# A. Quijada · A. Liston · P. Delgado A. Va´zquez-Lobo · E. R. Alvarez-Buylla

# Variation in the nuclear ribosomal DNA internal transcribed spacer (ITS) region of Pinus rzedowskii revealed by PCR-RFLP

Received: 25 August 1997 / Accepted: 19 September 1997

Abstract In the genus *Pinus* the internal transcribed spacers (ITS1 and ITS2) and the 5.8*s* region of the nuclear ribosomal DNA are approximately 3000 bp in length. ITS1 is considerably longer than ITS2 and partial sequences of ITS1 indicate that this region is evolving rapidly and exhibits intraspecific variation. The ITS2 and 5.8*s* regions are relatively conserved. We surveyed restriction fragment length variability of PCR-amplified fragments (PCR-RFLP) of the ITS region in four populations (86 individuals) of *Pinus* rzedowskii, a pine endemic to western Michoacán, Mexico. Five of the restriction endonucleases assayed revealed variation, with a total of 13 variants, most of which were length mutations of 300*—*900 bp. A moderate degree of population differentiation was detected. The average diversity (Shannon's index) of ITS fragment size patterns was 1.19, with 34% of the variation due to differences among populations and 66% due to differences among individuals within populations. The same individuals were assayed for nine polymorphic isozymes, which gave diversity measures similar to those of each restriction endonuclease.

Key words  $nrDNA \cdot ITS$  region  $\cdot$  Pines  $\cdot$ *Pinus rzedowskii* · Intraspecific variation · PCR-RFLP

Communicated by P. M. A. Tigerstedt

A. Quijada (⊠) · P. Delgado · A. Vázquez-Lobo E. R. Alvarez-Buylla Instituto de Ecología, Universidad Nacional Autónoma de México, México D.F., 04510, Mexico Fax: (525) 616-1976 E-mail: aquijada@miranda.ecologia.unam.mx

A. Liston

Department of Botany and Plant Pathology,

Oregon State University, Corvallis, Oregon, 97331, USA

## Introduction

The nuclear ribosomal DNA (nrDNA) has internal transcribed spacers located between the small subunit (16*s*-18*s*) and 5.8*s* rRNA coding regions (ITS1) and between the 5.8*s* and large subunit (23*s*-28*s*) rRNA coding regions (ITS2). The two spacers and the 5.8*s* subunit are collectively known as the internal transcribed spacer (ITS) region and have become an important nuclear locus for molecular systematic studies of flowering plants (Hamby and Zimmer 1992; Baldwin et al. 1995). The ITS of non-flowering seed plants has been shown to be longer than the typical angiosperm ITS (Bobola et al. 1992; Brown et al. 1993; Karvonen et al. 1993; Qu et al. 1993; Liston et al. 1996) and displays considerable length variation. Restriction-site analyses indicate that the ITS2 plus 5.8*s* subunit ranges from 350*—*425 bp, while ITS1 is responsible for most of the ITS length variation found in non-flowering seed plants (Liston et al. 1996). Partial ITS1 sequences for different species of *Pinus* (Alvarez-Buylla and Liston, unpublished) indicate that this region is evolving rapidly and there are no known restriction sites in this region that are conserved between the two recognized subgenera. This suggests that this region may be variable at the population level, possibly providing a valuable nuclear locus for studies of population variation at the DNA level in pines and permitting a microevolutionary approach to the study of homogenization mechanisms of nrDNA repeats.

Very few studies have evaluated within-population variability of nrDNA in pines and even fewer have compared such variation to that obtained with isozyme loci scored for the same individuals (Quijada 1996). Such comparative surveys are needed to evaluate the usefulness of nrDNA loci as markers for genetic studies of natural populations. As a part of a study on the genetic differentiation of the Mexican pine *P*. *rzedowskii* Madrigal et Caballero, we tested the hypothesis

that intraspecific restriction fragment variation occurs in the ITS region of this species. *P*. *rzedowskii* is endemic to the state of Michoacán, Mexico, and has a very restricted distribution, which is considered endangered by some authors (Perry 1991). We also compared levels and patterns of ITS and isozyme variation.

