DR. JOAQUIN ORTEGO (Orcid ID : 0000-0003-2709-429X)

Article type : From the Cover

Incorporating interspecific interactions into phylogeographic models: A case study with Californian oaks

Joaquín Ortego[1] and L. Lacey Knowles[2]

[1]Department of Integrative Ecology, Estación Biológica de Doñana, EBD-CSIC, Avda. Américo Vespucio 26, E-41092 Seville, Spain (joaquin.ortego@csic.es)

[2]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA (knowlesl@umich.edu)

Author for correspondence:
Joaquín Ortego
Estación Biológica de Doñana, EBD-CSIC,
Avda. Américo Vespucio 26, E-41092 Seville, Spain
E-mail: joaquin.ortego@csic.es
Phone: (+34) 954 232 340
Fax: (+34) 954 621 125

Running title: Species interactions and phylogeography

Abstract

It has been long assumed that abiotic factors (i.e., geography and climate) dominate the ecological and evolutionary processes underlying the distribution of species, lineages and genes at broad spatial scales and, as a result, the study of interspecific interactions has largely been overlooked in biogeography research and ignored entirely in phylogeographic inference. Here, we focus on plant-plant interactions and test whether their demographic consequences translate into broad-scale patterns of genomic variation in two Californian oak species. With our process-based analyses and statistical comparison of the likelihoods of alternative models, we show that spatial patterns of genomic variation are better explained by demographic scenarios incorporating interspecific interactions (including both competition and facilitation) than by null models that only consider heterogeneity of environmental suitability across the landscape. Collectively, our integrative approach supports the notion that the consequences of biotic interactions transcend much larger geographical and evolutionary scales than the traditional local focus.

KEYWORDS

Biotic interactions, competition, community ecology, demographic inference, facilitation, genetic variation, phylogeography

1. INTRODUCTION

The study of how organisms interact with landscape heterogeneity at contrasting spatiotemporal scales has figured prominently in our understanding of the ecological and evolutionary processes underlying geographical distribution of genetic variation, population divergence, and the formation of new species (Avise, 2000; Arbogast & Kenagy, 2001). Traditionally, phylogeography has focused on testing alternative hypotheses that link different abiotic (extrinsic) factors (e.g., barriers to dispersal, climate-driven range shifts, etc.) with population genetic structure (Avise, 2000; Knowles, 2009). More recently, conceptual frameworks have advocated for the importance of building and testing refined hypotheses that incorporate taxon-specific

traits (e.g., dispersal capacity, environmental niche, microhabitat preferences, etc.) to capture the biotic (intrinsic) factors structuring genetic variation (Papadopoulou & Knowles, 2016). By integrating the properties of organisms into alternative models, the relative support for the proximate biological processes underlying spatial patterns of genetic variation can be statistically evaluated and inferred, improving the predictive capacity of both distributional and phylogeographic models (Estrada, Morales-Castilla, Caplat, & Early, 2016; Papadopoulou & Knowles, 2016).

Despite these advances in biologically-informed models, an important biotic aspect has been essentially ignored in phylogeography research – namely, interspecific interactions (Wisz et al., 2013). As a result, we know virtually nothing about the role of this key biological component in structuring genetic variation. Only in taxa with highly specialized and tight relationships have studies attempted to address this question, and even among this class of interactors, we have very few examples (e.g., host-parasite interactions: Tsai & Manos, 2010; symbionts: James et al., 2011). The paucity of studies on the effect of species interactions on spatial patterns of genetic diversity and structure contrasts with the well-established demographic consequences of interspecific interactions within and across trophic levels from theoretical and empirical studies in classical ecological and evolutionary research (e.g., Godoy, Kraft, & Levine, 2014; Miriti, Wright, & Howe, 2001; Maynard, Wootton, Servan, & Allesina, 2019). This in part could reflect the arguments about the relative importance of species interactions beyond local spatial and temporal scales, with abiotic factors such as climate and geography presumably predominating at the large geographical extents at which species and population divergence occurs (Pearson & Dawson, 2003; Soberón, 2007). However, there is no reason to think that the demographic and evolutionary consequences of interspecific interactions observed at local spatial scales would not translate to broader geographical and temporal (i.e., evolutionary) scales (see Svenning et al., 2014; Godsoe, Jankowski, Holt, & Gravel, 2017) and accumulating empirical evidence point to their important role in determining species distributions (for a thorough review see Wisz et al., 2013). Moreover, it is also now broadly recognized that ignoring interspecific interactions (i.e. the community context) will most likely lead to misleading predictions about the impacts of global change on biodiversity (Gilman, Urban, Tewksbury, Gilchrist, & Holt, 2010).

The other factor that contributes to the paucity of studies on the effects of biotic interactions on population genetic structure is simply that there is not a straightforward, or obvious, approach for quantifying their potential role. Despite these challenges (and admittedly simplifying assumptions that will no doubt be necessary), it is also true that without a step toward integrating species interactions into demographic inference, we may not only be mis-ascribing their effects on genetic variation to other processes, but we may also be missing the opportunity to obtain realistic predictions about how populations, species and whole communities will respond to the many different components of ongoing global change from both a demographic (Espindola et al., 2012) and an adaptive perspective (Browne, Wright, Fitz-Gibbon, Gugger, & Sork, 2019).

Here, we focus on plant-plant interactions and their characterization by spatially explicit models for two oak species (genus *Quercus*) from the California Floristic Province (CFP) to test whether their demographic consequences translate into broad-scale patterns of genomic variation. We construct models aimed at capturing negative and positive species interactions given that both are key ecological processes that structure plant assemblages (Callaway & Walker, 1997), including forest communities (e.g., Leathwick & Austin, 2001; Cavender-Bares, Ackerly, Baum, & Bazzaz, 2004; Cavender-Bares, Kozak, Fine, & Kembel, 2009; Pollock, Bayly, & Vesk, 2015). For example, negative interactions (e.g., competition for limited resources, negative allelopathy, etc.) can reduce carrying capacities of subdominant species (e.g., Miriti et al., 2001), whereas positive interactions (e.g., nurse effects, enhancement of the chemical, physical or microbial environment, etc.) can facilitate seedling establishment and increase population growth rates and species expansion (Callaway, 1995). We also incorporate models that account for differences in species relatedness in mediating the direction and strength of interspecific interactions, which has been addressed by ecological studies that consider the phylogenetic context of interactions, albeit with mixed conclusions (i.e., phylogenetic niche conservatism; e.g., Valiente-Banuet & Verdú, 2007; Cahill, Kembel, Lamb, & Keddy, 2008; Godoy et al., 2014).

For competing models of genomic variation that integrate hypothetical positive and negative interactions, we specifically consider how other congeneric species (i.e. oak-oak interactions) impact the demography of two focal oak taxa widely distributed

in the CFP – *Quercus berberidifolia* (section *Quercus*) and *Quercus chrysolepis* (section *Protobalanus*) (Manos, 1997; Nixon & Muller, 1997; Nixon, 2002). The two species belong to different sections that also differ with respect to their species richness within the CFP, with 12 species in section *Quercus vs*. 4 species in section *Protobalanus* (Manos, 1997; Nixon & Muller, 1997; Denk, Grimm, Manos, Deng, & Hipp, 2017) (Figure 1). Because interspecific gene flow generally only takes place among species within the same section (Manos, Doyle, & Nixon, 1999; Nixon, 2002; Pham, Hipp, Manos, & Cronn, 2017), the two focal taxa also differ with respect to the number of closely related species they have the potential to hybridize with. Thus, by selecting these species, our tests can be used to examine the effects of phylogenetic relatedness (i.e., comparing hypothetical interactions exerted by oak species belonging to the same *vs*. different taxonomic sections than the focal taxa), as well as species-specific interactions (i.e., they provide independent tests of either the positive or negative effects of species interactions). This makes our study especially well-suited for testing alternative biogeographic scenarios from a comparative perspective about the potential role of phylogenetic relatedness on interspecific interactions, and how these impact range-wide patterns of genomic variation. Nevertheless, we acknowledge that there are a lot of unknowns and consequent assumptions that we must make in this study, the caveats of which are discussed thoroughly in the context of our findings and conclusions. As such, this work should be viewed as providing insights into the potential impact of species interactions on broad-scale genomic variation, which itself is novel and opens new avenues of research in phylogeographic inference from a community-level perspective. We discuss the utility of our analytical framework for stimulating future independent research aimed at corroborating the nature (i.e. underlying mechanisms) of the interspecific interactions we test here and whether the direction of these interactions (or lack of such interactions) depend upon the phylogenetic relatedness with the focal taxa.

2. MATERIALS AND METHODS

2.1. Population sampling and genomic library preparation and processing

Between 2010 and 2014, we sampled eight populations of California scrub oak (*Quercus berberidifolia*) (*n* = 63 individuals) and ten populations of canyon live oak (*Q. chrysolepis*) (*n* = 80 individuals) representative of their respective distributions in California (Manos, 1997; Nixon & Muller, 1997) (Figure 2; Table S1). We used a mixer mill to grind ~50 mg of frozen leaf tissue in tubes with a tungsten bead and performed DNA extraction and purification with NucleoSpin Plant II kits (Macherey-Nagel, Düren, Germany). We processed genomic DNA into genomic libraries using the double-digestion restriction-fragment-based procedure (ddRADseq) described in Peterson et al., (2012) (Methods S1) and used the different programs distributed as part of the STACKS v. 1.35 pipeline (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013) to filter and assemble our sequences into *de novo* loci and call genotypes (Methods S2).

2.2. Quantifying population genetic structure

We analysed population genetic structure of the two focal species using the Bayesian Markov chain Monte Carlo clustering method implemented in the program STRUCTURE v. 2.3.3 (Pritchard, Stephens, & Donnelly, 2000). We ran STRUCTURE assuming correlated allele frequencies and admixture without using prior population information. We conducted 15 independent runs with 200,000 MCMC cycles, following a burn-in step of 100,000 iterations, for each of the different possible *K* genetic clusters (from *K* = 1 to *K* = 10). We retained the ten runs having the highest likelihood for each value of *K* and inferred the number of populations best fitting the dataset using log probabilities [Pr(X|*K*)] (Pritchard et al., 2000) and the Δ*K* method (Evanno, Regnaut, & Goudet, 2005). To complement and confirm the results yielded by Bayesian clustering analyses (see Janes et al., 2017), we performed a principal component analysis (PCA) as implemented in the R v. 3.3.2 (R Core Team, 2020) package *adegenet* (Jombart, 2008). Before running PCAs, we scaled and centred allele frequencies and replaced missing data with mean allele frequencies using the *scaleGen* function as recommended by Jombart (2008).

