

## High levels of population differentiation for mitochondrial DNA haplotypes in *Pinus radiata*, *muricata*, and *attenuata*

S. H. Strauss, Y.-P. Hong\*, V. D. Hipkins

Department of Forest Science, FSL 020, Oregon State University, Corvallis, OR 97331-7501, USA

Received: 20 September 1992 / Accepted: 9 December 1992

**Abstract.** We analyzed mitochondrial (mt) DNA restriction fragment length polymorphisms (RFLPs) associated with cytochrome oxidase, subunit I (*coxI*)-related gene sequences in 268 trees derived from 19 natural populations of three species of pines from California (USA): Monterey pine (*Pinus radiata* D. Don), bishop pine (*P. muricata* D. Don), and knobcone pine (*P. attenuata* Lemm.). Total genomic DNA was digested with four restriction endonucleases and probed with a 750-bp fragment of the mitochondrial *coxI* gene amplified from *P. attenuata* via the polymerase chain reaction (PCR). The *coxI* gene is repeated at least 4 times in some populations, and all variants that we observed resulted from complex rearrangements rather than from point mutations. There was limited intrapopulation variation, but strong differentiation among populations. When applied to haplotype frequencies, Nei's gene diversity within populations ( $H_s$ ) averaged 7% ( $\pm 3$ ), and  $G_{st}$  varied from 75% for *P. radiata* to 96% for *P. muricata*. The high degree of population differentiation for mtDNA suggests that it can be a powerful marker of population differences, but its rapid rate of structural evolution appears to result from recombination among a limited number of repetitive elements—giving frequent homoplasious fragment phenotypes. The phylogenetic trees disagreed with results from chloroplast DNA, nuclear gene, and morphological studies.

**Key words:** Organelle – Pine – Mitochondrial DNA – Chloroplast DNA – Phylogeny

Communicated by P. M. A. Tigerstedt

\* Present address: Biotechnology Laboratory, University of British Columbia, Room 237 Westbrook Bldg., 6174 University Blvd., Vancouver, B.C., Canada V6T 1Z3

Correspondence to: S. H. Strauss

### Introduction

Little is known of the amount and distribution of mitochondrial DNA (mtDNA) diversity in natural populations of plants. Although there have been a number of studies of intraspecific variation (Palmer 1992), most have analyzed a limited number of cultivars, lines or accessions, rather than structured samples of natural populations. Unlike the chloroplast (cp), genomes of most plants and the mitochondrial genomes of animals, plant mitochondrial genomes have an abundance of dispersed repetitive DNA and an associated tendency to undergo rapid structural evolution (reviewed in Sederoff 1987). Combined with its very low rate of gene sequence evolution (Wolfe et al. 1987), this results in most polymorphisms resulting from genome rearrangements and length mutations rather than from point mutations. Although this feature limits its power for phylogenetic reconstruction (Palmer 1992), it may nonetheless be a useful tool for the analysis of population variability and differentiation.

To provide information on the natural distribution of mtDNA variation among and within populations, we studied a species-complex of closely related pines from California and nearby Oregon and Mexico (USA). Most populations in the complex are disjunct, particularly for the two coastal species – Monterey and bishop pine. The complex has been well studied for a number of other characters, including morphology, crossability, secondary compound chemistry, and allozyme frequencies (reviewed in Millar 1986, Millar et al. 1988). It was also the subject of a recent analysis of cpDNA diversity (Hong 1991). It therefore provides an opportunity to contrast mtDNA diversity with that of nuclear- and chloroplast-encoded characters.

