Algarrobilla (*Balsamocarpon brevifolium* Clos, Fabaceae) is a 1 to 2 m tall woody shrub endemic to Chile and with a restricted distribution between the south side of the Copiapó River (27°30′S) and north of the Elqui River (30°S), and from 300 to 2500 m elevation in the Andes Range. *Balsamocarpon brevifolium* is a conspicuous component of the arid region, particularly in the ecosystem of evergreen shrubs with microphyllous leaves that dominates the southern limits of the Atacama Desert of Chile (Luebert and Pliscoff, 2006). The arid climate is characterized by high temperatures and high solar radiation during the dry season in summer and scarce and scattered precipitation during winter (Aceituno et al., 2003). Every five to seven years, however, the El Niño Southern Oscillation carries intense precipitation to areas where *B. brevifolium* is present (Montecinos and Aceituno, 2003).

*Balsamocarpon brevifolium* frequently forms discontinuous populations, with small populations susceptible to alteration by human activities, and the species is under threat because of extensive charcoal production (Squeo et al., 2001, 2008; Estévez et al., 2010). In the past century, its yellow pods, which contain high tannin concentrations, were extensively harvested for use in the tanning industry (Wrann and Barros, 1987). Taking into account that ecological concentrations, were extensively harvested for use in the tanning industry (Montecinos and Aceituno, 2003).

Every five to seven years, however, the El Niño Southern Oscillation carries intense precipitation to areas where *B. brevifolium* is present (Montecinos and Aceituno, 2003).

Molecular markers are optimal tools for identification of species and study of their population genetics, and may provide less laborious methods for a more accurate delimitation of species than phenotypic evaluations (Arif et al., 2010). Several types of molecular markers have been developed, including random-amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), DNA barcoding, and single-nucleotide polymorphism (SNP) markers. However, simple sequence repeats (SSRs or microsatellites) remain a powerful molecular marker because of their ubiquity in plant genomes, relative simplicity for analysis, high levels of intraspecific polymorphism, and the potential of amplifying in related species (Queller et al., 1993). Today, through massive sequencing of genomic DNA, it is possible to obtain a large number of these markers to be used in the characterization of species complexes. In this paper, we present the first SSR markers developed for *B. brevifolium* and demonstrate their cross-amplification in related species, with the final aim of evaluating the genetic diversity of *B. brevifolium* through the study of the different populations of this endemic plant species of the Atacama Desert in Chile.

### METHODS AND RESULTS

Fifteen new microsatellite markers were developed for *B. brevifolium* and used to analyze three populations from the Atacama and Coquimbo regions in Chile. Microsatellites were highly polymorphic, with an average of 5.77 alleles per marker and an average level of expected heterozygosity of 0.72. These markers were evaluated and cross-amplified on two related species (*Senna cumingii* and *Caesalpinia angulata*) with partial success.

The development of this set of markers permits an extensive study of *B. brevifolium* populations for conservation purposes.

**KEY WORDS** *Balsamocarpon brevifolium; Caesalpinioideae; conservation genetics; Fabaceae; molecular markers; sustainable management.*
13 plants were available for this population because of the poor preservation conditions of plant material]. The populations were selected based on the distinct geographic distribution of the species. Representative samples of each population were deposited at the Herbarium of La Serena University (ULS), La Serena, Chile (Appendix 1). Heavy deforestation and changing climatic conditions have decimated most populations of this species, making it difficult to obtain a larger number of samples per population. Plant material was also obtained from individuals of *Senna cumingii* (Hook. & Arn.) H. S. Irwin & Barneby (voucher ULS 4693) and *Caesalpinia angulata* (Hook. & Arn.) Baill. (ULS 4708) (Appendix 1), species that are related and cohabit with *B. brevifolium* (Luebert and Pliscoff, 2006), to test for the cross-satellite identification at Ecogenics GmbH (Balgach, Switzerland).

