Comparative genetic structure in pines: evolutionary and conservation consequences

INTRODUCTION

The development of studies on the genetic structure in plants has been associated with that of the molecular markers. Studies using allozymes as markers started in the early 1970s and during the next three decades a great amount of data has accumulated (Hamrick & Godt 1990) that have...
shown some generalizations. The first one refers to the inverse relationship between outcrossing rates and genetic differentiation (Govindaraju 1988), so that species of tropical trees or trees in general show an outcrossing rate close to one and genetic differentiation estimates (usually estimated through allelic frequencies variance or $F_{ST}$) very close to zero (e.g., Furnier & Adams 1986, Eguiarte et al. 1993, Ledig 1998).

Hamrick and coworkers (e.g., Hamrick & Godt 1990) have made several generalizations between genetic structure and life history characteristics. An extraordinary homogeneity in expected heterozygosity has been observed. This pattern has been explained in light of Ohta’s nearly neutral theory of molecular evolution (Ohta 1995). Also, variation in outcrossing rates has shown a bimodal distribution. Schemske & Lande (1985) first described this bimodality. Other generalizations show that trees, perennial plants and widespread species have larger genetic variation. In particular, most pine species have been studied in terms of their genetic structure and the above results have been confirmed in this genus (Ledig 1998).

Some exceptions to the above generalizations are represented by species that are rare or that have passed through a bottleneck like Pinus torreyana or Pinus resinosa. Contrary to expectations based on their rarity (but see Comps et al. 2001), studies on the genetic structure on Mexican and North American pinyon pines have shown high genetic variation (heterozygosities between 0.216 and 0.220) except for $P. \ edulis$ (0.03, Premoli et al. 1994) and a high population differentiation ($F_{ST}$ between 0.18 and 0.25). This genetic structure coincides with the one found in other rare conifer species like Picea chihuahua (Ledig et al. 1997) and suggests that the outcrossing rate is less than one as found for the few species where it has been estimated (Ledig et al. 1999 for $P. \ maximartinezii$, Ledig et al. 2001 for $P. \ pinceana$). These data can now be studied under a phylogenetic approach. In particular, recent phylogenetic studies of pines using molecular markers (e.g., Liston et al. 1999) can be used as a historical framework to ask questions about the origin and adaptive value of the genetic structure.

Pines represent an interesting system to study the effect of different ecological characteristics on the genetic structure. This genus grows naturally in the northern hemisphere and species can be found as far north as in latitude 71º N (e.g., Pinus sylvestris), in ecosystems where annual rainfall is only 300-400 mm (e.g., $P. \ pinceana$), and longevities that reach 2,500 to 5,000 years (e.g., $P. \ aristata$ and $P. \ longaeva$). The great diversity of environments colonized by this genus and its local abundance make pines a key ecological element in most temperate forests of the northern hemisphere. Until now 110 species have been described. Seventy one of those species grow in America and about 50 in Mexico.

This paper presents a comparative analysis of pine genetic structure using allozymes and chloroplast microsatellites. In particular, the analysis is presented using a phylogenetic approach incorporating recent molecular systematic data. Furthermore, we explore the relevance and evolutionary implications of the use of microsatellite as a comparative tool to further our understanding of the causes and patterns of the genetic structure of these trees.

MATERIAL AND METHODS

Molecular markers

Different markers have been used in plants and in particular in trees to understand their genetic structure. How well is the genome sampled and how sensitive are they to estimate the genetic variation are the main questions that are usually asked to evaluate their ability to estimate genetic variability. The first answer depends on the way the markers are distributed in the genome and the number of loci that is being used. The second one depends directly on the number of base pairs that are sampled. In pines allozymes have been usually the markers used in most studies. Approximately 50 species have been studied with these markers. In some cases all three basic estimates of population genetic structure have been obtained (expected heterozygosity, genetic differentiation or $F_{ST}$ and outcrossing rates). In all studied species at least an estimate of expected heterozygosity was obtained. Recently, chloroplast microsatellites have been used to study these aspects of the genetic structure and data for at least 10 species have been published. However, no outcrossing rates estimates using nuclear microsatellites have been published. Because of the structure of the data, we restricted our analysis to data on allozymes and chloroplast microsatellites that now represent an important source of information for a comparative analysis.

