MS Editions

Blacpma

BOLETIN LATINOAMERICANO Y DEL CARIBE DE PLANTAS MEDICINALES Y AROMÁTICAS
19 (3): 300 - 313 (2020)
© / ISSN 0717 7917 / www.blacpma.ms-editions.cl

Artículo Invitado | Invited Article

Phylogenetic relationships of plant species from the flowering desert of the Atacama Region

[Relaciones filogenéticas de especies de plantas que crecen en el desierto florido de la Región de Atacama]

Roberto Contreras[1], Liesbeth van den Brink[2], Bernardo Sepúlveda[1], Fernanda Aguayo[1] & Vincenzo Porcile[1]

[1]*Centro Regional de Investigación y Desarrollo Sustentable de Atacama (CRIDESAT), Universidad de Atacama, Copiapó, Chile.*
[2]*Department of Plant Ecology, Universität Tübingen, Germany.*
Contactos | Contacts: **Roberto CONTRERAS - *E-mail address:*** **roberto.contreras@uda.cl**

Abstract: Every 3 to 7 year angiosperms species of the flowering desert appear in the Atacama Region of Chile, as a result of the climatic phenomenon "El Niño". Our objective was to evaluate the universality of *matK* and *rbcL* barcode markers of these species, and validate their taxon through phylogenetic relationships. *Argemone hunnemannii*, *Oenothera coquimbensis*, *Malesherbia humilis*, *Leucocoryne appendiculata*, *Loasa elongata*, *Nicotiana solanifolia*, *Stachys grandidentata*, *Aristolochia chilensis*, *Alstroemeria kingii* and *Adesmia eremophila*, almost all classified as endemic to Chile, were collected in Pan de Azúcar and Llanos de Challe National Park (Atacama Region, Chile) at the end of October 2017. The phylogeny of these ten angiosperm species from the flowering desert was analyzed using *rbcL* and *matK* markers with the maximum likelihood and Bayesian inference methods. The results showed that 70% of the species can be distinguished with the *matK* or *rbcL* locus, however, 100% were distinguished using both loci. The phylogenetic results showed that the species formed clades with high reliability and high support with both the *matK* and *rbcL* genes, when comparing our results with sequences obtained from GenBank. The *matK* and *rbcL* genes are efficient markers for analyzing phylogenetic relationships and validating the taxonomy of flowering species.

Keywords: Atacama Desert; *matK*; *rbcL*; DNA barcoding; Flowering Desert

Resumen: Las especies de angiospermas del Desierto Florido de la Región de Atacama de Chile aparecen cada 3 a 7 años, influenciado por el fenómeno climático "El Niño". Nuestro objetivo fue evaluar la universalidad de los marcadores de código de barra *matK* y *rbcL* de estas especies, y validar su taxón por medio de relaciones filogenéticas. Las especies *Argemone hunnemannii*, *Oenothera coquimbensis*, *Malesherbia humilis*, *Leucocoryne appendiculata*, *Loasa elongata*, *Nicotiana solanifolia*, *Stachys grandidentata*, *Aristolochia chilensis*, *Alstroemeria kingii* y *Adesmia eremophila* son clasificadas la mayoría endémicas de Chile. Estas especies fueron colectadas en el Parque Nacional Pan de Azúcar y Llanos de Challe, Región de Atacama, Chile. La colecta se realizó a fines de octubre de 2017. Con los marcadores *rbcL* y *matK* se analizó la filogenia con los métodos máxima verosimilitud e inferencia bayesiana en diez especies de angiosperma del Desierto Florido. Los resultados mostraron que el 70% de las especies pueden ser distinguidas con un locus *matK* o *rbcL*, sin embargo, el 100% se distinguió usando ambos locus. Los resultados filogenéticos mostraron que las especies formaron clados con alta fiabilidad y alto soporte tanto con los genes *matK* y *rbcL*, al comparar con accesos de secuencias obtenidas de GenBank. Lo genes *matK* y *rbcL* son marcadores eficientes para analizar relaciones filogenéticas y validar el taxón de las especies de flor.

Palabras clave: Desierto de Atacama; *matK*; *rbcL*; Código de barras de ADN; Desierto florido

Recibido | Received: December 4, 2019
Aceptado | Accepted: December 18, 2019
Aceptado en versión corregida | Accepted in revised form: December 24, 2019
Publicado en línea | Published online: May 30, 2020
Este artículo puede ser citado como / This article must be cited as: R Contreras, L van den Brink, B Sepúlveda, F Aguayo, V Porcile. 2020 Phylogenetic relationships of plant species from the flowering desert of the Atacama Region. **Bol Latinoam Caribe Plant Med Aromat** 19 (3): 300 – 313. **https://doi.org/10.37360/blacpma.20.19.3.18**

INTRODUCTION

Occasional rains in the Atacama Desert reveal a high biodiversity of plant species that germinate, flower and reproduce in a short period of time. The Atacama Desert is the driest and oldest desert on the planet, with constant hyper aridity for the last 10 million years (Sun *et al*., 2018). The flora that survives in the Atacama Desert has adapted to long periods of drought; where, plant structures such as bulbs and seeds can wait for years for a rain event to germinate, grow and bloom. In 2015 and 2017 the Atacama Desert received abundant rain for a short period, producing flash floods that caused serious damage to homes and infrastructure. In the following days the humidity generated the germination and growth of various flowering plant species, covering vast areas of the desert. This phenomenon is called "Flowering Desert" and is activated by a winter rainfall larger than 15 mm, associated with the El Niño-Southern Oscillation (ENSO) (Gutiérrez *et al*., 2008). Several, otherwise latent, plant species appear shortly after the rain, and flower with great intensity.

In 1830, Claudio Gay carried out the first systematic study of plants in the Atacama Region, and since 1885 more specimens were studied by Rodulfo Philippi and Federico Phillipi (Muñoz-Schick *et al*., 2012). These studies have identified around 980 native plant species, of which 200 species are associated with the flowering desert (Gutiérrez *et al*., 2008, Squeo *et al*., 2008, MMA, 2018). Due to the high diversity of species, the areas where the flowering occurs were declared priority sites for the Regional Biodiversity Strategy of Atacama in 2002, and highly valued by both the scientific community and tourism (MMA, 2018).

There are around 100 families of angiosperms associated with the Atacama Region and, some of these grow only when the desert is flowering. On a tour of the Flowering Desert in 2017, 10 species were found in Pan de Azúcar and Llanos de Challe National Parks, corresponding to the families (with total number of endemic species between brackets): Fabaceae (38 species), Solanaceae (22 species), Alstroemeriaceae (13 species), Amaryllidaceae (12 species), Loasaceae (7 species), Passifloraceae (6 endemic species), Papavereaceae (4 species and one extinct), Lamiaceae (4 species), Onagraceae (2 species), and Aristolochiaceae (2 species) (Peña, 2002; MMA, 2008; Squeo *et al*., 2008). However, not all of these species have been properly classified taxonomically so far. Within the species that grow when the desert flowers, there are species that have differences in the color of petals, classifying them, in the case of the genus *Rhodophiala* (añañuca), as two species: *R. bagnoldii* (Herb.) Traub (yellow flower) and *R. phycelloides* (Herb.) Hunz. (red flower). However, there are other species with color variation of petals which still classify as one species. In the family Alstroemeriaceae, for example, the species *Alstroemeria kingii* Phil. has individuals with cream and yellow petals. Genetic analyses could clarify if differences in petal color separates these individuals into different species. Unfortunately, the genetics of most of the plant species of the flowering desert have been poorly studied, and more molecular studies are needed to classify the taxonomy and phylogeny of these species, on a solid base.

