Electronic Journal of Biotechnology ISSN: 0717-3458
http://www.ejbiotechnology.info
DOI: 10.2225/vol16-issue2-fulltext-7 *RESEARCH ARTICLE*

DNA sequence variation of drought-response candidate genes in *Austrocedrus chilensis*

María F. Pomponio[1] · Susana Torales[1] ✉ · Leonardo A. Gallo[2] · Mario J. Pastorino[2,5] · Paula Marchelli[2,5] · María Teresa Cervera[3] · Susana Marcucci Poltri[4]

1 INTA, Instituto de Recursos Biológicos, Castelar, Argentina

2 INTA, Estación Experimental Agropecuaria Bariloche, Bariloche, Argentina

3 INIA, Centro de Investigación Forestal, Madrid, España

4 INTA, Instituto de Biotecnología, Castelar, Argentina

5 Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

✉ Corresponding author: storales@cnia.inta.gov.ar
Received June 5, 2012 / Accepted March 8, 2013
Published online: March 15, 2013

Abstract

Background: *Austrocedrus chilensis* (D. Don) Pic. Ser. et Bizzarri commonly known as Patagonian cypress is a member of the Cupressaceae family, characterized by a high adaptive potential for growing in marginal areas and good timber quality. The species grows over a wide area and under a wide range of rainfall. This study assessed adaptive genetic variation at SNP level in candidate genes involved in response to drought stress.
Results: A total of 18 single nucleotide polymorphisms (SNPs) were found among 1,428 bp. Average nucleotide diversity value (π = 0.00312) was similar to those previously reported in other Cupressaceae. The Fst average among genes and populations was 0.163 and the lowest differentiation was observed in continuous and humid populations. A number of neutrality tests were applied to find evidence of positive selection in our candidate gene set, but only *AcAQP2* gene in Pedregoso and San Ramón populations revealed significant departures from neutrality with positive values suggesting balancing selection.
Conclusions: In this study we report the levels of nucleotide diversity searched in some drought stress candidate genes in *Austrocedrus chilensis* and the selective factors that may be acting on this species.

Keywords: adaptive variation, balancing selection, Patagonian cypress, single nucleotide polymorphisms (SNPs).

INTRODUCTION

Genetic diversity in forest tree species plays an important role in adaptation to changes in environmental conditions (Krutovsky and Neale, 2005).

Recently, the increase in the number of sequences in public databases has allowed gathering information from sequences of functional candidate genes whose activity in other species has been described as key to the characterization of adaptive genetic diversity. In this way it is possible to perform association studies by relating changes at the nucleotide level with phenotypic variation in non-model species, such as some forest tree species (Gailing et al. 2009).

One of the strategies used to analyze the natural variability and to search for genes that are under selection is the study of the variability of molecular markers, such as single nucleotide polymorphisms (SNPs) associated with candidate genes, due to their abundance in the genome and their effectiveness in the association with diseases and adaptive traits (González-Martínez et al. 2006). This approach has been used in forest species such as *Cryptomeria japonica* (Kado et al. 2003; Fujimoto et al. 2010), *Pinus taeda* (González-Martínez et al. 2006; Eckert et al. 2010), *Populus nigra L*. (Chu et al. 2009), and forest Neotropical species such as *Pachira quinata* (Bombacaceae), *Virola sebifera* (Myristicaceae) and *Carapa guianensis* (Meliaceae) (Audigeos et al. 2010).

Austrocedrus chilensis (D. Don) Pic. Ser. et Bizzarri (Patagonian cypress) is a native Cupressaceae of the Andean-Patagonian forest and the only species within the genus *Austrocedrus*. It is characterized by its high adaptive potential, wood quality and durability outdoors, features that make it suitable for carpentry, being widely used in North Patagonia. The distribution of the species is discontinuous in the form of patches on both sides of the Andes. In Argentina, the natural distribution range covers a wide gradient of precipitation in the east-west direction, with mean annual values ranging from 2,500 mm in the west, on the border with Chile, to 400 mm in the east, in the Patagonian steppe (Pastorino and Gallo, 2002).

The types of forest vary along the rainfall gradient, with continuous populations in the humid areas of the west, and small fragmented groups in the arid areas of the steppe (Gallo et al. 2004). Since this species has a great capacity to adapt to the marginal, arid and semi-arid zones of North Patagonia, it represents an ideal species for studies of adaptive variability, constituting also a productive alternative for semi-arid temperate regions.

In this study, we analyzed the patterns of nucleotide diversity of candidate genes in natural populations, to infer the levels of polymorphism and the possible effects of natural selection and/or demographic processes on *Austrocedrus chilensis*.

MATERIALS AND METHODS

Plant material

The plant material was obtained from 37 trees belonging to four natural populations of *Austrocedrus chilensis*, selected on the basis of their contrasting rainfall regimes and their type of distribution: two from humid regions and continuous distribution (Llao Llao and Pedregoso) and two from xeric and isolated regions (Pilcañeu and San Ramón). A total of 7 to 10 individuals were analyzed in each population. Populations, geographic distribution, and the habitat characteristics are shown in Table 1 and Figure 1.

