# Limited hybridization between *Quercus lobata* and *Quercus douglasii* (Fagaceae) in a mixed stand in central coastal California[1]

## Kathleen J. Craft,[2] Mary V. Ashley,[2,4] and Walter D. Koenig[3]

[2]Department of Biological Sciences, University of Illinois at Chicago, 845 West Taylor Street, Chicago, Illinois 60607 USA; and
[3]Hastings Natural History Reservation, University of California, Berkeley, 38601 E. Carmel Valley Road,
Carmel Valley, California 93924 USA

Many oak species are interfertile, and morphological and genetic evidence for hybridization is widespread. Here we use DNA microsatellite markers to characterize hybridization between two closely related oak species in a mixed stand in central coastal California, *Quercus lobata* (valley oak) and *Q. douglasii* (blue oak) (Fagaceae). Genotypes from four microsatellite loci indicate that many alleles are shared between the two species. However, each species harbors unique alleles, and allele frequencies differ significantly. A Bayesian analysis of genetic structure in the stand identified two highly differentiated genetic clusters, essentially corresponding to species assignment based on morphology. Data from the four loci were sufficient to assign all 135 trees to one of the two species. In addition, five putative hybrid individuals having intermediate morphologies could be assigned genetically to one or the other species, and all but one had low probability of hybrid ancestry. Overall, only six (4.6%) trees showed >0.05 probability of hybrid ancestry, in all cases their probabilities for nonhybrid ancestry were substantially higher. We conclude that adult hybrids of *Q. douglasii* × *Q. lobata* are rare at this site and plasticity in morphological characters may lead to overestimates of hybridization among *Quercus* species.

**Key words:** Bayesian clustering; California oaks; Fagaceae; hybridization; microsatellites; *Quercus douglasii*; *Quercus lobata*.

Many species of the genus *Quercus* (Fagaceae) are interfertile, and the poor development of sterility barriers between oak species has been of interest to evolutionary biologists since Darwin (Darwin, 1859). Natural hybrids form between species that are clearly distinguishable in morphology, physiology, and ecology. The propensity for oaks to hybridize, together with extensive edaphic and clinal variation, have led species concepts to be challenged in *Quercus* (Burger, 1975; Van Valen, 1976). Clearly, infertility between species cannot fully explain why closely related oaks usually remain distinct even in areas of sympatry.

Here we used nuclear DNA microsatellite markers to study hybridization of two widely overlapping species, *Q. lobata* and *Q. douglasii*, growing in a mixed stand in California. These species are closely related and have been reported to hybridize (Little, 1979; Holstein, 1984), yet are distinct in morphology and ecology. At our study site in central coastal California, *Q. lobata* and *Q. douglasii* are found together in extensive mixed woodlands with occasional trees intermediate in appearance.

We applied DNA microsatellite markers to study hybridization between *Q. lobata* and *Q. douglasii*. Microsatellites, rather than other available molecular markers, were chosen for several reasons. First, cytoplasmic genomes such as chloroplast (cp) and mitochondrial (mt) DNA variants show high rates of introgression among plant species (Rieseberg and Soltis, 1991) and may occur in the absence of significant nuclear gene flow (Rieseberg, 1995). Whittemore and Schaal (1991) found extensive exchange of chloroplast genomes between white oak species (subgenus *Quercus*), whereas a nuclear ribosomal marker was species specific. Since then, many reports of cytoplasmic exchange among *Quercus* species have been made (Petit, Kremer, and Wagner, 1993; Bacilieri et al., 1996; Dumolin-Lapegue, Pemonge, and Petit, 1998; Dumolin-Lapegue, Kremer, and Petit, 1999), always within the same subgenus but not necessarily between species that are otherwise closely related (Belahbib et al., 2001). Generally, morphologically recognizable hybrids are rarer in nature than would be predicted from estimates of cytoplasmic gene flow (Whittemore and Schaal, 1991). Thus, cytoplasmic markers may provide a distorted view of genetic exchange among *Quercus* species.

Nuclear genes may be more indicative of natural introgression levels and thus more useful for studies of hybridization in plants. Certain nuclear markers, however, may be problematic in *Quercus* due to the nature of molecular evolution at these loci. Muir, Fleming, and Schlötterer (2001) report that the ITS2 regions of the ribosomal DNA (rDNA) of the European oaks *Q. robur* and *Q. petraea* show no species-specific differences. Three very divergent rDNA families exist in both species and indeed were found within single individuals, limiting their application to hybridization studies.

In contrast, the characteristics of microsatellite markers (highly variable, codominantly inherited length variants) make them well suited for studies of hybridization and introgression (Muir, Fleming, and Schlötterer, 2000). Their relatively rapid rate of mutation allows for differentiation between closely related species, and differences in allele frequencies as well as species-specific alleles can be used to estimate levels of interspecific gene flow. For example, microsatellite data distinguish between *Q. robur* and *Q. petraea*, whereas chloroplast markers, RAPDs, ITS sequences, and allozymes do not (Muir, Fleming, and Schlötterer, 2000, 2001).

TABLE 1.    Description of four microsatellite loci used in this study, including number of individuals scored ($N$), number of alleles, observed ($H_o$) and expected ($H_e$) heterozygosities, proportion of observed heterozygote deficiency ($D$), estimated frequency of null alleles ($r$), and number of private alleles (PA) for *Quercus lobata* and *Q. douglasii*. Dashes for $D$ and $r$ indicate loci without heterozygote deficiency.

| Locus | Species | $N$ | No. alleles | $H_o$ | $H_e$ | $D$ | $r$ | PA |
|---|---|---|---|---|---|---|---|---|
| QpZAG9[a] | *Q. lobata* | 109 | 14 | 0.779 | 0.833 | 0.065 | 0.033 | 2 |
| | *Q. douglasii* | 24 | 14 | 0.792 | 0.912 | 0.044 | 0.070 | 2 |
| QpZAG110[a] | *Q. lobata* | 109 | 18 | 0.844 | 0.818 | — | — | 8 |
| | *Q. douglasii* | 24 | 13 | 0.708 | 0.850 | 0.167 | 0.011 | 3 |
| MSQ13[b] | *Q. lobata* | 109 | 12 | 0.615 | 0.572 | — | — | 7 |
| | *Q. douglasii* | 23 | 14 | 0.913 | 0.866 | — | — | 9 |
| MSQ4[b] | *Q. lobata* | 108 | 21 | 0.759 | 0.902 | 0.159 | 0.086 | 10 |
| | *Q. douglasii* | 25 | 13 | 0.760 | 0.909 | 0.134 | 0.080 | 2 |

[a] Steinkellner et al. (1997).
[b] Dow, Ashley, and Howe (1995).