Allozymes

These markers have provided most of the data that we have now on the genetic structure of plants in general and particularly for trees. The polymorphisms are based on the differential mo-
Comparative Genetic Structure in Pines

Allozymes sample more with respect to the number of loci used and their sensitivity is about one fourth of the number of bases sampled. That is, only a base change in four is detected in a gel. For example, a study with 20 allozyme loci samples about (1000 x 20) / 4, or 5,000 nucleotides, assuming that an average gene has 333 amino acids. These are codominant markers.

Microsatellites

The single stranded repeats (SSRs) are DNA sequences that are repeated in tandem a number of times. The larger repeats (until 5 Mb) are called satellites. Intermediate repeats (the repeated unit is more than 10 bp and form blocks of 0.5 to 30 kb) are called minisatellites. Microsatellites have repeated units of 1-8 bp and form structures that have 20 to 100 bp. The number of bases sampled with microsatellites depends on the number of loci used but in general is very low because a region of high variation is sampled and to estimate the same amount of variation a much lower number of base pairs are needed.

Evolutionary models and their assumptions

The genetic estimates of heterozygosity and genetic differentiation depend on the genetic model that is assumed to have occurred during the evolution of the molecular markers. An infinite allele model is usually assumed if allozymes or RFLPs (restriction fragment length polymorphisms) are used but in the case of microsatellite there are good reasons to use a stepwise mutation model based on the way microsatellites are thought to mutate. In particular the model that is thought to be important in these markers is a model based on the slipped-strand misspairing mutation model (Li 1997) that produces an increase or a decrease of a repeat unit (usually a base). Recently many groups have questioned the application of this model based on many factors (Hedrick 1999, Balloux et al. 2000). The first is the possibility of homoplasy or allele convergence. The second refers to the possibility that the mutation mechanism produces alleles that are more than one repeat unit from the original allele. This could happen through insertions or deletions of more than one repeat unit. This process violates the assumptions of the model and makes the estimates biased in the same proportion. As a consequence, both estimates are regularly used, those based on a stepwise mutation model (Ohta & Kimura 1973, Valdés et al. 1993, Slatkin 1995) and on an infinite allele model (based on the estimation model originally developed by Wright [1949] and Kimura & Crow [1964] and further developed by Weir & Cockerham [1988]).

We used a comparative approach to analyze the genetic data on pines. First we used standard statistical tests to compare the estimates (Sokal & Rohlf 1994). Second we grouped the data with respect to their taxonomic status using the classification proposed by Price et al. (1998). Finally we used a published phylogenetic framework for pines (Liston et al. 1999) to further clarify historical patterns in genetic estimates.

RESULTS

Expected heterozygosity and genetic differentiation

Genetic variation estimated through the expected heterozygosity has been obtained for approximately 50 pine species. Ledig (1998) has published a recent review. For 38 species, both heterozygosity and genetic differentiation (using $F_{ST}$ or $G_{ST}$) have been published. In all cases allozymes were used as genetic markers. To these data we added some that were not included in Ledig’s review (Parker & Hamrick 1996, Delgado et al. 1999, Ledig 1999, Ledig et al. 1999, 2001, Molina-Freaner et al. 2001). These data show a distribution close to a Gaussian curve, but statistically, only the expected heterozygosity adjusts to a normal distribution (Fig. 1). Genetic differentiation does not adjust to such a distribution mainly because there is an excess of observed values close to zero. Averages for the genus are 0.198 for the expected heterozygosity and 0.129 for the genetic differentiation (Table 1). On the other hand genetic differentiation estimates are not significantly different ($t_{39} = 0.21, P > 0.5$) between hard (0.127) and soft pines (0.136). It is noteworthy the high variation among species in both the Pinus and the Ponderosas subsections for the expected estimates of heterozygosity. No significant correlation was detected ($r = 0.03, \text{d.f.} = 72, P > 0.05$) between hard (0.127) and soft pines (0.136). It is noteworthy the high variation among species in both the Pinus and the Ponderosas subsections for the expected estimates of heterozygosity. No significant correlation was detected ($r = 0.03, \text{d.f.} = 72, P > 0.05$) between the expected heterozygosity and genetic differentiation (Fig. 2). This correlation remains low when data are partitioned among soft ($r = 0.04, \text{d.f.} = 16, P > 0.05$) and hard pines($r = 0.07, \text{d.f.} = 72, P > 0.05$).

