

**Historical Separation and Present Gene Flow Through a Zone of Secondary Contact in Ponderosa Pine**



Robert G. Latta; Jeffrey B. Mitton

*Evolution*, Vol. 53, No. 3 (Jun., 1999), 769-776.

Stable URL:

<http://links.jstor.org/sici?sici=0014-3820%28199906%2953%3A3%3C769%3AHSAPGF%3E2.0.CO%3B2-G>

*Evolution* is currently published by Society for the Study of Evolution.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ssevol.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

## HISTORICAL SEPARATION AND PRESENT GENE FLOW THROUGH A ZONE OF SECONDARY CONTACT IN PONDEROSA PINE

ROBERT G. LATTA<sup>1</sup> AND JEFFRY B. MITTON

*Department of Environmental, Population and Organismic Biology, University of Colorado, Boulder, Colorado 80309*

**Abstract.**—I examined the effects of historical division and secondary contact between eastern and western varieties of ponderosa pine (*Pinus ponderosa* Laws Pinaceae) on extant patterns of genetic variation. Fossil and biogeographic evidence both indicate that the current point of contact between these two varieties represents secondary contact following historical separation during the Wisconsin glaciation. Current gene flow was assessed by observing the degree of introgression of paternally inherited cpDNA and maternally inherited mtDNA polymorphisms. Both seeds and pollen are wind dispersed in ponderosa pine. Introgression was primarily from west to east, the direction of the prevailing wind, for both organelles, but introgression of cpDNA far exceeded that of mtDNA. Thus pollen is the main agent of contemporary gene flow between the two varieties. Neither seeds nor pollen showed enough introgression since secondary contact to have homogenized the two gene pools. However, allozyme differentiation was minimal. This calls into question assumptions of selective neutrality for at least some of the markers. Theory predicts that nuclear markers will show a high locus-to-locus variance of  $F_{ST}$  following historical separation. This prediction is confirmed by the allozyme data for ponderosa pine, and may provide a useful means of identifying historical separations from allele frequency data.

**Key words.**—Allozymes, cpDNA,  $F_{ST}$ , introgression, mtDNA, *Pinus ponderosa*, population structure.

Received August 4, 1998. Accepted January 8, 1999.

The inverse relationship between gene flow and the genetic differentiation among populations is one of the central tenets of population genetics. The relationship

$$F_{ST} = \frac{1}{4Nm + 1} \quad (1)$$

(where  $F_{ST}$  is the proportion of genetic variation attributable to allele frequency differences among populations and  $Nm$  is the number of migrants exchanged among populations each generation; Wright 1951) has been frequently cited as a means of estimating the level of migration among populations (Slatkin 1994) because  $F_{ST}$  is easier to determine than direct estimates of  $Nm$ . Empirical evidence generally supports this relationship (Waples 1987; Hamrick and Loveless 1989). Species with a greater capacity for dispersal generally exhibit lower values of  $F_{ST}$ .

However, equation (1) assumes an equilibrium condition, which may not obtain in many natural populations.  $F_{ST}$  will approach its equilibrium value following a change in the level of gene flow with a time lag proportional to  $N$  (the effective population size) or  $1/m$ , whichever is greater (Slatkin 1994). With the increasing appreciation of historical changes in the range and migration patterns of species (Davis 1981; Betancourt et al. 1990; Avise 1992) comes the need to evaluate to what degree currently observed levels of differentiation reflect extant, as opposed to past, patterns of gene flow. A common situation in North America is that of recent biogeographic shifts brought about by climate change during the Wisconsin glaciation. Many currently widespread species were isolated into a small number of glacial refugia during the height of the glaciation, whereas other species that were widespread during the glaciation are now isolated into small

refugia of suitable habitat (Betancourt et al. 1990; Pielou 1991). In the former case, secondary contact between formerly isolated populations could create an increase in the level of gene flow, whereas recent fragmentation of populations into refugia, represents a decrease in gene flow. In either case, present gene flow may not yet be reflected in the level of genetic differentiation.

Numerous instances have been documented where continuously distributed species in fact show sharp genetic discontinuities, presumably where secondary contact has occurred (e.g., Avise 1992; Soltis et al. 1997). What is not known is whether current gene flow is eroding this genetic differentiation or whether the prior separation has created a barrier to gene exchange, such that current gene flow remains limited following contact. Essentially, the questions are whether enough time has elapsed for current levels of gene flow to have erased the imprint of historical separations and, if not, how we can most easily identify nonequilibrium conditions. The most direct approach to this problem is to study systems in which the historical patterns of migration can be inferred from fossil evidence and phylogeographic evidence, which is independent of allelic frequencies. This study examines the pattern of gene flow and allelic differentiation across a zone of secondary contact between historically separate varieties of ponderosa pine. The first goal will be to verify that the contact zone (described by Conkle and Critchfield 1988) does indeed represent a point of secondary contact between differentiated groups and to determine the degree of contemporary gene flow that has occurred between the two varieties since secondary contact. We use molecular markers that allow a separate assessment of the movement of seed and pollen, so as to obtain a complete accounting of gene flow via both these propagules.

The second goal of this study concerns finding ways to identify historical separations when independent evidence is lacking. Many studies employ allelic frequency data to infer

<sup>1</sup> Present address: Department of Biology, Dalhousie University, 1355 Oxford Street, Halifax, Nova Scotia B3H 4J1, Canada; E-mail: rglatta@is.dal.ca.