#### Materials and methods

We collected needles from a total of 86 individuals in four populations of *P*. *rzedowskii* from the following localities: El Pinabete (18*°* 39' 59" N, 103<sup>°</sup> 00' 24" W), Chiqueritos (18<sup>°</sup> 49' 17" N, 102<sup>°</sup> 55' 36" W), El Varaloso (18° 40<sup>'</sup> 39" N, 102° 54' 34" W), and El Fresno (18<sup>°</sup> 40" 04' N, 102<sup>°</sup> 59' 43" W).

Isolation of total genomic DNA followed the protocol of Murray and Thompson (1980) with some modifications. Needle tissue (10 g, fresh weight) was frozen with liquid nitrogen and quickly homogenized into crude powder. The powder was mixed with cold DNA extraction buffer [50 mM Tris, pH 8.0, 5 mM EDTA, 0.35 M Sorbitol, 10% (w/v) polyethylene glycol (Mt 3350), 0.1% 2-mercaptoethanol] and further homogenized with a Polytron homogenizer. After these steps the DNA extraction followed the procedure of Wagner et al. (1987). Each isolation used leaf tissue from a single individual.

Polymerase chain reaction (PCR) amplifications of the ITS region followed the protocol of Liston (1992) with the addition of 1% bovine serum albumin and 5% DMSO to each 100-µl reaction. The primers 26S-25R (5'-TAT GCT TAA ACT CAG CGG GT-3') (Nickrent et al. 1994) and a modification (5'-GGG AGG AGA AGT CGT AAC AAG G-3') of the ITS-5 primer of White et al. (1990) were used with a 55*°*C annealing step. Restriction digests of the ITS region were conducted with 250–500 ng (6–10 µl) of unpurified PCR product and the addition of the enzyme manufacturer's (New England Bioland) recommended buffer. The 16 enzymes assayed were *Aci*I, *Alu*I, *Bfa*I, *BsaJ*I, *Bst*ºI, *Dde*I, *Dpn*II, *Hae*III, *Hha*I, *Hin*fI, *MspI, RsI, Sau96I, ScrFI, TaqI, BsmAI, and FokI. Electrophoresis of* digested samples was conducted with Tris-Borate-EDTA buffer in 2% agarose gels at 100 V. Bands were visualized with ethidium bromide and photographed with a CCD video imager. The size of restricted products was estimated with a 100-bp ladder (Gibco/BRL).

Diversity in the frequency of fragment-size patterns from digests was apportioned within and among *P*. *rzedowskii* populations using Shannon's information-measure H (Hutchenson 1970). Diversity estimates of fragment size patterns were obtained using: (1) frequencies of fragment-size patterns for each enzyme in each population and (2) frequencies of combined fragment-size patterns for all enzymes per individual in each population. H was calculated for two levels of fragment-size pattern diversity.  $H_{pop}$  is the average diversity within populations and  $H_{sp}$  is the diversity within species. The proportion of diversity within populations is  $H_{pop}/H_{sp}$  and the proportion of diversity among populations is  $(H_{sp} - H_{pop})/H_{sp}$ . Further analyses of population structure were based on Weir and Cockerham's estimator of Fst (Alvarez-Buylla et al. 1996), where the value of p is the frequency of variable fragment size pattern per enzyme for each population. We tested the null hypothesis  $F_{ST} = 0$ by calculating  $c^2 = 2NF_{ST} (k-1)$ , where the degrees of freedom are given by  $(k-1)(s-1)$  and where k is number of fragment-size pattern variants per enzyme, s is number of populations, N is total sample size (King and Schaal 1989), and  $c^2$  is distributed as  $\chi^2$ . Gene flow among populations (Nm) was estimated by  $F_{ST} = 1/(4Nm[n/m)]$  $(n-1)$ } + 1, where n is the number of populations and m is the migration rate.

Nine polymorphic isozymes were assayed using the same individuals sampled for ITS variation. Electrophoretic procedures followed standard protocols (Mitton et al. 1979; Conkle et al. 1982; Delgado 1996). The polymorphic isozymes used were anodic peroxidase 1 (APX1), anodic peroxidase 2 (APX2), cathodic peroxidase (CPX), glutamate-oxaloacetate transaminase (GOT), menadione reductase (MNR), leucine aminopeptidase (LAP), esterase (EST), shikimate dehydrogenase (SKDH), and isocitrate dehydrogenase (IDH), Shannon's diversity index (H), Fst, and phenetic relationships were estimated as for the ITS data.