When multiple estimates have been published for a particular species, probably a species is better characterized by the largest published estimate due to subsampling of variation. That is, if a study reports a larger $F_{ST}$ estimate, this estimate
is probably closer to the species mean as some studies do not have a representative sample of both individuals and populations. The correlation between expected heterozygosity and $F_{ST}$ slightly improves when using these data ($r = 0.24$, d.f. = 39, $P > 0.05$) but it is statistically nonsignificant.

When we compared the means of the estimates for the genetic structure for the subsections in Table 1 some generalizations can be made. First, the expected heterozygosities in all subsections of soft pines are larger than the ones obtained for hard pines. Second, genetic differentiation in hard pines is in all cases smaller than the estimated for soft pines. In particular $P$. halepensis, a hard pine, shows an extreme value of 1%. This species is in subsection Halepenses. The estimate was obtained for populations in Greece (Loukas et al. 1983) and either represents a biased sample for the species or this species deserves future attention being such an extreme example of low population differentiation.

Estimates of genetic variation using microsatellites in pines range between 0.411 for $P$. heidrechii var. leucodermis (Powell et al. 1995) to 0.978 for $P$. sylvestris (Provan et al. 1998) with an average of 0.582 (Table 2).

On the other hand, estimates of $F_{ST}$ (using the infinite allele model) range between 0.023 for $P$. pinaster (Ribeiro et al. 2001) and 0.783 for $P$. pineana (Escalante 2001), with an average of 0.245. If we do not consider the estimate for $P$. pineana since probably represents a case of extreme differentiation through fragmentation, the average drops to 0.166. Estimates of $R_{ST}$ based on the stepwise mutation model published until now are only 6. The comparison of these estimates ($F_{ST}$ and $R_{ST}$) show that in half of the species $R_{ST}$ estimates are smaller than $F_{ST}$ (0.047 for $P$. nelsonii, Cuenca 2001; 0.068 for $P$. resinosa, Echt et al. 1998; and 0.212 for $P$. halepensis, Bucci et al. 1998) while in the rest $R_{ST}$ estimates are larger than $F_{ST}$ (0.068 for $P$. brutia subsp. eldarica, 0.289 for $P$. brutia, Bucci et al. 1998; and 0.783 for $P$. pineana, Escalante 2001). In fact both estimates are statistically correlated ($r = 0.985$, d.f. = 4, $P < 0.01$) which suggests that the differences in both models are not important when comparisons are made in these pine species. Furthermore, except for the estimate for $P$. pineana, a slightly negative correlation (statistically nonsignificant, $r = 0.688$, d.f. = 4, $P > 0.10$) is observed between genetic diversity and genetic differentiation as $F_{ST}$.

Microsatellites can be used to amplify the same locus in different species. Table 3 shows averages for six species from the subgenus Pinus and subgenus Strobus species (4) for 11 microsatellite loci (Cuenca 2001, Escalante 2001, Delgado, Vendramin & Piñero unpublished results). The average sizes of microsatellites are usually very low.
similar in the two subgenera except in one locus (Pt26081). Also, the variances for the two subgenera are, in general, quite similar.

**Population inbreeding and mating system**

Population inbreeding or in other words the deviation from Hardy-Weinberg equilibrium could have three causes. First, inbreeding could develop through self-fertilization. Secondly, mating could occur among genetically related individuals and this would produce inbreeding. Thirdly, genetic drift could also increase the level of homozygosity. The published estimates of inbreeding ($F_{IS}$) in pines range from estimates close to zero to 0.14 for *P. pineaena* (Ledig et al. 2001). On the other hand, outcrossing rates estimates range from 0.65 for *P. maximinoi* (Matheson et al. 1989) to values statistically equal to 100% outcrossing. Average of direct estimations using allozymes for 17 species (six from the subgenus *Strobus* and 11 from the subgenus *Pinus*) is 0.88 (SD = 0.10) (Ledig 1998, Ledig et al. 1999, 2001). This distribution is shown in Fig. 3 and fits a normal distribution (Kolmogorov-Smirnov test, $d = 0.16, P > 0.20$). Confidence intervals (95%) are rather narrow (0.82 and 0.93). Average estimates for hard and soft pines are not statistically different (0.89 and 0.86 respectively). For species in which a direct estimate of outcrossing rate has been obtained, we can calculate the expected inbreeding coefficient assuming that all inbreeding is due to the mating system ($f = [1-t] / [1+t]$) and compare both values (Table 4). All inbreeding estimates were obtained in seeds. In general, direct inbreeding estimates coincide with the expected ones based only on the mating system. In some cases like *P. sylvestris* (Muona & Szmidt 1991) and *P. pineaena* (Ledig et al. 2001) both estimates are clearly different. This is probably produced by natural selection in favor of heterozygotes or overdominance (see Ledig 1998 for a revision in pines). This effect probably has different intensity in different species and appears to act in earlier stages of the life cycle in some cases.

