

Comparative genetic structure in pines: evolutionary and conservation consequences

Estructura genética comparada en pinos: consecuencias evolutivas y para la conservación

PATRICIA DELGADO¹, ARGELIA CUENCA¹, ANA E. ESCALANTE¹, FRANCISCO MOLINA-FREANER²
& DANIEL PIÑERO¹

Departamento de Ecología Evolutiva, ¹Instituto de Ecología, Apartado 70-275, México Distrito Federal 04510, México & ²Estación Regional Noroeste, Departamento de Ecología Funcional y Aplicada, Instituto de Ecología, Apartado Postal 1354, Hermosillo, Sonora 83000, México;
e-mail: pinero@miranda.ecologia.unam.mx

ABSTRACT

Pines have been the focus of several studies that estimate population genetic parameters using both allozymes and chloroplast single sequence repeats (SSRs). Also, the genus has also been recently studied using molecular systematics so that we now have a more clear understanding of their evolutionary history. With this background we studied comparatively the genetic structure in pines. Expected heterozygosity is particularly constant with a 99 % confidence interval between 0.19 and 0.23 in species that have been studied until now using allozymes. There is a significant proportion of species (9/41) that show high population differentiation estimates (F_{ST} = or larger than 0.15) and five of these have large and wingless seeds probably associated with low densities, bird dispersal mechanisms and resistance to water stress. These species include the North American pinyon pines. Outcrossing rates are also constant among species from both subgenus *Pinus* and subgenus *Strobus*, which probably reflects a selective limit to the amount of deleterious alleles that can be maintained in pine species and this also affects inbreeding levels. We also explored the data published using microsatellites in pines and conclude that these markers uncover a higher proportion of variation and genetic differentiation as expected and that the evolutionary models that are used to derive the population genetic structure estimators should take into account other sources of mutation (point mutations, larger insertions and or deletions and duplications) to better understand the comparative applications of these molecular markers.

Key words: *Pinus*, genetic structure, outcrossing rates, microsatellites, allozymes.

RESUMEN

Los pinos han sido el objeto de varios estudios para estimar los parámetros genéticos de la población utilizando tanto aloenzimas como fragmentos repetidos de secuencia sencilla (RSSs) de cloroplasto. Este género también ha sido estudiado recientemente utilizando sistemática molecular de tal manera que ahora tenemos un entendimiento más claro de su historia evolutiva. Con estos antecedentes estudiamos comparativamente la estructura genética de pinos. La heterocigosis esperada es particularmente constante con límites de confianza al 99 % entre 0,19 y 0,23 en las especies que se han estudiado hasta ahora utilizando aloenzimas. Hay una proporción significativa de especies (9/41) que muestran estimados de diferenciación altos (F_{ST} = o mayor que 0,15). De ellos cinco especies tienen semillas grandes y sin alas asociadas probablemente con bajas densidades, dispersión de semillas por aves y resistencia a sequía. Estas especies incluyen a los pinos piñoneros de Norteamérica. Las tasas de entrecruzamiento también son constantes entre especies tanto del subgénero *Pinus* como del subgénero *Strobus* que refleja probablemente un límite selectivo a la cantidad de alelos deletéreos que pueden ser mantenidos en las poblaciones y que afecta también el nivel de consanguinidad. También exploramos los datos publicados usando microsatélites en pinos, concluyendo que estos marcadores muestran una mayor cantidad de variación y diferenciación genética como es esperado y que los modelos de evolución molecular utilizados para derivar los estimadores de la estructura genética de la población deben de tomar en consideración otras fuentes de mutación (mutaciones puntuales, inserciones y deleciones de mayor tamaño y la existencia de duplicaciones) para entender las aplicaciones que desde el punto de vista comparado se pueden hacer con este tipo de marcadores.

Palabras clave: *Pinus*, estructura genética, tasa de entrecruzamiento, microsatélites, aloenzimas.