Total DNA was extracted from young leaves using the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions. DNA was quantified by using BioSpec-nano (Shimadzu, Kyoto, Japan) and evaluated by electrophoresis on 0.8% agarose gels stained with ethidium bromide. DNA (50 μL, 450 ng/μL) from representative samples of *B. brevifolium* (vouchers ULS 4708, ULS 14346, ULS 14347; Appendix 1) was used for microsatellite enrichment. For this purpose, size-selected fragments from genomic DNA were enriched for the formation of an SSR-enriched library by using streptavidin-coated magnetic beads and biotin-labeled GATA and GTAT repeat oligonucleotides (Kijas et al., 1994; for a review, see Santana et al., 2009).

The SSR-enriched library was analyzed on an Illumina MiSeq (Illumina, San Diego, California, USA) at Microsynth AG (Balgach, Switzerland) using the Nano 2 × 250 v2 format. The resulting 284,104 sequences were stitched and assembled using MIRA 4.0.1 software (Chevreux et al., 1999), and candidate microsatellites were searched in 22,361 paired-end read assembled contigs using FINDER version 4.09 (Benson, 1999). Of the candidate microsatellites, 5983 contigs contained a microsatellite insert with a tetra- or a trinucleotide of at least six repeat units or a dinucleotide of at least 10 repeat units. Finally, Primer3 (Rozen and Skaletsky, 1999) was used to design PCR primers for the selected microsatellites using standard default values. Suitable primer design was possible in 2564 microsatellite candidates. Of them, a total of 430 microsatellites were further analyzed with size ranges from 80 bp to 250 bp; di-, tri-, and tetranucleotide motifs were identified, with seven to 22 repetitions per motif. These 430 sequences were deposited in GenBank (accession numbers MH052690–MH053105 and MF136749–MF136763; see Table 1). From this set, we selected 40 microsatellite markers for evaluation on three populations of *B. brevifolium*.

PCR amplifications contained, in a total volume of 12 μL, 30 ng of DNA, 2.4 μL of 5× colorless GoTaq Flexi buffer (Promega Corporation, Madison, Wisconsin, USA), MgCl₂, 1 mM, 250 μM of dNTPs, 0.4 μM of primers, 0.5 units (0.1 μL) of GoTaq DNA Polymerase Flexi (Promega Corporation), and completed with distilled H₂O. PCR cycling after an initial denaturation of 3 min at 94°C consisted of 35 cycles: 30 s at 95°C, annealing of 30 s at 60°C, and elongation of 60 s at 72°C. Finally, an incubation at 72°C for 10 min followed.

### TABLE 1. Microsatellites developed for *Balsamocarpren brevifolium*, an endemic xerophytic shrub from the Atacama Desert in Chile.

