Searching for molecular insight on hybridization in *Nothofagus* spp. forests at Lagunas de Epulauquen, Argentina

Estudio de la hibridación a nivel molecular en bosques de *Nothofagus* spp. en Lagunas de Epulauquen, Argentina

María M Azpilicueta **, Verónica A El Mujtar *, Leonardo A Gallo *

*Corresponding author: ^a Instituto Nacional de Tecnología Agropecuaria, Unidad de Genética Ecológica y Mejoramiento Forestal, Modesta Victoria 4450 CC 277, 8400 Bariloche, Río Negro, Argentina, phone 00 54 294 4422731, azpilicueta.maria@inta.gob.ar

SUMMARY

Lagunas de Epulauquen forests (36°49'S; 71°04'W) have been considered the northernmost population of *Nothofagus obliqua* in Argentina. Recently, however, its taxonomic status has been questioned due to its morphological, architectural and genetic distinctiveness. The convergence of migratory routes and hybridization with *Nothofagus alpina* are the two main hypotheses proposed to explain the distinctiveness of this population. Despite its unique characteristics, no evaluation of its putative origin has been carried out. In order to test the hypothesis of hybridization we analyzed nuclear *Adh* and *ITS* fragment sequences together with 12 nuclear microsatellites. These nuclear regions were selected based on their previously reported divergence between *N. obliqua* and *N. alpina*. The survey was conducted in 12 individuals from Lagunas de Epulauquen, and individuals from *N. obliqua* (8), *N. alpina* (1), *N. macrocarpa* (2), *N. glauca* (4) and *N. leonii* (2). By combining a phylogenetic analysis with a principal coordinate analysis of microsatellite data, we showed that Lagunas de Epulauquen individuals were closely related to other *N. obliqua* individuals, while hybridization with *N. alpina* was not detected. Our results indicate that hybridization processes could be discarded as the origin of the Lagunas de Epulauquen distinctiveness. The hypothesis of the convergence of two *N. obliqua* migratory routes at this latitude was reinforced based on detected patterns of genetic diversity. Notwithstanding, its high geographical isolation seems to be also a key evolutionary factor to explain the distinctiveness of this population. Lagunas de Epulauquen population could be therefore included within Argentinean *N. obliqua* domestication and conservation programs.

Key words: molecular systematics, southern beeches, Patagonia, SSR markers, hybridization.

RESUMEN

Los bosques en Lagunas de Epulauquen (36°49'S; 71°04'O) han sido considerados como la población más septentrional de *Nothofagus obliqua* en Argentina. En años recientes se ha cuestionado su verdadero estatus. La convergencia de rutas migratorias y la hibridación histórica con *Nothofagus alpina* fueron las hipótesis planteadas como posible origen de esta distinción. A pesar de sus características particulares, no fue realizado hasta el presente un estudio sobre su origen putativo. Con el objetivo de profundizar en la hipótesis de hibridación se analizó la secuencia del gen Adh y la región ITS, junto a 12 microsatélites nucleares que habían mostrado divergencia entre *N. obliqua* y *N. alpina* en estudios anteriores. Se analizaron 12 individuos de Lagunas de Epulauquen, junto a individuos de *N. obliqua* (8), *N. alpina* (1), *N. macrocarpa* (2), *N. glauca* (4) y *N. leonii* (2). La combinación de estudios filogenéticos y el análisis de coordenadas principales de microsatélites mostró que los individuos de Lagunas de Epulauquen se ubican cercanos a individuos de *N. obliqua* de otros orígenes, sin detectarse procesos de hibridación con *N. alpina*. Nuestros resultados sugieren descartar los procesos de hibridación como origen de esta distinción. Se refuerza la hipótesis previa de convergencia de dos rutas migratorias con base en los patrones de diversidad genética detectados, aunque no se debe descartar el efecto de procesos actuales de aislamiento geográfico como un factor evolutivo clave. Lagunas de Epulauquen debería incluirse en los programas de domesticación y conservación de *N. obliqua* en Argentina.

Palabras clave: sistemática molecular, Nothofagus, Patagonia, microsatélites, hibridación.

INTRODUCTION

The Nothofagaceae family has four recognized subgenera: *Fuscospora*, *Lophozonia*, *Nothofagus* and *Brassospora* (Manos 1997)¹. Based on phylogenetic analyses of pollen-type, flower and fruit characteristics, and the internal transcribed spacer region (*ITS*), the *Lophozonia* subgenus has been established to include three American

¹ Although a recently revised taxonomy defined these groups at the

primary rank of genus (Heenan and Smissen 2013), we maintained the previous taxonomic definitions since some of the taxa analysed in this study were not included in the last revision and considering the recommendation of Hill *et al.* (2015).

species, *Nothofagus obliqua* (Mirb.) Oerst., *Nothofagus alpina* (Phil.) Dimitri *et* Milano (= *N. nervosa*) and *Nothofagus glauca* (Phil.) Krasser (Manos 1997).

The wide latitudinal range of N. obliqua forests occurring from 33° to 41° S (Veblen et al. 1996) promotes a pronounced clinal variation in their morphological characters along its extensive distribution area in Chile (Donoso et al. 2004); covering an area of more than 1 million hectares. Three subspecies have been described within N. obliqua (ssp. obliqua, ssp. valdiviana, ssp. andina); and a new species, Nothofagus macrocarpa (A. DC.) F.M. Vázquez et R. Rodr., has been reported (Donoso et al. 2004). Notwithstanding, authors such as Vergara et al. (2014) still consider N. macrocarpa as a subspecies of N. obliqua capturing the clinal latitudinal variation described earlier by Donoso et al. 2004. Moreover, interspecific natural hybridization processes were reported between N. obliqua and N. alpina (e.g. Gallo et al. 1997, Marchelli and Gallo 2000) and N. obliqua and N. glauca (this hybrid was named N. leonii Espinosa, Donoso and Landrum 1979) based on morphological data. The existence of these entities co-occurring and hybridizing introduces extra complexity to the definition of their taxonomic status. Vergara et al. (2014) postulated hybridization processes with both N. alpina and N. glauca for the northern Chilean N. obliqua populations in order to explain the admixture found while analyzing molecular data of these three American taxa.

In Argentina, N. obliqua grows as pronounced naturally fragmented forests restricted to the West-East oriented glacial watersheds, over a latitudinal range from 36°50' S to 40°10' S (Sabatier et al. 2011) (figure 1A). The species is distributed in ca 33,800 hectares of forests along a heterogeneous environmental landscape (Sabatier et al. 2011). Notwithstanding, no identification at sub-species level has yet been done, though molecular data showed that Argentinean populations have a pronounced latitudinal distinction, with northern populations exhibiting high levels of genetic diversity (Azpilicueta et al. 2013) together with shared genetic variants with populations growing at the same latitudes at the West of Andes Mountains in Chile (Azpilicueta et al. 2009) (figure 1). Previous studies based on chloroplast DNA markers also revealed a genetic pattern where Chilean coastal N. obligua populations and current forests at North of the Andes and Rucañancu could be inferred as glacial refuges for the species (Azpilicueta et al. 2009), a hypothesis also supported by palynological data (Villagran 1991).

