Nuclear DNA diversity, population differentiation, and phylogenetic relationships in the California closed-cone pines based on RAPD and allozyme markers

J. Wu, K.V Krutovskii, and S.H. Strauss

Abstract: We studied nuclear gene diversity and population differentiation using 91-98 randomly amplified polymorphic DNA (RAPD) loci in the California closed-cone pines knobcone (*Pinus attenuata* Lemm.), bishop (*P. muricata* D. Don), and Monterey (*P. radiata* D. Don) pines. A total of 384 trees from 13 populations were analyzed for RAPDs and another sample of 242 trees from 12 of these 13 populations were analyzed at 32-36 allozyme loci, using a published data set. Twenty-eight of 30 (93%) comigrating RAPD fragments tested were found to be homologous by Southern hybridization in all three species. Using an enriched mitochondrial DNA (mtDNA) preparation and a chloroplast DNA (cpDNA) library as probes, two fragments of cpDNA origin, and one of mtDNA origin present among RAPD profiles were excluded from analysis of nuclear gene diversity. RAPD markers revealed moderately higher intrapopulation gene diversity and substantially higher total genetic diversity and population differentiation than did allozyme markers for each species. We performed a simulation study using allozyme data, which showed that the dominant and biallelic nature of RAPD markers could explain the differences observed in differentiation parameters, but not in gene diversity: RAPD phenotypes appear to represent more underlying gene diversity than do allozyme phenotypes. Results of joint phylogenetic analyses of both the RAPD and allozyme markers strongly supported a common ancestor for P. radiata and P. attenuata pines, and south-to-north migration histories for all three species.

Key words: allozymes, dominance, gene diversity, *Pinus attenuata*, *Pinus muricata*, *Pinus radiata*, phylogeny, RAPDs.

Resume : Les auteurs ont etudie la diversitd du genome nucleaire et la differenciation des populations **a** 1'aide de 91-98 loci RAPD (ADN polymorphe amplifie au hasard) chez les especes de pins A clines fermes de Californie *Pinus attenuata* Lemm., *P. muricata* D. Don et *P. radiata* D. Don. Au total. 384 arbres provenant de 13 populations ont ete analyses a I'aide des marqueurs RAPD et un autre echantillon comprenant 242 arbres provenant de 12 des 13 populations mentionnees plus haut a ete analyse a I'aide de 32-36 marqueurs alloenzymatiques pour lesquelles un ensemble de donnees avait ete publie anterieurement. Trente produits RAPD montrant une comigration ont ete testes par hybridation Southern et vingt-huit (93%) se sont averes homologues chez les trois especes. Au moyen de sondes constituees soit d'une preparation d'ADN enrichie en ADN mitochondrial (ADNmt) soit d'une banque de sequences d'ADN chloroplastique (ADNcp), il a ete possible d'exclure deux produits d'amplification d'origine chloroplastique et un produit d'origine mitochondriale afin de limiter 1'analyse de la diversite genique au seul genome nucleaire. Les marqueurs RAPD ont revdle un peu plus de diversite genique intrapopulation ainsi que beaucoup plus de diversite genetique globale et de differenciation des populations que ne 1'avaient fait les alloenzymes chez chacune de ces especes. Une simulation a ete realisee a Faide des donnees alloenzymatiques et celle-ci a montre que la nature diallelique et dominante des marqueurs RAPD pouvait expliquer les differences observees quant aux parametres de differenciation mais non quant a la diversite genique. Les phenotypes RAPD semblent refleter davantage la diversite genique que les phenotypes alloenzymatiques. Les resultats d'analyses phylogenetiques conjointes employant **a** la fois les RAPD et les alloenzymes appuyaient fortement I'hypothese d'un ancetre commun pour le *P. radiata* et le *P. attenuata* ainsi qu'un historique de migration du sud vers le nord pour les trois especes.

J. Wu, K.V. Krutovskii,' and S.H. Strauss.' Department of Forest Science, Oregon State University, 321 Richardson Hall, Corvallis, OR 97331-5752, U.S.A.

'Permanent address: Laboratory of Population Genetics, N.I. Vavilov Institute of General Genetics, Russian Academy of Sciences, 117809 GSP-1. Moscow B-333, Russia. zAuthor to whom all correspondence should be addressed (e-mail: strauss@fsl.orst.edu).

Mots cles : alloenzymes, dominance, diversite genique, *Pinus attenuata, Pinus muricata, Pinus radiata,* phylogenie, RAPD.

[Traduit par la Redaction]

Introduction

Assessment of genetic variation in plant populations has many potential uses for evolutionists, breeders, and conservation biologists. The amount and distribution of genetic diversity reflect the interaction of various evolutionary processes such as gene flow, mutation, genetic drift, and natural selection (e.g., Wright 1978). Knowledge of genetic variation within a species can help plant breeders to collect and utilize different genetic resources and predict potential genetic gain in breeding programs. Implementation of conservation policies also requires a careful analysis of genetic variation in target species.

Allozymes had been the markers of choice for quantifying population structure at the molecular level for many years (e.g., Hamrick and Godt 1989). They usually exhibit simple Mendelian inheritance and codominant expression, facilitating genetic interpretations. However, allozymes have some well-known limitations, such as highly biased genomic sampling (only detecting protein-encoded genes and a fraction of all mutational events) and a small number of available loci which may not adequately represent genome-wide diversity.

Random amplified polymorphic DNA (RAPD) has been widely used to study population genetic structure and phylogenetic relationships (Aagaard 1997; Aagaard et al. 1998a, 1998b). It provides a more representative sample of the genome than allozymes, and a very large number of loci can be readily studied. Several reports have compared genetic diversity and population differentiation using allozyme and RAPD markers (e.g., Baruffi et al. 1995; Isabel et al. 1995; Peakall et al. 1995; Vicario et al. 1995; Lanner-Herrera et al. 1996a; le Corre et al. 1997; Latta and Mitton 1997; Aagaard et al. 1998a), and have usually found RAPDs to show somewhat greater polymorphism and (or) differentiation than allozymes (e.g., Liu and Furnier 1993; Szmidt et al. 1996). RAPD markers have also been utilized extensively in the reconstruction of phenetic and cladistic relationships (reviewed in Wolfe and Liston 1997). There has been general concordance among the results derived from RAPDs and other techniques (e.g., Graham and McNicol 1995; Lifante and Aguinagalde 1996; Marillia and Scoles 1996).

However, RAPDs also have some significant limitations. First, the dominant allelic expression of RAPDs precludes direct estimates of allele frequencies from diploid material, and can thus bias the calculation of genetic diversity and population differentiation (Lynch and Milligan 1994; Isabel et al. 1995; Szmidt et al. 1996). Secondly, lack of homology among comigrating RAPD fragments has been found in some species confounding phenotypic interpretation (e.g " Thormann et al. 1994; Rieseberg 1996). Finally, RAPDs of organellar origin have been reported in several studies (e.g., Lorenz et al. 1994; Aagaard et al. 1995, 1998b), which skews the comparison of genetic parameters between allozymes and RAPDs.

The goals of this study were to estimate nuclear DNA diversity using RAPD markers and to compare it to results from allozyme and organelle DNA. The specific objectives were to: (i) test the homology of comigrating RAPD fragments among populations and species; (ii) identify the frequency of organellar DNA markers among RAPD profiles; (iii) verify biparental, Mendelian inheritance for a sample of putative nuclear markers; (iv) compare the genetic diversity and differentiation of RAPD and allozyme markers; (v) simulate the effects that dominance and biallelism would have on allozyme markers to facilitate comparison to RAPD markers; and (vi) study phylogenetic relationships among populations and species using the two markers separately and together.

The California closed-cone pines (CCCP) are composed of three species; there is one interior species, *Pinus attenuata* Lemm. (knobcone pine), and two maritime species, *P. muricata* D. Don (bishop pine) and *P. radiata* D. Don (Monterey pine). *P. attenuata* grows on interior sites of southern Oregon and California as disjunct populations. The two other species are distributed discontinuously along the California coast and on four islands (Fig. 1). Many characteristics of these species have been intensively studied for population genetic variation, including morphology, secondary compound chemistry, allozyme, chloroplast DNA (cpDNA), and mitochondrial DNA (mtDNA) (reviewed in Millar *1986;* Millar et al. *1988;* Hong et al. *1993a, 1993b;* Strauss et al. *1993; Wu* et al. *1998),* providing an ideal taxon for comparisons among genetic markers. For example, a recent phylogenetic study using chloroplast DNA markers showed unexpected phylogenetic affinities and variable rates of evolution (Hong et al. *1993b).* This result will be reexamined in this study using the large set of nuclear DNA markers provided by RAPDs.

Materials and methods

Plant materials

Trees were sampled from essentially contiguous natural populations or from gene conservation and genetic test plantations as descri4ed in Hong et al. (1993a). Two different collections contributed to this study (Fig. 1). These were Ano Nuevo, Cambria and Guadalupe populations of *P. radiata*, the Sierra Nevada and Santa Ana populations of *P. attenuata*, and the Santa Cruz population of *P. muricata*, that were primarily collected by Hong et al. (1993a). The other populations were collected specifically for this study. For *P. attenuata*, the Klamath population was sampled over a 6.0-km transect adjacent to the Lakehead exit on U.S. Interstate 5, California (latitude 40°55', longitude 122°30'), and the Oakland population was sampled over a 2.6-km transect along Flicker Ridge adjacent to the town of Moraga in the hills east of Oakland, California (tat 37°50', long 122°30'). For *P. muricata*, the San Vicente population was sampled in several small scattered populations along a road north of San Vicente that goes out to the town of Erendira, Mexico (tat 31°15', long 116°30'); the Monterey population was sampled over one linear mile in the woods of Samuel F.B. Morse Botanical Reserve located south of Monterey, California (tat 36°40', long 121°50'); the Marin population was sampled over a 1.7-km transect about 5.1 km southwest from the town of Inverness, California (tat 38°08', long 122°45'); the Mendocino popula

