

Article

Genetic, Morphological, and Environmental Differentiation of an Arid-Adapted Oak with a Disjunct Distribution

Bethany A. Zumwalde^{1,2,*}, Ross A. McCauley³, Ian J. Fullinwider³, Drew Duckett^{1,4}, Emma Spence¹ and Sean Hoban^{1,*}

¹ The Morton Arboretum, Center for Tree Science, 4100 Illinois 53, Lisle, IL 60532, USA; duckett.17@osu.edu (D.D.); emma@largelandscapes.org (E.S.)

² Department of Biology, University of Florida, Gainesville, FL 32611, USA

³ Department of Biology, Fort Lewis College, 1000 Rim Drive, Durango, CO 81301, USA; McCauley_R@fortlewis.edu (R.A.M.); ijfullinwider@fortlewis.edu (I.J.F.)

⁴ Department of Evolution, Ecology and Organismal Biology, The Ohio State University, 1315 Kinnear Rd., Columbus, OH 43212, USA

* Correspondence: bzumwalde@ufl.edu (B.A.Z.); shoban@mortonarb.org (S.H.)

Abstract: The patterns of genetic and morphological diversity of a widespread species can be influenced by environmental heterogeneity and the degree of connectivity across its geographic distribution. Here, we studied *Quercus havardii* Rydb., a uniquely adapted desert oak endemic to the Southwest region of the United States, using genetic, morphometric, and environmental datasets over various geographic scales to quantify differentiation and understand forces influencing population divergence. First, we quantified variation by analyzing 10 eastern and 13 western populations from the disjunct distribution of *Q. havardii* using 11 microsatellite loci, 17 morphological variables, and 19 bioclimatic variables. We then used regressions to examine local and regional correlations of climate with genetic variation. We found strong genetic, morphological and environmental differences corresponding with the large-scale disjunction of populations. Additionally, western populations had higher genetic diversity and lower relatedness than eastern populations. Levels of genetic variation in the eastern populations were found to be primarily associated with precipitation seasonality, while levels of genetic variation in western populations were associated with lower daily temperature fluctuations and higher winter precipitation. Finally, we found little to no observed environmental niche overlap between regions. Our results suggest that eastern and western populations likely represent two distinct taxonomic entities, each associated with a unique set of climatic variables potentially influencing local patterns of diversity.

Keywords: *Quercus havardii*; Fagaceae; genetic differentiation; morphometrics; bioclimatic associations



Citation: Zumwalde, B.A.; McCauley, R.A.; Fullinwider, I.J.; Duckett, D.; Spence, E.; Hoban, S. Genetic, Morphological, and Environmental Differentiation of an Arid-Adapted Oak with a Disjunct Distribution. *Forests* **2021**, *12*, 465. <https://doi.org/10.3390/f12040465>

Academic Editors: Mary Ashley and Janet R. Backs

Received: 20 February 2021

Accepted: 9 April 2021

Published: 10 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Genetic and phenotypic trait diversity of populations and species may be strongly influenced by environmental conditions at regional and local scales [1,2]. For example, favorable environmental conditions may promote higher genetic diversity by influencing the number of populations and migrants they produce, and thus increase their genetic connectivity [3]. Alternatively, unfavorable conditions could reduce genetic diversity and increase genetic drift by influencing population bottlenecks, inbreeding among close relatives, or local extinction–colonization dynamics [4]. Furthermore, environmental conditions can impose selective pressures and influence the rate at which populations diverge and speciation may occur [5–7].

An additional factor influencing genetic and trait divergence is large-scale spatial disjunctions, or discontinuities, in the range across which little or no migration occurs. Species with a once continuous range, but now consisting of disjunct populations, have long been of biogeographical interest [8–10], especially when those disjunctions have been

caused by climatic changes from events such as Pleistocene glacial cycles [11,12]. Disjunct distributions resulting in spatial isolation can influence the divergence of populations or ultimately result in parapatric speciation. This isolation may enable genetic changes that determine distinctive morphological and physiological characters [13].

Studies in *Quercus* L. (Fagaceae), one of the most economically and ecologically important woody genera of northern temperate zones, have been particularly informative for examining environmental [14,15], genetic [16,17], and trait-based [18,19] relationships in regard to geographic and biogeographic history [20,21]. Species in this genus have adapted to a remarkable range of habitats [22,23]; however, relatively few oaks in the United States are adapted to the arid environments of the Southwest [24,25], and in particular, to habitats with deep, shifting sand. A key research goal in understanding adaptation in desert oaks is to identify how this extreme environment has influenced genetic and morphological diversity, to better understand the range of conditions that oaks can survive in, and to apply this information to conserving oaks in situ.

To investigate the degree to which the environment contributes to genetic and phenotypic diversity, we examine Havard's oak, *Quercus havardii* Rydb., a keystone species that defines the sand shinnery ecosystem [26,27]. Commonly known as shinnery oak, *Q. havardii* is often found in harsh, arid environments as thickets or large shrubs in deep sands, which are not typically shared by other species of oak. This environment is quite extreme, and this growth habit is unusual among oaks. It has been reported as being highly clonal and rarely producing sexually via acorns [26,28,29]. As part of a recent evaluation of US oaks following IUCN Red List criteria [30], *Q. havardii* was listed as "Endangered" due to the increasing loss of habitat from agricultural practice and oil drilling, decreasing population size projections, and the total number of mature individuals.

Additionally, *Q. havardii* has a disjunct range: an eastern region ranging from western Oklahoma through to the Llano Estacado of west Texas and eastern New Mexico, and a western region encompassing the Navajo Basin in northeastern Arizona and southeastern Utah [31]. Few populations occur in the broad disjunction in central New Mexico, although the presence of this species in this area has been theorized to represent a continuous distribution in the past [32]. There has been no work to date examining the morphological and genetic diversity and structure of *Q. havardii* across the entirety of its range, nor any work on the range-wide population structure of any of the southwestern United States oaks (e.g., *Q. gambelii* Nutt., *Q. turbinella* Greene, *Q. ajoensis* C.H. Muller). Thus, we lack general knowledge of how oak populations are influenced by the harsh environmental conditions of desert habitats.

In this study, we aim to quantify genetic, morphological and environmental differentiation in populations of *Q. havardii*. We then summarize differentiation at both local and regional scales, and assess the timing at which population divergence potentially occurred. Finally, we evaluate how key environmental factors (i.e., precipitation and temperature) possibly contribute to this genetic variation within and across populations.

2. Materials and Methods

2.1. Sample Collection

Leaf tissue was collected in the summer of 2016 from a total of 489 samples from 23 populations for both morphological and genetic studies (Figure 1). From each individual, two leaves were collected and pressed for morphometric analysis and two to three leaves were placed on ice for DNA extraction. Heteroblastic leaf shape has been noted to occur in Fagaceae [33,34], in which leaf morphology may differ from various shoot types. Using a collection method implemented in similar morphometric studies of *Quercus* [32,35,36], we collected fully expanded and undamaged leaves from sun-exposed terminal shoots. A total of 23 populations were included in genetic analysis, while only 21 populations were included in morphological analysis. The two populations that were not included in the morphological analysis (E13 and E14) were sampled by collaborators and only included enough tissue for use in genetic analysis. Voucher specimens were produced for each

population and deposited at MOR and FLD with vouchers collected from the Navajo Nation additionally deposited in NAVA (herbarium acronyms follow Holmgren et al. [37]). Populations were chosen by considering timing and logistical issues of sampling while also attempting to represent the geographic range of the taxon. A list of possible sites was generated by contacting land managers of private and public land in the region, consulting GBIF (www.gbif.org; accessed 14 March 2015) and SEINet (<https://swbiodiversity.org/seinet>; accessed on 14 March 2015), and using suggestions from members of the International Oak Society. Additionally, populations were selected that were subject to little or no observed hybridization, although two sites located within the major disjunction (W11 and W12) documented occurrences of *Q. gambelii* and *Q. turbinella* in the vicinity.

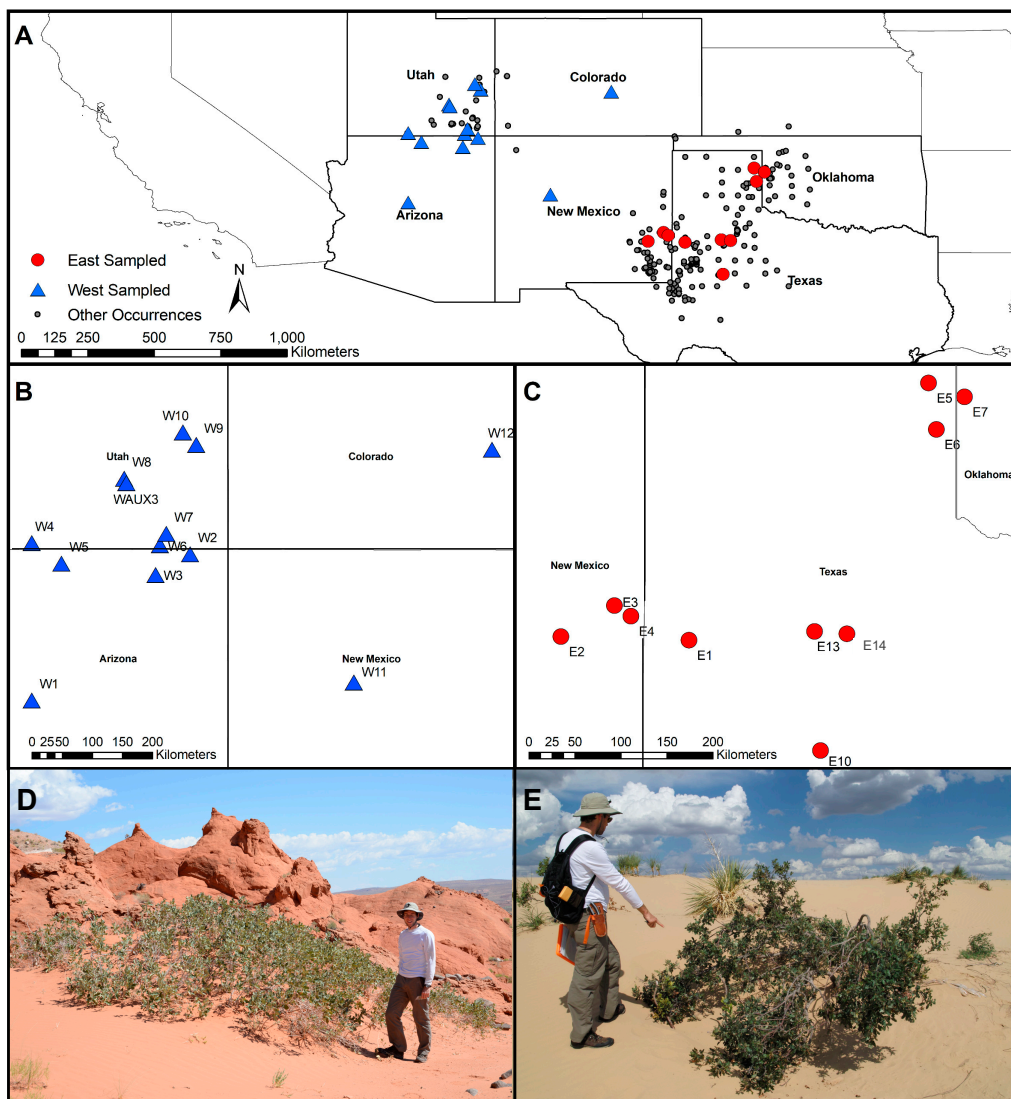


Figure 1. Sampling and distribution maps of *Quercus havardii*. (A) Known occurrence points for *Q. havardii* in grey circles and sampled populations used for genetic, morphological, and environmental analysis in blue triangles (western range) and red circles (eastern range); (B) sampling sites of 13 populations from the western range; (C) sampling sites of 10 populations from the eastern range; (D) example of *Q. havardii* habit from western region population W8; (E) example of *Q. havardii* in eastern region population E2.

2.2. DNA Extraction and Microsatellite Protocol

Genetic diversity was evaluated based on 11 nuclear microsatellite loci developed for other *Quercus* species (Appendix A Table A1). Genomic DNA was extracted from frozen leaf tissue using the E.Z.N.A. Plant DNA DS Kit (Omega Bio-tek, Norcross, GA, USA) and diluted to an approximate concentration of 5 ng/ μ L for PCR amplification. Microsatellite primers were organized using the software Multiplex Manager v.1.2 [38]. Multiplexed PCRs were carried out in a total volume of 10 μ L, containing 2 μ L of DNA template (5 ng/ μ L) and 8 μ L of master mix consisting of 1 \times reaction buffer, 0.5 mM total dNTPs, 1.5 mM MgCl₂, fluorescently tagged primer concentrations ranging from 0.02 to 0.28 μ M, 0.03 μ M of each reverse primer, 0.5 μ g/ μ L BSA, 0.025 U of GoTaq G2 DNA polymerase (Promega, Madison, WI, USA) and HPLC-grade water. Initial PCR conditions and reagents, including the use of BSA, were based upon similar microsatellite studies of congeneric species [39–42]. Forward primers were labeled with one of the 6-FAM, NED, VIC or PET fluorescent dyes (Applied Biosystems, Waltham, MA, USA). Cycling conditions were: an initial denaturation step at 94 °C for 5 min, then 30 cycles consisting of 30 s at 94 °C, 30 s at 50–56 °C (multiplex 1 at 50 °C, multiplex 2 at 54 °C, and multiplex 3 at 56 °C) and 30 s at 72 °C, followed by a final extension of 5 min at 72 °C. PCRs were ran on Eppendorf Mastercycler pro (Eppendorf, Hamburg, Germany) and C-1000 Touch (Bio-Rad, Hercules, CA, USA) machines. Amplification products were visualized using a 1.5% agarose, 1 \times TAE gel with ethidium bromide stain and sized with a 1 kb + ladder (ThermoFisher Scientific, Waltham, MA, USA). PCR products were run on an ABI 3730 Genetic Capillary Electrophoresis Sequencer (Applied Biosystems) and fragment sizes of alleles were scored using the software Geneious v.10.2.3 (Biomatters, Auckland, New Zealand), using the microsatellite plugin v.1.4.4. Alleles called and binned by hand were compared with suggested bins from Geneious. Micro-Checker v.2.2.3 [43] and INESTv.2.2 [44] were both used to check for null alleles across loci.

