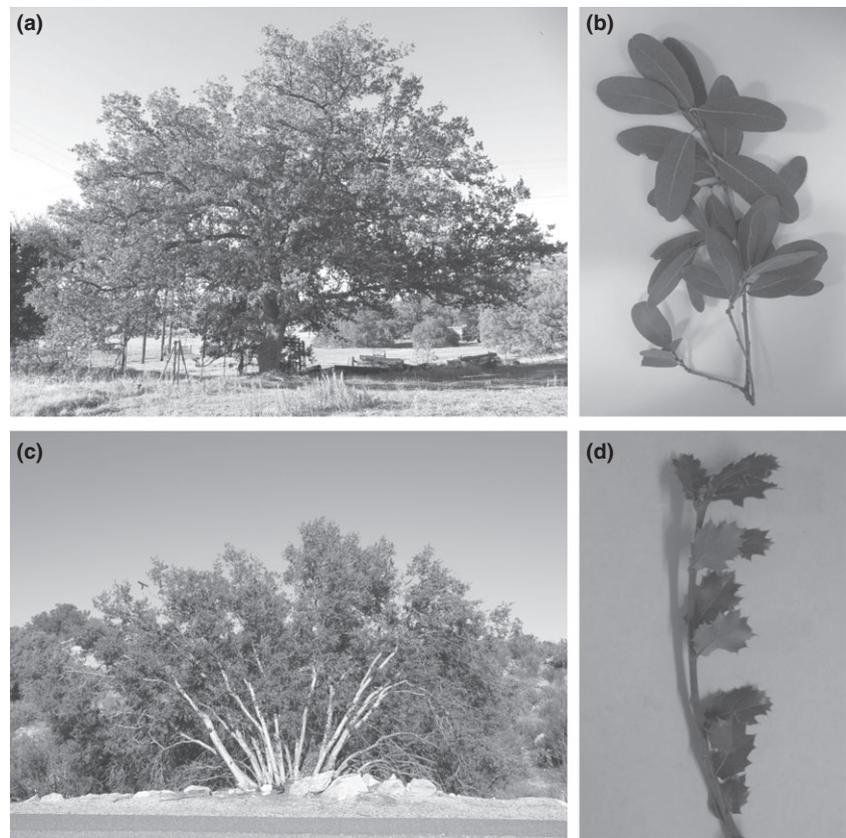
Figure 1 General appearance of the Californian white oak study species, Engelmann oak (*Quercus engelmannii*) and the scrub oak complex (*Quercus* spp.), with a more detailed picture of their respective leaves. (a) Engelmann oak, which is tree-like (up to 10 m), with (b) longer leaves and leaf margins generally entire or undulate. (c) Scrub oaks, which form multi-stemmed short (< 4 m) shrubs, with (d) smaller leaves and leaf margins typically toothed and spinose.

sympatry creates the opportunity for hybridization but climate shapes the extent to which hybrids persist in a given locality.

MATERIALS AND METHODS

Study species and sampling

The distribution range of Engelmann oak overlaps primarily with those of *Q. berberidifolia* and *Q. cornelius-mulleri*, but also to a small extent with that of *Q. durata* var. *gabrielensis* (Roberts, 1995). For the purpose of this study we treat all putative species belonging to the scrub oak complex as a single parental species (see also Introduction). We consider this a reasonable criterion given that STRUCTURE (Pritchard *et al.*, 2000; Falush *et al.*, 2003) analyses described in the results section indicate the presence of only two genetic clusters, one for Engelmann oak and one for scrub oaks. Further, phylogenetic analyses have shown that scrub oaks and Engelmann oak are paraphyletic and placed in very divergent clades within the white oak section (*Quercus*) (Pearse & Hipp, 2009). Phylogenetic analyses indicate that Engelmann oak is more related to the allopatric gray oak (*Q. grisea*), Sonoran blue oak (*Q. oblongifolia*) and Sonoran scrub oak (*Q. turbinella*) found in south-western USA and northern Mexico (Pearse & Hipp, 2009; eFloras, 2011). Scrub oaks are more related to blue oak (*Q. douglasii*) and valley oak (*Q. lobata*), both allopatric to Engelmann oak but co-distributed with scrub oaks in central and northern California (Roberts, 1995; eFloras, 2011).

We sampled Engelmann and scrub oaks across 31 localities covering the majority of the distribution range of Engelmann oak (Ortego *et al.*, 2012) (Fig. 2). During 2008–2011, we collected leaf tissue from 343 individuals from these localities (Table 1). The identification of Engelmann oak and scrub oaks during field sampling was based on their distinctive morphological appearance: Engelmann oak has a tree-like growth form with entire leaves whilst scrub oaks are multi-stemmed shrubs with toothed leaves (Fig. 1). Hybrids show intermediate leaf traits and tend to be more shrubby than tree-like, but can only be tentatively identified in the field on the basis of their morphological appearance (eFloras, 2011; V.L. Sork *et al.*, unpublished data) (Table 1). Spatial coordinates were registered using a Global Positioning System (GPS) and leaf samples were stored frozen (–20 °C) until needed for genetic analyses.

Microsatellite genotyping, admixture analyses and hybrid identification

We used nine polymorphic microsatellite markers previously developed for other *Quercus* species to genotype Engelmann and scrub oaks (see Appendix S1 in Supporting Information). DNA extraction and microsatellite amplification and genotyping were performed as described in Ortego *et al.* (2012).

We identified hybrid individuals using STRUCTURE 2.3.3 (Pritchard *et al.*, 2000; Falush *et al.*, 2003). In STRUCTURE,
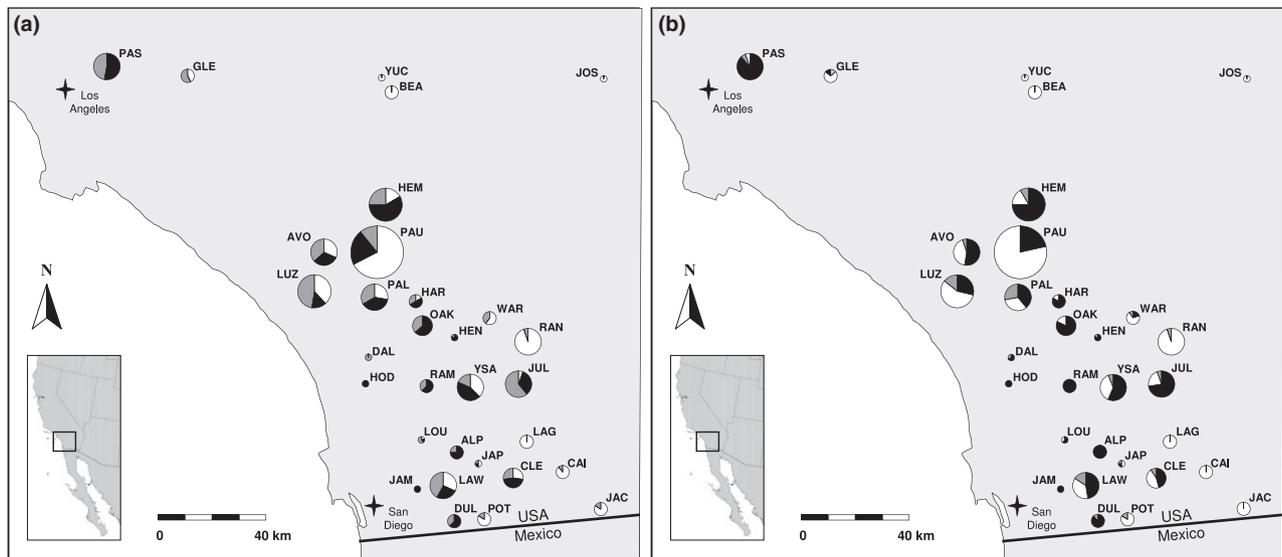
Figure 2 Map representing the location of white oak sampling sites in southern California. The proportion of Engelmann oaks, scrub oaks, and hybrids in each locality are represented using pie charts, with each colour indicating a different genotypic cluster (Engelmann oak: black; scrub oaks: white; hybrids: grey). Individuals were classified into these three groups according to two criteria: a threshold value of posterior probability (q) equal to 0.90 (panel a) and the 95% posterior probability interval around each individual's admixture proportion (panel b) (see text for further details). Circle area is proportional to the number of individuals sampled at each location. Population codes are described in Table 1.