To date, microsatellites studies in plants have been mostly used for studies of pollination and seed dispersal that involve parentage assignment (Chase et al., 1996; Dow and Ashley, 1996, 1998; Aldrich and Hamrick, 1998; Streiff et al., 1999) and studies assessing population genetic structure within species (Reusch, Stam, and Olsen, 2000; Friar et al., 2001) rather than for hybridization studies. Before conducting either mating system or population genetic studies in species that co-occur with potentially hybridizing species, however, it is necessary to characterize levels of microsatellite differentiation and exchange between species. This study was undertaken to obtain such information for *Q. douglasii* and *Q. lobata*.

Our study differs from others using genetic markers to study hybridization in plants in that we applied a Bayesian clustering approach to identify genetic structure in the mixed population in addition to more conventional analyses. That is, we tested for the presence of population structure in the combined multilocus genotype data without assigning individual trees to "*Q. lobata*" and "*Q. douglasii*" or "intermediate." We then asked how well the genetic structure of the mixed population corresponded to species assignments. This approach allows identification of hybrid individuals irrespective of their physical appearance. It also allows evaluation of the statistical significance of clusters, and the origin of individuals can be inferred by calculating the probability that individual multilocus genotypes belong to different genetic clusters, or alternatively, are hybrid in origin (Pritchard, Stephens, and Donnelly, 2000; Randi et al., 2001).

## MATERIALS AND METHODS

*Study organisms*—*Quercus lobata* is endemic to California, USA, with a distribution from Shasta County south through the Central Valley and lower-elevation foothills and valleys of the Sierra Nevada. It also occurs on Santa Cruz and Santa Catalina Islands (Griffin and Critchfield, 1972; Little, 1979). It once occurred as the only tree in extensive savannas, but this habitat has largely been cleared for farmland, vineyards, and other development. *Quercus lobata* also grows in riparian forest and in mixed oak stands in foothill woodlands. It is the largest North American oak with trees 10–25 m tall and 0.5–0.7 m diameter at breast height (dbh) (Munz, 1973). While valley oaks may live 500 yr, mature stands are typically 100–200 yr old.

*Quercus douglasii* is also endemic to California, with a distribution that circles the Central Valley, in valleys and foothills of the Coast Ranges and the Sierra Nevada. Communities where *Q. douglasii* occur range from open savanna to fairly dense woodland. It generally grows shorter and straighter than *Q. lobata*, ranging in height from 6 to 20 m and from 36 to 60 cm dbh (Burns and Honkala, 1990).

*Study area*—Trees were sampled at Hastings Reservation, a 900-ha reserve in the Santa Lucia Mountains of central coastal California. Hastings is located about 40 km inland at an elevation of 460–950 m above sea level and experiences a Mediterranean climate with an extended dry season from June to September. Three species of oak are common at Hastings, *Q. lobata*, *Q. douglasii*, and *Q. agrifolia*. *Quercus lobata* and *Q. douglasii* both belong to the "white oak" section *Quercus*, whereas *Q. agrifolia* belongs to the distantly related "red oak" section *Lobatae* (Nixon, 1993; Manos, Doyle, and Nixon, 1999). Thus only the two sampled species are candidates for hybridization. Based on vegetation surveys conducted throughout the Reservation, we estimate that *Q. douglasii*, which occur at densities averaging 69 trees/ha, are approximately three times more abundant than *Q. lobata* (W. D. Koenig, unpublished data).

*Sampling*—Mature leaves were collected from 109 *Q. lobata* and 26 *Q. douglasii* growing together in a mixed stand within an area of approximately 1.8 km². Sample sizes between species were different because *Q. lobata* was being exhaustively sampled for part of a pollination study. In addition to these trees, five individuals sampled were identified by one of us with extensive experience with oaks in this area (W. D. Koenig) as being of intermediate phenotype, indicated primarily by having bark that was not deeply furrowed or leaves that were somewhat bluish (as in *Q. douglasii*), but with leaves having the characteristic lobate shape of *Q. lobata*. This combination of characters is not common, and thus these trees were considered putative hybrids between the two species.

Both species appear to be regenerating poorly throughout much of California (Standiford, 2002), and saplings were in general not present within the study area. Thus, all individuals sampled were mature adults with the exception of two saplings, identified morphologically as *Q. lobata*, growing naturally within an exclosure present within the study area.

*Microsatellite data collection*—Approximately 1 g of frozen leaf material was ground to a fine powder in liquid nitrogen for extraction following Kleim et al. (1989). DNA was further purified using the Qiagen Tissue DNA Extraction Kit to remove proteins and other secondary compounds that inhibit polymerase chain reaction (PCR) reactions. Two of the microsatellite primers we used (MSQ13 and MSQ4) were developed for *Quercus macrocarpa* by Dow, Ashley, and Howe (1995) and Dow and Ashley (1996), and two (QpZAG9 and QpZAG110) were developed for *Q. petrea* by Steinkellner et al. (1997) (Table 1). One primer of each pair was fluorescently labeled. Polymerase chain reaction amplification was performed in reactions containing 0.2–0.4 µg genomic DNA, 100 µmol/L dNTP's, 0.15–0.6 µmol/L of each primer, 1.0 µg/µL bovine serum albumin, 2.0–3.0 mmol/L MgCl₂, PCR buffer, and 0.04 µL *Taq* polymerase (Promega). All PCR reactions consist of a 3-min preheat period at 95°C followed by 38 cycles of denaturing at 94°C for 30 s, annealing at 54°C for 30 s, and extension at 72°C for 45 s and a final extension at 72°C for 5 min. The PCR products were then run along with the fluorescent size standard Tamara 350 on an MJBaseStation automated se-

quencer (MJ Research, Waltham, Massachusetts, USA). Cartographer software (MJ Research) was used to size and score genotypes.

