Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)
Available online at www.inia.es/sjar
doi: http://dx.doi.org/10.5424/sjar/2012101-220-11

Spanish Journal of Agricultural Research 2012 10(1): 160-165
ISSN: 1695-971-X
eISSN: 2171-9292

Short communication: Development of a new polymorphic genetic marker in *Araucaria araucana* (Mol) K. Koch

F. Drake[1], M. A. Martin[2]*, A. Alvarez[2], J. R. Molina[3], J. B. Alvarez[2], M. A. Herrera[3] and L. M. Martin[2]

[1] *Departamento de Manejo de Bosques y Medioambiente, Facultad de Ciencias Forestales, Universidad de Concepción, Casilla 160-C, Concepción. Chile*

[2] *Departamento de Genética, ETSIAM, Edificio Gregor Mendel, Campus de Rabanales, Universidad de Córdoba, 14071 Córdoba. Spain*

[3] *Departamento de Ingeniería Forestal, ETSIAM, Edificio Leonardo Da Vinci, Universidad de Córdoba, 14071 Córdoba. Spain*

Abstract

Seed storage proteins have been used as genetic marker in forest species to evaluate genetic variability, demonstrating its effectiveness both in conifers and broad-leaved. In conifers, megagametophyte storage proteins are particularly useful because of their haploid nature. The aim of this study was to determine whether these proteins could be used as a new marker of genetic diversity in *Araucaria araucana,* one of the oldest conifers of South America and a representative symbol of Chilean forest biodiversity. For this, megagametophytes from two *A. araucana* populations were assessed to identify polymorphic bands and to obtain a preliminary estimation of the genetic diversity. The results revealed that globulin is the best fraction for measuring the variability in the species, due to their high level of variation (20 identified bands, 11 of them polymorphic). Both populations showed high genetic diversity, with more than 92% of the variation within populations. The study highlighted that these proteins can be used to measure the genetic diversity in *A. araucana*, providing good information to ensure the preservation of the species genetic resources.

Additional key words: Araucaria; genetic diversity; globulins; seed storage proteins.

Resumen

Comunicación corta. Desarrollo de un nuevo marcador genético polimórfico en *Araucaria araucana* (Mol) K. Koch

Las proteínas de reserva de la semilla se han usado como marcadores de la diversidad genética en especies forestales y han demostrado su efectividad tanto en coníferas como en frondosas. En coníferas, las proteínas de reserva del megagametofito son particularmente útiles debido a su naturaleza haploide. El objetivo de este estudio fue determinar si estas proteínas podrían emplearse para evaluar la diversidad genética de *Araucaria araucana,* una de las coníferas más antiguas de Sudamérica y un símbolo representativo de la biodiversidad del bosque chileno. Se analizaron megagametofitos procedentes de dos poblaciones de la especie para identificar las bandas polimórficas y tener una estimación preliminar de su diversidad genética. Los resultados revelaron que la fracción de globulinas fue la mejor para medir la variabilidad en la especie debido a su elevado grado de variación (20 bandas identificadas, 11 de ellas polimórficas). Las dos poblaciones mostraron una gran diversidad genética, con más de un 92% de la variación dentro de poblaciones. Este estudio ha puesto de manifiesto que estas proteínas pueden emplearse para evaluar la diversidad genética en *A. araucana* y que pueden proporcionar una información relevante para la preservación de sus recursos genéticos.

Palabras claves adicionales: Araucaria; diversidad genética; globulinas; proteínas de reserva de la semilla.

*Corresponding author: ge2macum@uco.es
Received: 10-05-11. Accepted: 11-01-12

Abbreviations used: AA (homozygotic tree for the absence of a band); CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora); Gst (genetic differentiation among populations); Hs (genetic diversity within populations); Ht (genetic diversity over all populations); PA (heterozygotic tree for the allele responsible of a band); PP (homozygotic tree for the allele responsible of a band).

Seed storage proteins have been used as genetic marker in plants for evaluating the genetic variability of many species. Their main advantages as markers are a high polymorphism level, simple genetic control, environmental independency and their economic, easy and fast analysis (Gepts, 1990). These proteins are present in seeds of all plant species, because they are sources of amino acids for synthesis processes that occur during germination. Moreover, their variations have no repercussion on any physiological function, thus allowing these random changes to be maintained.