DNA barcoding is a standardized procedure and optimal tool for monitoring biodiversity and conducting phylogeny and evolution studies of plant species (Pei *et al*., 2017), that uses universal DNA sequences. The Consortium for the Barcode of Life (CBOL Plant Working Group, 2009) proposed two chloroplast loci (*matK* and *rbcL*) as DNA barcode of plants. The most important challenge in the DNA barcoding of plants is the discrimination between taxa or specific genera (Seberg & Petersen, 2009). It has been shown that barcodes can be successfully applied to study diversity (Kesanakurti *et al*., 2011), classify herbarium species (de Vere *et al*., 2012) and discriminate taxa in families with complex taxonomy (Sass *et al*., 2007). This means that poorly studied desert plant species and their biodiversity could be evaluated and classified with DNA barcoding. The objective of this work is to evaluate the universality of the plant barcodes (*matK* and *rbcL*), using the phylogenetic relationships of the species from the flowering desert, which form geographically isolated populations as a result of that same desert.

MATERIALS AND METHODS

Plant sampling

Fresh leaves were collected from species that emerged in the flowering desert, after an abundant rainfall generated in 2017. The samples were collected in the Llanos de Challe (PLC) and Pan de Azúcar (PPA) national parks, located in the Atacama Region (Figure No. 1). The samples were georeferenced (see Table No. 1) and the collected

material was stored at -80°C.

DNA Extraction and Sequencing

DNA was extracted from fresh leaves using a modified CTAB method described by Contreras *et al*. (2018a). When the previous method produced low-quality or contaminated DNA, another method of extraction, CTAB-phenol-coloform-isoamyl alcohol-column of silica, was used (Contreras *et al*., 2018b). The DNA samples were stored at -80°C in the DNA bank of the University of Atacama. Two chloroplast primers, namely *rbcL* and *matK*, were used for the polymerase chain reaction (PCR). A partial region of the *rbcL* gene was amplified with the following primers:

5'-ATGTCACCACCACAACAGAGAGTAAAGC-3 '(forward) and
5'-CTTCTGCTACAAATAAGAATCGATCTC-3' (reverse),

as described by Kress & Erickson (2007). Similarly, a partial region of the *matK* gene was amplified using

390F 5'-CGATATTCATTCATTCAATTTC-3 '(forward) and
1326R 5'-TCTAGCACACGAAAGTCGAAGGAAGT3' (reverse),

Figure No. 1
Map showing the collection sites Pan de Azúcar and Llanos de Challe National Parks in the Atacama Region

as described by Cuenoud *et al*. (2002). The PCR mixture used contained: 10 µl of master mixture (Takara SapphireAmp 2X, Clontech), 1.5 µl of each primer (advance and rewind to 5 µM), 6 µl of genomic DNA (5 ng/µl) and 1 µl of nuclease-free water. The PCRs were performed in a Swift Max Pro thermal cycler (ESCO) using the following protocol: initial step 94°C for 4 min, then 40 denaturation cycles at 94°C for 30 seconds, annealing at 52°C or 55°C (for *rbcL* and *matK*, respectively) for 30

seconds, and polymerization at 72°C for 1 min, as well as a final extension stage at 72°C for 10 min. The PCR products obtained were separated on a 1% (w/v) agarose gel in a TBE buffer (100 V for 40 min), dyed using GelRedTM 10,000X (10 µL), and purified with the "Wizard® SV Gel and PCR Clean-Up" kit (Promega). The purified PCR products were sequenced in an ABI 3730xl DNA analyzer (Applied Biosystems Inc.) according to the manufacturer's protocols. For the sequencing of the purified products, the Genescan Service of Macrogen Inc (Seoul, South Korea) was used.

Table 1
Family, species name, code, origin, geographic coordinates, DNA barcode (*rbcL* and *matK*) and corresponding GenBank accession numbers

Family	Species	Code	Origen (*)	Latitud	Longitud	*rbcL*	*matK*
Papaveraceae	*Argemone hunnemannii* Otto	Ah699	Endemic	28°07'03.4"S	71°05'55.9"W	-	MK077752
Onagraceae	*Oenothera coquimbensis* Gay	Oc691	Endemic	26°27'30.9"S	70°40'51.6"W	-	MK077750
Passifloraceae	*Malesherbia humilis* Poepp.	Mh682	Native	26°06'18.7"S	70°38'09.6"W	MK077754	-
Amaryllidaceae	*Leucocoryne appendiculata* F.Phil.	La680	Endemic	26°06'33.3"S	70°38'34.1"W	MK089789	-
*Loasa*ceae	*Loasa elongata* Hook. & Arn.	Le675	Endemic	26°12'03.9"S	70°38'03.3"W	MK077755	MK077751
Solanaceae	*Nicotiana solanifolia* Walp.	Ns659	Endemic	26°11'46.2"S	70°37'08.7"W	MK077752	MK077749
Lamiaceae	*Stachys grandidentata* Lindl.	Sg671	Endemic	26°12'06.6"S	70°36'58.4"W	MK124961	MK124962
		Sg673	Endemic	26°12'06.2"S	70°36'57.2"W		
		Sg709	Endemic	27°52'06.3"S	70°59'30.2"W		
*Aristolochia*ceae	*Aristolochia chilensis* Bridges ex Lindl.	Ac704	Endemic	28°09'47.9"S	71°03'26.0"W	-	MK032806
*Alstroemeria*ceae	*Alstroemeria kingii (am)* Phil.	Ak703	Endemic	28°09'27.7"S	71°03'19.6"W	MK124959	MK124957
	Alstroemeria kingii (cr) Phil.	Ak705	Endemic	28°09'53.6"S	71°03'24.5"W	MK124960	MK124958
		Ak708	Endemic	27°52'06.2"S	70°59'30.1"W	MK124960	MK124958
Fabaceae	*Adesmia eremophila* Phil.	Ae712	Endemic	27°55'15.5"S	70°56'33.7"W	MK077756	-

DNA Barcode Analysis

The sequences "Forward" and "Reverse" were created by the software Chromas Pro v1 (Technelysium Pty, Ltd). The sequence fragments, obtained with the forward and reverse primers, were assembled using DNA Baser Sequence Assembler v4.10 (Biosoft, 2012). Automatic adjustments and analyses were used to edit the collection of contiguous DNA sequences (contigs) with predetermined parameters. Multiple sequence alignments were performed using MEGA 6 software (Tamura *et al*., 2013). All sequences obtained in this study were entered into GenBank, the international nucleotide sequence database (see Table No. 1 for access numbers). In addition, all *rbcL* and *matK* sequences obtained were validated by transcribing nucleotide sequences into protein chains. The *rbcL* and *matK* sequences that coincided with species from the flowering desert were downloaded from GenBank. To achieve this, the first five sequences with the highest score were consulted in GenBank (BLASTn with the predetermined parameters of NCBI), using the *rbcL* and *matK* sequences of each of the species we collected. These sequences, both from our study and GenBank, were then used together for the phylogenetic analysis.

Maximum likelihood (ML) and Bayesian inference (BI) methods were used to evaluate the phylogeny of the studied species. ML analysis was inferred based on the Tamura-Nei model (Tamura & Nei, 1993) using MEGA6 software (Tamura *et al*., 2013) with 1000 bootstrap (BS) replicas used for the

reliability of the analysis support. Bayesian phylogenetic inference analysis was performed with MrBayes 3.2 (Ronquist *et al.*, 2012). Two independent analyses of Monte Carlo Markov Chains (MCMC) were performed, each consisting of 1,000,000 generations, with an average frequency deviation of less than 0.01. Trees were sampled at a frequency of 200 generations, and 10% of the initial trees were discarded as burn-in. TRACER v. 1.5 (Rambaut & Drummond, 2007) was used to verify the stabilization of the overall probability and convergence between generations for each separate file. The reliability of clades in the Bayesian analysis was evaluated by means of posterior probability (PP). Values of PP lower than 0.8 were considered low support; between 0.8 and 0.89 were considered moderate support; and values of PP higher than 0.9 were considered high support. In the analysis, bootstrap (BS) values for internal nodes were estimated with 100 replicas, defining values greater than 70 to be reliable. Finally, the tree was visualized with the FigTree program (Rambaut, 2012).