***Genetic analyses***—Individuals were included in the analyses if they were scored for at least two of the four loci, so sample sizes are not equal across loci (Table 1). Descriptive statistics such as observed and expected heterozygosities were calculated using GDA Version 1.0 (d16c) (Lewis and Zaykin, 2001). We used GENEPOP (Raymond and Rousset, 1995) to implement the private allele method for estimating the number of migrants per generation (Slatkin, 1985; Barton and Slatkin, 1986), to test for homogeneity of allele distributions between species, and to test for deviations from Hardy-Weinberg equilibrium. GENEPOP calculates unbiased estimates of *P* values using a Markov chain method. Because null alleles had been observed in *Q. macrocarpa* for at least one of the loci used (Dow and Ashley, 1996), we estimated the proportion of observed heterozygote deficiencies (*D*) and the frequencies of null alleles (*r*) according to Chakraborty et al. (1992) using the expressions: $D = H_E - H_O/H_E$ and $r = (H_E - H_O)/(H_E + H_O)$.

We also used GENEPOP to calculate an estimate of Wright's $F_{ST}$ (θ, Weir and Cockerham, 1984). GENEPOP uses standardized allele sizes to determine variance components at each locus and calculates ρ, an estimator of $R_{ST}$ (Slatkin, 1995). $R_{ST}$ is an index of genetic differentiation estimated from the sum of squared number of repeat differences (Michalakis and Excoffier, 1996; Rousset, 1996). Although typically applied to population subdivision within species, these statistics were used here to assess genetic differentiation between the two species.

We also applied a Bayesian clustering approach to assign individual trees to species on the basis of their genotypes alone. This allowed us to determine how well the genetic structure in the entire data set corresponded to phenotypic assignment to species. This approach, implemented using the program *structure* (Pritchard, Stephens, and Donnelly, 2000), uses a model-based clustering method for inferring population structure from multilocus genotype data and to simultaneously assign individuals to clusters. It also allows for the identification of hybrid individuals and estimates the fraction of their alleles that are derived from each species.

To verify that our data set showed genetic differentiation between the two species, we ran five independent runs of the Gibbs sampler in *structure* at population numbers (*K*) at *K* = 1 (no structure), *K* = 2 (two genetic clusters), and *K* = 3 (three genetic clusters), using a burn-in period of 30 000 iterations and collected data for $10^6$ iterations. These initial runs were performed without using any information regarding species identification (USEPOPINFO = 0). Results at a given *K* were highly consistent across runs. Given *X,* the observed genotypes, ln Pr(*X*|*K*) = −2583 for *K* = 1, −2434 for *K* = 2, and −2463 for *K* = 3. The corresponding values for Pr (*K*|*X*) assuming a uniform prior on *K* = 1, 2, 3, yields a probability of about 0.0 for *K* = 1 and *K* = 3 and about 1.0 for *K* = 2. This confirms that the data set represents two clusters that are genetically strongly defined. To make sure our skewed sampling of species did not influence this result, we repeated the analyses with random subsets of 26 *Q. lobata* individuals, and each time two similar genetic clusters were identified. We thus continued our analyses to explore how well this structure corresponded to our assignment of individuals to species and to identify hybrid individuals, using prior information on species assignments (USEPOPINFO = 1). We ran the analysis at three values of the intergroup "immigration" rate *v* (in this context, "immigrant" means "hybrid"), as recommended by Pritchard, Stephens, and Donnelly (2000) to see whether results were robust to choice of *v*. Values used were *v* = 0.10, 0.05, and 0.01, which we considered to span the range of reasonable frequencies of *Q. lobata* and *Q. douglasii* hybridization. The *structure* program then estimates the posterior probabilities that an individual is purely from the other population (in this case mis-assigned to species) or has a given amount of ancestry in the other species.

## RESULTS

All four microsatellites were highly variable, with ≥12 alleles per locus and a mean expected heterozygosity of 0.833 across all loci (Table 1). *Quercus lobata* and *Q. douglasii*
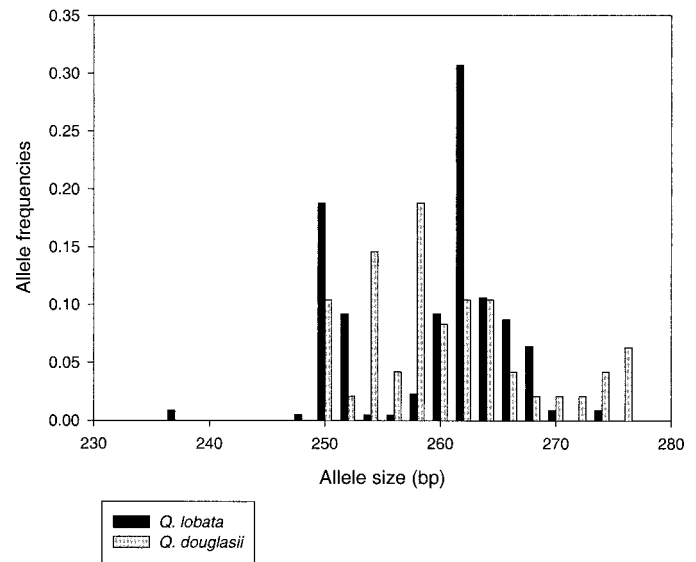
QpZAG9 allele frequencies



Fig. 1. Histogram showing the frequency distributions of microsatellite lengths of *Quercus lobata* (black bars) and *Q. douglasii* (gray bars) at microsatellite locus QpZAG9.

shared alleles at all loci but each harbored private alleles not found in the other species. Tests for homogeneity of allele distributions were highly significant (*P* < 0.0001) for all four loci. Aside from the general pattern of heterogeneity, the patterns varied extensively among loci. For QpZAG9 (Fig. 1), allele sizes ranged from 237 to 274 base pairs (bp) for *Q. lobata* and from 250 to 276 bp for *Q. douglasii*. The most common alleles for *Q. lobata* were 250 (19%) and 262 (31%). These two alleles were common in *Q. douglasii* but were not the most frequent. For QpZAG110 (Fig. 2), the most common allele in both species was 209, occurring at similar frequencies of 33%. There were other common alleles at this locus, however, that were unique to each species, including 207 in *Q. lobata* and 220 in *Q. douglasii*.