Fig. 1. Distribution of *Pinus radiata, P. attenuata, and P. muricata* (Hong et al. 1993*a*) and the origins of sampled populations (*). All populations were studied for chloroplast DNA polymorphism by Hong et al. (1993*a*).
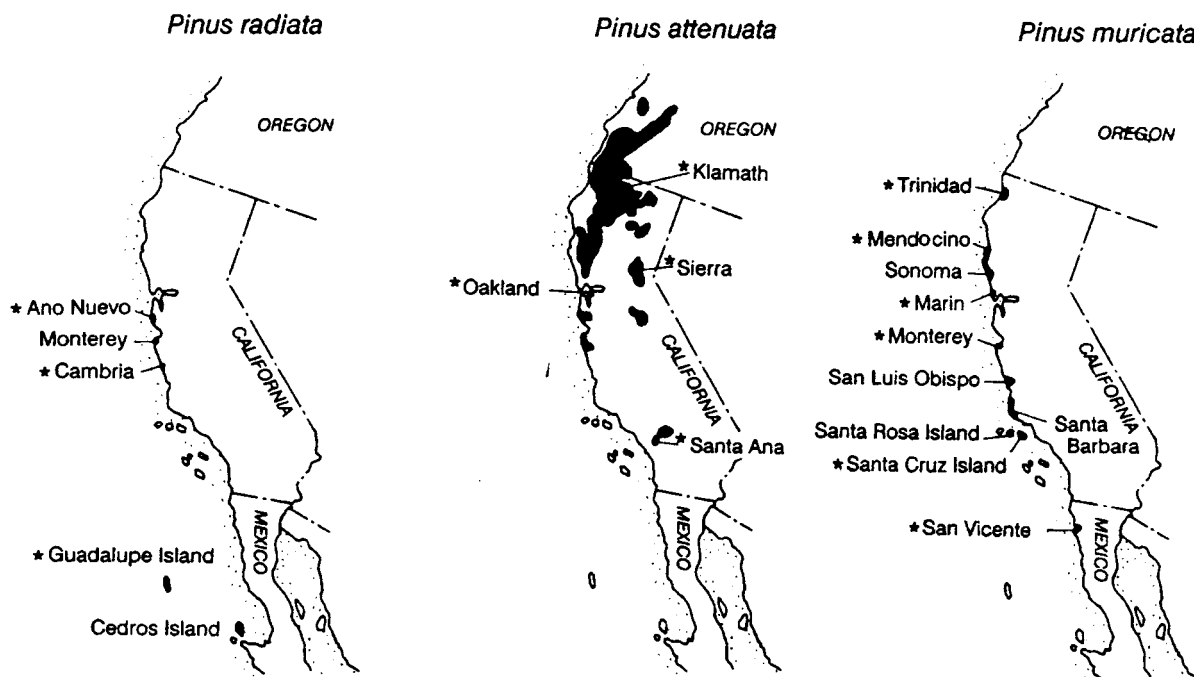
lion was sampled for 2.6 km on both sides of U.S. Highway 101, 7.5 km south of the town of Point Arena, California (lat 39°20', long 123°5Q'); and the Trinidad population was sampled over 1.7 linear kilometers along Fox Farm Road adjacent to the town of Trinidad, California Oat 41°05', long 124°10').

Allozyme data

We re-analyzed the allozyme data collected by Millar et al. (198R) as well as some additional data (R. Westfall, personal communication) for populations that were geographically matched to the populations used for the RAPD study (Fig. 1). The population identities and sample sizes were also shown in Table 1 for both markers. Allozyme data were not available for the San Vicente population.

DNA extraction and RAPD procedures

Total genomic DNA was extracted from megagametophytes and needles using a modified CTAB protocol described previously (Aagaard et al. 1995), followed by three phenol-chloroform purifications and a final ethanol precipitation. RAPD reactions were carried out in a volume *of* 25 pL with final concentrations *of* 10 mM Tris-HCl (pH 8.3); 1.8 mM MgC12; 100 pM of each dNTP, 0.2 1tM of primer; 1.3 mg/mL BSA; 2 ng template DNA; and one U of Taq polymerase (Stratagene, U.S.A.). Amplification was performed in a PTC-100 thermocycler (MJR Research, Inc., U.S.A.) using a period of 3 min of initial denaturation at 93°C, followed by 44 cycles of 1 min of denaturation at 93°C, 1 min of annealing at 37°C, 2 min of extension at 72°C, and a final extension step of 10 min at 72°C. All amplification products were electrophoresed using 2.0% agarose gels (TBE buffer pH 8.0) for 6 h. Gels were stained with 1 pg/mL of ethidium bromide for 30 min, then destained in distilled water for 2 h or more, and photographed under UV light.

A total of 500 primers from University of British Columbia (UBC, Vancouver, Canada) and Operon Technologies, Inc. (U.S.A.) were screened using one representative from each of three populations per species. Twenty-two primers that gave distinct, repeatable fragment patterns were selected for final analysis. These

were 17 UBC primers: #106 (CGTCTGCCCG), #137 (GGTCTCTCCC), #154 (TCCATGCCGT), #184 (CAAACGGCAC), #195 (GATCTCAGCG). #203 (CACGGCGAGT). #219 (GTGACCTCAG), #254 (CGCCCCCATT), #268 (AGGCCGCTTA), #299 (TGTCAGCGGT), #337 (TCCCGAACCG), #352 (CACAACGGGT), #429 (AAACCTGGAC), #485 (AGAATAGGGC), #503 (ATCGTCCAAC). #536 (GCCCCTCGTC), and #587 (GCTACTAACC); and 5 Operon primers; #OPAII (CAATCGCCGT), #OPC04 (CCGCATCTAC), #OPE17 (CTACTGCCGT), #OPG09 (CTGACGTCAC), and #OPY17 (GACGTGGTGA). Due to interpretation problems, primers UBC #154, UBC #203, and #OPA11 were scored only in *P. attenuata* and *P. radiata,* and #OPY17 was scored only in *P. radiata.* In order to avoid biasing the estimation of polymorphism, the selection of primers for full analysis was dependent only on the clearness and reproducibility of RAPD fragments, not on the level of polymorphism.

Homology of RAPD markers

We used Southern analysis to test homology among populations and species for 30 different RAPD markers. Some of these markers had an uneven distribution of fragment-staining intensities within or between species. The number of trees used for each test ranged from 30 to 150. RAPD fragments used as probes were excised individually from agarose gels, purified using the QIAquick gel extraction kit (QIAGEN, Inc., U.S.A.) or GENECLEAN kit (BIO101, Inc., U.S.A.), and then radioactively labeled with 3'-P by random hexamer labeling using Boehringer Mannheim's (Germany) random primed DNA labeling kit. RAPD fragments were first transferred to Zetabind nylon filters (CUNO, Inc.. U.S.A.) and then hybridized to isolated probes. Hybridization protocols followed Hong (1991), but were followed by final high-stringency washes (3 washes of 0.1 *x* SSC, 1 % SDS at 65°C).

Identification of organellar origin of RAPD markers

As described previously (Aagaard 1997), Southern hybridization of RAPD blots with organelle DNA was used to identify organellar DNA products among RAPD profiles. MtDNA was extracted from embryogenic suspension cultures of *P. radiata (J.* Aitken-Christie,

Table 1. Genetic diversity within populations and species of three California closed-cone pines based on RAPD and allozyme markers.

Species: population	RAPD					Allozyme				
	N [a]	A [b]	A_E [c]	P_{95} [d]	H_E [e]	N	A	A_E	P_{95}	H_E
P. attenuata	H_T [f] = 0.21±0.02 (98[g])					H_T = 0.13±0.02 (36)				
Klamath	27.9	1.47±0.05	1.26±0.04	46.94±5.07	0.15±0.02	50.0	1.83±0.14	1.22±0.05	58.33±8.33	0.15±0.03
Sierra Nevada	22.0	1.46±0.05	1.26±0.04	45.92±5.06	0.16±0.02	30.0	1.92±0.16	1.21±0.06	63.89±8.12	0.14±0.03
Oakland	29.0	1.67±0.05	1.32±0.03	67.35±4.76	0.20±0.02	10.0	1.25±0.07	1.11±0.04	25.00±7.32	0.07±0.03
Santa Ana	23.0	1.33±0.05	1.14±0.03	32.65±4.76	0.09±0.02	5.0	1.14±0.06	1.10±0.05	13.89±5.85	0.06±0.03
Average	25.5	1.48±0.07	1.24±0.04	48.21±7.16	0.15±0.02	23.8	1.55±0.20	1.16±0.03	40.28±12.29	0.11±0.02
P. radiata	H_T = 0.22±0.02 (91)					H_T = 0.15±0.02 (32)				
Año Nuevo	27.0	1.47±0.05	1.26±0.04	47.25±5.26	0.15±0.02	15.0	1.75±0.18	1.22±0.07	46.88±8.96	0.13±0.04
Cambria	28.0	1.48±0.05	1.30±0.04	48.35±5.27	0.18±0.02	15.0	1.81±0.17	1.24±0.07	50.00±8.98	0.14±0.03
Guadalupe	27.0	1.54±0.05	1.33±0.04	53.85±5.25	0.19±0.02	15.0	1.72±0.15	1.23±0.07	46.88±8.96	0.14±0.03
Average	27.3	1.50±0.02	1.30±0.02	49.82±2.04	0.17±0.01	15.0	1.76±0.03	1.23±0.00	47.92±1.04	0.14±0.04
P. muricata	H_T = 0.22±0.02 (98)					H_T = 0.15±0.02 (32)				
Trinidad	24.9	1.27±0.04	1.16±0.03	26.53±4.48	0.09±0.02	10.0	1.38±0.12	1.13±0.06	28.13±8.08	0.07±0.03
Mendocino	49.8	1.40±0.05	1.20±0.03	39.80±4.97	0.12±0.02	15.0	1.59±0.10	1.18±0.05	56.25±8.91	0.12±0.03
Marin	25.0	1.35±0.05	1.21±0.04	34.69±4.83	0.13±0.02	36.0	1.69±0.16	1.18±0.06	46.88±8.96	0.11±0.03
Monterey	25.0	1.39±0.05	1.21±0.03	38.78±4.95	0.13±0.02	12.0	1.28±0.08	1.17±0.06	28.13±8.08	0.11±0.03
Santa Cruz	25.0	1.53±0.05	1.27±0.04	53.06±5.07	0.16±0.02	29.0	1.53±0.12	1.27±0.09	43.75±8.91	0.14±0.04
San Vicente	49.7	1.52±0.05	1.29±0.04	52.04±5.07	0.17±0.02	–	–	–	–	–
Average [h]	33.2	1.41±0.04	1.22±0.02	40.82±4.17	0.13±0.01	–	–	–	–	–
Average [i]	29.9	1.39±0.04	1.21±0.02	38.57±4.31	0.13±0.01	20.4	1.49±0.07	1.19±0.02	40.63±5.50	0.11±0.01
Pooled species [j]	28.0	1.44	1.24	44.13	0.15	20.2	1.58	1.19	42.45	0.11

[a] N = average no. individuals analyzed per locus.
[b] A = average no. alleles per locus.
[c] A_E = average no. effective alleles per locus.
[d] P_{95} = percent of polymorphic loci (95% criteria).
[e] H_E = expected heterozygosity within populations or within species.
[f] H_T = total genetic diversity within species.
[g] No. loci studied in the species.
[h] Average over all six *P. muricata* populations.
[i] Average over five *P. muricata* populations shared by RAPD and allozyme studies (excluding San Vicente population).
[j] Genetic diversity parameters averaged over all populations in the three species.

unpublished) following the protocol of Aagaard (1997). The degree of enrichment of mtDNA was quantified using Southern hybridization: 0.2 ltg mtDNA extracted from embryogenic suspension cultures and 1 lig total genomic DNA extracted from needles were hybridized with a 32P-labeled mitochondrion-specific coxIII probe that had been amplified from *P. radiata* using a pair of universal primers (K. Krutovskii, unpublished data, Wu et al. 1998). Phosphorimage analysis (Phosphorimager model PSI-PC, Molecular Dynamics, U.S.A.) was used to quantify the ratio of hybridization signals of mtDNA and genomic DNA to estimate relative mtDNA enrichment.