2.3. Statistical Analysis of Genetic Data

Initial analyses were performed using a minimum of 10 individuals per population, which means that some initially sampled locations were not analyzed due to having too few individuals (E11, E13, and W2). We used R v.3.6.3 [45] for all R-based analyses. After the removal of clones from the dataset using the R package Poppr v.2.9.0 [46], we decided to decrease the minimum number of individuals per population to seven in order to maximize the number of populations included over the range of the taxon. The presence of null alleles was detected in several loci. Thus, analyses were performed both including and discarding these loci. The final dataset reported in this study used genetic data from eleven microsatellite loci for a total of 489 individuals from 23 populations, requiring a minimum of seven individuals per population (Table 1). Alternative analyses discarding loci with null alleles and adjusting minimum population requirements were performed and did not alter results; these can be run using the code on GitHub (doi:10.5281/zenodo.4453098). Genetic diversity statistics including allelic richness, heterozygosity, and F_{ST} were calculated using the R packages adegenet v.2.1.3 [47], hierfstat v.0.5-7 [48], and demerlate v.0.9-3 [49]. Tests for normality were performed using the Shapiro–Wilk test for normality and Q-Q plots, which showed no significant departures from normality for the three diversity and relatedness measures, but did show a significant non-normal distribution for F_{ST} . This is true for all subsets of the data tested. To compare populations from each region (East vs. West) for these basic summary statistics, we performed *t*-tests, as well as Wilcoxon tests, for these four statistics, for all subsets of the data. To determine patterns of isolation by distance and the distribution of genetic variation between and among populations, a Mantel Test and an Analysis of Molecular Variance (AMOVA; F_{ST} , 999 permutations) were performed using GenAlEx v.6.503 [50]. Genetic data were also visualized using principal coordinate analysis (PCoA) of pairwise distances between populations using GenAlEx. Genetic structure was visualized using STRUCTURE v.2.3.4 software [26,27]. STRUCTURE plots were examined for both range-wide and within regions [51]. *K* values were set from

1 to 40 with 20 iterations at 100,000 burn-ins and Markov Chain Monte Carlo (MCMC) repeats. Mean likelihood and ΔK values were obtained using Structure Harvester [52]. CLUMPP version 1.1.2 was used to resolve multimodality and DISTRUCT version 1.1 [53] was used to visualize results from STRUCTURE outputs. STRUCTURE clustering results were cross-checked with the model-free method discriminant analysis of principal components (DAPC) using the R package adegenet [47,54]. The number of K groups that best fit the data was chosen using the Evanno method [55].

Table 1. Population information and genetic diversity indices of the 23 populations of *Quercus havardii* using 11 microsatellite loci.

Population	Latitude (N)	Longitude (W)	<i>n</i>	MLG	Clones	<i>He</i>	<i>Ho</i>	<i>Na</i>	<i>Ar</i>	<i>F_{ST}</i>	<i>Rel</i>
E1	33.373	−102.617	20	18	2	0.554	0.524	58	3.61	0.088	0.215
E2	33.408	−103.867	25	22	3	0.562	0.496	56	3.42	0.090	0.250
E3	33.712	−103.344	27	20	7	0.520	0.418	57	3.44	0.094	0.239
E4	33.609	−103.183	19	13	6	0.542	0.433	49	3.58	0.082	0.149
E5	35.903	−100.276	32	17	15	0.543	0.516	60	3.55	0.095	0.265
E6	35.445	−100.199	14	11	3	0.654	0.570	60	4.26	0.078	0.113
E7	35.767	−99.926	56	40	16	0.615	0.554	91	4.16	0.072	0.156
E10	32.286	−101.332	9	9	0	0.649	0.581	55	4.38	0.066	0.054
E13	33.459	−101.391	8	8	0	0.634	0.648	53	4.29	0.073	0.158
E15	33.435	−101.074	24	21	3	0.704	0.591	84	4.82	0.061	0.009
W1	34.727	−111.985	39	25	14	0.553	0.478	58	3.37	0.129	0.275
W2	36.927	−109.618	11	7	4	0.681	0.494	53	4.57	0.077	−0.044
W3	36.616	−110.133	33	32	1	0.586	0.491	88	4.53	0.083	0.098
W4	37.096	−111.984	33	26	7	0.646	0.614	79	4.42	0.076	0.143
W5	36.786	−111.540	30	27	3	0.668	0.616	103	5.05	0.061	0.052
W6	37.064	−110.069	33	31	2	0.663	0.560	106	4.90	0.061	0.051
W7	37.234	−109.973	32	30	2	0.660	0.567	93	4.77	0.064	0.078
W8	38.065	−110.602	39	35	4	0.716	0.616	98	4.94	0.060	0.035
W9	38.573	−109.527	30	28	2	0.701	0.552	106	5.23	0.058	−0.001
W10	38.762	−109.725	29	22	7	0.696	0.603	90	5.14	0.066	0.021
W11	34.998	−107.172	20	19	1	0.647	0.584	84	4.80	0.086	0.089
W12	38.502	−105.106	18	15	3	0.601	0.453	69	4.38	0.093	0.081
WAUX3	38.000	−110.569	16	14	2	0.694	0.641	73	4.86	0.070	0.051
East Mean			23.4	17.9	5.5	0.598	0.533	62.3	3.95	0.080	0.161
West Mean			27.9	23.9	4.0	0.655	0.559	84.6	4.69	0.076	0.072
Overall Mean			26.0	21.3	4.7	0.630	0.548	74.1	4.79	0.077	0.110

n: sample size; MLG: No. multilocus genotype; Clones: No. sampled clones in the population; *He*: expected heterozygosity; *Na*: No. alleles; *Ar*: allelic richness by locus; *F_{ST}*: pairwise genetic differentiation; *Rel*: relatedness within a population.

2.4. Morphometric Methods

We used morphometric analysis, utilizing the characters defined in McCauley et al. [32], to assess patterns of morphological differentiation at local and regional scales (Table 2). Twelve continuous characters based on defined landmarks from 2 leaves per individual were measured manually on a total of 928 leaves. These characters included the number of lobes (LOBES), total leaf length (LENGTH), basal lobe blade width (BLBW), middle lobe blade width (MLBW), apical lobe blade width (ALBW), lower vein length (LVL), upper vein length (UVL), upper middle lobe sinus width (UMLS), lower middle lobe sinus width (LMLS), angle of lower lobe (LLA), angle of upper lobe (ULA), and angle of middle lobe from apex (MLA). Additionally, five composite variables derived from the principal measurements based on ratios of leaf length to width (LWR) and lobe depth (LODR) were included: LWR_1 (BLBW/Length), LWR_2 (MLBW/Length), LWR_3 (ALBW/Length), LODR_1 (MLBW-UMLS/MLBW), LODR_2 (MLBW-LMLS/MLBW). While specifically adapted for southwestern oaks, these measurements represent typical landmarks used in morphometric analysis in *Quercus* [56,57]. Normality of data was tested both across and within populations for individual characters using a Shapiro–Wilk test. As the data for many characters were not normally distributed, we used non-parametric Spearman’s correlation to identify potential correlation ($r > 0.70$) [58,59] among variables. As no correlation was observed, all variables were used in the subsequent analyses. Principal component analysis (PCA) was used to visualize differentiation of all individual samples and means from each population across the two disjunct regions with the first and second principal

components plotted using the R package ggplot2 v.3.3.3 [60]. To explore the partitioning of morphological variance in each of the measured variables across multiple levels, we performed a linear mixed effects model using the R package nlme v.3.1-152 [61]. Region was considered a fixed effect in the model, while population within each region and trees within each population were considered random effects. Residual plots were used to check for normality. The proportion of the total variance contributed by the fixed and combined random effects was determined using the marginal and conditional R^2 procedure of Nakagawa and Schielzeth [62] and calculated using the R package MuMIn v. 1.43.4 [63].

Table 2. Proportion of morphological variance in 12 directly measured and five composite variables in *Q. havardii* partitioned by East and West (regional) and by combined population and individual (local).

Character	Abbreviation	Regional	Local
Lobe Number	LOBES	0.182	0.345
Leaf Length	LENGTH	0.309	0.332
Basal Lobe Blade Width	BLBW	0.064	0.449
Middle Lobe Blade Width	MLBW	0.010	0.491
Apical Lobe Blade Width	ALBW	0.023	0.025
Lower Vein Length	LVL	0.105	0.145
Upper Vein Length	UVL	0.002	0.342
Upper Middle Lobe Sinus Width	UMLS	0.001	0.508
Lower Middle Lobe Sinus Width	LMLS	0.027	0.470
Lower Lobe Angle	LLA	0.383	0.202
Upper Lobe Angle	ULA	0.038	0.161
Middle Lobe Angle	MLA	0.038	0.205
Length to Width Ratio 1 (BLBW/Length)	LWR_1	0.058	0.442
Length to Width Ratio 2 (MLBW/Length)	LWR_2	0.403	0.264
Length to Width Ratio 3 (ALBW/Length)	LWR_3	0.040	0.223
Lobe Depth Ratio 1	LODR_1	0.026	0.402
Lobe Depth Ratio 2	LODR_2	0.076	0.411

2.5. Environmental Differentiation and Influence on Genetic Variation

We used PCA and calculated means of bioclimatic variables to investigate and visualize environmental relationships among our sampled populations. We utilized the 19 bioclimatic variables from the WorldClim version 2.1 database [64] and variables were downloaded at raster grids of 2.5 arc minutes (ca. 5 km) to compensate for the total area of the largest populations sampled. The full names and abbreviations for bioclimatic variables used for each environmental analysis are listed in Table A6. To avoid multicollinearity, one of each pair of highly correlated variables ($r > 0.7$) was removed [58], leaving 6 bioclimatic variables for analysis. Of these remaining variables, 4 were associated with temperature and 2 with precipitation. PCA was used to visualize climatic similarities and dissimilarities of the 23 sampled populations using the R package ggplot2.

In an additional analysis examining environmental differences between the regions, we assessed environmental niche overlap. To do so, we used a dataset that consisted of 303 occurrence points (243 from the East and 60 from the West) after accounting for spatial autocorrelation. Occurrence points were provided by collaborators who produced the IUCN Red List assessment for oaks in the United States [30]. The 19 bioclimatic variables were downloaded for 2.5 arc minutes resolution. Highly correlated variables ($r > 0.7$) were removed, leaving 7 bioclimatic variables for analysis, of which 4 were associated with temperature and 2 were associated with precipitation (Table A6). Niche differentiation in environmental space and also niche equivalency, similarity, and overlap were calculated using the R package ecospat v3.2 [65]. Niche overlap was measured by Schoener's D index, which ranges from 0 (no overlap) to 1 (full overlap; [66]).

To visualize specific environmental and genetic (i.e., heterozygosity, allelic richness, and relatedness) relationships within and among regions, we built models by multiple linear regressions for 3 genetic statistics (heterozygosity, allelic richness, and relatedness) using uncorrelated environmental variables at several geographic scales. Given the disjunct and patchy distribution of this species, and confirmation that different variables were found to be correlated within each region, we used three different sets of uncorrelated bioclimatic variables corresponding to: (1) the range-wide distribution including both eastern and western sampled populations; (2) eastern populations only; (3) western populations only. Spearman correlation analyses were performed for every subgroup and correlated variables were removed ($r > 0.7$). A final list of uncorrelated variables used as inputs for each model can be found in Table A6. The variables selected across all subgroups were subjected to a final check for collinearity using variance inflation factor ($VIF \leq 10$) [58]. We compared regression models using the all-subsets variable reduction approach (R function `regsubsets`, package `leaps` v.3.1 [67]) and selected the “best supported” model using Bayes’ BIC. Model performance was assessed using adjusted coefficients of determination (adjusted R^2) and the Benjamini and Hochberg [68] multiple comparisons correction for p -values.

2.6. Estimation of Divergence between East and West Regions Using ABC and Genetic Data

Modeling-based simulation analyses, such as approximate Bayesian computation (ABC), have been useful in elucidating the evolutionary histories of species [69–71]. We performed an ABC analysis to estimate the timing of divergence between the East and West regions. We compared two simple hypotheses: (1) splitting and isolation (the I model) or (2) splitting and continued migration (the IM model). We estimated the following parameters: splitting time (TS), mutation rate-scaled population sizes of each region, θ_1 and θ_2 (representing the East and West regions, respectively), and migration (M). We used the R code developed by Navascués [72], which runs simulations in the coalescent simulator `ms` [73] and uses random forest [74,75] for parameter estimation. This software simulates parameters in the form of θ , M and TS, which are scaled parameters rather than the parameters of biological interest here (i.e., population size and migration). An estimation of mutation rate is needed to transform these to population size, migration rate and splitting time on scales amenable to biological interpretation. We searched the literature for plant microsatellite mutation rates per site per generation and found four: wheat (2.4×10^{-4}), maize (7.7×10^{-4}), *Arabidopsis* (8.87×10^{-4}), and red cedar (6.3×10^{-4}) [76–79]. We then used the mean of the four values, 6.3×10^{-4} .