For locus MSQ13, only 5 of 21 alleles were shared between species, and each of these occurred at very low frequencies in *Q. lobata*. *Quercus lobata* alleles were generally in a smaller size range than *Q. douglasii* alleles at this locus (Fig. 3). Locus MSQ4 (Fig. 4) was more similar to QpZAG9 in that most, but not all, common alleles were shared between the two species. The mean frequency of private alleles over all loci was 0.049.

Global tests for heterozygote deficiencies revealed significant departures from Hardy-Weinberg expectations at three loci in both species, QpZag9, QpZAG110, and MSQ4 (*P* = 0.004, *P* = 0.001, and *P* < 0.0001, respectively). In the case of microsatellites, heterozygote deficiencies are likely indicative of null alleles present at these loci. Estimates of null allele frequencies (*r*) for loci having heterozygote deficiencies are given in Table 1. While the presence of null alleles may result in overestimation of the number of populations in a *structure* analysis (J. K. Pritchard, University of Chicago, personal communication), this did not seem to occur in our analysis because the presence of two genetic clusters was strongly supported.

Weir and Cockerham's (1984) estimator of population subdivision, θ, was 0.049, 0.040, 0.035, and 0.298 for QpZAG9, QpZAG110, MSQ13, and MSQ4, respectively, significant

### QpZAG110 allele frequencies



Fig. 2.   Histogram showing the frequency distributions of microsatellite lengths of *Quercus lobata* (black bars) and *Q. douglasii* (gray bars) at microsatellite locus QpZAG110.
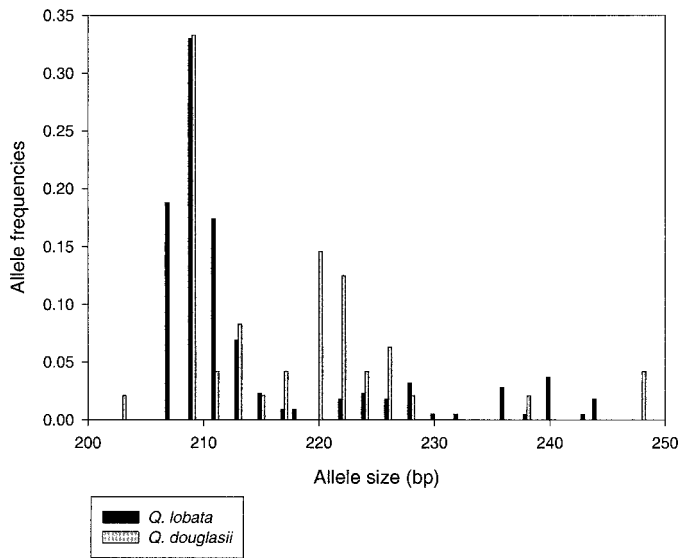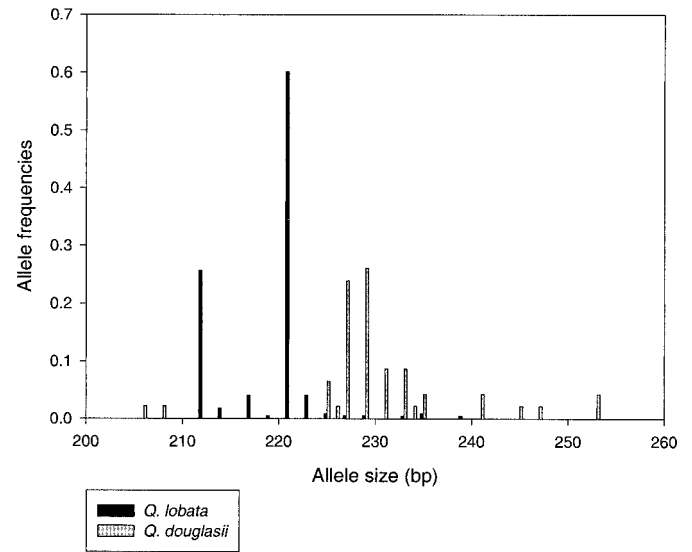
### MSQ13 allele frequencies



Fig. 3.   Histogram showing the frequency distributions of microsatellite lengths of *Quercus lobata* (black bars) and *Q. douglasii* (gray bars) at microsatellite locus MSQ13.

over all loci ($\theta = 0.106$, $P < 0.05$). The estimator of $R_{ST}$, $\rho$, was 0.004, 0.0135, 0.133, and 0.667 for QpZAG9, QpZAG110, MSQ13, and MSQ4, respectively, and was significant over all loci ($\rho = 0.216$, $P < 0.05$). Estimates of the number of migrants ($N_m$) based on $\theta = 2.11$, 0.91 for $\rho$, and 1.19 using the private allele method (Slatkin, 1985; Barton and Slatkin, 1986).

For the Bayesian analysis, we first used *structure* to cluster our data without using information regarding species assignment of individual trees. The two genetically distinct clusters found in the analysis correspond well to our assignment of individuals to *Q. lobata* and *Q. douglasii*. Of 109 *Q. lobata* genotyped, 98 (90%) had inferred ancestry in a "*lobata*" cluster with probability >0.90, and 22 of 26 *Q. douglasii* (85%) had >0.90 inferred ancestry in a "*douglasii*" cluster (Fig. 5). Of the five putative hybrid trees, two (1102H and 1050H) had >0.95 probability of ancestry in the "*douglasii*" cluster, whereas the other three, 1055H, 1123H, and 167H, had probabilities of ancestry in the "*lobata*" cluster of 0.954, 0.891, and 0.687, respectively.