For the cpDNA hybridizations, we probed with mixtures of cloned cpDNA fragments from *Pinus contorta* (Lidholm and Gustafsson 1991). Each hybridization used half of the 24 clones that together cover 92% of the chloroplast genome. Electrophoretically-separated RAPD products were blotted and sequentially hybridized with the mtDNA and cpDNA probes to determine the origin of RAPD markers. High (3 washes of 0.1 x SSC, 1 % SDS at 65°C) and low (3 washes of 2x SSC, 0.5% SDS at 65°C) stringency washes were used for the mtDNA and cpDNA probes, respectively. RAPD fragments that showed clear hybridization were excluded from nuclear DNA analyses.

Segregation analysis

Nuclear markers were also identified and confirmed by segregation analysis. Since mtDNA is normally maternally inherited and

cpDNA paternally inherited in pines (Hipkins et al. 1994), they should not segregate among haploid megagametophytes from seeds. It was therefore assumed that markers segregating among megagametophytes from the same trees were nuclear markers. We used seeds from two hybrid trees created from *P. attenuata* and *P. radiata,* three hybrid trees created from different populations of *P. radiata,* and 10 trees collected in eight natural populations for segregation analysis. A small number of megagametophytes (4-20) were used from each tree, because our goal was simply to identify segregating markers rather than test for quantitative conformance to Mendelian ratios. However, in order to conservatively select segregating markers, the homogeneity of segregation ratios among trees and the segregation ratio deviation for pooled samples taken from different trees (including hybrids and different species), were also tested using a G-test (Sokal and Rohlf 1981).

Genetic diversity analysis

RAPD markers show mainly dominant expression, usually with only a few percent showing codominance (e.g., Heun and Helentjaris 1993; Lu et al. 1995; Krutovskii et al. 1998). Thus, allele frequencies for RAPD loci were calculated from the phenotypic frequencies of null homozygotes assuming dominance and Hardy-Weinberg equilibrium. Wind-pollinated species, including the CCCP and other conifers, generally show only modest and transient departures from Hardy-Weinberg proportions (e.g., Plessas and Strauss 1986; Moran et al. 1988; Isabel et al. 1995).

Table 2. Estimates of population genetic diversity and differentiation in three California closed-cone pines based on three sets of data: all putative RAPD loci, loci selected based on segregation analysis, and loci selected using the polymorphism criteria of Lynch and Milligan (1994).

Pine species	$N_L{}^a$	A^b	$A_E{}^c$	$P_{95}{}^d$	$H_E{}^e$	$G_{ST}{}^f$
All loci						
P. attenuata	98	1.48±0.07	1.24±0.04	48.21±7.16	0.15±0.02	0.36
P. radiata	91	1.50±0.02	1.30±0.02	49.82±2.04	0.17±0.01	0.26
P. muricata	98	1.41±0.04	1.22±0.02	40.82±4.17	0.13±0.01	0.45
Loci selected based on segregation analysis						
P. attenuata	42	1.59±0.07	1.32±0.04	59.38±6.38	0.19±0.02	0.36
P. radiata	44	1.62±0.02	1.38±0.02	61.73±2.47	0.22±0.01	0.22
P. muricata	50	1.51±0.04	1.30±0.02	51.15±4.22	0.18±0.02	0.41
Loci selected using polymorphism criteria ^{g}						
P. attenuata	27	1.64±0.09	1.22±0.04	64.81±8.88	0.17±0.03	0.17
P. radiata	31	1.78±0.05	1.38±0.06	78.49±4.69	0.23±0.03	0.17
P. muricata	43	1.57±0.05	1.27±0.02	56.59±5.19	0.16±0.01	0.27

Note: All parameters of genetic diversity were averaged over all populations within each species for each data set.

$^a N_L$ = no. loci.

$^b A$ = no. alleles per locus.

$^c A_E$ = no. effective alleles per locus.

$^d P_{95}$ = percent of polymorphic loci (0.95 standard).

$^e H_E$ = Expected heterozygosity within populations.

$^f G_{ST}$ = Nei's G_{ST} (1986) unbiased for small sample size and population number.

^{g}Loci that meet the criterion recommended by Lynch and Milligan (1994) that number of null homozygotes is not less than three at each locus selected for analysis within each population.

Because of dominance, corrected allele frequencies (g) were used for estimation of gene diversity parameters (Lynch and Milligan 1994): q = $x^{1/2}[1 - Var(x)/8x^2]^{-1}$, where Var(x) = x(1 - x)/N, and x is the proportion of the N-sampled individuals that do not have amplified RAPD fragments (null-allele homozygous phenotype). Mean (A) and effective mean (AE) number of alleles per locus, number of private alleles (U) for each population, percentage of polymorphic loci (95% criterion, P95), Nei's (1978) unbiased expected heterozygosity (HE), and population differentiation (GsT), unbiased for sample size and population number (Nei 1986) were calculated using GENESTAT-PC 3.3 (Lewis 1994). To facilitate comparison of our data with that in a review of estimates of plant diversity based on allozymes, we also calculated GsT by averaging over polymorphic loci (Hamrick and Godt 1989). In addition, Wright's (1931) FsT, Weir and Cockerham's (1984) θ, and Lynch and Milligan's (1994) FsT were also calculated using Black's RAPDFST program (1996) for comparison with Nei's GsT. Hierarchical F-statistics describing subdivision between populations within species and between species were analyzed using the siosYS-i version 1.7 (Swofford and Selander 1989) that we recompiled for large sample sizes.

RAPD phenotypic (presence-absence) data were used to produce F-statistic analogs called (D-statistics (e.g., (DST = variance among populations/total variance within species) using the analysis of molecular variance procedure (WINAMOVA, Excoffier et al. 1992; Excoffier 1993). For each tree we created a binary vector of 1 s (presence of a band) and Os (absence of a band) for RAPD loci. The proportion of shared fragments was calculated for each possible pairwise comparison according to the formula proposed by Nei and Li (1985): S (similarity) = 2NAB/(NA + NB) (where NAB is the number of bands shared by individuals A and B; NA and NB are the number of bands in individuals A and B, respectively); the distance index is D = 1 - S. All similarity and distance indices were obtained using the RAPDPLOT program (Black 1996). An analysis of molecular variance was then performed on the resultant matrix for partitioning the total RAPD variation into the within- and among population variance components, and for calculating (DST.

Estimation of genetic parameters was performed for all RAPD loci, and again, separately for two selected sets of loci. First, we excluded those loci with segregation heterogeneity or distortion, and where segregation 3:1 (presence-absence) was only observed in 4 megagametophytes from a single tree. In addition, to limit bias resulting from the dominant biallelic properties of RAPDs, Lynch and Milligan (1994) suggested the restriction of RAPD analysis to loci with a null homozygote frequency greater than 3/N (N =the sample size of a population).

Allozyme data were analyzed using both allele frequencies listed in Millar et al. (1988) and unpublished data obtained from the authors (R. Westfall, personal communication). Allele frequencies at 32 loci were taken from Table 2 in Millar et al. (1988) for P. muricata and P. radiata and the Klamath and Sierra populations of P. attenuata, and from the original data files provided by R. Westfall for the Oakland and Santa Ana populations of P. attenuata. Allele frequencies of four additional loci (Aldl, Mnr3, Ugp3, and Ugp4) for P. attenuata were also inferred from the original data files.

Simulation analysis

We used a computer simulation (Krutovskii et al. 1999) to determine how the dominance and biallelism could affect the estimation of genetic parameters, and thus to aid our interpretation of comparisons between RAPDs and allozymes. The simulation also allowed us to determine how sample size affects the estimation of genetic parameters in the codominant and dominant data sets. We estimated genetic parameters from our CCCP allozyme data using both the codominant multiallelic data set and simulated data set where the allozyme genotypes were transformed into dominant recessive biallelic genotypes, thus imitating RAPD data. Allozyme allele frequencies of 12 populations of three species obtained from Millar et al. (1988) were first used to generate 12 populations of 1000 individuals, each having genotypes that maintained the original allele frequencies within each population. A total of 200 subpopulations of n individuals were then sampled with replacement from each of the 12 populations. For the codominant data set, ge

netic parameters *(Hs,* HT. and GST) were calculated from allele frequencies inferred directly from each set of 200 subpopulations. For the dominance analysis, one of the alleles at each locus was randomly selected as a dominant allele and others pooled as a null recessive allele. The genotype frequency of null homozygotes was then used to calculate null and dominant allele frequencies assuming Hardy-Weinberg equilibrium, and diversity and differentiation parameters were estimated based on these frequencies. Means and variances of *Hs,* HT. and GST were then calculated from the 200 sampled subpopulations.

Phylogenetic analysis

Nei's (1972) genetic distances between both populations and species were calculated from allele frequencies for both RAPD and allozyme markers using the BIOSYS-1 program (Swofford and Selander 1989) and GENEDIST program in the PHYLIP package (Felsenstein 1995). In addition, the pairwise Manhattan distances between populations (Prevosti distance in Wright 1978) were computed on the basis of dissimilarities in RAPD phenotypes (presence-absence of an amplified fragment) using the RAPDDIST program (Black 1996). A distance matrix or a set of matrices generated by bootstrapping over loci were produced using the RAPDDIST or RAPDBOOT programs (Black 1996). The distance matrices based on both allele frequencies and phenotypes were then used to produce phylogenetic trees using Neighbor-Joining Tree (NJTREE) and UPGMA clustering programs in PHYLIP (Felsenstein 1995). The SEQBOOT program was also used to generate bootstrap data sets, and the majority-rule and strict consensus trees were generated from bootstrap trees, using the program CONSENSE in the PHYLIP package. The phylogenetic trees were viewed and drawn using TREEVIEW program (Page 1996). We also combined both allozyme and RAPD loci for phylogenetic analysis.

Results

A total of **94-100** fragments were scored in each of *P. attenuata*, *P. muricata*, and *P. radiata*. More than 70% of the scored fragments showed polymorphism across populations within species. The percentage of polymorphic loci for each species ranged from 70% to **85%,** with the highest percentage for *P. muricata*. Seventy-six fragments were scored as homologous bands among all three species; one-third of these was tested for homology by Southern blot hybridization as described below.

Homology of RAPD markers

Southern blots confirmed homology for **28** of **30** (93%) randomly chosen RAPD markers tested by hybridizing comigrating DNA fragments to blots of DNA from different populations and species. For **2** of the 30 markers, some comigrating fragments did not give a hybridization signal. In addition to these 30 comigrating fragments, several other markers were also tested for hybridization because of their inconsistent band phenotypes. In two cases, the fragments comigrated but had very different intensities among individuals within populations; the homology tests showed that the fragments were nonhomologous, and controlled by at least two genetic loci. In two other cases, hybridization revealed that fragments with slight mobility differences were homologous and presumably controlled by two codominant alleles of the same locus. However, in two other cases, fragments of slightly different sizes were found to be completely nonhomologous. Thus, except in those cases with inconsistent band phenotypes, even when studying different species it ap

pears that RAPD markers can be treated as homologous with a modest error rate **(5-7%).**

Identification of organelle DNA among RAPD profiles

Quantification of mtDNA enrichment indicated that mtDNA extracted from embryogenic cultures was more than **50** times enriched for mtDNA compared to total genomic DNA (data not shown). However, of all of the RAPD fragments scored, only one fragment amplified by primer UBC **#268** hybridized strongly with the enriched mtDNA. Two fragments amplified by primer UBC **#337** were identified to be of putative chloroplast origin. These three fragments were excluded from further RAPD data analysis. Thus, **91** RAPD markers for P. radiata and **98** markers for each of P. attenuata and P. muricata pine were retained as putative nuclear loci.