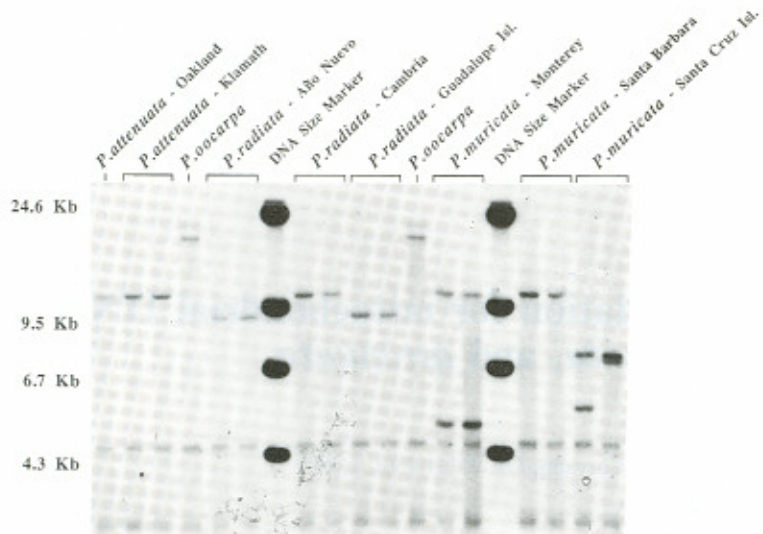


Fig. 1. An autoradiograph showing examples of intraspecific, interpopulation, and intrapopulation mtDNA diversity. Genomic DNA was digested with *Bam*HI

## Materials and methods

Trees were sampled from natural populations or from gene conservation plantations established from seed collected in natural populations, as described in Hong (1991). Total cellular DNA was extracted from needles using a CTAB-based DNA extraction method (Wagner et al. 1987). We modified this method only in that we worked in a cold room; used liquid nitrogen rather than extraction buffer for the initial grinding of needles; and shook the tubes vigorously during denaturation of the membranes in sarkosyl because it improved yield while having little adverse effect on DNA molecular weight.

DNA was digested with restriction enzymes according to the manufacturer's suggestions and fractionated according to size by electrophoresis in 0.8% agarose gels in a TAE buffer (80 mM TRIS, 16.6 mM sodium acetate, 2 mM EDTA, adjusted to pH 8.1 with glacial acetic acid). The gels were blotted in an alkaline denaturation solution onto Zetabind nylon membranes (Cuno Inc, Meriden, Conn.) and hybridized with a 750-bp portion of the *coxI* gene. *coxI* was amplified from knobcone pine via the polymerase chain reaction (PCR) with primers to highly conserved regions (Glaubitz and Carlson 1992). Forty PCR cycles with an annealing temperature of 55°C were performed using standard PCR chemistry as specified by the suppliers (Perkin Elmer Cetus). The PCR product was recovered from a 1.5% agarose gel under long-wave UV light, eluted with a Centrifuge micro-electroeluter (Amicon, Beverly, Mass.), and radioactively labelled to a specific activity of  $10^8$  cpm/ $\mu$ g with [ $^{32}$ P] by primer extension using random hexamers. Blotting, probe labelling, hybridization, washing, and stripping of blots were as described in Strauss and Doerksen (1990). Fragment sizes were estimated from autoradiograms using the Intelligenetics Suite program SIZER (IntelliGenetics, Inc, Mountain View, Calif.).

Diversity and subdivision statistics were calculated based on haplotype frequencies.  $G_{st}$  is Nei's fractional gene diversity statistic for population differentiation calculated using haplotype frequencies; it and its standard errors were calculated by the computer program HAPLO (Lynch and Crease 1990). Unbiased  $G_{st}$  estimates adjusted for sample size and population number were calculated by GENESTAT (Whitkus 1988).  $F_{st}$  is Wright's F-coefficient for population subdivision as estimated in Weir

and Cockerham (1984). It, and standard deviations derived from jackknifing over populations, were calculated from haplotype frequencies with a computer program described by B. S. Weir (1990). Hierarchical F-statistics describing haplotype subdivision in the complex as a whole were estimated using the computer program Biosys-1 (version 1.7; Swofford 1989).

Phylogenetic trees were derived from haplotype frequencies using Biosys-1 (version 1.7; Swofford 1989). UPGMA (unweighted pair-group with arithmetic-average clustering) trees were derived from modified Roger's distances and Nei's unbiased genetic identities; Distance Wagner trees were derived from Nei's and Prevosti distances.