RESULTS

A successful amplification and partial DNA sequence of most of the analyzed species was obtained with the *rbcL* and *matK* markers. However, when using only the *rbcL* marker, *Argemone hunnemannii* Otto, *Oenothera coquimbensis* Gay and *Aristolochia chilensis* Bridges ex Lindl., did not amplify as the sequence-end could not be determined. Likewise, with the *matK* gene promoters it was not possible to obtain PCR fragments nor DNA sequences of *Malesherbia humilis* Poepp, *Leucocoryne appendiculata* Phil. and *Adesmia eremophila* Phil.

In the following section we present the phylogenetic relationships based on partial sequences of the *rbcL* gene and *matK* of the flowering species that spontaneously germinated during the flowering desert event in 2017 (Atacama Region, Chile).

Phylogenetic relationships based on rbcL

Seven main clades were observed in the dendrograms in Figures No. 2 and No. 3, when we compared the different topology of the maximum likelihood (ML) and Bayesian inference (BI) methods, which grouped the individuals in the families: Fabaceae, Alstroemeriaceae, Amaryllidaceae, Loasaceae, Lamiaceae, Solanaceae and Passifloraceae.

All these families formed clades with reliable and high supporting values (BS > 81% and PP > 0.99), with the exception of the Fabaceae family clade (BS=51%, PP=0.76). The BI tree showed a more determined topology than the ML tree among the different species and families; therefore, the BI tree was explored and analyzed in greater detail. The Fabaceae family formed two groups, in one of these, *Adesmia eremophila* was grouped with *A. exilis* Clos (U74254.1), differing in four nucleotide replacements. In the Passifloraceae family, *Malesherbia humilis* formed a clade with *M. lirana* Gay (AB536557.1), with reliable and high support (BS=96%, PP=0.97), separating itself from the rest of species of the genus *Malesherbia*. However, our sample of *Malesherbia humilis* was not grouped (BS=100%, PP=1.00) with the known sequence of *M. humilis* (AB536556.1), differing in four nucleotide replacements. *Nicotiana solanifolia* Walp. was grouped with four species of the genus *Nicotiana (N. attenuata* Steud., *N. tabacum* L., *N. attenuata*, *N. sylvestris* Speg. & S. Comes), in the family Solanaceae, with high support (BS=100%, PP=1,00), differing from the other species in two nucleotide replacements. In the Lamiaceae family, three samples of *Stachys grandidentata* Lindl. collected in different areas formed a single clade, reliable and with high support (BS=77%, PP=0.97), separating itself from the rest of the species of the genus *Stachys*. *S. grandidentata* differed in only one nucleotide substitution with *S. chamissonis* Benth. (KU724138.1) and *S. lindenii* Benth. (AF502013.1), while it differed in three nucleotide replacements from the other species *S. sylvatica* L. and *S. coccinea* Ortega. The species from the *Loasa*ceae family formed two clades with reliable and high support (BS=81%, PP=0.99), one with the genus *Fendlera*, and one with the genera *Loasa* and *Nasa*, in which *Loasa elongata* Hook. & Arn. formed a clade with *Nasa loxensis* (Kunth) Weigend (U17876.1) with reliable and high support (BS=81%, PP=0.98), differing in the replacement of nine nucleotides. The species of the family Alstroemeriaceae and Amaryllidaceae were separated into two clades with maximum reliability and high support (BS=100%, PP=1,00). In the family Amaryllidaceae, *Leucocoryne appendiculata* formed a subgroup with *L. coquimbensis* F.Phil. ex Phil. (Z69199.1) and *Nothoscordum gracile* (Aiton) Stearn (HQ690454.1) with low reliability but high support (BS=42%, PP=0.98), however, *L. appendiculata* was closer to *L.*

coquimbensis (differing in the substitution of one nucleotide) than with the species of the other genus (differing in two nucleotides). In the family Alstroemeriaceae, three samples of *Alstroemeria kingii* collected in different zones and of different petals color formed a single clade, with reliability and high support (BS=92%, PP=0,92), separating them from the rest of species of the genus *Alstroemeria*, however, with strong support (BS=100%, PP=1,00) the known sequence *A. kingii* (JQ404678.1) did not nest into the clade of our *Alstroemeria kingii* samples, differing in two nucleotide replacements.