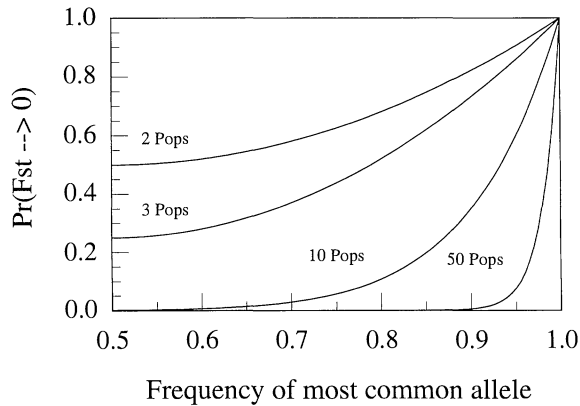


FIG. 1. The probability, equation (2), that  $F_{ST}$  for a diallelic locus will be small despite isolation given an initial frequency for the most common allele and the number of isolated populations into which an ancestral population is divided.

migration, because these are relatively easy to obtain for the most commonly used electrophoretic markers, primarily allozymes (Hamrick and Godt 1990). However, a deeper understanding of the effect of historical biogeographic movements on the geographic patterning of allelic frequencies is needed before we can disentangle current and historical effects from observed  $F_{ST}$ . For example, in a recent survey of allozyme and cpDNA markers across populations of *Senecio gallicus*, Comes and Abbott (1998) were unable to distinguish the action of contemporary gene flow from the historical biogeographical movements of this species. One possible approach to the problem is suggested by the theoretical work of Robertson (1975), who pointed out that if populations are historically structured, the variance of  $F_{ST}$  across loci would increase relative to the equilibrium situation. The intuitive understanding for this hinges on noting that restricted gene flow between two groups permits gene frequencies to drift independently in the two isolated populations. Thus, allelic frequencies may diverge, but need not necessarily do so. Independently drifting allelic frequencies may move in the same direction merely by chance. The probability of observing low differentiation in spite of isolation,  $\Pr(F_{ST} \rightarrow 0)$ , is approximately related to the number ( $n$ ) of subpopulations into which the ancestral population is divided and to the starting allelic frequencies ( $p$ ) before subpopulations became isolated:

$$\Pr(F_{ST} \rightarrow 0) = \sum p^n, \quad (2)$$

where the summation is across all alleles at a locus. This relationship is illustrated in Figure 1 for a diallelic locus. As the number of populations increases, the probability that the same allele will drift to high frequency in all populations (giving low  $F_{ST}$ ) decreases. In contrast, as the frequency of the most common allele increases, it becomes increasingly likely for that allele to drift to high frequency in all populations. Thus, for a species divided into two glacial refugia, the probability of any given locus giving the misleading result of low  $F_{ST}$  is at least 0.5, and frequently higher, such that there is a high probability that some loci will show low and some loci high values of  $F_{ST}$ . We applied this reasoning to RAPD and allozyme data (Latta and Mitton 1997) collected

from populations of limber pine along an elevational gradient in Colorado.  $F_{ST}$  values varied widely among loci from 0.00 to 0.36, and suggested a historical division between eastern and western populations during the Wisconsin glaciation. While Robertson's (1975) result has been observed in simulation studies (Holsinger and Mason-Gamer 1996), to date there has been no empirical test of Robertson's hypothesis, in a system where there is independent evidence of the historical separation. This study tests that hypothesis using the secondary contact zone in ponderosa pine.

## METHODS

### The Study System

Ponderosa pine (*Pinus ponderosa* Laws, Pinaceae) is widespread in the montane regions of the western United States, Mexico, and Canada. Pines are particularly well suited to the study of gene movement and historical separation for two reasons. First, the Pinaceae show opposite uniparental inheritance of mitochondrial (mtDNA) and chloroplast DNA (cpDNA). Mitochondrial polymorphisms are inherited maternally (Wagner 1992) and will be dispersed only by movement of seeds. In contrast, cpDNA polymorphisms are paternally inherited (Wagner 1992), and thus cpDNA will be dispersed through movement of both seeds and pollen. Therefore, comparing the relative amounts of differentiation at cpDNA and mtDNA allows an estimate of the relative rates of gene flow through seed and pollen (Ennos 1994; McCauley 1994) and permits a complete accounting of gene movement between the two varieties.

Second, there is biogeographic and fossil evidence suggesting that ponderosa pine consists of two historically separate groups. Two varieties have been identified: *P. p. ponderosa* Laws, which grows from southern California, up the Sierra Nevada, and throughout the northern plateau of Oregon and Idaho; and *P. p. scopulorum* Engelm., which grows throughout the interior Rocky Mountain ranges, western Great Plains, and in isolated mountain ranges of the desert southwest (Conkle and Critchfield 1988). These two varieties are separated by the Great Basin, but meet in a transition zone in west-central Montana (Fig. 2). This region corresponds to a climatic shift between a relatively maritime climate west of the Continental Divide and a more continental climate to the east (Rehfeldt 1984). Thus, gene flow between the two varieties will introduce genes into potentially different adaptive regimes. The varieties are interfertile, although they show some reduction of seed set on crossing (Conkle and Critchfield 1988). Fossil evidence suggests that ponderosa pine did not grow in the Great Basin during the most recent glaciation (Thompson 1990). Moreover, the spread of ponderosa pine northward was only relatively recent (Barnosky et al. 1987; Betancourt et al. 1990). For example, the range of ponderosa pine did not extend north of New Mexico until as recently as 10,000 yr ago. Thus it seems reasonable to postulate that the two varieties of ponderosa pine were historically separate during the most recent glaciation event and that the transition zone represents a region of recent secondary contact between these two. Phylogeographic evidence supports this notion, indicating that the Northern Plateau and Rocky Mountain populations of pon-