In this work, we focus on Lagunas de Epulauquen (LE) population, located at 36° 49' S - 71° 04' W. It has been considered the northern extreme of *N. obliqua* natural distribution in Argentina (Donoso *et al.* 2004), although much evidence discovered in the last decade has contributed to casting some doubt on its taxonomic identification. Lagunas de Epulauquen is considered an ecologically relevant area due to its high floristic biodiversity index and the presence of species with contrasting ecological niches

(Alfonso and Prina 2009). *Nothofagus* spp. forests at LE occur from ~ 1,500 - 1,700 m a.s.l. (the highest altitude for *N. obliqua* in Argentina), where mean annual precipitation values are ~ 2,500 mm year⁻¹, mainly occurring as snow, and mean annual temperature is 8.6 °C (AIC, 1997-2012 series). The LE population also exhibits pronounced geographical isolation, with the nearest con-specific forests growing at more than 220 km, in Moquehue basin in Argentina, whereas the closest Chilean population is located at less than 100 km but on the other side of the Andes.

The ecological characteristics and geographical isolation of LE suggest local adaptation. Recent analyses of leaf morphology, architectural features and growth conducted in common garden trials showed a clear differentiation of the LE population compared with other Argentinean ones (Azpilicueta et al. 2014, see figure 1C for examples of LE phenotypic differentiation). High leaf area values together with low specific leaf area characterized LE individuals, which also exhibited a distinct architectural feature similar to that found in N. alpina individuals. Within a provenances trial, LE individuals showed lower height, although exhibiting high relative growth rates (Azpilicueta et al. 2014). Meanwhile, in a study of the intraspecific variation of the resprouting capacity of N. obliqua species, LE was found to differ from other Argentinean populations in several key traits (Aparicio et al. 2015). For example, LE had the lowest rate of investment in shoot per unit of root seedling biomass; it showed the lowest capacity for recovering height, in a trade-off with multiple resprouting and the highest degree of episodic frost tolerance during resprouting. All these traits as a whole suggest divergence for the resprouting syndrome.

Previous studies focusing on the genetic characterization of N. obliqua also highlighted LE population distinctiveness. Isozyme (Azpilicueta and Gallo 2009), chloroplast (cpDNA; Azpilicueta et al. 2009) and simple sequence repeat (SSR or microsatellite; Azpilicueta et al. 2013) analyses revealed the presence of exclusive variants (alleles and haplotypes, figure 1B) together with higher genetic diversity parameters in this population than those found in other N. obliqua populations. Two cpDNA haplotypes were reported for LE, one fixed in N. obliqua populations north of Lanin volcano in Argentina (39° 30' S, haplotype II figure 1B) and another one found exclusively in Chilean populations (haplotype IV, figure 1B), including those located at a similar latitude to LE (Azpilicueta et al. 2009). This cpDNA variation pattern suggested the confluence of two migratory routes during the expansion that follows glacial retraction as the origin of LE (Azpilicueta et al. 2009). An introgression pathway from the West into that region has been proposed, which was dated ~ 5,000 years BP based on pollen fossil records (Markgraf et al. 2009); and reinforced by an isozyme analysis, since a high frequency of the isozyme allele Adh-2 was detected for both LE and northern coastal populations in Chile (Azpilicueta and Gallo 2009, Azpilicueta et al. 2014). However, considering

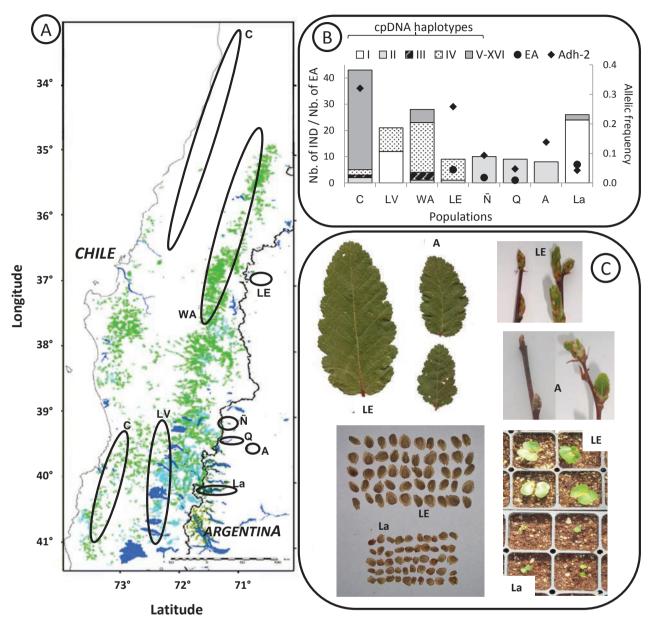


Figure 1. Distinctiveness of Lagunas de Epulauquen population. A) Entire natural distribution of *Nothofagus obliqua* forests in Argentina and Chile (green patches), showing the northernmost location of Lagunas de Epulauquen in Argentina. Ovals include *N. obliqua* populations studied by Azpilicueta *et al.* (2009, 2013) and Azpilicueta and Gallo (2009), supporting the distinctiveness of Lagunas de Epulauquen based on biochemical and molecular markers (figure 1B and in text). B) Biochemical and molecular distinctiveness of Lagunas de Epulauquen according to number of individuals showing cpDNA haplotypes, number of individuals (Nb. of IND), number of exclusive alleles for SSR (Nb. of EA) and allelic frequency of specific allele of Adh isozyme. C) Examples of phenotypic traits showing differentiation for Lagunas de Epulauquen (Azpilicueta *et al.* 2014). Population ID: C =Coast, LV = Longitudinal Valley, WA = West Andes, LE = Lagunas de Epulauquen. $\tilde{N} = \tilde{N}$ orquinco, Q = Quillén, A = Aluminé, La = Lácar. *Nothofagus obliqua* distribution map was provided by the Remote Sensing Laboratory at INTA EEA Bariloche.