#### Results

Of the 16 assayed restriction enzymes, five (*Hae*III, *Sau*96I, *Alu*I, *Rsa*I, *Dpn*II) revealed variation in the ITS region. Because the PCR primers were in the small subunit and large subunit of the rDNA repeat, they amplified a 3100-bp product, 100 bp longer than the actual ITS (3000 bp). Twelve fragment-size patterns were observed (Table 1), with differences characterized by the presence of one or two extra bands that yielded a total ITS length 300*—*900-bp longer than the normal 3000 bp. Estimated total lengths are in fact somewhat above or below the expected size due to errors in the estimation of fragment lengths. However, it is clear that banding patterns *b*, *c*, *e*, *f*, *h*, *j* and *l* from Table 1 clearly add up to more than 3100 bp. We only scored fragments that were easily visualized. Of 86 individuals analyzed, 71 showed extra less-intensely stained bands in at least one polymorphic enzyme, suggesting the presence of several different ITS repeats within some individuals. Replicate digests were performed for at least one pattern for each enzyme and population, and the extra faint bands appeared consistently.

Combined fragment-size patterns from each enzyme per individual produced 13 genotypes which varied in frequency among populations (Table 2). The *adgik* pattern is the only one for which all enzymes tested yielded fragments that total 3000 bp. This pattern was present in 17.4% of all individuals sampled and occurred primarily in the Chiqueritos population. All other patterns had extra bands for at least one of the enzymes and were found in all populations, including Chiqueritos. The most common patterns were *adhjk* (29%) and *adhik* (27.9%). Other patterns were rare, represented only by one individual in some populations.

Diversity, based on Shannon's index of combined fragment-size patterns averaged for the four populations, was 1.195. Thirty four percent of the diversity was attributable to differences among populations, while the remaining 66% was due to differences among individuals within populations (Table 3). When diversity estimates were obtained for each enzyme separately, average population diversity was 0.318, with 20.3% due to differences among populations and 79.7% due to differences among individuals within populations (Table 4). Absolute estimates of Shannon's index are different because combined fragment-size patterns per individual produced eight rare variants, while the **Table 1** Fragment sizes for digests of the ITS region in  $P$ . rzedowskii according 100-bp<br>ladder marker. Total fragment size pattern:  $a = 3170$ ,  $b = 3470$ ,<br>c = 4100, d = 3190, e = 3640,  $f = 3770 g = 3080, h = 3660,$ <br> $i = 3200, j = 3560, k = 3130,$  $1 = 3690$ 



Table 2 Distribution of the combined fragment-size patterns of the digested ITS region among P. rzedowskii populations



Populations	n	H
Pinabete	14	1.02
Varaloso	29	1.198
Fresno	17	1.325
Chiqueritos	26	1.124
Mean Hpop		1.195
H species	86	1.800
Hpop/Hspecies		0.66
$(Hspecies - Hpop)/Hspecies$		0.33

Table 4 Diversity estimates from fragment-size patterns of the ITS region and allelic isozyme variation between and within populations using Shannon's diversity index (H) in *P*. *rzedowskii* populations



number of rare variants per enzyme is no more than two. Nonetheless, in both diversity estimates, most of variability in ITS results from differences among individuals within a population, rather than among populations.

Isozyme data revealed greater levels of population variation and absolute diversity, but revealed a population genetic structure similar to that observed for ITS (Table 4). Average population diversity was 0.943, with 20.5% due to differences among populations and 79.5% due to differences among individuals within populations.  $F_{ST}$  estimates using both ITS and isozyme data coincide with diversity estimates, indicating moderate but significant differentiation among populations.  $F_{ST}$  for ITS was significantly different from zero (mean = 0.152, SD = 0.075,  $c^2 = 16.25$ ,  $P < 0.05$ ). Isozyme  $F_{ST}$  estimates were similar to those obtained with ITS and also significantly different from zero (mean  $= 0.107$ ; SD  $= 0.104$ ;  $c^2 = 17.05$ ,  $P < 0.05$ ).  $F_{ST}$  estimates obtained with ITS and enzyme data were