2.3. Incorporating interspecific interactions into phylogeographic models

Species interactions (positive, negative or neutral), as well as the magnitude of their effects (which may vary depending on the number of species that overlap in distribution with the focal taxa), were incorporated into a spatiotemporally explicit integrative distributional, demographic and coalescent (iDDC) modelling framework (He, Edwards, & Knowles, 2013) (Figure 1). To account for the impact of environmental heterogeneity across space and time on genomic variation, we translated current and last glacial maximum (LGM) suitability maps obtained for each focal taxon via environmental niche modelling (ENM) into layers of carrying capacities (see He et al., 2013). To model the effects of species-interactions (or their lack thereof) under different hypothetical scenarios, the local carrying capacities of the focal taxa across their respective distributions and time periods (LGM to present) remained unaltered (i.e., a null model of no species-interaction effects) or either increased (positive interactions) or decreased (negative interactions) in the presence of other oak species, whose distributions were also estimated through ENMs (see below for details).

Because the nature of species interactions may differ as a function of phylogenetic relatedness, we tested eight hypothetical interaction models (plus the null model) representing of a diverse suite of alternative scenarios that included the potential importance of phylogenetic relatedness (i.e., to belong or not to the same taxonomic section than the focal taxon) on the direction of interspecific interactions (Table 1; Figure S1). Note that the impact of species phylogenetic relatedness on the direction of interactions is mixed across different studies; some have supported, whereas others have rejected, the hypothesis that more distantly related species show lower niche overlap and compete less strongly than recently diverged species with more similar phenotypes and shared resource requirements (e.g., Cavender-Bares et al., 2004; Valiente-Banuet & Verdú, 2007; Cahill et al., 2008; Godoy et al., 2014; Narwani et al., 2017). As such, the specific models explored here consider (i) similar positive or negative interactions with all other oak species regardless of their phylogenetic relatedness (i.e., taxonomic section) with the focal taxon, and (ii) interactions in which co-distributed species belonging to either the same or different sections as the focal taxon exert contrasting effects (positive, negative or neutral) (Table 1; Figure S1).

The demographic consequences of species interactions (i.e., effects on local carrying capacities) and subsequent genetic expectations under each scenario were generated via spatiotemporally explicit coalescent-based simulations (1×10^6 simulations per model) as implemented in SPLATCHE2 (Ray, Currat, Foll, & Excoffier, 2010) and compared with empirical genomic data within an approximate Bayesian computation (ABC) framework (Beaumont, Zhang, & Balding, 2002) in order to determine the relative statistical support of each model and estimate the posterior distribution of the demographic parameters of the spatially explicit coalescent (e.g., Bemmels, Title, Ortego, & Knowles, 2016; He et al., 2013; Knowles & Massatti, 2017). In the next sections we provide all the specific details about the construction of the alternative phylogeographic scenarios, spatiotemporally explicit simulations, parameter estimation, and model testing and validation (also see illustrative summary of the general workflow in Figure 1).

2.3.1. Environmental niche modelling

We used the maximum entropy algorithm from MAXENT v. 3.3.3 (Elith et al., 2006, 2011; Phillips, Anderson, & Schapire, 2006) implemented in the R package *dismo* v. 1.1-4 (Hijmans, Phillips, & Elith, 2017) to build environmental niche models (ENMs) and generate suitability maps for both the present and the last glacial maximum (LGM, 21.5 ka) for each of our two focal taxa. We also built ENMs for each of the other oak species from California (Jensen, 1997; Manos, 1997; Nixon & Muller, 1997; Figure 1) and used projections of their geographical distributions during the present and the LGM to generate phylogeographic models of their potential hypothetical effects on our two focal taxa as detailed below. To build the models, we used species occurrence data from our own records, as well as those available at the Global Biodiversity Information Facility (http://www.gbif.org/), Calflora database (http://www.calflora.org/), the Consortium of California Herbaria (http://ucjeps.berkeley.edu/consortium/), the Consortium of Pacific Northwest Herbaria (http://www.pnwherbaria.org/) and the University of Arizona Herbarium (http://ag.arizona.edu/herbarium/) (Table S2). As environmental layers, we used the 19 bioclimatic variables available in WORLDCLIM v. 1.4 at 30 arc-second resolution (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) plus a layer of slope generated using

ARCMAP v. 10.2.1 (ESRI, Redlands, CA, USA) from a 30 arc-second digital elevation and bathymetry model (Becker et al., 2009). We conducted species-specific model parameter tuning using the R package *ENMeval* (Muscarella et al., 2014). Specifically, for each species, we tested a total of 248 models of varying complexity by combining a range of regularization multipliers (RM) (from 0 to 15 in increments of 0.5) with eight different feature classes (FC) combinations (L, LQ, LQP, H, T, LQH, LQHP, LQHPT, where L = linear, Q = quadratic, H = hinge, P = product and T = threshold; Muscarella et al., 2014). We compared MAXENT models with different settings using the Akaike information criterion corrected for small sample size (AICc) (Burnham & Anderson, 2002; Warren & Seifert, 2011). We performed a three-stage approach to select the species-specific set of environmental variables and model parameters (RM and FC) (Warren, Wright, Seifert, & Shaffer, 2014). In a first step, we built a full set of models including all variables, retained the model with the lowest AICc score, and among those variables that were spatially correlated (Pearson's correlation coefficient > 0.7, estimated using ENMTOOLS; Warren, Glor, & Turelli, 2010) we only retained for the next step the one with the highest percent contribution to the model. In a second step, we ran another full set of models with the subset of variables retained in the first step, selected the model with the lowest AICc score, and (if any) removed variables with zero percent contribution to the model. In a third step, we re-ran a final full set of models with the environmental variables retained in the previous step and used for all downstream analyses the model with the lowest AICc score. We projected final models for each species to the last glacial maximum (LGM) conditions derived from the Community Climate System Model v. 4 (CCSM4; Gent et al., 2011), which has been shown to perform well for predicting terrestrial climate conditions during this period (Harrison et al., 2014). To create maps of presence/absence for the species that may interact with the focal taxa, we converted the logistic output from MAXENT into binary maps (Figures S2 and S3) using the maximum training sensitivity plus specificity (MTSS) threshold values for occurrence obtained for each oak species (Table S2; see Liu, Berry, Dawson, & Pearson, 2005).

2.3.2. Translating ENMs into alternative phylogeographic models

We used information from ENMs to describe geographic variation in carrying capacities for our two focal species. For the null model of no species-interaction (Model A), the carrying capacities (*K*) of demes were scaled proportionally to logistic habitat suitability scores (ranging from 0 to 1) obtained from MAXENT for each focal species (e.g., Bemmels et al., 2016; González-Serna, Cordero, & Ortego, 2019; Knowles & Massatti, 2017; Massatti & Knowles, 2016). In models considering interspecific interactions, carrying capacities of the focal species were reduced (negative interactions) or increased (positive interactions) by the presence of other oak species in the same grid cell (Table 1). Specifically, the effect of each oak species projected to be present in the same grid cell (based on species-specific ENMs) as the focal taxon was modeled by either reducing (negative interaction) or increasing (positive interaction) the habitat suitability of the focal species by 0.05 (i.e., 5% from a maximum *K* of 100%). Although the magnitude of the potential effect of each oak species on the focal taxa is admittedly arbitrary, this value was selected because it is one that generated statistically distinguishable models of biological significance (see Papadopoulou & Knowles, 2016). Specifically, visual inspection of habitat suitability maps under the different scenarios suggested that smaller values did not result in any appreciable differences in the spatial distribution of carrying capacities among scenarios, whereas larger values would produce gaps in the distribution of the focal taxa when modeling negative interactions or resulted in little heterogeneity in local carrying capacities across the landscape when modelling positive interactions. In all models, the negative or positive impact of other oak species was always bounded within the range (0-1) of habitat suitability scores provided by the logistic output of MAXENT (i.e., the negative and positive effects of other oak species never increased the probability of occurrence of the focal species above one, or a k = 100%, or reduced it below zero, or a k = 0%, respectively). In other words, the parameter space in which the effects of overlapping with multiple species (as opposed to limited overlap) with the focal taxa was constrained. We recognize that our models do not capture other more complex interactions (e.g., multiplicative interactions or varying effects by species) and assume the positive or negative effects of potential interactions vary as a function of the number of species with distributional overlap (see Figure S1). Nevertheless, by capturing the potential effects of community composition on the

focal taxa, our models provide a good starting point for examining the potential effects of species interactions on broad-scale patterns of genetic variation of the focal taxa. It is in this spirit (and in recognition of all the assumptions about the nature of species interactions) that there is merit in the approach we apply here.

2.3.3. Spatiotemporally explicit simulations

We used the integrative distributional, demographic and coalescent (iDDC) framework (He et al., 2013), which applies SPLATCHE2 (Ray et al., 2010), to generate genetic expectations for the nine alternative models we test here (Table 1 and Figure S1) where the habitat suitabilities, and hence, carrying capacities, for the two focal species differ through time and across the landscape. For each model (see Figure 1), demographic simulations are carried out in which the suitability of the landscape varies across three temporal periods, i.e., input from ENMs incorporated based on bioclimatic/paleobioclimatic data for the LGM and present (e.g., Bemmels et al., 2016; Knowles & Massatti, 2017). For the intervening time period, we generated a new raster map with intermediate habitat suitability values between current and LGM layers obtained under each scenario. Habitat suitability bins corresponding to each of the three temporal periods (LGM, intermediate, current) were applied to one-third of the total number of simulated generations (see Figure 1 and 3).

To have a computationally tractable number of demes for demographic simulations, we statistically downscaled cell sizes to 5-arcminutes (~9 km^2) (e.g., Massatti & Knowles, 2016). Given that SPLATCHE2 requires a single raster file with positive integer numbers, we first categorized cell values (ranging continuously from 0 to 1) under each scenario and time period into 20 bins of equal magnitude (i.e., intervals of 0.05) with ARCMAP v. 10.2.1 and used a custom PYTHON script written by Q. He (deposited in Dryad; Bemmels et al., 2016) to convert the maps from the different time periods into a single raster map in which each category (LGM, intermediate, current) represents a unique combination of habitat suitability bins across the three time periods (e.g., He et al., 2013; Bemmels et al., 2016; Massatti & Knowles, 2016). Assuming a generation time of 50 years for oaks (Ortego, Noguerales, Gugger, & Sork, 2015; Bemmels et al., 2016; Ortego, Gugger, & Sork, 2018), a total of 430 generations from the LGM to present (21.5 ka) was modeled for each scenario with 1×10^6

simulations (9×10^6 simulations per species) generated using the same uniform priors for the three demographic parameters of the spatially explicit coalescent: the migration rate per deme per generation (m; range of log(m): -2.0, -0.2), the maximum carrying capacity of a deme (K_{MAX}, which is the value for demes with the highest suitability value; range of log(K_{MAX}): 2.9, 3.7), and the ancestral population size (N_{ANC}; range of log(N_{ANC}): 2.5, 5.5). The parameter space defined by the prior was chosen based on pilot runs across a broad parameter space which identifed parameters in which the colonization of the landscape within the time spanning from the LGM to the present generated genetic data within the range of observed empirical data (e.g., Bemmels et al., 2016). Demographic simulations were initialized 21.5 ka from hypothesized ancestral source populations for each focal species. These source populations corresponded to grid cells of the LGM map with habitat suitabilities higher than the 10th percentile of habitat suitability values of all grid cells of the current map containing an occurrence record (see Brown & Knowles, 2012). The carrying capacities of source populations were defined according to their habitat suitabilities during the LGM and categorized into the same bins described above for layers corresponding to each of the three temporal periods.