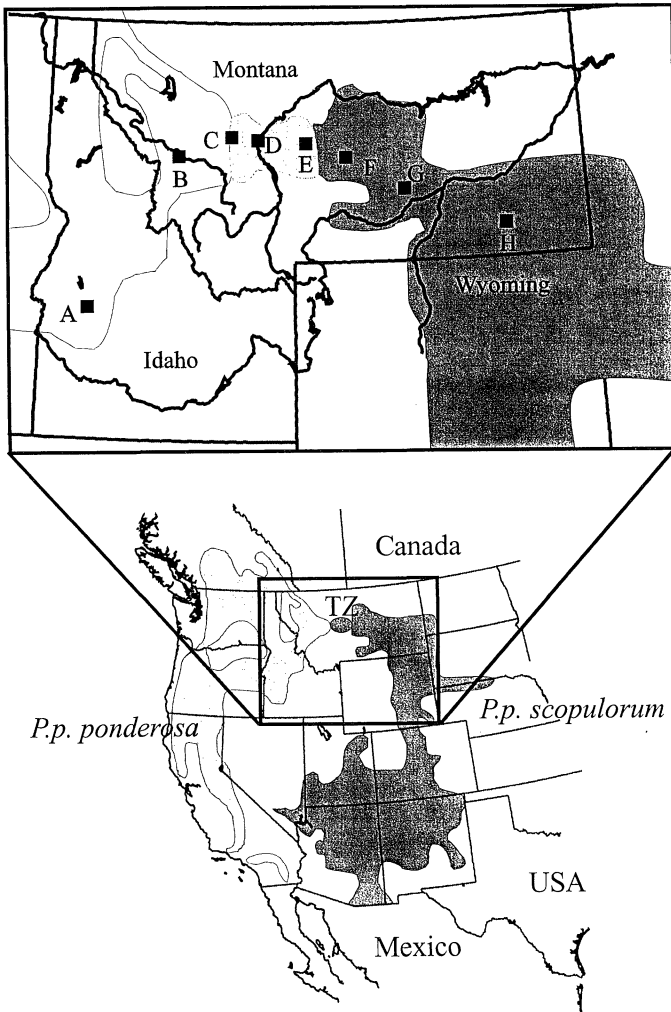


FIG. 2. Map of the approximate distribution of ponderosa pine in western North America. Locations of populations sampled in this study are given with letters that correspond to Table 1.

derosa pine represent separate phylogenetic lineages (S. Brunfeld, pers. comm.). Therefore the transition zone can be used as an example of secondary contact following historical isolation.

#### Genetic Data

In September 1996, we collected foliage samples from a transect of populations running through the transition zone. Eight populations were sampled (Fig. 2, Table 1), two west of, three within, and three east of the transition zone, as it was defined by Conkle and Critchfield (1988). Foliage tissue was collected from each tree sampled (sample sizes in Table 1), and returned to the laboratory within 1 week. Samples were stored at 4°C until DNA extraction.

We used length (insertion/deletion) variation in the cpDNA and mtDNA as molecular markers to track the degree of seed and pollen introgression through the transition zone since secondary contact. Variation in the mtDNA was assayed in the second intron of *nad1*, with the primers of Demesure et al. (1995), whereas cpDNA was found in the region between

TABLE 1. Number of each haplotype observed in each population sampled in a transect through the transition zone in ponderosa pine. Each observed haplotype is classified as being primarily eastern (east), primarily western (west), or unknown (unk). Map codes refer to the symbols in Figure 2. Haplotypes are illustrated in Figure 3.

Population	Location	Map code	mtDNA			cpDNA		
			a West	b West	c East	a West	b Unk	c East
Cascade	West	A	51	0	0	50	1	0
Blackfoot Range	West	B	0	46	0	41	6	0
Rogers Pass	Trans	C	0	21	0	18	2	2
Holter Dam	Trans	D	0	47	0	19	4	24
Monarch	Trans	E	0	4	35	10	3	32
Lewistown	East	F	0	0	50	5	3	39
Roundup	East	G	0	0	25	0	2	23
Lame Deer	East	H	0	0	49	0	8	40

*rbcl* and *accD* (formerly ORF106) using the primers of Arnold et al. (1991). Fragments were amplified in a PCR reaction, digested with *Rsa* I (mtDNA) or *Taq* I (cpDNA) to increase resolution of small size differences, and separated on polyacrylamide gels visualized with silver staining. Laboratory methods were as described for limber pine in Latta and Mitton (1997). Organellar haplotypes were treated as alleles at a single nonrecombining locus in the analyses.

Because of time and resource constraints, allozyme data were obtained from the literature rather than from the samples taken from the transect. Allozyme allelic frequencies for ponderosa pine are summarized by Conkle and Critchfield (1988) and Niebling and Conkle (1990) for 15 loci (Table 2). Data are available for three populations within the western variety, *P. p. ponderosa*, and of summary data for the eastern variety, *P. p. scopulorum*. Robertson's (1975) hypothesis predicts that  $F_{ST}$  among the three populations of *P. p. ponderosa* will show a significantly lower variance across loci than will  $F_{ST}$  between the two varieties.  $F_{ST}$  was calculated for each locus separately using BIOSYS (Swofford and Selander 1989). The significance of a difference between two variances can be tested statistically using an *F*-ratio test (Zar 1984). However, because  $F_{ST}$  is a proportion, it has a lower sampling variance near zero or one, than at intermediate values. To correct for the inflated sampling variance of  $F_{ST}$  in the between-variety estimates, we applied an arcsine square-root transformation to the  $F_{ST}$ -values before conducting the test.

