Evolutionary history and gene flow of an endemic island oak:

Quercus pacifica K. Nixon & C.H. Muller

Janet Rizner Backs and Mary V. Ashley

University of Illinois at Chicago, UIC Biological Sciences, M/C 066, 845 W. Taylor, Chicago, Illinois, 60607, United States of America
The authors thank S. Sillett, S. Ratay, J. Yoon, M. Desrosiers, M. Pesendorfer, M. Witter, T. Sagar, T. Valois, C. Anderson, and K. Lombardo for their help in collecting samples for this study. We also wish to acknowledge the reviewers whose comments and suggestions have been helpful in expanding the scope of this paper. Funding was through the University of Illinois at Chicago Hadley Grant. This manuscript was completed in partial fulfillment of the doctoral degree from the Graduate College at the University of Illinois at Chicago to JRB.

Author for correspondence (email: jbacks@uic.edu; alternate email: janet.rizner.backs@gmail.com)
ABSTRACT

Premise of the study: Understanding historical patterns of colonization and subsequent gene flow clarifies the evolutionary origins and history of endemic island species.

Methods: Here we use DNA microsatellite markers to characterize the genetic structure of the island endemic species *Quercus pacifica* K. Nixon & C.H. Muller, found on three of the California Channel Islands, and to examine its relationship to two mainland oaks, *Q. berberidifolia* and *Q. dumosa*.

Key results: We found that *Q. pacifica* is a genetically cohesive and differentiated evolutionary lineage, diverging from mainland scrub oaks in the Pleistocene with little subsequent introgression. Genetic differentiation of *Q. pacifica* among islands is small but significant. Both recent and historical gene flow was surprisingly high considering the disjunct distribution of *Q. pacifica* on islands separated by as much as 125 km of open ocean. Gene flow estimates were highest between the two northern islands and from the northern islands to Santa Catalina. While there is no evidence of recent bottlenecks, historical bottlenecks are indicated on each of the islands.

Conclusions: The genetic cohesiveness of the *Q. pacifica* species suggests allopatric speciation on the islands with subsequent gene flow that has maintained genetic continuity over great distances.

Key words: California scrub oaks; gene flow; island endemic; genetic structure; insular genetic isolation; microsatellites; *Quercus pacifica*. 
Pollin-mediated gene flow in trees, particularly wind-pollinated trees, has repeatedly been shown to be extensive and extend long distances, with the consequence that populations of trees typically show high connectivity and low levels of genetic differentiation (Hamrick and Godt, 1996; Petit and Hampe, 2006; Ashley, 2010). For the genus *Quercus*, even small, isolated populations are reported to receive substantial immigrant pollen (Muir et al., 2004; Craft and Ashley, 2007; Ortego et al., 2014). Studies to date, however, have looked at peripheral or fragmented populations of widespread oak species rather than species with a continuous history of geographic isolation. Here we examine population structure and gene flow in *Quercus pacifica* (island scrub oak), a species with a geographical range restricted to three oceanic islands. *Quercus pacifica* is endemic to the California Channel Islands, occurring on two of the northern islands, Santa Rosa and Santa Cruz, and the southern island of Santa Catalina. It is one of about 100 endemic plant species on the California Channel Islands, approximately 20 of which, like *Q. pacifica*, are shared between the Northern (Anacapa, San Miguel, Santa Rosa, and Santa Cruz) and Southern (San Clemente, San Nicolas, Santa Barbara, and Santa Catalina) Island groups (Fig. 1) (Raven, 1967; Philbrick, 1980; Moody, 2001). The taxonomic status of *Q. pacifica*, as an island endemic occurred fairly recently. Prior to 1994, *Q. pacifica* was considered part of the ‘*Q. dumosa complex’*, together with mainland scrub oak species including *Q. berberidifolia* and *Q. dumosa*, and was classified as *Quercus dumosa* Nuttall var. *polycarpa* Greene. At that time, based on morphology, ecology and distribution, the scrub oak growing on the Channel Islands was renamed *Q.*

The evolutionary history of *Q. pacifica* must be considered in light of the geologic and settlement history of the islands and of oak reproductive biology. Since their emergence, the Channel Islands (Fig. 1) have been sculpted by continued uplift, volcanism, varying sea levels due to changes in climate, soil formation and plant settlement, and erosion events caused by herbivory. The uplift of the Channel Islands began about five million years ago (Atwater, 1998; Schumann et al., 2012). Evidence of geological and geophysical activity over the last 2.6 million years indicates that there were no land bridges between the islands and the mainland, nor between the northern and southern islands (Junger and Johnson, 1980; Vedder and Howell, 1980). Thus colonization of the islands must have occurred through overwater dispersal of propagules; acorns in the case of oaks such as *Q. pacifica*. Glaciation events resulted in changes in sea level, and as recently as the Last Glacial Maximum (LGM) Santa Rosa and Santa Cruz were joined with Anacapa and San Miguel islands into a larger landmass, known as Santarosae, that was within 9 km of the mainland. In this period Santa Catalina was approximately 1.3 times its current area and was within 24 km of the mainland while remaining over 100 km from Santarosae. Therefore any gene flow occurring between *Q. pacifica* populations on Santa Catalina and the northern islands populations during this period required movement of pollen or acorns across large expanses of open ocean, whereas populations on the northern island were
continuous. Furthermore, both northern and southern island populations of *Quercus pacifica* have always been more proximate to mainland scrub oak populations than to each other.