Diferenciación de la población Lagunas de Epulauquen. A) distribución total de los bosques de *Nothofagus obliqua* en Argentina y Chile (parches verdes), destacando la población septentrional Lagunas de Epulauquen en Argentina. Los óvalos incluyen las poblaciones estudiadas por Azpilicueta *et al.* (2009, 2013) y Azpilicueta y Gallo (2009), mostrando la diferenciación de Lagunas de Epulauquen con base en información de marcadores bioquímicos y moleculares (figura 1B y en el texto). B) diferenciación a nivel bioquímico y molecular de acuerdo con haplotipos de cloroplasto, número de individuos (Nb. of IND), número de alelos exclusivos en microsatélites y alelos específicos en el locus isoenzimático Adh (Nb. of EA). C) Ejemplos de caracteres fenotípicos que muestran diferencias morfológicas en Lagunas de Epulauquen (Azpilicueta *et al.* 2014). ID poblacional: C = Costa, LV = Valle Longitudinal, WA = Andes Oeste, LE = Lagunas de Epulauquen, Ñ = Ñorquinco, Q = Quillén, A = Aluminé, La = Lácar. El mapa de la distribución de *Nothofagus obliqua* fue provisto por el Laboratorio de Teledetección de INTA EEA Bariloche.

that two alleles have been described for ADH locus, Adh-1 allele for *N. obliqua* and Adh-2 allele for *N. alpina* (Gallo *et al.* 1997), hybridization between both species has been postulated as the origin of LE distinctiveness based on its high frequency of Adh-2 allele; particularly historical hybridization events as *N. alpina* forests are currently absent in the area (Azpilicueta and Gallo 2009). The occurrence of natural hybridization with *N. glauca* - growing to the West of the Andean Cordillera (Chile) and situated at ~ 60 km – has been postulated by Donoso *et al.* (2004) as the origin of individuals with differential phenotypes at LE, although no analysis was made in order to validate this alternative hypothesis.

The hybridization hypothesis with N. alpina based on isozyme data requires verification. Particularly, the species-specificity of Adh alleles needs to be tested, as the isozyme analysis presents some weaknesses associated with the methodology itself and electro-morphs with similar electrical mobility could still exhibit differences in their amino acids or nucleotide constitution (e.g. Gaut and Clegg 1993). This means that it is possible that the isozyme Adh-2 allele detected in N. obliqua populations differs in its nucleotide constitution from that of N. alpina populations, and therefore isozyme heterozygote genotypes (Adh-1/Adh-2) should reveal intra-specific variability instead of signals of hybridization. A nucleotide sequence analysis of nuclear regions showed low to moderate intraspecific diversity combined with high divergence between N. obliqua and N. alpina, revealing a strong potential of nucleotide polymorphisms for the identification of species and hybrids (El Mujtar et al. 2014). Among low-copy nuclear genes, Adh has proven to be useful for inter and intraspecific analyses and to provide evidence of the hybrid origin for some species (Järvinen et al. 2004, Pan et al. 2007). Moreover, high divergence has been detected for Adh gene at nucleotide sequence between N. obligua and *N. alpina* (Nst = 0.98^2). Therefore a nucleotide sequence analysis of an Adh gene region can be used to test the hypothesis of hybridization as the origin of LE distinctiveness and also to verify the specificity of the Adh isozyme marker by sequencing individuals showing homozygote and heterozygote genotypes at isozyme Adh locus.

Phylogenetic studies in Nothofagaceae species based on the internal transcribed spacer (*ITS*) region showed correspondence with the four subgenera (Manos 1997) defined according to pollen-type and other flower and fruit characteristics, especially those related to cupule morphology (Manos 1997). Considering the utility and previous applications of *ITS* in Nothofagaceae species (Manos 1997, Acosta and Premoli 2010), we selected this marker to complement the *Adh* analysis. We also included twelve SSRs from those recently developed using next generation sequencing (NGS) technologies in *Lophozonia* subgenus; some of them showing a strong capacity to discriminate between *N. obliqua* and *N. alpina* (Torales *et al.* 2012, El Mujtar *et al.* 2014).

We focused this work on three objectives. The main objective was to evaluate the hybridization hypothesis between *N. alpina* and *N. obliqua* as the origin of the Lagunas de Epulauquen distinctiveness. Considering the use of discriminant nucleotide sequences and SSRs markers between *N. alpina* and *N. obliqua*, it is expected that hybrids will show alleles from both species in at least some of the functional and neutral analyzed genomic regions. Two secondary objectives were also defined: i) to test specificity of *Adh* isozyme alleles correlating molecular and biochemical markers and ii) to generate molecular data presently not available for *N. glauca*, *N. leonii* and *N. macrocarpa*, especially considering the alternative hypothesis of hybridization with *N. glauca*.

We expect that the results of these analyses will be useful to validate or not the inclusion of LE population within *N. obliqua* taxon and in case of confirming its inclusion, assist the domestication program of *N. obliqua* species.

METHODS

Considering our main objective, we sampled twelve individuals from LE and one individual of each putative parental species (N. obliqua and N. alpina) as molecular reference. A higher sample size was not required as genetic diversity of molecular makers from both species has been previously determined supporting the power of selected markers to discriminate both species (Manos 1997, El Mujtar et al. 2014, Sola et al. 2016). Samples representing putative parental species have homozygote genotypes for Adh isozyme alleles, N. obliqua Adh 1 - Adh 1 and N. alpina Adh 2 - Adh 2 and therefore they were also used to test species specificity of isozyme alleles. For this objective we additionally included one individual showing N. obliqua phenotypic characteristics and heterozygote isozyme genotype for Adh locus (Adh 1 - Adh 2), and six individuals belonging to a N. obliqua Chilean coastal population exhibiting high Adh-2 isozyme allele. Finally considering the third objective, the species entities N. glauca, N. leonii and N. macrocarpa were included. The sampled individuals and the species and populations they belonged to - including the previously defined isozyme genotype at Adh locus - are shown in table 1. Total DNA was extracted from leaf/bud tissue according to conditions described in Azpilicueta et al. (2013). The partial region of the Adh gene was amplified with the primers vF1:CTCAGGCAAAYGAAGTCCG and vR7n:AACAAAGTGGTAAATGGGCTG designed on conserved regions of sequence alignments of Fagaceae and Nothofagaceae species. This primer pair was targeted to one of the two loci of Adh detected in Fagaceae species, which seems to be closer to the unique Adh locus of Arabidposis thaliana, and designed to amplify the region between Exon2-4 (complete intron2-exon3-intron3) accor-

² V El Mujtar, unpublished data.

Table 1. Species and geographical location of sampled individuals. Genotype obtained for *Adh* locus based on isozyme analysis and sample size analyzed at each genetic marker are provided.

Especie y localización geográfica de los individuos muestreados. Se incluye la información del genotipo en el locus Adh basado en análisis isoenzimáticos y el tamaño de la muestra analizada en cada marcador.