Table 1. Populations of *Austrocedrus chilensis.*

Population	Latitude (S)	Longitude (W)	Precipitation (mm)	Distribution	Habitat
Llao Llao (L)	41º 03’	71º 32’	1500	Continuous	Humid
Pedregoso (G)	40º 37’	71º 35’	2000	Continuous	Humid
San Ramón (S)	41º 03’	71º 05’	600	Isolated	Xeric
Pilcañeu (P)	41º 13’	70º 42’	330	Isolated	Xeric

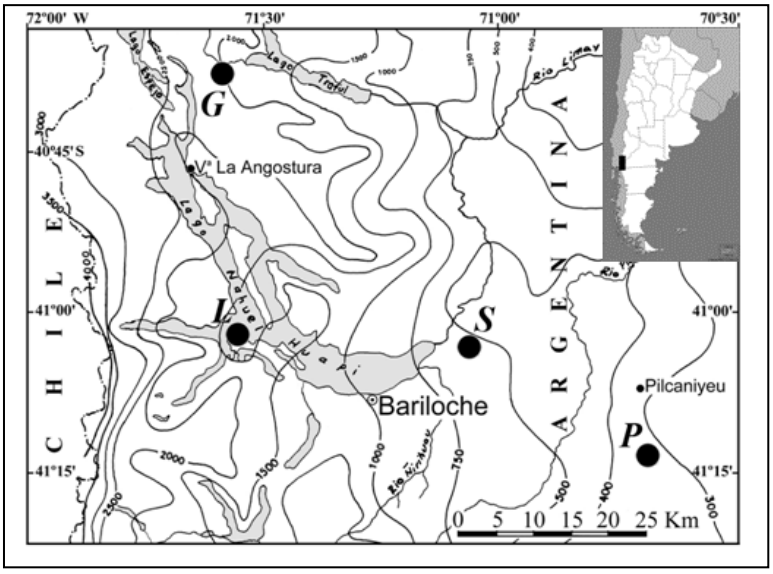

Fig 1. Distribution of *Austrocedrus chilensis* populations analyzed in this study. L: Llao Llao; G: Pedregoso; S: San Ramón; P: Pilcañeu.

Candidate gene selection

A set of 26 candidate genes involved in drought stress response was selected based on the information of genomic sequences and expressed sequence tags (ESTs) in model species such as *Arabidopsis thaliana*, related species such as *Pinus taeda*, *Pinus pinaster* (Pinaceae) and *Cryptomeria japonica* (Cupressaceae), and not related species such as *Prosopis juliflora* (Fabaceae), obtained from the database of the National Center of Biotechnological Information (NCBI) and the Institute for Genomic Research (TIGR). Primers were designed in conserved regions using the software Primer3 (Rozen and Skaletsky, 2000) and PriFi (Fredslund et al. 2005). A group of primers from bibliography was also used (González-Martínez et al. 2006).

DNA extraction, PCR amplification and sequencing

Genomic DNA was extracted from leaves, using the CTAB method (Doyle and Doyle, 1987). The quality and quantity of the DNA extracted was analyzed by spectrophotometry at 260 nm to 280 nm and by agarose gel electrophoresis.

The gene regions were amplified by the PCR technique using a high-fidelity polymerase (Platinum Taq DNA polymerase). PCR-amplified products were checked by agarose gel electrophoresis for presence of one single band before sequencing. Each fragment was sequenced in both directions using a capillary sequencer ABI 3100. The obtained sequences were revised with the Finch TV software version 1.4.0 (Geospiza, Inc.; Seattle, WA, USA; http://www.geospiza.com) to determine their quality, and the base pair discrepance between overlapping sequences. Alignment were performed using ClustalX (Thompson et al. 1994) coupled to the program Bioedit (Hall, 1999).

Validation of the orthologous sequences and haplotype reconstruction

To confirm the identity of the amplified fragments, the sequences were subjected to BLASTN and BLASTX algorithms search against the nucleotide sequences and proteins database (National Center for Biotechnology Information, NCBI), under an E-value threshold of $<10^{-5}$.

Since the DNA samples are diploid, the identification of the haplotypes may result ambiguous, especially in regions where heterozygous individuals are observed (Audigeos et al. 2010). The haplotype reconstruction was performed with the PHASE algorithm (Stephens and Donnelly, 2003) to produce two haploid sequences per individual. DnaSP software v 5.0 with 100 iterations, 1 thinning interval, 100 burn-in and a modeling option with recombination was used (Librado and Rozas, 2009).

The coding and non-coding regions were determined through the alignment of the obtained sequences and the mRNA reference from the NCBI database.

Analysis of nucleotide diversity, population differentiation and neutrality tests

The nucleotide diversity was estimated as θw and π; θw is based on the number of segregating sites (Watterson, 1975), whereas π is based on the average number of nucleotide differences per site between sequences (Nei and Li, 1979). The number of haplotypes (h) and the haplotype diversity (Hd) of Nei (Nei, 1987) were also calculated.

Population differentiation were estimated for each locus and between pairs of populations with the statistic Fst, based on the method of Hudson (Hudson et al. 1992). Although this statistic varies between 0 and 1, slightly negative values may be observed. Since these negative values do not have a biological interpretation, they were considered zero. The significance of Fst was estimated with Arlequin software v 3.1 (Excoffier et al. 2005).

The departures from the neutral model of evolution were estimated using the D tests of Tajima (1989), and F* and D* of Fu and Li (1993). Tajima's D reflects the differences between π and θw. At mutation-drift equilibrium, the expected value of D is close to zero. Negative and significant values of Tajima's test indicate an excess of low-frequency polymorphism, purifying selection or population expansion after a bottleneck. Significant and positive values indicate an excess of intermediate frequency polymorphism caused by balancing selection or population structure. The statistic D* of Fu and Li is based on the differences between the number of singletons (mutation which appears only once among the sequences) and the total number of mutations, whereas F* compares the differences between the number of singletons and the average of nucleotide differences between pairs of sequences (Fu and Li, 1993). A sliding window with a window length of 100 and step size of 25 sites was used to estimate D, F* and D*. All the analyses were carried out using DnaSP v5.0 software (Librado and Rozas, 2009).