Following each time-forward demographic simulation (see Figure 1), a spatially-explicit time-backward coalescent model informed by the deme-specific demographic parameters (K, m and N_{ANC}) was used to generate genetic data (Currat, Ray, & Excoffier, 2004; Ray et al., 2010). To make simulations computationally tractable, we randomly selected 1,250 loci for each focal taxon (e.g., Massatti & Knowles, 2016) and run an independent coalescent process to trace the genealogy for each locus from the present to the onset of population expansion from ancestral source populations 21.5 ka, with an additional period of 10^7 generations for all alleles to coalesce in a single ancestor (Ray et al., 2010). Simulated datasets were sampled from the same geographical locations (grid cells) from which the empirical genomic data were obtained (Table S1) and consisted of the same number of loci, number of individuals, and amount and pattern of missing data as the empirical data (see Massatti & Knowles, 2016). Finally, we used ARLSUMSTAT v. 3.5.2 to calculate a set of summary statistics for each empirical and simulated dataset, including the mean heterozygosity across loci for each population (H), the number of segregating sites for each population

(*S*), and the pairwise population F_{ST} -values (Excoffier & Lischer, 2010), for a total of 44 summary statistics for the eight populations of *Q. berberidifolia*, and 65 summary statistics for the ten populations of *Q. chrysolepis* (the different number of summary statistics reflects the larger number of sampled populations of *Q. chrysolepis*). All simulations were performed on the high-performance computing cluster from Centro de Supercomputación de Galicia (CESGA, Spain) and required ~432,000 hours of CPU time (i.e., ~24,000 CPU hours per model and species).

2.4. Model selection and parameter estimation

We used ABC for model selection and parameter estimation, as implemented in ABCTOOLBOX (Wegmann, Leuenberger, Neuenschwander, & Excoffier, 2010). We used the R (R Core Team, 2020) package *pls* v. 2.6-0 (Mevik & Wehrens, 2007) and the *findPLS* script (Wegmann et al., 2010) to extract partial least squares (PLS) components (with Box-Cox transformation) from the summary statistics of the first 10,000 simulations for each model and species (Wegmann, Leuenberger, & Excoffier, 2009). The first four PLSs extracted from the summary statistics were used for ABC analyses, as the root-mean-squared error (RMSE) of the three demographic parameters (K_{MAX}, m, N_{ANC}) for the two species did not decrease significantly with additional PLSs (Figures S4 and S5). The linear combinations of summary statistics obtained from the first 10,000 simulations for each model and species were used to transform all simulated datasets (Wegmann et al., 2010). For each model and species, the 5,000 simulations (0.5%) closest to empirical data were retained and used for model selection and to obtain posterior distributions of the parameters with an ABC-GLM adjustment (Leuenberger & Wegmann, 2010). We used Bayes factors (BF) for model selection (Jeffreys, 1961; Kass & Raftery, 1995).

2.5. Model validation

To evaluate the ability of each model to generate the empirical data, we calculated Wegmann's *p*-value from the 5,000 retained simulations (Wegmann et al., 2010). We also assessed the potential for a parameter to be correctly estimated by computing the

proportion of parameter variance that was explained (i.e., the coefficient of determination, R^2) by the retained PLSs (Neuenschwander et al., 2008). For the most supported model for each species, we determined the accuracy of parameter estimation using a total of 1,000 pseudo-observation datasets (PODs) generated from prior distributions of the parameters. If the estimation of the parameters is unbiased, posterior quantiles of the parameters obtained from PODs should be uniformly distributed (Wegmann et al., 2010). As with the empirical data, the posterior quantiles of true parameters for each pseudo run were calculated based on the posterior distribution of the regression-adjusted 5,000 simulations closest to each pseudo-observation.

3. RESULTS

3.1. Genomic data

After quality filtering, we retained a total of 102,086,259 reads for *Q. berberidifolia* (mean ± SD = 1,620,416 ± 328,146 reads per individual) and 119,011,704 reads for *Q. chrysolepis* (mean ± SD = 1,487,646 ± 259,978 reads per individual) (Figure S6). After filtering loci, the final datasets contained 3,589 SNPs for *Q. berberidifolia* and 2,977 SNPs for *Q. chrysolepis*. The proportion of missing data in individuals of *Q. berberidifolia* and *Q. chrysolepis* averaged 1.77 % and 1.52 %, respectively.

3.2. Genetic structure

For *Q. berberidifolia*, log probabilities [Pr(X|*K*)] from STRUCTURE analyses reached a plateau for $K = 2$ and the ΔK statistic indicated an "optimal" clustering for the same *K*-value (Figure S7a). The two genetic clusters presented some degree of genetic admixture and showed a latitudinal cline of genetic differentiation (Figure 2a), which was also supported by the PCA (Figure S8a) and previous microsatellite-based studies (Ortego et al., 2015; Ortego, Gugger, & Sork, 2017). For *Q. chrysolepis*, log probabilities [Pr(X|*K*)] reached a plateau for $K = 3$, a *K*-value also identified by the ΔK statistic as the "optimal" clustering solution (Figure S7b). As shown in previous studies on this species

(Bemmels et al., 2016; Ortego et al., 2018), the three genetic clusters were structured hierarchically and presented considerable genetic admixture in geographic areas of contact (Figure 2b). The two genetic clusters identified for K = 2 separated populations located north and south of the Transverse Ranges, whereas the third genetic cluster was mostly represented in the North Coast Ranges and in admixed populations from adjacent regions (northern Sierra Nevada and South Coast Ranges) (Figure 2b). PCA yielded analogous results. Namely, populations grouped into three main genetic clusters and populations with high admixed ancestry (HAS, SHA, and TAH; Figure 2b) occurred at intermediate positions along the main axes (PC1 and PC2) of genomic variation (Figure S8b).

3.3. Phylogeographic model testing and validation

Environmental niche models predicted well the current distribution of the different species (Figure S2; Table S2; Jensen, 1997; Manos, 1997; Nixon & Muller, 1997). As shown in previous studies on different Californian organisms (e.g., Ortego et al., 2015; Starrett, Hayashi, Derkarabetian, & Hedin, 2018), projections of ENMs to the LGM predicted that most species likely experienced local distributional shifts in response to Pleistocene glaciations (Figure S3).

Based on marginal densities calculated from the 5,000 simulations retained for each model and focal species, the best fitting model differed between taxa (Table 1). Specifically, for *Q. berberidifolia*, the model with a negative effect of all other oak species (Model B) was the best fit with the empirical data (Table 1; Figure 3a). The second and third most supported models were also those in which co-distributed species have a negative effect on the focal taxon (i.e., Model C and D, where the negative effect was associated with taxa from the same or a different taxonomic section as the focal taxon, respectively; Table 1). However, these two models had considerably lower marginal densities and a difference in Bayes factors > 25 with the most supported model in which all species negatively affect the focal taxon (Table 1), indicating strong support for Model B (Jeffreys, 1961; Kass & Raftery, 1995). Moreover, Model B was the only one in which the simulated genetic data were comparable with empirical data (Wegmann's p-value = 0.705), unlike the other models in which there

was a substantial difference between the likelihoods of the simulated data compared with the empirical data (Wegmann's *p*-values < 0.06; Table 1).

For *Q. chrysolepis*, the model that best explained the data was one in which co-distributed species from the same section had a positive effect on the focal taxon, whereas species from different sections had the opposite effect (Model I; Table 1; Figure 3b). However, three other models (Models D, F, B) also fit the data; small Bayes factors (BF<8) (Table 1) suggests that they are statistically indistinguishable from the best supported model (Kass & Raftery, 1995). Two of these models represent the individual components that Model I integrates; that is, negative effects of species from different sections (Model D) *vs.* positive effects of species within the same section (Model F). The third supported model was one in which all species negatively affect the focal taxon *Q. chrysolepis* (Model B) (Table 1). All of these models were capable of generating data compatible with empirical data (Wegmann's *p*-values > 0.1), in contrast with the very low Wegmann's *p*-values (< 0.05) obtained for the rest of models, which also were not probable models (BF > 5,000; Table 1).

Posterior distributions of parameters under the most probable models were considerably distinct from the prior, indicating that the simulated data contained information relevant to estimating the parameters (Figure 4). Comparison of the posterior distributions before and after the ABC-GLM also showed the improvement that this procedure had on parameter estimates (Figure 4). In the two focal species, the posterior distributions of maximum carrying capacity (K_{MAX}) and migration rates (m) were flatter than those obtained for ancestral population size (N_{ANC}), indicating that the former parameters were estimated at a comparatively lower precision (i.e., higher uncertainty). The coefficients of determination (R^2) between each demographic parameter and the four extracted PLS indicated that the employed summary statistics had a high potential to correctly estimate all the parameters (Table 1). However, the histograms of the posterior quantiles of m in *Q. berberidifolia,* and N_{ANC} in both focal taxa, significantly deviated from a uniform distribution, suggesting a potential bias in the estimation of these parameters (Figure S9).

4. DISCUSSION

Our process-based analyses indicate that spatial patterns of genomic variation in the two focal taxa are better explained by demographic models that incorporate interspecific interactions than by null models that only consider heterogeneity of environmental suitability across the landscape. In fact, models with no species interactions provided a very poor fit to our empirical data (Wegmann's *p*-values < 0.01), indicating that such models are not able to reproduce the demographic processes experienced by the focal taxa. Collectively, our results support the hypothesis that interactions with other congeneric taxa shape species´ distributions and range-wide patterns of genetic variation. Our study makes specific assumptions when modelling the potential effects of species interactions (e.g., it captures community-wide effects, but not taxon-specific or multiplicative interaction effects), which imposes constraints on making conclusions about the precise mechanisms involved (thoroughly discussed below). Nonetheless, our integrative approach provides empirical support not only for the demographic, but also the evolutionary consequences of interspecific interactions that transcend much larger geographical and evolutionary scales than the traditional local focus (Jablonski, 2008; Wisz et al., 2013; Araujo & Rozenfeld, 2014).