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 chihuahuana* (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 71° 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 ge-

nus 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-

bility on a particular electrophoretic support, usually starch. 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 $(1000 \times 20) / 4$, 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 mispairing 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 *Ponderosae* subsections for the expected estimates of heterozygosity. No significant correlation was detected ($r = 0.03$, 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 ($r = 0.04$, d.f. = 16, $P > 0.05$) and hard pines ($r = 0.07$, 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

TABLE 1

Average heterozygosities and genetic differentiation estimates (F_{ST}) in pines. Averages are presented for different taxonomic subsections as defined by Price et al. (1998). Standard deviations are shown in parentheses

Heterocigosidades y diferenciación genética (F_{ST}) en pinos. Los promedios se presentan para diferentes subsecciones como fueron definidas por Price et al. (1998). Entre paréntesis se muestran las desviaciones estándar

Taxonomic group	Number of species	Heterozygosity	Genetic differentiation (F_{ST})
Subgenus <i>Pinus</i>	29	0.186 (0.085)	0.127 (0.182)
<i>Strobus</i>	12	0.227 (0.068)	0.136 (0.099)
Subsection <i>Attenuata</i>	3	0.150 (0.009)	0.167 (0.050)
<i>Pinus</i>	7	0.207 (0.132)	0.100 (0.044)
Oocarpae	1	0.213	0.100
Ponderosae	6	0.172 (0.102)	0.116 (0.057)
<i>Australes</i>	7	0.217 (0.053)	0.076 (0.042)
Contortae	4	0.157 (0.047)	0.058 (0.010)
<i>Halepenses</i>	1	0.181	0.010
<i>Strobi</i>	4	0.197 (0.056)	0.150 (0.070)
<i>Cembroides</i>	3	0.219 (0.002)	0.153 (0.112)
Cembrae	4	0.234 (0.091)	0.110 (0.140)
Balfourianae	1	0.340	0.150
Total	41	0.198 (0.082)	0.129 (0.161)

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 ones 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. heidreichii* 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. pinceana* (Escalante 2001), with an average of

0.245. If we do not consider the estimate for *P. pinceana* 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. pinceana*, 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. pinceana*, 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

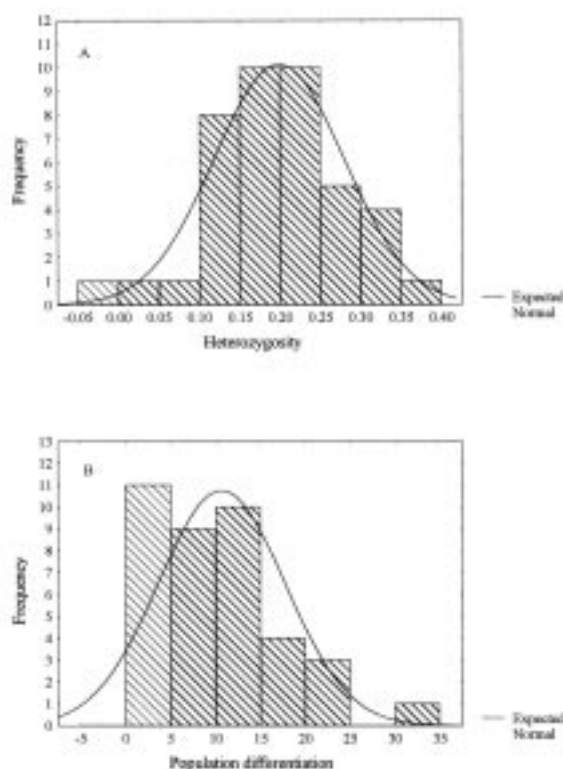


Fig. 1: Frequency distribution of expected heterozygosity (top) and genetic differentiation (bottom) in pine species from published data using allozymes. Differentiation estimates are expressed in percentage.

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.

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. pinceana* (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 % out-

crossing. 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 & Szmidi 1991) and *P. pinceana* (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.

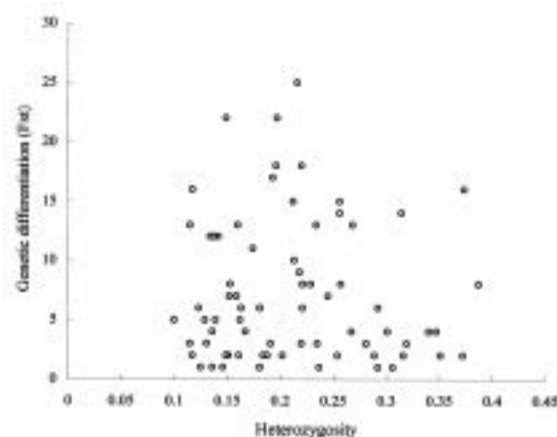


Fig. 2: Relation between expected heterozygosity and genetic differentiation in 71 studies in which both estimates were obtained for 38 pine species using allozyme markers. Circles are data for species in the subgenus *Pinus* (hard pines), squares are data for species in the subgenus *Strobus* (soft 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 cuadrados muestran los datos para las especies del subgénero *Strobus* (pinos blandos).