#### RESULTS

When the second intron of *nad1* was digested with *Rsa* I, two of the resulting fragments were polymorphic, giving three composite mtDNA haplotypes within the transect. These polymorphisms appear to represent length variation due to insertion or deletion mutations, because the same size differences were observed regardless of which restriction enzyme was used to digest the sample. The different haplotypes are illustrated in Figure 3a. Two of the mtDNA haplotypes (A and B) were observed primarily west of the transition zone (Table 1). Both of these haplotypes share the same size morph for the smaller fragment. The third haplotype (C) was observed primarily east of the transition zone and differed from the western haplotypes in both of the variable fragments.

TABLE 2. Differentiation of allelic frequencies at cpDNA, mtDNA, and 15 allozyme loci within *Pinus ponderosa ponderosa* and between *P. p. ponderosa* and *P. p. scopulorum*. Data are from Niebling and Conkle (1990).  $F_{ST}$ -values were arcsine square-root transformed before calculating the  $F$ -ratio.

	Within		Between	
	$F_{ST}$	$Nm$	$F_{ST}$	$Nm$
cpDNA <sup>1</sup>	0.033	7.33	0.652	0.13
mtDNA <sup>1</sup>	1.00	0.00	1.00	0.00
<i>Aap2</i>	na <sup>2</sup>		0.007	35.46
<i>Adh2</i>	0.032	7.56	0.212	0.92
<i>Gdh</i>	0.001	249.75	0.109	2.04
<i>Got1</i>	0.009	27.52	0.006	41.41
<i>Got2</i>	0.001	249.75	0.002	124.75
<i>Got3</i>	0.001	249.75	0.073	3.17
<i>Idh</i>	0.077	2.99	0.215	0.91
<i>Lap</i>	na		0.003	83.08
<i>Mdh2</i>	0.027	9.00	0.017	14.45
<i>Mdh3</i>	0.001	249.75	0.001	249.75
<i>Mdh4</i>	0.001	249.75	0.001	249.75
<i>Mpi</i>	na		0.026	9.36
<i>6Pgd2</i>	0.002	124.75	0.033	7.32
<i>6Pgd3</i>	0.004	62.25	0.001	249.75
<i>Pgi2</i>	0.041	5.84	0.001	249.75
<i>Pgm</i>	0.047	5.06	0.282	0.63
Mean	0.0187		0.0618	
Variance	0.000592		0.008595	

$F$ -ratio test for equal variances:  $F_{15,12} = 4.02$ ,  $P < 0.01$ .

<sup>1</sup> For comparison with the allozymes, populations in the transition zone were not included in calculating  $F_{ST}$  for organellar DNA.

<sup>2</sup> *Aap2*, *Lap*, and *Mpi* were monomorphic within *P. p. ponderosa*.

A fourth mtDNA haplotype (D) was identified in samples collected in Colorado (Latta et al. 1998), but was not observed in the transect. This haplotype displayed the same size morph as the smaller band as did the third haplotype (C). Thus, there appear to be characteristically "eastern" and "western" mtDNA haplotypes, which confirm the historical east-west division in ponderosa pine as described by Conkle and Critchfield (1988).

Only one of the restriction fragments in the cpDNA was variable, and again appeared to contain length variation due to insertion or deletion mutations (Fig. 3b). Three haplotypes were observed, one that occurred primarily west of the transition zone (A), one east (C), and one that was intermediate in size and was observed at low frequency within the transition zone (B). Some size variation was observed within each of these groupings, but this variation was too small to be reliably scored and was not great enough to overlap between categories.

The cline between eastern and western haplotypes in the transition zone is very steep for both mtDNA and cpDNA (Fig. 4). All but one population was monomorphic for mtDNA (Table 1). The only variable population was at Monarch, Montana, which is the easternmost population within the transition zone and contained a few copies of the western mtDNA type. The cline for paternally inherited cpDNA was less steep, with three of the populations polymorphic. The Holter Dam and Monarch populations in the transition zone contained intermediate frequencies of eastern and western types. In addition, the Lewistown population to the east of the transition zone showed some introgression of western

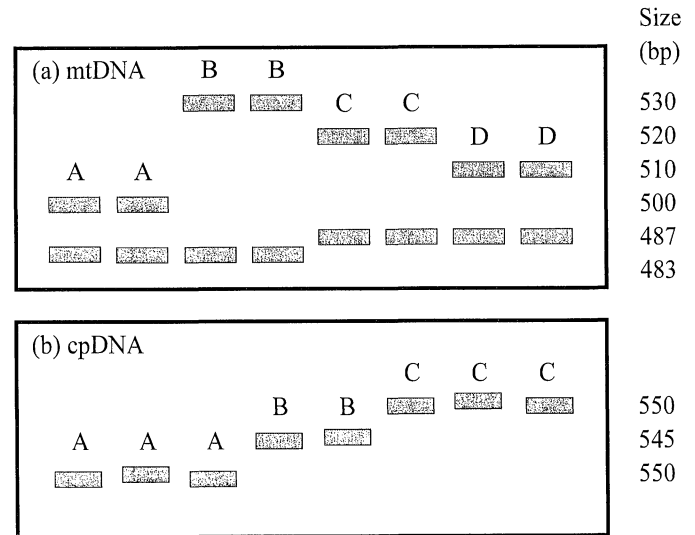


FIG. 3. Length variation at restriction fragments of (a) the second intron of *nad1* (mtDNA), and (b) the noncoding region between *rbcL* and *accD* (cpDNA).