**Fig. 1:** Frequency distribution of expected heterozygosity (top) and genetic differentiation (bottom) in pine species from published data using allozymes. Differentiation estimates are expressed in percentage.

Fig. 1: Distribución de frecuencias de la heterocigosidad esperada (arriba) y la diferenciación genética (abajo) en especies de pinos obtenida de datos publicados con marcadores isoenzimáticos. Los datos de diferenciación están multiplicados por 100.

**Fig. 2:** Relation between expected heterozygosity and genetic differentiation in 71 studies in which both estimates were obtained for 38 pine species using allozyme markers. Circles are data for species in the subgenus *Pinus* (hard pines), squares are data for species in the subgenus *Strobus* (soft pines).

Relación entre la heterocigosidad esperada y la diferenciación genética en 71 estudios en los que ambos estimados se han obtenido para 38 especies de pinos usando marcadores isoenzimáticos. Los círculos muestran los datos para especies de pinos del subgénero *Pinus* (pinos duros), los cuadros muestran los datos para las especies del subgénero *Strobus* (pinos blandos).
A historical framework can also be used to compare the mean outcrossing rates in different subsections. In all lineages, rates are in general high. Also, the lowest estimates correspond to *P. cembra* and *P. maximinoi*, (0.69 and 0.65, respectively), which belong to different subgenera (*Strobus* and *Pinus*, respectively).

### DISCUSSION

**Constancy in heterozygosity among species**

There are several sources of evidence that suggest that effective population size determines mostly the level of allozyme variation in a species or in other words, that within a species there is a correlation between expected heterozygosity and population size (Avise 1994). In the case of *Pinus* very few species show estimates of expected heterozygosity different from 0.20. These results are predicted if mutation rates are assumed relatively constant (Hedrick 1999). In particular for *Pinus*, 99% confidence limits for expected heterozygosity in 41 species are 0.189-0.227, which suggests that historically effective population sizes have been constant among species or that have been

### TABLE 2

Average genetic diversity and genetic differentiation estimates for pine species using chloroplast microsatellite loci (standard deviations are shown in parentheses)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Subsection</th>
<th>Diversity</th>
<th>F&lt;sub&gt;st&lt;/sub&gt;</th>
<th>R&lt;sub&gt;st&lt;/sub&gt;</th>
<th>Number of loci</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. sylvestris</em></td>
<td><em>Pinus</em></td>
<td>0.978</td>
<td></td>
<td></td>
<td></td>
<td>Provan et al. (1998)</td>
</tr>
<tr>
<td><em>P. pinaster</em></td>
<td><em>Pinus</em></td>
<td>0.885</td>
<td>0.023</td>
<td>0.068</td>
<td>6</td>
<td>Ribeiro et al. (2001)</td>
</tr>
<tr>
<td><em>P. resinosa</em></td>
<td><em>Pinus</em></td>
<td>0.618</td>
<td>0.121</td>
<td></td>
<td></td>
<td>Echt et al. (1998)</td>
</tr>
<tr>
<td><em>P. heidreichii var. leucodermis</em></td>
<td><em>Pinus</em></td>
<td>0.411</td>
<td>0.22</td>
<td></td>
<td></td>
<td>Powell et al. (1995)</td>
</tr>
<tr>
<td><em>P. halepensis</em></td>
<td><em>Halepenses</em></td>
<td>0.222</td>
<td>0.308</td>
<td>0.323</td>
<td>8</td>
<td>Bucci et al. (1998)</td>
</tr>
<tr>
<td><em>P. brutia</em></td>
<td><em>Halepenses</em></td>
<td>0.297</td>
<td>0.289</td>
<td>0.074</td>
<td>8</td>
<td>Bucci et al. (1998)</td>
</tr>
<tr>
<td><em>P. brutia ssp. Eldarica</em></td>
<td><em>Halepenses</em></td>
<td>0.271</td>
<td>0.068</td>
<td>0.131</td>
<td>4</td>
<td>Cuenca (2001)</td>
</tr>
<tr>
<td><em>P. nelsonii</em></td>
<td><em>Cembroides</em></td>
<td>0.728</td>
<td>0.131</td>
<td>0.930</td>
<td>4</td>
<td>Escalante (2001)</td>
</tr>
<tr>
<td><em>P. pineana</em></td>
<td><em>Cembroides</em></td>
<td>0.824</td>
<td>0.783</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>0.582</td>
<td>0.245</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.289)</td>
<td>(0.241)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean size (in bp) for different chloroplast microsatellite loci in both subgenera of *Pinus*. The loci identification corresponds to the notation of the chloroplast sequence of *Pinus thunbergii* (Wakasugi et al. 1994). Standard deviations are shown in parentheses