Population genetic structure

The genetic structure in the four populations was evaluated with eight nuclear microsatellite markers (SSRs) in 128 individuals (32 in each population) using data from previous studies in the same species (Arana et al. 2010). Multilocus genotypes were used to assess the genetic structure through cluster analysis, applying the Admixture model of the STRUCTURE software (Pritchard et al. 2000). The number of clusters (K) was one to four, with a burn-in period and run length of 10^5. Five repetitions were carried out with each value of K.

The Harvester program (Earl and Von Holdt, 2012) was used to determine the number of optimal clusters. In addition, the statistic R_{ST} was calculated using the Analysis of Molecular Variance (AMOVA). The significance was assessed through 1000 permutations with the GenAlEx program (Peakall and Smouse, 2006).

RESULTS

Amplification of 26 primer pairs designed on sequences of related and unrelated species to *Austrocedrus chilensis* were tested. Successful amplification conditions were found for 15 of them, representing a total 9,355 bp of sequences. Similarity searches were performed on the NCBI database by BLASTX and BLASTN methods. Significant similarity to known genes of stress was found in 3 of them: *pal-1* (phenylalanine ammonia lyase 1), *Aquaporin* and *Hak3P* totaling 1,428 bp. The remaining sequences (7,927 bp) showed no similarity to genes of known function.

Aquaporin gene produced two amplicons (*AcAQP1* and *AcAQP2*), while the sequences of the other genes are from single PCR fragments. In *AcAQP1* and *AcAQP2*, we amplified a small portion of exon 2, intron 2 and a large part of exon 3. The comparison between sequences of both fragments resulted in 43% of similarity, mainly on the exons. The introns presented different lengths and nucleotide composition.

For these reasons, both fragments of the gene *Aquaporin* were considered isoforms (see Discussion) so that the total number of genes analyzed in this work are 4 (Table 2).

Table 2. Candidate genes.

Candidate gene	Putative function	Primers (5'-3')	Source	Total (bp)	Exon (bp)	Intron (bp)
AcAQP1	Aquaporin, intrinsic membrane protein	F:ACGCAGTGCCCGAGACTCTCATG R:AGCAGCTCCAAGCTCCTTGCAGGG	1	316	144	172
AcAQP2				248	135	113
Hak3P	Potassium transporter	F:CCACATGTCAGAGCTGAGGA R:ATGCTGAAGTCCCCAACAAC	2	240	165	75
pal- 1	phenylalanine ammonium lyase	F:TTCCTGTGTTTGAAGCCGAGC R:TGTCTGCCACCCTTACATATTTCTG	3	624	624	-
Total				1428	1068	360

(1): Designed from sequences of *Cryptomeria japonica* from NCBI (This study); (2): Designed from sequences of *Prosopis juliflora* from NCBI; (3): Designed from sequences of *Pinus taeda* (González-Martínez et al. 2006).

Nucleotide diversity

We analysed 74 haploid sequences from three loci (*AcAQP1*, *AcAQP2* and *pal-1*) and 48 haploid sequences from one locus (*Hak3P*) corresponding to 37 and 24 trees respectively. A total of 18 SNPs in 1,428 bp were found; the average SNP frequency was one per 79 bp in all the regions sequenced (Table 3). The average frequency in the non-coding regions was one SNP every 60 bp, while that in the coding regions was one every 89 bp. Read lengths varied from 240 to 624 bp (Table 2).

Table 3. Distribution of SNPs observed in the four candidate genes.

Gene	Length (bp)	total No. of SNPs	Average frequency of SNPs (bp/SNP)	No. of SNPs in coding regions	No. of SNPs in non-coding regions	No. of silent SNPs*	No. of non-synonymous SNPs
AcAQP1	316	1	316	0	1	1	0
AcAQP2	248	7	35.4	2	5	7	0
pal-1	624	7	89.1	7	0	5	2
Hak3p	240	3	80	3	0	1	2
Total	1428	18	79.3	12	6	14	4

*Synonymous and non-coding.

The average nucleotide diversity for the whole set of genes was $\pi t = 0.00312$ and $\theta w = 0.00441$. The value found within the coding region in the synonymous positions (π syn) was 0.00761, while that for the non-synonymous positions (π ns) was 0.00016 (Table 4).

In general, the haplotypic diversity was moderate; the most extreme values were found in Pilcañeu population, with the lowest value (Hd = 0) in the *AcAQP1* gene and the highest value (Hd = 0.846) in the *pal-1* gene (Table 4).

Table 4. Nucleotide diversity.