## Results

### Nature of the fragment variants

The number and sizes of the fragments we observed suggest that probe-homologous *coxI* sequences are present in multiple copies. While the smaller fragments (<4 kb) were constant in all populations and species, the sizes of the larger fragments were highly variable (Table 1, Fig. 1). When the 0.5-kb band in the *Bam*HI digests, which may result from a fragment entirely internal to the probe, is ignored, there are three to four *coxI*-homologous bands revealed by *Bam*HI, two to three by *Eco*RI, two to three by *Eco*RV, and three to five by *Xba*I. If we assume that the one or two small (<4 kb), constant bands result from the presence of a single site within one of the copies of probe-homologous DNA, then the number of repeats remains three to four in *Bam*HI, but is reduced to one to two in *Eco*RI, one to two by *Eco*RV, and two to four by *Xba*I. *Bam*HI and *Xba*I appear to be the most sensitive with respect to differentiating types of repeats; the high degree of concurrence among them – for

**Table 1.** Restriction fragment phenotypes (kb) of nine observed haplotypes in mtDNA of the California closed cone pines

HT <sup>a</sup>	<i>Bam</i> HI			<i>Eco</i> RI				
I	0.5	3.2	4.6	9.2	1.3	2.5		
II	0.5	3.2	4.6	10.3	1.3	2.5		
III	0.5	3.2	4.6	7.3	10.3	1.3	2.5	
IV	0.5	3.2	4.6	10.3	1.3	2.5		
V	0.5	3.2	5.0	10.3	1.3	2.5		
VI	0.5	3.2	4.6	5.5	7.3	1.3	2.5	16.8
VII	0.5	3.2	4.6	5.5	7.3	1.3	2.5	16.8
VIII	0.5	3.2	4.6	5.5	7.3	1.3	2.5	16.8
IX	0.5	3.2	4.6	7.0	7.3	1.3	2.5	16.8

HT <sup>a</sup>	<i>Eco</i> RV		<i>Xba</i> I					
I	2.7	20.0	1.8	1.9	18.7			
II	2.7	24.0	1.8	1.9	12.2			
III	2.7	24.0	1.8	1.9	10.0	12.2		
IV	2.7	20.0	1.8	1.9	12.2			
V	2.7	24.0	1.8	1.9	12.2			
VI	2.7	24.0	1.8	1.9	10.0	14.4		
VII	2.7	18.5	24.0	1.8	1.9	10.0	14.4	
VIII	2.7	15.3	24.0	1.8	1.9	10.0	14.4	
IX	2.7	15.3	24.0	1.8	1.9	8.5	10.0	14.4

<sup>a</sup> HT, Haplotype

example, in both showing a large number of bands in haplotypes III, and VI through IX – supports our interpretation that band number results primarily from dispersed repeat number. Note that our estimates are minimums in that only repeats differentiated in the location of restriction sites relative to probe-homologous sequences will be observed. Nonetheless, results from *Bam*HI and *Xba*I indicate that at least four distinct *cox*I-containing repeats exist in some populations.

While variation in fragment number among haplotypes was correlated in the different enzymes, the presence or absence of specific fragments was not necessarily so. Fragments whose presence and absence were correlated among haplotypes were the 10.3-, 7.3-, 7.0-, and 9.2-kb fragments of *Bam*HI and the 16.8-kb fragment of *Eco*RI with the 12.2-, 10.0-, 8.5-, 18.7-, and 14.4-kb fragments of *Xba*I, respectively. Fragments whose presence versus absence had no perfect correlate were the 4.6-, 5.0-, and 5.5-kb fragments of *Bam*HI, and the 15.3-, 18.5-, 20.0-, and 24.0-kb fragments of *Eco*RV. This substantial degree of independence suggests that the enzymes did not always reflect the same rearrangements.