3. Results

3.1. Genetic Variation and Differentiation

Genetic analyses showed an overall pattern of differentiation between the East and West regions of *Q. havardii*. Summary statistics for all populations as well as means and standard deviation across populations are shown in Table 1. Genetic diversity was higher in western populations ($He = 0.655 \pm 0.04$) than eastern populations ($He = 0.598 \pm 0.061$). Using a Mantel Test (Figure A1), isolation by distance was significant when comparing among regions, as well as in the West ($R^2 = 0.256$, $p < 0.001$), but not in the East. The results of the AMOVA show that most of the molecular variance (81%) was found within and among individuals; however, more genetic variance occurred among regions (15%) than among populations (4%). When comparing between regions (Table A2), there is significantly higher allelic richness ($p = 0.0003$), higher heterozygosity ($p = 0.015$), and lower relatedness in West populations ($p = 0.009$; Figure 2) than in East populations. With the exception of population W1, which had significantly higher clones than other western populations, we detected a higher frequency of clones within eastern populations. Patterns from the first three axes of separate PCoA analysis among individuals (Figure A2) and populations (Figure 3A) revealed genetic distinction between the two regions and subdivisions of populations within the West, with W1 being the most separated from other populations. Bayesian clustering analyses from STRUCTURE (Figure 3D–F) and

a discriminant analysis of principal components using the R package *adegenet* showed highly similar relationships among populations. For all three analyses (i.e., range-wide, eastern region, and western region), $K = 2$ was the best-supported assumption for each dataset. The next best supported number of clusters ($K = 3$) for the range-wide dataset is also shown (Figure 3D).

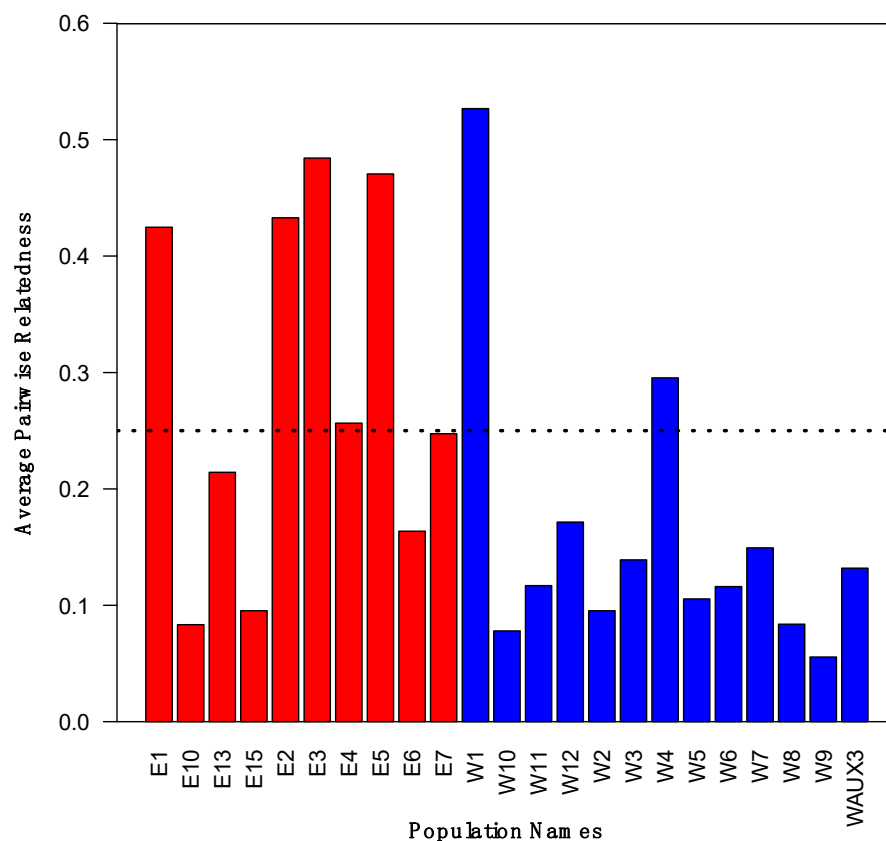


Figure 2. Mean pairwise relatedness within each *Quercus havardii* population, with blue and red indicating western and eastern populations, respectively. The expected value of relatedness for full siblings is shown by the horizontal dotted red line.

3.2. Morphological Variation

Similarly, the morphological analysis showed the presence of two groupings corresponding to the two regions. This was most pronounced when using population means, which indicated a level of differentiation similar to that seen in the genetic analysis (Figure 3B). The first two principal components represented 62.5% of the variation (Table A3; PCA plot loading scores are given in Table A4). Examination of population means showed most individual populations to cluster relatively close together in the West but with more dispersion of populations in PCA space than in the East (Figures A3 and A4). One western population, W4, was found to be highly divergent from the remaining populations. However, analysis at the individual-level (Figures A5 and A6) showed a high degree of morphological variation occurring across both regions and interdigitation between the East and West regions with a high degree of variation seen within most populations. The partitioning of morphological variance indicated that factors associated with leaf length and lower lobe angle (Table A5) contributed the most to separation between the two regions while high levels of variation at population- and tree-levels, strongest for factors associated with leaf width, indicated high levels of within-population variation (Table 2). Generally, leaves in the East were characterized by having longer leaves with slightly fewer lobes and more acute lobes, and leaves in the West were characterized by having shorter leaves that are widest at the middle lobe of the leaf blade.

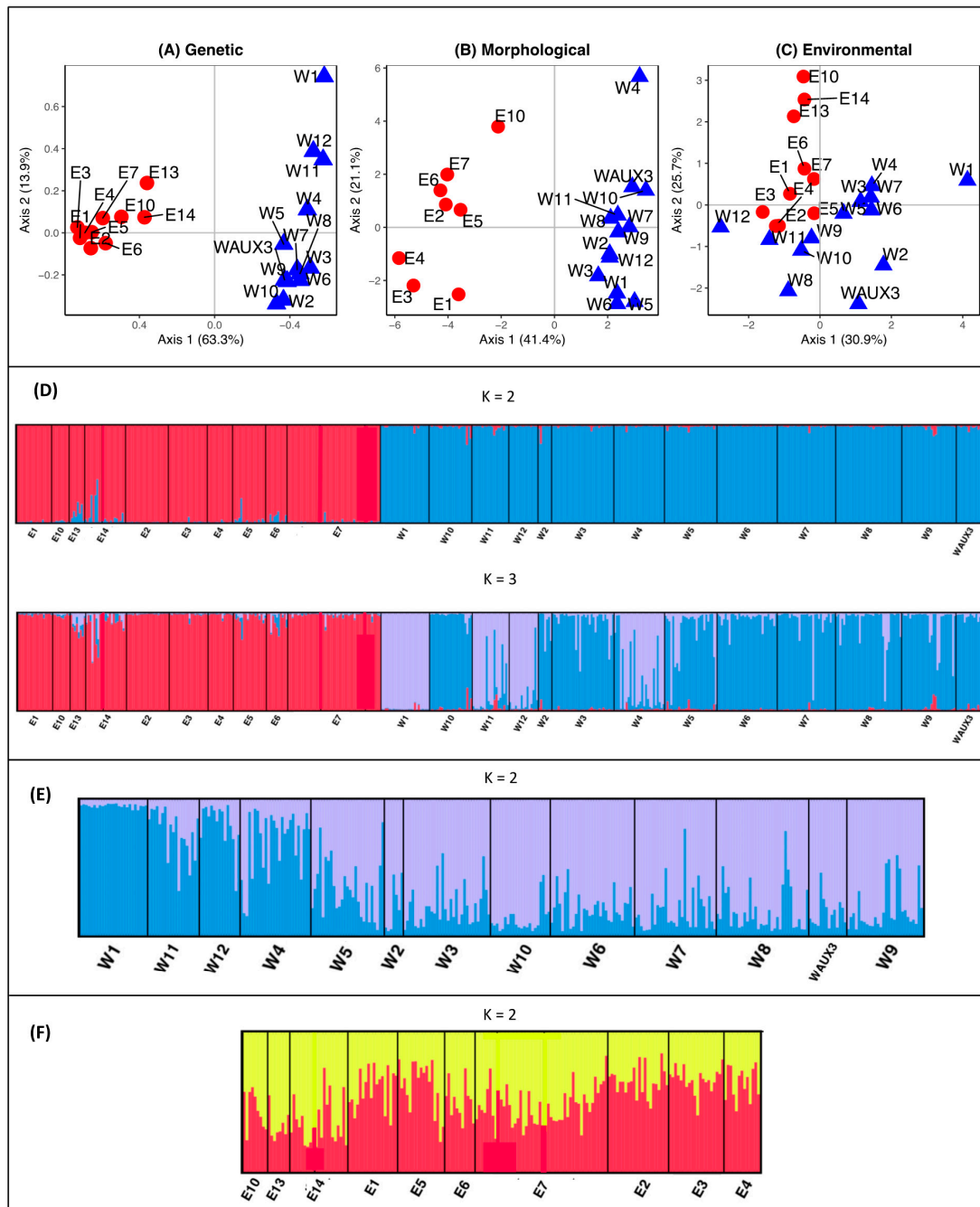


Figure 3. Principal coordinate analysis (PCoA) and principal component analyses (PCA) from 23 sampled populations from western (blue triangles) and eastern (red circles) populations of *Quercus havardii*. The percent of variation explained by each principal component is included on each axis. (A) PCoA of genetic variation among sampled populations shown by scores on the first two coordinates of variation at 11 microsatellite loci in 489 genotyped individuals; (B) PCA plot of mean population-level morphological variation on the first two coordinates derived from 12 characters in 928 measured leaves, consisting of 2 leaves per individual; (C) PCA plot of environmental variation on the first two components using 13 uncorrelated bioclimatic variables; (D) population genetic structure plots by STRUCTURE for range-wide with $K = 2$; (E) $K = 3$, within-region for eastern populations using optimal clustering value, $K = 2$; (F) within-region for western populations showing optimal clustering value, $K = 2$.

3.3. Environmental Differentiation, Variation, and Influence on Genetic Diversity

Principal component analysis of the six uncorrelated bioclimatic variables showed significant separation between populations from the East and West regions, with the first two components representing 56.6% of the variation (Figure 3C). More specifically, PC1 represented 30.9% of variation while PC2 represented 25.7% (Table A9). Maximum temperature of warmest month (BIO5), mean temperature of driest quarter (BIO9), and precipitation of coldest quarter (BIO19) showed almost equal contributions to the separation of populations along PC1, while annual temperature range (BIO7) and precipitation of the driest month (BIO14) showed along PC2 (Table A8). Mean temperature of the driest quarter was most important for the separation of West populations, while most East populations were primarily separated by mean diurnal range and precipitation of the driest and coldest months. We found higher average temperatures and precipitation in the East when compared to western populations (Table A7). Additionally, there is less seasonal temperature variability in the eastern region. Notably, one western population, W1, was found to be more climatically similar to populations in the East.

We found little evidence of environmental niche overlap ($D = 0.002$; Figure A7) between the eastern and western regions of *Q. havardii*. Additionally, western populations were observed to occupy a larger climatic niche space than eastern populations. Bioclimatic variables representing environmental extremes (e.g., precipitation of the driest month, mean temperature of the warmest quarter, and maximum temperature of the warmest month) contributed most to the separation of the East and West regions in environmental space. We found no significant differences for either niche equivalency or similarity between the two regions.

For each of the regression models with the lowest BIC, using a Benjamini–Hochberg procedure was found to be significant ($p < 0.05$) between genetic statistics and environmental variables, with the exception of the model of relatedness within the eastern region (Table 3). When examining the significant regressions across the overall geographic space (i.e., analyzing the East and West populations together), we found that the selected models showed a strong relationship mostly between temperature-related variables and all genetic diversity metrics especially with bioclimatic variables representing extreme or limiting environmental factors. For the range-wide model, allelic richness and relatedness were mostly found to be associated with seasonality-based variables, while heterozygosity was associated with variables pertaining to environmental extremes. Within the East, we found that heterozygosity and relatedness were primarily correlated with temperature extremes (mean temperature of the driest quarter; BIO9), while allelic richness was primarily correlated with precipitation extremes (precipitation of the warmest quarter; BIO18). Within the West, we found very strong correlations between both heterozygosity and allelic richness with environmental variables (both with adjusted $R^2 > 0.93$). More specifically, heterozygosity and relatedness were associated with extreme or annual trends of precipitation (precipitation of the driest month and annual precipitation; BIO14 and BIO12, respectively), while allelic richness was associated with temperature extremes (mean temperature of wettest quarter; BIO8).

3.4. Approximate Bayesian Computation (ABC) Estimate of Divergence Time

The support for the IM model over the I model is 0.758/0.232, equaling a Bayes Factor of 3.26, which is “moderate” but not “very strong” evidence [80,81]. The out-of-bag prior error rate is 42%, suggesting that it is difficult to distinguish between the two models. After untransformed values (Table A10) were converted to usable parameters (Table A11), the mean divergence time using the I model is 12,380 years, with confidence intervals ranging from 567 to 59,500 years. This suggests that divergence possibly occurred on a timescale of tens of thousands of years. We also present posterior parameter estimates based on all simulations run over both models (Table A11). When both models are included, the median estimate of the divergence time is 31,000 years ago, though the confidence intervals run from a few thousand to more than ten million. This wider confidence range seems logical

because the divergence time estimate will vary with the migration rate. Additionally, the estimated population size in the West was consistently larger than in the East.

Table 3. Genetic–environmental correlations for each of the best-supported multiple linear regression (MLR) models for *Quercus havardii* at regional and local geographic levels. The *p*-values for the variable coefficients are shown, in addition to the adjusted R^2 and *p*-values for each reported model.

	Range-Wide MLR			East MLR			West MLR		
	<i>He</i>	<i>Ar</i>	<i>Rel</i>	<i>He</i>	<i>Ar</i>	<i>Rel</i>	<i>He</i>	<i>Ar</i>	<i>Rel</i>
BIO2: Mean Diurnal Range	0.0473	0.0013	0.0105	0.0049	0.0016	0.0194			
BIO3: Isothermality							0.0118		
BIO5: Max Temp of Warmest Month	0.1034								
BIO7: Annual Temp Range	0.0002	0.0019	0.0153					0.0008	
BIO8: Mean Temp of Wettest Quarter							0.0004	0.0501	0.0011
BIO9: Mean Temp of Driest Quarter	0.0181			0.1306	0.0451	0.0775			
BIO10: Mean Temp of Warmest Quarter								0.0023	
BIO12: Annual Precipitation							0.0001		0.0635
BIO14: Precipitation of Driest Month	0.0041						0.1390	0.0006	
BIO15: Precipitation Seasonality							0.0231	0.0002	
BIO18: Precipitation of Warmest Quarter					0.1484				
BIO19: Precipitation of Coldest Quarter								0.0000	
Model Adjusted R^2	0.6121	0.5019	0.3457	0.6286	0.7527	0.5093	0.9303	0.9665	0.7257
Model <i>p</i> -value	0.0026	0.0007	0.0111	0.0259	0.0275	0.0687	0.0005	0.0003	0.0012

He: expected heterozygosity; *Ar*: allelic richness by locus; *Rel*: relatedness within a population.