Gene	P	N	S	H	Hd (DS)	πt	πsyn	πns	πsil	θwtot	θwsyn	θwns	θwsil
AcAQP1	L	20	1	2	0.189 (0.108)	0.00060	0	0	0.00088	0.00089	0	0	0.00132
	G	20	1	2	0.100 (0.088)	0.00032	0	0	0.00046	0.00089	0	0	0.00131
	S	20	1	2	0.189 (0.108)	0.00060	0	0	0.00088	0.00089	0	0	0.00131
	P	14	0	1	0	0	0	0	0	0	0	0	0
	Total	74	1	2	0.128 (0.051)	0.00040	0	0	0.00060	0.0065	0	0	0.00096
AcAQP2	L	20	5	3	0.511 (0.091)	0.00753	0.01840	0	0.1186	0.00568	0.01920	0	0.00896
	G	20	5	3	0.637 (0.064)	0.00881	0.01713	0	0.01534	0.00568	0.00958	0	0.00895
	S	20	3	3	0.574 (0.055)	0.00632	0.01766	0	0.00996	0.00341	0.00957	0	0.00537
	P	14	2	3	0.385 (0.149)	0.00164	0	0	0.00258	0.00254	0	0	0.00398
	Total	74	7	8	0.701 (0.037)	0.00872	0.01810	0	0.01373	0.00579	0.01391	0	0.00912
pal-1	L	20	2	3	0.589 (0.093)	0.00108	0.00424	0	0.00416	0.00090	0.00354	0	0.00348
	G	20	1	2	0.100 (0.088)	0.00016	0.00063	0	0.00062	0.00045	0.00177	0	0.00174
	S	20	5	6	0.621 (0.109)	0.00132	0.00457	0.00022	0.00449	0.00226	0.00709	0.00061	0.00696
	P	14	4	6	0.846 (0.061)	0.00234	0.00729	0.00063	0.00729	0.00221	0.00645	0.00075	0.00645
	Total	74	7	9	0.575 (0.063)	0.00134	0.00468	0.00019	0.00465	0.00253	0.00702	0.00097	0.00697
Hak3P	L	18	1	2	0.209 (0.116)	0.00087	0.00395	0	0.00387	0.00121	0.00549	0	0.00538
	G	16	1	2	0.400 (0.114)	0.00167	0.00755	0	0.00741	0.00126	0.00569	0	0.00558
	S	8	1	2	0.429 (0.169)	0.00179	0.00809	0	0.00794	0.00161	0.00728	0	0.00714
	P	6	3	2	0.333 (0.003)	0.00418	0.00633	0.00368	0.00621	0.00550	0.00832	0.00483	0.00816
	Total	48	3	3	0.414 (0.065)	0.00204	0.00766	0.00046	0.00752	0.00283	0.00428	0.00249	0.0042
Average		270	5	6	0.454	0.003120	0.00761	0.00016	0.00665	0.00441	0.00630	0.00086	0.00531

P: Populations; L: Llao Llao; G: Pedregoso; S: San Ramón; P: Pilcañeu. N: number of sequences per population and total; S: number of segregating sites; H: number of haplotypes; Hd: haplotype diversity; Genetic diversity: πt total diversity for all the sites; πns: for non-synonymous sites; πsyn: for synonymous sites; πsil: for silent sites; θwtot: for all the sites; θwns: for non-synonymous sites; θwsyn: for synonymous sites; owsil: for silent sites

Population differentiation

Fst values between pairs of populations varied between 0.017 and 0.322 among loci with an average value of Fst = 0.163 (Table 5).

Table 5. Estimate of the Fst between pairs of populations for each locus.

Fst	*Hak3P*	AcAQP1	AcAQP2	*pal- 1*	Average
L-P	0.00^{ns}	0.052^{ns}	0.602*	0.212*	0.216
L-S	0.544*	0.00^{ns}	0.222*	0.00^{ns}	0.191
L-G	0.003^{ns}	0.00^{ns}	0.00^{ns}	0.100^{ns}	0.025
G-S	0.330*	0.00^{ns}	0.173*	0.035^{ns}	0.134
G-P	0.00^{ns}	0.000^{ns}	0.533*	0.258*	0.197
P-S	0.280^{ns}	0.052^{ns}	0.404*	0.152*	0.222
Average	0.19	0.017	0.322	0.126	0.163

Statistical significance are given by * P < 0.05; ^{ns} not significant. Populations: L: Llao Llao, G: Pedregoso; S: San Ramón; P: Pilcañeu.

The differentiation between the humid and continuous populations distribution was low and non significant in all loci, whereas genetic differentiation in the xeric and fragmented populations was high, with a substantial variation across the genes (Table 5). Significant Fst values between xeric populations and between humid and xeric populations were found at *AcAQP2*, *Hak3P* and *pal-1* genes.

The analysis with the STRUCTURE software using 8 nuclear SSR revealed that the optimal grouping was found with a K = 2, one group formed by Pilcañeu and the other one by the rest of the populations (Figure 2 and Table 6). In line with these findings, the AMOVA revealed significant differentiation between populations (R_{ST} = 0.087, p = 0.001). However, when the analysis was repeated excluding the population of Pilcañeu, the differences were not significant (R_{ST} = 0.011, p = 0.108).

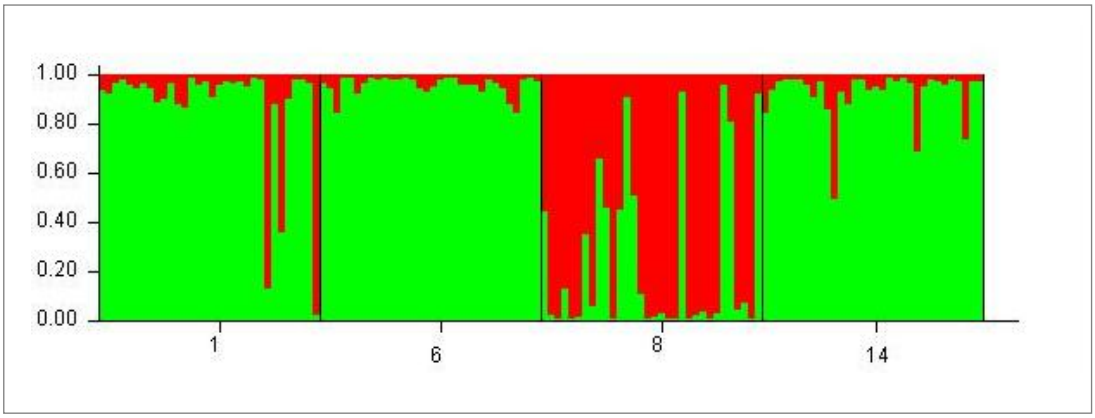

Fig 2. Analysis of population structure. Populations: (1): Llao Llao; (6): San Ramón; (8): Pilcañeu: (14): Pedregoso.