The seed storage proteins are classified in four fractions according to their solubility: albumins, globulins, prolamins and glutelins (Osborne, 1924). These proteins have been widely studied in cereals due to its relations to bread-making quality (Payne *et al.*, 1982). These studies have shown high levels of genetic variability, highlighting their possibility in the study of genetic resources (Metakovsky *et al.,* 2000) and varietal identification (Zillman & Bushuk, 1979).

The role of these proteins as genetic marker in forest species has been scarcely explored despite their demonstrated effectiveness in evaluating genetic diversity both in broad-leaved and conifers. In this respect, the protein fraction used as a marker is different depending of the species: albumins are used in *Abies pinsapo* Boiss. (Martín *et al.*, 2010a), globulins in *Pinus pinea* L. (Alvarez *et al.*, 2004) and *Castanea sativa* Mill., (Alvarez *et al.*, 2003) and glutelins in *Quercus ilex* L. (Martín *et al.*, 2009) and *Nothofagus* spp. (Martín *et al.*, 2010b). Furthermore, in conifers this technique is particularly favourable given that seed storage proteins constitute an important component of the megagametophyte, a haploid tissue derived from the female gametophyte that surrounds the embryo in a mature seed. Thus, analysis of the segregation pattern in parental trees may enable the determination of the genetic control and linkage relationships without making crosses (Neale & Adams, 1981; El-Kassaby *et al.*, 1987).

Araucaria araucana (Mol) K. Koch, or monkey-puzzle tree, is one of the oldest conifers of South America and is part of the impressive landscape in native forest ecosystems because of its umbrella-shaped adult crown. Moreover, it is considered a representative symbol of Chilean forest biodiversity given its endemicity and longevity (Hoffmann *et al.*, 2001). The current distribution of the species is fragmented and occurs mainly in the Andean region at the frontier of Chile and Argentina and in two disjunct populations in the Coastal Cordillera (Nahuelbuta Cordillera) of Chile. Thus, the current populations are remnants of a more extensive former distribution that was severely diminished by the intense exploitation of the species in the past, the impact of domestic herbivores, seed harvest and natural perturbations (Veblen, 1982; Burn, 1993). For these reasons, the species is included in Appendix I of the Convention on International Trade of Endangered Species of Wild Flora and Fauna (CITES) and is listed as a vulnerable species by the International Union for Conservation of Nature (IUCN, 1996).

Until now, there is no a complete diagnosis about the situation and distribution of the genetic resources of the species and the available information is contradictory, in some cases. On the one hand, studies using markers such as isozymes (Gallo, 2003; Ruiz *et al.*, 2007) and RAPDs (Bekessy *et al.*, 2002, 2003) have indicated that genetic diversity is high and unaffected by the ecological heterogeneity of the species. Conversely, a recent study based on chloroplast and mitochondrial markers reported the low variation in the species (Marchelli *et al.*, 2010). On the other hand, an expedition carried out in the distribution range of the species revealed high fragmentation in some Araucaria forests, indicating low or nonexistent regeneration (Drake *et al.*, 2009) and corroborating the hypothesis that these Araucaria stands are in decline phase and will be lost in a short period of time (Drake, 2004).

To assess the complex situation of this species, it is necessary to first implement the information on its genetic diversity using all available means. Thus, the aim of this study was to determine whether seed storage proteins could be used as new marker of genetic diversity in the species.

Twenty-one trees from two populations of *A. araucana* from both the Chilean and Argentinean sides of the Andean Cordillera were collected for the study. The Chilean population was situated in the Cunco District (38° 57’ S; 71° 40’ W) and the Argentinean in Lanin National Park (39° 47’ S; 71° 40’ W). This material comes from a collecting mission carried out in 2007 in the distribution range of the species (Drake *et al*., 2009). Eight megagametophytes per tree were used for the analysis, because this is an optimum number of megagametophytes for estimating the homozygotic or heterozygotic nature of each locus in conifers (Morris & Spieth, 1978).