the posterior probability (q) describes the proportion of an individual genotype originating from each cluster. We ran STRUCTURE assuming correlated allele frequencies and admixture and conducted 10 independent runs for each value of $K = 1$–10 to estimate the 'true' number of clusters with 10^6 Markov chain Monte Carlo (MCMC) cycles, following a burn-in period of 100,000 iterations (Pritchard *et al.*, 2000; Falush *et al.*, 2003). We expected an optimal value of $K = 2$, with one genetic cluster corresponding to Engelmann oak and another to scrub oak, but we ran STRUCTURE considering $K = 1$–10 to allow log probabilities [Pr($X|K$)] to stabilize and to make sure that we were accurately inferring the optimal number of genetic clusters (Evanno *et al.*, 2005). The number of populations best fitting the data set was defined both using log probabilities [Pr($X|K$)] (Pritchard *et al.*, 2000) and the ΔK method (Evanno *et al.*, 2005), as implemented in STRUCTURE HARVESTER 0.9.93 (Earl & vonHoldt, 2012).

Microsatellite genotypes were tested for departure from Hardy–Weinberg equilibrium (HWE) within each sampling population/species at each locus using an exact test (Guo & Thompson, 1992) based on 900,000 Markov chain iterations as implemented in the program ARLEQUIN 3.1 (Excoffier *et al.*, 2005). We also used ARLEQUIN 3.1 to test for linkage equilibrium between each pair of loci and population/species using a likelihood-ratio statistic, whose distribution was obtained by a permutation procedure (Excoffier *et al.*, 2005). Sequential Bonferroni corrections were applied to account for multiple comparisons (Rice, 1989). We assessed the diagnostic power of each marker by calculating allele frequency differentials (δ; Shriver *et al.*, 1997) based on individuals

showing pure genotypes according to STRUCTURE analyses. For this specific analysis, we considered as pure individuals those with a low degree of admixed ancestry using a threshold value of q equal to 0.90. Measures of allelic richness (A_R) were standardized for sample size using the program HP-RARE 1.1 (Kalinowski, 2005).

We identified hybrid individuals and assessed support for possible instances of mixed ancestry in two ways. First, we used a threshold value of $q = 0.90$ to classify each individual as pure-bred or hybrid. A value of $q \geq 0.90$ was classified as a pure-bred genotype, and a value of $q < 0.90$ was classified as an introgressed genotype (e.g. Vähä & Primmer, 2006; Lepais *et al.*, 2009; Ortego & Bonal, 2010; Cullingham *et al.*, 2012). Hereafter, we refer to this criterion as the 'q threshold criterion'. Second, we used the ANCESTDIST option in STRUCTURE to compute the 95% posterior probability interval (the Bayesian analogue of a confidence interval) around each individual's admixture proportion (Pritchard *et al.*, 2000; Falush *et al.*, 2003). According to this criterion, individuals for which the probability interval of the genome proportion derived from the cluster of origin did not include q equal to one are likely to have experienced introgression (e.g. Kronforst *et al.*, 2006; Blair & Hufbauer, 2010; Wheeldon *et al.*, 2010). Hereafter, we refer to this criterion as the '95% interval criterion'. It should be noted that when the number of employed markers is relatively low (e.g. the typical panel of 7–10 microsatellite markers typed here and in most previous studies on hybridization; e.g. Lepais *et al.*, 2009; Peñaloza-Ramírez *et al.*, 2010; Cullingham *et al.*, 2012), the 95% interval criterion is likely to be more restrictive than the q threshold criterion to identify hybrid individuals but at

Table 1 Geographical location of white oak sampling sites in southern California. Number of Engelmann oaks (N_{ENG}), scrub oaks (N_{SCR}), and hybrids (N_{HYB}) classified according to STRUCTURE analyses are indicated for each locality. Individuals were classified in these three groups according to two criteria: (A) a threshold value of posterior probability (q) equal to 0.90 and (B) the 95% posterior probability interval around each individual's admixture proportion (see text for further details). We also indicate the *a priori* assignment of individuals during field sampling. These data should be exclusively considered as a reference and interpreted with caution given that hybrid individuals were only tentatively identified during field sampling based on a quick assessment of their morphological appearance.

Locality	Code	Latitude	Longitude	Field classified			A			B		
				N_{ENG}	N_{SCR}	N_{HYB}	N_{ENG}	N_{SCR}	N_{HYB}	N_{ENG}	N_{SCR}	N_{HYB}
Glendora	GLE	34.177483	−118.095033	3	4	0	0	3	4	1	5	1
Pasadena	PAS	34.134079	−118.098914	17	0	0	9	0	8	15	1	1
Yucaipa	YUC	34.038817	−117.021700	0	4	0	0	4	0	0	4	0
Joshua Tree National Park	JOS	34.017380	−116.167374	0	5	0	0	5	0	0	5	0
Beaumont	BEA	33.909783	−116.983217	0	7	0	0	7	0	0	7	0
Hemet	HEM	33.628262	−117.012864	15	5	4	14	4	6	18	4	2
Avocado Mesa	AVO	33.513735	−117.308909	5	5	9	6	6	7	10	8	1
Pauba Ranch	PAU	33.508552	−117.088208	3	11	23	8	25	4	8	29	0
De Luz	LUZ	33.423553	−117.321380	12	6	3	3	8	10	6	12	3
Palomar Mountains	PAL	33.390607	−117.039250	12	3	3	7	5	6	7	6	5
Harolds	HAR	33.302025	−116.892959	3	0	3	3	1	2	5	1	0
Oak Knoll Ranch	OAK	33.298210	−116.922127	11	0	0	7	0	4	9	0	2
Lake Henshaw	HEN	33.276442	−116.855038	5	0	0	4	0	1	4	0	1
Warner Springs	WAR	33.275230	−116.624060	2	6	2	0	6	4	2	7	1
Ranchita	RAN	33.211081	−116.485548	0	18	1	0	18	1	0	18	1
Daley Ranch	DAL	33.165990	−117.047007	3	0	0	0	0	3	2	0	1
Santa Ysabel	YSA	33.102790	−116.669374	6	6	4	7	6	3	9	6	1
Julian	JUL	33.074770	−116.549129	10	6	2	6	1	11	13	4	1
Lake Hodges	HOD	33.074700	−117.118072	3	0	0	3	0	0	3	0	0
Ramona	RAM	33.029917	−116.823050	8	0	0	5	0	3	8	0	0
Louis A. Stelzer County Park	LOU	32.881655	−116.901225	1	3	1	1	1	3	3	2	0
Laguna Mountain	LAG	32.849683	−116.485150	0	6	0	0	6	0	0	6	0
Japatul	JAP	32.823380	−116.627470	1	3	0	1	2	1	1	2	1
Alpine	ALP	32.814090	−116.772395	8	0	0	6	0	2	8	0	0
Cleveland National Forest	CLE	32.776504	−116.494784	5	5	1	3	3	5	5	5	1
Mc Cain Valley Road	CAI	32.770260	−116.258620	0	9	0	0	8	1	0	9	0
Lawson Valley Road	LAW	32.744610	−116.805661	8	6	5	5	6	8	9	7	3
Jamul	JAM	32.730587	−116.875713	3	0	0	3	0	0	3	0	0
Dulzura	DUL	32.631651	−116.761523	8	0	0	5	0	3	7	0	1
Jacumba	JAC	32.622233	−116.218267	0	6	0	0	5	1	0	6	0
Potrero	POT	32.597267	−116.554867	0	5	1	0	5	1	0	5	1

the expense of classifying many individuals with a high degree of admixed ancestry as pure-bred (e.g. Blair & Hufbauer, 2010; Wheeldon *et al.*, 2010). For this reason, we present all subsequent analyses using classifications of individuals based on both criteria.