Fig. 1A,B Phenograms generated by UPGMA clustering of isozyme variation data (A) and ITS fragment-size patterns (B) of *P*. *rzedowskii* populations

not significantly different ( $t = 0.400$ , P  $> 0.05$ ). Likewise, gene flow (Nm) estimates were similar  $(t = 1.15,$  $P > 0.05$ ), indicating relatively high levels of gene flow both for ITS (mean  $= 2.45$ ; SD  $= 1.57$ ) and isozyme data (mean =  $4.06$ ; SD = 2.86). Clustering patterns obtained by UPGMA analyses (Sneath and Sokal 1973) of isozyme and ITS variation were slightly different (Fig. 1).

## **Discussion**

Our results show that there is considerable intraspecific variation in the ITS region of *P*. *rzedowskii*. Because previous studies (Liston et al. 1996) showed that most length and sequence variation in the nrDNA of the Pinaceae is due to ITS1 variation, while the 5.8*s* and ITS2 are more conserved, we attribute the populationlevel variation observed in *P*. *rzedowskii* to ITS1. Therefore this region may be useful for estimating levels and patterns of intraspecific genetic variation. Additionally, the ITS and isozymes display similar levels and patterns of variation.

Previous documentation of nrDNA variation in the Pinaceae has revealed within-individual and withinpopulation variation (Bobola et al. 1992; Karvonen and Savolainen 1993). The presence of bands of lower intensity in some individuals of *P*. *rzedowskii* suggests intra-individual ITS polymorphism. We also observed intense extra bands that yield total ITS lengths greater than 3000 bp. Similar patterns were obtained for several enzymes by Karvonen and Savolainen (1993). The intensity of these bands suggests that they occur at the same frequency as those bands of the 'standard' repeats. The fact that the fragment size is constant for each particular enzyme used and that such bands were repeatedly obtained in replicate digestions indicates that they are not the result of partial digestions. Different enzyme sensitivity to methylation could also explain these extra bands, but only two (*Sau*96I, *Dpn*II) of the five polymorphic restriction enzymes used are known to be sensitive to methylation.

There are two possible explanations for these intense extra bands. The first possibility is that they correspond to a different repeat variant with an extra restriction site in one of the larger bands of the 'standard' pattern, but that the other complementary band cannot be resolved for one of two reasons. Either it is of the same size as one of the other bands or it is so small that it can not be seen in the gel. Another possiblity is that the extra bands observed are a consequence of interlocus length heterogeneity due to low rates of concerted evolution in *P*. *rzedowskii*. Length variation in other regions of nrDNA has been reported previously. For example, non-transcribed spacers (NTS) exhibit wide length heterogeneity in flowering plants both at the intra-population and intra-individual levels (Appels and Dvorak 1982; Flavell et al. 1986; Schaal et al. 1987; King and Schaal 1989; Capposela et al. 1992). Spacerlength heterogeneity appears in several instances to be due to variability in the number of sub-repeating elements in the NTS, possibly arising from unequal crossing over among arrays of sub-repeats (Appels and Dvorak 1982; Flavell et al. 1986; King and Schaal 1989).

We observed levels of ITS variation considerably higher than those reported for angiosperms (Appels and Dvorak 1982; Soltis and Kuzoff 1993). The rate of homogenization of repeats by concerted evolution, and hence the level of variability, depend on the number of copies, the number of loci and chromosomes on which the gene copies are found, the location of the loci along the chromosomes and the rate of DNA exchange within and between chromosomes (Arheim et al. 1980; Wendel et al. 1995). In pines, and in gymnosperms in general, slower rates of homogenization are expected for several reasons. Concerted evolution forces have long been known to homogenize repeats within individual arrays, but there are few robust demonstrations of the concerted evolution of repeats among different arrays (Wendel et al. 1995). Therefore, based on the few studies available, which show that *Pinus* and *Picea* have 10–12 nucleolus organizer regions (NORs) and a larger number of nrDNA loci than angiosperms, we expect slower rates of homogenization in pines than in angiosperms which have only one or two NORs (Karvonen and Savolainen 1993; Wendel et al. 1995). Nonetheless, further research is needed in this area because a recent study of cotton accessions has inferred bidirectional inter-locus concerted evolution based on ITS sequence data and phylogenetic analysis (Wendel et al. 1995). Finally, the localization of the arrays of tandem repeats along chromosome arms also affects the homogenization rate (Saghai-Maroof et al. 1984).