Segregation analysis

To establish a data set of nuclear loci showing evidence of Mendelian segregation, we analyzed inheritance in several trees. A total of **48, 54,** and **58** loci were segregating among megagametophytes in *P. attenuata*, *P. radiata*, and *P. muricata*, respectively, including several hybrid trees. Altogether, **92** loci and **180** tree-locus combinations were tested in three species. Of these loci, seven loci showed statistically significant (P <_ 0.05) segregation distortion in pooled samples and were excluded from our segregation data set. Of 52 loci that had segregation data for more than one tree and thus provided an opportunity to test segregation ratio homogeneity among trees, three loci demonstrated significant segregation ratio heterogeneity. Those loci were also excluded from the data set. Seven loci that were analyzed in only four megagametophytes from a single tree and that showed only one "absence" phenotype were also discarded because the null phenotype could have resulted from failed amplification. Thus, a total of 17 loci were excluded for the three species combined, and 42, 44, and 50 loci were retained for estimation of genetic parameters for *P. attenuata, P. radiata,* and *P. muricata,* respectively. The overall ratio (pooled for all trees and loci) of presence vs. absence of bands was 629:653, thus showing no tendency for an excess of null phenotypes because of poor amplification (x21:1 = 0.45, P > 0.05). Deviations from 1:1 segregation ratio were apparently stochastic; the ratio favored the dominant allele at 36 loci and favored the null allele at 38 loci, while the remaining 18 loci had an exact 1:1 ratio.

RAPD genetic diversity and differentiation

Different locus sets

Estimates of genetic diversity within populations derived from the set of normally segregating loci were slightly higher than those based on all RAPD loci (Table 2). However, differentiation among populations estimated from only segregating loci was nearly the same as that obtained from all RAPD loci. When many monomorphic and weakly polymorphic loci were excluded from the data set of all loci following Lynch and Milligan's (1994) recommendation, genetic diversities were higher and GST values decreased substantially. Using their criterion, only one-third to approx

imately one-half of all loci could be retained for the different species (Table 2).

Diversitv

Based on the data set with all loci (Table 1), among populations the average number of alleles per locus (A) ranged from 1.3 to 1.7, the effective number of alleles per locus (AE) from 1.2 to 1.3, the percent of polymorphic loci $(P95)$ from 26.5% to 67.4%, and the expected heterozygosity (HE) from 0.09 to 0.20. Among the three species, P. *muricata* had the lowest average number of alleles per locus (1.4), percent of polymorphic loci (38.6%) and expected heterozygosity (0.13). Total genetic diversities within species were very similar among the three species $(HT = 0.21\text{-}0.22$, Table 1).

Differentiation

Allele frequencies were highly variable at many loci among populations within each species (data not shown). For example, there were 14 private alleles for *P. dttenuata* (unique to different populations); 1I of those 14 alleles belonged to the Oakland population. In P. *radiata* 15 private alleles were detected, however, only one marker was strictly population specific in our sample (monomorphic in one population, absent elsewhere, UBC primer #219, 0.52 kb). Nine of those 15 alleles were present in the Guadalupe Island population. Similar results were observed with *P. muricata;* 15 private alleles were detected, most of them belonging to the southern California mainland population (5 alleles, San Vicente) and an island population (6 alleles, Santa Cruz). One marker (UBC primer #268, 1.82 kb) was specific to the San Vicente population sample.

Populations within each species generally showed high differentiation from one another. In contrast to intrapopulation diversity, *P. muricata* revealed the highest population differentiation due to its large regional differences $(GST = 0.45$ compared to 0.36 for P. *attenuata* and 0.26 for *P. radiata,* Table 2). Several different methods of analysis gave congruent trends of differentiation, although their absolute values sometimes differed substantially. Nei's (1986) GST values were nearly the same as Lynch and Milligan's (I 994) FST and Weir and Cockerham's (1984) 0 (Table 3). Hamrick and Godt's (1989) GST and Wright's (1931) FST values were also very similar, but both were lower than Nei's (1986) $GST.$ The F-statistic analog obtained from WINAMOVA analysis, $(DST,$ was the highest among all population differentiation parameters. When the species complex was considered as a unit, as expected the differentiation among all populations was substantially increased $(GST = 0.60$, Table 3), and differentiation among species in the complex was also high $(GST = 0.42)$. When all species and populations were analyzed hierarchically using BIOSYS-1 (Swofford and Selander 1989), FST among species was decreased to 0.31 (F-coefficients were 0.39 for populations within species and 0.58 for populations relative to the total complex).

RAPDs, the estimates of genetic diversity $(HE = 0.06\text{-}0.15)$ were highly variable between different populations. However, unlike RAPDs, the population sample sizes were also highly variable. Averaged over all populations, the estimates of A, $AE, P95,$ and HE were 1.58, 1.19, 42.45%, and 0.11, respectively. Similar to RAPDs, allozyme diversity was highest in the populations of P. *radiata.* Total diversities $(HT = 0.13\text{-}0.15)$ within species were very similar among the three species.

Differentiation

Allele frequencies varied greatly among species and among populations within species (data not shown). For example, allele *Lap2100* frequencies were high in *P. attenuata,* low in *P. radiata,* and moderate in P. *muricata. Got] 114* was present only in P. *muricata,* while *Got' 14* and *Got 157* were observed only in *P. radiata.* All three species had geographic trends in allelic variation. Nei's (1986) GST was 0.22 for *P. attenuata,* 0.12 for P. *radiata,* and 0.32 for *P. muricata* (Table 3). Differentiation among all populations in the three species complex was 0.44 and the differentiation among species was 0.31. When the complex was analyzed hierarchically using the BIOSYS-1 program (Swofford and Selander 1989), FST among species was reduced to 0.23, while FST coefficients were 0.25 for populations within species and 0.43 for populations relative to the total complex (Table 3).

Simulation analysis

Estimates of population gene diversity and differentiation were obtained from allozyme allele frequencies inferred from both the codominant multiallelic data set and the transformed dominant biallelic data set in simulated samples of different population sizes (Fig. 2). Diversities within populations and species were substantially lower for the simulated dominant data set than for the codominant data set over a large range of sample sizes. At our average sample size of approximately 25, the dominant markers were decreased in their absolute diversity values (Hs) by 4-6% within populations for each species. The simulated bias in population differentiation values (GST) for each species depended strongly on sample size, decreasing as sample size increased. At our sample size of 25, a 5-10% upward change of absolute GST values was predicted (Fig. 2).

Comparison of genetic diversity and differentiation between RAPDs and allozymes

Both diversity and differentiation for RAPDs were higher than for allozymes (Tables 1 and 3). Averaged within population diversity over all populations for RAPDs was nearly 30% higher than that for allozymes (0.15 vs. 0.11, Table 1). Total diversity within species averaged 0.21 over three species for RAPDs, more than one-third greater than for allozymes (0.13). Population differentiation within each species $(GST = 0.26\text{-}0.41)$ for RAPDs was substantially larger than observed for allozymes $(GST = 0.12\text{-}0.32$, Table 3). Differentiation among species also differed between RAPDs and allozymes (0.42 vs. 0.31, respectively).

When estimates of within-population diversity (Hs) for RAPD markers were compared to those based on the dominant data set that we simulated based on allozymes, the diversities were more than twice as high as those of trans

Allozyme genetic diversity and differentiation

Diversitv

Allozyme diversity within populations (HE) averaged for each species ranged from 0.11 to 0.14 (Table 1). Similar to

Table 3. Parameters of genetic differentiation among populations and species of three California closed-cone pines based on RAPD, allozyme, mitochondrial (mtDNA), and chloroplast (cpDNA) RFLP markers.

Parameter	P. attenuata	P. radiata	P. muricata[a]	P. muricata[b]	Species	Populations in all species
RAPDs						
G_{ST} [c]	0.36	0.26	0.45	0.41	0.42	0.60
G_{ST} [d]	0.24±0.03	0.18±0.02	0.34±0.03	0.29±0.03		0.51±0.03
F_{ST} [e]	0.23±0.03	0.18±0.02	0.35±0.03	0.29±0.03		
θ [f]	0.37±0.01	0.26±0.01	0.45±0.01	0.41±0.01		
F_{ST} [g]	0.37±0.03	0.26±0.03	0.45±0.03	0.42±0.03		
Φ_{ST} [h]	0.41	0.32	0.52	0.44		
F_{ST} [i]					0.31	0.58
Allozymes						
G_{ST} [c]	0.22	0.12		0.32	0.31	0.44
G_{ST} [d]	0.13±0.03	0.08±0.02		0.17±0.03		0.28±0.04
F_{ST} [i]					0.23	0.43
mtDNA [j]						
G_{ST}[c]	0.79	0.79	0.75		0.22	0.78
θ [k]	0.78±0.13	0.79±0.14	0.77±0.06		0.21	0.78
Φ_{ST} [h]	0.90	0.91	0.95		0.26	0.91
cpDNA [l]						
G_{ST} [c]	–[m]	–	0.88		0.84	0.93
θ [k]	–	–	0.82		0.82	0.93
N_{ST} [n]	–	–	0.98		0.88	0.99

[a]Based on all *P. muricata* populations.
[b]Based on five *P. muricata* populations (without the San Vicente population) that were shared by RAPD and allozyme studies.
[c]G_{ST} (Nei 1986) unbiased for both sample size and population no. and calculated using GENESTAT-PC (Lewis 1994).
[d]G_{ST} (Nei 1986) averaged over polymorphic loci (Hamrick and Godt 1989) using GENESTAT-PC (Lewis 1994).
[e]F_{ST} (Wright 1931) averaged over polymorphic loci using RAPDFST (Black 1996).
[f]θ (Weir and Cockerham 1984) estimated using RAPDFST (Black 1996).
[g]F_{ST} (Lynch and Milligan 1994) calculated using RAPDFST (Black 1996).
[h]Φ_{ST} (Excoffier et al. 1992) calculated using WINAMOVA (Excoffier 1993).
[i]F_{ST} (Wright 1978) calculated when all species and populations were analyzed hierarchically using BIOSYS-1 program (Swofford and Selander 1989).
[j]Complete mtDNA data set in Wu et al. (1998).
[k]θ (Weir and Cockerham 1984) calculated using Genetic Data Analysis program (GDA, Lewis and Zaykin 1996).
[l]CpDNA restriction site data were obtained from Hong et al. (1993a).
[m]– = No polymorphism to allow estimates.
[n]N_{ST} calculated using HAPLO program (Lynch and Crease 1990).

formed allozymes at N = 25 (Table 1; Fig. 2). However, observed RAPD population differentiation (GsT) was similar to that based on the transformed data set. RAPD estimates were either moderately higher than those for the transformed allozymes (GsT = 0.36 vs. 0.30 for *P. attenuata*, 0.26 vs, 0.19 for *P. radiata*), or slightly lower (0.41 vs. 0.44 in *P. muricata*). Thus, unlike differentiation values (GsT), the difference in the diversity values (Hs and *HT*) observed between allozyme and RAPD markers was not in agreement with the simulated estimates. Diversity estimates were higher with RAPDs, though they were expected to be dramatically lower.