TABLE 2

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

Estimados de la diversidad y la diferenciación genética usando loci de microsatélites de cloroplasto en especies de pinos (las desviaciones estándar se muestran entre paréntesis)

Taxon	Subsection	Diversity	F _{st}	R _{st}	Number of loci	Reference
<i>P. sylvestris</i>	<i>Pinus</i>	0.978				Provan et al. (1998)
<i>P. pinaster</i>	<i>Pinus</i>	0.885	0.023		6	Ribeiro et al. (2001)
<i>P. resinosa</i>	<i>Pinus</i>	0.618	0.121	0.068		Echt et al. (1998)
<i>P. heidreichii</i> var. <i>leucodermis</i>	<i>Pinus</i>	0.411	0.22			Powell et al. (1995)
<i>P. halepensis</i>	<i>Halepenses</i>	0.222	0.308	0.212	8	Bucci et al. (1998)
<i>P. brutia</i>	<i>Halepenses</i>	0.297	0.289	0.323	8	Bucci et al. (1998)
<i>P. brutia</i> ssp. <i>Eldarica</i>	<i>Halepenses</i>	0.271	0.068	0.074	8	Bucci et al. (1998)
<i>P. nelsonii</i>	<i>Cembroides</i>	0.728	0.131	0.047	4	Cuenca (2001)
<i>P. pinceana</i>	<i>Cembroides</i>	0.824	0.783	0.930	4	Escalante (2001)
Total		0.582 (0.289)	0.245 (0.241)			

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 (Avice 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 3

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

Tamaño promedio para diferentes loci de microsatélite de cloroplasto en ambos subgéneros de *Pinus*. Los loci corresponden a la notación de la secuencia del cloroplasto de *Pinus thunbergii* (Wakasugi et al. 1994). Las desviaciones estándar se muestran entre paréntesis

Locus	Subgenus <i>Pinus</i>	Subgenus <i>Strobus</i>	All species
<i>Pt1254</i>	66.5 (1.22)	67.5 (1.73)	66.9 (1.45)
<i>Pt1520</i>	122.8 (2.48)	122 (3.65)	122.5 (2.84)
<i>Pt9383</i>	88 (3.35)	88.3 (3.59)	88.1 (3.25)
<i>Pt26081</i>	142.3 (1.75)	156.3 (1.26)	147.9 (7.34)
<i>Pt36480</i>	144.7 (2.88)	142.8 (2.06)	143.9 (2.64)
<i>Pt41093</i>	77.3 (0.82)	76.8 (2.63)	77.1 (1.66)
<i>Pt4821</i>	88.5 (0.55)	87.5 (0.58)	88.1 (0.74)
<i>Pt5187</i>	79.2 (5.42)	85.8 (2.06)	81.8 (5.41)
<i>Pt63718</i>	93.2 (0.41)	91 (0.82)	92.3 (1.25)
<i>Pt71936</i>	145.8 (1.60)	142.8 (2.06)	144.6 (2.31)
<i>Pt87268</i>	163.7 (2.73)	167.5 (1.00)	165.2 (2.90)

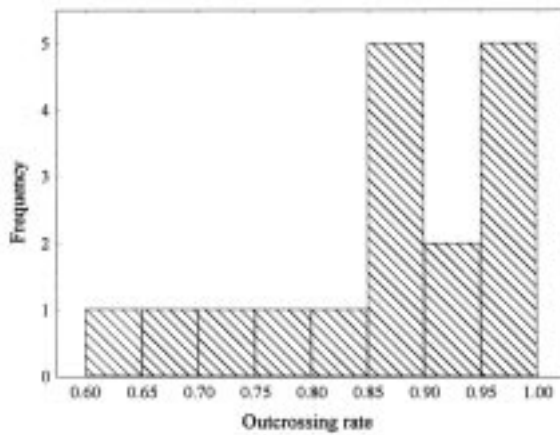


Fig. 3: Frequency distribution of outcrossing rates in pine species from published data using allozymes.