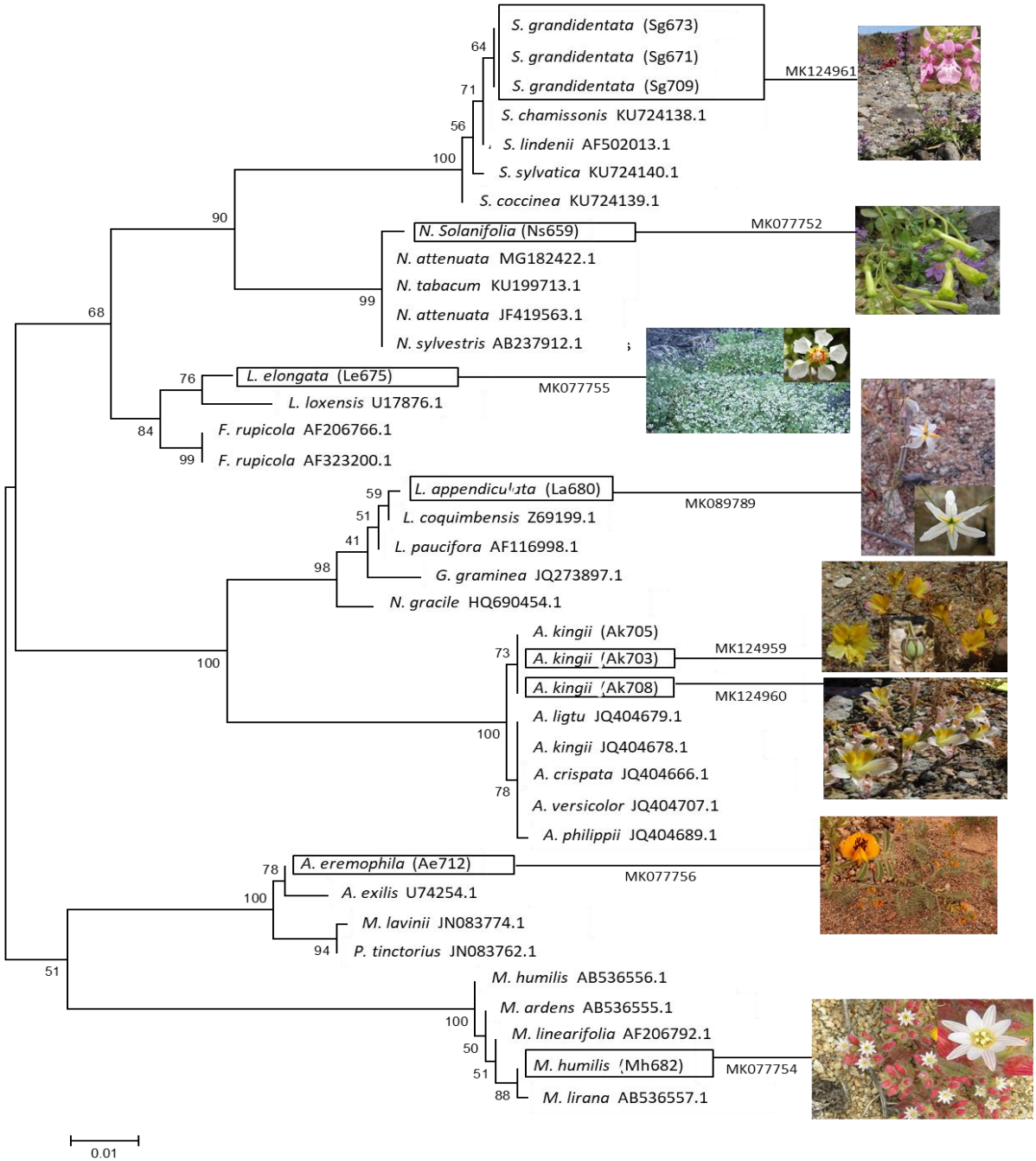

Figure No. 2
Maximum likelihood (ML) tree with bootstrap (BS) values of *rbcL* sequences of seven species from the Atacama Region, using the Tamura-Nei evolution model

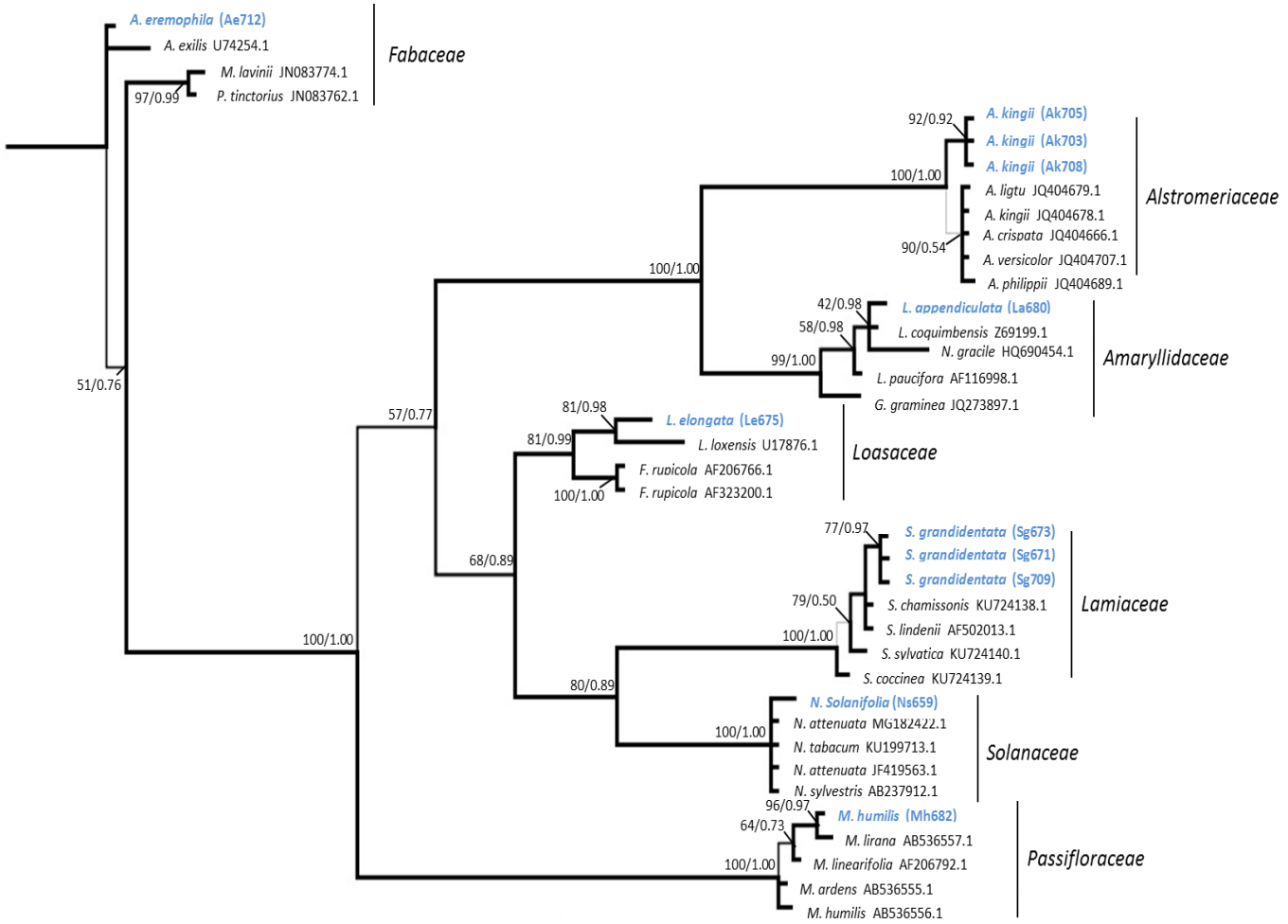

Figure No. 3
Bayesian inference tree (BI) inferred from the *rbcL* data set of seven species in the Atacama Region. Bootstrap (BS) values are presented on the left, and posterior probabilities (PP) are shown on the right

Phylogenetic relationships based on matK

The dendrograms in Figures No. 4 and No. 5 show seven main clades as well, when we compared the different topology of maximum likelihood (ML) and Bayesian inference (IB), and group in the families: Solanaceae, Alstroemeriaceae, Lamiaceae, Aristolochiaceae, Papaveraceae, Onograceae and Loasaceae. All these families formed clades with maximum reliability and high support (BS=100% and PP=0.99). Between the two phylogenetic trees (ML and BI), the BI tree showed a stronger topology among the different species and families, therefore, we explored this tree further. In the Solanaceae family four groups were formed, and in one of them, the *Nicotiana solanifolia* sample formed a clade with already known *N. solanifolia* (AB039990.1), differing in three nucleotide substitutions. The three samples of *Stachys grandidentata* (Lamiaceae) collected in different areas formed a single clade with the other species of the genus *Stachys*, with a robust support (BS=100%, PP=1,00). However, differences in more than two nucleotide replacements were observed with the species S. *coccinea* (KU724139.1) and *S. sylvatica* (KU724140.1). The samples of *Alstroemeria kingii*, from the Alstroemeriaceae family, collected in different areas and with different petal color formed clades with high reliability and support (BS=99%, PP=0.91). Due to differences in more than five nucleotide replacement they separated themselves from the other clade that includes the rest

of species of the genus *Alstroemeria*, with maximum reliability and support (BS=100%, PP=1.00). The Aristolochiaceae family was divided into two clades with maximum reliability and support (BS=100%, PP=1.00); in one of these clades, the *Aristolochia chilensis* sample was grouped with *A. anguicida* Jacq. (KP998777.1), however, with low reliability and low support (BS=49%, PP=0.51), differing with seven nucleotide replacements. Four species of the genus *Argemone* (Papaveraceae) were grouped together and