4. Discussion

4.1. General Patterns of Genetic, Morphological, and Environmental Variation

We observed strong differentiation of genetic, environmental, and morphological datasets coinciding with contemporary geographic separation of populations. Across its entire geographic range, we found that *Q. havardii* has relatively moderate levels of genetic diversity and high differentiation when compared to other oaks [41,42,82,83]. While morphological differentiation, particularly population means, showed separation between the two regions principally driven by differences in leaf length, the total variance in the dataset was shown to be very high within populations. This suggests that morphological diversity and plasticity can be quite great at the individual and population level, which is a pattern typically seen and documented in oaks of different groups attributed to patterns of interspecific hybridization [84] or canopy position and light exposure [85].

Climate is likely playing a role in shaping genetic diversity within and among populations and regions of *Q. havardii*. We found stark differences in regional climate between the East and West regions (e.g., mean annual precipitation in eastern populations was nearly double that of western populations). Overall, in western populations, allelic richness is 25% higher and heterozygosity is 11% higher, while relatedness is <50%, compared to eastern populations. We also observed several correlations within regions of population genetic diversity statistics and environmental variables that have potential biological relevance to population size and persistence. Specifically, measures of genetic diversity (i.e., heterozygosity, allelic richness, and relatedness) were correlated with climatic variables that represent seasonality or extremes (e.g., heterozygosity and allelic richness were correlated

with mostly temperature seasonality or extremes in the East; while heterozygosity, allelic richness, and relatedness were correlated with both precipitation and temperature extremes in the West). This suggests that these factors may be influencing long-term population size and stability and thus genetic diversity via genetic drift. Studies in other desert plants [86] and oak species [87,88] correlating neutral microsatellite markers and climatic variables showed a similar pattern of differences at a regional level. Additional studies within oaks similarly noted seasonality-related variables as possible key factors shaping genetic diversity [89,90] and the increased rates of evolution for leaf habit [22]. Of course, these are correlations; further research on the effects of climate on oak diversity will be needed to elucidate the importance of these mechanisms. Possible adaptive trait differentiation corresponding to morphological, genetic and environmental differentiation patterns should be explored further.

An alternative influence on genetic diversity could be the relatively higher rates of human disturbance in the East. The Llano Estacado region of Texas and New Mexico has high rates of ranching and oil and gas development [30]. In such disturbed areas, population extinction and colonization may be higher, which could cause lower genetic diversity, higher relatedness, and higher genetic differentiation due to bottleneck effects. Additionally, we observed high rates of clonality, but much lower rates than were suggested by earlier allozyme data for *Q. havardii* [91]. Future work should investigate in more detail the spatial distribution of clones within populations, which might have profound effects on genetic differentiation, genotypic diversity and perhaps the long-term viability of the species. Thus, we cannot fully untangle the influences of genetic diversity in this system, but our results do present specific hypotheses to be investigated in more detail.

We found that the divergence of eastern and western populations most likely occurred on the scale of tens of thousands of years ago, according to our results from the ABC analysis. Although confidence intervals are broad, this divergence estimate implies that these populations have been isolated for hundreds of generations and are likely sister species or subspecies; however, this assumes an evolutionary relationship between these populations that cannot be confirmed by our work. Our estimates align with a broader study of the American oak clade, suggesting that *Q. havardii* likely arose in the early- to mid-Miocene, and also suggested a significant adaptive transition in ecological space, perhaps due to niche specialization [22]. The history of these populations was likely influenced by glacial cycles and consequent effects such as drying of the deserts of the Southwest and loss of suitable habitat. To resolve questions of evolutionary history, possible hybrid origins, and phylogeography, we are currently working to incorporate samples from both geographic regions into the phylogeny of North American oaks [92].

4.2. Possible Implications for Conservation and Taxonomy

Quercus havardii is an important keystone species that defines the sand shinnery ecosystem and provides a habitat to vulnerable species such as the lesser prairie chicken and dunes sagebrush lizard. The conservation status is currently listed as “Endangered” by the IUCN Red List, which notes that populations are facing increased threats of habitat loss and fragmentation arising from development for oil and gas industries and agriculture [30]. We observed relatively moderate genetic diversity and a low degree of clonality in this study suggesting that the species, and most of its populations, are not in imminent danger of genetic problems like inbreeding. However, in our study, the population W1, a putative lagging-edge population [93], was identified as extremely vulnerable due to high clonality, high inbreeding and low genetic diversity, and thus should be considered as a possible priority for conservation. Other isolated or lagging-edge populations should be investigated as well.

Additionally, the taxonomic uncertainty of *Q. havardii* remains as it is currently treated as a singular species over the full extent of its range. However, this has been debated over the years as some authors have called the western region its own taxon, *Q. havardii* var. *tuckerii* Welsh or *Q. welshii* R.A. Denham, and has also been speculated as a putative

introgressed hybrid with *Q. gambelli* [31,94,95]. Our work provided evidence that both eastern and western entities are genetically and morphologically distinct, and occupy unique environmental niche space with little to no overlap. However, future phylogenetic work may still be needed to elucidate the evolutionary relationships of these taxa and to support any taxonomic changes. If taxonomic changes are deemed valid, a reassessment of the conservation status of *Q. havardii* and *Q. welshii* will be needed to incorporate regional-specific information such as population genetic data and modified area of occurrences.

4.3. Caveats

There are several important caveats that should be mentioned in this work. It must be noted that sampling was not systematic within and across each region, nor designed a priori to evenly cover different environmental variables but rather focused on the known locations of populations. It is possible that the relationships observed in this study could be artifacts of sampling or reflect a relationship with another unknown underlying variable. Moreover, we do not incorporate data on some variables that may be most important in determining a species demographic and genetic stability, (e.g., soil attributes). Additionally, the ABC analysis has extremely wide confidence intervals and assumes very simple models. An alternative hypothesis that these data do not support, but should not yet be ruled out yet, is that the divergence is deeper in time and that these are not sister taxa but rather have evolved similar morphology, habitat and habitat preference via convergent evolution.

Of additional importance, hybridization is possible among many oak taxa, and hybrids have been observed in many studies. Although oaks have been deemed taxonomically difficult due to high rates of intraspecific gene exchange, genetic investigations show that contemporary hybridization is usually observed at relatively low levels, such that species may maintain coherence [96]. We do not examine hybridization in this study, and only focus on individuals morphologically and ecologically consistent with *Q. havardii*; however, morphologies ascribed to *Q. eastwoodiae* or the *Q. x undulata* complex were observed (FLD and MOR herbarium specimens McCauley 700, 701, and 702). Additionally, several collected populations (W1, W4, W11, and W12) did not occur in deep, shifting sand and were noted as having other *Quercus* species located only a few kilometers away. It is likely that the highly specific and unique sand habitat of most populations of *Q. havardii* could be playing a role in observed diversity patterns by significantly limiting the opportunity for interspecific gene flow. Nonetheless, hybridization and its possible influence on diversity and population persistence is an important future area to investigate for this species.

5. Conclusions

We have provided the first range-wide genetic study of one of the arid-adapted oaks of the Southwest USA and one of the relatively few integrated genetic, ecological, and morphological characterization of oaks in general. *Q. havardii* has somewhat lower genetic diversity and higher genetic differentiation when compared with other widespread and important oaks [97], as might be expected with this species' highly specialized and fragmented habitat. We find strong differentiation between eastern and western regions, potentially at the level of evolutionary significant units (ESUs) or higher, suggesting important conservation designation. Within this larger regional context, correlations with environmental variables suggest that temperature seasonality, precipitation extremes, and other key factors may be influencing local levels of genetic diversity. Future studies of arid-adapted oaks are needed as climates continue to rapidly change and make habitats increasingly vulnerable, thus emphasizing the importance of these species as possible sources of genetic resources for breeding or restoration.

Author Contributions: Conceptualization, S.H., D.D. and R.A.M.; methodology, B.A.Z., D.D., E.S.; software, B.A.Z., S.H., R.A.M.; validation, B.A.Z., S.H., R.A.M., D.D., E.S.; formal analysis, B.A.Z., S.H., R.A.M., D.D.; investigation, B.A.Z., S.H., R.A.M., I.J.F., D.D., E.S.; resources, S.H. and R.A.M.; data curation, B.A.Z., R.A.M., S.H.; writing—original draft preparation, B.A.Z. and S.H.; writing—review and editing, B.A.Z., S.H., R.A.M., I.J.F., D.D., E.S.; visualization, B.A.Z., R.A.M., E.S., S.H.;

supervision, S.H.; project administration, S.H.; funding acquisition, S.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the U.S. Forest Service and American Public Gardens Association Tree Gene Conservation Partnership and also in part by The Morton Arboretum.

Data Availability Statement: The data and code for analyses used in this study are openly available on the Github repository *Quercus_havardii_Forests_Manuscript* at doi:10.5281/zenodo.4453098.

Acknowledgments: We would like to thank the three anonymous reviewers for their insightful comments on the manuscript. We thank Mackenzie Codon and Cindy Johnson for their assistance with lab work, and Kevin Feldheim and Isabel Distefano for providing their technical support on fragment analysis equipment and protocols. Additionally, we thank the Field Museum for providing us access to equipment for microsatellite analysis. The authors would like to thank the following permit-granting agencies: Bureau of Land Management, Texas Parks and Wildlife Department, New Mexico Department of Fish and Game, The Nature Conservancy, US Forest Service, Navajo Nation Department of Fish and Wildlife. We also thank Robert D. Cox, Jamie Baker, Brandon Childers, and Mike Melendrez for their assistance with the collection of material. The authors also thank Drs. Jamie Gillooly and Al Meyers for their valuable comments that greatly improved the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. Microsatellite markers used for *Quercus havardii* population genetics.

Locus	Source [Citation]	Fluorescent Dye	Repeat Unit
QS03797	Chatwin et al. 2014 [98]	6-FAM	(CA) ₇
MSQ13	Dow et al. 1995 [99]	6-FAM	(TC) ₁₄
QrZAG20	Kampfer et al. 2004 [100]	VIC	(TC) ₁₈
QpZAG110	Steinkellner et al. 1997 [101]	NED	(AG) ₁₅
QS00314	Chatwin et al. 2014 [98]	PET	(GAA) ₆
QS1904	Chatwin et al. 2014 [98]	6-FAM	(TC) ₁₀
MSQ4	Dow et al. 1995 [99]	VIC	(AG) ₁₇
QS00562	Steinkellner et al. 1997 [101]	6-FAM	(GA) ₇
QrZAG87	Kampfer et al. 2004 [100]	NED	(TC) ₂₀
QpZAG1.5	Steinkellner et al. 1997 [101]	PET	(GT) ₅ (GA) ₉
QM69.2M1	Isagi and Suhandono 1997 [102]	PET	(TGG) ₆ (CGG)(TGG) ₂

Table A2. Genetic summary statistics by region.

Parameter	East Mean	West Mean	T-test <i>p</i> -Value	Wilcoxon <i>p</i> -Value
Allelic richness	47.160	58.877	0.0003	0.0013
Heterozygosity	0.587	0.653	0.0150	0.0148
Pairwise F_{ST}	0.085	0.079	0.3263	0.1111
Relatedness	0.176	0.077	0.0063	0.0093

Table A3. Eigenvalues and percent variability for each principal component for the PCA using 17 uncorrelated morphological variables presented in Figure 3B.

	Eigenvalue	Variance (%)	Cumulative Variance (%)
PC1	7.039	41.4	41.4
PC2	3.584	21.1	62.5
PC3	3.290	19.4	81.8
PC4	1.459	8.6	90.4
PC5	0.679	4.0	94.4
PC6	0.494	2.9	97.3
PC7	0.178	1.0	98.4
PC8	0.093	0.5	98.9
PC9	0.075	0.4	99.4
PC10	0.045	0.3	99.6
PC11	0.032	0.2	99.8

Table A3. Cont.

	Eigenvalue	Variance (%)	Cumulative Variance (%)
PC12	0.021	0.1	99.9
PC13	0.007	0.0	100.0
PC14	0.003	0.0	100.0
PC15	0.001	0.0	100.0
PC16	0.000	0.0	100.0
PC17	0.000	0.0	100.0

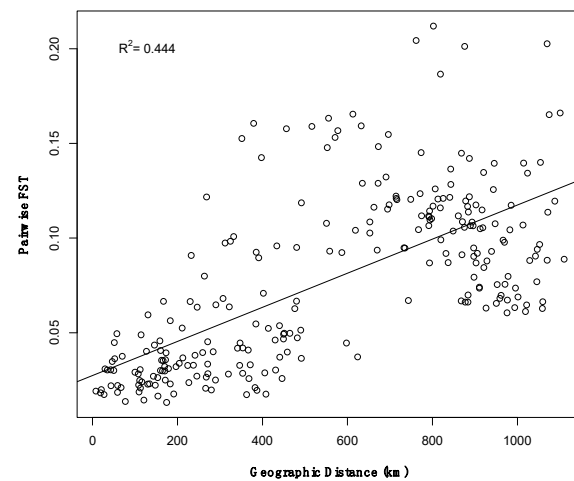


Figure A1. Mantel test to evaluate isolation by distance within *Quercus havardii*. Pairwise F_{ST} is compared to geographic distance (km) between each pair of populations.