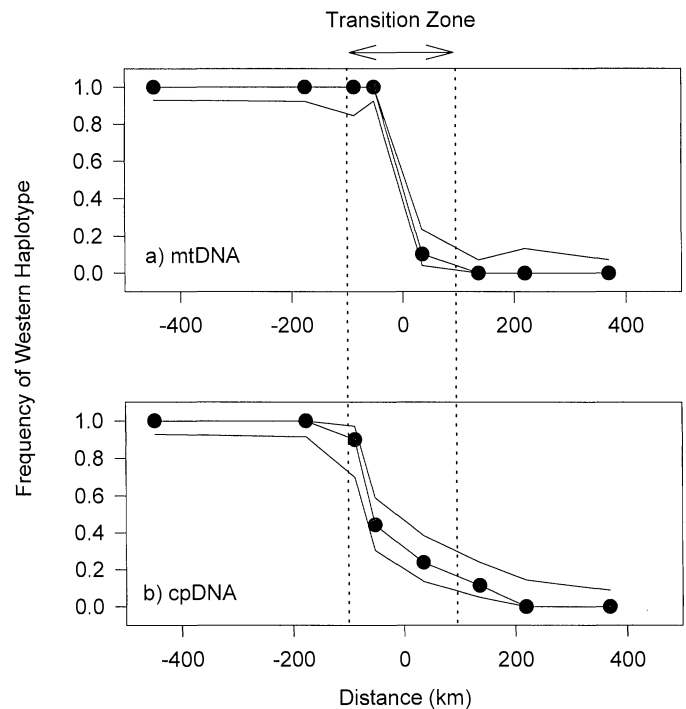


FIG. 4. Clines of haplotype frequencies through the transition zone for (a) maternally inherited mtDNA, and (b) paternally inherited cpDNA. Black circles indicate observed frequency of the western haplotypes, and lines represent the 95% confidence limits (Zar 1984). The dashed vertical lines represent the limits of the transition zone as defined by Conkle and Critchfield (1988). The western edge of the transition zone corresponds to the Continental Divide. Population locations are given as distance (km) east of the middle of the transition zone. cpDNA haplotype B could not be assigned to either the "eastern" or "western" group (Table 1) and therefore was omitted from this figure.

cpDNA types. These patterns are consistent with higher gene movement via pollen than seed and suggest that propagule movement has been greater with the prevailing wind (west to east) than against it. Although a few pollen grains appear to have moved from east to west (i.e., a few eastern cpDNA haplotypes were observed at Holter Dam and Rogers Pass), this was less frequent than west-to-east movement, and no eastern haplotypes were observed west of the Continental Divide (Fig. 4).

Several allozyme loci show differentiation across the transition zone, but overall differentiation is significantly lower for allozymes than for organellar DNA (Table 2).  $F_{ST}$  for the organellar DNA markers falls outside the 95% confidence limits derived for the allozymes. Although  $F_{ST}$  was higher between varieties than within,  $F_{ST}$  for allozyme loci was low (Table 2) and for none of the loci was more than 30% of the variance attributable to differences between the varieties. However, the variance of  $F_{ST}$  across loci is greater for the between-varieties measures than for differentiation among regions within *P. p. ponderosa* (Table 2). There was significantly greater locus-to-locus variance in  $F_{ST}$  values between the two historically separate groups than within ( $F_{15,12} = 4.02$ ,  $P < 0.01$ ).

#### DISCUSSION

Allozymes and organellar DNA show contrasting levels of differentiation between the two varieties of ponderosa pine. Organellar DNA shows very little introgression across the transition zone (Fig. 4). Clearly, movement of both seed and pollen since secondary contact has been relatively limited and certainly has not been sufficient to homogenize the two historically separate groups of ponderosa pine in the time since they came back into contact. In contrast, allelic frequency differentiation is minimal for most allozyme loci, a result that would be consistent with much higher levels of gene exchange. Nevertheless, the historical separation of the two groups does appear to have resulted in a higher interlocus variance in  $F_{ST}$  between as opposed to within varieties (Table 2).

Gene flow patterns inferred from the organellar DNA variation are consistent with expectations based on the movement of propagules. Both seeds and pollen are wind dispersed in ponderosa pine. Consequently, both cpDNA and mtDNA results suggest that movement with the prevailing wind (west to east) has been greater than movement upwind (Fig. 4), although this may instead reflect clines of earlier to later flowering times (J. E. Rehfeldt, pers. comm.). In addition, pollen movement has been much more extensive than seed movement, which is consistent with the expected greater dispersal ability of the much lighter pollen. Seed movement has been negligible through the transition zone, with the only evidence of seed movement being the few western mtDNA haplotypes observed in the eastern half of the transition zone. In seven populations of limber pine, Latta and Mitton (1997) found much greater differentiation of mtDNA than cpDNA, which is consistent with pollen being the main agent of gene flow. Dong and Wagner (1994) found a similar result among populations of jack pine and lodgepole pine. Several studies in angiosperms, in which only maternally inherited organelles

are available, have documented higher differentiation of the maternally inherited organelles than of biparentally inherited allozymes, which is again consistent with pollen being the main agent of gene flow (Ennos 1994; El Mousadik and Petit 1996).

The very limited movement of propagules, particularly seeds, through the transition zone stands in contrast with the apparently very rapid northward expansion of *P. p. scopulorum* following the glacial retreat. Fossil estimates place the northern limit of *P. ponderosa* at approximately northern New Mexico as recently as 10,000 yr ago (Betancourt et al. 1990). Thus, seed movement was sufficiently great during range expansion to move the northern limit of ponderosa pine from New Mexico to Montana in less than 10,000 yr. This corresponds to roughly 100–200 m/yr, a rate comparable to that observed in other tree species (Davis 1981). In contrast, seed movement between established stands of the two varieties has been sufficiently limited to have introgressed less than 100 km in the time since the varieties came back into contact. It is possible that seed movement into and establishment in empty habitat is less restricted than movement into and establishment in an already existing population, simply because recruitment is enhanced in empty habitat. Alternatively, the introgression of genes from one variety into the other may be constrained by selection; *P. p. ponderosa* is characterized by higher relative growth rate, but lower cold tolerance than *P. p. scopulorum* (Rehfeldt 1984). Thus, restricted seed movement may represent local adaptation and reduced fitness of immigrants. Under such an interpretation, the apparently greater pollen movement would represent the ability of  $F_1$  hybrids to tolerate conditions intermediate between the two parents and, thus, establish further from the transition zone.