<table>
<thead>
<tr>
<th>Locus</th>
<th>Subgenus <em>Pinus</em></th>
<th>Subgenus <em>Strobus</em></th>
<th>All species</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pt1254</em></td>
<td>66.5 (1.22)</td>
<td>67.5 (1.73)</td>
<td>66.9 (1.45)</td>
</tr>
<tr>
<td><em>Pt1520</em></td>
<td>122.8 (2.48)</td>
<td>122 (3.65)</td>
<td>122.5 (2.84)</td>
</tr>
<tr>
<td><em>Pt9383</em></td>
<td>88 (3.35)</td>
<td>88.3 (3.59)</td>
<td>88.1 (3.25)</td>
</tr>
<tr>
<td><em>Pt26081</em></td>
<td>142.3 (1.75)</td>
<td>156.3 (1.26)</td>
<td>147.9 (7.34)</td>
</tr>
<tr>
<td><em>Pt36480</em></td>
<td>144.7 (2.88)</td>
<td>142.5 (2.06)</td>
<td>143.9 (2.64)</td>
</tr>
<tr>
<td><em>Pt41093</em></td>
<td>77.3 (0.82)</td>
<td>76.8 (2.63)</td>
<td>77.1 (1.66)</td>
</tr>
<tr>
<td><em>Pt4821</em></td>
<td>88.5 (0.55)</td>
<td>87.5 (0.58)</td>
<td>88.1 (0.74)</td>
</tr>
<tr>
<td><em>Pt5187</em></td>
<td>79.2 (5.42)</td>
<td>85.8 (2.06)</td>
<td>81.8 (5.41)</td>
</tr>
<tr>
<td><em>Pt63718</em></td>
<td>93.2 (0.41)</td>
<td>91 (0.82)</td>
<td>92.3 (1.25)</td>
</tr>
<tr>
<td><em>Pt71936</em></td>
<td>145.8 (1.60)</td>
<td>142.5 (2.06)</td>
<td>144.6 (2.31)</td>
</tr>
<tr>
<td><em>Pt87268</em></td>
<td>163.7 (2.73)</td>
<td>167.5 (1.00)</td>
<td>165.2 (2.90)</td>
</tr>
</tbody>
</table>
maintained above levels at which increases in population sizes would marginally increase the level of genetic variation.

Other aspects of the life history that could affect the level of genetic variation is individual longevity. Although species have been described (particularly in subsection Balfourianae) in which individual longevity could reach thousands of years, it is more common within the range from tens of years to a few hundred years. Longevity could be one of the causes of the data observed, like the level of expected heterozygosity estimated for *P. longaeva* (0.340, Hiebert & Hamrick 1983).

In this context of relative constancy of estimates for expected heterozygosity, it is relevant to analyze in some detail those species that show a significant deviation from the mean. In general the distribution showed in Fig. 1 shows that those species that have expected heterozygosities below 0.1 are very few and all of them (for example, *P. torreyana*, Ledig & Conkle 1983) are probably the consequence of a bottleneck. This phenomenon has also been described for coulter pine (*P. coulteri*, Ledig 2000) at the intraspecific level.

**Heterogeneity in population differentiation among species**

Genetic differentiation shows a contrasting pattern from that found for expected heterozygosity. The distribution does not fit a normal distribution and is probably due to different causes. Theoretically, both migration and genetic drift determine population differentiation. In some cases, attention has been given to the fact that the estimation of genetic differentiation could be biased if the molecular markers show a high mutation rate (Hedrick 1999). Other aspects like the ecological components of migration and genetic drift could also affect the level of population differentiation. These include, for example, population density and migration agents of both gametes and embryos (Alvarez-Buylla et al. 1996). These factors are quite variable in pines and would explain variation in estimates of genetic differentiation. For example pinyon pines in North America share a low population density, relatively isolated populations and seed dispersal by birds. All these factors cause a relatively high population differentiation.