Phylogenetic relationships

UPGMA and Neighbor-Joining (NJ) cluster analyses were carried out to estimate phylogenetic relationships among the 12 (allozymes) or 13 (RAPDs) populations using genetic distance matrices. For RAPDs, phenogram topologies were very similar regardless of whether distance matrices were derived from allele frequencies or direct phenotypic data (presence-absence), except that the *P. attenuata* subgroup and the group including both *P. attenuata* and *P. radiata* spe

cies from NJ were better supported by bootstrap values using the latter data set (bootstrap values of 93% and 94%, respectively).

In both UPGMA and NJ analyses of RAPDs, 13 populati9ns were clustered into three main groups, clearly corresponding to the three species. *P. attenuata and P. radiata* were more similar to one another (D = 0.20, Table 4) than were *P. attenuata* to *P. muricata* (D = 0.27) or P. radiata to *P. muricata* (D = 0.32), and they were considered a monophyletic group at high confidence in UPGMA (100%, data not shown) and modest confidence in NJ (73%) trees.

Among populations, the mainland populations (Ano Nuevo and Cumbria) of P. radiata were closer to each other than to the island population (Guadalupe) based on RAPD data. The Santa Ana population of *P. attenuata* was most distinct from the other, northern populations. The three regional races of *P. muricata* formed proximal phylogenetic groups (South = San Vicente and Santa Cruz, Central = Marin and Monterey, North = Mendocino and Trinidad). In all cases the most southern populations) of each species were most basal (Santa Ana, Guadalupe, and San Vicente). In the genetically distinct populations of *P. muricata*, the

Fig. 2. Genetic diversity (H_S) and differentiation (G_{ST}, Nei 1986) values averaged over populations of each California closed-cone pine species for codominant multiallelic allozyme and dominant biallelic markers simulated in the samples of different sizes. Standard deviations (shown as error bars) were calculated from the variance among 200 resamples simulated for each species. Observed RAPD values are also shown as a star.
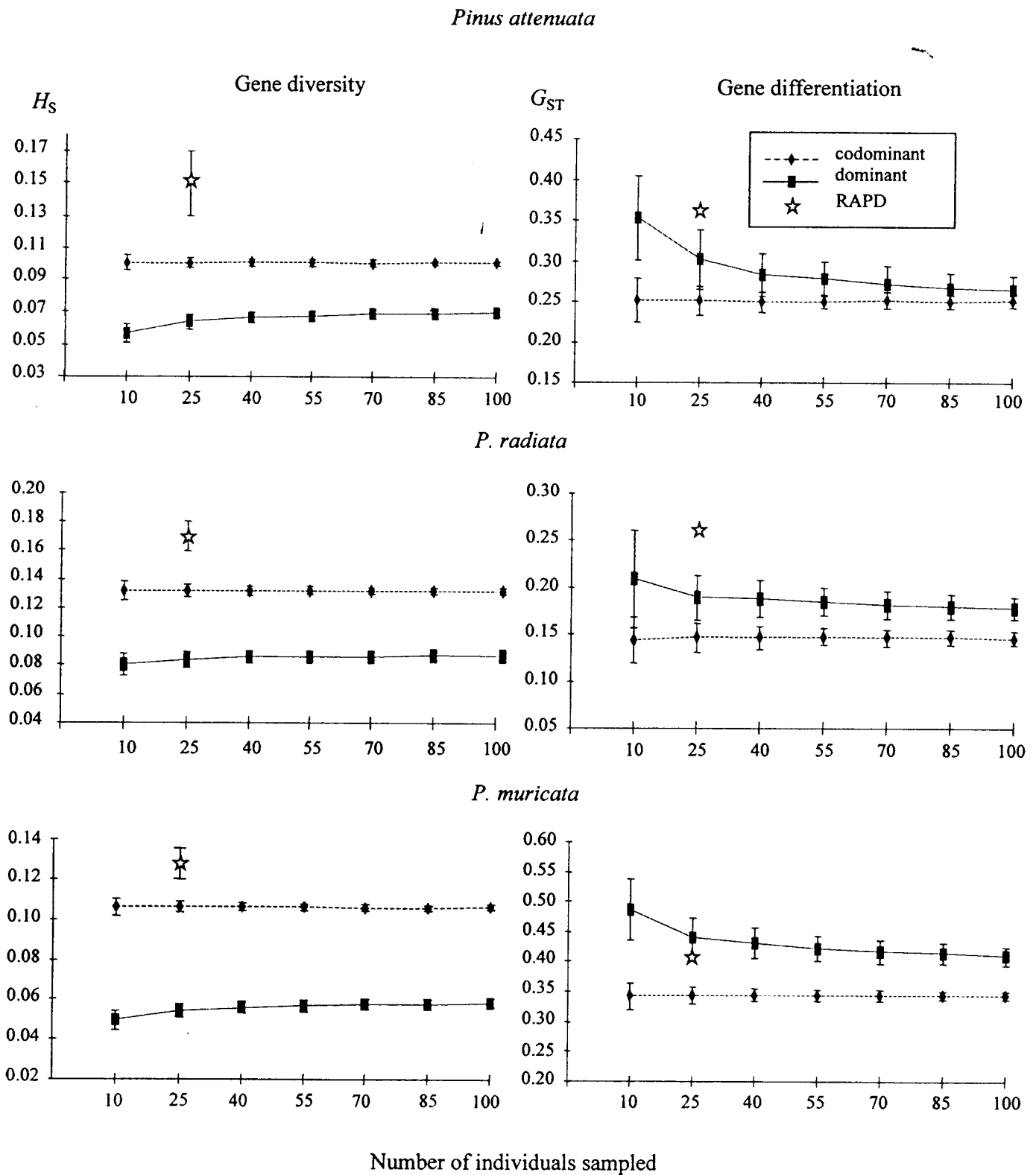

Pinus attenuata

P. radiata

P. muricata

Number of individuals sampled

Table 4. Nei's (1978) unbiased genetic distance within and among species based on allele frequencies of allozyme and RAPD markers. The values in the parentheses are the ranges of genetic distance between populations within species.

Species	P. attenuata	P. radiata	P. muricata
RAPDs (76 loci)			
P. attenuata	0.070 (0.031–0.095)		
P. radiata	0.204 (0.178–0.233)	0.050 (0.029–0.067)	
P. muricata	0.271 (0.166–0.337)	0.319 (0.231–0.381)	0.111 (0.037–0.209)
Allozymes (32 loci)			
P. attenuata	0.035 (0.009–0.054)		
P. radiata	0.144 (0.089–0.222)	0.021 (0.013–0.030)	
P. muricata	0.143 (0.071–0.249)	0.122 (0.077–0.179)	0.062 (0.034–0.112)

populations were arrayed in a clear south (basal) to north (derived) phylogenetic sequence. UPGMA and NJ phenograms based on allozymes had many minor differences from those based on RAPDs, though all were supported by low to modest bootstrap values in the NJ tree (Figs. 3 and 4).

When the RAPD loci were combined with the allozyme loci for phylogenetic analysis, both UPGMA and NJ phenograms (Fig. 4) were similar to those based only on RAPD loci (Fig. 3). The main differences were that most clusters were supported at higher bootstrap levels. Most notably, *P. attenuatta* and *P. radiata* formed a monophyletic group at high confidence (96-100%) in NJ and UPGMA trees, respectively.

Discussion

Homology of comigrating RAPD fragments

A large majority of comigrating RAPD markers among CCCP were homologous, so similarity of fragment size appears to be a reasonably reliable predictor of homology, at least among closely related populations or species. These results are similar to those recently published by Rieseberg (1996) who found that 200 of 220 comigrating fragments (91%) were homologous among three species of sunflower *(Helianthus)* based on Southern hybridization and (or) restriction enzyme analysis. Williams et al-. (1993) also found nine of 10 comigrating fragments (90%) were homologous among several *Glycine* species from Southern hybridization. Lanner-Herrera et al. (1996b) observed an even higher degree of homology in 10 wild *Brassica* species, as only 3 of 250 RAPD comigrating markers tested did not crosshybridize as expected. A lower level of homology, however, was reported by Thormann et al. (1994), who revealed that 3 of IS comigrating fragments were not homologous among *six Brassica* species, although all fragments showed the expected hybridization within species. Thus, it appears that most RAPD markers can be used for analysis of closely related species. However, homology tests are very helpful for excluding some doubtful RAPD fragments such as those with large differences in intensity or slight differences in mobility. Without these tests, it is advisable to exclude any markers whose phenotypes are not identical.

Genomic origin of RAPD fragments

Based on Southern hybridization with organellar DNA as probes, we found only one mtDNA and two cpDNA frag

ments among more than 200 tested in these pines. This is consistent with the results obtained by Lug et al. (1995) and Heun and Helentjaris (1993), who didn't detect any RAPD fragments of cytoplasmic origin in *Pinus sylvestris* L., and *Zea mays* by segregation analysis of Fi progeny, respectively. However, Lorenz et al. (1994) found that a number of fragments in "genomic" RAPD patterns of *Beta vulgaris* L. originated from mtDNA, and to a lesser extent from cpDNA. In *Brassica,* results of Southern hybridization indicated that among the 142 scored fragments, seven (4.9%) were of mitochondrial origin, and one (0.7%) was of chloroplast origin (Thormann et al. 1994). The most extreme case occurred in Douglas-fir *(Pseudotsuga menziessi* (Mirb.) Franco), in which 45% of scored RAPD markers showed maternal inheritance and were considered to be mtDNAderived fragments (Aagaard et al. 1995). However, based on Southern analysis, Aagaard et al. (1998b) recently reestimated the mtDNA value to be 16%, with 3% of chloroplast origin. They surmised that the higher proportion of RAPD bands found in the 1995 study may have been a result of their selective study of those fragments that showed distinct inter-varietal differentiation and strong band intensity.

Variation in the frequency of cytoplasmic fragments in RAPD profiles among species could be explained by variable contributions of organellar genomes to total cellular DNA among species (Williams et al. 1990; Heun and Helentjaris 1993). This can result from varying cellular states of development at the time of DNA extraction and variation in relative genome sizes. In addition, there will be varying numbers of inverted repeat sequences (necessary for single-primer amplifications) in cytoplasmic genomes relative to nuclear genomes (Lorenz et al. 1994). This would be expected to range widely, depending greatly on relative genome sizes and complexity. Finally, values based on segregation from intraspecific or intra-varietal crosses are expected to provide only a minimum estimate because of the modest levels of intrapopulation polymorphism expected for plant organelle genomes as compared to nuclear genomes.