Species	Sampling location	ID	Adh genotype ^b	N (Adh)	N (ITS)	N (SSR)
?	Lagunas de Epulauquen	LE	-	12	8	6
Nothofagus obliqua	Northern Pacific coast ^a	No-CP	-	6	4	4
	Field identification	No	11	1	1	1
	Field identification	No	12	1	1	1
Nothofagus alpina	Field identification	Na	22	1	1	1
Nothofagus glauca	Field identification ^a	Ng	-	4	2	2
Nothofagus leonii	Field identification ^a	Nl	-	2	2	2
Nothofagus macrocarpa	Field identification ^a	Nm	-	2	2	2

N: number of analyzed individuals.

^a samples from Chile; all other samples were collected in Argentina.

^b Adh genotype was extracted from Marchelli and Gallo (2000) and Azpilicueta and Gallo (2009). -: no information.

ding to A. thaliana gene structure description. Touchdown (TD) PCR was used for amplifying fragments according to the following conditions: an initial denaturating step for 3 min at 94 °C followed by 2 cycles at 94 °C for 1 min, a primer annealing step at 65 °C for 1 min, an extension step at 72 °C for 1 min, followed by 18 cycles of denaturation at 93 °C for 45 s, primer annealing at 64 °C for 30 s (- 0.5 °C per cycle), extension at 72 °C for 1 min, then 20 cycles of denaturation at 92 °C for 30 s, primer annealing at 55 °C for 30 s, and extension at 72 °C for 1 min, and finally a 15 min extension step at 72 °C. PCR reactions were carried out using a 1X Buffer with GoTaq DNA polymerase, containing a final concentration of 0.5 µM of each primer, 200 µM dNTPs, 3.0 mM of MgCl₂, 0.3 mg mL⁻¹ of BSA, and 0.05 U µL⁻¹ of GoTaq DNA polymerase (Promega, USA). For ITS including 5.8S region amplification, we used primers CY1 and CY3 (Wright et al. 2006) and the PCR conditions reported in Acosta and Premoli (2010). MyCycler (BIORAD) PCR thermocycler was used for all TD-PCR reactions. ITS and Adh amplicons were purified by precipitation with sodium acetate and isopropanol and directly sequenced with an ABI 3730 XL DNA Analyzer (Applied Biosystems) at the Sequencing and Genotyping Services of CNIA (INTA, Argentina). Sequences were aligned with BIOEDIT v.7.0.5.3, and polymorphic sites revised with CodonCode v.4.0.4 (CodonCode Corporation).

Twelve SSRs (table 2) were selected from those developed by Torales *et al.* (2012) and El Mujtar *et al.* (2014), considering the presence of species-specific alleles for *N. obliqua* and *N. alpina* and the number of detected alleles. In addition, 16 samples from Lácar watershed (Argentina) previously genotyped at ten out of these

12 SSRs were included in the analysis: eight natural hybrids between N. obliqua and N. alpina (NoxNa), four N. obligua and four N. alpina individuals were selected based on their genotypes, with hybrids showing alleles from both species at more than one loci (Sola et al. 2016). PCR reactions of SSRs developed from transcriptome 454 NGS (table 2) were carried out using a 1X Buffer with Go-Taq DNA polymerase, containing a final concentration of 300 μ M dNTPs, 3.0 mM of MgCl₂, and 0.15 U μ L⁻¹ of Go-Taq DNA polymerase (Promega, USA), as single or multiplex PCR (table 2). PCR reactions of SSRs developed from genome 454 NGS (table 2) were carried out using a 1X Buffer with GoTaq DNA polymerase, containing a final concentration of 100 µM dNTPs, 3.0 (or 2.0) mM of MgCl₂, 0.3 mg mL⁻¹ of BSA and 0.05 U µL⁻¹ of GoTaq DNA polymerase (Promega, USA), as single or multiplex PCR (table 2) MyCycler (BIORAD) PCR thermocycler was used for TD-PCR, according to conditions detailed in table 2. Amplified loci were pooled and genotyped at the INTA service mentioned previously. The SSR profiles were examined and scored using GENEMARKER version 1.95 (SoftGenetics). GENALEX v.6.5 was used for principal coordinate analysis (PCoA).

The software package Phylip v3.573c was used for the phylogenetic analysis of the complete combined sequence obtained by *Adh* and *ITS* amplified gene regions, applying SEQBOOT (to produce 1,000 bootstrapped data sets), DNAPARS for the parsimony analysis using multiple data (1,000), jumble (100 times) and outgroup identification, and CONSENSE for the identification of the consensus tree. Drawgram from Phylip v3.573c was applied for tree graphics editing. IUPAC codification was used for poly-

Lista de microsatélites utilizados en este estudio, incluyendo condiciones de la PCR y del genotipado.

SSR name	SSR Development	Dye ^a	Forward ^b	Reverse ^b	PCR	PCR condition	Genotyping
Notho226 e,f	genome NGS	FAM	0.025	0.1	PCR1	TD 64-55 °	PoolA
Notho224 $^{\rm f}$	genome NGS	FAM	0.025	0.1	Multiplex1	TD 64-55 °	PoolA
Notho218 $^{\rm f}$	genome NGS	HEX	0.025	0.1	Multiplex1	TD 64-55 °	PoolA
Notho214 $^{\rm f}$	genome NGS	HEX	0.025	0.1	Multiplex1	TD 64-55 °	PoolA
32SN ^g	transcriptome NGS	HEX	0.35	0.35	Multiplex2	TD 60-50 ^d	PoolB
IN0230a ^g	transcriptome NGS	FAM	0.35	0.35	Multiplex2	TD 60-50 ^d	PoolB
IN0192b ^g	transcriptome NGS	NED	0.25	0.25	Multiplex2	TD 60-50 ^d	PoolB
8SN ^g	transcriptome NGS	HEX	0.25	0.25	PCR2	TD 60-50 ^d	PoolB
IN0597 a ^g	transcriptome NGS	FAM	0.25	0.25	PCR3	TD 55-45 ^d	PoolB
23SN ^g	transcriptome NGS	NED	0.25	0.25	Multiplex3	TD 60-50 ^d	PoolC
13SN ^g	transcriptome NGS	FAM	0.25	0.25	Multiplex3	TD 60-50 ^d	PoolC
19SN ^g	transcriptome NGS	HEX	0.25	0.25	Multiplex3	TD 60-50 ^d	PoolC

^a For SSRs from genome NGS dyes were incorporated using M13 method, with 0.1 μ M of tagged modified forward primer added at the PCR reaction. For SSRs from transcriptome NGS dyes were directly incorporated into forward or reverse primers. ^b Concentration of each primer in the PCR reaction, indicated as μ M. ^c PCR condition according to El Mujtar *et al.* (2014). ^d PCR conditions according to Torales *et al.* (2012). ^c PCR for this locus was carried out with 0.2 mM MgCl₂. ^f Primers from El Mujtar *et al.* (2014) ^g Primers from Torales *et al.* (2012). Locus name in bold indicates species-specific markers for *N. obliqua* and *N. alpina*.