**Microsatellites: comparative approaches**

Several studies have explored the patterns of microsatellite evolution. Although at the intraspecific level the interpretation of genetic variation is straightforward (slipped-strand misspairing producing one step mutations at high rates), at the interspecific level, particularly when more diver-

### TABLE 4

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Subsection</th>
<th><em>t_m</em></th>
<th><em>f</em></th>
<th><em>F_{IS}</em></th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. contorta</em></td>
<td>Contortae</td>
<td>0.95</td>
<td>0.03</td>
<td>0</td>
<td>Wheeler &amp; Guries (1982)</td>
</tr>
<tr>
<td><em>P. ponderosa</em></td>
<td>Ponderosae</td>
<td>0.96</td>
<td>0.02</td>
<td>0.02</td>
<td>Farris &amp; Mitton (1984)</td>
</tr>
<tr>
<td><em>P. sibirica</em></td>
<td>Cembrae</td>
<td>0.89</td>
<td>0.06</td>
<td>-0.01</td>
<td>Goncharenko et al. (1992)</td>
</tr>
<tr>
<td><em>P. sylvestris</em></td>
<td>Pinus</td>
<td>0.94</td>
<td>0.03</td>
<td>-0.14</td>
<td>Muona &amp; Szmidt (1991)</td>
</tr>
<tr>
<td><em>P. maximartinezii</em></td>
<td>Cembroides</td>
<td>0.82</td>
<td>0.10</td>
<td>0.08</td>
<td>Ledig et al. (1999)</td>
</tr>
<tr>
<td><em>P. pinceana</em></td>
<td>Cembroides</td>
<td>0.86</td>
<td>0.07</td>
<td>0.14</td>
<td>Ledig et al. (2001)</td>
</tr>
</tbody>
</table>
gent species are compared, other factors could be involved. This is because different studies have reported other mechanisms that should be taken into account (Kruglyak et al. 1998, Karhu et al. 2000). Although slipped-strand misspairing was first described as the main mechanism for microsatellite mutation, it has been demonstrated that base substitutions could account for a significant portion of observed substitutions when different species have been compared (Kruglyak et al. 1998). Mutation rates of these two kinds of events differ for various orders of magnitude. While point mutations occur at rates between \(10^{-10} -10^{-8}\), mutations in microsatellites that either increase or decrease the number of repeats have been estimated between \(10^{-5} -10^{-2}\). In particular, SSRs mutation rates of \(10^{-5}\) have been found in pines (Provan et al. 1998). These results show that there must be a difference of about 4 orders of magnitude between the average allelic substitution times for mutations based on the number of repeat copies in a microsatellite and those based on point mutations. This will produce that the transient polymorphisms of both mutation mechanisms will be present in different populations but allele fixation originated from point mutations will occur with higher frequency among divergent species. The comparison of alleles among divergent species has shown this fact (Kahru et al. 2000) but also the existence of a third kind of substitutions (besides size mutations and point mutations), duplications that in highly divergent species will also be present. These mutation mechanisms should be better studied in order to interpret the observed polymorphisms.

The relative proportion of mutation rates with respect to migration rates has been used to explain the incongruence seen in \(F_{ST}\) estimates using uniparental and biparental markers, in particular to explain lower \(F_{ST}\)’s for uniparental markers (Bollux et al. 2000). In pines, differentiation would be expected to be highest in maternally inherited markers (mitochondria) than in paternally inherited markers (chloroplast) (Furnier & Stine 1995). The lowest differentiation would then be expected in nuclear markers. There are in the literature reports of genetic differentiation using nuclear allozymes and chloroplast microsatellites. These data show that there is a positive relation (but statistically non significant, \(r = 0.86,\) d.f. = 2, \(P > 0.10\)) between genetic differentiation using these markers when nuclear data is used as the independent variable for \(P.\) pinaster, \(P.\) pinceana, \(P.\) halepensis and \(P.\) heidrechii var. leucoderma (Table 2, Ledig 1998, Ledig et al. 2001). They also show that differentiation estimates using allozymes are always lower than those obtained from chloroplast microsatellites as was also found for \(P.\) glauca by Furnier & Stine (1995).