4.1. Predominance of negative species-interactions

Although previous research suggests that niche partitioning can minimize negative interactions among closely related taxa (e.g., Cavender-Bares et al., 2004, 2018), our analyses indicate that such interactions can still play an important role in limiting species´ distributions and shaping their range-wide patterns of genomic variation. The most supported models for each of the two focal taxa were dominated by the negative effects of co-distributed species, which in our framework are modelled as reductions in local population sizes. Different mechanisms can explain the inferred reduction of local carrying capacities of the focal taxa exerted by other congeneric species, including competition for resources in limited supply (Craine & Dybzinski, 2013), alteration of biotic and abiotic soil properties that reduce their competitive performance (Bennett & Cahill 2016), and increased impact of phytophagous insects and infectious diseases shared with closely related species in the community (Yguel et al., 2011). In wind

pollinated trees separated by weak reproductive barriers, the genetic neighbourhood can be several orders of magnitude larger than the ecological neighbourhood and, as a result, interspecific interactions are not limited to narrow local scales (Levin, 2006). Accordingly, hybridization could reduce species performance and abundance through reproductive interference (Levin, 2006; Pollock et al., 2015) and genetic or demographic swamping by the most abundant congener (Levin, Francisco-Ortega, & Jansen, 1996; Rhymer & Simberloff, 1996; Louthan, Doak, & Angert, 2015). It should be noted that the two focal taxa studied here are keystone and dominant species in different ecosystems from the CFP and, thus, negative interactions are expected to play even a more prominent role in shaping the distribution of genetic variation in subdominant species such as herbs or small scrubs (DeBach, 1966).

4.2. Taxon-specific interactions and corroborative evidence from other studies

Our model-based comparative phylogeography framework has also proven useful to unravel taxon-specific effects of interspecific interactions. Interpreted in the light of the contrasting life-histories and ecologies of the taxa involved, such results can provide important biological insights into the processes structuring genomic variation (Papadopoulou & Knowles, 2016) and, ultimately, may help to forecast the idiosyncratic demographic responses of species to environmental change (Gilman et al., 2010; Estrada et al., 2016). Although demographic models that best fitted empirical genomic data for the two focal taxa were mostly dominated by negative interspecific interactions, the two taxa also presented some notable differences. For example, although the most supported model for the California scrub oak (*Q. berberidifolia*) was the one considering a negative effect of all other oak species, the spatial distribution of genomic variation in the canyon live oak (*Q. chrysolepis*) was best explained by a scenario combining a negative impact of species from different sections and positive effects by closely related species within the same section. These differences are especially intriguing when the natural histories of the focal taxa are considered. *Quercus berberidifolia* is a scrubby oak (< 2 m of height) that is often the dominant species in chaparral formations and the margins of coastal sage scrub habitats where tree life-forms are absent (Nixon & Muller, 1997). In tree-dominated habitats, *Q.*

berberidifolia only persist in forest margins or becomes a subdominant understory species at very low densities, suggesting that it experiences competitive displacement (DeBach, 1966). This species has also been recorded to hybridize with most Californian white oaks, including trees (Ortego et al., 2015, 2017; Kim et al., 2018; Nixon & Muller, 1997). Although hybridization with other oak trees from the same section could assist gene flow of our focal species (Potts & Reid 1988), it might not compensate for the negative effects of competitive exclusion (Craine & Dybzinski, 2013) or, as mentioned above, could be responsible for reducing local carrying capacities through reproductive interference (Levin, 2006; Pollock et al., 2015) or demographic swamping in suboptimal habitats dominated by tree oaks (Levin et al., 1996; Rhymer & Simberloff, 1996). In contrast, *Q. berberidifolia* is mostly allopatric or parapatric with respect to the rest of scrub oak taxa from the CFP (Nixon & Muller, 1997), suggesting that any impact on the demography of this focal species is likely to be limited, beyond perhaps sporadic hybridization in narrow contact zones (Ortego et al., 2015, 2017). The only exception is the sister species of *Q. berberidifolia*, the serpentine-soil specialist leather oak (*Q. durata*) (Nixon & Muller, 1997; Ortego et al., 2017). The broad-scale distribution of *Q. durata* is similar to that of *Q. berberidifolia* and the two species are often found living in close geographical proximity, but rarely in the same patches, with the former growing in scattered serpentine outcrops, whereas the latter is unable to form stable populations in such areas (Nixon & Muller, 1997). Hybridization between these two species is common and coalescent-based migration models have supported asymmetric gene flow from *Q. durata* into *Q. berberidifolia*, which has been interpreted as a consequence of low hybrid performance in serpentine soils (Ortego et al., 2017). Thus, *Q. durata* could negatively impact *Q. berberidifolia* through reproductive interference and maladaptive gene flow even if the two species occupy well differentiated edaphic niches (Ting & Cutter, 2018). Although beyond the scope of this study, incorporating more mechanistic models for comparison to the ones considered here would provide a potential way to corroborate the long-term consequences of interspecific gene flow (i.e., demonstrate its impact on range-wide levels of genetic variation).

Evaluation of the relative support of the different demographic scenarios for *Q. chrysolepis* revealed that four models were statistically indistinguishable from each

other (BF < 20) and able to generate data compatible with empirical genomic data (Wegmann's *p*-values > 0.1). These models represent different sides of the same coin and collectively highlight the impact of phylogenetic relatedness (same *vs*. different taxonomic sections) on the inferred interspecific interactions: a positive effect of species within the same section *vs.* negative effects exerted by species from different sections than the focal taxon. An exception is the strong relative support for the model considering a negative effect of all other oak species (Model B). However, given that there are only three other oak species belonging to the same section as *Q. chrysolepis* with somewhat limited geographic and/or ecological overlap (i.e., the narrow endemic Channel Island oak, *Q. tomentella*, the Palmer oak, *Q. palmeri,* and the huckleberry oak, *Q. vaccinifolia*; Manos, 1997), the fit of this model is not entirely unexpected. That is, the expectations in terms of carrying capacities of a model considering a negative effect of all oak species are pretty similar to those from a model in which essentially all but two taxa are modelled to exert a negative effect (see Figure S1). *Quercus chrysolepis* can become large trees (>20 m) and it is often the dominant species in its specific microhabitats (mountain ridges, canyons and moist slopes), whereas the two other species from section *Protobalanus* distributed in continental California have a shrubby life form (Manos, 1997). *Quercus palmeri* is ecologically isolated from *Q. chrysolepis* and interspecific hybridization between the two species has not been recorded in California, suggesting that interactions between these two taxa are probably very limited (Tucker, 1980; Nixon, 2002; Ortego et al., 2018). In contrast, *Q. chrysolepis* is often sympatric with *Q. vaccinifolia* in northern and eastern California where the distribution of the two species overlap and the presence of intermediate individuals resulting from hybridization between them is fairly frequent (Manos, 1997; Nixon, 2002; Ortego et al., 2018). *Quercus vaccinifolia* presents a low spreading scrubby life form (up to 1.5 m) and is often an understory species (Manos, 1997; Mohr, Whitlock, & Skinner, 2000). As a result, it probably receives a massive pollen rain from *Q. chrysolepis*, which could explain anecdotal evidence of asymmetric gene flow from *Q. chrysolepis* into *Q. vaccinifolia* (Ortego et al., 2018). Given that *Q. vaccinifolia* is a cold adapted species living at high elevations (up to 2,800 m; Mohr et al., 2000; Briles, Whitlock, Skinner, & Mohr, 2011), one possibility is that our focal species has benefited from assisted dispersal and postglacial colonization through hybridization with this

637 closely related species (see Potts & Reid, 1988; Petit, Bodenes, Ducousso, Roussel, &
638 Kremer, 2004). Likewise, previous studies on Californian oaks have demonstrated
639 facilitative relationships between shrubs and tree oak seedlings (Callaway, 1992). Thus,
640 another non-mutually exclusive explanation for the observed positive effects is that *Q.*
641 *vaccinifolia* facilitates seedling establishment and increases recruitment rates of *Q.*
642 *chrysolepis* through different nursing effects, including the improvement of the
643 physical environment, protection against herbivores, and enhanced nutrient uptake
644 (Cavender-Bares et al., 2018).
645
646 **4.3. Limitations and future directions**
647
648 It is also important to acknowledge some of the limitations of our model-based
649 framework. First, our approach does not provide mechanistic insights (i.e., we cannot
650 speak about the relative likelihood of different specific processes invoked in the
651 interpretations of our results) because the effects are expressed through the
652 demographic parameter of the focal species – the local carrying capacity.
653 Nevertheless, given that species distributions vary spatially, the demographic
654 consequences of co-distributed species, and hence patterns of genetic variation, as
655 modelled here are fairly specific. For example, changing the relationship between a
656 focal taxon's local population size and the environment (Brown & Knowles, 2012) by
657 itself would not produce similar genetic consequences to those associated with
658 species-interactions. We also caution that conclusions about the relative statistical
659 support of alternative demographic scenarios, including whether models with or
660 without interactions explain better patterns of genomic variation across the landscape,
661 need to always consider uncertainty regarding the strength and nature of the
662 interactions that are modelled here. Likewise, our models ignored many other
663 interspecific interactions, including some recognized in oaks such as
664 competition/facilitation by other non-oak trees (Petritan, Marzano, Petritan, & Lingua,
665 2014), interactions with seed dispersers and predators (Pesendorfer, Sillett, Morrison,
666 & Kamil, 2016), infectious diseases (Rizzo, Garbelotto, Davidson, Slaughter, & Koike,
667 2002), and multiple complex non-mutually exclusive interconnections among them
668 (Shi, Gao, Zheng, & Guo, 2017). In the same line, ENMs are unlikely to capture all

environmental constraints (e.g., adaptive/non-adaptive processes) that plants are responding to (Hampe, 2004), some of which could be spatially correlated with the presence of other oaks species from the community, and which might potentially get confounded with positive/negative interactions in our tested models (Keitt, Bjornstad, Dixon, & Citron-Pousty, 2002; Koenig, 1999). Finally, our approach assumed interspecific interactions to be constant across space and time and of equal magnitude across species within sections, when their intensity is expected to change across environmental gradients and be context- and species-specific (Wisz et al., 2013). However, it must be noted that with an almost infinite number of alternative scenarios that might be tested, which includes incorporating other types of interactions and species-specific strengths and directions, the analyses would become computationally intractable and the selection of one model over another would probably be difficult to interpret and provide few biological insights (Massatti & Knowles, 2016). An interesting line of future research would be to explore how the expectations of alternative joint species distribution models (JSDM) that simultaneously consider a wider range of species-interactions (Pollock et al., 2014) fit to genomic data in comparison with only environment-based niche models. Nevertheless, at this point, the lack of information about species co-occurrence in the past would limit such tests to temporally static models (i.e., one snapshot in time related to the current species distribution; see He et al., 2013). Yet, such an approach could still be useful and worth exploring in highly stable and species-rich regions such as the tropics (Costa et al., 2018).