**Table 2.** Haplotype frequencies (%) in sampled populations and number of trees assayed (*n*)

Population	Haplotype										<i>n</i>
	I	II	III	IV	V	VI	VII	VIII	IX		
<i>Monterey pine</i>											
Año Nuevo	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	28
Cambria	0.0	90.5	9.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	21
Monterey	77.8	0.0	22.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9
Guadalupe	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8
Cedros	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4
										<i>Total</i>	72
<i>Knobcone pine</i>											
Sierra	0.0	81.0	0.0	9.0	0.0	0.0	0.0	0.0	0.0	0.0	21
Santa Ana	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22
Klamath	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	23
Oakland	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	23
										<i>Total</i>	89
<i>Bishop pine</i> <sup>a</sup>											
Mendocino	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	4
Trinidad	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	2
Marin	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	4
Monterey	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	11
Sonoma	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	6
Santa Barbara	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14
Santa Cruz	0.0	0.0	0.0	0.0	0.0	83.3	8.3	5.6	2.8	0.0	36
San Luis Obispo	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9
Santa Rosa	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	13
San Vicente	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8
										<i>Total</i>	107
										<i>Total sample</i>	268

<sup>a</sup> Northern (above) and southern (below) groups of bishop pine used in the analyses of Table 3 are separated

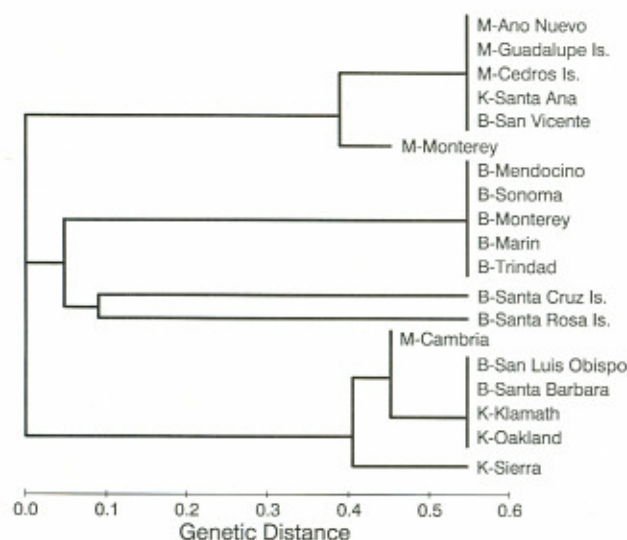


Fig. 2. Distance Wagner tree derived from Prevosti distances. M = Monterey pine, K = Knobcone pine, and B = bishop pine. The tree was rooted at midpoint between the pair of taxa with the greatest patristic distance. Percent standard deviation relative to distance matrix is 10.4; the cophenetic correlation is 0.996

#### Phylogenetic analyses

Fragment phenotypes often differed among populations within species, yet certain populations had phenotypes identical to those prevalent in populations of different species. For example, most of the knobcone pine populations had a phenotype that was the same as that prevalent in the Cambria population of Monterey pine and in the Santa Barbara and San Luis Obispo populations of bishop pine (Fig. 1, Table 2).

Consequently, all Distance Wagner and UPGMA phylogenetic trees gave results suggesting that species and populations were often polyphyletic (e.g., Fig. 2). All of the topologies had four major branches, and two of the branches included all three species near their termini. The third and fourth branches contained bishop pine only, with one containing the 2 southern island populations of Santa Cruz and Santa Rosa, and the other the northern populations. Results were similar when the frequencies of each fragment, rather than of haplotypes, were used to incorporate phenotypic similarity of the different haplotypes into the distance measure (not shown).

#### Intrapopulation diversity

Diversity within populations was limited. Most populations in the complex showed a single haplotype, and where polymorphism was observed a single haplotype was dominant (Table 2). The frequency of the common haplotype was never less than 78%, and the mean intrapopulation frequency was 92% for Monterey pine, 95% for knobcone pine, and 98% for bishop pine. All 27 trees sampled in the northern group of bishop pine had

Table 3. Hierarchical partitioning of mtDNA diversity based on haplotype frequencies