| Locus | Primer sequences (5′–3′) | Repeat motif | Allele size range (bp) | GenBank accession no. |
|-------|--------------------------|--------------|------------------------|-----------------------|
| BBR-005* | F: TGGCTCTCAAGTGGTCCATGCG  
R: CCTATTCCTAGAAGATGCCCC | (GAA)₁₃ | 190 | MF136749 |
| BBR-007* | F: TCGTACAGAAAGGGCTAAGAAAG  
R: CCTCTTTATCAAAATGCCC | (AGA)₁₂ | 135 | MF136750 |
| BBR-008* | F: TGCAAGCATACCTCAAGAGC  
R: TCGATCATCCTGCCACCTC | (TG)₁₂ | 158 | MF136751 |
| BBR-009  | F: CTGACGTTCAATTCTCCCCTC  
R: GCAGCACCTCTGTTTGGG | (TTT)₂₅ | 150–190 | MF136752 |
| BBR-010  | F: GTTCTCACACACTCAAGCG  
R: TTCTCGAGCACTTATTCCCC | (AGA)₁₂ | 175–220 | MF136753 |
| BBR-014  | F: GCTTCCGCTAATGCCTCTTC  
R: GCCACCCCAAGTGATTTCCTC | (CT)₁₁ | 160–190 | MF136754 |
| BBR-017* | F: TGGCAAGCTATGCTTTTCTTG  
R: GTCCCTTTTCTAGACGTGGT | (TCTT)₁₅ | 170 | MF136755 |
| BBR-018  | F: GTCCTCAAAGCCTAAATCCTTTTC  
R: GACCTGTGGGTATTGCTATTTG | (ATA)₁₁ | 170–215 | MF136756 |
| BBR-022  | F: TCCGGGATACCGTCTCCGTC  
R: TCTCCGGGGAGGAGGAGG | (TCTT)₁₀ | 120–130 | MF136757 |
| BBR-026  | F: GTGTTAGAGCTGATACATGAAATGC  
R: TGCCCTCTTTTATTTGCTTTAG | (CATA)₁₀ | 130–150 | MF136758 |
| BBR-039* | F: AAAGTTGGGTCGGGAAATGAC  
R: CGGATTTGGAATCCTAGGCC | (TTT)₂₃ | 170–178 | MF136759 |
| BBR-041  | F: ACCTACAGCAATGCTCTCTAATCG  
R: AGCCGGGAGAGTACTCATCG | (ACAT)₁₀ | 190–200 | MF136762 |
| BBR-043  | F: CTCCGAGTGTATTGTCCCC  
R: AAGCTGTGGCAAAATTTTGGG | (TG)₁₀ | 174–220 | MF136763 |
| BBR-078  | F: TGAGGGTTTTCATCATACCTGC  
R: AACTTACCGATTGAGGACGC | (AC)₁₅ | 244–282 | MH052718.1 |

*Annealing temperatures were the same for all loci (60°C).
Monomorphic markers.
min was included. PCR products were separated in 6% polyacrylamide gels and visualized by silver staining as described by Narváez et al. (2001).

Genetic diversity parameters including effective number of alleles ($A_e$) and observed and expected levels of heterozygosity ($H_o$ and $H_e$, respectively) were estimated using GenAlEx 6.5 (Peakall and Smouse, 2012). For each SSR marker, $A_e$ and $H_e$ were estimated for each population, and $H_o$ corresponded to the frequency of heterozygous individuals calculated for each population.

Of the 40 primer sets analyzed, 15 markers (10 polymorphic and five monomorphic) that showed clear amplification were analyzed in more detail (Table 1). Twenty-five primers showed no amplification, or a complex pattern, and were not further analyzed. Sequences of the 15 markers that showed clear amplification have been deposited in GenBank (Table 1); sequences of the remaining 25 primers are available upon request. For the 10 polymorphic primers, allele number ranged from three to nine (average of 5.77 alleles per marker in the populations), $A_e$ ranged from 2.27 to 6.00, and the expected and observed levels of heterozygosity varied from 0.56 to 0.83 and from 0.53 to 1.00, respectively (Table 2). These are the first expected and observed levels of heterozygosity varied from 0.56 to 0.83 and from 0.53 to 1.00, respectively (Table 2). These are the first expected and observed levels of heterozygosity varied from 0.56 to 0.83 and from 0.53 to 1.00, respectively (Table 2).

CONCLUSIONS

We have identified 15 new microsatellite markers for algarrobilla, 10 of which had high levels of polymorphism, representing the first markers developed in B. brevifolium. Some of these markers were also useful in the related species C. angulata and S. cumingii, which share the same habitat as B. brevifolium. Analyses of B. brevifolium individuals from different locations in the Atacama and Coquimbo regions in Chile would help characterize algarrobilla populations for conservation purposes and be the basis for future genetic studies for this species and possibly other endangered and related legumes endemic to northern Chile.