In subsequent analyses, we used two indices to describe the genetic characteristics of sampled individuals. The first index is the assignment scores (q) provided by STRUCTURE. This index ranges from 0 (a 'pure' Engelmann oak genotype) to 1 (a 'pure' scrub oak genotype) (see Results). Hybrid individuals in which Engelmann and scrub oak genotypes are equally represented would have a score of 0.5 (see Chatfield *et al.*, 2010). Second, we developed a 'hybridity index' (H_{GEN}), calculated using the equation: $H_{GEN} = 2 \times (0.5 - |(q - 0.5)|)$. Values close to 1 indicate hybrid individuals in which Engelmann and scrub oak genotypes are equally represented. Values close to 0 indicate pure-bred individuals of either Engelmann or scrub oak. Thus, the main property of

this index is that it scores pure-bred individuals with the same value ($H_{GEN} = 0$) irrespective of whether they belong to one or the other parental species.

Climate niche modelling

We used niche modelling to predict the geographical distribution of climatically suitable habitat for Engelmann and scrub oaks and analyse whether current climatic conditions and niche overlap are associated with observed patterns of hybridization. We modelled the current climate-based distribution of Engelmann and scrub oaks using a maximum entropy algorithm, MAXENT 3.3.3 (Phillips *et al.*, 2006; Phillips & Dudik, 2008). MAXENT calculates probability distributions based on incomplete information and does not require absence data, making it appropriate for modelling species distributions based on presence-only records (Elith *et al.*, 2006; Phillips *et al.*, 2006). Additionally, MAXENT has high

performance with spatially biased datasets (Loiselle *et al.*, 2008).

Models were trained using species occurrence data obtained from herbarium records and from genetically identified Engelmann and scrub oaks field sampling locations using STRUCTURE software and the *q* threshold criterion, i.e. the most accurate and restrictive criterion to identify genetically pure individuals (see above for details). Herbarium record localities were compiled from specimens collected from 1950 to 2011 using the Consortium of California Herbaria database (http://ucjeps.berkeley.edu/consortium/). For the scrub oak complex, we compiled records from the following three species: *Q. berberidifolia*, *Q. cornelius-mulleri*, *Q. durata* var. *gabrielensis*. Because *Q. berberidifolia* has a broad distribution in California, herbarium records for this species were restricted to latitudes overlapping with *Q. engelmannii* and the other scrub oak species (32.56–34.38° N). All records were carefully checked to exclude species misidentifications, errors in georeferencing and duplicate records. For models, all duplicate locations falling within the same grid cell were also removed, resulting in 175 localities for Engelmann oak and 381 localities for the scrub oak complex. Models trained using only herbarium data (i.e. excluding genetically identified individuals from sampling sites) provided analogous distribution models and results in subsequent analyses (data not shown).

To construct the models, we used 19 climate variables from the WorldClim 1.4 dataset (http://www.worldclim.org/) interpolated to 30-arcsec (*c.* 1-km) resolution (Hijmans *et al.*, 2005). We chose WorldClim because it is the highest resolution dataset available that spans the USA–Mexico political boundary and covers the Mexican portion of the ranges of both Engelmann and scrub oaks. This resolution is also the most detailed available at the spatial scale at which climatically based selection is likely to determine the establishment of oak seedlings (e.g. Borchert *et al.*, 1989). Final model variables were selected to minimize correlations among variables and maximize variable contribution to model predictions. If two variables were highly correlated (Pearson's correlation coefficient, $r > 0.85$), we retained only the variable with the maximum contribution to the model. Variables with a high contribution to model performance, measured by the training gain, and those whose absence led to a strong drop in training gain were retained. Thus, variable importance was determined from the percentage contribution of each variable to the model and the loss of predictive power when each variable was excluded. Model evaluation statistics were produced from 10 cross-validation replicate model runs. Overall model performance was evaluated using the area under the receiver operating characteristic curve (AUC), which ranges from 0.5 (random prediction) to 1 (maximum prediction). Climate suitability maps were created using the logistic output of MAXENT, which that consists of a gridded map with each cell having an index of suitability between 0 and 1. Low values indicate that conditions are unsuitable for the species to occur, whereas high values indicate that conditions are suitable. Model predictions were visualized in ARCMAP 9.3 (ESRI, Redlands, CA, USA). We have only modelled the climatic niche of the studied species, which does not preclude the impact of other abiotic (e.g. exposure, soil type, etc.) or biotic factors (e.g. interspecific competition, parasitism, etc.) on the spatial distribution of the studied species (e.g. Pearson & Dawson, 2003; Kozak *et al.*, 2008; Wiens *et al.*, 2009).

Analyses of hybrid zones

We used an information-theoretic model selection approach to analyse H_{GEN} in relation to presence of Engelmann oak in a given locality (included as fixed factor) and climatic niche suitability indices for Engelmann and scrub oaks (included as covariates). This analysis allowed us to test the role of climatic suitability of each parental species on hybrid occurrence beyond any simple association with species co-occurrence. In particular, if the establishment of hybrids is mostly constrained by climatic suitability for the most restricted Engelmann oak, we would expect that the degree of admixed ancestry of individuals is higher in stands where Engelmann oak is present and increases with climatic suitability for this species. These analyses were performed at the individual level using generalized linear mixed models (GLMM) with a normal error structure and identity link function as implemented in the R 3.0.0 package LME4 (R Core Team, 2012). Locality was included as random effect to control for the expected non-independence of data due to genetic and environmental spatial autocorrelation (Krackow & Tkadlec, 2001). We ranked the resulting models following a model-selection approach on the basis of Akaike's information criterion corrected for small sample size (AIC$_c$; Burnham & Anderson, 1998). AIC$_c$ values for each model were rescaled (ΔAIC$_c$) calculating the difference between the AIC$_c$ value of each model and the minimum AIC$_c$ obtained among all competing models (i.e. the best model has ΔAIC$_c$ = 0). Models with ΔAIC$_c \leq 2$ were considered equivalent (Burnham & Anderson, 1998). Model selection was performed using the R package AICcMODAVG (R Core Team, 2012).