Level of analysis	Average gene diversity <sup>a</sup>	$G'_{st}$	$F'_{st}$
<i>Pooled total complex</i>			
Within populations	6.6 ( $\pm 2.8$ )		
Among populations	71.4 (ud)	91.6 (ud)	87.7 ( $\pm 5.3$ )
<i>Species</i>			
Within species	57.0 ( $\pm 10.8$ )		
Among species	21.7 (ud)	27.6 (ud)	27.0 ( $\pm 1.9$ )
<i>Populations within species</i>			
<i>Monterey pine</i>			
Within populations	11.4 ( $\pm 6.7$ )		
Among populations	35.1 ( $\pm 34.6$ )	75.5 ( $\pm 15.2$ )	83.3 ( $\pm 51.8$ )
<i>Knobcone pine</i>			
Within populations	8.1 ( $\pm 7.6$ )		
Among populations	48.2 ( $\pm 38.2$ )	85.6 ( $\pm 11.9$ )	86.3 ( $\pm 57.0$ )
<i>Bishop pine</i>			
Within populations	3.4 ( $\pm 3.2$ )		
Among populations	76.9 (ud)	95.8 (ud)	88.2 ( $\pm 11.5$ )
<i>Regional - Bishop pine</i>			
Within regions	36.8 ( $\pm 36.7$ )		
Among regions	63.2 <sup>a</sup> ( $\pm 1.2$ )	63.2 <sup>b</sup> (ud)	53.0 (ud)
<i>Southern group</i>			
Within populations	6.0 ( $\pm 5.7$ )		
Among populations	83.1 (ud)	93.2 (ud)	85.2 ( $\pm 13.4$ )

<sup>a</sup> Diversity estimates are given as percentages; standard errors ( $G'_{st}$ ) and standard deviations from jackknifing over populations ( $F'_{st}$ ), in parentheses; ud, undefined (incalculable by HAPLO). Populations included in regional analysis of bishop pine are shown in Table 2

<sup>b</sup> GENESTAT estimate was 31.6%

<sup>c</sup> GENESTAT estimate was 46.3%

a single haplotype, and only 1 of the 5 southern populations showed polymorphism.

On average, gene diversity within populations was 6.6% (Table 3). This ranged from a high value of 11.4% in Monterey pine, a consequence of significant polymorphism seen in the Cambria and Monterey populations, to a low of 3.4% in bishop pine, a consequence of only 1 of its populations showing any polymorphism. In all cases, intrapopulation gene diversities calculated from GENESTAT differed from those in Table 2 by less than 1%.

#### Population differentiation

Haplotype frequencies often differed strongly among populations (Table 2). The Santa Ana population of knobcone pine had a haplotype distinct from that found in all other knobcone pine populations. The Cambria population of Monterey pine lacked the haplotype prevalent in the other mainland populations and in the Guadalupe and Cedros Island populations. The northern populations of bishop pine all contained a single haplotype that was absent in the southern populations. Among the southern populations, populations often differed widely in their common haplotypes. San Vicente, Santa Cruz, and Santa Rosa each had haplotypes that were

common in no other bishop pine populations; Santa Barbara and San Luis Obispo shared a haplotype found in no other populations of bishop pine, but which was the prevalent form in knobcone pine.

We calculated G- and F-statistics to estimate the degree of population differentiation relative to diversity within populations. Both parameters can be interpreted as the fraction of total genetic diversity expressed as population differences compared to the total amount of genetic diversity in the reference group. When the species complex is considered as a unit, a large majority of the genetic diversity present – approximately 90% – occurs as population differences. Interestingly, when species are considered as groups, the average diversity among groups drops to 27%, indicating that the large diversity among populations in the complex is not primarily a reflection of species differences, but of differences among populations within species. When the complex was analyzed hierarchically – thus correcting for population variance when comparing species – the F coefficient for species was diminished to 13% (Swofford 1989) (F coefficients were 91% for populations within species and 92% for populations relative to the total complex). Differentiation among populations within species was 75–83% for Monterey pine, 86% in knobcone pine, and 88–96% in bishop pine (Table 3). Subdividing bishop pine into a northern and southern region reduced differentiation to 46–63% because most of the population diversity in the species occurred as differences among the southern populations. Analysis of the southern group alone indicated that 85–93% of the total genetic diversity resulted from population differences. Except as noted,  $G'_{st}$  estimates from GENESTAT were within 5% of those presented (Table 3).