Although isozyme data revealed greater absolute levels of genetic variation than ITS data, both predict that most of this variation is found within, rather than among, populations. The slight differences observed between UPGMA-generated trees are expected due to differences in the number of loci and characters involved. Similar results have been obtained for other nrDNA studies (Appels and Dvorak 1982; Flavell et al. 1986; King and Schaal 1989; Karvonen and Savolainen 1993). For example, the population structure of *Pinus sylvestris* in Finland was similar when inferred from variation of the entire nrDNA repeat and from isozyme variation (Muona and Harju 1989; Karvonen and Savolainen 1993; Karhu et al. 1996). These, together with the results presented in this paper, suggest that the processes that govern the distribution of ITS and isozyme variation may be similar and that the genetic mechanisms affecting homogenization rates of repeats in the nrDNA multigene family do not bias estimates of genetic differentiation.

Acknowledgments We thank Manuel Zapien and Seferino Moreno for field support and assistance and the people of Dos Aguas and Varaloso for friendly support. Alejandra Covarrubias provided advice and support in her laboratory during some parts of this work. Glenn R. Furnier, Daniel Pińero and Luis Eguiarte made valuable suggestions on the first draft of manuscript. Rosalinda Tapia kindly assisted during the PCR-RFLP analyses. William Robinson, América Castañeda, and Ángeles Cortez kindly helped with various important tasks and provided technical assistance. This study was financed by CONACYT (4715N) and PAPIIT (DGAPA IN-212995, UNAM) grants, and a CONACYT scholarship to A.Q.