Acknowledging the limitations inherent to any model-based approach, our integrative framework demonstrates that interspecific interactions leave signals on spatial patterns of genomic variation that can be informative to unravel the evolutionary and ecological processes determining species distributions and community assembly beyond local scales. Collectively, this study opens new avenues of research to integrate the community-context in which species respond to landscape heterogeneity (and shifts in the environment), which is especially relevant to questions where such context has been identified to be a critical factor, as for forecasting the impact of ongoing climate change at different biodiversity levels (Gilman et al., 2010).

ACKNOWLEDGEMENTS

We wish to thank to Paul F. Gugger and Victoria L. Sork for logistic support and help during fieldwork, Amparo Hidalgo-Galiana for genomic library preparation, Sergio Pereira (The Centre for Applied Genomics) for Illumina sequencing, and three anonymous referees for constructive comments on an earlier version of this article. We also thank the US National Park Service and California State Parks for providing sampling permits, Laboratorio de Ecología Molecular from Estación Biológica de Doñana (LEM-EBD) for logistic support, and Centro de Supercomputación de Galicia (CESGA) and Doñana's Singular Scientific-Technical Infrastructure (ICTS-RBD) for access to computer resources. Research was funded by a 2017 Leonardo Grant for Researchers and Cultural Creators, BBVA Foundation (grant number IN17_CMA_CMA_0019). The BBVA Foundation accepts no responsibility for the opinions, statements and contents included in the project and/or the results thereof, which are entirely the responsibility of the authors.

AUTHOR CONTRIBUTIONS

J.O. and L.L.K. conceived the study and designed the research. J.O. collected the samples and produced and analyzed the data. J.O. led the writing with inputs from L.L.K.

DATA ACCESSIBILITY STATEMENT

Raw Illumina reads have been deposited at the NCBI Sequence Read Archive (SRA) under BioProject PRJNA639507. Input files for all analyses (STRUCTURE, PCA, environmental niche modelling, and SPLATCHE2) are available for download from Figshare (https://doi.org/10.6084/m9.figshare.12388781).

ORCID

Joaquín Ortego https://orcid.org/0000-0003-2709-429X

L. Lacey Knowles https://orcid.org/0000-0002-6567-4853

REFERENCES

Araujo, M. B., & Rozenfeld, A. (2014). The geographic scaling of biotic interactions. *Ecography, 37*(5), 406-415. doi:10.1111/j.1600-0587.2013.00643.x

Arbogast, B. S., & Kenagy, G. J. (2001). Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography, 28*(7), 819-825.

Avise, J.C. (2000). *Phylogeography: The History and Formation of Species*. Cambridge, MA, USA: Harvard University Press.

Beaumont, M. A., Zhang, W. Y., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics, 162*(4), 2025-2035.

Becker, J. J., Sandwell, D. T., Smith, W. H. F., Braud, J., Binder, B., Depner, J., . . . Weatherall, P. (2009). Global bathymetry and elevation data at 30 arc seconds resolution: SRTM30_PLUS. *Marine Geodesy, 32*(4), 355-371. doi:10.1080/01490410903297766

Bemmels, J. B., Title, P. O., Ortego, J., & Knowles, L. L. (2016). Tests of species-specific models reveal the importance of drought in postglacial range shifts of a Mediterranean-climate tree: insights from integrative distributional, demographic and coalescent modelling and ABC model selection. *Molecular Ecology, 25*(19), 4889-4906. doi:10.1111/mec.13804

Bennett, J. A., & Cahill, J. F. (2016). Fungal effects on plant-plant interactions contribute to grassland plant abundances: evidence from the field. *Journal of Ecology, 104*(3), 755-764. doi:10.1111/1365-2745.12558

Briles, C. E., Whitlock, C., Skinner, C. N., & Mohr, J. (2011). Holocene forest development and maintenance on different substrates in the Klamath Mountains, northern California, USA. *Ecology, 92*(3), 590-601. doi:10.1890/09-1772.1

Brown, J. L., & Knowles, L. L. (2012). Spatially explicit models of dynamic histories: examination of the genetic consequences of Pleistocene glaciation and recent

climate change on the American Pika. *Molecular Ecology, 21*(15), 3757-3775. doi:10.1111/j.1365-294X.2012.05640.x

Browne, L., Wright, J. W., Fitz-Gibbon, S., Gugger, P. F., & Sork, V. L. (2019). Adaptational lag to temperature in valley oak (*Quercus lobata*) can be mitigated by genome-informed assisted gene flow. *Proceedings of the National Academy of Sciences of the United States of America, 116*(50), 25179-25185. doi:10.1073/pnas.1908771116

Burnham, K. P. & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York, NY, USA: Springer.

Cahill, J. F., Kembel, S. W., Lamb, E. G., & Keddy, P. A. (2008). Does phylogenetic relatedness influence the strength of competition among vascular plants? *Perspectives in Plant Ecology Evolution and Systematics, 10*(1), 41-50. doi:10.1016/j.ppees.2007.10.001

Callaway, R. M. (1992). Effect of shrubs on recruitment of *Quercus douglasii* and *Quercus lobata* in California. *Ecology, 73*(6), 2118-2128. doi:10.2307/1941460

Callaway, R. M. (1995). Positive interactions among plants. *Botanical Reviews*, *61*(4), 306-349. doi:10.1007/bf02912621

Callaway, R. M., & Walker, L. R. (1997). Competition and facilitation: A synthetic approach to interactions in plant communities. *Ecology, 78*(7), 1958-1965. doi:10.1890/0012-9658(1997)078[1958:cafasa]2.0.co;2

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). STACKS: an analysis tool set for population genomics. *Molecular Ecology, 22*(11), 3124-3140. doi:10.1111/mec.12354

Cavender-Bares, J., Ackerly, D. D., Baum, D. A., & Bazzaz, F. A. (2004). Phylogenetic overdispersion in Floridian oak communities. *American Naturalist, 163*(6), 823-843.

Cavender-Bares, J., Kozak, K. H., Fine, P. V. A., & Kembel, S. W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters, 12*(7), 693-715. doi:10.1111/j.1461-0248.2009.01314.x

Cavender-Bares, J., Kothari, S., Meireles, J. E., Kaproth, M. A., Manos, P. S., & Hipp, A. L. (2018). The role of diversification in community assembly of the oaks

(*Quercus* L.) across the continental U.S. *American Journal of Botany, 105*(3), 565-586. doi:10.1002/ajb2.1049

Costa, G. C., Hampe, A., Ledru, M. P., Martinez, P. A., Mazzochini, G. G., Shepard, D. B., . . . Carnaval, A. C. (2018). Biome stability in South America over the last 30 kyr: Inferences from long-term vegetation dynamics and habitat modelling. *Global Ecology and Biogeography, 27*(3), 285-297. doi:10.1111/geb.12694

Craine, J. M., & Dybzinski, R. (2013). Mechanisms of plant competition for nutrients, water and light. *Functional Ecology, 27*(4), 833-840. doi:10.1111/1365-2435.12081

Currat, M., Ray, N., & Excoffier, L. (2004). SPLATCHE: a program to simulate genetic diversity taking into account environmental heterogeneity. *Molecular Ecology Notes, 4*(1), 139-142. doi:10.1046/j.1471-8286.2003.00582.x

DeBach, P. (1966). Competitive displacement and coexistence principles. *Annual Review of Entomology, 11*, 183-212. doi:10.1146/annurev.en.11.010166.001151

Denk, T., Grimm, G.W., Manos, P.S., Deng, M. & Hipp, A.L. (2017). An updated infrageneric classification of the oaks: review of previous taxonomic schemes and synthesis of evolutionary patterns. In: *Oaks Physiological Ecology. Exploring the Functional Diversity of Genus Quercus L., Tree Physiology, vol 7* (eds. Gil-Pelegrín, E., Peguero-Pina, J.J. & Sancho-Knapik, D.). Springer, Cham, pp. 13-38.

Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Guisan, A., . . . Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography, 29*(2), 129-151. doi:10.1111/j.2006.0906-7590.04596.x

Elith, J., Phillips, S. J., Hastie, T., Dudik, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MAXENT for ecologists. *Diversity and Distributions, 17*(1), 43-57. doi:10.1111/j.1472-4642.2010.00725.x

Espindola, A., Pellissier, L., Maiorano, L., Hordijk, W., Guisan, A., & Alvarez, N. (2012). Predicting present and future intra-specific genetic structure through niche hindcasting across 24 millennia. *Ecology Letters, 15*(7), 649-657. doi:10.1111/j.1461-0248.2012.01779.x

825 Estrada, A., Morales-Castilla, I., Caplat, P., & Early, R. (2016). Usefulness of species
826 traits in predicting range shifts. *Trends in Ecology & Evolution, 31*(3), 190-203.
827 doi:10.1016/j.tree.2015.12.014
828 Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of
829 individuals using the software STRUCTURE: a simulation study. *Molecular Ecology,*
830 *14*(8), 2611-2620. doi:10.1111/j.1365-294X.2005.02553.x
831 Excoffier, L., & Lischer, H. E. L. (2010). ARLEQUIN suite ver 3.5: a new series of programs
832 to perform population genetics analyses under Linux and Windows. *Molecular*
833 *Ecology Resources, 10*(3), 564-567. doi:10.1111/j.1755-0998.2010.02847.x
834 Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., . . .
835 Zhang, M. H. (2011). The Community Climate System Model Version 4. *Journal*
836 *of Climate, 24*(19), 4973-4991. doi:10.1175/2011jcli4083.1
837 Gilman, S. E., Urban, M. C., Tewksbury, J., Gilchrist, G. W., & Holt, R. D. (2010). A
838 framework for community interactions under climate change. *Trends in Ecology*
839 *& Evolution, 25*(6), 325-331. doi:10.1016/j.tree.2010.03.002
840 Godoy, O., Kraft, N. J. B., & Levine, J. M. (2014). Phylogenetic relatedness and the
841 determinants of competitive outcomes. *Ecology Letters, 17*(7), 836-844.
842 doi:10.1111/ele.12289
843 Godsoe, W., Jankowski, J., Holt, R. D., & Gravel, D. (2017). Integrating biogeography
844 with contemporary niche theory. *Trends in Ecology & Evolution, 32*(7), 488-499.
845 doi:10.1016/j.tree.2017.03.008
846 González-Serna, M. J., Cordero, P. J., & Ortego, J. (2019). Spatiotemporally explicit
847 demographic modelling supports a joint effect of historical barriers to dispersal
848 and contemporary landscape composition on structuring genomic variation in a
849 red-listed grasshopper. *Molecular Ecology, 28*(9), 2155-2172.
850 doi:10.1111/mec.15086
851 Hampe, A. (2004). Bioclimate envelope models: what they detect and what they hide.
852 *Global Ecology and Biogeography, 13*(5), 469-471. doi:10.1111/j.1466-
853 822X.2004.00090.x
854 Harrison, S. P., Bartlein, P. J., Brewer, S., Prentice, I. C., Boyd, M., Hessler, I., . . . Willis,
855 K. (2014). Climate model benchmarking with glacial and mid-Holocene climates.
856 *Climate Dynamics, 43*(3-4), 671-688. doi:10.1007/s00382-013-1922-6