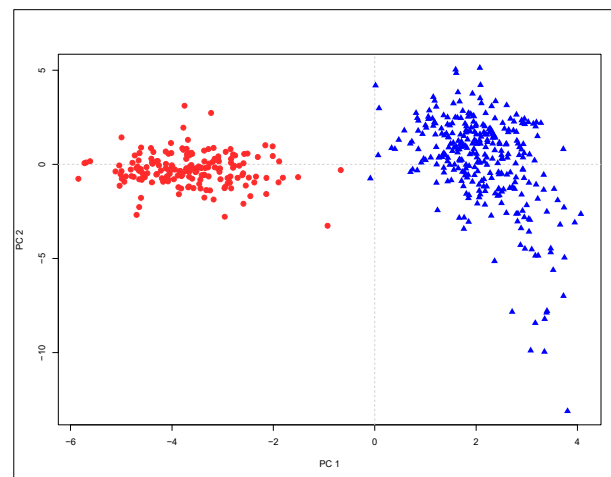


Figure A2. PCoA comparing genetic characteristics among 489 individuals of *Quercus havardii*. Western and eastern individuals are represented by blue triangles and red circles, respectively.

Table A4. Plot loading scores of principle components axes with eigenvalues greater than 1 for the PCA using 17 uncorrelated morphological variables presented in Figure 3B. Important loadings that contribute more than one variable's worth of information are in bold.

Character	PC1	PC2	PC3	PC4
Lobes	0.32	0.07	-0.20	0.19
Length	-0.35	0.08	-0.18	-0.08
BLBW	-0.26	-0.21	-0.25	0.09
MLBW	0.08	-0.15	-0.51	-0.12

Table A4. *Cont.*

Character	PC1	PC2	PC3	PC4
ALBW	−0.22	−0.14	−0.02	−0.59
LVL	−0.31	0.18	−0.19	−0.04
UVL	−0.04	0.19	−0.49	−0.05
UMLS	−0.08	−0.43	−0.22	−0.19
LMLS	−0.20	−0.37	−0.20	0.12
LLA	0.35	−0.07	−0.10	−0.15
ULA	−0.19	0.31	0.16	−0.17
MLA	0.26	0.09	−0.06	−0.41
LWR_1	0.20	−0.36	0.01	0.25
LWR_2	0.35	−0.17	−0.11	0.03
LWR_3	0.20	−0.23	0.25	−0.42
LODR_1	0.17	0.33	−0.28	0.11
LODR_2	0.24	0.29	−0.22	−0.24

The bolding here highlights the characters that most influence PCA's.

Table A5. Correlation of variables with the first two axes in a discriminant analysis of morphological variation across the full range of *Q. havardii*. The three most influential characters for each axis are indicated in bold.

Character	Abbreviation	CV1	CV2
Lobe Number	LOBES	−0.332	−0.068
Leaf Length	LENGTH	0.052	0.025
Basal Lobe Blade Width	BLBW	0.018	−0.300
Middle Lobe Blade Width	MLBW	−0.135	0.299
Apical Lobe Blade Width	ALBW	0.058	−0.392
Lower Vein Length	LVL	0.041	−0.057
Upper Vein Length	UVL	0.018	−0.084
Upper Middle Lobe Sinus Width	UMLS	0.008	0.174
Lower Middle Lobe Sinus Width	LMLS	0.065	−0.075
Lower Lobe Angle	LLA	−0.034	0.007
Upper Lobe Angle	ULA	0.014	0.013
Middle Lobe Angle	MLA	0.001	−0.013
Length to Width Ratio 1 (BLBW/Length)	LWR_1	−0.332	−0.068
Length to Width Ratio 2 (MLBW/Length)	LWR_2	0.052	0.025
Length to Width Ratio 3 (ALBW/Length)	LWR_3	0.018	−0.300
Lobe Depth Ratio 1	LODR_1	−0.135	0.299
Lobe Depth Ratio 2	LODR_2	0.058	−0.392

The bolding here highlights the characters that most influence PCA's.

Table A6. Uncorrelated bioclimatic variables used as inputs for the principle component analysis (Figure 1), building multiple linear regressions (MLR) models for each of the three regional and local geographic level datasets (Table 3), and variables used for the environmental niche overlap analysis (Figure A7).

Bioclimatic Variable	Code	Range-Wide MLR and PCA	East MLR	West MLR	Niche Overlap
Annual Mean Temperature	BIO1				
Mean Diurnal Range	BIO2	X	X		
Isothermality	BIO3			X	
Temperature Seasonality	BIO4		X		
Maximum Temperature of Warmest Month	BIO5	X			X
Minimum Temperature of Coldest Month	BIO6				
Temperature Annual Range	BIO7	X		X	X
Mean Temperature of Wettest Quarter	BIO8		X	X	X
Mean Temperature of Driest Quarter	BIO9	X	X		X
Mean Temperature of Warmest Quarter	BIO10			X	
Mean Temperature of Coldest Quarter	BIO11				
Annual Precipitation	BIO12			X	
Precipitation of Wettest Month	BIO13				
Precipitation of Driest Month	BIO14	X		X	X
Precipitation Seasonality	BIO15			X	X
Precipitation of Wettest Quarter	BIO16				
Precipitation of Driest Quarter	BIO17				
Precipitation of Warmest Quarter	BIO18		X		X
Precipitation of Coldest Quarter	BIO19	X		X	

Table A7. Summary of the bioclimatic variable means from the 23 populations of *Quercus havardii* sampled in this study.

Pop	BIO 1	BIO2	BIO3	BIO4	BIO5	BIO6	BIO7	BIO8	BIO9	BIO10	BIO 11	BIO 12	BIO 13	BIO 14	BIO 15	BIO 16	BIO 17	BIO 18	BIO 19
E1	15	17.1	44	801.7	34.1	-4.8	38.9	23.8	5	24.7	5	463	69	12	61.6	201	42	198	42
E2	15.1	17.2	43.8	801.9	34.3	-4.9	39.2	24.1	5.1	24.8	5.1	385	70	8	68.8	190	31	179	31
E3	14.2	17.7	44.8	797.2	33.5	-5.9	39.4	23.8	5.9	23.8	4.3	412	74	10	69.9	199	32	199	32
E4	14.3	17.9	45.1	798.2	33.8	-5.9	39.7	23.9	4.3	23.9	4.3	421	74	9	69	200	32	200	32
E5	14.5	15	36.8	929.4	34.7	-5.9	40.7	23.5	2.9	25.9	2.9	528	89	9	58.1	219	42	195	42
E6	14.6	14.8	37	906	34.6	-5.4	40	23.4	3.3	25.7	3.3	568	93	13	56	228	49	198	49
E7	14.5	14.5	35.9	927.2	34.6	-5.9	40.5	18.9	2.9	25.8	2.9	570	96	12	54.8	229	51	193	51
E10	17.5	14.8	39.7	808.5	35.6	-1.6	37.2	26.5	7.1	27	7.1	492	87	16	55.6	184	53	157	53
E13	15.9	14.7	39.5	820.7	34.1	-3	37.1	24.1	5.6	25.8	5.6	541	78	14	55.1	208	52	203	52
E15	16.4	15.1	39.6	837	35.3	-2.7	38.1	24.7	5.9	26.6	5.9	543	79	17	53	201	57	196	57
W1	17.1	19.7	47.6	793.9	37.8	-3.5	41.3	26.7	20.2	27.1	7.7	377	53	7	41.3	137	38	103	99
W2	12.2	17.1	38.9	956	33.1	-10.9	44	19.1	16.5	24.1	0.5	214	27	6	29.4	72	32	51	51
W3	11.8	15.9	38.7	906.2	32.1	-9.1	41.1	22.4	15.8	23.3	0.9	233	31	7	34.3	82	32	65	53
W4	12.3	16.4	40.4	876.8	32.4	-8.2	40.6	22.7	16.1	23.5	1.9	262	35	8	32.4	84	35	68	69
W5	12.1	16	39.1	903.7	32.4	-8.6	41	22.7	16.2	23.5	1.3	229	32	7	33.1	79	32	64	54
W6	12.3	14.2	35.1	943.8	32.1	-8.3	40.3	19	16.5	24.2	1	208	26	7	32.5	73	30	55	45
W7	11.9	13.2	34	926.1	30.8	-8	38.7	18.6	15.9	23.6	0.8	233	28	8	29.8	79	34	60	56
W8	12.2	16	36.3	996.7	33.2	-10.7	43.9	19	-0.1	24.6	-0.1	179	22	7	31.3	64	33	47	33
W9	12.6	16.3	36.5	1002.6	33.7	-10.8	44.5	19.6	22.3	25.1	0.2	245	31	9	24.9	74	51	53	52
W10	12	17	36.3	1041.7	33.8	-13.1	46.9	19.2	-1.1	24.8	-1.1	226	29	9	25.3	71	47	50	47
W11	12.5	19.4	45.6	847.6	33.4	-9.1	42.5	22.3	7.7	23.1	2.2	239	48	10	62.9	120	32	98	35
W12	9.1	15	41.7	784.5	27.4	-8.6	35.9	19.3	0	19.3	0	364	65	9	58.2	161	33	161	33
WAUX3	12.3	16.2	36.5	1003.8	33.6	-10.8	44.4	19.1	0	24.8	0	177	21	7	29.9	62	33	46	33
East Mean	15.2	15.9	40.6	842.8	34.5	-4.6	39.1	23.7	4.8	25.4	4.6	492.3	80.9	12.0	60.2	205.9	44.1	191.8	44.1
East ± SD	1.1	1.4	3.5	55.5	0.7	1.6	1.3	1.9	1.4	1.1	1.5	68.2	9.7	3.1	6.7	15.1	9.7	13.9	9.7
West Mean	12.3	16.3	39.0	921.8	32.8	-9.2	41.9	20.7	11.2	23.9	1.2	245.1	34.5	7.8	35.8	89.1	35.5	70.8	50.8
West ± SD	1.7	1.8	4	80.8	2.3	2.3	2.9	2.4	8.6	1.8	2.2	60.6	13	1.2	11.8	30.5	6.3	32.5	18.1

Table A8. Plot loading scores of principle components axes with eigenvalues greater than 1 for the PCA using bioclimatic data from 6 uncorrelated variables presented in Figure 3C. Important loadings that contribute more than one variable's worth of information are in bold.

Bioclimatic Variable	PC1	PC2	PC3
BIO2: Mean Diurnal Range	-0.113	-0.268	0.331
BIO5: Max Temp of Warmest Month	0.410	0.062	0.663
BIO7: Annual Temp Range	0.101	-0.635	0.375
BIO9: Mean Temp of Driest Quarter	0.623	-0.002	-0.389
BIO14: Precipitation of Driest Month	-0.334	0.610	0.364
BIO19: Precipitation of Coldest Quarter	0.556	0.385	0.164

The bolding here highlights the characters that most influence PCA's.

Table A9. Eigenvalues and percent variability for each principal component for the PCA using bioclimatic data from 6 uncorrelated variables presented in Figure 3C.

	Eigenvalue	Variance (%)	Cumulative Variance (%)
PC1	1.852	30.9	30.9
PC2	1.542	25.7	56.6
PC3	1.322	22.0	78.6
PC4	0.921	15.3	94.0
PC5	0.275	4.6	98.5
PC6	0.088	1.5	100.0

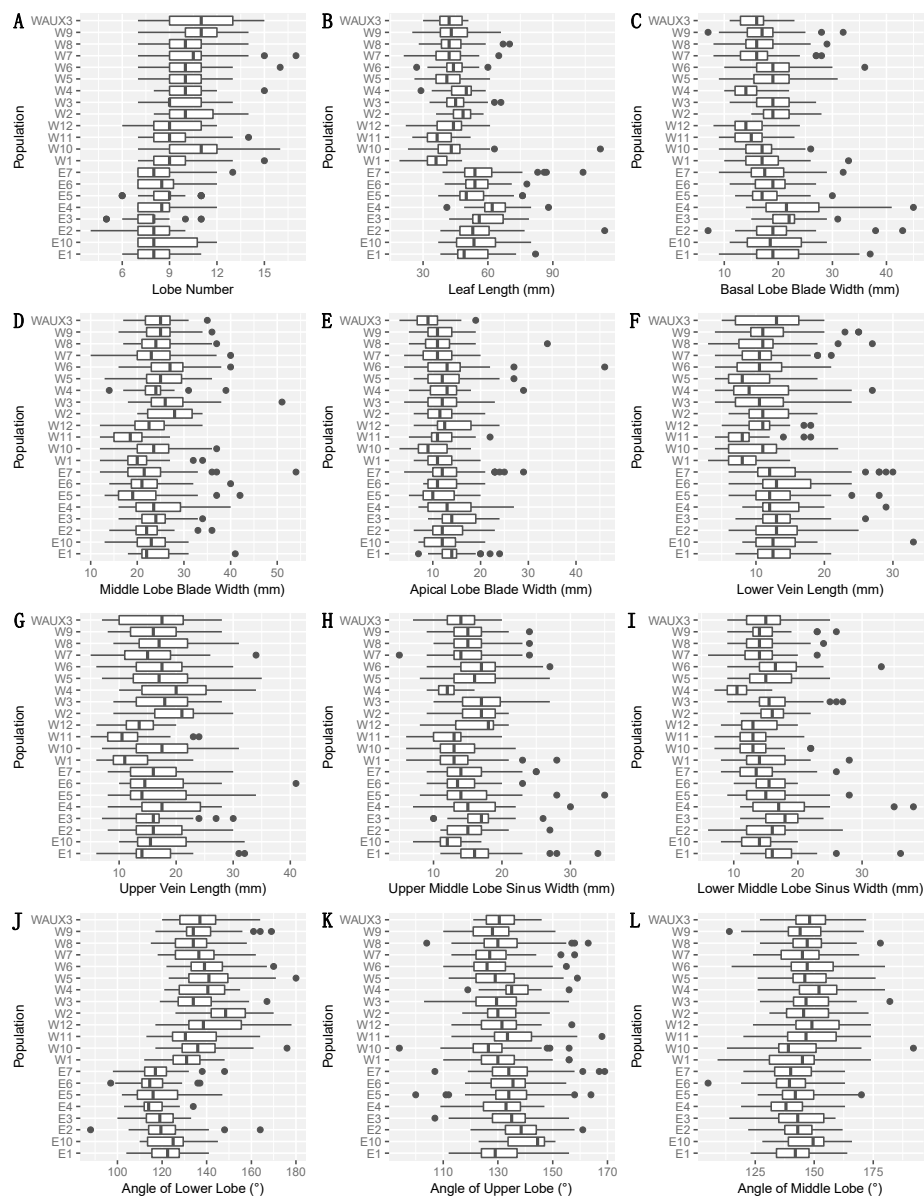


Figure A3. Box plots of 12 directly measured variables in 21 populations of *Quercus havardii* showing (A) leaf lobe number; (B) leaf length; (C) basal lobe blade width; (D) middle lobe blade width; (E) apical lobe blade width (F) lower vein length; (G) upper vein length; (H) upper middle lobe sinus width; (I) lower middle lobe sinus width; (J) angle of lower lobe; (K) angle of upper lobe; and (L) angle of middle lobe. Boxes depict standard Tukey representations. Population names follow Table 1.