Table 6. Estimated parameters on each repetition of the assumed K.

# K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	5	-44.804.000	18.152	NA	NA	NA
2	6	-42.926.333	14.855	187.766.667	199.513.333	134.308.518
3	5	-43.043.800	137.123	-11.746.667	30.833.333	2.248.590
4	5	-43.469.600	750.862	-42.580.000	15.700.000	0.209093
5	5	-43.738.400	1.476.204	-26.880.000	93.480.000	0.633246
6	5	-44.942.000	738.200	-120.360.000	NA	NA

NA: not available.

Neutrality tests

Neutrality tests were applied to assess whether signatures of positive selection were present in the candidate loci. The results of these tests are shown in Table 7.

Table 7. Neutrality tests.

	AcAQP1			AcAQP2			*pal- 1*			*Hak3p*		
	D	D*	F*	D	D*	F*	D	D*	F*	D	D*	F*
L	-0.591	0.649	0.367	0.996	0.386	0.642	0.457	0.866	0.867	-0.5289	0.6668	0.4045
G	-1.164	-1.539	-1.647	**2.185***	1.186	**1.690***	-1.164	-1.539	-1.647	0.6499	0.6882	0.7720
S	-0.591	0.649	0.367	**2.283***	1.006	1.55	-1.265	-1.212	-1.416	0.3335	0.8877	0.8252
P	0	0	0	-0.959	-0.445	-0.655	0.188	1.164	1.036	-1.0233	-1.2601	-1.3483

D: Tajima test; F* and D*: Fu & Li test. Populations: L: Llao Llao; G: Pedregoso; S: San Ramón; P: Plicañeu. Statistical significance are given by *P < 0.05**.**

*Hak3*P, *pal-1* and *AcAQP1* had non-significant departures from neutral equilibrium expectations. Only one locus, *AcAQP2,* showed positive and significant values in Pedregoso and San Ramón populations, which could be indicative of balancing selection or population structure. A sliding-window analysis of the mutation-drift equilibrium tests at Pedregoso population (Figure 3) shows that most of the significant values were detected in windows centered at nucleotides 76-175 (midpoint 125) and 101-200 (midpoint 150), close to the predicted intron splicing site. In the same way, San Ramón revealed positive and significant values for Tajima's D, also focused on nucleotides 76-175.

DISCUSSION

In the present study, we analyzed patterns of nucleotide diversity in *Austrocedrus chilensis*. The transfer of information from candidate genes from model species to non-model species is a challenge, particularly in conifers, due to the extremely large size of their genomes and to their structural complexity, with a high percentage of repeated sequences, large gene families and many pseudogenes (Ahuja and Neale, 2005; Dolgosheina et al. 2008). In addition, it must be considered that *A. chilensis* is the only species of the genus and there are no studies of genome sequencing.

Although in this study we used highly conserved regions of species belonging to families of conifers, such as the pine (*Pinus taeda, Picea abies*) and cypress (*Cryptomeria japonica*) for the primer design, the percentage of genes transferred to *Austrocedrus chilensis* was low (15%). Limited transfer ability was also observed in other conifers, *e.g.*, a small percentage (0.5%) of candidate genes from *Pinus taeda* were transferred to the species *Sequoia sempervirens* (Ritland, 2012).

The two amplicons observed in *Aquaporin* gene suggest the presence of isoforms as was observed in other plants, like aquaporins in *Arabidopsis thaliana* (Quigley et al. 2002; Javot et al. 2003), *Zea mays* (Maurel et al. 2008) and *Oryza sativa* (Sakurai et al. 2005). These Ac*AQP* fragments could also correspond to pseudogenes. However we did not found evidence, such as stop codons, deletions that disrupt the reading frame of the protein, or lack of splicing sites of exons that could indicate the presence of non-functional alleles.

The average level of nucleotide diversity ($\pi = 0.00312$) were lower than those observed in *Pinus mugo Turra* ($\pi = 0.0081$, Mosca et al. 2012), *Larix decidua* ($\pi = 0.0078$) and *Abies alba* ($\pi = 0.0059$, Mosca et al. 2012), a slightly higher than those found in other Cupressaceae such as *Chamaecyparis pisifera* ($\pi = 0.00289$; Kado et al. 2008), *Cryptomeria japonica* ($\pi = 0.0025$; Kado et al. 2003) *Chamaecyparis obtusa* ($\pi = 0.0024$; Kado et al. 2008), and much higher than those of *P. sylvestris* ($\pi = 0.0014$; Dvornyk et al. 2002).