Finally, we tested the null hypothesis that hybrids and parental species are distributed in an identical climatic space. For this purpose we built a climatic niche model for genetically identified hybrids with MAXENT and considering the same bioclimatic layers and parameters used for parental species as described in the previous section. We quantified niche overlap between parental species and hybrids using Schoener's *D* (Schoener, 1968) and Warren's *I* (Warren *et al.*, 2008) metrics. Both measures range from 0 (no niche overlap) to 1 (identical niches). Then, we performed niche-identity analyses with 200 pseudo-replicates to test whether niche models generated for both parental species and hybrids are more different than expected if they were drawn from the same underlying distribution. These analyses were performed using program ENMTOOLS 1.4.3 (Warren *et al.*, 2010). We built niche models and conducted identity tests
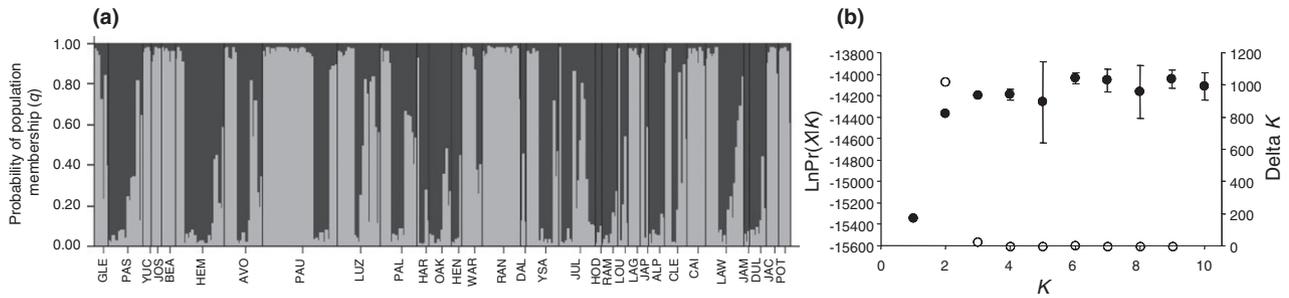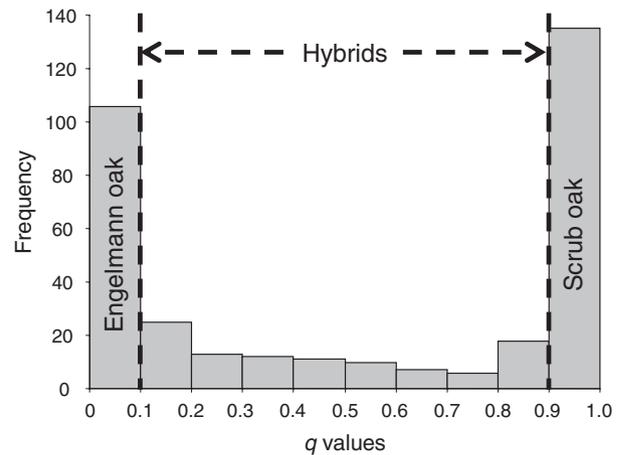
Figure 3 (a) Results of genetic assignments of Engelmann oaks, scrub oaks and hybrids from southern California ($n = 343$ individuals in total) based on the Bayesian method implemented in the program STRUCTURE considering two genetic clusters. Each individual is represented by a thin vertical line, which is partitioned into coloured segments that represent the individual's probability of belonging to the cluster with that colour (Engelmann oak: dark grey; scrub oaks: light grey). Individuals are grouped according to sampling localities (codes are described in Table 1). (b) Results of Bayesian clustering analyses in STRUCTURE. Plots show the mean ($\pm$ SD) log probability of the data (ln Pr($X|K$)) over 10 runs (left axis, black dots and error bars) for each value of K. The magnitude of ΔK as a function of K indicates the most likely number of genetic clusters ($K = 2$) in the STRUCTURE analyses (right axis, open dots).

considering hybrids classified according to the q threshold criterion and the 95% interval criterion.

RESULTS

Admixture analyses and hybrid identification

Log probabilities [Pr($X|K$)] of STRUCTURE analyses sharply increased from $K = 1$ to $K = 2$, but only slightly from $K = 2$ until reaching a plateau at $K = 6$ (Fig. 3b). Accordingly, the ΔK method (Evanno et al., 2005) indicated that the best-supported number of clusters was $K = 2$ (Fig. 3b). The two groups correspond to Engelmann oak and the scrub oak species complex (Figs 2 & 3; Table 1). Considering a threshold value of $q = 0.90$ to classify each individual as pure or hybrid, we identified 106 Engelmann oaks, 135 scrub oaks, and 102 individuals with some degree of genetic admixture of which only 21 showed approximately 50% of admixture ($0.4 < q < 0.6$) for each of the two parental species (Table 1, Figs 3 & 4). Twenty-eight out of these 102 individuals with some degree of admixed ancestry had 95% posterior probability intervals around the individual's admixture proportions that did not include q equal to one (Table 1) and 19 (68%) of them showed approximately 50% of admixture ($0.4 < q < 0.6$) for each of the two parental species. We found that some specimens tentatively identified in the field as either Engelmann or scrub oaks had a high degree of admixed ancestry and q scores closer to the other parental species. This resulted in the assignment of some individuals, morphologically identified as Engelmann oaks (two individuals from Glendora, one individual from Pasadena and six individuals from De Luz) to pure scrub oaks, and the assignment of some individuals, morphologically identified as scrub oaks (two individuals from Julian), to pure Engelmann oaks, using the 95% interval criterion (Table 1) (see also Cavender-Bares & Pahlich, 2009 for a similar genetic and morphological discordance in two hybridizing oaks).

Figure 4 Frequency distribution of q values according to genetic assignments by STRUCTURE analyses of Engelmann oaks, scrub oaks and hybrids from southern California ($n = 343$ individuals in total). Dashed vertical lines indicate threshold values of posterior probabilities (q) used to classify individuals as pure or hybrids.

Polymorphism characteristics of microsatellite loci used to type Engelmann and scrub oaks based on genetically identified individuals using assignment scores from STRUCTURE ($q < 0.1$ for Engelmann oak and $q > 0.90$ for scrub oaks) are indicated in Appendix S1. They contained 7–31 alleles across the nine loci and provided a high to moderate discriminatory power, with allele frequency differentials ranging from $\delta = 0.44$ to $\delta = 0.89$ (Appendix S1). After applying sequential Bonferroni corrections to compensate for multiple statistical tests, no locus deviated from HWE in any of the sampled populations/ species. Furthermore, we did not find any evidence of linkage disequilibrium at any pair of loci in any population/species.

Climate niche modelling

The climate variables included in the final habitat suitability models for both Engelmann and scrub oaks were temperature

seasonality (BIO4), maximum temperature of the warmest month (BIO5), minimum temperature of the coldest month (BIO6), mean annual precipitation (BIO12), precipitation seasonality (BIO15), and precipitation of the driest quarter (BIO17). The AUC for the test data was above 0.9 for both Engelmann and the scrub oak complex (mean $\pm$ SD, Engelmann oak: 0.959 ± 0.009; scrub oak: 0.914 ± 0.012; $n = 10$ replicate model runs), indicating a high correspondence of the modelled distribution to the observed current distribution (Fielding & Bell, 1997; Phillips *et al.*, 2006). The predicted current distributions for Engelmann and scrub oaks (Fig. 5) are consistent with the observed distribution of both species within the distribution range of Engelmann oak (Scott, 1991; Roberts, 1995). Niche modelling analyses have shown that the climatic niche envelope of Engelmann oak is geographically embedded within the climatic niche envelope of scrub oaks, which is expected based on the broader observed scrub oak distribution range (Fig. 5) (Roberts, 1995; eFloras, 2011).

Analyses of hybrid zones

In testing the co-occurrence hypothesis, our most parsimonious model revealed that H_{GEN} was significantly higher in localities where Engelmann oak is present ($t = -3.70$, $P < 0.001$; estimate $\pm$ SE $= -0.146 \pm 0.039$) (Fig. 6). All other tested models had a ΔAIC_c value > 4 (Table 2). It should be noted that all the analysed independent variables were highly intercorrelated and after the exclusion from the model of the variable 'presence of Engelmann oak', H_{GEN} was positively associated with niche suitability for Engelmann oak ($t = 2.62$, $P = 0.009$; estimate $\pm$ SE $= 0.183 \pm 0.070$). However, H_{GEN} was not associated with niche suitability for scrub oak even when this variable was included alone into the model ($t = 0.43$, $P = 0.667$), indicating that the lack of correlation was not due to interactions among independent variables.