Segregation analysis

To eliminate organellar and unreliable fragments, we studied segregation among haploid megagametophytes. More than 40 loci segregated among megagametophytes of several hybrid and (or) natural cones for each species. Lu et al. (1995) studied 8 parents and 80 Fl progeny of Pinus

Fig. 3. Neighbor Joining trees based on Nei's (1972) genetic distance among populations of three California closed-cone pine species estimated from allele frequencies of 76 RAPD and 32 allozyme loci. Numbers near relevant nodes are percentage of 200 bootstrap replicates.
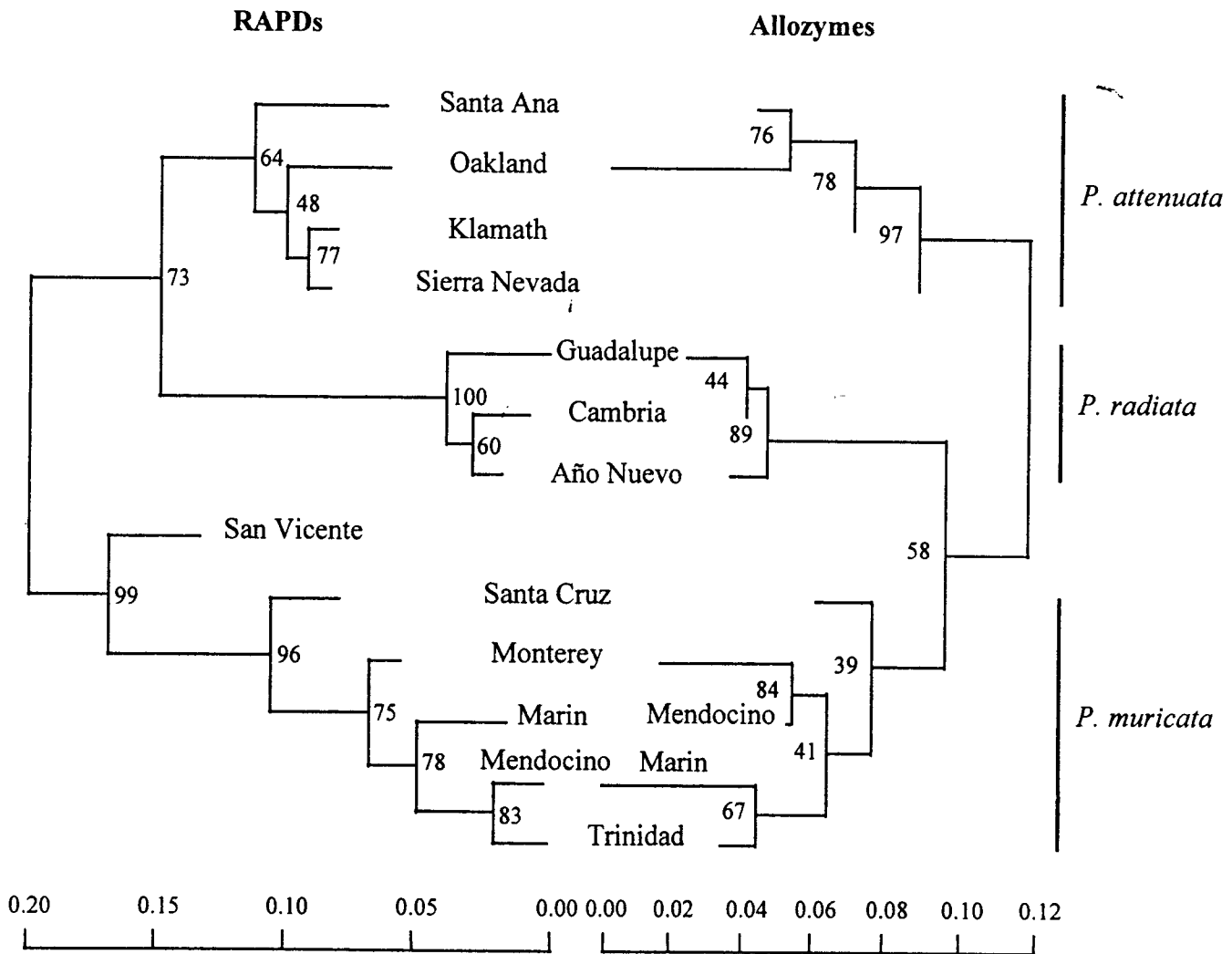
sylvestris, and found that the segregation of all polymorphic fragments was consistent with the biparental, diploid mode of inheritance expected for a dominant trait. Similar results have also been reported in other organisms (e.g., Heun and Helentjaris 1993; Kazan et al. 1993; Krutovskii et al. 1998). However, the proportion of Mendelian loci expressed as a fraction of the total number of amplification products can be as low as 33% *(Picea abies,* Bucci and Menozzi 1993). Because of the rarity of suitable trees or crosses, selection of loci based on segregation analysis is difficult for weakly polymorphic or monomorphic loci. Therefore, segregating loci will rarely represent a random set of RAPD loci, and can bias the estimation of genetic diversity.

RAPDs vs. allozymes

The genetic data obtained at both RAPD and allozyme loci allowed us to make a direct comparison of genetic diversity within populations, and differentiation among populations for these two types of nuclear loci. RAPD markers generally detected more intrapopulation diversity than did allozymes in terms of effective number of alleles per locus,

proportion of polymorphic loci, and expected heterozygosity, but not for number of alleles per locus. Allozymes are expected to have a larger number of alleles per locus because many allozyme loci are multiallelic, although in most cases only one or two alleles are common and others are rare. Population and species differentiation were also significantly higher for RAPDs than for allozymes.

However, dominance of RAPDs is expected to bias estimates of null-allele frequency when the null homozygotes are rare within populations (Lynch and Milligan 1994). For example, only one null homozygote per 25 individuals is expected, even when null-allele frequency is 0.20. Such rare homozygotes can go easily undetected in small samples, biasing frequency estimates downward and potentially affecting estimates of genetic diversity and differentiation. When there is an excess of heterozygotes and Hardy-Weinberg equilibrium is violated, the bias will be even higher. Szmidt et al. (1996) inferred allele frequencies of RAPDs directly for maternal trees from segregation patterns in haploid megagametophytes, and indirectly estimated allele frequencies based on frequency of null homozygotes among diploid

Fig. 4. Neighbor-Joining and UPGMA trees based on Nei's (1972) genetic distance among populations of three California closed-cone pine species estimated from allele frequencies of both 76 RAPD and 32 allozyme loci. Numbers near relevant nodes are percentage of 200 bootstrap replicates.
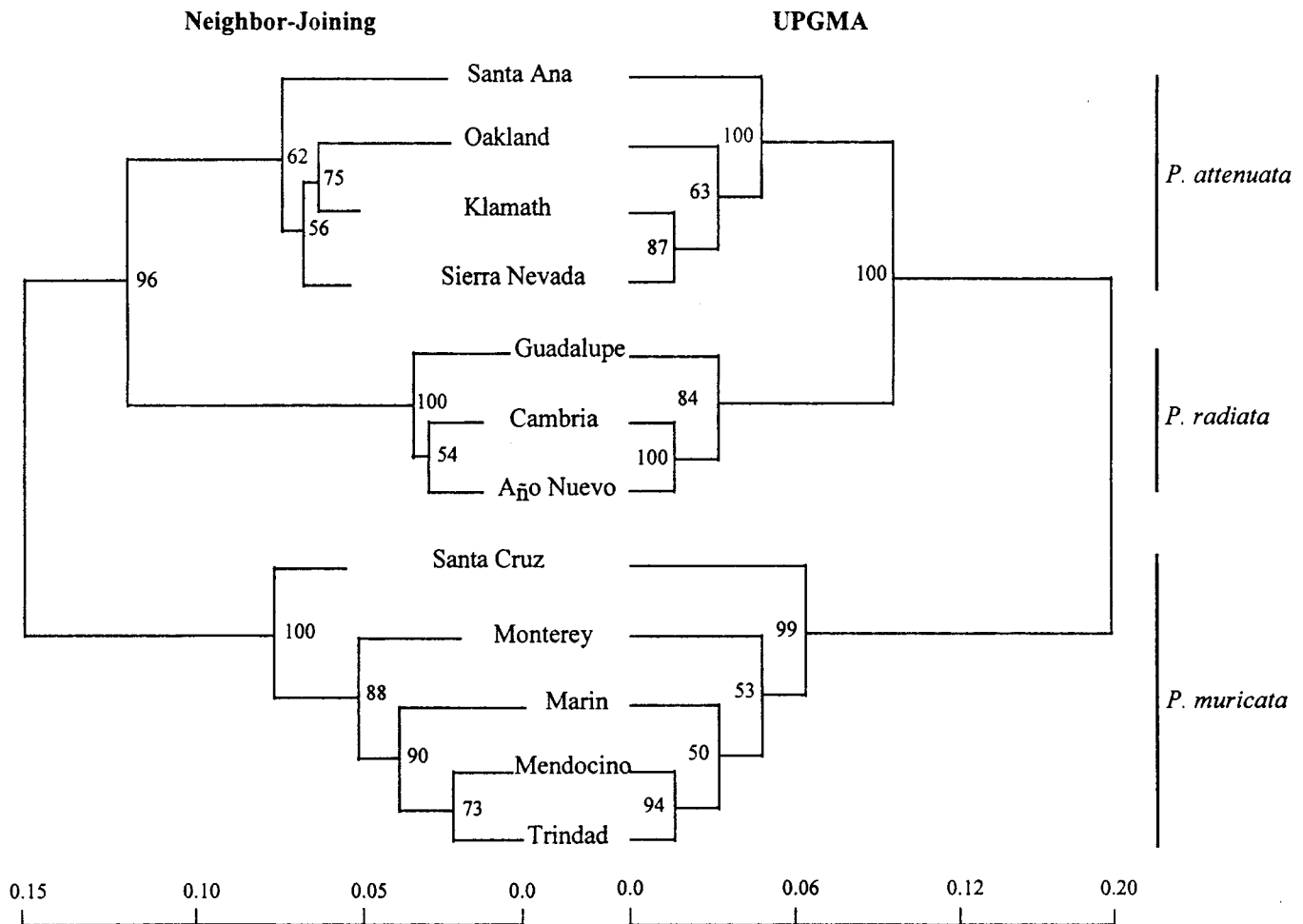
phenotypes converted from haploid material in Scots pine, *Pinus sylvestris.* The comparison of genetic parameters based on these two sets of allele frequencies indicated that genetic diversity within populations was substantially higher for direct RAPDs than for indirect RAPDs *(Hs = 0.37* vs. 0.22), whereas differentiation among populations was 3-fold higher for indirect RAPDs than for direct RAPDs (GsT = 0.06 vs. 0.02). Not surprisingly, dominance of RAPDs can affect calculation of genetic parameters, although the small sample sizes and low number of loci in this study may have increased the extent of bias. A lower expected heterozygosity and inflated among-population differentiation were also observed with predicted diploid phenotypes for RAPDs in black spruce, *Picea mariana* (Isabel et al. 1995).

The simulation analysis allowed us to estimate the effect of RAPD dominance and biallelism and thus make a more accurate comparison of genetic parameters between RAPDs and allozymes. Genetic diversities within populations based on allozyme allele frequencies that were transformed into biallelic dominant data were significantly lower than for nontransformed multiallelic codominant data, while population differentiation was biased upwardly in each species (Fig. 2). Nonetheless, RAPD markers detected more than

twice the intrapopulation diversity than the transformed dominant. allozymes. Population differentiation was similar *(P. muricata)* or only moderately larger *(P. radiata* and *P. attenuata)* for RAPDs compared to transformed allozymes. This suggests that RAPDs are predictably underestimating diversity, yet should be more sensitive markers for detecting variation in the level of genetic diversity than are allozymes. However, comparisons of population differentiation between RAPDs and allozymes can be seriously biased in a complex manner that depends on allele frequencies and sample sizes.