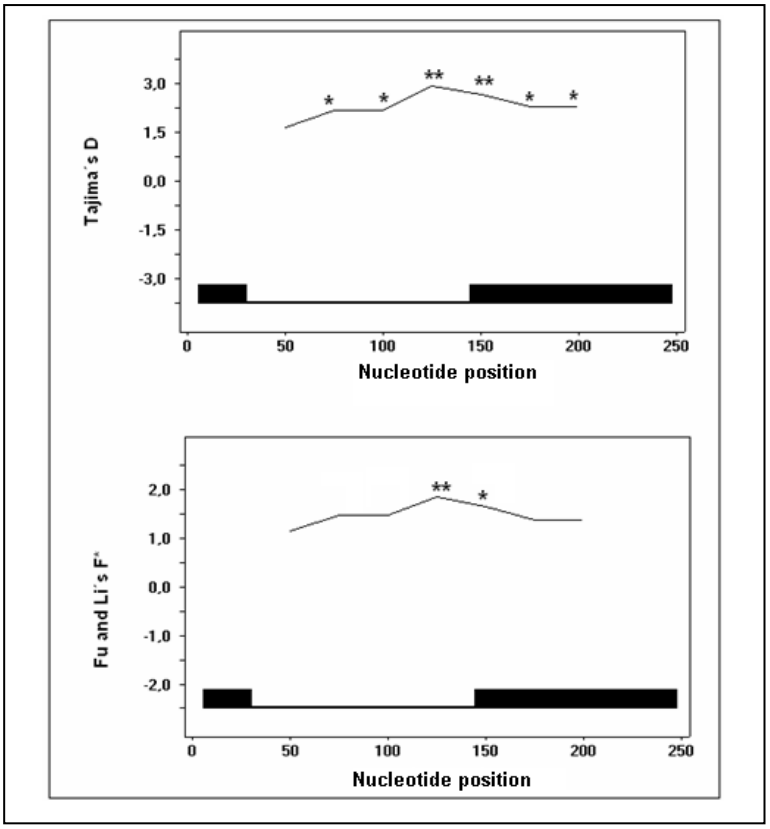

Fig 3. Sliding window for the Tajima's D and Fu and Li's F* test for the AcAQP2 gene in Pedregoso population. The stars indicate statistical significance (*P < 0.10; **P < 0.05). The exons are represented as black boxes and the introns as lines.

The level of genetic differentiation among the populations (Fst = 0.163) was higher compared with that previously reported in this species using isozymes (Fst = 0.066, Pastorino et al. 2004; Fst = 0.12, Souto et al. 2011) and using microsatellite markers (Fst = 0.074, Arana et al. 2010). The differences in the Fst values can be partly attributed to the sampling and to the levels of variation between the different types of markers (Charlesworth, 1998). Discrepancy between the different types of markers has been also reported in *Picea abies*, where the Fst values found using nucleotide sequences (0.117) were much higher than those previously found for isozymes (0.052), AFLP and SSRs (0.02) markers (Heuertz et al. 2006).

The low differentiation in all loci observed in the continuous populations could indicate the same genes were selected in similar environments. In the fragmented populations, we did not find a clear trend in the level of Fst values, given that there was a substantial variation across the genes.The significant differentiation based on neutral markers (SSR) obtained for the Pilcañeu population may be attributed to genetic drift, given the small population (Gallo and Pastorino, 2010).

The pattern of genetic variation observed in three out of four genes analyzed is compatible with the expectations of the neutral model of evolution, while positive and significant departures from the neutral model were found for the *AcAQP2* gene in San Ramón and Pedregoso populations, which suggests either balancing selection or a population structure.

Deducing the type of evolutionary forces acting on populations based only on the departures from the mutation-drift balance is difficult, given that the changes in population sizes and other demographic events may show patterns of genetic diversity similar to those of natural selection (Tajima, 1989; Depaulis et al. 2003). The analyses carried out with neutral markers in the species do not support the presence of demographic process such as population structure, thus we can tentatively conclude that the departures from equilibrium at *AcAQP2* loci are due to balancing selection. In San Ramón and

Pedregoso populations, the observed departures from neutrality seem to be mostly situated in the exon-intron limit, suggesting selection in the intron splicing sites or perhaps, other regulatory functions by intronic sequences, such as those previously observed in the forest tree species *Eperua falcata* in *Aquaporin genes* (Audigeos et al. 2010).

CONCLUDING REMARKS

This exploratory study identified genes potentially involved in the adaptive response to drought stress; nucleotide diversity was analyzed and neutrality tests were applied to search for evidence of adaptive selection in *Austrocedrus chilensis*.

The main limitations of this study include: the low number of samples analyzed, the lack of genome resources in the species, and the low transferability, all of which significantly reduced the number of genes analyzed.

Although most of the genes studied did not present departures from the neutral model of evolution, positive and significant departures were found for the *AcAQP2* gene. Because the evolutionary patterns vary between different regions of the gene, the complete sequencing of the genes, including the region of the promoter, would improve the prediction of neutrality.

Despite these limitations, the data analyzed allowed us to obtain a trend regarding the levels of nucleotide diversity and selective factors that may act on genes. $_{C}$DNA gene libraries are currently being developed in *A. chilensis*, which will enable researchers to obtain molecular markers of great importance in genetic studies of populations and association with adaptive characteristics.

ACKNOWLEDGMENTS

We would like to thank to Lic. Marcos Rodríguez for providing us designed primer sequences from *Prosopis juliflora.*

Financial support: This research was supported by the INTA-PE 242421 and PPR 242001.

REFERENCES

AHUJA, M.R. and NEALE, D.B. (2005). Evolution of genome size in conifers. *Silvae Genetica*, vol. 54, no. 3, p. 126-137.