Table A10. Untransformed parameter estimates from approximate Bayesian computation analysis for splitting time (TS), migration (M), and theta 1 and theta 2 (representing the East and West regions, respectively) under the isolation with migration (IM) and isolation only (I) models.

Model	Parameter	0.025 CI	Median	0.975 CI
IM	Theta 1	0.396	2.17	16.9
IM	Theta 2	2.35	4.1	75.5
IM	M	1.75×10^{-10}	0.177	5.52
IM	TS	0.0157	0.181	76.7
I	Theta 1	0.1816	2.105	40.96
I	Theta 2	1.938	3.47	10.43
I	M	1.75×10^{-10}	0.1766	5.518
I	TS	0.003393	0.07413	0.3564

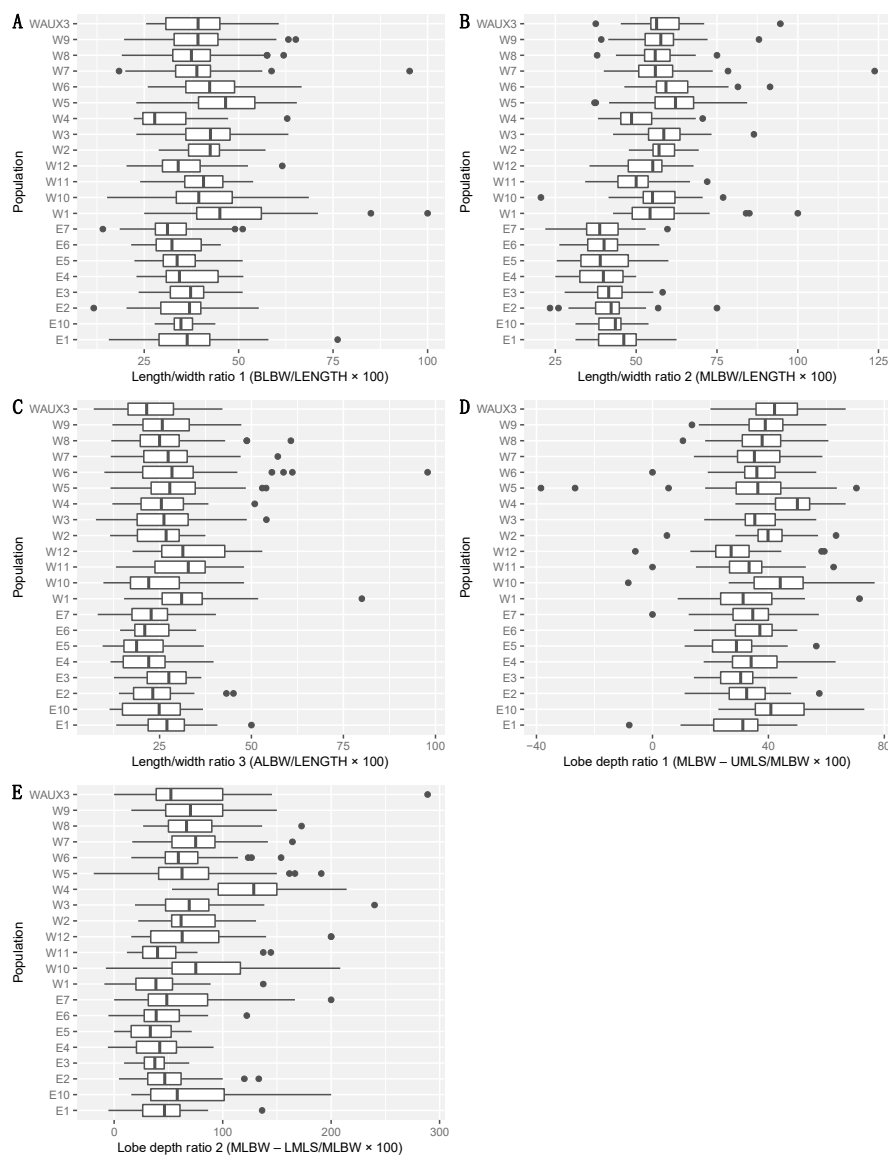


Figure A4. Box plots of five ratio variables derived from directly measured variables in 21 populations of *Quercus havardii* showing (A) length/width ratio 1 (BLBW/LENGTH); (B) length/width ratio 2 (MLBW/LENGTH); (C) length/width ratio 3 (ALBW/LENGTH); and (D) lobe depth ratio 1 (MLBW—UMLS/MLBW); (E) lobe depth ratio 2 (MLBW—LMLS/MLBW). Boxes depict standard Tukey representations. Population names follow Table 1.

Table A11. Transformed parameter estimates from approximate Bayesian computation analysis for splitting time (TS), migration (M), and theta 1 and theta 2 (representing the East and West regions, respectively) under the isolation with migration (IM) and isolation only (I) models.

Model	Parameter	0.025 CI	Median	0.975 CI
IM	Theta 1	157	863	6710
IM	Theta 2	930	1630	29,900
IM	M	5.07×10^{-10}	5.12×10^{-5}	1.60×10^{-3}
IM	TS	2710	31,100	13,200,000
I	Theta 1	72	835	16,300
I	Theta 2	769	1377	4140
I	M	5.23×10^{-10}	5.29×10^{-5}	1.65×10^{-3}
I	TS	567	12,380	59,500

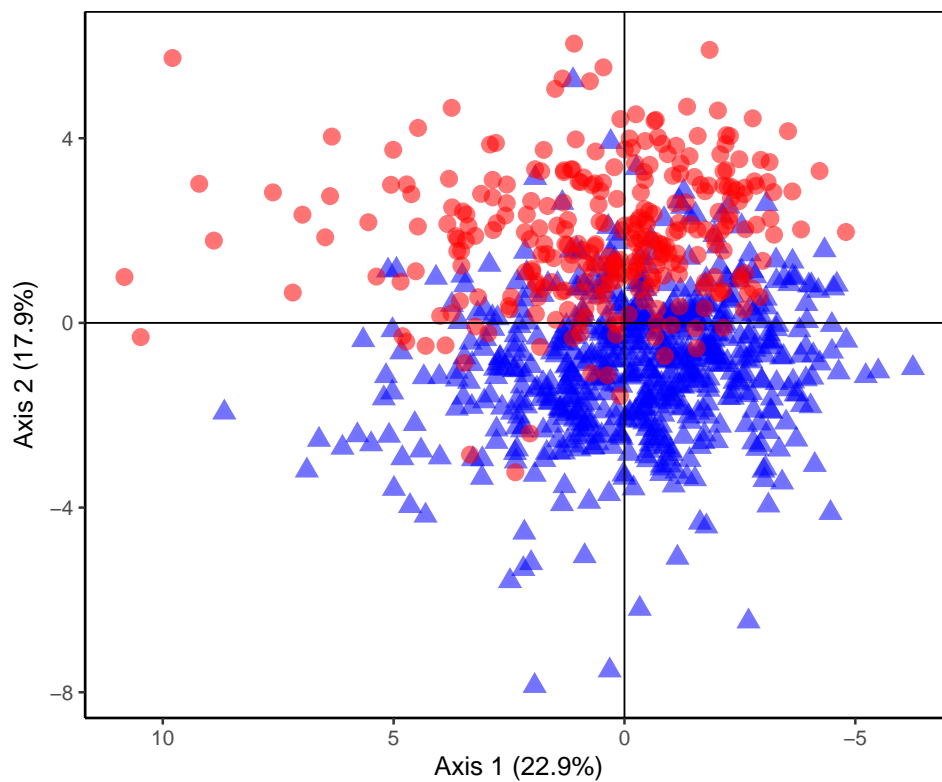


Figure A5. PCA comparing morphological characteristics among all sampled individuals of *Quercus havardii*, with the first two components explaining 40.8% of variation. Individuals from the western and eastern regions are represented by blue triangles and red circles, respectively.

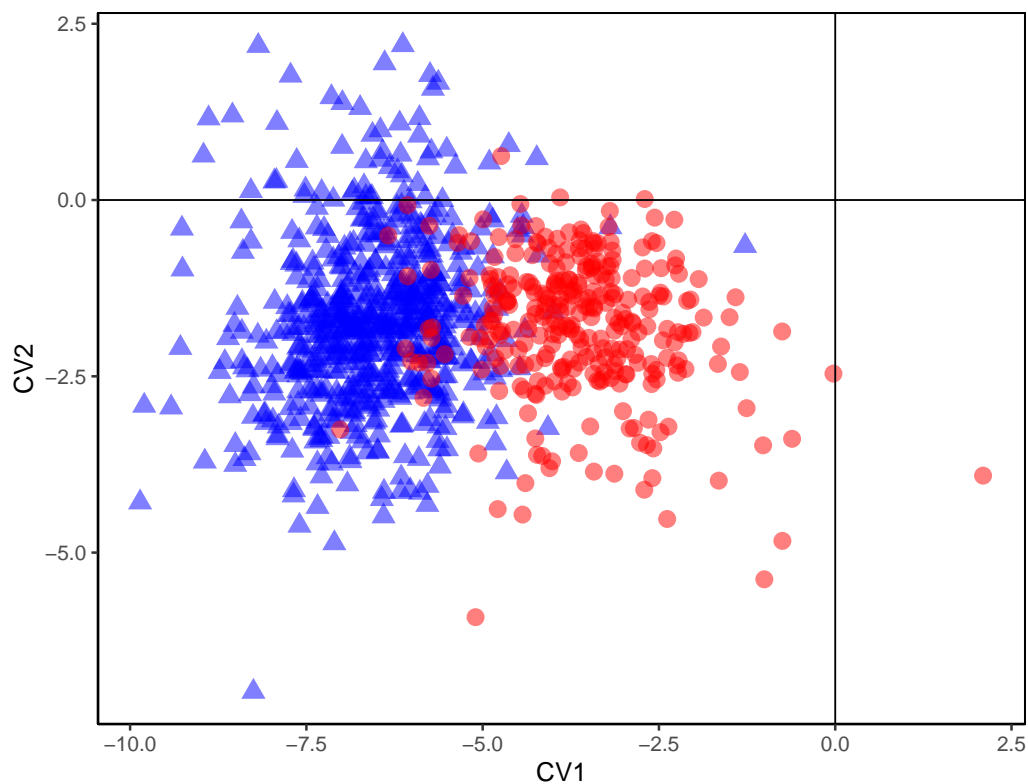


Figure A6. CVA plot of the first two canonical variates axes depicting segregation between western and eastern groups of *Q. havardii*. Western and eastern individuals are represented by blue triangles and red circles, respectively.

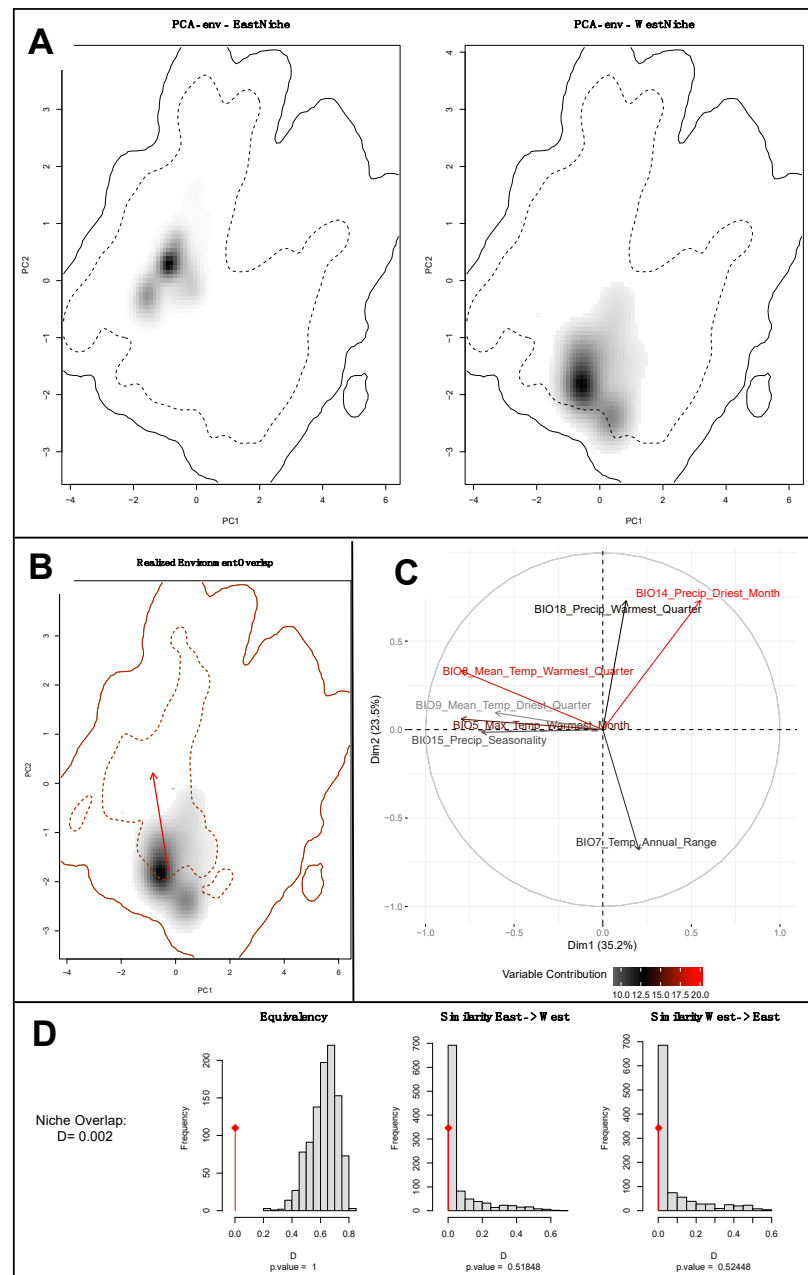


Figure A7. Environmental niche overlap analysis of *Quercus havardii* populations from the eastern and western regions of the species' disjunct range showing: (A) plot of the occupancy of the environment by populations located in the East (left) and West (right); (B) area of niche unique to eastern populations (red), western populations (blue), and estimated area of shared overlap (green); (C) correlation circle of the loadings of individual environmental variables to the two PCA axes; (D) estimated niche overlap values for the East and West regions in terms of Schoener's D, and niche similarity and equivalency tests with corresponding *p*-values.