Distribución de frecuencias de las tasas de entrecruzamiento en especies de pinos de datos publicados usando aloenzimas.

maintained above levels at which increases in population sizes would marginally increase the level of genetic variation.

Other aspect 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 be-

low 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 mispairing producing one step mutations at high rates), at the interspecific level, particularly when more diver-

TABLE 4

Inbreeding coefficients (F_{IS}) and outcrossing rate (t_m) derived from inbreeding estimates (f) for pine species

Coefficientes de consanguinidad (F_{IS}) y estimados de consanguinidad (f) derivados de la tasa de entrecruzamiento (t_m) para diferentes especies de pinos

Taxon	Subsection	t_m	f	F_{IS}	Reference
<i>P. contorta</i>	<i>Contortae</i>	0.95	0.03	0	Wheeler & Guries (1982)
<i>P. ponderosa</i>	<i>Ponderosae</i>	0.96	0.02	0.02	Farris & Mitton (1984)
<i>P. sibirica</i>	<i>Cembrae</i>	0.89	0.06	-0.01	Goncharenko et al. (1992)
<i>P. sylvestris</i>	<i>Pinus</i>	0.94	0.03	-0.14	Muona & Szmidi (1991)
<i>P. maximartinezii</i>	<i>Cembroides</i>	0.82	0.10	0.08	Ledig et al. (1999)
<i>P. pincheana</i>	<i>Cembroides</i>	0.86	0.07	0.14	Ledig et al. (2001)

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 mispairing 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. 1999). 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 (Balloux 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. heidreichii* var. *leucodermis* (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 *Picea 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

- AVISE JC (1994) Molecular markers, natural history and evolution. Chapman & Hall, New York, New York. 511 pp.
- ÁLVAREZ-BUYLLA E, A CHAOS, D PIÑERO & A GARAY (1996) Demographic genetics of a pioneer tropical tree species: patch dynamics, seed dispersal and seed banks. *Evolution* 50: 1155-1166.
- BALLOUX F, H BRÜNNER, N LUGON-MOULIN, J HAUSSE & J GOUDET (2000) Microsatellites can be misleading: an empirical and simulation study. *Evolution* 54: 1414-1422.
- BERMEJO-VELÁZQUEZ B (1983) Genetic diversity and the mating system in *Pinus engelmannii* Carr. Ph.D. Thesis (Forestry), University of Wisconsin, Madison, Wisconsin. 91 pp.
- COMPS B, D GÖMÖRY, J LETOUZEY, B THIÉBAUT & RJ PETIT (2001) Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics* 157: 389-397.
- CUENCA A (2001) Variación y estructura genética de una especie de pino endémica de México (*Pinus nelsonii* Shaw.). Bachelor in Science Thesis, Facultad de Ciencias, Universidad Nacional Autónoma de México, Ciudad de México, México. 99 pp.
- DELGADO P, D PIÑERO, A CHAOS, N PÉREZ-NASSER & E ÁLVAREZ-BUYLLA (1999) High population differentiation and genetic variation in the endangered Mexican pine *Pinus rzedowskii* (Pinaceae). *American Journal of Botany* 86: 669-676.
- ECHE CS, LL DEVERNO, M ANZIDEI & GG VENDRAMIN (1998) Chloroplast microsatellites reveal population genetic diversity in red pine, *Pinus resinosa* Ait. *Molecular Ecology* 7: 307-316.
- EGUIARTE LE, A BÚRQUEZ, J RODRÍGUEZ, M MARTÍNEZ-RAMOS, J SARUKHÁN & D PIÑERO (1993) Direct and indirect estimates of neighborhood and effective population size in a tropical palm, *Astrocaryum mexicanum*. *Evolution* 47: 75-87.
- ESCALANTE AE (2001) Estructura genética de poblaciones de *Pinus pinceana* G. Gordon & Glendinning usando como marcadores moleculares microsatélites de cloroplasto (cpSSR's). Bachelor in Sciences Thesis, Facultad de Ciencias, Universidad Nacional Autónoma de México, Ciudad de México, México. 95 pp.
- FARJON A & BT STYLES (1997) *Pinus* (Pinaceae). Flora Neotropica Monograph 75. The New York Botanical Garden, Bronx, New York. 291 pp.