## Discussion

### *Evidence that we are studying mtDNA*

A number of observations and experiments support our contention that hybridization of our blots of total genomic DNA with PCR-amplified *coxI* DNA resolved mtDNA fragments, and not those of nuclear or chloroplast origin:

1) Hybridization of the *coxI* probe to blots with parents and progeny from two crosses between *Pinus banksiana* Lamb. and *P. contorta* Dougl. showed that all 16 progeny contained a fragment similar to that found in the female parent, but unlike that found in the male parent (unpublished data). The predominantly maternal inheritance of mtDNA is expected in pines, whereas cpDNA is paternally inherited and nuclear DNA is biparentally inherited (Wagner et al. 1991).

2) Hybridization of our blots with a similarly sized PCR fragment (770 bp) amplified from the *rbcL* gene of

knobcone pine and labelled to the same specific activity gave a single band that was 20- to 30-fold more intense than any seen for *coxI*. This is expected given that our cellular DNAs were prepared from vigorous, green needles, and thus are likely to have many times more chloroplast genomes per cell than mitochondrial genomes (Bendich 1987).

3) The multiple fragment phenotypes we observed with *coxI* contrasts with those observed for cpDNA when similarly sized probes, such as *rbcL* described above, are used. Hybridization of digested DNA with small, coding-region-specific cpDNA probes only very rarely give multiple bands when restriction enzymes with a 6-bp recognition site are used (unpublished data).

4) The nature of the variants observed (complex rearrangements) and their high frequency contrast with those seen for cpDNA. Diversity among and within populations is extremely low for cpDNA, and nearly all variants are small insertions (ca. 100 bp) or point mutations (Hong 1991). Moreover, despite moderately frequent mtDNA polymorphisms in some populations, no evidence of heterozygotes were ever observed – as would be expected for nuclear DNA.

The combination of maternal inheritance, repetitive banding patterns, gene dosage, frequency of polymorphisms, haploidy, and the complex nature of the variants strongly argue that we are studying mtDNA.

### *Phylogenetic interpretations*

Two results of the study are of great interest: that a high level of population differentiation for mtDNA exists, and that like fragment phenotypes appear to have evolved independently several times over a short evolutionary time span.

Certain aspects of our phylogenetic trees disagree strongly with those based on morphology and allozymes. Allozymes have strongly confirmed the distinctness of the species, the correct placement of questionable populations within species (Millar et al. 1988), and the similarity of the mainland populations of Monterey pine when compared to the island populations. Allozymes, as well as terpenes (Mirov et al. 1966) and cpDNA (Hong 1991), have confirmed the strong differentiation of the northern versus southern populations of bishop pine. In contrast, our phylogenetic analyses based on mtDNA would suggest that the species are all polyphyletic, as is the southern bishop pine group. They also suggest that the mainland populations of Monterey pine at Año Nuevo and Monterey are more closely related to the Guadalupe and Cedros Island populations than to the mainland population at Cambria. In light of the very substantial amount of data suggesting these phylogenetic patterns are incorrect, we interpret these results as being consequences of convergent mtDNA structural evolution. Dong and

Wagner (1993) also observed some discordant phylogenetic results in their studies of mtDNA variation among lodgepole pine populations.

#### Nature of polymorphisms

All the polymorphisms that we observed appeared to result from structural rearrangements rather than from point mutations, which is in agreement both with results from lodgepole and jack pines (Dong and Wagner 1993) and with most other observations of plant mtDNA evolution (Palmer 1992). The stability of fragment phenotypes within populations suggests that phenomena such as heteroplasmy or recent somatic mutations – which are expected to give transient changes in phenotypes – are not important causes of the complex phenotypes observed.