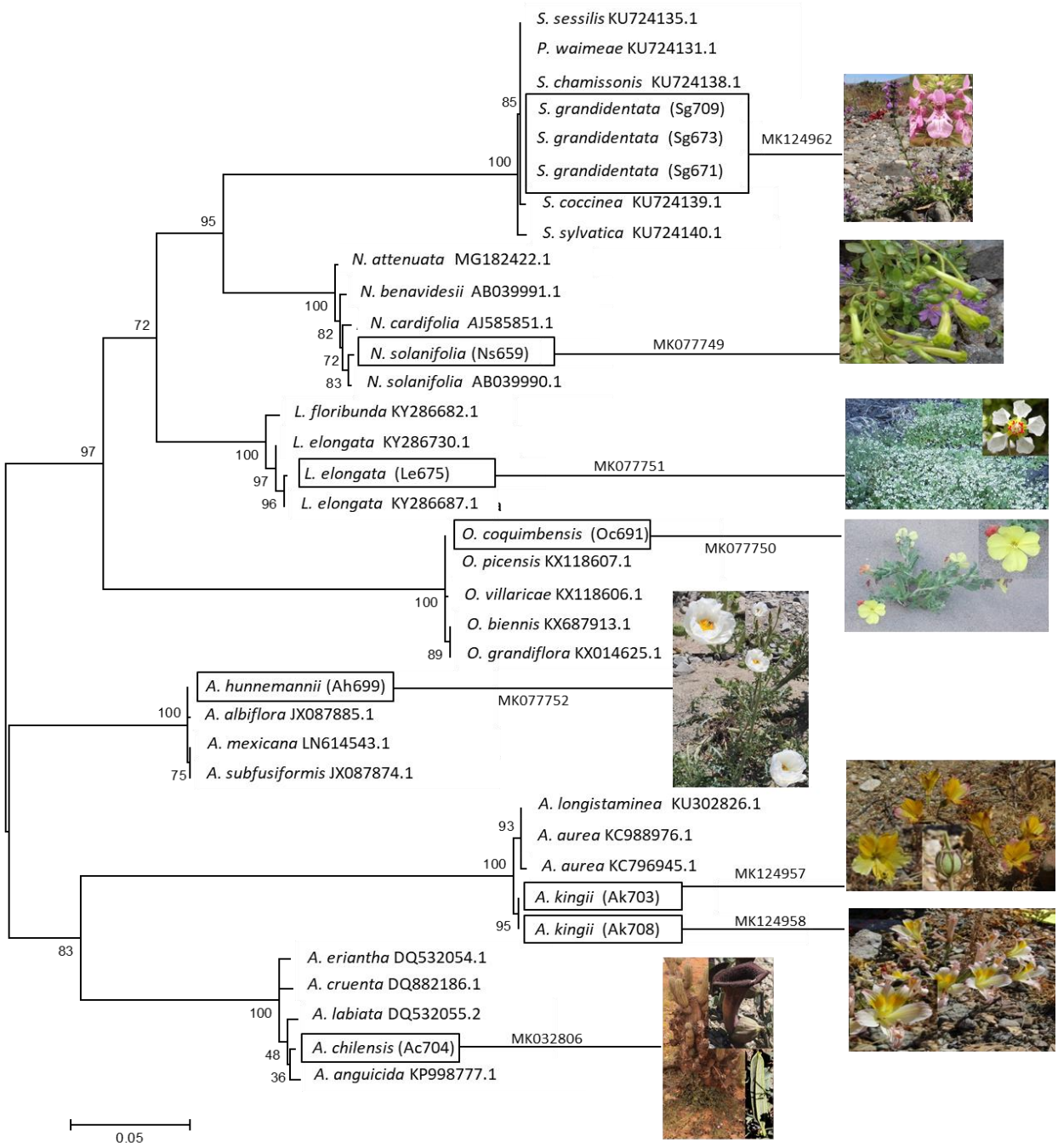

Figure No. 4
Maximum likelihood (ML) tree with bootstrap (BS) values of *matK* sequences of seven species from the Atacama Region, using the Tamura-Nei evolution model

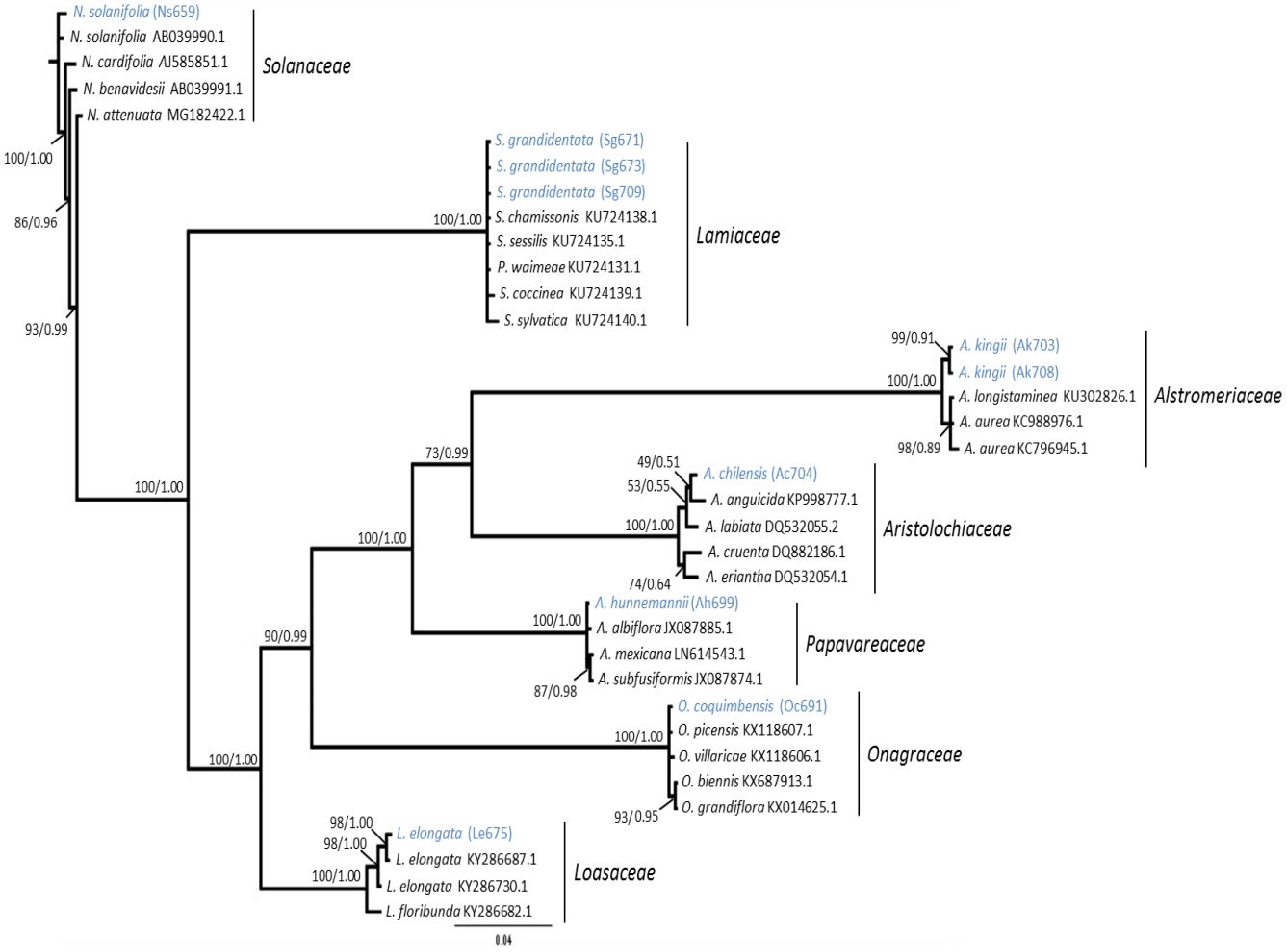

Figure No. 5
Bayesian inference tree (BI) obtained from the *matK* data set of seven species from the Atacama Region. Bootstrap values (BS) are presented on the left, and posterior probabilities (PP) are shown on the right

two formed a subclade, locating *A. hunnemannii* close to *A. albiflora* Hornem. (with the substitution of a nucleotide), while the other two species (*A. mexicana* L. and *A. subfusiformis* Ownbey) that formed the other subclade had reliable and high support (BS=87%, PP=0,98). In the Onagraceae family, five species of the *Oenothera* genus formed clades, where two of these (*Oenothera biennis* L. and *O. biennis f. grandiflora* (L`Her) D.S. Carp.) formed a subclade, with high reliability and support (BS=93%, PP=0.95), and *O. coquimbensis* ended up near *O. picensis* Phil. and *O. villaricae* W. Dietr. The *Loasa elongata* sample (Loasaceae) formed a reliable subclade with *L. elongata* (KY286687.1) with strong support (BS=98%, PP=1,00), differing in one nucleotide substitution; however, these two sequences were separated from a third *L. elongata* sequence (KY286730.1), also with high reliability and high support (BS=98%, PP=1,00).