While most plant species were present on the islands well before humans arrived (Rick et al., 2014; Ortego et al., 2015), settlers may have played a role in gene flow of *Quercus* species and certainly were responsible for habitat transformation on the islands. Native Americans and more recently non-native settlers have continuously occupied the California Channel Islands for over 10,000 years (Erlandson et al., 2008). The Chumash and the Gabrielino, sophisticated maritime cultures, made use of the resources of the islands for millennia (Collins, 1991). Marine travel followed established routes and facilitated exchange of goods, including acorns (Arnold, 1992). Humans likely affected the natural plant communities before European settlement, possibly transplanting oaks as well as other organisms (Erlandson et al., 2004). While acorns, including scrub oak acorns (Bainbridge, 1987), were a primary food source for mainland Chumash (Anderson, 2005), archaeological excavations and isotopic evidence suggest that acorns may not have been a major food or trade item among the island Chumash (Fauvelle, 2013). However acorn exchange during rituals such as mourning ceremonies may have occurred (Hollimon, 2001).

Europeans settlement of the islands began in the mid-19th century, accompanied by the introduction of non-native herbivores to Santa Rosa, Santa Cruz, and Santa Catalina. Over the past 200 years, introduced sheep, cattle, pigs, goats, mule deer and bison have led to over-grazing and dramatic losses of shrub land and woodland habitats on the islands (Westman, 1983; Knowlton et al., 2007; Knapp, 2010; Knapp, 2014). Recently, however,
efforts to restore native ecosystems through removal of non-indigenous herbivores have been successfully undertaken, resulting in some recovery of native communities (Klinger et al., 2003; Beltran et al., 2014).

*Quercus pacifica*, like all oaks, is monoecious and wind-pollinated, and self-pollination is rare (Yacine and Bouras, 1997; Buiteveld et al., 2001). Both maternal seed and paternal pollen play a part in gene flow. Acorns can be transported by rodents (Miyaki and Kikuzawa, 1988; Iida, 1996; Li and Zhang, 2003), birds (Darley-Hill and Johnson, 1981; Gómez, 2003; Scofield et al., 2010), humans (Toumi and Lumaret, 1998; Anderson, 2005; Abrams and Nowacki, 2008) or mechanical means as when gravity moves them down slopes. As mentioned above, long distance gene flow through pollen is common in the genus *Quercus* (Dow and Ashley, 1998; Streiff et al., 1999; Koenig and Ashley, 2003; Dutech et al., 2005; Nakanishi et al., 2008; Ashley et al., 2010; Chybicki and Burczyk, 2010), even at distances of more than 80 km (Buschbom et al., 2011). In addition to sexual reproduction, some oak species may also reproduce clonally, so that stands of trees which may appear to be multiple individuals have only one or a few unique genotypes (Ashley et al., 2010; Backs et al., 2015).

The overall objective of this research was to examine the evolutionary history of *Q. pacifica* by evaluating genetic variation and genetic differentiation and inferring historical patterns of colonization and subsequent gene flow among islands. Specifically, we addressed three questions: 1) Does *Q. pacifica* on all three islands comprise a genetically cohesive species, differentiated from its mainland relatives, specifically its closest relative *Q. dumosa* and the
widespread scrub oak *Q. berberidifolia*? 2) Are the disjunct island populations of *Q. pacifica* genetically distinct from each other, or rather connected by ongoing or historical gene flow? 3) What can we infer about the demographic history and speciation of this island endemic? With regard to the last question, we were interested in evidence for population bottlenecks, possibly associated with colonization events, and inferring the evolutionary history of colonization and gene flow in *Q. pacifica*, both among the different islands and relative to mainland scrub oaks.

**MATERIALS AND METHODS**

**Study Species**— *Q. pacifica*—A key component of the scrub oak chaparral, it is found at elevations between 50 – 150 m on ridges and open slopes as well as in canyons and can cover relatively large areas. The species occurs in both a shrub and tree form up to 5 m in height. As an island endemic *Q. pacifica* is a conservation concern and is listed as ‘vulnerable’ in the Red List of Oaks (Oldfield and Eastwood, 2007). Evolving for some time without native herbivores other than insects, it exhibits significantly reduced morphological defenses as exemplified by fewer and shorter spines on leaves compared with those of its closely related mainland counterparts (Bowen and Vuren, 1997). Current threats to the species include episodic oak dieback, root rot fungus, and herbivory by insects and non-native animals (Knapp, 2002). *Quercus pacifica* has benefitted from conservation efforts to rid the islands of non-native herbivores, monitor oak health (Knapp, 2002), and restore the chaparral ecosystem (Stratton, 2001).

*Q. berberidifolia* and *Q. dumosa*— Revisions of taxonomic designations of California oaks have been ongoing for over a century (Fryer, 2012). During this time the name ‘Quercus
*dumosa*’ was applied to a number of shrub oaks and the ‘*Quercus dumosa complex*’ included at least five species that are now recognized as separate taxa based on acorn morphology, leaf vestiture, and habitat (Nixon, 2002). *Quercus berberidifolia* (common name: California scrub oak) is widespread in the Central and South Coast ranges of California at elevations of 100-1800 m. It grows on a variety of soils in chaparral plant communities. From Santa Barbara southward, it does not descend to the low elevations where *Q. dumosa* is found. *Quercus dumosa* Nutt. sensu stricto (common name: coastal sage scrub oak) is found only in Southern California and northern Baja California, growing in open chaparral on coastal bluffs and hillsides near the sea, often on loose sandstone or granitics. Its habitat is dwindling due to human encroachment into the desirable real estate locations it occupies. Misclassification of other species as ‘*dumosa*’ may have contributed to a lack of appreciation of the scarcity of this species.