TABLE 3. Transferability of 10 polymorphic SSR markers developed for Balsamocarpus brevifolium in two Atacama Desert shrubs, Caesalpinia angulata and Senna cumingii (Fabaceae).*

| Locus | 1    | 2    | 3    | 4    | 5    | 1    | 2    | 3    | 4    | 5    | Balsamocarpus brevifolium |
|-------|------|------|------|------|------|------|------|------|------|------|--------------------------|
|       | 1/2  | 3/4  | 5/6  | 7/8  | 9/10 | 11/12| 13/14| 15/16| 17/18| 19/20|                          |
| BBR-09 | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−                      |
| BBR-02 | −      |      |      |      |      |      |      |      |      |      |                          |
| BBR-04 | 115–170 | +    | +    | +    | +    | +/−  | +/−  | +/−  | +/−  | +/−  | +/−                      |
| BBR-05 | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−                      |
| BBR-08 | −      |      |      |      |      |      |      |      |      |      | −                        |
| BBR-11 | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−                      |
| BBR-12 | −      |      |      |      |      |      |      |      |      |      | +/−                      |
| BBR-13 | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−                      |
| BBR-14 | −      |      |      |      |      |      |      |      |      |      | −                        |
| BBR-15 | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−  | +/−                      |
| Note: + = clear and strong PCR signal; ++ = weak signal; − = no amplification. |

The estimated allele sizes for each primer/species combination are indicated for sample 1 of each species.

http://www.wileyonlinelibrary.com/journal/AppsPlantSci © 2019 Ravest et al.
ACKNOWLEDGMENTS

This research was carried out as part of the Agreement of the Center for Disclosure, Research and Cultivation of Algarrobilla, established between the Instituto de Investigaciones Agropecuarias (INIA) and Barrick Pascua-Lama in order to support the preservation of this species.

DATA ACCESSIBILITY

Sequence information for the developed primers has been deposited to the National Center for Biotechnology Information (NCBI); GenBank accession numbers are provided in Table 1.

LITERATURE CITED

Acetino, P., H. Fuenzalida, and B. Rosenbluth. 1993. Climate along the extratropical west coast of South America. In H. Mooney, E. Fuentes, and B. Kronberg [eds.], Earth system responses to global change: Contrasts between North and South America, 61–69. Academic Press, San Diego, California, USA.

Arif, I. A., M. A. Bakir, H. A. Khan, A. H. Al Farhan, A. A. Al Homaidan, A. H. Bahkali, M. A. Sadoon, and M. Shobrak. 2010. A brief review of molecular techniques to assess plant diversity. International Journal of Molecular Sciences 11: 2079–2096.

Benson, G. 1999. Tandem repeats finder: A program to analyze DNA sequences. Nucleic Acids Research 27: 573–580.

Chevreux, B., T. Wetter, and S. Suhai. 1999. Genome sequence assembly using trace signals and additional sequence information. In Proceedings of the German Conference on Bioinformatics (GCB) 99, Hannover, Germany, 4–6 October 1999, pp. 45–56.

Estévez, R. A., F. A. Squeo, G. Arancio, and M. B. Erazo. 2010. Producción de carbón vegetal a partir de arbustos nativos en la Región de Atacama, Chile. Gayana Botanica 67: 213–222.

Gagnone, E., A. Bruneau, C. E. Hughes, L. Paganucci de Queiroz, and G. P. Lewis. 2016. A new generic system for the pantropical Caesalpinia group (Leguminosae). PhytoKeys 71: 1–160.

Kijas, J. M., J. C. Fowler, C. A. Garbett, and M. R. Thomas. 1994. Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. BioTechniques 16: 656–660, 662.

LPGW (The Legume Phylogeny Working Group). 2017. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. Taxon 66: 44–77.

Luebert, F., and P. Plisoff. 2006. Sinopsis bioclimática y vegetacional de Chile. Editorial Universitaria, Santiago, Chile.

Montecinos, A., and P. Aceituno. 2003. Seasonality of the ENSO-related rainfall variability in Central Chile and associated circulation anomalies. Journal of Climate 16: 281–296.

Narváez, C., M. H. Castro, J. Valenzuela, and P. Hinrichsen. 2001. Patrones genéticos de los cultivares de vides de vinificación más comúnmente usados en Chile basados en marcadores de microsatélites [Fingerprinting of wine grape cultivars most commonly grown in Chile based on microsatellite markers]. Agrícola Técnica (Chile) 61: 249–261.