^a Para SSRs derivados de datos genómicos la marca fluorescente fue incorporada usando el método de M13, adicionando a la reacción de PCR el primer directo modificado (incluyendo la secuencia de M13) y marcado en concentración 0.1 µM en la reacción de PCR. Para SSRs derivados de transcriptoma, la marca fluorescente fue incorporada directamente a los primers *forward* o *reverse*. ^b Concentración de los primers en la reacción de PCR (µM). ^c Condiciones de PCR según El Mujtar *et al.* (2014). ^d Condiciones de PCR según Torales *et al.* (2012). ^e La PCR para este locus se realizó con 0.2 mM MgCl₂. ^f Primers tomados de El Mujtar *et al.* (2014) ^g Primers tomados de Torales *et al.* (2012). El nombre del locus en negrita indica marcador especie-específico para *N. obliqua y N. alpina.*

morphic sites and indels were considered as a fifth state (as a single polymorphism independently of the indel's length). Individuals representing unique variants for the combined nuclear amplified regions per analyzed taxon or LE population were retained for phylogenetic studies. Outgroup sequences for *Adh* and *ITS* were D63463.1 from *A. thaliana* and GQ863264.1 from *N. pumilio* respectively, obtained from Genbank database of National Center for Biotechnology Information (NCBI 2016).

RESULTS

Adh gene analyses. The Adh amplified region has 665 bp in *N. alpina*, while a size of 682 bp was found in the remainder of the sampled taxa. Twenty-four substitutions were detected between *N. obliqua* and *N. alpina* but all individuals of the LE population showed the allele of *N. obliqua* at each of the polymorphic sites (figure 2). A similar result was found comparing *N. obliqua* and *N. glauca*, although only two polymorphic sites were detected. Positions 44 and 47 of the sequence alignment showed polymorphism for *N. glauca*, with the major allele (G and C, respectively) differing from that fixed for all other taxa

596

including also samples from LE (A and T, figure 2). Therefore, *Adh* nucleotide sequences do not reveal signs of hybridization with *N. alpina* nor *N. glauca* in LE individuals.

Two polymorphic sites (41, 45) were detected considering Adh sequences of N. obliqua individuals with isozyme genotypes Adh 1-Adh 1 (No[11]) and Adh 1-Adh 2 (No[12]). The individual No[11] was homozygotic at both sites (T, G), whereas the individual No[12] was heterozygotic at these positions (Y=T/C, K=G/T). The individual of N. alpina with isozyme genotype Adh 2 - Adh 2 (Na[22]) was homozygotic for both sites (T, C), but differed from the nucleotide combination of No[11]. The two variants detected in No[11] and No[12] individuals were also present in the remaining sequenced samples from N. obliqua and the LE population. They may be considered as two haplotypes (T,G and Y,K) since recombination between these sites was not detected (e.g. T,K; T,Y; G,K or G,Y); in addition, one individual of LE showed the alternative homozygotic haplotype (C,T). The haplotype of Na[22] was not detected among samples of N. obliqua or the LE population. Furthermore, we did not detect evidence of hybridization with N. alpina in any nucleotide position of sequences from No[12] or LE individuals.

ITS analyses. ITS amplified sequence (795 bp) was obtained for 21 out of 29 individuals. Nucleotide sequences for N. obliqua, N. alpina and N. glauca were in agreement with the sequences reported by Manos (1997) and Acosta and Premoli (2010), although several indels and polymorphisms were observed in Manos's sequences. Two nucleotide substitutions (positions 16 and 27) were detected between N. obliqua and N. alpina considering the available sequences (including those reported by the previously mentioned authors), whereas all individuals of the LE population showed N. obliqua allele at these positions (figure 2). On the other hand, one fixed nucleotide substitution (A in position 18, figure 2) was detected between N. glauca and all other taxa (C fixed allele), with the exception of both N. leonii individuals that were heterozygous at this position (M=C/A). All individuals of the LE population were homozygous for C allele (figure 2). Therefore, although a lower divergence was observed among species for ITS than for Adh nucleotide sequences, no evidence of hybridization with N. alpina or N. glauca in LE individuals was detected.

Phylogenetic analysis. A phylogenetic analysis of fourteen unique variants of the combined *Adh* and *ITS* sequence showed the LE population together with *N. obliqua*, *N. leonii* and *N. macrocarpa* (figure 3). *N. glauca* joined this clade, whereas *N. alpina* was clearly separated from all other taxa (figure 3). This last result was expected, based on the high number of nucleotide differences between *N. alpina* and the other species. *N. obliqua* individuals from No-CP

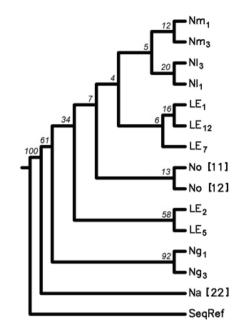


Figure 3. Consensus tree (over 1,000 trees) based on combined Adh and ITS sequence analysis. Bootstrap values (%) are given over the tree branches.

Árbol consenso (a partir de 1.000 árboles) basado en el análisis de las secuencias combinadas de Adh e ITS. Los valores de *bootstraping* (%) se muestran sobre las ramas.

Adh-ITS combined sequences LE1 = LE3 = LE4 = LE8; LE2 = LE9 SeqRef: outgroup sequence Adh-ITS secuencia combinada LE1 = LE3 = LE4 = LE8; LE2 = LE9 SeqRef: secuencia outgroup

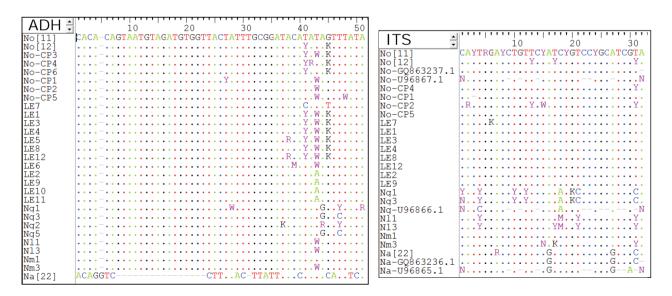


Figure 2. Number and location of informative nucleotide polymorphisms between nuclear sequences of *N. obliqua*, *N. alpina* and *N. glauca*.

Número y localización de polimorfismos de nucleótido entre las secuencias nucleares de N. obliqua, N. alpina and N. glauca.

No-GQ863237.1 and Na-GQ863236.1 reference sequences from Acosta and Premoli (2010). No-U96867.1, Ng-U96866.1 and Na-U96865.1 reference sequences from Manos (1997). No-GQ863237.1 y Na-GQ863236.1 secuencia de referencia tomada de Acosta y Premoli (2010). No-U96867.1, Ng-U96866.1 y Na-U96865.1 secuencia de referencia tomada de Manos (1997). population were not included in this analysis because intraspecific diversity could influence the results of phylogenetic relationships between species, based mainly on the low nucleotide differentiation detected among *N. obliqua*, *N. leonii*, *N. macrocarpa* and LE individuals. However, the same result was obtained when individuals from the northern Pacific coast (No-CP) were included (result not shown).

