Primer Note

Characterization of 23 polymorphic SSR markers in *Salix humboldtiana* (Salicaceae) using next-generation sequencing and cross-amplification from related species[1]

Jorge A. Bozzi[2,6], Sascha Liepelt[3], Sebastian Ohneiser[3], Leonardo A. Gallo[2], Paula Marchelli[2,4], Ilona Leyer[3,5], Birgit Ziegenhagen[3], and Christina Mengel[3]

[2]Unidad de Genética Ecológica y Mejoramiento Forestal, Instituto Nacional de Tecnología Agropecuaria (INTA), E.E.A. Bariloche, Modesta Victoria 4450, Bariloche 8400, Argentina; [3]Faculty of Biology, Conservation Biology Group, University of Marburg, Karl-von-Frisch-Str. 8, 35043 Marburg, Germany; [4]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Rivadavia 1917, Ciudad Autónoma de Buenos Aires C1092AAR, Argentina; and [5]Institute of Botany, Plant Ecology and Nature Conservation, University of Geisenheim, Von-Lade-Str. 1, 65366 Geisenheim, Germany

- *Premise of the study:* We present a set of 23 polymorphic nuclear microsatellite loci, 18 of which are identified for the first time within the riparian species *Salix humboldtiana* (Salicaceae) using next-generation sequencing.
- *Methods and Results:* To characterize the 23 loci, up to 60 individuals were sampled and genotyped at each locus. The number of alleles ranged from two to eight, with an average of 4.43 alleles per locus. The effective number of alleles ranged from 1.15 to 3.09 per locus, and allelic richness ranged from 2.00 to 7.73 alleles per locus.
- *Conclusions:* The new marker set will be used for future studies of genetic diversity and differentiation as well as for unraveling spatial genetic structures in *S. humboldtiana* populations in northern Patagonia, Argentina.

Key words: next-generation sequencing; nuclear microsatellite marker; river margins; Salicaceae; *Salix humboldtiana*.

The diploid *Salix humboldtiana* Willd. (Salicaceae) is the only native *Salix* L. species in the southern hemisphere (Becerra et al., 2009). This dioecious species forms dense natural stands on wet sand banks of river margins (Tortorelli, 2009). Its natural distribution range, one of the widest among Argentinean native woody species, reaches from Mexico in the northern hemisphere to Argentina and Chile on the 45th parallel in the southern hemisphere (Tortorelli, 2009). In northern Patagonia, floodplain forests structured by *S. humboldtiana* have been displaced by mixed forests dominated by Eurasian invasive willows and poplars (Thomas and Leyer, 2014). Together with landscape fragmentation and alterations of the hydrological regime caused by dam construction, these invasion processes represent a serious threat to this species in northern Patagonian riparian ecosystems. In addition, there is genetic evidence of hybridization between *S. humboldtiana* and invasive willows (Bozzi et al., unpublished). Programs for conserving its genetic resources and genetic diversity should be established. Therefore, knowledge is needed about nuclear genetic diversity. To be able to perform these analyses, we developed markers for ongoing population genetic research in *S. humboldtiana*. Addi-

tionally, we tested cross-amplification in *S. humboldtiana* of markers previously developed for related species.

METHODS AND RESULTS

For primer testing and diversity assessment, a total of 60 *S. humboldtiana* individuals were sampled at the Río Negro in Argentina. DNA was extracted from dried leaves following the protocol by Dumolin et al. (1995). A novel set of microsatellite markers specific for *S. humboldtiana* was developed using a next-generation sequencing approach (Table 1). 454 sequencing was performed by Ecogenics (Zurich-Schlieren, Switzerland) as follows: DNA was enriched for two repeat motifs (CT and GT) representing all enrichable dinucleotide motifs, and sequenced on a GS FLX (Roche Applied Science, Indianapolis, Indiana, USA) after library construction. A FASTA file was provided with a total of 14,714 reads covering 2.23 Mb and exhibiting an average read length of 152 bases and a mode of 101 bases (sequence data available upon request). We used the software QDD (Meglécz et al., 2010) to assemble the reads and screen for di-, tri-, tetra-, penta-, and hexanucleotide repeats. The screening was performed using default settings except for the minimum length of PCR product, which was set to 80 bp. The software Primer3 (Rozen and Skaletsky, 1999), included in QDD, was used to design primer pairs for PCR for 67 sequences containing microsatellite motifs with the specified characteristics. We discarded all loci (eight sequences) showing compound or interrupted simple sequence repeats (SSRs). Moreover, from the 59 sequences that showed a perfect microsatellite motif, 17 loci had to be removed because they showed undesirable properties such as poor sequence quality in the flanking region or SSR stretches too close to the end of the read. Primer pairs for the remaining 42 loci were ordered from Metabion (Martinsried, Germany). Additionally, 24 extra primer pairs suggested by Ecogenics, and 24 primer pairs previously developed for related species were tested (*Salix alba* L. [King et al., 2010], *S. lanata* L. [Stamati et al., 2003], *Populus trichocarpa* Torr. & A. Gray [Tuskan et al., 2004; International *Populus* Genome Consortium: http://www.ornl.gov/sci/ipgc/ssr_resource.htm],