Results shown in Table 3 support a more complex microsatellite evolutionary model than expected. From a model with genetic drift and mutation, the expected size for microsatellites from different species would diverge at a rate proportional to both mutation rate and population size and as a consequence variance among species would be larger. Microsatellite average sizes for different hard and soft pine species are equal in six of the 11 loci. In the other 5 loci, there is a significant difference but in two of those loci (Pt26081 and Pt87268) is larger in pines from the subgenus Strobus while in the other three the average size is larger in pines from the subgenus Pinus (Pt4821, Pt63718 and Pt71936). The other prediction would be that the variance for both groups of pines taken together would be larger than the variances for each one of the subgenera. Variances are statistically larger at the 5 % limit (Bartlett’s test) in these five loci when all species are taken together. This probably shows that size mutation appears to be controlling the evolutionary dynamics of microsatellites at these five loci but it also suggests that other factor should explain both the variance homogeneity and also the homogeneity of the sizes of microsatellites at the other six loci. Selection would explain both homogeneities and further studies should explore this possibility.

The most representative parameter of the genetic structure in plants is the rate of outcrossing, estimated as the proportion of seed produced through self-fertilization with respect to those produced through cross-fertilization (Govindaraju 1988). Mating system has consequences on inbreeding levels and through that on the different factors that produce inbreeding depression. For example, a species with a high outcrossing rate, usually has a high frequency of recessive deleterious and lethal alleles and inbreeding depression will be high if self-fertilization increases. These species will normally have decreasing inbreeding coefficients (an increase in heterozygote proportions) through life stages until adulthood.

Plasticity within some pine species or their ability to modify outcrossing rates (Ledig 1998) is probably due to adaptation to contrasting population densities that favour colonization to new environments in which trees self-fertilize more frequently, as has been described for \(P.\) radiata (Bannister 1965 cited in Ledig 1998). This flexibility, on the other hand, has not been strong enough as to modify the mating system in such a way that a species would be predominantly self-fertilized. In fact mating system in pines is so
open that interspecific hybridization is quite common in sympatric species so that hybridization has frequently been proposed as a speciation mechanism in this genus (e.g., Bucci et al. 1998).

Mating system plasticity and morphological adaptation

One of the main conclusions of this work is that pines show a striking homogeneity in their genetic system that includes homogeneity of expected heterozygosity but most of all nearly constant outcrossing rates for species that diverged 135 million years (see for example Hamrick & Godt 1990). Apparently there has been a strong limitation to the reduction of outcrossing rates in pines and probably those populations or species in which this has happened have probably gone extinct.

These conclusions about the genetic structure probably do not apply to morphological characters. For example, while there is low differentiation in molecular markers (allozymes, RAPDs or microsatellites in *P. sylvestris* in Finland, Kahru et al. 1996) there is a strong morphological differentiation probably as a result of strong natural selection on morphological characters.

In view of the knowledge on the genus *Pinus* with respect to their uses, ecological services, morphological adaptations and its conservation status, the group can be considered as model to establish conservation strategies in tree species. Until now, proposals to establish conservation strategies in this genus could be summarized as follows. First, species that show fragmented distributions have higher genetic differentiation and outcrossing rates different from 100%. North American pinyon pines belong to this group of species and require both in situ and ex situ conservation strategies with emphasis in reforestation with young trees. The second strategy could be generated from a phylogenetic perspective, conservation of abundant tree species is as important as conservation of relictual lineages so that speciation and adaptation processes continue. These conclusions should be put in practice in countries like Mexico with high deforestation rates. Finally, it has become clear that knowledge of comparative genetic structure in this group of trees is an excellent tool to develop strategies for their conservation.

ACKNOWLEDGMENTS

We would like to thank Rocío Esteban for her technical assistance and Tom Ledig for making available a copy of his paper on *P. pinceana* before publication. This work was supported by grant 31777-N from Comisión Nacional de Ciencia y Tecnología to D.P. and IN205998 to D.P. and David Gernandt.

LITERATURE CITED


KIMURA M & JF CROW (1964) The number of alleles that can be maintained in a finite population. Genetics 49: 725-738.


OHTA T & M KIMURA (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genetic Research 22: 201-204.


Associate Editor: R. Medel

Received April 30, 2001; accepted January 8, 2002