## References

- Alvarez-Buylla ER, Chaos A, Piñero D, Garay A (1996) Evolutionary consequences of patch dynamics, seed dispersal, and seed banks: demographic genetics of *Cecropia obtusifolia*, a pioneer tropical tree species. Evolution 50 : 1155*—*1166
- Appels R, Dvorak J (1982) The wheat ribosomal DNA spacer region: its structure and variation in populations and among species. Theor Appl Genet 63 : 337*—*348
- Arnheim N, Krystal M, Schmickel R, Wilson G, Ryder O, Zimmer E (1980) Molecular evidence for genetic exchanges among ribosomal genes on non-homologous chromosomes in man and apes. Proc Natl Acad Sci USA 77 : 7323*—*7327
- Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ (1995) The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann Miss Bot Gard 82 : 247*—*277
- Bobola MS, Smith DE, Klein AS (1992) Five major nuclear ribosomal repeats represent a large and variable fraction of the genomic DNA of *Picea rubens* and *P*. *mariana*. Mol Biol Evol 9 : 125*—*137
- Brown GR, Amarasinghe V, Kiss G, Carlson JE (1993) Preliminary karyotype and chromosomal localization of ribosomal DNA sites in white spruce using flourescence in situ hybridization. Genome 36 : 310*—*316
- Capossela A, Silander JA, Jansen RK, Bergen B, Talbot DR (1992) Nuclear ribosomal DNA variation among ramets and genets of white clover. Evolution 46 : 1240*—*1247
- Conkle MT, Hodskiss PD, Nunnally LB, Hunter SC (1982) Starchgel electrophoresis of conifer seeds: a laboratory manual. Gen Tech Rep PSW-64, Berkeley, California, USA
- Delgado P (1996) Estructura genética y demográfica de una especie del género Pinus (P. rzedowskii Madrigal & Caballero) endémica del Estado de Michoacán, México. MSc thesis. Universidad Nacional Autónoma de México, México
- Flavell RB, O'Dell M, Sharp P, Nevo E, Beiles A (1986) Variation in intergenic spacer of ribosomal DNA of wild wheat, Triticum *dicoccoides* in Israel. Mol Biol Evol 3 : 547*—*558
- Hamby RK, Zimmer EA (1992) Ribosomal RNA as a phylogenetic tool in plant systematics. In: Soltis PS, Soltis DE, Doyle JJ (eds) Molecular systematics of plants. Chapman and Hall, New York pp 50*—*91
- Hutchenson K (1970) A test for comparing diversities based on the Shannon formula. J Theor Biol 29 : 151*—*154
- Karhu A, Hurme P, Karjalainen M, Karvonen P, Karkkainen K, Neale D, Savolainen O (1996) Do molecular markers reflect patterns of differentiation in adaptive traits of conifers? Theor Appl Genet 93 : 215*—*221
- Karvonen P, Savolainen O (1993) Variation and inheritance of ribosomal DNA in *Pinus sylvestris* L. (Scots pine). Heredity 71 : 614*—*622
- Karvonen P, Karjalainen M, Savolainen O (1993) Ribosomal RNA genes in Scots pine (*Pinus sylvestris* L.): chromosomal organization and structure. Genetica 88 : 59*—*68
- King LM, Schaal BA (1989) Ribosomal DNA variation and distribution in *Rudbeckia missouriensis*. Evolution 43 : 1117*—*1119
- Liston A (1992) Variation in the chloroplast genes *rpo* C1 and *rpo* C2 of the genus *Astragalus* (Fabaceae): evidence from restriction-site mapping of a PCR-amplified fragment. Am J Bot 79 : 953*—*961
- Liston A, Robinson WA, Oliphant JM, Alvarez-Buylla ER (1996) Length variation in the nuclear ribosomal DNA internal transcribed spacer of non-flowering seed plants. Systematic Bot 21 : 109*—*120
- Mitton JB, Linhart YB, Sturgeon K, Hamrick JL (1979) Allozyme polymorphisms detected in mature needle tissue of ponderosa pine. J Hered 70 : 86*—*89
- Muona O, Harju A (1989) Effective population sizes, genetic variability and mating system in natural stands and seed orchards of *Pinus sylvestris*. Silvae Genet 38 : 221*—*228
- Murray MG, Thompson WF (1980) Rapid isolation of high-molecular-weight plant DNA. Nucleic Acids Res 8 : 4321*—*4325
- Nickrent DC, Schuette KP, Starr EM (1994) A molecular phylogeny of *Arceuthobium* (Viscaceae) based on nuclear ribosomal DNA internal transcribed spacer sequences. Am J Bot 81 : 1149*—*1160
- Perry JP (1991) The pines of Mexico and Central America. Timber Press, Portland, Oregon
- Qu L-H, Xi B-Q, Shi S-H, Zhang H-D (1993) A characteristic size of the ITS in gymnosperm rDNA. In: XV<sup>th</sup> Int Bot Congr Abstracts. International Union of Biological Sciences, Yokohama, Japan, pp 201
- Quijada A (1996) La región de los espaciadores internos transcritos (ITS) del DNA ribosomal del núcleo (nrDNA): detección de hibridación y variación intraespecifica en pinos. MSc thesis, Universidad Nacional Autónoma de México, México
- Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81 : 8014*—*8018
- Schaal BA, Leverich WJ, Nieto-Sotelo J (1987) Ribosomal DNA variation in the native plant *Phlox divaricata*. Mol Biol Evol 4 : 611*—*621
- Sneath PHA, Sokal FJ (1973) Numerical taxonomy. Freeman, San Francisco
- Soltis PS, Kuzoff RK (1993) ITS sequence variation within and among populations of *Lomation grayi* and *L. laevigatum*. Mol Phylog Evol 2 : 166*—*170
- Wagner DB, Furnier GR, Saghai-Maroof MA, Williams SM, Dancik BP, Allard RW (1987) Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. Proc Natl Acad Sci USA 84 : 2097*—*2100
- Wendel JF, Schnabel A, Seelanen T (1995) Bi-directional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc Natl Acad Sci USA 92 : 280*—*284
- White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M, Gelfand D, Sninsky J, White T (eds) PCR protocols: a guide to methods and applications. Academic Press, San Diego, pp 315*—*322