He, Q. X., Edwards, D. L., & Knowles, L. L. (2013). Integrative testing of how environments from the past to the present shape genetic structure across landscapes. *Evolution, 67*(12), 3386-3402. doi:10.1111/evo.12159

Hijmans, R. J., Phillips, S., & Elith, J. L. (2017). *dismo*: Species Distribution Modeling. R package version 1.1-4. https://CRAN.R-project.org/package=dismo

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*(15), 1965-1978. doi:10.1002/joc.1276

Hipp, A. L., Manos, P. S., Gonzalez-Rodriguez, A., Hahn, M., Kaproth, M., McVay, J. D., . . . Cavender-Bares, J. (2018). Sympatric parallel diversification of major oak clades in the Americas and the origins of Mexican species diversity. *New Phytologist, 217*(1), 439-452. doi:10.1111/nph.14773

Jablonski, D. (2008). Biotic interactions and macroevolution: Extensions and mismatches across scales and levels. *Evolution, 62*(4), 715-739. doi:10.1111/j.1558-5646.2008.00317.x

James, P. M. A., Coltman, D. W., Murray, B. W., Hamelin, R. C., & Sperling, F. A. H. (2011). Spatial genetic structure of a symbiotic beetle-fungal system: Toward multi-taxa integrated landscape genetics. *PLoS One*, *6*(10), e25359. doi:10.1371/journal.pone.0025359

Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The *K*=2 conundrum. *Molecular Ecology, 26*(14), 3594-3602. doi:10.1111/mec.14187

Jeffreys, H. (1961). *Theory of Probability*. Oxford, UK: Oxford University Press.

Jensen, R. J. (1997). *Quercus* Linnaeus sect. *Lobatae* Loudon. In: *Flora of North America, North of Mexico. Vol. 3* (ed. Flora of North America Editorial Committee), pp. 447-468. New York, NY, USA: Oxford University Press.

Jombart, T. (2008). *Adegenet*: A R package for the multivariate analysis of genetic markers. *Bioinformatics, 24*(11), 1403–1405. doi.org/10.1093/bioinformatics/btn129

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773-795. doi:10.1080/01621459.1995.10476572

888 Keitt, T. H., Bjornstad, O. N., Dixon, P. M., & Citron-Pousty, S. (2002). Accounting for
889 spatial pattern when modeling organism-environment interactions. *Ecography,*
890 *25*(5), 616-625. doi:10.1034/j.1600-0587.2002.250509.x
891 Kim, B. Y., Wei, X. Z., Fitz-Gibbon, S., Lohmueller, K. E., Ortego, J., Gugger, P. F., & Sork,
892 V. L. (2018). RADseq data reveal ancient, but not pervasive, introgression
893 between Californian tree and scrub oak species (*Quercus* sect. *Quercus*:
894 Fagaceae). *Molecular Ecology, 27*(22), 4556-4571. doi:10.1111/mec.14869
895 Knowles, L. L. (2009). Statistical phylogeography. *Annual Review of Ecology, Evolution*
896 *and Systematics*, 40, 593-612.
897 Knowles, L. L., & Massatti, R. (2017). Distributional shifts - not geographic isolation - as
898 a probable driver of montane species divergence. *Ecography, 40*(12), 1475-
899 1485. doi:10.1111/ecog.02893
900 Koenig, W. D. (1999). Spatial autocorrelation of ecological phenomena. *Trends in*
901 *Ecology & Evolution, 14*(1), 22-26. doi:10.1016/s0169-5347(98)01533-x
902 Leathwick, J. R., & Austin, M. P. (2001). Competitive interactions between tree species
903 in New Zealand's old-growth indigenous forests. *Ecology, 82*(9), 2560-2573.
904 doi:10.2307/2679936
905 Leuenberger, C., & Wegmann, D. (2010). Bayesian computation and model selection
906 without likelihoods. *Genetics, 184*(1), 243-252.
907 doi:10.1534/genetics.109.109058
908 Levin, D. A. (2006). The spatial sorting of ecological species: Ghost of competition or of
909 hybridization past? *Systematic Botany, 31*(1), 8-12.
910 doi:10.1600/036364406775971831
911 Levin, D. A., Francisco-Ortega, J., & Jansen, R. K. (1996). Hybridization and the
912 extinction of rare plant species. *Conservation Biology, 10*(1), 10-16.
913 doi:10.1046/j.1523-1739.1996.10010010.x
914 Liu, C. R., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of
915 occurrence in the prediction of species distributions. *Ecography, 28*(3), 385-
916 393. doi:10.1111/j.0906-7590.2005.03957.x
917 Louthan, A. M., Doak, D. F., & Angert, A. L. (2015). Where and when do species
918 interactions set range limits? *Trends in Ecology & Evolution, 30*(12), 780-792.
919 doi:10.1016/j.tree.2015.09.011

Manos, P. S. (1997). *Quercus* Linnaeus sect. *Protobalanus* (Trelease) A. Camus. In: *Flora of North America, Vol. 3* (ed. Flora of North America Editorial Committee), pp. 468-471. New York, NY, USA: Oxford University Press.

Manos, P. S., Doyle, J. J., & Nixon, K. C. (1999). Phylogeny, biogeography, and processes of molecular differentiation in Quercus subgenus Quercus (Fagaceae). *Molecular Phylogenetics and Evolution*, *12*(3), 333-349.

Massatti, R., & Knowles, L. L. (2016). Contrasting support for alternative models of genomic variation based on microhabitat preference: species-specific effects of climate change in alpine sedges. *Molecular Ecology, 25*(16), 3974-3986. doi:10.1111/mec.13735

Maynard, D. S., Wootton, J. T., Servan, C. A., & Allesina, S. (2019). Reconciling empirical interactions and species coexistence. *Ecology Letters, 22*(6), 1028-1037. doi:10.1111/ele.13256

Mevik, B. H., & Wehrens, R. (2007). The *pls* package: Principal component and partial least squares regression in R. *Journal of Statistical Software, 18*(2), 1-23.

Miriti, M. N., Wright, S. J., & Howe, H. F. (2001). The effects of neighbors on the demography of a dominant desert shrub (*Ambrosia dumosa*). *Ecological Monographs, 71*(4), 491-509. doi:10.1890/0012-9615(2001)071[0491:teonot]2.0.co;2

Mohr, J. A., Whitlock, C., & Skinner, C. N. (2000). Postglacial vegetation and fire history, eastern Klamath Mountains, California, USA. *Holocene, 10*(5), 587-601. doi:10.1191/095968300675837671

Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (2014). *ENMeval*: An R package for conducting spatially independent evaluations and estimating optimal model complexity for MAXENT ecological niche models. *Methods in Ecology and Evolution, 5*(11), 1198-1205. doi:10.1111/2041-210x.12261

Narwani, A., Bentlage, B., Alexandrou, M. A., Fritschie, K. J., Delwiche, C., Oakley, T. H., & Cardinale, B. J. (2017). Ecological interactions and coexistence are predicted by gene expression similarity in freshwater green algae. *Journal of Ecology, 105*(3), 580-591. doi:10.1111/1365-2745.12759

951 Neuenschwander, S., Largiader, C. R., Ray, N., Currat, M., Vonlanthen, P., & Excoffier, L.
952 (2008). Colonization history of the Swiss Rhine basin by the bullhead (*Cottus*
953 *gobio*): inference under a Bayesian spatially explicit framework. *Molecular*
954 *Ecology, 17*(3), 757-772.
955 Nixon, K. C. & Muller, C. H. (1997). *Quercus* Linnaeus sect. *Quercus*. In: *Flora of North*
956 *America, Vol. 3* (ed. Flora of North America Editorial Committee), pp. 471-506.
957 New York, NY, USA: Oxford University Press.
958 Nixon, K. C. (2002). *The Oak (Quercus) Biodiversity of California and Adjacent Regions*.
959 General Technical Reports PSW-GTR-184. USDA Forest Service.
960 Ortego, J., Noguerales, V., Gugger, P. F., & Sork, V. L. (2015). Evolutionary and
961 demographic history of the Californian scrub white oak species complex: an
962 integrative approach. *Molecular Ecology, 24*(24), 6188-6208.
963 doi:10.1111/mec.13457
964 Ortego, J., Gugger, P. F., & Sork, V. L. (2017). Impacts of human-induced environmental
965 disturbances on hybridization between two ecologically differentiated
966 Californian oak species. *New Phytologist, 213*(2), 930-943.
967 doi:10.1111/nph.14182
968 Ortego, J., Gugger, P. F., & Sork, V. L. (2018). Genomic data reveal cryptic lineage
969 diversification and introgression in Californian golden cup oaks (section
970 *Protobalanus*). *New Phytologist, 218*(2), 804-818. doi:10.1111/nph.14951
971 Papadopoulou, A., & Knowles, L. L. (2016). Toward a paradigm shift in comparative
972 phylogeography driven by trait-based hypotheses. *Proceedings of the National*
973 *Academy of Sciences of the United States of America, 113*(29), 8018-8024.
974 doi:10.1073/pnas.1601069113
975 Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the
976 distribution of species: are bioclimate envelope models useful? *Global Ecology*
977 *and Biogeography, 12*(5), 361-371. doi:10.1046/j.1466-822X.2003.00042.x
978 Pesendorfer, M. B., Sillett, T. S., Morrison, S. A., & Kamil, A. C. (2016). Context-
979 dependent seed dispersal by a scatter-hoarding corvid. *Journal of Animal*
980 *Ecology, 85*(3), 798-805. doi:10.1111/1365-2656.12501
981 Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double
982 digest RADseq: An inexpensive method for *de novo* SNP discovery and

genotyping in model and non-model species. *PLoS One, 7*(5), e37135. doi:10.1371/journal.pone.0037135

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

Petritan, I. C., Marzano, R., Petritan, A. M., & Lingua, E. (2014). Overstory succession in a mixed *Quercus petraea-Fagus sylvatica* old growth forest revealed through the spatial pattern of competition and mortality. *Forest Ecology and Management, 326*, 9-17. doi:10.1016/j.foreco.2014.04.017

Pham, K. K., Hipp, A. L., Manos, P. S., & Cronn, R. C. (2017). A time and a place for everything: phylogenetic history and geography as joint predictors of oak plastome phylogeny. *Genome*, *60*(9), 720-732. doi:10.1139/gen-2016-0191