The higher population diversity and differentiation of RAPDs in this study are consistent with the results reported in other organisms. Liu and Furnier (1993) found that RAPD diversity *(Hs)* is much higher than allozymes and RFLPs within populations of bigtooth aspen *(Populus grandidentata) (Hs = 0.31* for RAPDs, 0.08 for allozymes and 0.13 for RFLPs), despite the dominance of RAPDs. Diversity within populations of *Pinus sylvestris* is also substantially higher for haploid-based RAPDs than for allozymes *(Hs = 0.36* vs. 0.26). Peakall et al. (1995) transformed allozyme data into binary presence-absence (1, 0) vectors to enable comparison to RAPDs and then computed distance matrices

for both markers and analyzed them using AMOVA (Excoffier et al. 1992). RAPD markers provided higher population differentiation than allozymes (58.4% vs. 45.2%).

The greater population diversity and divergence for RAPDs than for allozymes may result from several causes (summarized in Aagaard et al. 1995). First, only functional genes encoding soluble enzymes are studied for allozymes, whereas RAPDs anneal randomly to inverted repetitive DNA, and are expected to reflect primarily random variation of noncoding repetitive DNA. It is well known that the repetitive component of genomes can diverge and change in sequence rapidly (Cabot et al. 1993; Charlesworth et al. 1994). Second, allozymes often fail to detect many kinds of genetic variation because only nucleotide differences that lead to electrophoretically detectable changes in amino acid composition of the gene product are observed. RAPDs, however, should be sensitive to different types of mutations, including point mutations at primer annealing sites, inversions flanking the annealing sites, secondary structure constraints, and insertions that cause a greater length than can be amplified with routine polymerase chain reaction (PCR) techniques. Third, it has been hypothesized (e.g., Black et al. 1992) that highly repetitive regions of genomes such as telomeres and minisatellites evolve more rapidly than allozyme loci, and are preferably represented in RAPD fragments. Finally, allozymes may be subject to stronger natural selection since they represent functional gene products. Directional selection may reinforce the action of genetic drift in reducing the variability at protein level (Begun and Aquadro 1993; Baruffi et al. 1995). It has also been suggested that allozymes are under the influence of balancing selection (Altukhov 1990, 1991; Karl and Avise 1992; Mitton 1997), which helps to maintain intrapopulation polymorphism, but could reduce population differentiation.

Population differentiation based only on loci selected by segregation analysis was similar to that based on all RAPD loci, though the genetic diversities were inflated. Thus, segregation analysis could be used to select nuclear RAPD markers to estimate differentiation when enriched organellar DNA is unavailable for excluding RAPD markers of cytoplasmic origin.

For diploid data, Lynch and Milligan (1994) recommended that unbiased estimates of population genetic parameters can be obtained with RAPDs if the analysis is limited to loci with dominant alleles that are not prevalent (recessive nullallele homozygote occurs not less than in three individuals). This restriction can exclude many population-specific loci and highly differentiated loci. Loci that are polymorphic in some populations or species, but monomorphic in others, will also be excluded. It is therefore not surprising that our estimates of population differentiation based on this limited set of loci were much lower compared to those based on all loci (Table 2). Szmidt et al. (1996) also found that population differentiation was substantially underestimated based on loci selected by Lynch and Milligan's (1994) criteria compared to estimates from all RAPD loci.

The degree of population differentiation for nuclear markers (RAPDs and allozymes) was also compared with that for cytoplasmic DNA RFLP markers (mitochondrial and chloroplast DNA) in the same groups (Table 3). The levels of population subdivision for mtDNA length mutations and

cpDNA site mutations greatly exceeded those observed with RAPD and allozyme markers. This is consistent with the theoretical predictions for both organelle genomes, which have small effective population size, low sequence mutation rate, and limited gene flow for maternally inherited mtDNA. The strong differentiation for cpDNA, which is paternally inherited in pines, may result from a combination of the strongly disjunct nature of most populations iii the CCCP, limiting effective pollen dispersal, and haploidy promoting genetic drift.

Phylogenetic analysis

Cluster analysis of RAPD markers resulted in unambiguous distinction of the three species, *P. attenuata, P. radiata,* and *P. muricata.* These results are in agreement with the allozyme study of Millar et al. (1988), where three species were distinguished clearly and each species was about equally differentiated from the other two. In their results there was no obvious phylogenetic affinity between pairs of species. However, our NJ analysis of the joint RAPDallozyme data provided strong support for a common ancestor for *P. attenuata* and *P. radiata.* CpDNA analysis, on the other hand, divided the CCCP into three major groups: (i) a very strongly supported clade with *P. attenuata; (ii)* a weakly supported clade with both *P. radiata* and the southern race of *P. muricata;* and (iii) a very strongly supported clade with the northern and intermediate races of *P. muricata* (Hong et al. 1993b). Rates of cpDNA evolution were also highly variable. It appears that the chloroplast and nuclear genomes have different phylogenetic origins in the CCCP, a result also reported in several other plant taxa (Rieseberg and Soltis 1991).

Geographic clines of RAPD were apparent within the three species. RAPD relationships of *P. attenuata* populations were similar to allozyme patterns (Millar et al. 1988) and to latitudinal clines observed in growth and morphological characteristics (Newcomb 1962). Clinal patterns of variation in the NJ tree appear to follow a gene flow path from Santa Ana northward along Sierra Nevada through the Klamath mountains, and then southward along the coast mountain ranges. The large genetic distance of the Oakland population from the other three northern populations is likely to have forced it to its basal positions among those populations in the UPGMA tree (Fig. 4).

In *P. radiata,* RAPD analysis indicated that the two mainland populations are more closely related to each other than to the Guadalupe Island population. This is also consistent with allozyme variation (Plessas and Strauss 1986; Millar et al. 1988; Moran et al. 1988), and many morphological and biochemical traits such as needle number. stem form, seed proteins, and terpenes (reviewed in Millar 1986). Based on cone morphology (Axelord 1980) and allozymes, the island populations were inferred to be most ancestorlike, consistent with our RAPD results.

Separate taxonomic designations have been proposed for the northern and southern groups of *P. muricata* based on studies of a number of traits, including morphology (Fielding 1961; Shelbourne et al. 1982), monoterpene composition (Forde and Blight 1964; Mirov et al. 1966), and crossability (Critchfield 1967: Millar and Critchfield 1988). They are also supported by frequencies at some specific

allozyme loci (Millar 1983: Millar et al. 1988) and cpDNA restriction site variation (Hong et al. 1993b). However, both groups still clustered tightly within *P. muricata* based on allozymes, and genetic distances between groups were not much higher than among distant populations in *P. attenuata* and *P. radiata*. Phenetic analysis of RAPD data in the present study confirms that populations of *P. muricata* are monophyletic though they are highly differentiated. Both RAPDs and allozymes indicate clinal patterns of divergence from north to south, with the southern populations appearing most ancestral (similar to outgroups) and the northern mainland populations being most divergent (Millar et al. 1988). The several phylogenetic discrepancies between RAPDs and allozymes in our analysis are very likely to be the result of small sample size in several populations in the allozyme data set (Millar et al. 1988).

The broad relationships of populations and species based on RAPDs agrees well with those based on allozymes (Millar et al. 1988) and many morphological and biochemical traits. This suggests that RAPDs are appropriate molecular markers for phylogenetic analysis at this level. RAPD markers have been widely used in the reconstruction of phylogenetic relationships for many organisms (reviewed by Wolfe and Liston 1997), and there has been general concordance among the results derived from RAPDs and other techniques. These and our own results confirm that despite the complications of dominance and phenotype interpretation, for which RAPDs are infamous, the large number of loci resolved can make them valuable phylogenetic markers within and between closely related species.

Acknowledgements

References

Aagaard, J.E. *1997*. Genetic diversity and differentiation in Douglas-fir from RAPD markers of nuclear and mitochondrial origin. M.S~c. Thesis, Oregon State University, Corvallis, Or.

Aagaard, J.E., Vollmer, S.S., Sorensen, F.C., and Strauss, S.H. *1995*. Mitochondrial DNA products among RAPD profiles are frequent and strongly differentiated between races of Douglasfir. Mol. Ecol. *4: 441-447*.

Aagaard, I.E., Krutovskii, KX, and Strauss, S.H. *1998a*. RAPDs and allozymes exhibit similar levels of diversity and differentiation among populations and races of Douglas-fir. Heredity, 81: *69-78*.

Aagaard, J.E., Krutovskii, KX, and Strauss, S.H. *19986*. RAPD markers of mitochondrial origin exhibit lower population diver

sity and higher differentiation than RAPDS of nuclear origin in Douglas-fir. Mol. *Ecol. 7: 801-812*.

Altukhov, Yu.P. *1990*. Population Genetics, Diversity and Stability. Harwood Academic Publishers. London.

Altukhov, *Yu.P. 1991*. The role of balancing selection and overdominance in maintaining allozyme polymorphisms. Genetica, 85: *79-90*.

Axelord, D.I. *1980*. History of the maritime closed-cone pines, Alta and Baja California. Univ. Calif. Press, Calif.

Baruffi, L., Damian, G., Guglielmino, C. R., Bandis, C., Malacrida, A. R., and Gasperi, G. *1995*. Polymorphism within and between populations of *Ceratitis capitata,* comparison between RAPD and multilocus enzyme electrophoresis data. Heredity, *74: 425-437*.

Begun, D.J., and Aquadro, C.F. *1993*. African and North American populations of *Drosophila melanogaster* are very different at the DNA level. Nature, 365: *548-550*.

Black, W.C., IV. 1996. RAPDBOOT 1.0, RAPDDIST 1.0, RAPDFST 4.0 RAPDPLOT 2.4., Department of Microbiology, Colorado State Universoty, Fort Collins, Colo. (ftp lamar.colostate.edu in pub/wcb4).

Black, W.C., IV, Duteau, N.M., Puterka, G.J., Nechols, M.1., and Pettorini, J.M. *1992*. Use of the random amplified polymorphic DNA polymerase chain reaction (RAPD-PCR) to detect DNA polymorphism in aphids *(Homoptera,* Aphididae). Bull. Entomol. Res. 82: *151-159*.

Bucci, G., and Menozzi, *P. 1993*. Segregation analysis of random amplified polymorphic DNA (RAPD) markers in *Picea abies* Karst. Mol. Ecol. 2: 227-232.

Cabot, E.L., Doshi, P., Wu, M.L.. and Wu, C.I. 1993. Population genetics of tandem repeats in centrometric heterchromatinunequal crossing over and chromosomal divergence at the responder locus of *Drosophila melanogaster.* Genetics, 135: 477487.

Charlesworth, B., Sniegowski, P., and Stephen, W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature, 371: 215-220.

Critchfield, W.B. 1967. Crossability and relationships of the California closed-cone pines. Silvae Genet. 16: 89-97.

Excoffier. L. 1993. WINAMOVA. Genetics and Biometry Laboratory, University of Geneva, Carouge, Switzerland (ftp: 129.194.113.13/ftp/comp/win/ or ftp: anthropologie.unige.ch/ pub/comp/win/).