ARANA, M.V.; GALLO, L.A.; VENDRAMIN, G.G.; PASTORINO, M.J.; SEBASTIANI, F. and MARCHELLI, P. (2010). High genetic variation in marginal fragmented populations at extreme climatic conditions of the Patagonian Cypress *Austrocedrus chilensis. Molecular Phylogenetics and Evolution*, vol. 54, no. 3, p. 941-949. **[CrossRef]**

AUDIGEOS, D.; BUONAMICI, A.; BELKADI, L.; RYMER, P.; BOSHIER, D.; SCOTTI-SAINTAGNE, C.; VENDRAMIN, G.G. and SCOTTI, I. (2010). Aquaporins in the wild: Natural genetic diversity and selective pressure in the *PIP* gene family in five Neotropical tree species. *BMC Evolutionary Biology*, vol. 10, no. 202. **[CrossRef]**

CHARLESWORTH, B. (1998). Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution*, vol. 15, no. 5, p. 538-543. **[CrossRef]**

CHU, Y.; SU, X.; HUANG, Q. and ZHANG, X. (2009). Patterns of DNA sequence variation at candidate gene loci in black poplar (*Populus nigra L*.) as revealed by single nucleotide polymorphisms. *Genetica*, vol. 137, no. 2, p. 141-150. **[CrossRef]**

DEPAULIS, F.; MOUSSET, S. and VEUILLE, M. (2003). Power of neutrality test to detect bottlenecks and hitchhiking. *Journal of Molecular Evolution*, vol. 57, no. 1, p. S190-S200. **[CrossRef]**

DOLGOSHEINA, E.V.; MORIN, R.D.; AKSAY, G.; SAHINALP, S.C.; MAGRINI, V.; MARDIS, E.R.; MATTSSON, J. and UNRAU, P.J. (2008). Conifers have a unique small RNA silencing signature. *RNA Society*, vol. 14, no. 8, p. 1508-1515. **[CrossRef]**

DOYLE, J.A. and DOYLE, J.L. (1987). A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochemistry Bulletin,* vol. 19, no. 1, p. 11-15.

DVORNYK, V.; SIRVIO, A.; MIKKONEN, M. and SAVOLAINEN, O. (2002). Low nucleotide diversity at the *pal1* locus in the widely distributed *Pinus sylvestris*. *Molecular Biology and Evolution*, vol. 19, no. 2, p. 179-188. **[CrossRef]**

EARL, D.A. and von HOLDT, B.M. (2012). Structure Harvester: A website and program for visualizing structure output and implementing the Evanno method. *Conservation Genetics Resources*, vol. 4, no. 2, p. 359-361. **[CrossRef]**

ECKERT, A.; HEERWAADEN, J.V.; WEGRZYN, J.L.; NELSON, C.D.; ROSS-IBARRA, J.; GONZÁLEZ-MARTÍNEZ, S.C. and NEALE, D.B. (2010). Patterns of population structure and environmental associations to aridity across the range of Loblolly pine (*Pinus taeda L*., Pinaceae). *Genetics,* vol. 185, no. 3, p. 969-982. **[CrossRef]**

EXCOFFIER, L.; LAVAL, G. and SCHENEIDER, S. (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* no. 1, p. 47-50.

FREDSLUND, J.; SCHAUSER, L.; MADSEN, L.H.; SANDAL, N. and STOUGAARD, J. (2005). PriFi: Using a multiple alignment of related sequences to find primers for amplification of homologs. *Nucleic Acids Research,* vol. 33, p. 516-520. **[CrossRef]**

FU, Y.X. and LI, W.H. (1993). Statistical tests of neutrality of mutations. *Genetics*, vol. 133, no. 3, p. 693-709.

FUJIMOTO, A.; KADO, T.; YOSHIMARU, H.; TSUMURA, Y. and TACHIDA, H. (2010). Adaptative and slightly deleterious evolution in a conifer, *Cryptomeria japonica*. *Journal of Molecular Evolution,* vol. 67, no. 2, p. 201-210. **[CrossRef]**

GAILING, O.; VORNAM, B.; LEINEMANN, L. and FINKELDEY, R. (2009). Genetic and genomic approaches to assess adaptive genetic variation in plants: Forest trees as a model. *Physiologia Plantarum*, vol. 137, no. 4, p. 509-519. **[CrossRef]**

GALLO, L.; PASTORINO, M.J. and DONOSO, C. (2004). Variación en *Austrocedrus chilensis* (D. Don) Pic. Ser et Bizzarri (Ciprés de la Cordillera). In: DONOSO, C.; PREMOLI, A.; GALLO L. and IPINZA, R. eds. Variación intraespecífica en las especies arbóreas de los bosques templados de Chile y Argentina. Chile, Editorial Universitaria, Chapter 10, p. 233-252.

GALLO, L. and PASTORINO, M. (2010). Evidence of genetic drift in neutral and adaptive genome. In: Evoltree Conference "Forest Ecosystem Genomics and Adaptation", San lorenzo del Escorial, Madrid, España INIA.

GONZÁLEZ-MARTÍNEZ, S.C.; ERSOZ, E.; BROWN, G.R.; WHEELER, N.C. and NEALE, D.B. (2006). DNA Sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L. *Genetics*, vol. 172, no. 3, p. 1915-1926. **[CrossRef]**

HALL, T.A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, vol. 41, p. 95-98.

HEUERTZ, M.; DE PAOLI, E.; KÄLLMAN, T.; LARSSON, H.; JURMAN, I.; MORGANTE, M.; LASCOUX, M. and GYLLENSTRAND, N. (2006). Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of norway spruce [*Picea abies* (L.) Karst]. *Genetics*, vol. 174, no. 4, p. 2095-2105. **[CrossRef]**

HUDSON, R.R.; SLATKIN, M. and MADDISON, W.P. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics*, vol. 132, no. 2, p. 583-589.