Analyses of niche identity showed that observed niche overlap estimates between Engelmann and scrub oaks were significantly lower than those obtained from the randomized distribution of 200 pseudoreplicates (I: $P < 0.01$; D: $P < 0.01$; Fig. 7a), rejecting the null hypothesis of niche identity between both parental species. The climatic niche of hybrids identified using the q threshold criterion (AUC, mean $\pm$ SD: 0.903 ± 0.065; $n = 10$ replicate model runs) was significantly different from Engelmann (I: $P = 0.07$; D: $P = 0.04$; Fig. 7b) and scrub oaks (I: $P < 0.01$; $D = P < 0.01$; Fig. 7d). The climatic niche of hybrids identified using the 95% interval criterion (AUC, mean $\pm$ SD: 0.891 ± 0.091; $n = 10$ replicate model runs) differed from Engelmann oak (I: $P = 0.03$; D: $P < 0.01$; Fig. 7c) but only marginally from scrub oaks (I: $P = 0.06$; D: $P = 0.1$; Fig. 7e). Assignment scores (q) for hybrid individuals tended to be lower than 0.5 (mean $\pm$
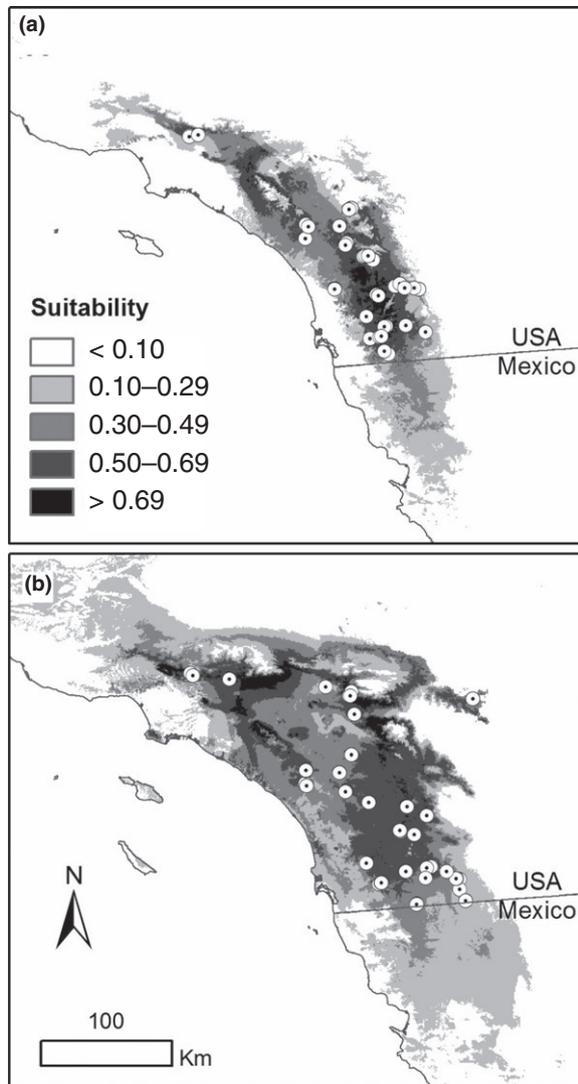

Figure 5 MAXENT predictions of the present distribution for (a) Engelmann oak and (b) scrub oaks. Sampling sites for each species are indicated by white 'bullseyes'. Suitability values indicate logistic probability of presence and range from 0 to 1, with increasingly darker shades of grey with increasing habitat suitability.

Table 2 Model selection to assess the association between the 'hybridity index' (H_{GEN}) and presence of Engelmann oak (A), niche suitability for Engelmann oak (B), and niche suitability for scrub oak (C) in southern California.

Model no.	Model	K	AIC_c	ΔAIC_c	ωi
1	A	4	−456.00	0.00	0.82
2	B	4	−451.38	4.62	0.08
3	A + C	5	−450.41	5.59	0.05
4	A + B	5	−449.70	6.30	0.04
5	C	4	−446.17	9.83	0.01
6	B + C	5	−445.80	10.20	0.01
7	A + B + C	6	−444.15	11.85	0.00

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; ωi, AIC_c weight. Locality was included as random effect in all the generalized linear mixed models (GLMM).
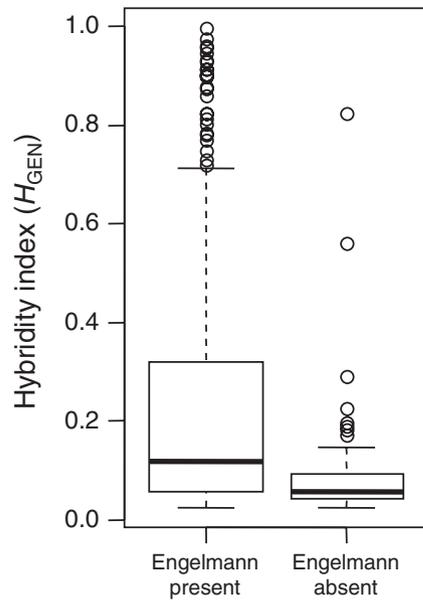
Figure 6 Boxplots of 'hybridity index' (H_{GEN}) for localities where Engelmann oak is present versus absent in southern California. Horizontal black lines show the median and the bottom and top of the boxes show the 25 and 75 percentiles, respectively. The whiskers extend to the highest and lowest values, excluding outliers (open circles).

SE = 0.449 ± 0.026; t-test: $t = 1.97$, $P = 0.051$), indicating that hybrid individuals carry a higher proportion of Engelmann oak than scrub oak alleles. We performed additional analyses only considering hybrid individuals with $0.6 > q > 0.4$ ($n = 21$), i.e. individuals with a high admixed ancestry probably belonging to F_1 hybrids or crosses between them. Genetic assignment scores (q) of this subset of hybrid individuals with a high admixed ancestry did not significantly differ from 0.5 (mean ± SE = 0.503 ± 0.060; t-test: $t = 0.25$, $P = 0.805$). With this reduced dataset we also found that the climatic niche of hybrids (AUC, mean ± SD: 0.922 ± 0.022; $n = 10$ replicate model runs) differed from Engelmann (I: $P < 0.01$; D: $P < 0.01$; Fig. 7f) and scrub oaks (I: $P < 0.01$; D: $P < 0.01$; Fig. 7g).

DISCUSSION

Our data indicate that hybridization between Engelmann and scrub oaks is widespread across all the distribution range of Engelmann oak (Figs 2 & 3). Contrary to other studies analysing two or more species that occupy contrasting climatic niches and hybridize in discrete suture zones of secondary contact (e.g. Swenson, 2006; Chatfield et al., 2010), niche modelling applied to our study system has shown that the climate niche envelope of Engelmann oak is geographically embedded within the climate niche envelope of scrub oaks. Despite the extraordinarily high dispersal potential of pollen in oaks (e.g. Dow & Ashley, 1998; Buschbom et al., 2011), the degree of individual genetic admixture is much higher in areas where Engelmann oak is present and our data indicate

that hybrids occur in environments with climates that are not identical to those occupied by either parental species (Fig. 7). This suggests that sympatry creates the opportunity for hybridization but climate-based selection determines the extent to which hybrids persist in a given locality (Swenson et al., 2008). Overall, our results indicate that the location of hybrid zones is characterized by the presence of particular climatic environments (in terms of both temperature and rainfall) and scattered Engelmann oak patches within a matrix broadly occupied by scrub oaks, a pattern that resembles a mosaic hybrid structure (e.g. Harrison, 1986; Howard et al., 1997; Dodd & Afzal-Rafii, 2004; Cullingham et al., 2012) rather than a simple cline (e.g. Chatfield et al., 2010; Peñaloza-Ramírez et al., 2010; Carling & Zuckerberg, 2011). This mosaic pattern is similar to that previously reported for a contact zone between Q. grisea and Q. gambelii, suggesting that both geographical overlap and environmental heterogeneity are key factors determining the location of hybrid zones in oaks (Howard et al., 1997; Swenson et al., 2008).