References

1. Shafer, A.B.A.; Wolf, J.B.W. Widespread Evidence for Incipient Ecological Speciation: A Meta-Analysis of Isolation-by-Ecology. *Ecol. Lett.* **2013**, *16*, 940–950. [[CrossRef](#)] [[PubMed](#)]
2. Coyne, J.A. Speciation in a Small Space. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 12975–12976. [[CrossRef](#)] [[PubMed](#)]
3. Slatkin, M. Gene Flow and the Geographic Structure of Natural Populations. *Science* **1987**, *236*, 787–792. [[CrossRef](#)]
4. Ellstrand, N.C.; Elam, D.R. Population Genetic Consequences of Small Population Size: Implications for Plant Conservation. *Annu. Rev. Ecol. Syst.* **1993**, *24*, 217–242. [[CrossRef](#)]

5. Jain, S.K.; Bradshaw, A.D. Evolutionary Divergence among Adjacent Plant Populations I. The Evidence and Its Theoretical Analysis. *Heredity* **1966**, *21*, 407–441. [[CrossRef](#)]
6. Schluter, D. Ecology and the Origin of Species. *Trends Ecol. Evol.* **2001**, *16*, 372–380. [[CrossRef](#)]
7. Rundle, H.D.; Nasil, P. Ecological Speciation. *Ecol. Lett.* **2005**, *8*, 336–352. [[CrossRef](#)]
8. Thorne, R.F. Major Disjunctions in the Geographic Ranges of Seed Plants. *Q. Rev. Biol.* **1972**, *47*, 365–411. [[CrossRef](#)]
9. Honnay, O.; Verheyen, K.; Butaye, J.; Jacquemyn, H.; Bossuyt, B.; Hermy, M. Possible Effects of Habitat Fragmentation and Climate Change on the Range of Forest Plant Species. *Ecol. Lett.* **2002**, *5*, 525–530. [[CrossRef](#)]
10. Beatty, G.E.; Provan, J. Post-Glacial Dispersal, Rather than in Situ Glacial Survival, Best Explains the Disjunct Distribution of the Lusitanian Plant Species *Daboecia cantabrica* (Ericaceae). *J. Biogeogr.* **2012**, *40*, 335–344. [[CrossRef](#)]
11. Hewitt, G.M. Post-Glacial Re-Colonization of European Biota. *Biol. J. Linn. Soc.* **1999**, *68*, 87–112. [[CrossRef](#)]
12. Zhao, M.; Chang, Y.; Kimball, R.T.; Zhao, J.; Lei, F.; Qu, Y. Pleistocene Glaciation Explains the Disjunct Distribution of the Chestnut-Vented Nuthatch (*Aves*, Sittidae). *Zool. Scr.* **2018**, *48*, 33–45. [[CrossRef](#)]
13. Grant, V. *Plant Speciation*, 2nd ed.; Columbia University Press: New York, NY, USA, 1981.
14. Ortego, J.; Riordan, E.C.; Gugger, P.F.; Sork, V.L. Influence of Environmental Heterogeneity on Genetic Diversity and Structure in an Endemic Southern Californian Oak. *Mol. Ecol.* **2012**, *21*, 3210–3223. [[CrossRef](#)] [[PubMed](#)]
15. Ülker, E.D.; Tavşanoğlu, Ç.; Perktas, U. Ecological Niche Modelling of Pedunculate Oak (*Quercus robur*) Supports the ‘Expansion–Contraction’ Model of Pleistocene Biogeography. *Biol. J. Linn. Soc.* **2018**, *123*, 338–347. [[CrossRef](#)]
16. Di Pietro, R.; Conte, A.L.; Di Marzio, P.; Fortini, P.; Farris, E.; Gianguzzi, L.; Müller, M.; Rosati, L.; Spampinato, G.; Gailing, O. Does the Genetic Diversity among Pubescent White Oaks in Southern Italy, Sicily and Sardinia Islands Support the Current Taxonomic Classification? *Eur. J. For. Res.* **2020**, *140*, 355–371. [[CrossRef](#)]
17. Fortini, P.; Antonecchia, G.; Di Marzio, P.; Maiuro, L.; Viscosi, V. Role of Micromorphological Leaf Traits and Molecular Data in Taxonomy of Three Sympatric White Oak Species and Their Hybrids (*Quercus* L.). *Plant Biosyst. Int. J. Deal. All Asp. Plant Biol.* **2015**, *149*, 546–558.
18. Fortini, P.; Di Marzio, P.; Di Pietro, R. Differentiation and Hybridization of *Quercus frainetto*, *Q. petraea*, and *Q. pubescens*, (Fagaceae): Insights from Macro-Morphological Leaf Traits and Molecular Data. *Plant Syst. Evol.* **2015**, *301*, 375–385. [[CrossRef](#)]
19. Fortini, P.; Viscosi, V.; Maiuro, L.; Fineschi, S.; Vendramin, G. Comparative Leaf Surface Morphology and Molecular Data of Five Oaks of the Subgenus *Quercus* Oerst (Fagaceae). *Plant Biosyst.* **2009**, *143*, 543–554. [[CrossRef](#)]
20. Manos, P.S.; Doyle, J.J.; Nixon, K.C. Phylogeny, Biogeography, and Processes of Molecular Differentiation in *Quercus* Subgenus *Quercus* (Fagaceae). *Mol. Phylogenetics Evol.* **1999**, *12*, 333–349. [[CrossRef](#)] [[PubMed](#)]
21. Cavender-Bares, J.; González-Rodríguez, A.; Eaton, D.A.; Hipp, A.A.; Beulke, A.; Manos, P.S. Phylogeny and Biogeography of the American Live Oaks (*Quercus* Subsection *Virentes*): A Genomic and Population Genetics Approach. *Mol. Ecol.* **2015**, *24*, 3668–3687. [[CrossRef](#)]
22. Hipp, A.L.; Manos, P.S.; González-Rodríguez, A.; Hahn, M.; Kaproth, M.; McVay, J.D.; Avalos, S.V.; Cavender-Bares, J. Sympatric Parallel Diversification of Major Oak Clades in the Americas and the Origins of Mexican Species Diversity. *New Phytol.* **2017**, *217*, 439–452. [[CrossRef](#)] [[PubMed](#)]
23. Cavender-Bares, J. Diversification, Adaptation, and Community Assembly of the American Oaks (*Quercus*), a Model Clade for Integrating Ecology and Evolution. *New Phytol.* **2018**, *221*, 669–692. [[CrossRef](#)] [[PubMed](#)]
24. Poulos, H.M. A Review of the Evidence for Pine–Oak Niche Differentiation in the American Southwest. *J. Sustain. For.* **2009**, *28*, 92–107. [[CrossRef](#)]
25. Felger, R.S.; Johnson, M.B. Biodiversity and Management of the Madrean Archipelago: The Sky Islands of Southwestern United States and Northwestern Mexico. In *Biodiversity and Management of the Madrean Archipelago: The Sky Islands of Southwestern United States and northern Mexico*; The Station: Fort Collins, CO, USA, 1995; Volume 264, pp. 71–77.
26. Dhillon, S.S.; Mills, M.H. The sand shinnery oak (*Quercus havardii*) communities of the Llano Estacado: History, structure, ecology, and restoration. In *Savannas, Barrens, and Rock Outcrop Plant Communities of North America*; Cambridge University Press: Cambridge, UK, 1999; pp. 262–274.
27. Peterson, R.; Boyd, C.S. *Ecology and Management of Sand Shinnery Communities: A Literature Review*; U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station: Fort Collins, CO, USA, 2000.
28. Dhillon, S.S. Environmental Heterogeneity, Animal Disturbances, Microsite Characteristics, and Seedling Establishment in a *Quercus havardii* Community. *Restor. Ecol.* **1999**, *7*, 399–406. [[CrossRef](#)]
29. Davis, W.J. *Shin-Oak (Quercus havardii, Rydb.; Fagaceae) Rhizome Shoot Production: Possibilities for Use in Restoration*. M.S.; Texas Tech University: Lubbock, TX, USA, 2013.
30. Jerome, D.; Beckman, E.; Kenny, L.; Wenzell, K.; Kua, C.; Westwood, M. *The Red List of US Oaks*; The Morton Arboretum: Lisle, IL, USA, 2017.
31. Tucker, J.M. Studies in the *Quercus undulata* Complex. IV. The Contribution of *Quercus havardii*. *Am. J. Bot.* **1970**, *57*, 71–84. [[CrossRef](#)]
32. McCauley, R.A.; Christie, B.J.; Ireland, E.L.; Landers, R.A.; Nichols, H.R.; Schendel, M.T. Influence of Relictual Species on the Morphology of a Hybridizing Oak Complex: An Analysis of the *Quercus* x *undulata* Complex in the Four Corners Region. *West. N. Am. Nat.* **2012**, *72*, 296–310. [[CrossRef](#)]

33. Jarčuška, B.; Milla, R. Shoot-Level Biomass Allocation Is Affected by Shoot Type in *Fagus sylvatica*. *J. Plant Ecol.* **2012**, *5*, 422–428. [[CrossRef](#)]
34. Cicák, A. Estimation of Morphological Parameters for European Beech Leaves on Spring Shoots Using the Method of Calculation Coefficients. *Oecol* **2003**, *30*, 131–140.
35. Viscosi, V.; Lepais, O.; Gerber, S.; Fortini, P. Leaf Morphological Analyses in Four European Oak Species (*Quercus*) and Their Hybrids: A Comparison of Traditional and Geometric Morphometric Methods. *Plant Biosyst.* **2009**, *143*, 564–574. [[CrossRef](#)]
36. Albarrán-Lara, A.L.; Mendoza-Cuenca, L.; Valencia-Avalos, S.; González-Rodríguez, A.; Oyama, K. Leaf Fluctuating Asymmetry Increases with Hybridization and Introgression between *Quercus magnoliifolia* and *Quercus resinosa* (Fagaceae) through an Altitudinal Gradient in Mexico. *Int. J. Plant Sci.* **2010**, *171*, 310–322. [[CrossRef](#)]
37. Holmgren, P.K.; Holmgren, N.H. Index Herbariorum. *Taxon* **1991**, *40*, 687–692. [[CrossRef](#)]
38. Holleley, C.E.; Geerts, P.G. Multiplex Manager 1.0: A Cross-Platform Computer Program That Plans and Optimizes Multiplex PCR. *BioTechniques* **2009**, *46*, 511–517. [[CrossRef](#)]
39. Dzialuk, A.; Chybicki, I.; Burczyk, J. PCR Multiplexing of Nuclear Microsatellite Loci in *Quercus* Species. *Plant Mol. Biol. Rep.* **2005**, *23*, 121–128. [[CrossRef](#)]
40. Abraham, S.T.; Zaya, D.N.; Koenig, W.D.; Ashley, M.V. Interspecific and Intraspecific Pollination Patterns of Valley Oak, *Quercus lobata*, in a Mixed Stand in Coastal Central California. *Int. J. Plant Sci.* **2011**, *172*, 691–699. [[CrossRef](#)]
41. Backs, J.R.; Terry, M.; Ashley, M.V. Using Genetic Analysis to Evaluate Hybridization as a Conservation Concern for the Threatened Species *Quercus hinckleyi* C.H. Muller (Fagaceae). *Int. J. Plant Sci.* **2016**, *177*, 121–131. [[CrossRef](#)]
42. Craft, K.J.; Ashley, M.V. Landscape Genetic Structure of Bur Oak (*Quercus macrocarpa*) Savannas in Illinois. *For. Ecol. Manag.* **2007**, *239*, 13–20. [[CrossRef](#)]
43. Oosterhout, C.V.; Hutchinson, W.F.; Wills, D.P.; Shipley, P. Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. *Mol. Ecol. Notes* **2004**, *4*, 535–538. [[CrossRef](#)]
44. Chybicki, I.J.; Burczyk, J. Simultaneous Estimation of Null Alleles and Inbreeding Coefficients. *J. Hered.* **2008**, *100*, 106–113. [[CrossRef](#)] [[PubMed](#)]
45. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.
46. Kamvar, Z.N.; Tabima, J.F.; Grünwald, N.J. Poppr: An R Package for Genetic Analysis of Populations with Clonal, Partially Clonal, and/or Sexual Reproduction. *PeerJ* **2014**, *2*, e281. [[CrossRef](#)]
47. Jombart, T. Adegenet: A R Package for the Multivariate Analysis of Genetic Markers. *Bioinformatics* **2008**, *24*, 1403–1405. [[CrossRef](#)]
48. Goudet, J. Hierstat, a Package for R to Compute and Test Hierarchical F-Statistics. *Mol. Ecol. Notes* **2005**, *5*, 184–186. [[CrossRef](#)]
49. Kraemer, P.; Gerlach, G. Demerelate: Calculating Interindividual Relatedness for Kinship Analysis Based on Codominant Diploid Genetic Markers Using R. *Mol. Ecol. Resour.* **2017**, *17*, 1371–1377. [[CrossRef](#)] [[PubMed](#)]
50. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research—an Update. *Bioinformatics* **2012**, *28*, 2537–2539. [[CrossRef](#)] [[PubMed](#)]
51. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **2000**, *155*, 945–959. [[PubMed](#)]
52. Earl, D.A.; vonHoldt, B.M. STRUCTURE HARVESTER: A Website and Program for Visualizing STRUCTURE Output and Implementing the Evanno Method. *Conserv. Genet. Resour.* **2011**, *4*, 359–361. [[CrossRef](#)]
53. Rosenberg, N.A. Distruct: A Program for the Graphical Display of Population Structure. *Mol. Ecol. Notes* **2003**, *4*, 137–138. [[CrossRef](#)]
54. Jombart, T.; Devillard, S.; Balloux, F. Discriminant Analysis of Principal Components: A New Method for the Analysis of Genetically Structured Populations. *BMC Genet.* **2010**, *11*, 94. [[CrossRef](#)]
55. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the Number of Clusters of Individuals Using the Software Structure: A Simulation Study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [[CrossRef](#)]
56. Uslu, E.; Bakış, Y. Morphometric Analyses of the Leaf Variation within *Quercus* L. Sect. *Cerris* Loudon in Turkey. *Dendrobiology* **2013**, 109–117. [[CrossRef](#)]
57. Jensen, J.S. Variation of Growth in Danish Provenance Trials with Oak (*Quercus robur* L. and *Quercus petraea* Mattuschka Liebl.). *Ann. Des Sci. For.* **1993**, *50*, 203s–207s. [[CrossRef](#)]
58. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J.R.G.; Gruber, B.; Lafourcade, B.; Leitão, P.J.; et al. Collinearity: A Review of Methods to Deal with It and a Simulation Study Evaluating Their Performance. *Ecography* **2012**, *36*, 27–46. [[CrossRef](#)]
59. Chatterjee, S.; Hadi, A.S.; Price, B. Simple Linear Regression. In *Regression Analysis by Example*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2006; pp. 21–51.
60. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016; ISBN 978-3-319-24277-4.
61. Pinheiro, J.; Bates, D.; DebRoy, S.; Sarkar, D.; Team, R.C. Linear and Nonlinear Mixed Effects Models. *R package version* **2007**, *3*, 1–89.
62. Nakagawa, S.; Schielzeth, H. A General and Simple Method for Obtaining R² from Generalized Linear Mixed-Effects Models. *Methods Ecol. Evol.* **2012**, *4*, 133–142. [[CrossRef](#)]