**Sampling**— Leaf samples from a total of 133 *Quercus pacifica* trees were collected across the three islands on which it is found (Appendix S1, see Supplemental Data with the online version of this article); 46 from Santa Rosa, 39 from Santa Cruz, and 48 from Santa Catalina. Of these, 106 were characterized by form: 30 as a shrub and 76 as a tree. Leaf samples were also collected from mainland oaks: 60 *Quercus berberidifolia* from the western Santa Monica Mountains and near Topanga; 24 *Quercus dumosa* from the Point Loma Peninsula and Torrey Pines State Park (Fig. 1). Samples were stored at room temperature with Drierite desiccant (W.A. Hammond Drierite Co., Ltd., Xenia, Ohio, USA) until DNA extraction.

**Microsatellite genotyping**— Approximately 20 mg of dry leaf material was homogenized to a fine powder using Talboys High Throughput Homogenizer (Henry Troemner LLC, Thorofare,
New Jersey, USA). DNA was extracted with DNAeasy Plant MiniKit (Qiagen, Germantown, Maryland, USA). Nuclear microsatellite loci and primers used in this study were developed and applied successfully with other oaks in the white oak group (Quercus subgenus Quercus) (Dow et al., 1995; Isagi and Suhandono, 1997; Craft and Ashley, 2007; Ashley et al., 2010). Each tree was genotyped at eight neutral microsatellite loci: QpZAG1/5, QpZAG15, and QpZAG110, originally developed for Q. petraea (Steinkellner et al., 1997); QpZAG15 and QpZAG11, developed for Q. robur (Kampfer et al., 1998); MSQ4 and MSQ13, developed for Q. macrocarpa (Dow et al., 1995); and QM69–2M1, developed for Q. myrsinifolia (Isagi and Suhandono, 1997). PCR amplification followed the protocol described previously by Abraham et al. (2011) using labeled forward primers as described by Schuelke (2000). PCR products (0.9 – 1.5 µL) were genotyped on an ABI 3730 DNA Analyzer using GeneScan™ - 500 LIZ® Size Standard (Applied Biosystems, Foster City, California, USA). All genotypes were scored using Applied Biosystems GeneMapper, version 3.7.

Data analyses—Clone identification—Testing for clones was done using ALLELEMATCH version 2.03 (Galpern et al., 2012), an R-package designed to identify unique multilocus genotypes (UMGs) when the number of individuals is not known a priori and where limited genotyping error and missing data may be present. Probability that matches represent unique genotypes is given by P_{sib}, the probability that a sample is a sibling, and not a duplicate, of a UMG.

Genetic Diversity—Following removal of duplicate clones, allele frequency (Na), observed (Ho) and expected (He) heterozygosity, and fixation index (F_{IS}) were calculated with GenAlEx
6.501 (Peakall and Smouse, 2006; Peakall and Smouse, 2012). The R-package DEMEtics (Gerlach et al., 2010) was used to examine population differentiation giving fixation index $D_{\text{Jost}}$ (Jost, 2008). To reduce the possibility of Type I errors caused by multiple pairwise comparisons of a single data set, it uses bootstrapping to calculate Bonferroni corrected $p$-values. This R-package also checks for Hardy-Weinberg Equilibrium (HWE) in the input populations, and uses either alleles or genotypes as appropriate based on HWE for a given locus (Goudet et al., 1996). We calculated allelic richness and private allele richness for each population of *Q. pacifica* HP-RARE 1.0 (Kalinowski, 2004, 2005).

Interspecific differentiation and gene flow—*Q. pacifica, Q. berberidifolia* and *Q. dumosa*—Population structure among these species was analyzed using STRUCTURE (Pritchard et al., 2000; Falush et al., 2003, 2007; Hubisz et al., 2009) and Principal Coordinates Analysis (GenAlEx 6.501). STRUCTURE uses Bayesian clustering algorithms based on genotypes of individual samples to infer population structure. Principal Coordinates Analysis (PCoA) provides a visual representation of patterns resulting from a multivariate analysis of multiple loci and multiple samples.

STRUCTURE was run under the Admixture Model with LOCPRIOR using the three putative species as location identifiers and a 50K burnin and 100K MCMC for potential K of 1 to 7 for 10 reps each. Best K was determined by calculating $\ln(K)$ and $\Delta K$ (Evanno et al., 2005) using STRUCTURE HARVESTER (Earl and vonHoldt, 2012). Output from STRUCTURE HARVESTER was run through CLUMPP (Jakobsson and Rosenberg, 2007) to group individuals into genotypes and those results were input into DISTRUCT (Rosenberg, 2004),
producing a summary image of the runs. Using the best K value, the STRUCTURE project was rerun with 50K burnin and 250K MCMC to determine $q$-values (posterior probability), representing the percentage of individual genotypes in each K cluster.

To determine the relative importance of genetic drift versus mutation in genetic differentiation among the species, we compared $R_{ST}$ and $F_{ST}$. $R_{ST}$ is based on allele sizes and is expected to be larger than $F_{ST}$ if mutations have contributed to genetic differentiation, assuming that the mutation process follows at least partially a SMM (stepwise mutation model) (Hardy et al., 2003). We used SPAGeDi 1.5a (Hardy and Vekemans, 2002) to compute $F_{ST}$ and $R_{ST}$ values and to analyze $R_{ST}$ vs. $F_{ST}$ using Hardy’s permutation of allele sizes.