DISCUSSION

Chile is one of the 35 hotspots of biodiversity, making its conservation a priority in the world (Mittermeier *et al.*, 2011). The hyper arid environment of the Atacama Desert harbours areas where the flowering desert occurs (Arroyo *et al.*, 2004), that have a great diversity of plants, and a high number of endemic plants. The species from this area have been identified taxonomically with morphological keys for 200 years. However, their

identification and/or validation through genetic tools has been studied little. Therefore, DNA barcoding is a good method to identify or validate these species and confirm their taxonomy.

Our results showed that 70% of the species we studied can be distinguished with a barcode locus (*matK* or *rbcL*), and 100% can be determinated using both *matK* and *rbcL* genes. According to Bolson *et al*. (2015), the most efficient way to determinate a species from a DNA barcode is to check the grouping of species, genus and family using phylogenetic analyses.

We constructed phylogenetic trees of plant species that emerge when the desert flowers, in the national parks Llanos de Challe and Pan de Azúcar, based on the DNA barcoding techniques using the *rbcL* and *matK* genes. Using the *rbcL* marker showed that the species *Adesmia eremophila* and *A. exilis* were grouped according to their corresponding genus *Adesmia*, separating themselves from other genera of the same family. However, between *A. eremophila* and *A. exilis*, we observed differences in nucleotides, probably due to their biogeographic evolutionary differences. *Malesherbia humilis* grouped with good support with *M. lirana* Gay, using the marker *rbcL*, even though *M. lirana* is endemic to La Rioja, Argentina (Darwinion Botanical Institute, 2019), separating the species by more than 500 km and the Andes mountain range. The evolution of *M. humilis* has possibly been constrained, due to the barrier imposed by the Atacama Desert. Curiously *M. humilis* did not grouped with the AB536556.1 GenBank record of *M. humilis*. However, this entry does not indicate the location where it was collected, making comparison between the samples difficult. It is possible that the difference between these two sequences of *M. humilis* is due to the fact that they belong to different varieties, as *M humilis* forms a complex of five varieties (Gengler-Nowak, 2002). Therefore, a thorough review of this species in the area should be carried out.

Nicotiana solanifolia grouped with other species of the same genus, both with the *rbcL* and marK genes; however, only the *matK* gene differentiated the species with high support. According to Clarkson *et al*. (2004) the species of the genus *Nicotiana* of the Paniculatae section to which *N. solanifolia* belongs are difficult to distinguish, both vegetatively and floreally. However, the sequence of the *matK* gene from the individual collected in Pan de Azúcar national park was a good indicator to successfully group, with good support, with the record described by the same author.

As for *Stachys grandidentata*, the sequences of individuals from the two different national parks presented no differences in nucleotides, in spite of the 400 km of separation. Although this species and other records of the same genus were grouped together with the genes *matK* and *rbcL*, with the latter gene, high nucleotide differences were observed when comparing *S. grandidentata* with *S. sylvatica* (Iflora, 2019) and *S. coccinea* (Berumen-Cornejo *et al*., 2017). But those species have even larger geographical differences; the latter two being endemic to Europe and North America, respectively. In the case of *Loasa elongata*, this species was successfully grouped with the same species (KY286687.1) with the *matK* gene, and differentiated from *N. loxensis*, with the *rbcL* gene. Even though, *L. elongata* and *N. loxensis*, form a single group, they present large nucleotide differences, perhaps because *N. loxensis* is endemic to Ecuador (Hempel *et al*., 1995). The species of the family Amaryllidaceae showed great evolutionary nearness between *Leucocoryne appendiculata* and *L. coquimbensis*, probably because they are distributed in the same habitat and both species are limited to the same geographical region. However, *L. appendiculata* is not endemic to the Atacama Region only, as its distribution range includes all of "big north" of Chile (Antofagasta, Tarapacá Region and Arica-Parinacota, and part of Atacama Region) (Pinto & Luebert, 2009).

On the other hand, in the family Alstroemeriaceae the sequences of two variants of *Alstroemeria kingii* (cream) and *Alstroemeria kingii* (yellow) formed a separate clade from other species of the same genus, both with the *matK* and *rbcL* genes. However, no differences were found between their two variants, with either of the genes. There are cases of petals color combination; for example, in species of *Lactucinae* and *Crepidinae* that could have been the result of a possible hybridization (Wang *et al*., 2014). According to Xiao & Simpson (2017), species of the genus *Meconopsis* of the Aculeatae and Primulinae section have petals in more than one color, and these differences were confirmed by phylogenetic analysis with *matK*, *ndhF*, *rbcL* and *trnL-trnF* markers. Using other barcode DNA markers, could give a more definite result of whether

there are differences between the cream and yellow variants of *A. kingii*.

Aristolochia chilensis formed a weakly supported group with two species of the same genus, However, it differed with good support from other species such as *A. cruenta* Barringer, endemic to Costa Rica (Barringer, 1983) and *A. eriantha* Mart., endemic to Brazil (González, 2011). Nevertheless, the lack of expertise in the classification of this taxon could lead to identification errors, making it possible to misclassify individuals as *Aristolochia bridgesii* (Klotzsch) Duch., as the latter has its distribution range in the same sampling area of *A. chilensis* (Saldivia & Faúndez, 2015). The inconsistencies in the identification were demonstrated by Saldivia & Faúndez (2015), whose morphological comparisons made it possible to differentiated between three species (*A. bridgesii*, *Aristolochia pearcei* Phill. and *A. chilensis*) of this genus. There is no doubt that in future studies these three species could be confirmed using the DNA barcoding technique.

Within the Papavareaceae family, the species *Argemone hunnemannii* was closer related to *A. albiflora* than to other species of the same genus, even though, the latter is endemic to North America (Schwarzbach & Kadereit, 1999). Despite the geographical distance, nucleotide differences were very low. *Oenothera coquimbensis*, present in a large part of the "Big North" and central Chile, is phylogenetically very closely related to *O. picensis* (central Chile, González *et al*., 2011). Their close relationship is not surprising, as they share the same geographical range. However, *O. coquimbensis* was also close phylogenetically to *O. villaricae*, which has its distribution in central and southern Chile, as well as southern Argentina (Plants of the world online, 2019). Therefore, we suspect that the *matK* gene was probably conserved in the evolution of *Oenothera* species despite the distance and diversity of climates they inhabit.