Next, we incorporated species information to improve clustering and to look for first- or later-generation hybrid ancestry. We assigned each of the putative hybrids to either *Q. douglasii* or *Q. lobata* as indicated by the earlier analysis. Incorporating species identification improved the accuracy of assigning individuals to clusters. Four of the five putative hybrids had very small probabilities of being first generation hybrids, even at *v* = 0.10 (Table 2). One, 167H, had probability of 0.231 of being a hybrid and thus was our mostly likely hybrid candidate. Even this tree, however, had a greater probability (0.520) of being "pure" *Q. lobata*. Only five other trees (all shown in Table 2) had >0.05 probability of being first-generation hybrids; these were all individuals phenotypically assigned to *Q. lobata*. These five trees had probabilities of "pure" *Q. lobata* ancestry that were at least twice that of their probabilities

of hybrid ancestry at all values of *v*. One of these (1117L) was one of the two saplings included in the analysis.

### DISCUSSION

*Quercus lobata* and *Q. douglasii* are closely related oak species that have broad areas of sympatry in California and are reported to hybridize (Little, 1979; Holstein, 1984). Occasional individuals with intermediate morphologies, at Has-
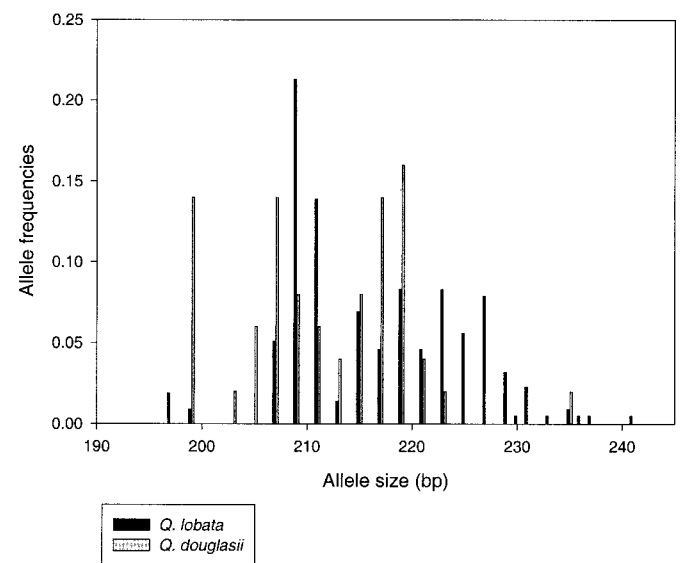
### MSQ4 allele frequencies



Fig. 4.   Histogram showing the frequency distributions of microsatellite lengths of *Quercus lobata* (black bars) and *Q. douglasii* (gray bars) at microsatellite locus MSQ4.
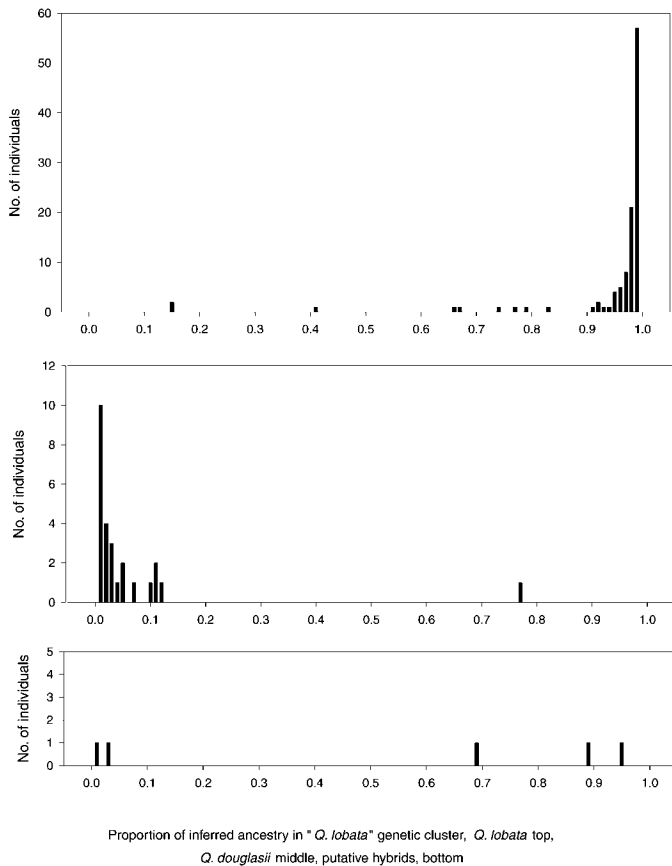
Fig. 5. Upper histogram shows the mean value of $q_1^{(i)}$ (the proportion of ancestry in population 1, the "*lobata*" cluster) over a single run of the Gibbs sampler in *structure* for trees identified in the field as *Quercus lobata*. Middle histogram shows $q_1^{(i)}$ for trees identified in the field as *Q. douglasii*. Lower histogram shows $q_1^{(i)}$ for trees identified as putative hybrids based on intermediate morphology.

TABLE 2. Test results showing whether particular individuals are hybrids. Values represent probabilities.