Some of our findings indicate that hybridization could be contributing to the progressive extinction of some peripheral and decimated populations of Engelmann oak where the scrub oak is widespread. In these localities, we found few tree-like individuals with the morphological appearance of Engelmann oak (e.g. RAN, POT; Fig. 2) and genetic analyses have revealed that all these individuals are hybrids. The proportion of scrub oak genome in most of these hybrid individuals was very high (>35%), suggesting that extensive pollen swamping and dilution of the genome of the rare species could have resulted in local extinction of Engelmann oak in these localities (Lepais et al., 2009). Thus, one possibility is that introgressed individuals present in these peripheral populations are the result of historical hybridization events followed by a progressive extinction of the rare Engelmann oak whose alleles only persist in hybrid individuals (Dodd & Afzal-Rafii, 2004; Lepais et al., 2009). Scrub oaks probably sustain much higher local effective population sizes (N_e) than Engelmann oak due to their higher population densities, a fact that is also supported by the higher levels of genetic diversity observed in scrub oaks in comparison with Engelmann oak across the nine microsatellite makers analysed (paired t-test, A_R: $t = 2.69$, $P = 0.028$; Appendix S1) (see also Ortego et al., 2010). An alternative to the 'extinction through hybridization' hypothesis to explain the presence of hybrid individuals in these peripheral localities is that pollen dispersal from adjacent Engelmann oak populations has resulted in the establishment of some hybrids in marginal areas with low suitability for Engelmann oak or where this species went extinct due to other factors (Dodd & Afzal-Rafii, 2004; Petit et al., 2004).

A limitation of most microsatellite-based studies is that individuals can be only broadly assigned to pure-bred or hybrids due to the inherent difficulties to differentiate first/second generation hybrids and backcrosses with the typical panel of 7–10 microsatellite markers generally employed in most studies (e.g. Lepais et al., 2009; Peñaloza-Ramírez et al.,
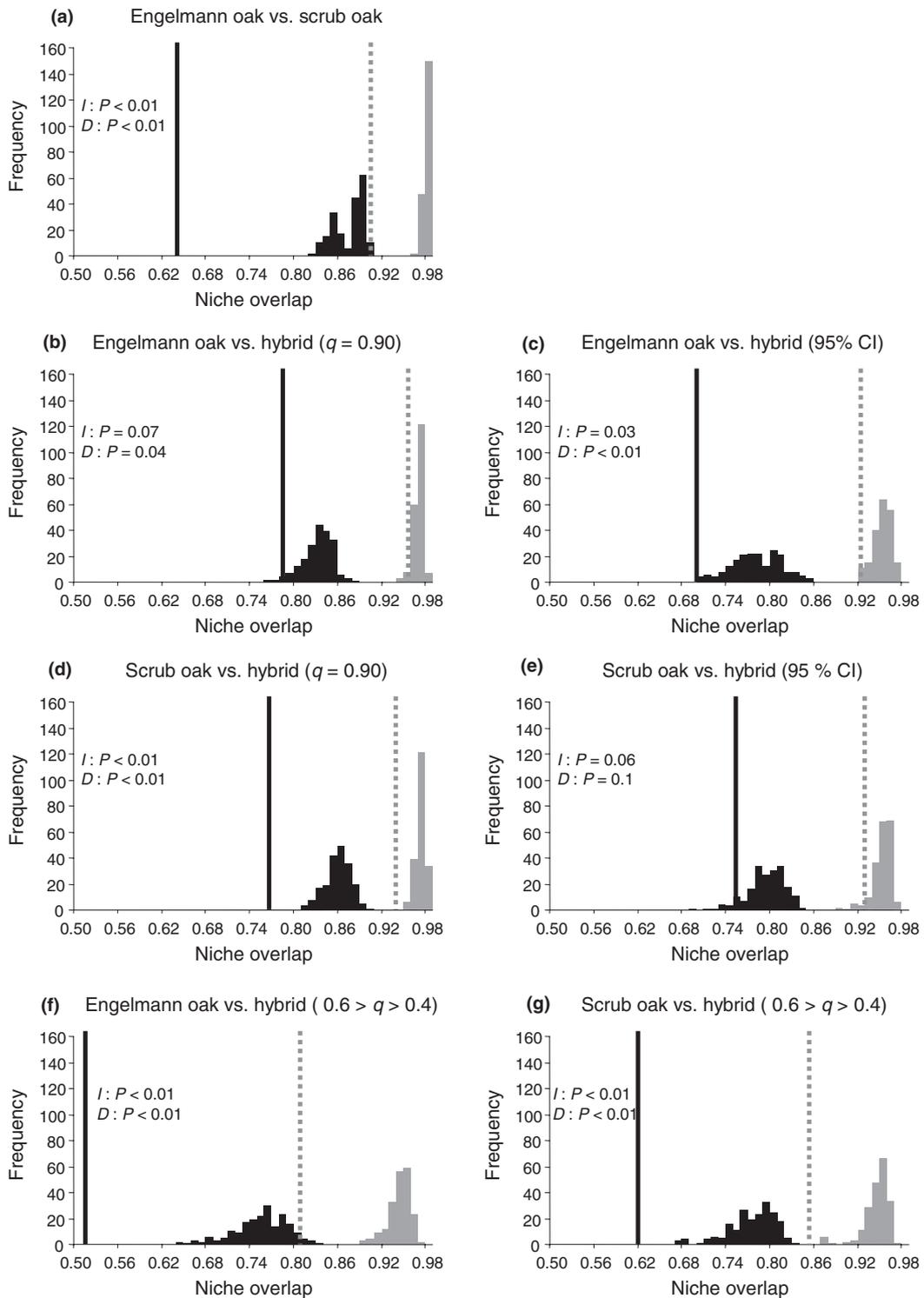
Figure 7 Results of niche identity tests for pairwise comparisons involving Engelmann oak, scrub oaks and hybrids from southern California defined according to different criteria. Vertical lines represent the estimated value of niche overlap calculated from empirical occurrence data using *I* (grey dotted line) and *D* (black solid line) niche similarity metrics. Histograms represent the distribution of niche overlap obtained by random re-sampling of occurrence points (*I*: grey bars; *D*: black bars). The null hypothesis of no significant niche differentiation is rejected if the estimated value falls outside the 95% probability threshold of the permuted distributions (*P*-values are indicated).

2010; Cullingham *et al.*, 2012). Despite this limitation, our data suggest that only a small portion of the analysed individuals have a high degree of admixed ancestry (i.e. with genotypes of both species similarly represented; $0.6 > q > 0.4$) typical of F_1 hybrids or crosses between them (see also Lepais *et al.*, 2009). Both the low frequency of F_1

progeny and the sharp bimodal distribution of parental genotypes (Fig. 4) indicate a low rate of ongoing hybridization, which suggests selection against hybrids and/or strong pre-mating or post-mating pre-zygotic barriers to hybridization (Zeng *et al.*, 2011). Many hybrid individuals (mostly identified using the q threshold criterion) may be backcrosses with a low degree of admixed ancestry, suggesting that F_1 hybrids are preferentially pollinated by one of the parental species and/or increased viability of backcrosses in comparison with first- and second-generation hybrids (Lepais & Gerber, 2011). This could be contributing to avoid total collapse into a hybrid swarm in sympatric areas and favour the recovery of pure-bred genotypes within a few generations and the maintenance of the species genetic identity despite recurrent gene flow between them (Lepais & Gerber, 2011). Thus, the hybrid state could be a transitory phase that may allow the exchange of some alleles with potential evolutionary benefits (Morjan & Rieseberg, 2004; Fitzpatrick *et al.*, 2010; Lepais & Gerber, 2011).