63. Barton, K. MuMIn: Multi-Model Inference. R Package Version 1. 0. 0. 2009. Available online: <http://r-forge.r-project.org/projects/mumin/> (accessed on 13 March 2021).
64. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [[CrossRef](#)]
65. Broennimann, O.; Fitzpatrick, M.C.; Pearman, P.B.; Petitpierre, B.; Pellissier, L.; Yoccoz, N.G.; Thuiller, W.; Fortin, M.J.; Randin, C.; Zimmermann, N.E.; et al. Measuring ecological niche overlap from occurrence and spatial environmental data. *Glob. Ecol. Biogeogr.* **2012**, *21*, 481–497. [[CrossRef](#)]
66. Schoener, T.W. Nonsynchronous Spatial Overlap of Lizards in Patchy Habitats. *Ecology* **1970**, *51*, 408–418. [[CrossRef](#)]
67. Lumley, T. Based on R Code by A. M. Leaps: Regression Subset Selection, R Package Version 3.0. Available online: <https://cran.r-project.org/package=leaps> (accessed on 10 March 2021).
68. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc.* **1995**, *57*, 289–300. [[CrossRef](#)]
69. Yang, J.-Q.; Hsu, K.-C.; Liu, Z.-Z.; Su, L.-W.; Kuo, P.-H.; Tang, W.-Q.; Zhou, Z.-C.; Liu, D.; Bao, B.-L.; Lin, H.-D. The Population History of *Garra orientalis* (Teleostei: Cyprinidae) Using Mitochondrial DNA and Microsatellite Data with Approximate Bayesian Computation. *BMC Evol. Biol.* **2016**, *16*. [[CrossRef](#)]
70. Zhang, X.-W.; Li, Y.; Zhang, Q.; Fang, Y.-M. Ancient East-West Divergence, Recent Admixture, and Multiple Marginal Refugia Shape Genetic Structure of a Widespread Oak Species (*Quercus acutissima*) in China. *Tree Genet. Genomes* **2018**, *14*, 88. [[CrossRef](#)]
71. Feng, L.; Zhang, Y.-P.; Chen, X.-D.; Yang, J.; Zhou, T.; Bai, G.-Q.; Yang, J.; Li, Z.-H.; Peng, C.-I.; Zhao, G.-F. Allopatric Divergence, Local Adaptation, and Multiple Quaternary Refugia in a Long-Lived Tree (*Quercus spinosa*) from Subtropical China. *BioRxiv* **2017**. [[CrossRef](#)]
72. Navascués, M. MicrosatABC-IM: Approximate Bayesian Computation Inferences under the Isolation with Migration Model from Microsatellite Data. Available online: <https://hal.inrae.fr/hal-02789612> (accessed on 20 February 2021). [[CrossRef](#)]
73. Hudson, R.R. Generating Samples under a Wright-Fisher Neutral Model of Genetic Variation. *Bioinformatics* **2002**, *18*, 337–338. [[CrossRef](#)]
74. Raynal, L.; Marin, J.-M.; Pudlo, P.; Ribatet, M.; Robert, C.P.; Estoup, A. ABC Random Forests for Bayesian Parameter Inference. *Bioinformatics* **2018**, *35*, 1720–1728. [[CrossRef](#)]
75. Pudlo, P.; Marin, J.-M.; Estoup, A.; Cornuet, J.-M.; Gautier, M.; Robert, C.P. Reliable ABC Model Choice via Random Forests. *Bioinformatics* **2016**, *32*, 859–866. [[CrossRef](#)]
76. Thuillet, A.-C.; Bru, D.; David, J.; Roumet, P.; Santoni, S.; Sourdille, P.; Bataillon, T. Direct Estimation of Mutation Rate for 10 Microsatellite Loci in Durum Wheat, *Triticum turgidum* (L.) Thell. ssp. *Durum* Desf. *Mol. Biol. Evol.* **2002**, *19*, 122–125. [[CrossRef](#)]
77. Vigouroux, Y.; Jaqueth, J.S.; Matsuoka, Y.; Smith, O.S.; Beavis, W.D.; Smith, J.S.C.; Doebley, J. Rate and Pattern of Mutation at Microsatellite Loci in Maize. *Mol. Biol. Evol.* **2002**, *19*, 1251–1260. [[CrossRef](#)]
78. Marriage, T.N.; Hudman, S.; Mort, M.E.; Orive, M.E.; Shaw, R.G.; Kelly, J.K. Direct Estimation of the Mutation Rate at Dinucleotide Microsatellite Loci in *Arabidopsis thaliana* (Brassicaceae). *Heredity* **2009**, *103*, 310–317. [[CrossRef](#)]
79. O’Connell, L.M. Somatic Mutations at Microsatellite Loci in Western Red Cedar (*Thuja plicata*: Cupressaceae). *J. Hered.* **2004**, *95*, 172–176. [[CrossRef](#)] [[PubMed](#)]
80. Kass, R.E.; Raftery, A.E. Bayes Factors. *J. Am. Stat. Assoc.* **1995**, *90*, 773–795. [[CrossRef](#)]
81. Jarosz, A.F.; Wiley, J. What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. *J. Probl. Solving* **2014**, *7*, 2. [[CrossRef](#)]
82. Ashley, M.V.; Backs, J.R.; Kindsvater, L.; Abraham, S.T. Genetic Variation and Structure in an Endemic Island Oak, *Quercus tomentella*, and Mainland Canyon Oak, *Quercus chrysolepis*. *Int. J. Plant Sci.* **2018**, *179*, 151–161. [[CrossRef](#)]
83. Backs, J.R.; Terry, M.; Klein, M.; Ashley, M.V. Genetic Analysis of a Rare Isolated Species: A Tough Little West Texas Oak, *Quercus hinckleyi* C.H. Mull. *J. Torrey Bot. Soc.* **2015**, *142*, 302–313. [[CrossRef](#)]
84. Gonzalez-Rodriguez, A.; Oyama, K. Leaf Morphometric Variation in *Quercus affinis* and *Q. laurina* (Fagaceae), Two Hybridizing Mexican Red Oaks. *Bot. J. Linn. Soc.* **2005**, *147*, 427–435. [[CrossRef](#)]
85. Kusi, J.; Karsai, I. Plastic Leaf Morphology in Three Species of *Quercus*: The More Exposed Leaves Are Smaller, More Lobated and Denser. *Plant Species Biol.* **2019**, *35*, 24–37. [[CrossRef](#)]
86. Trejo, L.; Alvarado-Cárdenas, L.O.; Scheinvar, E.; Eguiarte, L.E. Population Genetic Analysis and Bioclimatic Modeling in *Agave striata* in the Chihuahuan Desert Indicate Higher Genetic Variation and Lower Differentiation in Drier and More Variable Environments. *Am. J. Bot.* **2016**, *103*, 1020–1029. [[CrossRef](#)]
87. Sork, V.L.; Davis, F.W.; Westfall, R.; Flint, A.; Ikegami, M.; Wang, H.; Grivet, D. Gene Movement and Genetic Association with Regional Climate Gradients in California Valley Oak (*Quercus lobata* Née) in the Face of Climate Change. *Mol. Ecol.* **2010**, *19*, 3806–3823. [[CrossRef](#)]
88. Riordan, E.C.; Gugger, P.F.; Ortego, J.; Smith, C.; Gaddis, K.; Thompson, P.; Sork, V.L. Association of Genetic and Phenotypic Variability with Geography and Climate in Three Southern California Oaks. *Am. J. Bot.* **2016**, *103*, 73–85. [[CrossRef](#)]
89. Yang, J.; Vázquez, L.; Feng, L.; Liu, Z.; Zhao, G. Climatic and Soil Factors Shape the Demographical History and Genetic Diversity of a Deciduous Oak (*Quercus liaotungensis*) in Northern China. *Front. Plant Sci.* **2018**, *9*, 9. [[CrossRef](#)]
90. Ortego, J.; Gugger, P.F.; Sork, V.L. Climatically Stable Landscapes Predict Patterns of Genetic Structure and Admixture in the Californian Canyon Live Oak. *J. Biogeogr.* **2014**, *42*, 328–338. [[CrossRef](#)]

91. Mayes, S.G.; McGinley, M.A.; Werth, C.R. Clonal Population Structure and Genetic Variation in Sand-Shinnery Oak, *Quercus havardii* (Fagaceae). *Am. J. Bot.* **1998**, *85*, 1609–1617. [[CrossRef](#)] [[PubMed](#)]
92. Hipp, A.L.; Eaton, D.A.R.; Cavender-Bares, J.; Fitzek, E.; Nipper, R.; Manos, P.S. A Framework Phylogeny of the American Oak Clade Based on Sequenced RAD Data. *PLoS ONE* **2014**, *9*, e93975. [[CrossRef](#)]
93. Hampe, A.; Petit, R.J. Conserving Biodiversity under Climate Change: The Rear Edge Matters. *Ecol. Lett.* **2005**, *8*, 461–467. [[CrossRef](#)]
94. Muller, C.H. Ecological Control of Hybridization in *Quercus*: A Factor in the Mechanism of Evolution. *Evolution* **1952**, *6*, 147. [[CrossRef](#)]
95. Hipp, A.L.; Manos, P.S.; Hahn, M.; Avishai, M.; Bodénès, C.; Cavender-Bares, J.; Crawl, A.A.; Deng, M.; Denk, T.; Fitz-Gibbon, S.; et al. Genomic Landscape of the Global Oak Phylogeny. *New Phytol.* **2019**, *226*, 1198–1212. [[CrossRef](#)]
96. Lepais, O.; Gerber, S. Reproductive Patterns Shape Introgression Dynamics and Species Succession within the European White Oak Species Complex. *Evolution* **2010**, *65*, 156–170. [[CrossRef](#)]
97. Spence, E.S.; Fant, J.; Gailing, O.; Griffith, M.P.; Havens, K.; Hipp, A.L.; Kadav, P.; Kramer, A.; Thompson, P.; Toppila, R.; et al. Comparing genetic diversity in three threatened oaks. *Forests* **2021**, *12*.
98. Chatwin, W.B.; Carpenter, K.K.; Jimenez, F.R.; Elzinga, D.B.; Johnson, L.A.; Maughan, P.J. Microsatellite Primer Development for Post Oak, *Quercus stellata* (Fagaceae). *Appl. Plant Sci.* **2014**, *2*, 1400070. [[CrossRef](#)] [[PubMed](#)]
99. Dow, B.D.; Ashley, M.V.; Howe, H.F. Characterization of Highly Variable (GA/CT) n Microsatellites in the Bur Oak, *Quercus macrocarpa*. *Theor. Appl. Genet.* **1995**, *91*, 137–141. [[CrossRef](#)] [[PubMed](#)]
100. Kampfer, S.; Lexer, C.; Glössl, J.; Steinkellner, H. Characterization of (GA) n Microsatellite Loci from *Quercus robur*. *Hereditas* **2004**, *129*, 183–186. [[CrossRef](#)]
101. Steinkellner, H.; Lexer, C.; Turetschek, E.; Glössl, J. Conservation of (GA) n Microsatellite Loci between *Quercus* Species. *Mol. Ecol.* **1997**, *6*, 1189–1194. [[CrossRef](#)]
102. Isagi, Y.; Suhandono, S. PCR Primers Amplifying Microsatellite Loci of *Quercus myrsinifolia* Blume and Their Conservation between Oak Species. *Mol. Ecol.* **1997**, *6*, 897–899. [[CrossRef](#)] [[PubMed](#)]