*Quercus pacifica*—Population structure and gene flow—Population structure within and across islands was analyzed using three tools: STRUCTURE, GENELAND (Guillot et al., 2005b; Guillot et al., 2005a; Guillot, 2008; Guillot et al., 2008), and Principal Coordinates Analysis (GenAlEx 6.501). Like STRUCTURE, GENELAND uses Bayesian clustering algorithms based on genotypes of individual samples, but it also includes geographic locations of samples to infer population structure and potential physical barriers to gene flow, even in populations with low genetic differentiation (Blair et al., 2012; Vernesi et al., 2012; Rieux et al., 2013; Bech et al., 2014).

For STRUCTURE analysis, the procedure was similar to that described above for interspecific comparisons, with the exception that the Admixture Model with NOLOCPRIOR was used. With GENELAND, the GUI interface was used to run the correlated frequency, spatial model at 100K MCMC, thin rate 100, and burnin 200 for 10 runs at K of 1 to 10. This
procedure was repeated five times. For each repetition, the run with the highest posterior probability was chosen, and from these five results best K was inferred from the modal value of K with the highest posterior probability.

To examine the role of mutation as a factor in the genetic differentiation among the islands, we also compared $R_{st}$ and $F_{st}$ as described above.

To assess patterns of recent migration (within 2-3 generations), we used BayesAss v. 3.0 (Wilson and Rannala, 2003). BayesAss relaxes some assumptions on which long-term migration estimators are based, such as deviations from HWE in populations. We ran $10^7$ MCMC iterations, $10^6$ burnin iterations, sampling frequency of 100, and 0.30 acceptance adjustments for allele frequency and inbreeding coefficients. Convergence was verified by repeating the run three times with different initial seed numbers and examination of trace output using the Tracer program (Rambaut and Drummond, 2007). We used 0.099 as the cut-off value to distinguish true migration from background noise (López-Vinyallonga et al., 2015).

We used Migrate-n version 3.6.11 (Beerli and Felsenstein, 1999, 2001) to investigate long-term effective population sizes ($\Theta$) and migration rates ($Nm$) among the three islands. Migrate-n uses coalescence theory to model population sizes and migration rates, and mutation models to explain change of alleles at sites over time. We ran in Bayesian mode using the infinite allele option, which is recommended when the mutation model is unknown. Uniform prior for $\Theta$ was set to min: 0.0, max: 100.0, delta: 10.0. Uniform prior for migration was set to 0.0, max: 1000.0, delta: 100.0. Posterior distributions were generated using the Metropolis-Hasting algorithm. Parameter starting values were based on $F_{st}$ estimates. We used an UPGMA starting tree and
static heating set to 4 chains with temperatures 1.00, 1.50, 3.00, 10,000,000.0. Search strategies were set to 1 long chain with 10,000 recorded genealogies, a sampling increment of 100 and a burnin of 10,000. Output from five runs with different random seeds was compared to ensure convergence. A final run of a longer chain was used to report results as suggested by Beerli (2009). We set this final search strategy to 500,000 genealogies, a sampling increment of 100, and a burnin of 100,000.

Excess heterozygosity may indicate recent bottleneck events (Cornuet and Luikart, 1996). We used values for observed and expected heterozygosity for each locus as computed in GenAlEx 6.501 (Table 1) and conducted a one-tailed Wilcoxon signed-rank test with significance level of 0.01 to check for significant differences between these two values for the eight loci in the populations on each of the three islands.

We also examined possible bottlenecks using M-ratio (Garza and Williamson, 2001; Excoffier et al., 2005), where $M = k/r$ with $k =$ total number of alleles and $r =$ overall range in allele size. M-ratio is less sensitive to effects of null alleles or Wahlund effects than heterozygosity comparisons and retains indications of a bottleneck over longer periods of time (Garza and Williamson, 2001; Spear et al., 2006; Peery et al., 2012). We used Arlequin v3.5.1.2 (Excoffier and Lischer, 2010) to determine the M-ratio (Garza-Williamson index). A ratio below 0.68 is considered to be indicative of a bottleneck (Hoban et al., 2010; Ortego et al., 2010; Pollegioni et al., 2011; Martin et al., 2014; Duncan et al., 2016).
RESULTS

**Clones** — ALLELEMATCH (psib < .001) showed that the original set of 60 *Q. berberidifolia* samples contained 43 UMGs. No clones were found among the 24 *Q. dumosa* and 133 *Q. pacifica* samples. Clones were collapsed for *Q. berberidifolia* and from that point only unique genotypes were used for analysis.

**Genetic Diversity** — All loci were polymorphic and the level of genetic diversity is comparable to values found for mainland California oaks (Craft et al., 2002; Dutech et al., 2005; Abraham et al., 2011) as shown in Table 1.

*Q. pacifica* — Mean Na (number of alleles) per locus across the islands varies from 15.3 to 16.8, mean Ho (observed heterozygosity) from 0.795 to 0.828, and mean He (expected heterozygosity) from 0.832 to 0.862 (Table 1). The fixation index (FIS) was low for each island and no significant deviations from HWE were found.