- FARRIS MA & JB MITTON (1984) Population density, outcrossing rate, and heterozygote superiority in ponderosa pine. *Evolution* 38: 1154-1156.
- FURNIER GR & WT ADAMS (1986) Mating system in natural populations of Jeffrey pine. *American Journal of Botany* 73: 1009-1015.
- FURNIER GR & M STINE (1995) Interpopulation differentiation of nuclear and chloroplast loci in white spruce. *Canadian Journal of Forest Research* 25: 736-742.
- GONCHARENKO GG, VEPADUTOV & AE SILIN (1992) Population structure, gene diversity, and differentiation in natural populations of cedar pines (*Pinus subsect. Cembrae*, Pinaceae in the USSR. *Plant Systematics and Evolution* 182: 121-134.
- GOVINDARAJU DR (1988) A note on the relationship between outcrossing rate and gene flow in plants. *Heredity* 61: 401-404.
- HAMRICK JL & MJW GODT (1990) Allozyme diversity in plant species. In: Brown AHD, MT Clegg, AL Kahler & BS Weir (eds) *Plant population genetics, Breeding and Genetic Resources*: 43-63. Sinauer Associates, Sunderland, Massachusetts.
- HEDRICK PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* 53: 313-318.
- HIEBERT RD & JL HAMRICK (1983) Patterns and levels of genetic variation in the Great Basin bristlecone pine, *Pinus longaeva*. *Evolution* 37: 302-310.
- KARHU A, J-H DIETERICH & O SAVOLAINEN (2000) Rapid expansion of microsatellite sequences in pines. *Molecular Biology and Evolution* 17: 259-265.
- KARHU A, P HURME, M KARJALAINEN, P KARVONEN, KARKKAINEN, D NEALE & O SAVOLAINEN (1996) Do molecular markers reflect patterns of differentiation in adaptive traits of conifers. *Theoretical and Applied Genetics* 93: 215-221.
- KIMURA M & JF CROW (1964) The number of alleles that can be maintained in a finite population. *Genetics* 49: 725-738.
- KRUGLYAK S, RT DURRETT, MD SCHUG & CF AQUADRO (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proceedings of the National Academy of Sciences USA* 95: 10774-10778.
- LATTA RG & JB MITTON (1997) A comparison of population differentiation across four classes of gene markers in limber pine (*Pinus flexilis* James). *Genetics* 146: 1153-1163.
- LEDIG FT (1998) Genetic variation in *Pinus*. In: Richardson DM (ed) *Ecology and biogeography of Pinus*: 251-280. Cambridge University Press, Cambridge, United Kingdom.
- LEDIG FT (1999) Founder effects and genetic structure in Coulter pine. *Journal of Heredity* 91: 307-315.
- LEDIG FT & MT CONKLE (1983) Gene diversity and genetic structure in a narrow endemic, Torrey pine (*Pinus torreyana* Parry ex Carr). *Evolution* 37: 79-85.
- LEDIG FT, V JACOB-CERVANTES, PD HODGSKISS & T EGUILUZ-PIEDRA (1997) Recent evolution and divergence among populations of a rare Mexican endemic, Chihuahua spruce, following Holocene climatic warming. *Evolution* 51: 1815-1827.
- LEDIG FT, MT CONKLE, B BERMEJO, T EGUILUZ, P HODGSKISS, DR JOHNSON & WS DVORAK (1999) Evidence for an extreme bottleneck in a rare Mexican pinyon: Genetic diversity, disequilibrium and the mating system in *P. maximartinezii*. *Evolution* 53: 91-99.
- LEDIG FT, M CAPÓ-ARTEAGA, PD HODGSKISS, H SBAY, C FLORES-LÓPEZ, MT CONKLE & B BERMEJO-VELÁZQUEZ (2001) Genic diversity and the mating system of a rare Mexican piñon, *Pinus pinceana*, and a comparison with *Pinus maximartinezii* (Pinaceae). *American Journal of Botany* 88: 1977-1987.
- LI W-H (1997) *Molecular evolution*. Sinauer, Sunderland, Massachusetts. 487 pp.
- LISTON A, WA ROBINSON, D PIÑERO & ER ÁLVAREZ-BUYLLA (1999) Phylogenetics of *Pinus* (Pinaceae) based on nuclear ribosomal DNA internal transcribed spacer region sequences. *Molecular Phylogenetics and Evolution* 11: 95-109.
- LOUKAS M, Y VERGINI & CB KRIMBAS (1983) Isozyme variation and heterozygosity in *Pinus halepensis*. *Biochemical Genetics* 21: 497-509.
- MATHESON AC, JC BELL & RD BARNES (1989) Breeding systems and genetic structure in some Central American pine populations. *Silvae Genetica* 38: 107-113.
- MOLINA-FREANER F, P DELGADO, D PIÑERO, N PÉREZ-NASSER & E ÁLVAREZ-BUYLLA (2001) Do rare pines need different conservation strategies? evidence from three Mexican rare species. *Canadian Journal of Botany* 79: 131-138.
- MUONA O & AE SZMIDT (1991) A multilocus study of natural populations in *Pinus sylvestris*. In: Gregorius HE (ed) *Population genetics in forestry, Lecture Notes in Biomathematics No. 60*: 226-240. Springer-Verlag, Berlin, Germany.
- OHTA T (1995) Synonymous and nonsynonymous substitutions in mammalian genes and the nearly neutral theory. *Journal of Molecular Evolution* 40: 56-63.
- 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.
- PARKER KC & JL HAMRICK (1996) Genetic variation in sand pine (*Pinus clausa*). *Canadian Journal of Forest Research* 26: 244-254.
- POWELL W, M MORGANTE, R MCDEVITT, GG VENDRAMIN & JA RAFALASKI (1995) Polymorphic simple sequence repeat regions in the chloroplast genome: Applications to the population genetics of pines. *Proceedings of the National Academy of Sciences USA* 92: 7759-7763.
- PRICE RA, A LISTON & SH STRAUSS (1998) Phylogeny and systematics of *Pinus*. In: Richardson DM (ed) *Ecology and biogeography of Pinus*: 49-68. Cambridge University Press, Cambridge, United Kingdom.
- PROVAN J, N SORANZO, NJ WILSON, JW McNICOL, GI FORREST, J COTTRELL & W POWELL (1998) Gene pool variation in Caledonian and European Scots pine (*Pinus sylvestris* L.) revealed by chloroplast simple sequence repeats. *Proceedings of the Royal Society of London Series B* 265: 1697-1705.