Megagametophytes were isolated by removing both the seed coat and the embryo. The lipids were removed with diethyl ether (50:1 v/w, 1 h, 4°C), followed by a second extraction with acetone (50:1 v/w, 1 h, 4°C).

The sequential extraction process was carried out according to the protocol proposed by Fonseca *et al.* (1997) for *Q. suber* L. Albumin were extracted using 500 L of 10 mM Tris-HCl pH 7.5 + 0.1% (w/v) dithiothreitol + 10 mM $MgCl_2$ + 10 mM $CaCl_2$ and incubated for 2 h at 4°C. Globulins were extracted with 500 L of 10 mM Tris-HCl pH 7.5 + 0.1 % (w/v) dithiothreitol + 10 mM EDTA + 10 mM ethylene glycol bis (ß-aminoethyl ether, EGTA) + 1 M NaCl, incubated for 2 h at 4°C. The samples were centrifuged at 14,000 g for 10 min and the supernatant was transferred to a new tube. For prolamins 75% (v/v) ethanol was added and samples were incubated for 2 h at 4°C. Finally, glutelins were extracted using 500 L of sodium borate buffer pH 10 + 1% (v/v) dithiothreitol + 1% (v/v) SDS, incubated for 2 h at 4°C. Each fraction was precipitated with 1 mL of cold-acetone, and the dried pellets were solubilised in buffer containing 625 mM Tris-HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 0.02% (w/v) bromophenol blue, and 2% (w/v) dithiothreitol in ratio 1:5 (w/v).

Proteins were fractionated in vertical SDS-PAGE slabs in a discontinuous Tris-HCl-SDS buffer system (pH 6.8/8.8) at a 12% polyacrylamide concentration (w/v, C=1.28%). The Tris-HCl/glycine buffer system of Laemmli (1970) was used. Electrophoresis was performed at a constant current of 30 mA/gel at 18°C for 10 min after the tracking dye migrated off the gel. Gels were stained overnight with 12% (w/v) trichloroacetic acid solution containing 5% (v/v) ethanol and 0.05% (w/v) Coomassie Brilliant Blue R-250. Destaining was carried out with tap water.

The Mendelian inheritance of each polymorphic band in the megagametophytes from all polymorphic trees for each band was tested for goodness of fit to the expected 1:1 (presence:absence) segregation using the chi-square test. The eight megagametophytes from each tree were used to determine if the mother tree was homozygotic for the allele responsible for the band (PP), homozygotic for the absence of the band (AA) or heterogygotic (PA). Genetic variability parameters as the percentage of polymorphic bands (*P*), total genetic diversity over all populations (*Ht*), the average genetic diversity within populations (*Hs*) and genetic differentiation between populations (*Gst*) were calculated according to Nei (1973) using Popgene software, version 1.32 (Yeh *et al.*, 1997).

Among the four protein fractions, globulins showed the highest and the clearest resolution between bands with no detected ghost bands and other contaminations. Moreover, the method was highly reliable and repetitive, and these bands were consistently present. These results agree with those reported in the study of other conifers that indicated globulins as the main stored fraction in *P. pinea* (Allona *et al.*, 1994) and as the best fraction for the analysis of the genetic diversity in the species (Alvarez *et al.*, 2004).

Five zones were identified in the gel, classified according to a molecular weight marker used as reference: A (upper 66 kDa), B (66-45 kDa), C (45-30 kDa), D (30-14 kDa) and E (below 14 kDa) (Fig. 1). The identified bands were named with the letter corresponding to the zone where they were located and a number in function of their mobility.

The analysis of the evaluated accessions for globulins showed up to 20 consistently scored bands, 11 of which were polymorphic (55%). The differences between bands were mainly due to mobility although these differences were also due to intensity in some cases. Four, two and five polymorphic bands were identified in zones C, D and E, respectively, whereas bands were tenuous in zones A and B and no polymorphism was detected (Fig. 1). The number of detected bands in the species was low compared with other conifers such as *P. pinea* and *A. pinsapo*; however, the percentage of polymorphic bands was significantly higher (Alvarez *et al.*, 2004; Martín *et al.*, 2010a).