Microsatellite analyses. PCoA based on SSR data showed clear discrimination of *N. alpina*. Individuals from the LE population and *N. obliqua*, together with the other sampled

taxa, formed another cloud (figure 4A). A similar result was obtained when previously genotyped individuals were included in the analysis (figure 4B). SSR data showed an enormous capacity to separate *N. alpina* and NoxNa hybrids, while individuals from the LE population, *N. obliqua*, and the other taxa remain as a single cloud.

DISCUSSION

The haplotypes described for *Adh* are located at intron3 of the gene, and therefore could not be directly associated

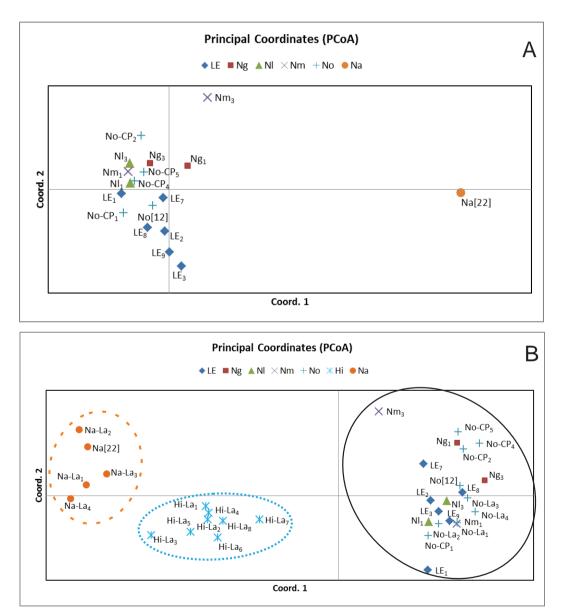


Figure 4. Principal coordinate analysis (PCoA) based on the genotypes obtained at 12 nuclear microsatellite loci. A) Using individuals reported in table 1. B) Including previously genotyped individuals at Lácar watershed: No-La, Na-La and Hi-La corresponding to *N. obliqua*, *N. alpina* and hybrid (NoxNa) individuals, respectively.

Análisis de coordenadas principales basado en los genotipos obtenidos en 12 *loci* microsatélites. A) utilizando individuos informados en el cuadro 1. B) incluyendo individuos de la cuenca Lácar previamente genotipados: No-La, Na-La e Hi-La corresponden a individuos de *N. obliqua*, *N. alpina* e híbrido (NoxNa), respectivamente.

with polymorphisms in translated sequences. However, a hypervariable region of the Adh gene has been described for A. thaliana, including parts of intron3 and exon4 and showing two predominant haplotypes among 37 analyzed ecotypes (Hanfstingl et al. 1994). Moreover, extensive linkage disequilibrium has been described for the Adh gene in Populus tremula (Ingvarsson 2005). Mutagenesis and structural studies indicate that this hypervariable region is functionally important, and that amino acid replacement polymorphisms at exon4 correlate with mobility differences between isozyme Adh electrophoretic alleles (Hanfstingl et al. 1994, Vinichenko et al. 2004). Our results suggest the occurrence of different sequence variants in the Adh gene that explains the patterns of isozyme Adh alleles at LE. That is, isozyme Adh 2 allele of N. obliqua and N. alpina have the same electrophoretic mobility, though they differ in their nucleotide constitution. Therefore isozyme genotype Adh 1 - Adh 2 at LE reveals intra-specific variability but not hybridization with N. alpina. Hence, these results indicate constraints in the application of Adh isozyme as a biochemical marker to identify hybrids between N. obliqua and N. alpina, since the contribution of intraspecific variation and hybridization in the genotype Adh 1 - Adh 2 cannot be resolved. That is in sympatric populations of the hybridizing species N. obliqua and N. alpina, we cannot exclude the presence of hybrids having the isozyme genotype $Adh 1_{N_0}$ - $Adh 2_{N_a}$, which could not be separated from isozyme genotype Adh 1_{N_0} - Adh 2_{N_0} . The isozyme analysis needs to be replaced therefore by a nucleotide sequence analysis of candidate genes, which has proven its capacity to discriminate clearly between these species.

Phylogenetic analyses show LE population clearly separated from *N. alpina*, without evidence of hybridization. Even more, SSRs analyses reveal the absence of LE individuals located in the quadrant of hybrids *N. obliqua* x *N. alpina* hybrids. Globally our results indicate that the hypothesis of hybridization between *N. alpina* and *N. obliqua* as the origin of distinctiveness of LE should be refuted.

Excluding N. alpina, phylogenetic and SSRs analyses of molecular data reveals low power to discriminate among N. leonii, N. macrocarpa, N. glauca and N. obliqua. Considering the alternative hypothesis of hybridization between N. obliqua and N. glauca as origin of distinctiveness of Lagunas de Epulauquen population, evidence of hybridization was not detected regarding nucleotide polymorphism at Adh and ITS; that is, specific alleles of N. glauca were not detected on individuals of Lagunas de Epulauquen. However N. glauca, N. obliqua and Lagunas de Epulauquen individuals were located within a single cloud that also included N. leonii and N. macrocarpa individuals; this result indicates that SSR alleles are mostly shared between these species. Therefore, it is actually difficult to make strong conclusions given the low divergence showed by the applied markers.

The two haplotypes of *Adh* sequence described for *N. obliqua* were also found at Lake Lácar watershed in 25 individuals of this species, although a low frequency of the haplotype Y,K was detected (El Mujtar *et al.* 2012). While among *N. alpina* individuals, only the haplotype T,C was found. On the other hand, both LE and Chilean populations exhibited a high frequency of haplotype Y,K. Although no C,T haplotype was found in the Chilean population - probably due to the low sample size - the Chilean and LE populations exhibited a similar frequency pattern, which was different from that found at Lácar (table 3).

Considering the correlation between *Adh* haplotypes and isozyme genotypes previously described, the trend detected for haplotype frequency among the Chilean, LE and

Table 3. Comparison of *Adh* frequency of sequence haplotypes and isozyme genotypes among Chile, Lagunas de Epulauquen (LE) and Lácar populations.