Excoffier, L., Smouse, P. E., and Quattro. J. M. 1992. Analysis of ,molecular variance inferred from metric distances among DNA haplotypes, application to human mitochondrial DNA restriction data. Genetics, 131: 479-491.

Felsenstein, J. 1995. PHYLIP: A Phylogeny Inference Package, version 3.57c. Department of Genetics, University of Washington, Seattle, Washington (http://evolution.genetics.washington.edu/ phylip/software.html).

Fielding, J.M. 1961. Provenances of monterey and bishop pines. Aust. For. Timb. Bureau Bull. 38: 1-30.

Forde, M.B., and Blight, M.M. 1964. Geographic variation in the turpentine of bishop pine. N. Z. J. Bot. 2: 44-52.

Graham, J., and McNicol, R.J. 1995. An examination of the ability of RAPD markers to determine the relationships within and between *Rubus* species. Theor. Appl. Genet. 90: 1128-1132.

Hamrick, J.L., and Godt, M.J. 1989. Allozyme diversity in plant species. *In:* A.H.D. Brown, M.T. Clegg, A.L. Kahler, and B.S. Weir (Eds.). Plant Population Genetics, Breeding and Genetic Resources, Sinauer Associates, Inc., Sunderland, Mass. pp. 4363.

Heun, M., and Helentjaris, T. 1993. Inheritance of RAPDs in FI hybrids of corn. Theor. Appl. Genet. 85: 961-968.

Hipkins, V.D., Krutovskii, K.V, and Strauss, S.H. 1994. Organelle genomes in conifers: Structure, evolution, and diversity. Forest Genetics, 1: 179-189.

Hong, Y.P. 1991. Chloroplast DNA variability and phylogeny in the California closed-cone pines. Ph.D. Dissertation, Oregon State University, Corvallis, Or.

Hong, YP., Hipkins, V.D., and Strauss, S.H. 1993a. Chloroplast DNA diversity among trees, populations and species in the California closed-cone pines. Genetics, 135: 1187-1196.

Hong, Y.P., Krupkin, A.B., and Strauss, S.H. 19936. Chloroplast DNA transgresses species boundaries and evolves at variable rates in the California closed-cone pines (Pinus radiata, P. muricata, and P. attenuata). Mol. Phylogenet. and Evol. 2: 322-329. 1

Isabel, N., Beaulieu, J., and Bousquet, J. 1995. Complete congruence between gene diversity estimates derived from genotypic data at enzyme and random amplified polymorphic DNA loci in black spruce. Proc. Natl. Acad. Sci. U.S.A. 92: 6369-6373.

Karl, S.A., and Avise, J.C. 1992. Balancing selection at allozyme loci in oysters, implications from nuclear RFLPs. Science, 256: 100-102.

Kazan, K., Manners, J.M.. and Cameron, D.F. 1993. Inheritance of random amplified polymorphic DNA markers in an interspecific cross in the genus Srvlosanthes. Genome, 36: 50-56.

Krutovskii, K.V., Vollmer, S.S., Sorensen, F.C., Adams, W.T., Knapp, S.J., and Strauss, S.H. 1998. RAPD genome maps of Douglas-fir. J. Heredity, 89: 197-205.

Krutovskii, K.V., Erofeeva, S.Y., Aagaard, J.E., and Strauss, S.H. 1999. Simulation of effects of dominance on estimates of population genetic diversity and differentiation. J. Heredity, 90: 499502.

Lanner-Herrera, C.. Gustafsson, M., Falt, A. S., and Bryngelsson, T. 1996a. Diversity of wild Brassica oleraceae as estimated by isozyme and RAPD analysis. Genetic Resources and Crop Evolution, 43: 13-23.

Lanner-Herrera, C., Bryngelsson, T., and Gustafsson, M. 19966. Genetic validity of RAPD markers at the intra- and interspecific level in wild Brassica species with n = 9. Theor. Appl. Genet. 93: 9-14.

Latta, R.G., and Mitton, J.B. 1997. A comparison of population differentiation across four classes of gene marker in limber pine Winus flexilis James). Genetics, 146: 1153-1163.

le Corre, V., Dumolin-Lapegue, S., and Kremer, A. 1997. Genetic variation at allozyme and RAPD loci in sessile oak Quercus petraea (Matt.) Liebl.: The role of history and geography. Mol. Ecol. 6: 519-529.

Lewis, P.O. 1994. cENEsTAT-rc 3.3. Department of Statistics, North Carolina State University, Raleigh, N.C. (ftp: alleyn.unm.edu/genestat/).

Lewis, P.O., and Zaykin, D.V. 1996. Genetic Data Analysis (GDA). Dept. of Statistics. North Carolina State University, Raleigh. N.C. (ftp: alleyn.eeb.uconn.edu/pub/gda or http://alleyn.eeb.uconn.edu/gda).

Lidholm, J., and Gustafsson, P. 1991. The chloroplast genome of the gymnosperm Pinus cnntorta, a physical map and a complete collection of overlapping clones. Curr. Genet. 20: 161-166.

Lifante, Z.D., and Aguinagalde, I. 1996. The use of random amplified polymorphic DNA (RAPD) markers for the study of taxonomical relationships among species of Asphodelus sect. Uerinea (Asphodelaceae). Am. Bot. 83: 949-953.

Liu, Z.. and Furnier, G.R. 1993. Comparison of allozyme, RFLP and RAPD markers for revealing genetic variation within and between trembling and bigtooth aspen. Theor. Appl. Genet. 87: 97-105.

Lorenz, M., Weihe, A., and Borner, T. 1994. DNA fragments of organellar origin in random amplified polymorphic DNA (RAPD) patterns of sugar beet (Beta vulgaris L.). Theor. Appl. Genet. 88: 775-779.

Lu, M.Z., Szmidt, A.E., and Wang, X.R. 1995. Inheritance of RAPD fragments in haploid and diploid tissues of Pinus sylvestris (L.). Heredity, 74: 582-589.

Lynch, M.. and Crease, T.J. 1990. The analysis of population survey data on DNA sequence variation. Mol. Biol. Evol. 7: 377394.

Lynch, M., and Milligan, B.G. 1994. Analysis of population structure with RAPD markers. Mol. Ecol. 3: 91-99.

Marillia, E.F., and Scoles, G.J. 1996. The use of RAPD markers in Hordeum phylogeny. Genome, 39: 646-654.

Millar, C.I. 1983. A steep cline in Pinus muricata. Evolution, 37: 311-319.

Millar, C.I. 1986. The California closed-cone pines (Subsection Oocarpae Little and Critchfield), A taxonomic history and review. Taxon, 35: 657-670.

Millar, C.I., and Critchfield, W.B. 1988. Crossability and relationships of bishop pine. Madroiio, 35: 39-53.

Millar, C.I., Strauss, S.H., Conkle, M.T., and Westfall, R. 1988. Allozyme differentiation and biosystematics of the California closed-cone pines. Syst. Bot. 13: 351-370.

Mirov, N.T., Zavarin, E., Snajberk, K., and Costello, K. 1966. Further studies of *Pinus muricata* in relation to its taxonomy. Phytochemistry, 5: 343-355.

Mitton, J.B. 1997. Selection in Natural Populations. Oxford University Press.

Moran, G.F., Bell, J.C., and Eldridge, K.G. 1988. The genetic structure and conservation of the five natural populations of *Pinus radiata*. Can. J. For. Res. 18: 506-514.

Nei, M. 1972. Genetic distance between populations. Am. Nat. 106: 283-292.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics, 89: 583590.

Nei, M. 1986. Definition and estimation of fixation indices. Evolution, 40: 643-645.

Nei, M., and Li, W.H. 1985. Mathematical model for studying genetic variation in terms of restriction endonuclease. Proc. Nail. Acad. Sci. U.S.A. 76: 5269-5273.

Newcomb, G.B. 1962. Geographic Variation in *Pinus attenuata* Lemmon. Ph.D. Dissertation, University of California at Berkeley, Calif.

Page, R.D.M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences, 112: 357-358. (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).

Peakall, R., Smouse, P.E., and Huff, D.R. 1995. Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloe dachtvloides*. Mol. Ecol. 4: 135-147.

Plessas, M., and Strauss, S.H. 1986. Allozyme differentiation among populations, stands and cohorts in Monterey pine. Can. J. For. Res. 16: 1155-1164.

Rieseberg, L.H. 1996. Homology among RAPD fragments in interspecific comparisons. Mol. Ecol. 5: 99-105.

Rieseberg, L.H., and Soltis, D.E. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evol. Trends Plants, 5: 6584.

Shelbourne, C.I.A., Bannister, M.H., and Wilcox, M.D. 1982. Early results of provenance studies on Pinus muricata in New Zealand. N. Z. J. Forestry, 27: 50-66.

Sokal, R.R., and Rohlf, F.J. 1981. Biometry, 2nd ed. W.H. Freeman and Co., New York.

Strauss, S.H., Hong, YP., and Hipkins, V.D. 1993. High levels of population differentiation for mitochondrial DNA haplotypes in Pinus radiata, P. muricata, and P. attenuata. Theor. Appl. Genet. 86: 605-611.

Swofford, D.L., and Selander, R.B. 1989. alosYS-1.7: A computer program for the analysis of allelic variation in population genetics and biochemical systematics. Illinois Natural History Survey, Champaign, Ill. (ftp: lamar.colostate.edu in pub/wcb4).

Szmidt, A.E., Wang, X., and Lu, M. 1996. Empirical assessment of allozyme and RAPD variation in, Pinus sylvestris (L.) using haploid tissue analysis. Heredity, 76: 412-1120.

Thormann, C.E., Ferreira, M.E., Camargo, L.E.A., Tivang, J.G., and Osborn, T.C. 1994. Comparison of RFLP and RAPD markers for estimating genetic relationships within and among cruciferous species. Theor. Appl. Genet. 88: 973-980.

Vicario, F., Vendramin, G.G., Rossi, P., Lio, P., and Gianini, R. 1995. Allozyme, chloroplast DNA and RAPD markers for determining genetic relationships between Abies alba and the relic

population of Abies nebrodensis. Theor. Appl. Genet. 90: 1012-1018.

Weir, B.S., and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. Evolution, 38: 1358-1370.

Williams. J.G.K., Kubelik, A.R., Livak, K.J., Rafalski. J.A., and Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18: 6531-6535.

Williams, J.G.K., Hanafey, M.K., Rafalski, J.A., and Tingey, S.V. 1993. Genetic analysis using random amplified polymorphic DNA markers. Methods in Enzymology, 218: 704-740.

Wolfe, A.D., and Liston, A. 1997. Contribution of the polymerase reaction to plant systematics. In P.S. Soltis, D.E. Soltis, and J.J. Doyle (Eds.). Molecular systematics of plants. Chapman and Hall, New York.

Wright, S. 1931. Evolution in Mendelian populations. Genetics, 16: 97-159.

Wright, S. 1978. Evolution and Genetics of Populations. Vol. 4. Variability Within and Among Natural Populations. University of Chicago Press, Chicago.

Wu, J., Krutovskii, KX, and Strauss, S.H. 1998. Abundant mitochondrial genome diversity, population differentiation, and evolution in pines. Genetics, 150: 1605-1614.