The presence or absence of each polymorphic band in the megagametophytes of all heterozygotic trees fit a 1:1 ratio, with no significant deviation from the expected segregation in any case (results not shown). For each band, the genotype of the studied trees was obtained along with the corresponding frequency (Table 1). All polymorphic bands were present in both populations; the exception was band 5E, which was absent in the Lanin population. In general, all bands displayed a wide distribution among the populations, except for band 3C that showed the lowest frequency (19.0%); however, the frequency of this band was more than double in the Cunco District than in the Lanin National Park (27.3% *vs* 10.0%) (Table 1).

A preliminary evaluation of the genetic diversity was carried out in two populations of the species. The total genetic diversity was considerably high (*Ht* = 0.429), being this value slightly higher in the Cunco District than in the Lanin National Park (0.431 and 0.374, respectively). The differentiation among populations (*Gst*) was 7.1%, indicating that 92.9% of the variation was within populations. When comparing the results obtained using this marker with those achieved with

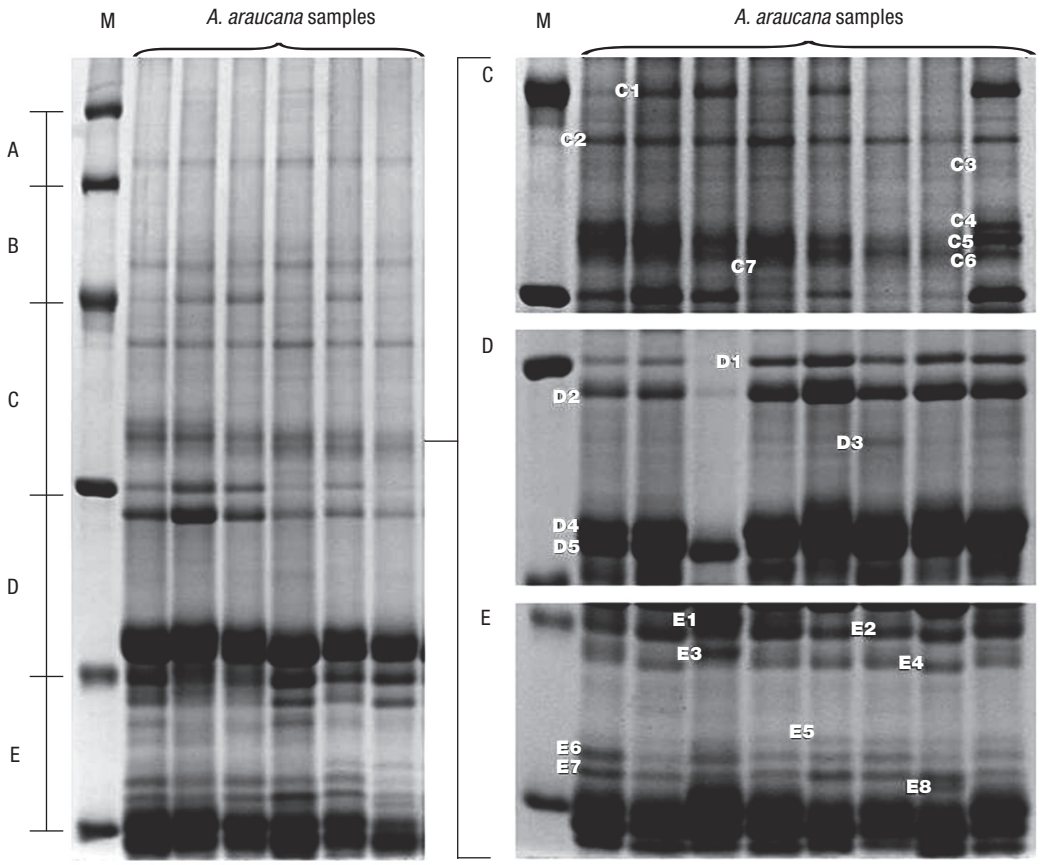

Figure 1. SDS-PAGE of representative samples of the variation found in *Araucaria araucana.* M, molecular marker used as reference. A-E, identified zones in the gel classified according to the molecular marker. C1-E8, the total bands found in the study.