	Frequency of <i>Adh</i> region sequence (haplotypes at polymorphic sites)				Frequency of <i>Adh</i> genotypes (at isozyme level)			
Population	Ν	TG	YK	СТ	Ν	11	12	22
Chile	6	0.50	0.50	0.00	53	0.29	0.41	0.30
LE	12	0.42	0.50	0.08	120	0.55	0.38	0.07
Lácar	25	0.88	0.12	0.00	91	0.93	0.05	0.02

Comparación de la frecuencia de los haplotipos de Adh y los genotipos isoenzimáticos entre poblaciones de Chile, Lagunas de Epulauquen (LE) y Lácar.

N= number of individuals

Isozyme genotype frequency at Adh gene region was extracted from Azpilicueta and Gallo (2009) and Azpilicueta *et al.* (2014). Mean frequency values taken from six populations within the watershed are presented for Lácar population.

Lácar Adh sequences (El Mujtar et al. 2012).

N= número de individuos

La frecuencia de genotipos isoenzimáticos en la región génica Adh fue tomada de Azpilicueta and Gallo (2009) y Azpilicueta *et al.* (2014). Se presenta el valor medio de la frecuencia de seis poblaciones de la cuenca Lácar.

Secuencias de Adh en la cuenca Lácar (El Mujtar et al. 2012).

Lácar populations resembles those reported for isozyme alleles at *Adh* locus. These results suggest a common origin for Chilean and LE populations and therefore reinforce the hypothesis of convergence of migratory routes around the geographic area of LE. Even more, new alleles at SSRs were detected at two of the analyzed loci for LE and *N. obliqua*, *N. macrocarpa*, *N. glauca* and *N. leonii* Chilean individuals suggesting a higher genetic diversity for Chilean populations and LE, which also supports the hypothesis of convergence at LE.

Taking into account the fitness exhibited by LE individuals in some characters of commercial interest, its inclusion in *N. obliqua* breeding programs should be seriously considered. Notwithstanding, some considerations must be stated based on the presence of exclusive molecular variants at LE population. For restoration activities, a sub-zone compound of LE population was distinguished from the northern group (Azpilicueta *et al.* 2014). Based on this, restoration activities at LE should use propagation material - seeds or plants - belonging to the same origin in order to maintain its original genetic structure, mainly taking into account that LE population constitutes a provincial protected area.

CONCLUSIONS

The results of this work argue against the hypothesis of hybridization between *N. obliqua* and *N. alpina*. The results suggest a common origin for LE and Chilean *N. obliqua* populations and support the hypothesis of convergence of two migratory routes at LE.

The distinctiveness of Lagunas de Epulauquen population could therefore be the combined result of genetic variation, convergence and the influence of evolutionary processes related to its particular ecological conditions and marked geographical isolation (*i.e.* divergence). Further studies with large genome coverage will be useful to identify the genomic bases of the distinctiveness of the LE population. From this point on, Lagunas de Epulauquen population could be included within the *Nothofagus obliqua* taxon, and considered for Argentinean *N. obliqua* domestication and conservation programs.

We highlight the importance of molecular markers combination such as the sequencing of low and high copy number nuclear gene and SSR genotyping in order to evaluate different hypotheses, and for the analysis of related species and putative hybrids.

We cannot rule out hybridization with *N. glauca*, although evidence was not detected at *Adh* or *ITS* sequences. The low differentiation found among *N. obliqua*, *N. macrocarpa* and *N. leonii* - for all analyzed markers (*Adh*, *ITS* and SSRs) - indicates that further studies are necessary in order to identify highly divergent molecular markers among these taxa and to achieve a higher degree of accuracy in their taxonomic identification, including individuals sampled from different natural populations.

ACKNOWLEDGEMENTS

Special thanks to Paula Marchelli who encouraged us to do these analyses. Thanks to Alejandro Aparicio and Mario Pastorino for collecting the samples of the Lagunas de Epulauquen population and Georgina Sola for providing SSR data from individuals at Lácar watershed. The authors would also like to thank Viviana Becerra and Mario Paredes from Instituto de Investigaciones Agropecuarias (INIA, Chile), Quilamapu and Braulio Gutiérrez from Instituto Forestal de Chile for kindly providing Chilean samples within the project "Genetic characterization of populations of N. obliqua (Mirb. et Oerst.) and N. alpina ((Phil.) Dim. et Mil. through molecular and isozyme markers" BID-FONTAGRO-IICA. We are indebted to Alejandro Aparicio who critically revised and improved our manuscript. This study was funded by Subprograma Nothofagus - PROMEF (Argentina) and PNFOR1104062 and PNFOR1104063 from Instituto Nacional de Tecnología Agropecuaria (INTA, Argentina).

María M Azpilicueta and Verónica A El Mujtar contributed equally to the paper.

REFERENCES

- Acosta MC, AC Premoli. 2010. Evidence of chloroplast capture in South American *Nothofagus* (subgenus *Nothofagus*, Nothofagaceae). *Molecular Phylogenetics and Evolution* 54: 235-242. DOI: 10.1016/j.ympev.2009.08.008.
- Alfonso G, A Prina. 2009. Catálogo florístico de la Reserva lagunas de Epu Lauquen, Departamento Minas, Provincia de Neuquén, Argentina. *Ernstia* 19: 109-136
- Aparicio AG, SM Zuki, MM Azpilicueta, FA Barbero, MJ Pastorino. 2015. Genetic versus environmental contributions to variation in seedling resprouting in *Nothofagus obliqua*. *Tree Genetics & Genomes* 11: 23.
- Azpilicueta MM, LA Gallo. 2009. Shaping forces modelling genetic variation patterns in the naturally fragmented forests of a South American Beech. *Biochemical Systematic and Ecology* 37: 290-297. DOI: 10.1016/j.bse.2009.05.004.
- Azpilicueta MM, P Marchelli, LA Gallo. 2009. The effects of Quaternary Glaciations in Patagonia as evidenced by chloroplast DNA phylogeography of Southern beech *Nothofa*gus obliqua. Tree Genetics and Genomes 5: 561-571. DOI: 10.1007/s11295-009-0209.
- Azpilicueta MM, LA Gallo, M van Zonneveld, E Thomas, C Moreno, P Marchelli. 2013. Management of *Nothofagus* genetic resources: Definition of genetic zones based on a combination of nuclear and chloroplast marker data. *Forest Ecology and Management* 302: 414-424. DOI: 10.1016/j. foreco.2013.03.037.
- Azpilicueta MM, MJ Pastorino, J Puntieri, F Barbero, A Martínez- Meier, P Marchelli, LA Gallo. 2014. Robles in Lagunas de Epulauquen, Argentina: previous and recent evidence of their distinctive character. *Revista Chilena de Historia Natural* 87: 24-35. DOI: 10.1186/s40693-014-0024-0.
- Donoso C, A Premoli, L Gallo, R Ipinza. 2004. Variación intraespecífica en las especies arbóreas de los bosques templados de Chile y Argentina. Santiago, Chile. Editorial Universi-

taria. 419 p.