*Q. berberidifolia and Q. dumosa* — For *Q. berberidifolia*, Na (number of alleles) varies from 14 to 24 per locus, Ho (observed heterozygosity) varies from 0.558 to 0.977, and He (expected heterozygosity) from 0.812 to 0.928; for *Q. dumosa*, Na varies from 5 to 16 per locus, Ho from 0.542 to 0.917, and He from 0.597 to 0.903 (Table 1). Although some loci show homozygous excess, others show heterozygous excess; overall, there is no consistent pattern across species or loci. Mean FIS (fixation index) for *Q. berberidifolia* is 0.039 and for *Q. dumosa* is 0.029. These low values suggest an absence of inbreeding.

**Interspecific differentiation** — Comparing *Q. pacifica* to the two mainland species, the species all show significant levels of genetic differentiation (Table 2). The largest distance is
between *Q. berberidifolia* and *Q. dumosa*, with $D_{\text{JOST}} = 0.341$ and $F_{\text{ST}} = 0.053$, while differentiation between *Q. berberidifolia* and *Q. pacifica* ($D_{\text{JOST}} = 0.219$, $F_{\text{ST}} = 0.028$) is less than that between *Q. dumosa* and *Q. pacifica* ($D_{\text{JOST}} = 0.274$, $F_{\text{ST}} = 0.040$). Each species has alleles not found in the other species: *Q. berberidifolia* has 26, *Q. dumosa* has five, and *Q. pacifica* has 49.

Delta K results from STRUCTURE analysis peaked sharply at $K=3$, indicating three distinct genetic clusters which corresponded closely to the three named species (Fig. 2). The proportion of inferred ancestry of each pre-defined population (LOCPRIOR) in each of the three clusters was high: 0.905 for *Q. berberidifolia*, 0.849 for *Q. dumosa*, and 0.938 for *Q. pacifica*.

Among the three species, $R_{\text{st}}$ values are substantially higher than $F_{\text{st}}$ values and allele size permutation tests demonstrate that shifts in average allele sizes contribute significantly to species differentiation. This result suggests these species have been isolated long enough for accumulation of mutations with limited introgression (Hardy et al., 2003), supporting the taxonomic status of *Q. pacifica* as an endemic island species. (see Table 2).

*Quercus pacifica*—Population structure and gene flow—Pairwise differentiation as measured using $D_{\text{JOST}}$ and $F_{\text{ST}}$ for *Q. pacifica* on Santa Rosa, Santa Cruz and Santa Catalina showed low but significant values (Table 2). Genetic differentiation between the shrub and tree forms was also low but significant with $D_{\text{JOST}} = 0.075$, Bonferroni corrected $p = 0.001$, and $F_{\text{ST}} = 0.012$ (GenAlEx 6.501).

STRUCTURE results showed $\Delta K$ peaked sharply at $K=3$, indicating three genetic clusters,
but the clusters did not coincide with the separate islands. Furthermore, each individual included substantial percentages of each genetic cluster (Fig. 3). GENELAND identified five clusters within *Q. pacifica*. Members of the same cluster were found on Santa Rosa, Santa Cruz and Santa Catalina, and two clusters were found on both Santa Cruz and Santa Catalina. Both Santa Rosa and Santa Cruz had clusters found on no other island. Principal Coordinates Analysis (Fig. 4) showed individuals from each island intermixed across the coordinates, with no discernible grouping. PCoA analysis of the two phenotypes indicated no grouping by growth form (data not shown).

Within islands *R*\(_{st}\) values are close to *F*\(_{st}\) values and allele size permutation tests do not reveal a significant contribution of stepwise mutations to population differentiation. While mutations may have spread among the three islands when they occurred, the current low levels of genetic differentiation as shown by *D*\(_{JOST}\) and *F*\(_{st}\) indicates mutation was not sufficient to overcome succeeding levels of gene flow (Hardy et al., 2003) (see Table 2).

The results of the BayesAss analysis indicated detectable recent gene flow from Santa Rosa to Santa Cruz (*m* = 0.3111) and from Santa Rosa to Santa Catalina (*m* = 0.1184). The proportion on non-migrants was highest on Santa Rosa, followed by Santa Catalina and finally Santa Cruz (Table 3).

Although confidence intervals were large and included zero at the lower bound, Migrate-n indicated detectable historical asymmetric gene flow (*Nm*) among all of the islands (Table 4). The highest rates of migration were from Santa Cruz to its neighbor Santa Rosa, which was at one time connected to it, (*Nm* = 4.92), and from Santa Cruz to Santa Catalina (*Nm* = 3.08).
While effective gene flow was less from south to north, there is evidence of historical flow from Santa Catalina to both Santa Rosa ($Nm = 2.42$) and Santa Cruz ($Nm = 1.08$).

The Wilcoxon signed-rank test results showed that the difference between observed and expected heterozygosity for each island is not significant. For Santa Rosa ($W = 9$) the critical value of $W$ is 1; for Santa Cruz ($W = 9$), the critical value of $W$ is 1; and for Santa Catalina ($W = 5$), the critical value of $W$ is 1. If the observed value of $W$ exceeds the critical value, we do not reject the null hypothesis that the median difference between the paired data is zero. Based on these results there is no evidence for recent bottlenecks on any of the islands as would be indicated by excess heterozygosity. In contrast, the $M$-ratio (Garza-Williamson index) was $0.407 \pm 0.117$ for Santa Rosa, $0.479 \pm 0.188$ for Santa Cruz, and $0.470 \pm 0.139$ for Santa Catalina. Each of these values falls below the 0.68 cutoff ratio for a bottleneck, suggesting a bottleneck had occurred on each island at some time in the past.