The sequences obtained in this study and sequences downloaded from GenBank of the same species did not always form a glade. For example, the *rbcL* sequence of species *A. kingii* of the record JQ404678.1 was not grouped together with the sequences we obtained from the same species; likewise, the *rbcL* sequence of *M. humilis* (record AB536556.1) was not grouped together with the sample of *M. humilis* from this study. However, a common denominator might justify these inconsistencies. According to de Vere *et al*. (2012) the use of herbarium specimens for DNA sequence analysis presents disadvantages, when comparing their results to fresh material. The sequences of records JQ404678.1 and AB536556.1 were indeed obtained from herbarium material, while the sequences of this study were obtained from fresh samples. According to de Vere *et al*. (2012), herbarium samples often require more amplification attempts with more combinations of primers. This potentially increases the chance of obtaining incorrect sequences, through increased chances of sample mixing or contamination. De Vere *et al*. (2012) showed that for *matK* this is not the case, but, for *rbcL* the levels of incorrect sequences were higher using herbarium material than freshly collected samples. This is probably due to the greater universality of the primer and the ease of amplification. In addition, some studies suggest using the *matK* region as a tool for phylogenetic analysis because of this greater ease of sequence retrieval (Ipek *et al*., 2014).

On the other hand, as indicated at the beginning, some species amplified weakly with the *matK* and *rbcL* markers. Therefore, other barcoding markers such as *ITS* (Schultz & Wolf, 2009), *ycf* (Dong *et al*., 2015) and *trnH-psbA* (Kress *et al*., 2005) were tested in order to strengthen the information and the statistical analyses. However, we did not include these markers in the final analyses because they showed multiple copies of PCR fragments, inconsistencies in sequence alignment and additionally, with *ITS*, some species returned DNA sequences from bacteria or fungi. Thus, in order to identify species from the flowering desert, the DNA sequences obtained from the *matK* and *rbcL* markers were more successful.

The current information of barcode DNA sequences, of flowering desert species of Atacama, obtained from the present work and their phylogenetic differences with other plant species will contribute to the acknowledgement of these specimens. It is the first time that *rbcL* sequences of *A. eremophila* (MK077756), *L. appendiculata* (MK089789), *L. elongata* (MK077755), *S. grandidentata* (MK124961) and *N. solanifolia* (MK077752), and *matK* sequences of *S. grandidentata* (MK124962), *A. kingii* (MK124957 and MK124958), *A. chilensis* (MK032806), *A. hunnemannii* (MK077752) and *O. coquimbensis*

(MK077750) are available.

Our results, both with BI and ML analyses, showed that the flowering desert species formed glades, with high support values, based on genus and family when we used the regions *rbcL* and *matK*. However, in other papers the potential for species discrimination has been variable; for example, in neotropical species some authors report discrimination ranges between 75 and 95% using the barcode markers *rbcL*, *matK* and *trnH-psbA* (Kress *et al*., 2009), while other papers failed to discriminate between species, from, e.g. families such as Lauraceae, Myrtaceae and Sapindaceae, even when combining up to three barcoding loci, including *matK* and *rbcL* (Costion *et al*., 2011). Our results demonstrate that these two genes are highly efficient at reconstructing phylogenetic relationships for plant species from the flowering desert, and that the method could facilitate accurate identification of species when their taxonomy is not certain. Without a doubt, the results obtained in this study, support the effectiveness of barcoding to identify the taxonomic richness of the species in the flowering Atacama Desert.

ACKNOWLEDGEMENTS

This research was financed by the Regional Innovation Fund for Regional Competitiveness (FIC Regional, 2018) of the Regional Government of Atacama, Code BIP 40013338-0. In addition, we thank the National Forestry Corporation (CONAF) of the Region of Atacama for the authorization to sample plant species of the Flowering Desert (Permit N° 106/2017).

REFERENCES

Arroyo MTK, Marquet PA, Marticorena C, Simonetti JA, Cavieres L, Squeo F, Rozzi TR. 2004. **Chilean winter rainfall-Valdivian forests**. *In*: Mittermeier RA, Gil PR, Hoffmann M, Pilgrim J, Brooks T, Mittermeier CG, Lamoreux J, da Fonseca GAB, eds. Hotspots Revisted: Earth's Biologically Wealthiest and most Threatened Ecosystems, CEMEX, México DF, Mexico.

Barringer K. 1983. Notes on Central American *Aristolochia*ceae. **Brittonia** 35: 171 - 174. **https://doi.org/10.2307/2805959**

Berumen-Cornejo AM, Charlotte L, Pérez EM, Siqueiros ME. 2017. Phylogeny of the *Stachys coccinea* (Lamiaceae) complex based on molecular and morphological data. **Syst Bot** 42: 484 - 493. **https://doi.org/10.1600/036364417x696113**

Biosoft H. 2012. **DNA Baser Sequence Assembler**, v4.10. Heracle BioSoft SRL Romania. **http://www.dnabaser.com/download/download.html**

Bolson M, Smidt EC, Brotto ML, Silva- Pereira V. 2015. ITS and trnH- psbA as efficient DNA barcodes to identify threatened commercial woody angiosperms from Southern Brazilian Atlantic rainforests. **Plos One** 10: e0143049. **https://doi.org/10.1371/journal.pone.0143049**

CBOL Plant Working Group. 2009. A DNA barcode for land plants. **Proc Nat Acad Sci USA** 106: 12794 - 12797.

Clarkson JJ, Knapp S, Aoki S, Garcia VF, Olmstead RG, Chase MW. 2004. Phylogenetic relationships in *Nicotiana* (Solanaceae) inferred from multiple plastid DNA regions. **Mol Phylogen Evol** 33: 75 - 90. **https://doi.org/10.1016/j.ympev.2004.05.002**

Contreras R, Aguayo F, Guerra A, Tapia F, Porcile V. 2018a. Genetic characterization of centennial olive trees from northern Chile: The case of extra virgin olive oil from Huasco in the process of designation of origin. **Chil J Agric Anim Sci** 34:126 - 139. **https://doi.org/10.4067/s0719-38902018005000402**

Contreras R, Porcile V, Aguayo F. 2018b. Genetic diversity of Geoffroea decorticans, a native woody leguminous species from Atacama Desert in Chile. **Bosque** 39: 321 - 332. **https://doi.org/10.4067/s0717-92002018000200321**

Costion C, Ford A, Cross H, Crayn D, Harrington M, Lowe A. 2011. Plant DNA barcodes can accurately estimate species richness in poorly known floras. **Plos One** 6: e26841. **https://doi.org/10.1371/journal.pone.0026841**

Cuenoud P, Savolainen V, Chatrou LW, Powell M, Grayer RJ, Chase MW. 2002. Molecular phylogenetics of caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, atpb, and *matK* DNA sequences. **Am J Bot** 89: 132 - 44. **https://doi.org/10.3732/ajb.89.1.132**

de Vere N, Rich TCG, Ford CL, Trinder SA, Long C, Moore CW, Satterthwaite D, Davies H, Allainguillaume J,

Ronca S, Tatarinova T, Garbett H, Walker K, Wilkinson MJ. 2012. DNA barcoding the native flowering plants and conifers of Wales. **Plos One** 7: 1 - 12. **https://doi.org/10.1371/journal.pone.0037945**

Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, Cheng T, Guo J, Zhou S. 2015. ycf1, the most promising plastid DNA barcode of land plants. **Scientific Rep** 5: 8348. **https://doi.org/10.1038/srep08348**

Gengler-Nowak K. 2002. Phenetic analyses of morphological traits in the *Malesherbia humilis* complex (*Malesherbia*ceae). **Taxon** 51: 281 - 293. **https://doi.org/10.2307/1554898**

González F. 2011. A new pseudostipule-bearing species of *Aristolochia* (*Aristolochia*ceae) from Bahía and Espíritu Santo, Brazil. **Brittonia** 63: 430 - 435. **https://doi.org/10.1007/s12228-011-9184-2**

González I, Cortes A, Neaman A, Rubio P. 2011. Biodegradable chelate enhances the phytoextraction of copper by *Oenothera picensis* grown in copper-contaminated acid soils. **Chemosphere** 84: 490 - 496. **https://doi.org/10.1016/j.chemosphere.2011.03.015**

Gutiérrez JR. 2008. **El Desierto Florido de la Región de Atacama**. En: Squeo FA, Arancio G, Gutiérrez JR, eds. Libro Rojo de la Flora Nativa y de los Sitios Prioritarios para su conservación: Región de Atacama. Ediciones Universidad de La Serena, La Serena, Chile.