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling, 190*(3-4), 231-259. doi:10.1016/j.ecolmodel.2005.03.026

Pollock, L. J., Bayly, M. J., & Vesk, P. A. (2015). The roles of ecological and evolutionary processes in plant community assembly: The environment, hybridization, and introgression influence co-occurrence of *Eucalyptus*. *American Naturalist, 185*(6), 784-796. doi:10.1086/680983

Pollock, L. J., Tingley, R., Morris, W. K., Golding, N., O'Hara, R. B., Parris, K. M., . . . McCarthy, M. A. (2014). Understanding co-occurrence bymodelling species simultaneously with a Joint Species DistributionModel (JSDM). *Methods in Ecology and Evolution, 5*(5), 397-406. doi:10.1111/2041-210x.12180

Potts, B. M., & Reid, J. B. (1988). Hybridization as a dispersal mechanism. *Evolution, 42*(6), 1245-1255. doi:10.2307/2409008

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

R Core Team (2020). R: *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ray, N., Currat, M., Foll, M., & Excoffier, L. (2010). SPLATCHE2: a spatially explicit simulation framework for complex demography, genetic admixture and recombination. *Bioinformatics, 26*(23), 2993-2994. doi:10.1093/bioinformatics/btq579

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

Rizzo, D. M., Garbelotto, M., Davidson, J. M., Slaughter, G. W., & Koike, S. T. (2002). Phytophthora ramorum as the cause of extensive mortality of *Quercus* spp. and *Lithocarpus densiflorus* in California. *Plant Disease, 86*(3), 205-214. doi:10.1094/pdis.2002.86.3.205

Shi, N. N., Gao, C., Zheng, Y., & Guo, L. D. (2017). Effects of ectomycorrhizal fungal identity and diversity on subtropical tree competition. *Journal of Plant Ecology, 10*(1), 47-55. doi:10.1093/jpe/rtw060

Soberon, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters, 10*(12), 1115-1123. doi:10.1111/j.1461-0248.2007.01107.x

Starrett, J., Hayashi, C. Y., Derkarabetian, S., & Hedin, M. (2018). Cryptic elevational zonation in trapdoor spiders (Araneae, Antrodiaetidae, *Aliatypus janus* complex) from the California southern Sierra Nevada. *Molecular Phylogenetics and Evolution, 118*, 403-413. doi:10.1016/j.ympev.2017.09.003

Svenning, J. C., Gravel, D., Holt, R. D., Schurr, F. M., Thuiller, W., Munkemuller, T., . . . Normand, S. (2014). The influence of interspecific interactions on species range expansion rates. *Ecography, 37*(12), 1198-1209. doi:10.1111/j.1600-0587.2013.00574.x

Ting, J. J., & Cutter, A. D. (2018). Demographic consequences of reproductive interference in multi-species communities. *BMC Ecology, 18*, 46. doi:10.1186/s12898-018-0201-0

Tsai, Y. H. E., & Manos, P. S. (2010). Host density drives the postglacial migration of the tree parasite, *Epifagus virginiana*. *Proceedings of the National Academy of Sciences of the United States of America, 107*(39), 17035-17040. doi:10.1073/pnas.1006225107

Tucker, J. M. (1980). Taxonomy of California oaks. In: Plumb, T.R., technical coordinator. *Proceedings of the Symposium on the Ecology, Management and Utilization of California Oaks, June 26–28, 1979, Claremont, California*. Berkeley, CA, USA: US Department of Agriculture Forest Service, Pacific

Southwest Forest and Range Experiment Station, General Technical Report PSW-144, 19-29

Valiente-Banuet, A., & Verdu, M. (2007). Facilitation can increase the phylogenetic diversity of plant communities. *Ecology Letters, 10*(11), 1029-1036. doi:10.1111/j.1461-0248.2007.01100.x

Warren, D. L., Glor, R. E., & Turelli, M. (2010). ENMTOOLS: a toolbox for comparative studies of environmental niche models. *Ecography, 33*(3), 607-611. doi:10.1111/j.1600-0587.2009.06142.x

Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in MAXENT: the importance of model complexity and the performance of model selection criteria. *Ecological Applications, 21*(2), 335-342. doi:10.1890/10-1171.1

Warren, D. L., Wright, A. N., Seifert, S. N., & Shaffer, H. B. (2014). Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. *Diversity and Distributions, 20*(3), 334-343.

Wegmann, D., Leuenberger, C., & Excoffier, L. (2009). Efficient approximate Bayesian computation coupled with Markov Chain Monte Carlo without likelihood. *Genetics, 182*(4), 1207-1218. doi:10.1534/genetics.109.102509

Wegmann, D., Leuenberger, C., Neuenschwander, S., & Excoffier, L. (2010). ABCTOOLBOX: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics, 11*, 116. doi:10.1186/1471-2105-11-116

Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., . . . Svenning, J. C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews, 88*(1), 15-30. doi:10.1111/j.1469-185X.2012.00235.x

Yguel, B., Bailey, R., Tosh, N. D., Vialatte, A., Vasseur, C., Vitrac, X., . . . Prinzing, A. (2011). Phytophagy on phylogenetically isolated trees: why hosts should escape their relatives. *Ecology Letters, 14*(11), 1117-1124. doi:10.1111/j.1461-0248.2011.01680.x

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.Legends to figures

FIGURE 1 Workflow illustrating the integrative distributional, demographic and coalescent framework (iDDC; He et al., 2013) employed in this study to test alternative phylogeographic models, but modified here to incorporate interspecific interactions. We use Californian oaks as a case study to illustrate the workflow. Here we illustrate by reference to the canyon live oak (*Quercus chrysolepis*) as the focal taxon and the hypothetical neutral (0), negative (-) or positive (+) effects exerted by the rest of oak species. We used ENMs to translate such interactions into nine phylogeographic models (described in Table 1), where the nature of the interaction may differ depending upon the phylogenetic relationships among oak taxa (Hipp et al., 2018; Ortego et al., 2018), as indicated by taxonomic sections. Note that the small black boxes in the schematic correspond to the specific subsections in the Materials and Methods detailing each step. LGM, last glacial maximum; PLS, partial least square; BF, Bayes Factor; K_{MAX}, carrying capacity of the deme with highest suitability; *m*, migration rate per deme per generation; N_{ANC}, ancestral population size.

FIGURE 2 Studied populations of (a) California scrub oak (*Quercus berberidifolia*) and (b) canyon live oak (*Q. chrysolepis*). Pie charts show the probability of membership of the studied populations to each of the most likely number of genetic clusters inferred by the Bayesian method implemented in the program STRUCTURE. Bar plots at the bottom show individual probabilities of membership to each genetic cluster, with thin vertical black lines separating different populations. Grey shading shows the current distribution of each species based on an environmental niche model (ENM). Dashed lines on map from panel (a) illustrate the location of the main mountain ranges of the region (text in italics). Population codes are described in Table S1.

FIGURE 3 Spatiotemporally explicit demographic scenarios most supported for (a) California scrub oak (*Quercus berberidifolia*) (Model B) and (b) canyon live oak (*Q. chrysolepis*) (Model I). Local carrying capacities (*K*, colored scale bar) change across the

landscape and three time periods (from the last glacial maximum to present), with each snapshot used for one-third (7.2 ka) of the total number (21.5 ka) of simulations. Local carrying capacities for the focal species range from 0 (minimum) to 1 (maximum) and were scaled based on habitat suitabilities estimated from environmental niche models (ENMs) and considering interspecific interactions (Model B: negative effect of all other oak species; Model I: positive effect of other species within the same section + negative effect of species from different sections). ka, thousands years ago

FIGURE 4 Posterior distribution (solid black line) and mode (vertical dotted black line) of parameter estimates (K_{MAX}, m, N_{ANC}) for the most supported model for (a) California scrub oak (*Quercus berberidifolia*) (Model B) and (b) canyon live oak (*Q. chrysolepis*) (Model I) based on a general linear model (GLM) regression adjustment of the 5,000 retained simulations (0.5%) closest to empirical data. The comparison of posterior distributions before (blue shading) and after (solid black line) the ABC-GLM shows the improvement that this procedure had on parameter estimates. The comparison of prior (red shading) and posterior (solid black line and blue shading) distributions demostrates that the data contained information relevant to estimating the parameters. Note that y-axes are scaled differently. K_{MAX}, carrying capacity of the deme with highest suitability; m, migration rate per deme per generation; N_{ANC}, ancestral population size.

TABLE 1 Statistics from the ABC procedure used for evaluating the relative support of each model in the two focal species. A higher marginal density corresponds to a higher model support and a high (i.e., non-significant) Wegmann`s *p*-value ($p > .05$) indicates that the model is able to generate data in agreement with empirical data. Bayes factors represent the degree of relative support for the model with the highest marginal density (in bold) over the other models. Bayes factors >20 indicate strong support, while those >150 indicate very strong support (Kass & Raftery, 1995). R^2 is the coefficient of determination from a regression between each demographic parameter (K_{MAX}, m, N_{ANC}) and the four partial least squares (PLS) extracted from all summary statistics.

Model - Interactions by other oak species		Marginal density	Wegmann´s *p*-value	Bayes factor	R^2		
					K_{MAX}	m	N_{ANC}
(a) *Quercus berberidifolia*							
A	Null	1.33×10^{-09}	<0.001	2.87×10^{06}	0.80	0.95	0.87
B	**Negative (by all species)**	$\mathbf{3.81 \times 10^{-03}}$	**0.705**	**—**	**0.81**	**0.94**	**0.92**
C	Negative (by species within the same section)	9.22×10^{-05}	0.029	41	0.84	0.95	0.90
D	Negative (by species from different sections)	1.43×10^{-04}	0.055	27	0.86	0.95	0.91
E	Positive (by all species)	1.39×10^{-18}	<0.001	2.75×10^{15}	0.65	0.93	0.81
F	Positive (by species within the same section)	3.59×10^{-15}	<0.001	1.06×10^{12}	0.72	0.94	0.85
G	Positive (by species from different sections)	1.09×10^{-15}	<0.001	3.51×10^{12}	0.68	0.94	0.84
H	Negative (same section) + Positive (different sections)	1.13×10^{-08}	0.001	3.38×10^{05}	0.81	0.95	0.87
I	Positive (same section) + Negative (different sections)	3.22×10^{-15}	<0.001	1.18×10^{12}	0.78	0.94	0.87
(b) *Quercus chrysolepis*							
A	Null	8.27×10^{-07}	0.007	5.56×10^{03}	0.52	0.78	0.85
B	Negative (by all species)	5.81×10^{-04}	0.839	7.92	0.81	0.88	0.89
C	Negative (by species within the same section)	3.76×10^{-16}	<0.001	1.22×10^{13}	0.56	0.77	0.84
D	Negative (by species from different sections)	3.06×10^{-03}	0.989	1.50	0.80	0.88	0.89
E	Positive (by all species)	4.04×10^{-08}	0.002	1.14×10^{05}	0.35	0.75	0.79