**DISCUSSION**

With high levels of gene flow and genetic connectivity of populations being characteristic of oak species (Craft and Ashley, 2007; Ashley, 2010; Buschbom et al., 2011; Gerber et al., 2014), the disjunct and isolated range of *Quercus pacifica* provides an interesting system to challenge these patterns. Given that its populations are distributed allopatrically on separate islands, does *Q. pacifica* comprise a single, genetically cohesive species? To answer this, we compared all three island populations to two closely related mainland species, *Q. berberidifolia* and *Q. dumosa*. *Quercus berberidifolia* was chosen because it is the most broadly distributed California scrub oak species and its range includes coastal Southern California. *Quercus dumosa* also occurs (or
did so historically) on the coastal mainland and has recently been reported to be the sister species to *Q. pacifica* (Ortego et al., 2015; Sork, 2016). The ranges of these two mainland species partially overlap and they are sometimes sympatric. Other species of California scrub oaks (*Q. durata, Q. cornelius-mulleri, Q. john-tuckeri*) occur further north or inland. We found *Q. pacifica, Q. berberidifolia* and *Q. dumosa* to be well differentiated from each other as indicated by measures of genetic distance (*D_{JOST}, F_{ST}, R_{ST},* Table 2). Furthermore, Bayesian analysis identified three distinct clusters corresponding to the three species, with high proportions of ancestry in each (Figure 2). Each species harbored private alleles not found in the other species (Table 1) and *R_{ST}* values were significantly greater than *F_{ST}* values. These findings indicate that sufficient time has elapsed since speciation for mutations to accumulate and that introgression rates have not exceeded mutation rates.

Our results lead us to different conclusions than another recent study on California scrub oaks by Ortego et al. (2015). We found the genetic distances between *Q. pacifica* and *Q. berberidifolia* (*D_{JOST}, F_{ST}, R_{ST},* Table 2) were slightly smaller than between *Q. pacifica* and *Q. dumosa*, in contrast to the much closer relationship between *Q. pacifica* and *Q. dumosa* reported by Ortego et al. (2015). These authors suggest a three-way split between *Q. dumosa*, Santa Catalina *Q. pacifica* and Santa Cruz/Santa Rosa *Q. pacifica* that took place around the LGM, with subsequent episodes of admixture. Our study, on the other hand, suggests that *Q. pacifica* diverged from mainland scrub oaks through allopatric speciation, and with little subsequent interspecific introgression between islands and mainland scrub oaks. These differing results may reflect sampling differences (our study included a larger sample size of *Q. pacifica*, 133 vs 53),
markers used (Ortego et al. 2015 used more loci, including chloroplast microsatellites), or analytical approaches which did not overlap. To summarize our conclusions regarding our first question, we found support for the taxonomic status of *Quercus pacifica* as a genetically distinct and cohesive species that has been evolving independently from mainland scrub oaks for some time.

Our second question involved the level of connectivity among the disjunct island populations of *Q. pacifica*, which we examined through genetic clustering approaches and gene flow estimations. While the islands are separated by more than 100 km of open ocean, genetic distance measures (*D*_JOST and *F*_ST, Table 2) indicate only low genetic differentiation among them. We found no evidence that each island formed a distinct genetic cluster. While Bayesian analysis using STRUCTURE found three genetic clusters, these were not associated with specific islands. Spatial analysis using GENELAND found five genetic clusters, but these were shared among individuals on different islands. The PCoA also showed that individuals from all of the islands were intermixed across coordinates.

While recent divergence of the island populations of *Q. pacifica* may in part explain the low levels of genetic differentiation among islands, there is also evidence that long-distance genetic continuity has been maintained within and among the islands through gene exchange. We used several methods to better understand the level and direction of gene flow. Historic gene flow estimates indicate that the islands exchanged one to five individuals per generation (Table 4), and estimates of more recent gene flow suggest a significant portion of individuals on Santa Cruz and Santa Catalina are immigrants (Table 3). These levels of gene flow are somewhat remarkable, especially between Santa Catalina and the two northern islands, given Santa
Catalina’s geographic isolation. How could this occur? Birds act as acorn dispersers (Scofield et al., 2010; Caldwell et al., 2013) and the island scrub-jay (*Aphelocoma insularis*) is a known hoarder of *Q. pacifica* acorns (Pesendorfer et al., 2014; Pesendorfer et al., 2016). Island scrub-jays are currently found only on Santa Cruz but may have been on Santa Rosa in the past and played a role in genetic exchange between those two islands. However, this species is not known from Santa Catalina and generally does not transport acorns over long distances.

Pollen-mitigated gene flow is a more likely factor in both recent gene flow and historical genetic migration, especially between the northern islands and Santa Catalina. Pollen has been found to travel long distances and studies have found a large percentage of pollen comes from outside oak stands (Dow and Ashley, 1998; Ashley et al., 2010; Buschbom et al., 2011). Pollen-mediated gene flow across open water has been reported across the Strait of Gibraltar in a study of pine species (Burban and Petit, 2003) and from Cuba to Central America in a study of Cuban oaks (Gugger and Cavender-Bares, 2013). Prevailing winds are in the northwest to southeast direction, and we did find greater historical migration from north to south, but there is also evidence of gene flow from Santa Catalina to both Santa Rosa and Santa Cruz. This northerly migration against prevailing winds raises the possibility of more recent human-mediated acorn exchange among the islands. Given the sophisticated maritime cultures of the Native Americans who inhabited the islands for over 10,000 years, a human role in moving acorns from island to island as part of trade or ceremonial exchanges is not implausible. Interestingly however, allelic and private allelic richness were highest on Santa Catalina, followed by Santa Cruz and Santa Rosa. Centers of high levels of genetic diversity have been used to infer species origins.
(Rosenbom et al., 2014). This suggests that *Q. pacifica* may have originated on Santa Catalina with subsequent colonization of the northern islands, but this occurred long before human occupation of the islands.