Table 1. Number of trees of each genotype for each band and allele frequency of each band in the two populations evaluated

Band	Cunco District (N = 11)				Lanin National Park (N = 10)				Total
	PP	PA	AA	%	PP	PA	AA	%	%
2C	5	6	0	72.7	2	7	1	55.0	64.3
3C	1	4	6	27.3	0	2	8	10.0	19.0
4C	0	10	1	45.5	0	9	1	45.0	45.2
6C	3	8	0	63.6	0	10	0	50.0	57.1
3D	0	8	3	36.4	1	7	2	45.0	40.5
4D	3	8	0	63.6	1	9	0	55.0	59.5
1E	1	10	0	54.5	4	6	0	70.0	61.9
3E	1	10	0	54.5	5	5	0	75.0	64.3
4E	1	10	0	54.5	5	5	0	75.0	64.3
5E	2	6	3	45.5	0	0	10	0.0	23.8
8E	1	8	2	45.5	1	7	2	45.0	45.2

N, number of trees; P, presence of the band; A, absence of the band. PP, homozygotic tree for the allele responsible for the band; PA, heterozygotic tree for the allele responsible for the band; AA, homozygotic tree for the absence of the band.

isozymes, the genetic diversity was similar to that described by Gallo (2003) for the species in Argentina but significantly higher than that detected by Ruiz *et al.* (2007) for the species in Chile.

In conclusion, the current study highlighted that the globulin fraction is a good marker to measure the genetic diversity in *A. araucana.* Its use may provide an additional tool to complement the information provided by other genetic markers and to set up lines of action to ensure the preservation of the species genetic resources.

Acknowledgments

This research was partially support by grant No. 207-141-018-1.0 from the Research Service of Concepción University (Chile) and by grants A/023099/09 and A/030789/10 of the Spanish Agency of International Cooperation for the Development (AECID) from the Spanish Ministry for Foreign Affairs and Cooperation.

References

Allona I, Collada C, Casado R, Aragoncillo C, 1994. Electrophoretic analysis of seed storage proteins from gymnosperms. Electrophoresis 15: 1062-1067.

Alvarez JB, Muñoz C, Martín MA, López S, Martín LM, 2003. Cotyledon storage proteins as markers of the genetic diversity in *Castanea sativa* Miller. Theor Appl Genet 107: 730-735.

Alvarez JB, Toledo MJ, Abellanas B, Martín LM, 2004. Use of megagametophyte storage proteins as markers of the genetic diversity in stone pine (*Pinus pinea* L) in Andalusia, Spain. Genet Res Crop Evol 51: 621-627.

Bekessy SA, Allnut TR, Premoli AC, Lara A, Ennos RA, Burgman MA, Cortes M, Newton AC, 2002. Genetic variation in the vulnerable and endemic monkey puzzle tree, detected using RAPDs. Heredity 88: 243-249.

Bekessy SA, Ennos RA, Burgman MA, Newton AC, Ades PR, 2003. Neutral DNA markers fail to detect genetic divergence in an ecologically important trait. Biol Conserv 110: 267-275.

Burns BR, 1993. Fire-induced dynamics of *Araucaria araucana-Nothofagus antartica* forest in the southern Andes. J Biogeogr 20: 669-685.

Drake F, 2004. Uso sostenible en bosques de *Araucaria araucana* (Mol.) K. Koch. Aplicación de modelos de gestión. Doctoral thesis, Universidad de Córdoba, Spain. [In Spanish].

Drake F, Martín MA, Herrera MA, Molina JR, Drake-Martin F, Martín LM, 2009. Networking sampling of *Araucaria araucana* (Mol.) K. Koch in Chile and the bordering zone of Argentina: implications for the genetic resources and the sustainable management. iForest 2: 207-212.

El-Kassaby YA, Meagher MD, Parkinson J, Portlock FT, 1987. Allozyme inheritance, heterozygosity and outcrossing rate among *Pinus monticola* near Ladysmith, British Columbia. Heredity 58: 173-181.

Fonseca PA, Ferreira RB, Texeira AR 1997. Seed proteins from *Quercus suber*. J Agr Food Chem 45: 3443-3447.