JAVOT, H.; LAUVERGEAT, V.; SANTONI, V.; MARTIN-LAURENT, F.; GÜÇLÜ, J.; VINH, J.; HEYES, J.; FRANCK, K.I.; SCHÄFFNER, A.R.; BOUCHEZ, D. and MAUREL, C. (2003). Role of a single aquaporin isoform in root water uptake. *The Plant Cell*, vol. 15, no. 2, p. 509-522. **[CrossRef]**

KADO, T.; YOSHIMARU, H.; TSUMURA, Y. and TACHIDA, H. (2003). DNA variation in a conifer, *Cryptomeria japonica* (Cupressaceae sensu lato). *Genetics*, vol. 164, no. 4, p. 1547-1559.

KADO, T.; MATSUMOTO, A.; UJINO-IHARA, T. and TSUMURA, Y. (2008). Amounts and patterns of nucleotide variation within and between two Japanese conifers, sugi (*Cryptomeria japonica*) and hinoki (*Chamaecyparis obtusa*) (Cupressaceae *sensu lato*). *Tree Genetics & Genomes*, vol. 4, no. 1, p. 133-141. **[CrossRef]**

KRUTOVSKY, K.V. and NEALE, D.B. (2005). Forest genomics and new molecular genetic approaches to measuring and conserving adaptive genetic diversity in forest trees. In: GEBUREK, T. and TUROK, J. eds. *Conservation and Management of Forest Genetic Resources in Europe*. Arbora Publishers, Zvolen, p. 369-390.

LIBRADO, P. and ROZAS, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, vol. 25, no. 11, p. 1451-1452. **[CrossRef]**

MAUREL, C.; VERDOUCQ, L.; LUU, D.T. and SANTONI, V. (2008). Plant aquaporins: Membrane channels with multiple integrated functions. *Annual Reviews of Plant Biology*, vol. 59, no. 1, p. 595-624. **[CrossRef]**

MOSCA, E.; ECKERT, A.J.; LIECHTY, J.D.; WEGRZYN, J.L.; LA PORTA, N.; VENDRAMIN, G.G. and NEALE, D.B. (2012). Contrasting patterns of nucleotide diversity for four conifers of Alpine European forests. *Evolutionary Applications*, vol. 5, no. 7, p. 762-775. **[CrossRef]**

NEI, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York. 512 p. ISBN 9780231063210.

NEI, M. and LI, W.H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 76, no. 10, p. 5269-5273.

PASTORINO, M.J. and GALLO, L.A. (2002). Quaternary evolutionary history of *Austrocedrus chilensis*, a cypress native to the Andean-Patagonian forest. *Journal of Biogeography*, vol. 29, no. 9, p. 1167-1178. **[CrossRef]**

PASTORINO, M.J.; GALLO, L.A. and HATTEMER, H.H. (2004). Genetic variation in natural populations of *Austrocedrus chilensis*, a cypress of the Andean-Patagonian forest. *Biochemical Systematics and Ecology*, vol. 32, no. 11, p. 993-1008. **[CrossRef]**

PEAKALL, R. and SMOUSE, P.E. (2006). Genalex 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, vol. 6, no. 1, p. 288-295. **[CrossRef]**

PRITCHARD, J.; STEPHENS, M. and DONNELLY, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, vol. 155, no. 2, p. 945-959.

QUIGLEY, F.; ROSENBER, G.J.M.; SHACHAR-HILL, Y. and BOHNERT, H.J. (2002). From genome to function: The *Arabidopsis* aquaporins. *Genome Biology*, vol. 3, no. research0001-research0001.17. **[CrossRef]**

RITLAND, K. (2012). Genomics of a phylum distant from flowering plants: Conifers. *Tree Genetics & Genomes*, vol. 8, no. 3, p. 573-582. **[CrossRef]**

ROZEN, S. and SKALETSKY, H.J. (2000). Primer3 on the WWW for general users and for biologist programmers. In: KRAWETZ, S. and MISENER, S. eds. *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, p. 365-386.

SAKURAI, J.; ISHIKAWA, F.; YAMAGUCHI, T.; UEMURA, M. and MAESHIMA, M. (2005). Identification of 33 rice *aquaporin* genes and analysis of their expression and function. *Plant and Cell Physiology*, vol. 46, no. 9, p. 1568-1577. **[CrossRef]**

SOUTO, C.P.; HEINEMANN, K.; KITZBERGER, T.; NEWTON, A.C. and PREMOLI, A.C. (2011). Genetic diversity and structure in *Austrocedrus chilensis* populations: Implications for dryland forest restoration. *Restoration Ecology,* vol. 20, no. 5, p. 568-575. **[CrossRef]**

STEPHENS, M. and DONNELLY, P. (2003). A comparison of bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, vol. 73, no. 5, p. 1162-1169. **[CrossRef]**

TAJIMA, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, vol. 123, no. 3, p. 585-595.

THOMPSON, J.D.; HIGGINS, D.G. and GIBSON, T.J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, vol. 22, no. 22, p. 4673-4680. **[CrossRef]**

WATTERSON, G.A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, vol. 7, no. 2, p. 256-276. **[CrossRef]**

How to reference this article:

POMPONIO, M.F.; TORALES, S.; GALLO, L.A.; PASTORINO, M.J.; MARCHELLI, P.; CERVERA, M.T. and MARCUCCI POLTRI, S. (2013). DNA sequence variation of drought-response candidate genes in *Austrocedrus chilensis. Electronic Journal of Biotechnology*, vol. 16, no. 2. http://dx.doi.org/10.2225/vol16-issue2-fulltext-7