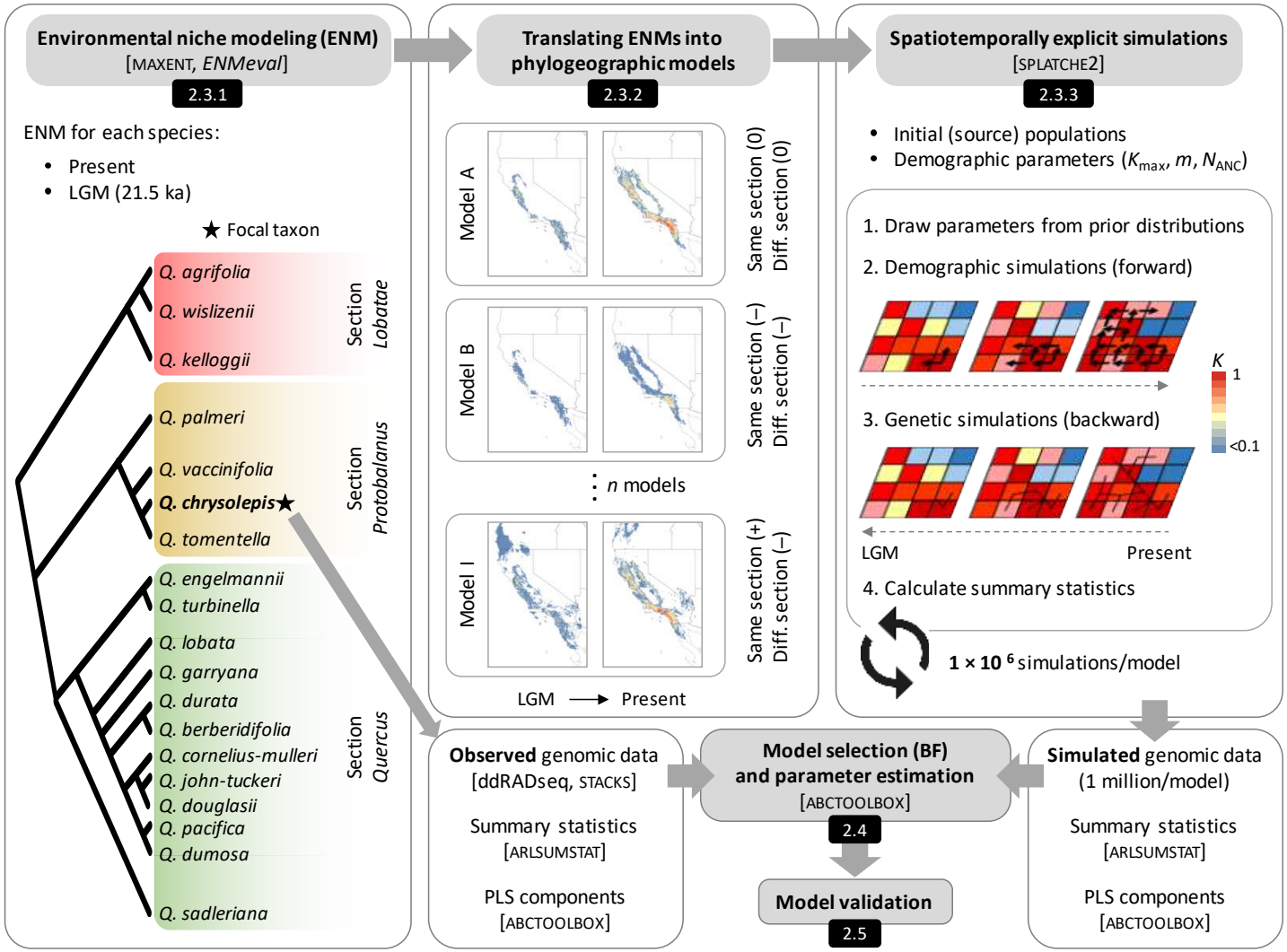
F	Positive (by species within the same section)	8.17×10^{-04}	0.112	5.63	0.48	0.79	0.84
G	Positive (by species from different sections)	3.80×10^{-08}	0.001	1.21×10^{05}	0.36	0.74	0.82
H	Negative (same section) + Positive (different sections)	1.66×10^{-07}	0.003	2.77×10^{04}	0.38	0.74	0.82
I	**Positive (same section) + Negative (different sections)**	$\mathbf{4.60 \times 10^{-03}}$	**0.998**	**—**	**0.74**	**0.87**	**0.88**

K_{MAX}, carrying capacity of the deme with highest suitability

m, migration rate per deme per generation

N_{ANC}, ancestral population size

FIGURE 1



1145

FIGURE 2

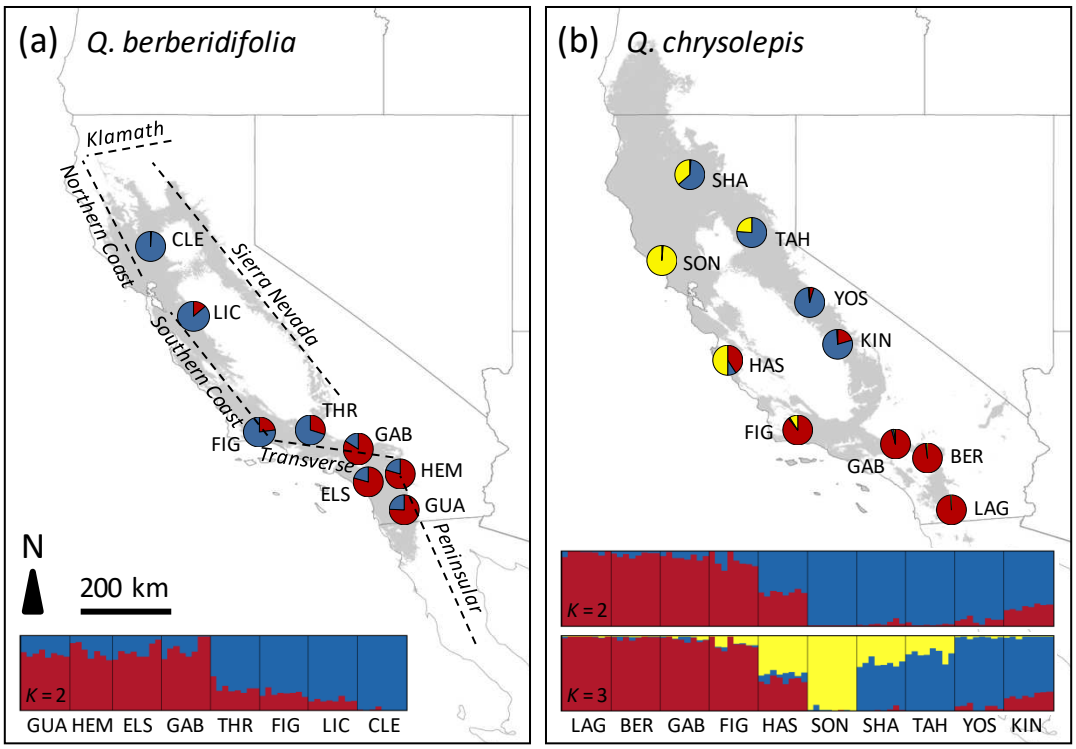


FIGURE 3

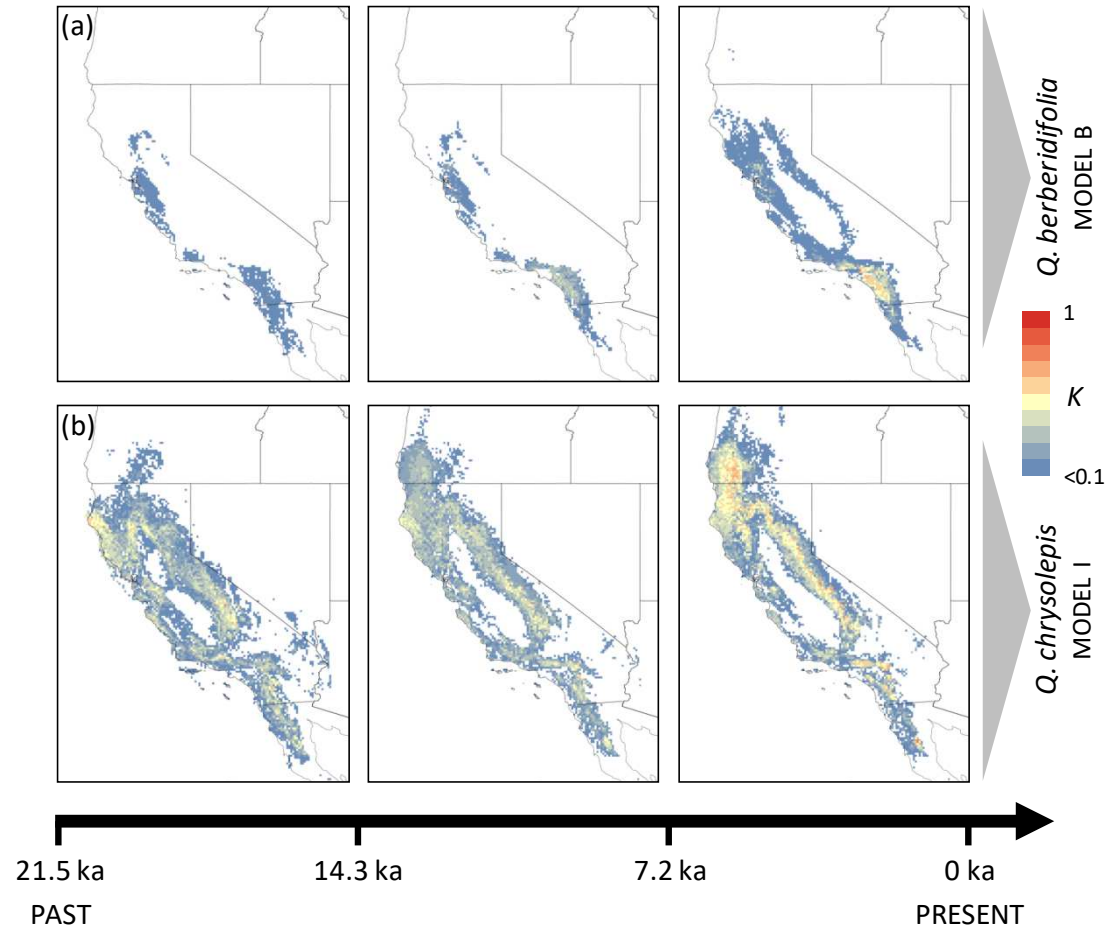


FIGURE 4