- PROVAN J, N SORANZO, NJ WILSON, DB GOLDSTEIN & W POWELL (1999) A low mutation rate for chloroplast microsatellites. *Genetics* 153: 943-947.
- RIBEIRO MM, C PLOMION, R PETIT, GG VENDRAMIN & AE SZMIDT (2001) Variation in chloroplast single sequence repeats in Portuguese maritime pine (*Pinus pinaster* Ait.). *Theoretical and Applied Genetics* 102: 97-103.
- SCHEMSKE DW & R LANDE (1985) The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. *Evolution* 39: 41-52.
- SLATKIN M (1995) A measure of population subdivision based on microsatellite allele frequency. *Genetics* 139: 457-462.
- SOKAL RR & FJ ROHLF (1994) *Biometry: the principles and practice of statistics in biological research*. Third edition. W.H. Freeman & Company, San Francisco, California. 887 pp.
- VALDÉS AM, M SLATKIN & NB FREINER (1993) Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics* 133: 737-749.
- WHEELER NC & RP GURIES (1982) Population structure, genic diversity, and morphological variation in *Pinus contorta* Dougl. *Canadian Journal of Forest Research* 12: 595-606.
- WRIGHT S (1949) Adaptation and selection. In: Jepson GL, GG Simpson & E Mayr (eds) *Genetics, paleontology and evolution*: 365-389. Princeton University Press, Princeton, New Jersey.

Associate Editor: R. Medel

Received April 30, 2001; accepted January 8, 2002