Conkle and Critchfield (1988) report reduced seed set when *P. p. ponderosa* was hand-pollinated with pollen from *P. p. scopulorum* as compared to pollen from *P. p. ponderosa*. If such an impediment to crossing occurs in nature, this could present a significant barrier to gene flow, especially by pollen, because pollen must contact native trees to effect gene flow. However, most migrants are pollen grains that have mated with native trees upon arrival and produced viable progeny. There is no evidence of assortative mating in the organellar DNA data presented here—the four migrant (western) mtDNA haplotypes observed at Monarch (Table 1) carried the nonmigrant (eastern) cpDNA haplotype. Thus, the original immigrant seed(s) must have established and been pollinated by native trees to produce offspring that were sampled in this survey.

The availability of both paternally and maternally inherited organellar markers in the Pinaceae provides a useful check on assumptions about gene flow and population structure. In most plants, only maternally inherited organelles are present. Thus, levels of pollen movement must be derived from the difference between the amount of genetic differentiation observed at nuclear markers and that at the maternally inherited organelle (Ennos 1994; McCauley 1994). Allozyme allelic frequencies give a misleading impression of extensive gene flow across the transition zone. Estimates of  $Nm$  between the varieties (Table 2) range from 0.63 (or about one migrant every other generation; Pgm; see Table 2), to several hundred, with a median of 14.5. Had only a maternally inherited mark-

er been available, we would have concluded that gene flow via migrant pollen is extensive. However, the steep cline in cpDNA haplotype frequencies indicates that this is not the case. Pollen movement is more extensive than seed movement, but is still limited. Therefore, the two varieties do not appear to be in migration/drift equilibrium, making it difficult to infer levels of migration from  $F_{ST}$  of nuclear loci. A similar finding was reported by Hong et al. (1994) in Bishop pine (*Pinus muricata*) along a north-south transect in California. Both cpDNA and mtDNA showed very high levels of differentiation, but nuclear encoded allozymes were relatively undifferentiated in this species (Millar et al. 1988).

Paradoxically, however, allozymes show less differentiation than expected, rather than more. It is apparent from both the fossil and the organellar DNA evidence that the transition zone represents a zone of recent secondary contact between the *P. p. scopulorum* and *P. p. ponderosa*. Thus, contemporary gene flow, although limited, is higher than it was historically. In this case, we would expect genetic differentiation to be higher than what is consistent with the level of contemporary gene flow. However, allozyme  $F_{ST}$  is considerably lower than what is consistent with the gene flow patterns inferred from the organellar DNA. There are two possible explanations for this inconsistency, both of which relate to the rate of divergence during the historical separation. First, some form of selective sweep may have accelerated the differentiation of the organellar haplotype frequencies. Alternatively, balancing selection may have constrained or slowed the differentiation of allozyme allelic frequencies.

It is not known how long the two varieties were isolated into separate refugia during the Pleistocene. If a generation time of 100 yr is assumed and separation lasted 100,000 yr (the duration of the Wisconsin glaciation), then separation could have been as brief as 1000 generations. Very much longer separation is possible if generation times are shorter or if there was no contact established during the previous (Sangamon) interglacial. Because haploid organelles have a smaller effective population size, it is expected that their haplotype frequencies will drift more rapidly than allelic frequencies at a diploid nuclear locus. However, in a monoecious species, the difference in effective population size is only twofold, which seems unlikely to be enough to account for the pronounced discordance between the allozyme and organellar results. Because the organellar DNA does not recombine, the rise of a novel beneficial mutation (Begun and Aquadro 1992) or the elimination of deleterious mutations (Charlesworth 1994) can accelerate the fixation of a particular haplotype. In either case, this may have caused the organellar haplotype frequencies to differentiate faster than the neutral expectation and, thus, become fixed for alternate haplotypes before the allozymes.

Alternatively, it is possible that differentiation of allozyme allelic frequencies was retarded by some form of balancing selection. This need not imply overdominant or frequency-dependent selection at allozyme loci. Instead, it is sufficient that: (1) different alleles have better metabolic function in different environments; and (2) that the environmental conditions fluctuate through time. This is less restrictive than the case for an equilibrium balanced polymorphism and simply requires that fixation of one allele over the other be

slowed long enough that it does not occur before secondary contact. There is good evidence that different allozyme alleles can have different metabolic activity in different environments (Hilbish and Koehn 1985; Powers et al. 1993) and the fluctuation of the environment through time (particularly with respect to climate) is well established (Pielou 1991).

These two hypotheses make contrasting predictions about the distribution of coalescence times within and between varieties. A selective sweep at organellar DNA would cause coalescence times to be very much smaller within varieties than between varieties. In contrast, balancing selection at allozyme loci would create coalescence times that were greater within varieties than between.

#### *Identifying Historical Divisions from Allelic Frequencies*

Theoretical results suggest that historical separation between populations will result in a high locus-to-locus variance in  $F_{ST}$  (Robertson 1975, Fig. 1). To our knowledge, this result has never been empirically tested in a system for which there is independent evidence of a historical separation. In ponderosa pine, morphological (Conkle and Critchfield 1988), fossil (Barnosky et al. 1987; Betancourt et al. 1990), and phylogeographic (S. Brunsfeld, pers. comm.) evidence all point to the transition zone being a zone of secondary contact between historically separate varieties.  $F_{ST}$  calculated for 15 loci across the transition zone show a significantly higher variance than does  $F_{ST}$  calculated for the same loci among populations within one variety (Table 2). Thus, allozyme loci in ponderosa pine provide evidence in support of Robertson's (1975) theory.