The high frequency of homoplasy we observed is not surprising given the repetitive nature of mtDNA. If mtDNA structural evolution results primarily from recombination among a finite number of repeat elements (Palmer 1992), then the evolution of similar recombinant genotypes is expected, especially with the largest, most recombinogenic repeats. If *coxI* is associated with such repeats in pines, then it may be a particularly poor locus for phylogenetic studies. Other mtDNA regions need to be studied, as some may show less tendency for convergence.

#### Intrapopulation diversity

MtDNA intrapopulation haplotype diversities averaged 7%, which indicates that if genomes could combine at random to form diploids about 1 in 14 trees would be heterozygotes, with the remaining 13 trees nearly always homozygotes for the common haplotype. In contrast, average gene diversities for allozymes (15%: Millar et al. 1988) and cpDNA length mutations (17%: Hong 1991) were two- to threefold greater. Compared to cpDNA site mutations, mtDNA gene diversities were only marginally greater (7 vs. 6%: Table 3, and Hong 1991). However, mtDNA haplotypes were more phenotypically distinct, and thus far easier to differentiate, than cpDNA site mutations – which required analysis with about 100 kb of gene probes and 10–20 restriction enzymes for detection of differences.

#### MtDNA differentiation compared to other genomes

The large majority of the genetic variation we observed was a consequence of strong differences in haplotype frequencies among populations within species.  $G'_{st}$  and  $F_{st}$  values indicated that more than three-quarters of the gene diversity present in all three species was a result of population differences. Although some of these values may be slightly inflated owing to small samples, particu-

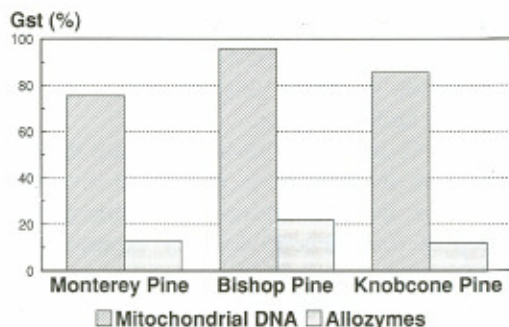


Fig. 3. Differences in magnitude of population subdivision between mtDNA and allozymes as measured by  $G_{st}$  in the three species

larly in bishop pine, it is clear that it is not the main cause of the high subdivision observed. Strong differences in haplotype frequencies among our population samples were observed in cases where our samples were the largest – including among the Año Nuevo and Cambria populations in Monterey pine; Santa Ana versus other populations of knobcone pine; Santa Cruz versus other southern populations of bishop pine; and between the pooled northern versus pooled southern populations of bishop pine. Moreover, when populations with less than 10 trees were excluded from the analysis, which left in the analysis 2 populations of Monterey pine, all 4 populations of knobcone pine, and 4 populations of bishop pine,  $G_{st}$  values either increased, as in the case of Monterey pine (to 91%), or decreased to less than 5% from their values in Table 3. High levels of population differentiation for mtDNA haplotype frequencies have also been observed in rangewide studies of lodgepole and jack pines ( $G_{st} \geq 50\%$ : Dong and Wagner 1993).

The degree of population differentiation for mtDNA far exceeds that observed in this group with nuclear gene markers (allozymes), where only 12–22% of the gene diversity present within species is associated with population differences in allele frequencies (Fig. 3; Millar et al. 1988). The high degree of mtDNA subdivision is similar, however, to that for cpDNA restriction site mutations, where strong differences among species ( $G'_{st} > 84\%$ ) and populations of bishop pine ( $G'_{st} > 87\%$ ) were observed (Hong 1991). These results are also in agreement with cpDNA differentiation found in oaks and with theoretical predictions for organelle genomes in the absence of substantial migration (discussed in Hong 1991).

#### MtDNA markers in conservation and breeding

The strong differences among populations that we observed for mtDNA suggests that it could be an efficient and useful indicator of population differentiation for breeders and conservation biologists. Although we have sampled only a small part of the genome with our *coxI*

probe, the use of additional gene probes would be expected to differentiate more – if not most – populations. The use of additional probes should also reduce the amount of homoplasy by providing a more complete picture of genomic differences. With very small samples per population breeders may be able to sort populations into broad classes when designing crossing and testing schemes. When populations are sampled for potential conservation purposes, mtDNA, in combination with geographic information, could help to identify unique populations that should receive a high priority for conservation. In contrast, nuclear gene markers would require much larger samples of both individuals and loci to detect statistically meaningful differences in allele frequencies among populations.