Gallo L, 2003. Conservación, manejo y uso sustentable de los recursos genéticos de la *Araucaria araucana* en Argentina, Comunidades Aucapan y Chiuquilihuin, San Carlos de Bariloche. Report of the Project "Conservation, Management and Sustainable Use of Forest Genetic Resources with reference to Brazil and Argentina". International Plant Genetic Resources Institute, Rome, Italy. [In Spanish].

Gepts P, 1990. Genetic diversity of seed storage proteins in plants. In: Plant population genetics, breeding and genetic resources (Brown AHD, Clegg MT, Kahler AL, Weir BS eds.). Sinauer Assoc Inc Publ, Suderland, MA, USA. pp: 64-82.

Hoffmann A, Sierra M, Prosser C, Valle M, 2001. Enciclopedia de los bosques chilenos. Gráfica Andes, Santiago de Chile. [In Spanish].

IUCN, 1996. World list of threatened trees. International Union for the Conservation of Nature, Gland, Switzerland. Available in: http://www.iucn.org/about/work/programmes/species/red_list.htm. [11 January 2012].

Laemmli UK, 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

Marchelli P, Baier C, Mengel C, Ziegenhagen B, Gallo LA, 2010. Biogeographic history of the threatened species *Araucaria araucana* (Molina) K. Koch and implications for conservation: a case study with organelle DNA markers. Conserv Genet 11: 951-963.

Martín MA, Navarro-Cerrillo R, Ortega P, Alvarez JB, 2009. The use of cotyledon proteins to assess the genetic diversity in sweet holm oak. J For Sci 55: 526-531.

Martín MA, Alvarez JB, Martín LM, 2010a. Genetic diversity of Spanish fir (*Abies pinsapo* Boiss.) populations by means of megagametophyte storage proteins. Ann For Sci 67: art. 603 (7 p).

Martín MA, Muñoz S, Muñoz F, Uribe M, Molina JR, Herrera MA, Martín LM, Alvarez JB, 2010b. Primeros resultados en el desarrollo de un marcador genético basado en las proteínas de reserva en especies del género *Nothofagus*. Bosque 31, 252-257. [In Spanish].

Metakovsky EV, Gómez M, Vázquez F, Carrillo JM, 2000. High genetic diversity of Spanish common wheats as judged from gliadin alleles. Plant Breeding 119: 37-42.

Morris RW, Spieth PT, 1978. Sampling strategies for using female gametophytes to estimate heterozygosity in conifers. Theor Appl Genet 51: 217-222.

Neale D, Adams WT, 1981. Inheritance of isozyme variants in seed tissues of balsam fir (*Abies balsamea*). Can J Bot 59: 1285-1291.

Nei M, 1973. Analysis of gene diversity in subdivided populations. P Natl Acad Sci USA 70: 3321-3323.

Osborne TB, 1924. The vegetable proteins. Longmans, Green & Co. London.

Payne PI, Holt LM, Worland AJ, Law CN, 1982. Structural and genetical studies on the high-molecular-weight subunits of wheat glutenin. 3. Telocentric mapping of the subunits genes on the long arms of the homoeologous group 1 chromosomes. Theor Appl Genet 63: 129-138.

Ruiz E, Gonzalez F, Torres-Diaz C, Fuentes G, Mardones M, Stuessy T, Rosabelle S, Becerra J, Silva M, 2007. Genetic diversity and differentiation within and among Chilean populations of *Araucaria araucana* (*Araucariaceae*) based on allozyme variability. Taxon 56: 1221-1228.

Veblen TT, 1982. Regeneration patterns in *Araucaria araucana* forest in Chile. J Biogeog 9: 11-28.

Yeh FC, Yang RC, Boyle TBJ, Ye ZH, Mao JX, 1997. Popgene ver 1.32. The user-friendly software for population genetic analysis. Mol Biol Biotechnol Cent, Univ Alberta, Canada.

Zillman RR, Bushuk W, 1979. Wheat cultivar identification by gliadin electropherograms. III. Catalogue of electropherogram formulas of Canadian wheat cultivars. Can J Plant Sci 59: 287-298.