Cavalli-Sforza (1966) and Lewontin and Krakauer (1973) originally proposed that a high locus-to-locus variance in  $F_{ST}$  could be taken as evidence that natural selection was shaping population differentiation in at least some of the loci. It was in response to this suggestion that Robertson (1975) developed the theory concerning historical divisions and  $F_{ST}$  to demonstrate that natural selection need not be invoked to explain discordant patterns among nuclear loci. As a greater appreciation of the role of historical migrations in shaping extant population structure has developed, it has become more important that we be able to distinguish between past and present gene flow when interpreting surveys of electrophoretic variation among populations. Therefore, we suggest that the Lewontin Krakauer (1973) test could instead be used to infer historical separation rather than natural selection.

Of course, a high locus-to-locus variance of  $F_{ST}$  could result from either selection or historical division. The key difference between the two would lie in the concordance of the division outlined by those loci showing differentiation. For example, Latta and Mitton (1997) found four exceptionally differentiated RAPD loci in a survey of seven populations of limber pine. All other loci (10 allozymes and five RAPDs) were relatively undifferentiated among populations. Because all of the differentiated RAPD loci (as well as mtDNA variation) gave the same division among populations (the westernmost population vs. the rest), Latta and Mitton (1997) inferred that the differentiation was the result of an historical division, most likely into eastern and western glacial refugia. Similarly, Karl and Avise (1992) showed marked

variation in  $F_{ST}$ -values between allozyme and anonymous nuclear DNA loci in the American oyster (*Crassostrea virginica*) in Florida. Although in both limber pine and oysters the contrast between protein and DNA markers is puzzling, in both cases the loci with high  $F_{ST}$ -values reveal a concordant break dividing the populations into two groups (Hare and Avise 1996), which is consistent with a hypothesis of secondary contact following historical separation. In contrast, in *Drosophila*  $F_{ST}$ -values vary widely among allozyme loci (Singh and Rhomberg 1987), but do not reveal concordant geographical divisions and are not concordant with differentiation inferred from mtDNA phylogeny (Hale and Singh 1991). Thus, Hale and Singh (1991) conclude that "natural selection... remains the best explanation" for the variation in  $F_{ST}$  across loci in *Drosophila*. A detailed analysis of a large database of allozyme surveys (e.g., Hamrick and Godt 1990) might yield useful insights into the relative frequency with which historical separation and natural selection create discordant levels of genetic differentiation across loci.

## ACKNOWLEDGMENTS

RGL is very grateful to R. G. Olmstead for encouraging him to think about historical factors influencing population structure and to J. B. Mitton, Y. B. Linhart, J. Rehfeldt, D. J. Schoen, and G. Stirling for comments on an earlier draft of the manuscript. Funding for this research was provided by a National Science Foundation Doctoral Dissertation Improvement Grant.