What else can we infer about the demographic and establishment history of this island endemic? We found no evidence for recent bottlenecks, suggesting that the overgrazing and vegetation loss of the 20th century did not have a negative genetic impact on *Q. pacifica*. We did find evidence that bottlenecks took place at some time in the past, perhaps associated with the original colonization of the islands. M-ratio, on which we base this conclusion, is most likely to identify a bottleneck that lasted over multiple generations but from which the population made a demographic recovery (Williamson-Natesan, 2005). The current high level of genetic variability across the islands supports a demographic recovery from any initial bottlenecks and contradicts the prediction that low genetic diversity characterizes island populations and especially island endemics (DeJoode and Wendel, 1992; Frankham, 1997).

In summary, our data suggest that *Q. pacific* originated through an allopatric speciation event on one or more of the Channel Islands, followed by sufficient gene flow among the islands to maintain genetic continuity. *Quercus pacifica* is a genetically differentiated species, distinct from two mainland species, *Q. berberidifolia* and *Q. dumosa*, with which it has been closely linked. While there is evidence of historical bottlenecks on the islands, the current low levels of genetic differentiation among the island populations reinforces the importance of gene flow is maintaining the cohesiveness of this important island endemic.
Backs, Ashley:  **Endemic Channel Islands oak: Quercus pacifica** K. Nixon & C.H. Muller (Fagaceae)
Literature Cited


ATWATER, T. M. 1998. Plate tectonic history of southern California with emphasis on the western Transverse Ranges and northern Channel Islands. Contributions to the geology of the


multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567-1587.


**Frankham, R.** 1997. Do island populations have less genetic variation than mainland populations? *Heredity* 78: 311-327.


**Gerlach, G., A. Jueterbock, P. Kraemer, J. Deppermann, and P. Harmand.** 2010. Calculations of
population differentiation based on GST and D: forget GST but not all of statistics!

*Molecular Ecology* 19: 3845-3852.


Backs, Ashley: Endemic Channel Islands oak: *Quercus pacifica* K. Nixon & C.H. Muller (Fagaceae)


Jakobsson, M., and N. A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation


forests based on molecular and historic evidences. *Tree genetics & genomes* 10: 839-851.


Long-term human impacts on genetic structure of Italian walnut inferred by SSR markers.

Tree genetics & genomes 7: 707-723.


Backs, Ashley:  Endemic Channel Islands oak: *Quercus pacifica* K. Nixon & C.H. Muller (Fagaceae)
Table 1. Descriptive statistics for *Q. pacifica*, *Q. dumosa*, and *Q. berberidifolia*. Mean values for N = number of samples, A = number of alleles, Ar = allelic richness, PAr = private allele richness, Ho = observed heterozygosity, He = expected heterozygosity, FIS = fixation index. (GenAlEx 6.501, allelic richness using HP-RARE 1.0)

<table>
<thead>
<tr>
<th>Species</th>
<th>Population</th>
<th>N</th>
<th>A</th>
<th>Ar</th>
<th>PAr</th>
<th>Ho</th>
<th>He</th>
<th>FIS</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Q. pacifica</em></td>
<td>Overall</td>
<td>43.5</td>
<td>15.8</td>
<td>5.54</td>
<td>2.41</td>
<td>0.808</td>
<td>0.851</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td>S. Rosa</td>
<td>45.6</td>
<td>15.4</td>
<td>5.53</td>
<td>2.45</td>
<td>0.828</td>
<td>0.832</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>S. Cruz</td>
<td>37.8</td>
<td>15.3</td>
<td>5.75</td>
<td>2.63</td>
<td>0.800</td>
<td>0.862</td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>S. Catalina</td>
<td>47.0</td>
<td>16.8</td>
<td>5.85</td>
<td>2.83</td>
<td>0.795</td>
<td>0.860</td>
<td>0.075</td>
</tr>
<tr>
<td><em>Q. dumosa</em></td>
<td></td>
<td>24.0</td>
<td>11.1</td>
<td>4.90</td>
<td>1.86</td>
<td>0.776</td>
<td>0.802</td>
<td>0.029</td>
</tr>
<tr>
<td><em>Q. berberidifolia</em></td>
<td></td>
<td>42.8</td>
<td>18.6</td>
<td>5.72</td>
<td>2.66</td>
<td>0.846</td>
<td>0.878</td>
<td>0.039</td>
</tr>
</tbody>
</table>
Table 2. *Q. pacifica* island level and *Q. pacifica, Q. berberidifolia, and Q. dumosa* interspecies level results for $D_{JOST}$, Bonferroni corrected $p=0.003$ (Jost, 2008), and $F_{ST}$, $R_{ST}$, and $pR_{ST}$, the mean permutation value (based on allele size permutation tests using SPAGeDi 1.5a).