- Donoso C, LR Landrum. 1979. Nothofagus leonii Espinosa, a natural Irbid between Nothofagus obliqua (Mirb.) Oerst. and Nothofagus glauca (Phil.) Krasser. New Zealand Journal of Botany 17: 353-360.
- El Mujtar V, Gallo L, Garnier-Géré P. 2012. Nucleotide diversity patterns at genes potentially involved in Nothofagus species divergence along altitudinal gradient. IUFRO Conference "Genetics of Fagaceae and Nothofagaceae", 9-12 October, Bordeaux, France.
- El Mujtar V, L Gallo, T Lang, P Garnier-Géré. 2014. Development of genomic resources for Nothofagus species using next-generation sequencing data. *Molecular Ecology Resources* 14(6): 1281-1295. DOI: 10.1111/1755-0998.12276.
- Gallo LA, P Marchelli, A Breitembücher. 1997. Morphological and allozymic evidence of natural hybridisation between two Southern beeches (*Nothofagus* spp) and its relation to heterozygosity and height growth. *Forest Genetics* 4: 15-23.
- Gaut BS, MT Clegg. 1993. Nucleotide polymorphism in the *Adh1* locus of pearl millet (*Pennisetum glaucum*) (Poaceae). *Genetics* 135: 1091-1097.
- Hanfstingl U, A Berry, EA Kellogg, JT Costa, W Rudigert, FM Ausubel. 1994. Haplotypic Divergence Coupled with Lack of Diversity at the Arabidopsis thaliana Alcohol Dehydrogenase Locus: Roles for Both Balancing and Directional Selection? *Genetics* 138: 811-828.
- Heenan PB, RD Smissen. 2013. Revised circumscription of Nothofagus and recognition of the segregate genera Fuscospora, Lophozonia, and Trisyngyne (Nothofagaceae). Phytotaxa 146: 1-31. DOI: 10.11646/phytotaxa.146.1.1
- Hill RS, GJ Jordan J, Macphail MK. 2015. Why we should retain Nothofagus sensu lato. *Australian Systematic Botany* 28: 190-193
- Ingvarsson P. 2005. Nucleotide Polymorphism and Linkage Disequilibrium Within and Among Natural Populations of European Aspen (*Populus tremula* L., Salicaceae). *Genetics* 169: 945-953. DOI: 10.1534/genetics.104.034959.
- Järvinen P, A Palme, LO Morales, M Lännenpää, M Keinänen, T Sopanen, M Lascoux. 2004. Phylogenetic relationships of *Betula* species (Betulaceae) based on nuclear *ADH* and chloroplast *MATK* Sequences. *American Journal of Botany* 91: 1834-1845. DOI: 10.3732/ajb.91.11.1834.
- NCBI (National Center for Biotechnology Information, US). 2016. The Nucleotide database is a collection of sequences from several sources, including GenBank, RefSeq, TPA and PDB. Genome, gene and transcript sequence data provide the foundation for biomedical research and discovery. Available in http://www.ncbi.nlm.nih.gov/nuccore
- Manos PS. 1997. Systematics of *Nothofagus* (Nothofagaceae) based on rDNA spacer sequences (ITS): taxonomic congruence with morphology and plastid sequences. *American Journal of Botany* 84: 1137-1155. DOI: 10.2307/2446156.
- Marchelli P, LA Gallo. 2000. Genetic analysis of isozyme variants

in open pollinated families of southern beech *Nothofagus nervosa* (Phil.) Dim. *et* Mil. *Silvae Genetica* 49: 90-98.

- Markgraf V, C Whitlock, RS Anderson, A Garcia. 2009. Late Quaternary vegetation and fire history in the northernmost forest region: Mallín Vaca Lauquen, Neuquén province, Argentina. *Journal of Quaternary Science* 24: 248-258. DOI: 10.1002/jqs.1233.
- Pan J, D Zhang, T Sang. 2007. Molecular phylogenetic evidence for the origin of a diploid hybrid of *Paeonia* (Paeoniaceae). *American Journal of Botany* 94: 400-408. DOI: 10.3732/ ajb.94.3.400.
- Sabatier Y, MM Azpilicueta, P Marchelli, M González-Peñalba, L Lozano, L García, A Martínez, LA Gallo, F Umaña, D Bran, MJ Pastorino. 2011. Distribución natural de *Nothofagus alpina* y *Nothofagus obliqua* (Nothofagaceae) en Argentina, dos especies de primera importancia forestal de los bosques templados norpatagónicos. *Boletín de la Sociedad Argenina de Botánica* 46: 131-138.
- Sola G, El Mujtar V, Tsuda Y, Vendramin GG, Gallo LA, 2016. The effect of silvicultural management on the genetic diversity of a mixed *Nothofagus* forest in Lanín National Reserve, Argentina. *Forest Ecology and Management* 363: 11-20.
- Torales SL, M Rivarola, MF Pomponio, P Fernández, CV Acuña, P Marchelli, S Gonzalez, MM Azpilicueta, HE Hopp, LA Gallo, NB Paniego, SN Marcucci Poltri. 2012. Transcriptome survey of Patagonian southern beech *Nothofagus nervosa* (= *N. alpina*): assembly, annotation and molecular marker discovery. *BMC Genomics* 13: 291 DOI: 10.1186/1471-2164-13-291.
- Veblen TT, C Donoso, T Kitzberger, AJ Rebertus. 1996. Ecology of southern Chilean and Argentinean *Nothofagus* forests. *In* Veblen TT, RS Hill, J Read eds. The ecology of biogeography of *Nothofagus* forests. New Haven, USA. Yale University Press. p. 293-353.
- Vergara R, MA Gitzendanner, DE Soltis, P Soltis. 2014. Population genetic structure, genetic diversity, and natural history of the South American species of *Nothofagus* subgenus *Lophozonia* (Nothofagaceae) inferred from nuclear microsatellite data. *Ecology and Evolution* 4(12): 2450-2471. DOI: 10.1002/ece3.1108.
- Villagrán C. 1991. Historia de los bosques templados del sur de Chile durante el Tardiglacial y Postglacial. *Revista Chilena de Historia Natural* 64: 447-460.
- Vinichenko NA, KA Golovnina, AG Blinov, OO Sokolova (Antonova), EV Levites. 2004. Molecular Differences between Alleles Adh1-F and Adh1-S in Sugar Beet *Beta vulgaris* L. *Russian Journal of Genetics* 40: 172-177. DOI: 10.1023/b: ruge.0000016991.44399.7d.
- Wright S, J Keeling, L Gillman. 2006. The road from Santa Rosalia: a faster tempo of evolution in tropical climates. *Proceeding of the National Academy of Science. USA* 103: 7718-7722. DOI: 10.1073/pnas.0510383103.

Recibido: 26.11.15 Aceptado: 16.08.16