## LITERATURE CITED

- ARNOLD, M. L., C. M. BUCKNER, AND J. J. ROBINSON. 1991. Pollen-mediated introgression and hybrid speciation in Louisiana irises. *Proc. Natl. Acad. Sci. USA* 88:1398-1402.
- AVISE, J. C. 1992. Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos* 63:62-76.
- BARNOSKY, C. W., P. M. ANDERSON, AND P. J. BARTLEIN. 1987. The northwestern U.S. during deglaciation: vegetational history and paleoclimatic implications. Pp. 289-321 in W. F. Ruddiman and H. E. Wright Jr., eds. *North America and adjacent oceans during the last deglaciation*. Geological Society of America, Boulder, CO.
- BEGUN, D. J., AND C. F. AQUADRO. 1992. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* 356:519-520.
- BETANCOURT, J. L., T. R. VAN DEVENDER, AND P. S. MARTIN. 1990. *Packrat middens: the last 40,000 years of biotic change*. Univ. of Arizona Press, Tucson.
- CAVALLI-SFORZA, L. 1966. Population structure and human evolution. *Proc. Roy. Soc. Lond. B Biol. Sci.* 164:362-379.
- CHARLESWORTH, B. 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genet. Res. Camb.* 62:213-227.
- COMES, H. P., AND R. J. ABBOTT. 1998. The relative importance of historical events and gene flow on the population structure of a Mediterranean ragwort, *Senecio gallicus* (Asteraceae). *Evolution* 52:355-367.
- CONKLE, M. T., AND W. B. CRITCHFIELD. 1988. Genetic variation and hybridization of ponderosa pine. Pp. 27-43 in D. M. Baumgartner and J. E. Lotan, eds. *Ponderosa pine: the species and its management*. Washington State Univ. Press, Pullman.
- DAVIS, M. B. 1981. Quaternary history and the stability of forest communities. Pp. 132-153 in D. C. West, H. H. Shugart, and D. B. Botkin, eds. *Forest succession: concepts and application*. Springer Verlag, New York.
- DEMESURE, B., N. SODZI, AND R. J. PETIT. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Mol. Ecol.* 4:129-131.
- DONG, J., AND D. B. WAGNER. 1994. Paternally inherited chloroplast polymorphism in *Pinus*: estimation of diversity and population subdivision and tests of disequilibrium with a maternally inherited mitochondrial polymorphism. *Genetics* 136:1187-1194.
- EL MOUSADIK, A., AND R. J. PETIT. 1996. Chloroplast DNA phylogeography of the argan tree of Morocco. *Mol. Ecol.* 5:547-555.
- ENNOS, R. A. 1994. Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* 72:250-259.
- HALE, L. R., AND R. S. SINGH. 1991. A comprehensive study of genic variation in natural populations of *Drosophila melanogaster*. IV. mitochondrial DNA variation and the role of history vs. selection in the genetic structure of geographic populations. *Genetics* 129:103-117.
- HAMRICK, J. L., AND M. J. W. GODT. 1990. Allozyme diversity in plant species. Pp. 43-63 in A. H. D. Brown, M. T. Clegg, A. L. Kahler, and B. S. Weir, eds. *Plant population genetics, breeding and genetic resources*. Sinauer, Sunderland, MA.
- HAMRICK, J. L., AND M. D. LOVELESS. 1989. The genetic structure of tropical tree populations: associations with reproductive biology. Pp. 129-146 in J. H. Bock and Y. B. Linhart, eds. *Plant evolutionary ecology*. Westview Press, Boulder, CO.
- HARE, M. P., AND J. C. AVISE. 1996. Molecular genetic analysis of a stepped multilocus cline in the American oyster (*Crassostrea virginica*). *Evolution* 50:2305-2315.
- HILBISH, J. T., AND R. K. KOEHN. 1985. The physiological basis of natural selection at the LAP locus. *Evolution* 39:1302-1317.
- HOLSINGER, K. E., AND R. J. MASON-GAMER. 1996. Hierarchical analysis of nucleotide diversity in geographically structured populations. *Genetics* 142:629-639.
- HONG, Y. P., V. D. HIPKINS, AND S. H. STRAUSS. 1994. Chloroplast DNA diversity among trees, populations and species in the California closed cone pines (*Pinus radiata*, *P. muricata* and *P. attenuata*). *Genetics* 135:1187-1196.
- KARL, S. A., AND J. C. AVISE. 1992. Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. *Science* 256:100-102.
- LATTA, R. G., AND J. B. MITTON. 1997. A comparison of population structure across four classes of gene marker in limber pine. *Genetics* 146:1153-1163.
- LATTA, R. G., Y. B. LINHART, D. FLECK, AND M. ELLIOT. 1998. Direct and indirect estimates of seed versus pollen movement in ponderosa pine. *Evolution* 52:61-67.
- LEWONTIN, R. C., AND J. KRAKAUER. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* 74:175-195.
- MCCAULEY, D. E. 1994. Contrasting the distribution of chloroplast DNA and allozyme polymorphism among local populations of *Silene alba*: implications for studies of gene flow in plants. *Proc. Natl. Acad. Sci. USA* 91:8127-8131.
- MILLAR, C. I., S. H. STRAUSS, M. T. CONKLE, AND R. D. WESTFALL. 1988. Allozyme differentiation and biosystematics of the Californian closed cone pines (*Pinus* subsect. *Oocarpae*). *Syst. Bot.* 13:351-370.
- NIEBLING, C. R., AND M. T. CONKLE. 1990. Diversity of Washoe pine and comparisons with allozymes of ponderosa pine races. *Can. J. For. Res.* 20:298-308.
- PIELOU, E. C. 1991. *After the ice age: the return of life to glaciated North America*. Univ. Chicago Press, Chicago.
- POWERS, D. A., M. SMITH, I. GONZALEZ-VILLASENOR, L. DiMICHELE, D. CRAWFORD, G. BERNARDI, AND T. LAUERMAN. 1993. A multidisciplinary approach to the selectionist/neutralist controversy using the model teleost, *Fundulus heteroclitus*. Pp. 49-109 in D. Futuyama and J. Antonovics, eds. *Oxford Surveys in Evolutionary Biology*. Vol. 9. Oxford Univ. Press, Oxford.
- REHFELDT, J. E. 1984. Microevolution of conifers in the northern Rocky Mountains: a view from common gardens. Pp. 132-146

- in R. M. Lanner, ed. Proceedings of the eighth North American forest biology workshop. Utah State Univ. Press, Logan, UT.
- ROBERTSON, A. 1975. Gene frequency distributions as a test of selective neutrality. *Genetics* 81:775-785.
- SINGH, R. S., AND L. R. RHOMBERG. 1987. A comprehensive study of genic variation in natural populations of *Drosophila melanogaster*. II. estimates of heterozygosity and patterns of geographic differentiation. *Genetics* 117:255-271.
- SLATKIN, M. 1994. Gene flow and population structure. Pp. 3-17 in L. A. Real, ed. *Ecological genetics*. Princeton Univ. Press, Princeton, NJ.
- SOLTIS, D. E., M. A. GITZENDANNER, D. D. STRENGE, AND P. S. SOLTIS. 1997. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Pl. Syst. Evol.* 206:353-373.
- SWOFFORD, D. L., AND R. K. SELANDER. 1989. BIOSYS-1: A computer program for the analysis of allelic variation in population genetics and biochemical systematics. Release 1.7. Univ. of Illinois Press, Urbana.
- THOMPSON, R. S. 1990. Late Quaternary vegetation and climate in the great Basin. Pp. 200-239 in J. L. Betancourt, T. R. Van Devender and P. S. Martin, eds. *Packrat middens: the last 40,000 years of biotic change*. Univ. of Arizona Press, Tucson.
- WAGNER, D. B. 1992. Nuclear, chloroplast and mitochondrial DNA polymorphisms as biochemical markers in population genetic analyses of forest trees. *New Forests* 6:373-390.
- WAPLES, R. S. 1987. A multispecies approach to the analysis of gene flow in marine shore fishes. *Evolution* 41:385-400.
- WRIGHT, S. 1951. The genetical structure of populations. *Ann. Eugen.* 15:323-354.
- ZAR, J. H. 1984. *Biostatistical analysis*. 2d ed. Prentice Hall, Englewood Cliffs, NJ.

Corresponding Editor: O. Savolainen