| Individual | Species assignment[a] | $\nu$[b] | No hybrid ancestry | Mis-assigned to species | $F_1$ hybrid | $F_2$ hybrid |
|---|---|---|---|---|---|---|
| 167H | Hybrid | 0.01 | 0.915[c] | 0.000 | 0.042 | 0.042 |
| | | 0.05 | 0.687 | 0.002 | 0.152 | 0.159 |
| | | 0.10 | 0.520 | 0.002 | 0.231 | 0.246 |
| 1050H | Hybrid | 0.01 | 0.994[d] | 0.000 | 0.001 | 0.005 |
| | | 0.05 | 0.972 | 0.000 | 0.003 | 0.025 |
| | | 0.10 | 0.943 | 0.000 | 0.005 | 0.052 |
| 1055H | Hybrid | 0.01 | 0.993[c] | 0.000 | 0.002 | 0.005 |
| | | 0.05 | 0.972 | 0.000 | 0.003 | 0.025 |
| | | 0.10 | 0.922 | 0.000 | 0.020 | 0.058 |
| 1102H | Hybrid | 0.01 | 0.998[d] | 0.000 | 0.001 | 0.011 |
| | | 0.05 | 0.944 | 0.000 | 0.005 | 0.051 |
| | | 0.10 | 0.894 | 0.000 | 0.009 | 0.097 |
| 1123H | Hybrid | 0.01 | 0.994[c] | 0.000 | 0.001 | 0.005 |
| | | 0.05 | 0.967 | 0.000 | 0.008 | 0.025 |
| | | 0.10 | 0.926 | 0.001 | 0.018 | 0.055 |
| 1083L | *Quercus lobata* | 0.01 | 0.982 | 0.000 | 0.005 | 0.013 |
| | | 0.05 | 0.909 | 0.000 | 0.026 | 0.065 |
| | | 0.10 | 0.819 | 0.000 | 0.053 | 0.128 |
| 1100L | *Quercus lobata* | 0.01 | 0.942 | 0.005 | 0.027 | 0.026 |
| | | 0.05 | 0.750 | 0.022 | 0.117 | 0.112 |
| | | 0.10 | 0.569 | 0.040 | 0.201 | 0.190 |
| 1117L | *Quercus lobata* | 0.01 | 0.857 | 0.060 | 0.024 | 0.059 |
| | | 0.05 | 0.531 | 0.195 | 0.077 | 0.197 |
| | | 0.10 | 0.354 | 0.262 | 0.107 | 0.278 |
| 1122L | *Quercus lobata* | 0.01 | 0.920 | 0.000 | 0.037 | 0.043 |
| | | 0.05 | 0.696 | 0.000 | 0.138 | 0.165 |
| | | 0.10 | 0.522 | 0.000 | 0.216 | 0.262 |
| 1156L | *Quercus lobata* | 0.01 | 0.972 | 0.000 | 0.008 | 0.020 |
| | | 0.05 | 0.875 | 0.002 | 0.035 | 0.088 |
| | | 0.10 | 0.774 | 0.003 | 0.063 | 0.160 |

[a] Initial assignment based on appearance of tree and leaves.
[b] $\nu$ is the probability of mixed ancestry.
[c] Assigned to *Q. lobata* based on genetic analysis.
[d] Assigned to *Q. douglasii* based on genetic analysis.

tings and elsewhere, support the possibility of hybridization. At our study site, the species show similar flowering phenologies (W. D. Koenig, unpublished data) and patterns of annual acorn production that are both strongly correlated with each other (Koenig et al., 1994) and with conditions during the spring flowering period (Koenig et al., 1996). Our goal was to characterize microsatellite variability in these species and to use microsatellite data to gain accurate estimates of hybridization frequency.

We applied both standard genetic differentiation approaches as well as Bayesian clustering methods to characterize hybridization. Recently, the appropriateness of different measures of genetic differentiation for microsatellite loci (e.g., $F_{ST}$ and $R_{ST}$) has come under scrutiny (Gaggiotti et al., 1999; Hedrick, 1999; Balloux and Lugon-Moulin, 2002). In our case, both estimates of genetic differentiation were significant between *Q. lobata* and *Q. douglasii*, although values of both estimators ranged widely across loci. Over all loci, the estimate of $R_{ST}$ was more than twice the value of $F_{ST}$, indicating that the two species do not just differ in distributions of allele frequencies, but also differ in allele sizes. Such shifts likely occur over longer divergence times than changes in frequencies (Slatkin, 1995), suggesting historically low levels of introgression. The existence of many private alleles further indicates little gene flow between the two species (Slatkin, 1985). *Quercus doug-*

*lasii* had private alleles at all four loci, even though the number of individuals sampled was relatively small. *Quercus lobata* had private alleles at three loci, although some of these alleles may have been found at low frequency in *Q. douglasii* but not sampled.

Previous work on hybridization in plants has relied upon a priori assignment of individual trees to groups (for example, species "A," species "B," or "hybrid") based on phenotype. Hybridization in oaks is commonly inferred on the basis of morphologically intermediate trees. The classification by the researcher, however, may not reflect the underlying genetic structure of the population and may not present an accurate picture regarding introgression and hybridization. In our initial Bayesian analysis, we used a purely genetical analysis, using no external information, to infer population structure. The correspondence between genetic structure and phenotypic classification to species was extremely close, although the few exceptions (Table 2) are informative. The Bayesian clustering approach has the further advantage of assigning individuals to populations according to likelihoods based on allele frequencies and to identify individuals whose genetic makeup may be drawn from two populations (hybrids). This allowed us to rule out hybrid ancestry for nearly all sampled individuals. Four of five phenotypically intermediate trees (1050H, 1055H, 1102H, and 1123H, Table 2) had very low probabilities of hybrid ancestry. Three trees assigned to *Q. lobata* had modest

probabilities of hybrid ancestry (1083L, 1117L, and 1156L, Table 2); one of these (1117L) was one of the two saplings sampled. Even the three trees that had the highest probability of hybrid ancestry (167H, 1100L, and 1122L, Table 2) had probabilities of pure *Q. lobata* ancestry that were twice as great. Our results suggests that the frequency of hybridization across all trees of these species is no higher than 4% (6/141), but could be much closer to zero.

Our results caution against the assumption that phenotypically intermediate trees are necessarily hybrids. Four phenotypically intermediate trees at our site showed little evidence of mixed ancestry; two were assigned to *Q. lobata* (1055H and 1123H, Table 2) and two to *Q. douglasii* (1050H and 1102H). Of the four trees with the highest probability of hybrid ancestry, only one was identified as intermediate in appearance. These findings suggest that apparently intermediate phenotypes between these two species are not necessarily hybrids and that true hybrids are not necessarily intermediate in phenotype.

The high level of differentiation between *Q. lobata* and *Q. douglasii* for microsatellite loci suggests that these markers may provide improved resolution for studying hybridization in oaks. Other molecular markers have often proven unsatisfactory. For example, chloroplast DNA may be commonly exchanged between species that are distantly related and show only limited ability to hybridize (Whittemore and Schaal, 1991; Belahbib et al., 2001). Other markers, such as allozymes, cannot distinguish among closely related oak species. Manos and Fairbrothers (1987) found that five species of red oaks had a high degree of allozymic similarity, with a mean genetic identity of 0.958 between species, and concluded that isozymes "were of little value in identifying species since marker alleles were not found" (p. 365). Nason, Ellstrand, and Arnold (1992) had better success using allozymes to study hybridization between two black oaks, *Q. kelloggi* and *Q. wislizenii*, and found these species differed in numbers and frequencies of alleles at allozyme loci. In contrast to our results, the putative hybrid trees in their study did correspond genetically to $F_1$ hybrids, bearing alleles unique to alternative parental species at different loci. Howard et al. (1997) were successful in finding 14 species-specific RAPD markers for the closely related white oaks *Q. grisea* and *Q. gambelii*, but this was only through screening 700 primers.