Hempel AL, Reeves PA, Olmstead RG, Jansen RK. 1995. Implications of *rbcL* sequence data for higher order relationships of *Loasa*ceae and the anomalous aquatic plant Hydro*Stachys* (Hydrostachyaceae). **Plant Syst Evol** 194: 25 - 37. **https://doi.org/10.1007/bf00983214**

Instituto de Botánica Darwinion. 2019. ***Malesherbia lirana* Gay**. CONICET, Argentina. **http://www2.darwin.edu.ar/Proyectos/FloraArgentina/DetalleEspecie.asp?forma=&variedad=lirana&subespecie=&especie=lirana&genero=malesherbia&espcod=28479**

iFlora. 2019. ***Stachys sylvatica*** L. **http://www.i-flora.com/en/fact-sheets/phylogenetic-tree/art/show/*Stachys*-sylvatica.html**

Ipek M, Ipek A, Simon PW. 2014. Testing the utility of *matK* and ITS DNA regions for discrimination of Allium species. **Turkish J Bot** 38: 203 - 212. **https://doi.org/10.3906/bot-1308-46**

Kesanakurti PR, Fazekas AJ, Burguess KS, Percy DM, Newmaster SG, Graham SW, Barret SC, Hajibabaei M, Husband BC. 2011. Spatial patterns of plant diversity below-ground as revealed by DNA barcoding. **Mol Ecol** 20: 1289 - 1302. **https://doi.org/10.1111/j.1365-294x.2010.04989.x**

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. 2005. Use of DNA barcodes to identify flowering plants. **Proc Nat Acad Sci USA** 102: 8369 - 8374. **https://doi.org/10.1073/pnas.0503123102**

Kress WJ, Erickson DL. 2007. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding trnH-psbA spacer region. **Plos One** 2: e508. **https://doi.org/10.1371/journal.pone.0000508**

Kress WJ, Erickson DL, Jones FA, Swenson NG, Perez R, Sanjur O, Bermingham E. 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. **Proc Nat Acad Sci USA** 106: 18621 - 18626. **https://doi.org/10.1073/pnas.0909820106**

Mittermeier RA, Turner WR, Larsen FW, Brooks TM, Gascon C. 2011. **Global biodiversity conservation: the critical role of hotspots.** In: Zach FE, Habel JC, eds. Biodiversity hotspots: distribution and protection of conservation priority areas. Berlin, Heidelberg, Springer, Germany. **https://doi.org/10.1007/978-3-642-20992-5_1**

MMA. 2018. [Ministerio de Medio Ambiente] **Informe de localización espacial y caracterización de núcleos de biodiversidad en el sitio prioritario Desierto Florido, Región de Atacama**. Ministerio de Medio Ambiente, Chile**. http://brdnap.mma.gob.cl/recursos/SINIA/Biblio%20SP-6470.pdf**

Muñoz-Schick M, Morales V, Moreira-Muñoz A. 2012. La colección de tipos de plantas vasculares del Herbario Nacional de Chile (SGO): Análisis histórico, temporal y espacial. **Gayana Bot** 69: 123 - 135. **https://doi.org/10.4067/s0717-66432012000100012**

Pei N, Chen B, Kress WJ. 2017. Advances of community-level plant DNA barcoding in China. **Front Plant Sci** 8: 225. **https://doi.org/10.3389/fpls.2017.00225**

Peña R. 2002. El género Argemone (Papaveraceae) en Chile. **Chloris Chilensis** 5 (2): **http://www.chlorischile.cl**

Pinto R, Luebert F. 2009. Datos sobre la flora vascular del desierto costero de Arica y Tarapacá, Chile, y sus relaciones fitogeográficas con el sur de Perú. **Gayana Bot** 66: 28 - 49.

https://doi.org/10.4067/s0717-66432009000100004

Plant of the World online. 2019. ***Oeonothera villaricae*** **W. Dietr**. Kewscience. **http://www.plantsoftheworldonline.org/taxon/urn:lsid:ipni.org:names:173524-2**

Rambaut A, Drummond AJ. 2007. **Tracer v1.5**. **https://github.com/evolvedmicrobe/beast-mcmc/tree/master/release_tracer/Windows/Tracer%20v1.5**

Rambaut A. 2012. **FigTree: Tree Figure Drawing Tool Version 1.4**. University of Edinburgh, Edinburgh, Scotland. **http://tree.bio.ed.ac.uk/software/figtree**

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. **Syst Biol** 61: 539 - 542. **https://doi.org/10.1093/sysbio/sys029**

Saldivia P, Faúndez L. 2015. Revalidation of *Aristolochia pearcei* Phil. and notes on the identity of *A. bridgesii* (Klotzsch) Duch. (*Aristolochia*ceae). **Gayana Bot** 72: 367 - 372. **https://doi.org/10.4067/s0717-66432015000200016**

Sass C, Little DP, Stevenson DW, Specht CD. 2007. DNA Barcoding in the Cycadales: Testing the potential of proposed barcoding markers for species identification of cycads. **Plos One** 2: e1154. **https://doi.org/10.1371/journal.pone.0001154**

Schultz J, Wolf M. 2009. ITS2 sequence-structure analysis in phylogenetics: a how-to manual for molecular systematics. **Mol Phylog Evol** 52: 520 - 523. **https://doi.org/10.1016/j.ympev.2009.01.008**

Schwarzbach AE, Kadereit JW. 1999. Phylogeny of prickly poppies, *Argemone* (Papaveraceae), and the evolution of morphological and alkaloid characters based on ITS nrDNA sequence variation. **Plant Syst Evol** 218: 257 - 279. **https://doi.org/10.1007/bf01089231**

Seberg O, Petersen G. 2009. How many loci does it take to DNA barcode a crocus? **Plos One** 4: e4598. **https://doi.org/10.1371/journal.pone.0004598**

Squeo F, Arancio G, Gutiérrez J, Letelier L, Arroyo MK, León P, Retería L. 2008. **Flora amenazada de la Región de Atacama y estrategias para su conservación**. Eds. Universidad de La Serena, La Serena, Chile.

Sun T, Bao H, Reich M, Hemming, SR. 2018. More than ten million years of hyper-aridity recorded in the Atacama Gravels. **Geochim Cosmochim Acta** 227: 123 - 132. **https://doi.org/10.1016/j.gca.2018.02.021**

Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. **Mol Biol Evol** 10: 512 - 526. **https://doi.org/10.1093/oxfordjournals.molbev.a040023**

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. **Mol Biol Evol** 30: 2725 - 2729. **https://doi.org/10.1093/molbev/mst197**

Wang GY, Meng Y, Deng T, Yang YP. 2014. Molecular phylogeny of Faberia (Asteraceae: Cichorieae) based on nuclear and chloroplast sequences. **Phytotaxa** 167: 223 - 234. **https://doi.org/10.11646/phytotaxa.167.3.1**

Xiao W, Simpson BB. 2017. A New Infrageneric Classification of Meconopsis (Papaveraceae) Based on a Well-supported Molecular Phylogeny. **Syst Bot** 42: 226 - 233. **https://doi.org/10.1600/036364417x695466**