The populations that we studied were drawn from a rangewide sample, were isolated by substantial distances, and the differentiation observed was often associated with populations drawn from different varieties or races. An important question for future studies is to determine the minimum degree of isolation necessary for useful levels of mtDNA differentiation to develop.

*Acknowledgements.* We thank John Carlson and Jeffrey Glaubitz for providing the *cox1* PCR primers; Dave Wagner and Jinsheng Dong for lending their inheritance blots for hybridization, and for sharing their unpublished data and manuscript; Lisa Balduman for excellent technical assistance during the many blot hybridizations; Al Doerksen for technical and photographic aid; the CSIRO Division of Plant Industry, Canberra, Australia, where the manuscript was prepared during SHS's sabbatical; the National Institutes of Health (GM36367) and the National Science Foundation (BSR 8957023) for grant support; and the Institute of Forest Genetics, U.S. Forest Service, Berkeley, California, and Bill Libby for use of their plantations.

## References

- Bendich AJ (1987) Why do chloroplasts and mitochondria have so many copies of their genome? *Bioessays* 6:279–282
- Dong J, Wagner DB (1993) Taxonomic and population differentiation of mitochondrial DNA diversity in *Pinus banksiana* and *Pinus contorta*. *Theor Appl Genet* (in press)
- Glaubitz JC, Carlson JE (1992) RNA editing in the mitochondria of a conifer. *Curr Genet* 22:163–165
- Hong Y-P (1991) Chloroplast DNA variability and phylogeny in the California closed cone pines. PhD thesis, Oregon State University, Corvallis, Ore., USA
- Lynch M, Crease TJ (1990) The analysis of population survey data on DNA sequence variation. *Mol Biol Evol* 7:377–394
- Millar CI (1986) The Californian closed cone pines (subsection *Oocarpae* Little and Critchfield): A taxonomic history and review. *Taxon* 35:657–670
- Millar CI, Strauss SH, Conkle MT, Westfall R (1988) Allozyme differentiation and biosystematics of the Californian closed-cone pines. *Syst Bot* 13:351–370
- Mirov NT, Zavarin E, Snajberk K, Costello K (1966) Further studies of *Pinus muricata* in relation to its taxonomy. *Phytochemistry* 5:343–355
- Palmer JP (1992) Mitochondrial DNA in plant systematics: applications and limitations. In: Soltis PS, Soltis DE, Doyle JJ (eds) *Molecular systematics of plants*. Chapman & Hall, New York, pp 36–49
- Sederoff RR (1987) Molecular mechanisms of mitochondrial genome evolution in higher plants. *Am Nat* 130:S30–S45
- Strauss SH, Doerksen AH (1990) Restriction fragment analysis of pine phylogeny. *Evolution* 44:1081–1096
- Swofford D (1989) BIOSYS-1, Release 1.7. Illinois Natural History Survey, University of Illinois, Champaign, Ill., USA
- Wagner DB, Furnier GR, Saghai-Marroof MA, Williams SM, Dancik BP, Allard RW (1987) Chloroplast DNA polymorphisms in lodgepole pine and their hybrids. *Proc Natl Acad Sci USA* 84:2097–2100
- Wagner DB, Dong J, Carlson MR, Yanchuk AD (1991) Paternal leakage of mitochondrial DNA in *Pinus*. *Theor Appl Genet* 82:510–514
- Weir BS (1990) Intraspecific differentiation. In: Hills DM, Moritz C (eds) *Molecular systematics*. Sinauer Assoc. Sunderland, Mass., pp 373–410
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370
- Whitkus R (1988) Modified version of GENESTAT: a program for computing genetic statistics from allelic frequency data. *Plant Genet Newsl* 4:10
- Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci USA* 84:9054–9058