The best studied oak system thus far, and the only other system where microsatellite markers have been applied to the question of hybridization, is that of the European *Q. robur*-*Q. petraea*-*Q. pubescens* complex. These species are highly interfertile (Bacilieri et al., 1996) and lack diagnostic chloroplast markers (Ferris et al., 1993; Dumolin-Lapegue, Kremer, and Petit, 1999), allozymes (Zanetto, Roussel, and Kremer, 1994), and ITS sequences (Muir, Fleming, and Schlötterer, 2001). Although the species show small but significant differences in allele frequencies at microsatellite loci (Muir, Fleming, and Schlötterer, 2000), they are clearly not as well differentiated as *Q. lobata* and *Q. douglasii*. For example, assignment tests using data from 20 microsatellite loci could assign only 78% of individuals to species (Muir, Fleming, and Schlötterer, 2000). In our study, data from only four microsatellite loci were sufficient to assign nearly all individuals to one species or another. Bruschi et al. (2000) reports no significant differences for *Q. petraea* and *Q. pubescens* at two of three nuclear microsatellite loci and genetic differentiation ($R_{ST}$) between the species was only 0.048, compared to 0.216 for *Q. lobata* and *Q. douglasii*.

Hybridization and species boundaries among oak species remain complicated and compelling issues. Here we show that genetic introgression may be rare, even between closely related species occurring in a mixed stand. Because we primarily genotyped adult oaks, we do not know whether there are strong fertility barriers between *Q. lobata* and *Q. douglasii* or whether hybrid seedlings or saplings do not survive to adulthood. Indeed, the result that one of the two saplings we were able to sample was a potential hybrid raises the possibility that hybrids among acorns and saplings might be more common than among adults. Genotyping of *Q. lobata* acorns, planned as part of a pollination study, will allow us to determine if *Q. douglasii* pollen fertilizes *Q. lobata* acorns and whether the proportion of hybrid individuals declines between seed and adult stages.

## LITERATURE CITED

ALDRICH, P. R., AND J. L. HAMRICK. 1998. Reproductive dominance of pasture trees in a fragmented tropical forest mosaic. *Science* 281: 103–105.

BACILIERI, R., A. DUCOUSSON, R. J. PETIT, AND A. KREMER. 1996. Mating system and asymetric hybridization in a mixed stand of European oaks. *Evolution* 50: 900–908.

BALLOUX, F., AND N. LUGON-MOULIN. 2002. The estimation of population differentiation with microsatellite markers. *Molecular Ecology* 11: 155–165.

BARTON, N. H., AND M. SLATKIN. 1986. A Quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56: 409–415.

BELAHBIB, N., M.-H. PEMONGE, A. OUASSOU, H. SBAY, A. KREMER, AND R. J. PETIT. 2001. Frequent cytoplasmic exchanges between oak species that are not closely related: *Quercus ruber* and *Q. ilex* in Morocco. *Molecular Ecology* 10: 2003–2012.

BRUSCHI, P., G. G. VENDRAMIN, F. BUSSONTTI, AND P. GROSSONI. 2000. Morphological and molecular differentiation between *Quercus petraea* (Matt.) Liebl. and *Quercus pubescens* Willd. (Fagaceae) in northern and central Italy. *Annals of Botany* 85: 325–333.

BURGER, W. C. 1975. The species concept in *Quercus. Taxon* 24: 45–50.

BURNS, R. M., AND B. H. HONKALA. 1990. Silvics of North America, vol. 2. Hardwoods. Agricultural Handbook 654, Department of Agriculture, Forest Service, Washington D.C., USA.

CHAKRABORTY, R., M. DE ANDRADE, S. P. DAIGER, AND B. BUDOWLE. 1992. Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Annals of Human Genetics* 56: 45–57.

CHASE, M. R., C. MOLLER, R. KESSELI, AND K. S. BAWA. 1996. Distant gene flow in tropical trees. *Nature* 383: 398–399.

DARWIN, C. 1859. On the origin of species by means of natural selection. John Murray, London, UK.

DOW, B. D., AND M. V. ASHLEY. 1996. Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa. Molecular Ecology* 5: 120–132.

DOW, B. D., AND M. V. ASHLEY. 1998. High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. *Journal of Heredity* 89: 62–70.

DOW, B. D., M. V. ASHLEY, AND H. F. HOWE. 1995. Characterization of highly variable (GA/CT)ₙ microsatellites in the bur oak, *Quercus marcrocarpa. Theoretical and Applied Genetics* 91: 137–141.

DUMOLIN-LAPEGUE, S., A. KREMER, AND R. J. PETIT. 1999. Are chloroplast and mitochondrial DNA variation species independent in oaks? *Evolution* 53: 1406–1413.

DUMOLIN-LAPEGUE, S., M.-H. PEMONGE, AND R. J. PETIT. 1998. Association between chloroplast and mitochondrial lineages in oaks. *Molecular Biology and Evolution* 15: 1321–1331.

FERRIS, C., R. P. OLIVER, A. J. DAVY, AND G. M. HEWITT. 1993. Native oak chloroplasts reveal an ancient divide across Europe. *Molecular Ecology* 2: 337–344.

FRIAR, E. A., D. L. BOOSE, T. LADOUX, E. H. ROALSON, AND R. H. ROBI-

CHAUX. 2001. Population structure in the endangered Mauna Loa silversword, *Argyroxiphium kauense* (Asteraceae), and its bearing on reintroduction. *Molecular Ecology* 10: 1657–1663.

GAGGIOTTI, O. E., O. LANGE, K. RASSMAN, AND C. GLIDDON. 1999. A comparison of two indirect methods for extimating average levels of gene flow using microsatellite data. *Molecular Ecology* 8: 1513–1520.