<table>
<thead>
<tr>
<th>Island level</th>
<th>$D_{JOST}$</th>
<th>$F_{ST}$</th>
<th>$R_{ST}$</th>
<th>$pR_{ST}$ (95% C.I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. Rosa</td>
<td>S. Cruz</td>
<td>0.106</td>
<td>-0.001 NS</td>
<td>0.014 (-0.005—0.045)</td>
</tr>
<tr>
<td>S. Rosa</td>
<td>S. Catalina</td>
<td>0.155</td>
<td>0.013 NS</td>
<td>0.032 (-0.002—0.104)</td>
</tr>
<tr>
<td>S. Cruz</td>
<td>S. Catalina</td>
<td>0.188</td>
<td>0.002 NS</td>
<td>0.028 (-0.002—0.083)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species level</th>
<th>$Q. berberidifolia$</th>
<th>$Q. dumosa$</th>
<th>$Q. pacifica$</th>
<th>$Q. dumosa$</th>
<th>$Q. pacifica$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q. berberidifolia$</td>
<td>0.341</td>
<td>0.053</td>
<td>0.133 *</td>
<td>0.050 (0.001—0.120)</td>
<td></td>
</tr>
<tr>
<td>$Q. berberidifolia$</td>
<td>0.219</td>
<td>0.028</td>
<td>0.109 **</td>
<td>0.031 (0.002—0.086)</td>
<td></td>
</tr>
<tr>
<td>$Q. dumosa$</td>
<td>0.274</td>
<td>0.040</td>
<td>0.095 *</td>
<td>0.039 (0.001—0.098)</td>
<td></td>
</tr>
</tbody>
</table>

P values testing if $R_{ST} > pR_{ST}$: NS, not significant; *$P<0.05$; **$P<0.01$
Table 3. Mean recent migration rates (m) with 95% confidence intervals for three island populations, estimated from eight microsatellites using BayesAss 3.0. Underlined values indicate the proportion of individuals in each generation that are not migrants (resident populations). Instances where mean m is below 0.099 are assumed to indicate no migration. Highlighted values show informative migration rates.

<table>
<thead>
<tr>
<th>To</th>
<th>From</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Santa Rosa</td>
<td>Santa Cruz</td>
<td>Santa Catalina</td>
</tr>
<tr>
<td>Santa Rosa</td>
<td>0.9741</td>
<td>0.0093</td>
<td>0.0166</td>
</tr>
<tr>
<td></td>
<td>(0.939-1.009)</td>
<td>(0.000-0.027)</td>
<td>(0.000-0.048)</td>
</tr>
<tr>
<td>Santa Cruz</td>
<td><strong>0.3111</strong></td>
<td>0.6761</td>
<td>0.0128</td>
</tr>
<tr>
<td></td>
<td>(0.282-0.340)</td>
<td>(0.658-0.694)</td>
<td>(0.000-0.036)</td>
</tr>
<tr>
<td>Santa Catalina</td>
<td><strong>0.1184</strong></td>
<td>0.0096</td>
<td><strong>0.872</strong></td>
</tr>
<tr>
<td></td>
<td>(0.030-0.207)</td>
<td>(0.000-0.027)</td>
<td>(0.783-0.961)</td>
</tr>
</tbody>
</table>
Table 4. Migrate-n Bayesian modes of effective population size ($\Theta = 4N_e\mu$) and bidirectional gene flow ($Nm =$ immigrants per generation). Numbers in parentheses lower 2.5% and upper 97.5% of posterior distribution.

<table>
<thead>
<tr>
<th>Source</th>
<th>$\Theta$</th>
<th>Santa Rosa</th>
<th>Santa Cruz</th>
<th>Santa Catalina</th>
</tr>
</thead>
<tbody>
<tr>
<td>Santa Rosa</td>
<td>3.633 (0.733, 7.467)</td>
<td>1.58 (0, 6.17)</td>
<td>1.58 (0.0, 6.17)</td>
<td></td>
</tr>
<tr>
<td>Santa Cruz</td>
<td>1.167 (0.000, 3.533)</td>
<td>4.92 (0.00, 10.17)</td>
<td>3.08 (0.0, 14.83)</td>
<td></td>
</tr>
<tr>
<td>Santa Catalina</td>
<td>8.967 (4.867, 11.667)</td>
<td>2.42 (0.00, 7.17)</td>
<td>1.08 (0.00, 5.50)</td>
<td></td>
</tr>
</tbody>
</table>
FIGURE LEGENDS

Fig. 1. *Quercus berberidifolia, Q. dumosa, and Q. pacifica* collection sites. Source: “Southern California collection sites”. 33°31’37.71” N and 118°42’35.11” W. GOOGLE EARTH. April 9, 2013. July 7, 2014.

Fig. 2. *Q. berberidifolia, Q. dumosa,* and *Q. pacifica* inferred admixture proportions in each genetic cluster for individual plants (columns). Admixture Model with LOCPRIOR grouped by species (STRUCTURE 2.3.4). Best K = 3 (per Evanno method). Species: 1 = *Q. berberidifolia*, 2 = *Q. dumosa*, 3 = *Q. pacifica*

Fig. 3. *Quercus pacifica* inferred admixture proportions in each genetic cluster for individual plants (columns). Admixture Model with NOLOCPRIOR (STRUCTURE 2.3.4). Best K= 3 (per Evanno method).

Fig. 4. *Q. pacifica* Principal Coordinates Analysis of variance. Principal coordinate 1 and 2 account for 5.48% and 5.37% of the variation, respectively (GenAlEx 6.501).