Peakall, R., and P. E. Smouse. 2012. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—An update. Bioinformatics 28: 2537–2539.

Queller, D. C., J. E. Strassmann, and C. R. Hughes. 1993. Microsatellites and kinship. Trends in Ecology and Evolution 8: 285–288.

Rozen, S., and H. Skaletsky. 1999. Primer3 on the WWW for general users and for biologist programmers. In S. Misener and S. A. Krawetz [eds.], Methods in molecular biology, vol. 132: Bioinformatics: Methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

Santana, Q. C., M. P. A. Coetzee, E. T. Steenkamp, O. X. Mlonenyi, G. N. A. Hammond, M. J. Wingfield, and B. D. Wingfield. 2009. Microsatellite discovery by deep sequencing of enriched genomic libraries. BioTechniques 46: 217–223.

Squeo, F. A., G. Arancio, L. Cavieres, J. R. Gutiérrez, M. Muñoz, and C. Marticorena C. 2001. Análisis del estado de conservación de la flora nativa de la IV Región de Coquimbo. In F. A. Squeo, G. Arancio, and J. Gutiérrez [eds.], Libro Rojo de la flora nativa y los sitios prioritarios para su conservación: Región de Coquimbo, 53–63. Ediciones Universidad de La Serena, La Serena, Chile.

Squeo, F. A., G. Arancio, L. Letelier, A. Marticorena, M. Muñoz-Schick, P. León-Lobos, and M. T. K. Arroyo. 2008. Estado de conservación de la flora nativa de la Región de Atacama. In F. A. Squeo, G. Arancio, and J. Gutiérrez [eds.], Libro Rojo de la flora nativa y los sitios prioritarios para su conservación: Región de Atacama, 45–59. Ediciones Universidad de La Serena, La Serena, Chile.

Ulibarri, E. A. 2008. Los géneros de Caesalpinioideae (Leguminosae) presentes en Sudamérica. Darwiniana 46: 69–163.

Wrann, H. J., and R. D. Barros. 1987. Ensayos de reforestación por siembra directa con algarrobilla (Balsamocarpus brevifolium Clos) en la zona de Vallenar [Trials of reforestation by direct sowing with algarrobilla (Balsamocarpus brevifolium) around Vallenar]. Ciencia e Investigación Forestal (Chile) 1: 45–55.

APPENDIX 1. Herbarium vouchers and geographical coordinates for representative specimens of Balsamocarpus brevifolium, Caesalpinia angulata, and Senna cumingii from Chile used for the development of microsatellite markers.*

| Species | Population code | Voucher no. | Geographical coordinates | Elevation (m) | N |
|---------|----------------|-------------|-------------------------|--------------|---|
| Balsamocarpus brevifolium Clos | Population 1 | ULS 14346 | 28°13’S, 70°42’W | 426 | 21 |
| Population 2 | ULS 14345 | 28°53’S, 70°46’W | 1043 | 18 |
| Population 3 | ULS 14347 | 28°08’S, 70°51’W | 1096 | 13 |
| Caesalpinia angulata (Hook. & Arn.) Baill. | ND | ULS 4693 | 29°7’S, 70°53’W | 750 | >10 |
| Senna cumingii [(Hook. & Arn.) H. S. Irwin & Barneby] | ND | ULS 4708 | 28°38’S, 70°49’W | 435 | >10 |

*Note: N = number of individuals; ND = population not defined.

**GPS and elevation ranges for all individuals within each population: Population 1 = 28°5’S, 70°36’W to 28°13’S, 70°43’W; 399–540 m; Population 2 = 28°53’S, 70°52’W to 28°55’S, 70°51’W, 863–2130 m; Population 3 = 29°8’S, 70°36’W to 29°11’S, 70°52’W, 1091–1999 m.

**Voucher specimens were deposited at the herbarium of Universidad de La Serena (ULS), La Serena, Chile.