GRIFFIN, J. R., AND W. B. CRITCHFIELD. 1972. The distribution of forest trees in California. Pacific Southwest Forest and Range Experiment Station, Berkeley, California, USA.

HEDRICK, P. W. 1999. Highly variable loci and their interpretation in evolution and conservation. *Evolution* 53: 313–318.

HOLSTEIN, G. 1984. California riparian forests: deciduous islands in an evergreen sea. *In* R. E. Warner and K. M. Hendrix [eds.], California riparian systems: ecology, conservation, and productive management: proceedings of a conference, 2–22. University of California Press, Berkeley, California, USA.

HOWARD, D. J., R. W. PRESZLER, J. WILLIAMS, S. FENCHEL, AND W. J. BOECKLEN. 1997. How discrete are oak species? Insights from a hybrid zone between *Quercus grisea* and *Quercus gambelii*. *Evolution* 51: 747–755.

KEIM, P., K. N. PAIGE, T. G. WHITHAM, AND K. G. LARK. 1989. Genetic analysis of an interspecific hybrid swarm of populus: occurrence of unidirectional introgression. *Genetics* 123: 557–565.

KOENIG, W. D., J. M. H. KNOPS, W. J. CARMEN, M. T. STANBACK, AND R. L. MUMME. 1996. Acorn production by oaks in central coastal California: influence of weather at three levels. *Canadian Journal of Forest Research* 26: 1677–1683.

KOENIG, W. D., R. L. MUMME, W. J. CARMEN, AND M. T. STANBACK. 1994. Acorn production by oaks in central coastal California: variation within and among years. *Ecology* 75: 99–109.

LEWIS, P. O., AND D. ZAYKIN. 2001. Genetic data analysis: computer program for the analysis of allelic data. http://lewis.eeb.uconn.edu/lewishome/software.htm.

LITTLE, E. L., JR. 1979. Checklist of United States trees (native and naturalized). Agricultural Handbook 541, U.S. Department of Agriculture, Forest Service, Washington D.C., USA.

MANOS, P. S., J. J. DOYLE, AND K. C. NIXON. 1999. Phylogeny, biogeography, and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). *Molecular Phylogenetics and Evolution* 12: 333–349.

MANOS, P. S., AND D. E. FAIRBROTHERS. 1987. Allozyme variation in populations of six Northeastern American Red Oaks (Fagaceae: *Quercus* subg. *Erythrobalanus*). *Systematic Botany* 12: 365–373.

MICHALAKIS, Y., AND L. EXCOFFIER. 1996. A generic estimation of population subdivision using distances between alleles with special reference to microsatellite loci. *Genetics* 142: 1061–1064.

MUIR, G., C. C. FLEMING, AND C. SCHLÖTTERER. 2000. Species status of hybridizing oaks. *Nature* 405: 1016.

MUIR, G., C. C. FLEMING, AND C. SCHLÖTTERER. 2001. Three divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. *Molecular Biology and Evolution* 18: 112–119.

MUNZ, P. A. 1973. A California flora and supplement. University of California Press, Berkeley, California, USA.

NASON, J. D., N. C. ELLSTRAND, AND M. L. ARNOLD. 1992. Patterns of hybridization and introgression in populations of oaks, manzanitas, and irises. *American Journal of Botany* 79: 101–111.

NIXON, K. C. 1993. Infrageneric classification of *Quercus* (Fagaceae) and typification of sectional names. *Annales des Sciences Forestieres Supplement 1 (Paris)* 50: 25s–34s.

PETIT, R., A. KREMER, AND D. WAGNER. 1993. Geographic structure of chloroplast DNA polymorphisms in European oaks. *Theoretical and Applied Genetics* 87: 122–128.

PRITCHARD, J. K., M. STEPHENS, AND P. DONNELLY. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

RANDI, M., M. PIERPAOLI, M. BEAUMONT, AND B. RAGNI. 2001. Genetic identification of wild and domestic cats (*Felix silvestris*) and their hybrids using Bayesian clustering methods. *Molecular Biology and Evolution* 18: 1670–1693.

RAYMOND, M., AND F. ROUSSET. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248–249.

REUSCH, T. B. H., W. T. STAM, AND J. L. OLSEN. 2000. A microsatellite-based estimation of clonal diversity and population subdivision in *Zostera marina*, a marine flowering plant. *Molecular Ecology* 9: 127–140.

RIESEBERG, L. H. 1995. The role of hybridization in evolution: old wine in new skins. *American Journal of Botany* 82: 944–953.

RIESEBERG, L. H., AND D. E. SOLTIS. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* 5: 65–84.

ROUSSET, F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics* 142: 1357–1362.

SLATKIN, M. 1985. Rare alleles as indicators of gene flow. *Evolution* 39: 53–65.

SLATKIN, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 130: 457–462.

STANDIFORD, R. B. 2002. California oak woodlands. *In* W. J. McShea and W. M. Healy [eds.], Oak forest ecosystems, 280–303. Johns Hopkins Press, Baltimore, Maryland, USA.

STEINKELLNER, H., S. FLUCH, E. TURETSCHEK, C. LEXER, R. STREIFF, A. KREMER, K. BURG, AND J. GLÖSSL. 1997. Identification and characterization of (GA/CT)$_n$-microsatellite loci from *Quercus petraea*. *Plant Molecular Biology* 33: 1093–1096.

STREIFF, R., A. DUCOUSSO, C. LEXER, H. STEINKELLNER, J. GLOESSL, AND A. KREMER. 1999. Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. *Molecular Ecology* 8: 831–841.

VAN VALEN, L. 1976. Ecological species, multispecies, and oaks. *Taxon* 25: 233–239.

WEIR, B., AND C. COCKERHAM. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.

WHITTEMORE, A. T., AND B. A. SCHAAL. 1991. Interspecific gene flow in sympatric oaks. *Proceedings of the National Academy of Sciences, USA* 88: 2240–2544.

ZANETTO, A., G. ROUSSEL, AND A. KREMER. 1994. Geographic variation of interspecific differentiation between *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. *Forest Genetics* 2: 111–123.