Despite the fact that hybridization can compromise the long-term persistence of rare species or have negative fitness consequences due to the disruption of co-adapted gene complexes (e.g. Rhymer & Simberloff, 1996; Johansen-Morris & Latta, 2008), a moderate degree of allele exchange can also potentially have both additive and non-additive genetic benefits (Edmands, 1999; Barrett & Schluter, 2008). Collective evolution at major loci via trans-specific spread of favourable alleles can enhance within species adaptive variation and increase the chance of long-term population persistence under changing environmental conditions (Becker *et al.*, 2013). On the other hand, some degree of inter-specific allele exchange can also increase fitness via heterosis in certain genes with large effects, which could be particularly advantageous for genetically depauperate species (Edmands, 1999; Kirk *et al.*, 2005). We have found that individuals with moderate levels of admixed ancestry $(0.4 > q > 0.1)$ that were closer to Engelmann oak showed higher levels of allelic richness in comparison with pure-bred Engelmann oaks $(q < 0.1)$ across the nine microsatellite loci analysed (paired t-test; $P < 0.001$) but a comparable analysis was not significant for scrub oaks $(P = 0.243)$. These findings suggest that moderate introgression may increase genetic variability in the genetically less diverse Engelmann oak (see also Nettel *et al.*, 2008). Thus, heterozygote advantage and/or the acquisition of novel favourable alleles from moderate levels of interspecific gene flow could benefit Engelmann oak (Edmands, 1999; Kirk *et al.*, 2005).

Overall, our data showed that hybrid zones follow a mosaic pattern mostly constrained by the presence of scattered patches occupied by Engelmann oak but characterized by particular climatic environments that differ from those occupied by either parental species. Our study adds empirical evidence for the role of environment in promoting hybridization and supports previous research indicating that introgression could contribute to widen parental species ecological niches and occupy new or marginal habitats (Choler *et al.*,

2004; Kunte *et al.*, 2011). The risk of extinction through hybridization in some decimated Engelmann oak populations should be considered in future management programmes and restoration projects. However, the low rates of ongoing hybridization and the bimodal distribution of parental genotypes found here suggest strong selection against hybrids and/or barriers to hybridization, indicating that the studied populations do not represent a hybrid swarm and that hybridization is not threatening Engelmann oak. Instead, it is possible that introgression can be a means of introducing novel alleles that may increase the adaptive potential of Engelmann oak.

REFERENCES

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.

Allendorf, F.W., Leary, R.F., Spruell, P. & Wenburg, J.K. (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology and Evolution*, **16**, 613–622.

Anderson, E. (1948) Hybridization of the habitat. *Evolution*, **2**, 1–9.

Barrett, R.D.H. & Schluter, D. (2008) Adaptation from standing genetic variation. *Trends in Ecology and Evolution*, **23**, 38–44.

Barton, N.H. & Hewitt, G.M. (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Becker, M., Gruenheit, N., Steel, M., Voelckel, C., Deusch, O., Heenan, P.B., McLenachan, P.A., Kardailsky, O., Leigh, J.W. & Lockhart, P.J. (2013) Hybridization may facilitate *in situ* survival of endemic species through periods of climate change. *Nature Climate Change*, **3**, 1039–1043.

Benson, L., Phillips, E.A. & Wilder, P.A. (1967) Evolutionary sorting of characters in a hybrid swarm. I. Direction of slope. *American Journal of Botany*, **54**, 1017–1026.

Blair, A.C. & Hufbauer, R.A. (2010) Hybridization and invasion: one of North America's most devastating invasive plants shows evidence for a history of interspecific hybridization. *Evolutionary Applications*, **3**, 40–51.

Borchert, M.I., Davis, F.W., Michaelsen, J. & Oyler, L.D. (1989) Interactions of factors affecting seedling recruitment of blue oak (*Quercus douglasii*) in California. *Ecology*, **70**, 389–404.

Burnham, K.P. & Anderson, D.R. (1998) *Model selection and inference: a practical information-theoretic approach.* Springer-Verlag, New York.

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

Carling, M.D. & Zuckerberg, B. (2011) Spatio-temporal changes in the genetic structure of the *Passerina* bunting hybrid zone. *Molecular Ecology*, **20**, 1166–1175.

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.

Chapman, M.A. & Burke, J.M. (2007) Genetic divergence and hybrid speciation. *Evolution*, **61**, 1773–1780.

Chatfield, M.W.H., Kozak, K.H., Fitzpatrick, B.M. & Tucker, P.K. (2010) Patterns of differential introgression in a salamander hybrid zone: inferences from genetic data and ecological niche modelling. *Molecular Ecology*, **19**, 4265–4282.

Choler, P., Erschbamer, B., Tribsch, A., Gielly, L. & Taberlet, P. (2004) Genetic introgression as a potential to widen a species' niche: insights from alpine *Carex curvula*. *Proceedings of the National Academy of Sciences USA*, **101**, 171–176.

Cullingham, C.I., James, P.M.A., Cooke, J.E.K. & Coltman, D.W. (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, A.L., Gailing, O. & Finkeldey, R. (2007) Evidence for hybridization and introgression within a species-rich oak (*Quercus* spp.) community. *BMC Evolutionary Biology*, **7**, 218.

Dodd, R.S. & Afzal-Rafii, Z. (2004) Selection and dispersal in a multispecies oak hybrid zone. *Evolution*, **58**, 261–269.

Dow, B.D. & Ashley, M.V. (1998) High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. *Journal of Heredity*, **89**, 62–70.

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

Edmands, S. (1999) Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. *Evolution*, **53**, 1757–1768.

Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Falush, D., Stephens, M. & Pritchard, J.K. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.

Fitzpatrick, B.M., Johnson, J.R., Kump, D.K., Smith, J.J., Voss, S.R. & Shaffer, H.B. (2010) Rapid spread of invasive genes into a threatened native species. *Proceedings of the National Academy of Sciences USA*, **107**, 3606–3610.

eFloras (2011) *Flora of North America.* Missouri Botanical Garden/Harvard University Herbaria, St. Louis, MO/Cambridge, MA. Available at: http://www.efloras.org

González-Rodríguez, A. & Oyama, K. (2005) Leaf morphometric variation in *Quercus affinis* and *Q. laurina* (Fagaceae), two hybridizing Mexican red oaks. *Botanical Journal of the Linnean Society*, **147**, 427–435.

Guo, S.W. & Thompson, E.A. (1992) A Monte Carlo method for combined segregation and linkage analysis. *American Journal of Human Genetics*, **51**, 1111–1126.

Harrison, R.G. (1986) Pattern and process in a narrow hybrid zone. *Heredity*, **56**, 337–349.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Howard, D.J., Preszler, R.W., Williams, J., Fenchel, S. & Boecklen, W.J. (1997) How discrete are oak species? Insights from a hybrid zone between *Quercus grisea* and *Quercus gambelii*. *Evolution*, **51**, 747–755.

IUCN (2010) *IUCN Red List of Threatened Species.* Version 2010.4. Available at: http://www.iucnredlist.org (accessed 17 April 2011).

Johansen-Morris, A.D. & Latta, R.G. (2008) Genotype by environment interactions for fitness in hybrid genotypes of *Avena barbata*. *Evolution*, **62**, 573–585.

Kalinowski, S.T. (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187–189.

Kirk, H., Vrieling, K. & Klinkhamer, P.G.L. (2005) Maternal effects and heterosis influence the fitness of plant hybrids. *New Phytologist*, **166**, 685–694.

Kozak, K.H., Graham, C.H. & Wiens, J.J. (2008) Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology and Evolution*, **23**, 141–148.

Krackow, S. & Tkadlec, E. (2001) Analysis of brood sex ratios: implications of offspring clustering. *Behavioral Ecology and Sociobiology*, **50**, 293–301.

Kronforst, M.R., Young, L.G., Blume, L.M. & Gilbert, L.E. (2006) Multilocus analyses of admixture and introgression

among hybridizing *Heliconius* butterflies. *Evolution*, **60**, 1254–1268.

Kunte, K., Shea, C., Aardema, M.L., Scriber, J.M., Juenger, T.E., Gilbert, L.E. & Kronforst, M.R. (2011) Sex chromosome mosaicism and hybrid speciation among tiger swallowtail butterflies. *PLOS Genetics*, **7**, e1002274.

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, R.J., Guichoux, E., Lavabre, J.E., Alberto, F., Kremer, A. & Gerber, S. (2009) Species relative abundance and direction of introgression in oaks. *Molecular Ecology*, **18**, 2228–2242.

Loiselle, B.A., Jørgensen, P.M., Consiglio, T., Jiménez, I., Blake, J.G., Lohmann, L.G. & Montiel, O.M. (2008) Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? *Journal of Biogeography*, **35**, 105–116.

Morjan, C.L. & Rieseberg, L.H. (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology*, **13**, 1341–1356.

Muller, C.H. (1952) Ecological control of hybridization in *Quercus*: a factor in the mechanism of evolution. *Evolution*, **6**, 147–161.

Nettel, A., Dood, R.S., Afzal-Rafii, Z. & Tovilla-Hernandez, C. (2008) Genetic diversity enhanced by ancient introgression and secondary contact in East Pacific black mangroves. *Molecular Ecology*, **17**, 2680–2690.

Nixon, K. (2002) The oak (*Quercus*) biodiversity of California and adjacent regions. *Proceedings of the Fifth Symposium on Oak Woodlands: Oaks in California's Changing Landscape, October 22–25, 2001, San Diego, California* (ed. by R.B. Standiford, D. McCreary and K.L. Purcell), pp. 3–20. USDA Forest Service General Technical Report PSW-GTR-184, San Diego, CA.

Nixon, K.C. & Muller, C.H. (1994) *New names in California oaks. Novon*, **4**, 391–393.

Nixon, K.C. & Steele, K.P. (1981) A new species of *Quercus* (Fagaceae) from southern California. *Madroño*, **28**, 210–219.

Ortego, J. & Bonal, R. (2010) Natural hybridisation between kermes (*Quercus coccifera* L.) and holm oaks (*Q. ilex* L.) revealed by microsatellite markers. *Plant Biology*, **12**, 234–238.

Ortego, J., Bonal, R. & Muñoz, A. (2010) Genetic consequences of habitat fragmentation in long-lived tree species: the case of the Mediterranean holm oak (*Quercus ilex*, L.). *Journal of Heredity*, **101**, 717–726.

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

Pearse, I.S. & Hipp, A.L. (2009) Phylogenetic and trait similarity to a native species predict herbivory on non-native oaks. *Proceedings of the National Academy of Sciences USA*, **106**, 18097–18102.

Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.

Peñaloza-Ramírez, J.M., González-Rodríguez, A., Mendoza-Cuenca, L., Caron, H., Kremer, A. & Oyama, K. (2010) Interspecific gene flow in a multispecies oak hybrid zone in the Sierra Tarahumara of Mexico. *Annals of Botany*, **105**, 389–399.

Petit, R.J., Bodenes, C., Ducousso, A., Roussel, G. & Kremer, A. (2004) Hybridization as a mechanism of invasion in oaks. *New Phytologist*, **161**, 151–164.

Phillips, S.J. & Dudik, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

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

R Core Team (2012) *R: a language and environment for statistical computing*. Version 3.0.2. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.r-project.org.

Rhymer, J.M. & Simberloff, D. (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83–109.

Rice, W.R. (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

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

Schoener, T.W. (1968) *Anolis* lizards of Bimini: resource partitioning in a complex fauna. *Ecology*, **49**, 704–726.

Scott, T.A. (1990) Conserving California's rarest white oak: the Engelmann oak. *Fremontia*, **18**, 26–29.

Scott, T.A. (1991) The distribution of Engelmann oak (*Quercus engelmannii*) in California. *Proceedings of the Symposium on Oak Woodlands and Hardwood Rangeland Management* (ed. by R.B. Standiford), pp. 351–359. US Department of Agriculture, Forest Service Pacific Southwest Forest & Range Experimental Station, Berkeley, CA.

Shriver, M.D., Smith, M.W., Jin, L., Marcini, A., Akey, J.M., Deka, R. & Ferrell, R.E. (1997) Ethnic-affiliation estimation by use of population-specific DNA markers. *American Journal of Human Genetics*, **60**, 957–964.

Stebbins, G.L., Jr, Matzke, E.B. & Epling, C. (1947) Hybridization in a population of *Quercus marilandica* and *Quercus ilicifolia*. *Evolution*, **1**, 79–88.

Swenson, N.G. (2006) GIS-based niche models reveal unifying climatic mechanisms that maintain the location of avian hybrid zones in a North American suture zone. *Journal of Evolutionary Biology*, **19**, 717–725.

Swenson, N.G. & Howard, D.J. (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *The American Naturalist*, **166**, 581–591.

Swenson, N.G., Fair, J.M. & Heikoop, J. (2008) Water stress and hybridization between *Quercus gambelii* and *Quercus grisea*. *Western North American Naturalist*, **68**, 498–507.

Tucker, J.M. (1952) Taxonomic interrelationships in the *Quercus dumosa* complex. *Madroño*, **11**, 234–251.

Tucker, J.M. (1953) The relationship between *Quercus dumosa* and *Quercus turbinella*. *Madroño*, **12**, 49–60.

Vähä, J.P. & Primmer, C.R. (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.

Van Valen, L. (1976) Ecological species, multispecies, and oaks. *Taxon*, **25**, 233–239.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868–2883.

Warren, D.L., Glor, R.E. & Turelli, M. (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, **33**, 607–611.

Wheeldon, T.J., Patterson, B.R. & White, B.N. (2010) Sympatric wolf and coyote populations of the western Great Lakes region are reproductively isolated. *Molecular Ecology*, **19**, 4428–4440.

Wiens, J.A., Stralberg, D., Jongsomjit, D., Howell, C.A. & Snyder, M.A. (2009) Niches, models, and climate change: Assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences USA*, **106**, 19729–19736.

Williams, J.H., Boecklen, W.J. & Howard, D.J. (2001) Reproductive processes in two oak (*Quercus*) contact zones with different levels of hybridization. *Heredity*, **87**, 680–690.

Zeng, Y.F., Liao, W.J., Petit, R.J. & Zhang, D.Y. (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 in the online version of this article:

Appendix S1 Polymorphism characteristics of microsatellite loci used to type Engelmann and scrub oaks.

BIOSKETCH

Joaquín Ortego is currently a postdoctoral fellow at Estación Biológica de Doñana (CSIC) in Seville (Spain). His main research interests focus on the study of genetic variability in natural populations at different spatio-temporal and phylogenetic scales.

The main research interest of the team of authors is centred on the study of population genetics, phylogeography, genomics, palaeoecology and distribution of trees (https://www.eeb.ucla.edu/Faculty/Sork/Sorklab/).

Author contributions: J.O. and V.L.S conceived the ideas; J.O., PFG and VLS collected the data; J.O. and E.C.R. analysed the data; and J.O. led the writing. All authors made significant comments on and improvements to the manuscript.

Editor: Brent Emerson