TABLE 1. Characteristics of 18 nuclear microsatellite loci developed in *Salix humboldtiana*.

Locus	Primer sequences (5′–3′)	Repeat motif	Allele size range (bp)	GenBank accession no.
Shum_002	F: ACTTGCAGGGGTGTCTACTG R: AGGTCAAAACATTGACATCCAAATTC	$(AC)_{13}$	100–106	KP208969
Shum_006	F: CAACACAACACAACAACGCA R: GGAGAAAGATCTCCGCTTTG	$(AACAC)_8$	102–117	KP208970
Shum_029	F: TAAGCTACCCCTGACAACCC R: AGCCCCATGAATAATCCCCG	$(AC)_{13}$	107–109[a]	KP208971
Shum_032	F: ACCAAGCTGGCAATATGGAG R: CCGTTTGGAACTTTGTGATG	$(AGC)_6$	101–128	KP208972
Shum_033	F: AATGAGCAGTGCCTTTTGAC R: GAACATGTGGGTCGTTCTCC	$(CA)_{19}$	105–109[a]	KP208973
Shum_047	F: TGCAAATCCATAATGACTTCTTTC R: GCCTAGGCCACTTTGTGTTC	$(AC)_{19}$	121–133	KP208974
Shum_049	F: TATCCATCTTTCGAGCTGGC R: TCTCGCTCTATCTGCCATCA	$(AC)_{13}$	237–241[a]	KP208975
Shum_060	F: TGACACGCATCCCTTCTGTG R: ACAGTTCTGAATGCCAGTCTC	$(AC)_{20}$	110–130	KP208976
Shum_061	F: TATTTTGATTCGAGCCCCCG R: TTTCGTCCACTCTGGCTTCC	$(GT)_{19}$	108–116	KP208977
Shum_062	F: TTTAAGAACGATGGTGGGGG R: TCCTTGTACCCGAGTTCTGC	$(AC)_{17}$	162–168[a]	KP208978
Shum_064	F: ATGTCCAAGAGTGCGCTATG R: GACTAGTTGTGCAGTACACGC	$(GT)_{16}$	99–103	KP208979
Shum_066	F: ATTTGATCGCGGAGGTCACG R: ACCTTATGTTTCCTTTTAATGTTGAG	$(AC)_{16}$	115–129	KP208980
Shum_067	F: TCAAATGCGCCGGGATAAATAG R: AGCTCATACCAACCACATCTAC	$(GT)_{16}$	100–112[a]	KP208981
Shum_070	F: ATCGGATGGATCGGGCATAG R: GAGGGGAGTACACTCTAAAACC	$(AC)_{15}$	161–167	KP208982
Shum_071	F: GTAACAGACTTGGCAACCCG R: TTGGCGGCTTCCATTACATC	$(CA)_{14}$	244–248	KP208983
Shum_074	F: TTCCAGCCTTAGATTGCTTGC R: GTCAACTCAGCTGCCATTCG	$(AC)_{13}$	82–86	KP208984
Shum_076	F: TATCTGATCCACCCCATGCC R: TTACAACTCTGCAATAGTAAGATCC	$(CA)_{13}$	141–153	KP208985
Shum_077	F: AGTAGTTTTCGCATACGCTG R: ATGTCACTGGTAGAGGACGC	$(TG)_{13}(AC)_{13}$	180–192	KP208986

[a] Fragment size calculated by discounting 18 bp belonging to the M13 fragment.

S. burjatica Nasarow [Hanley et al., 2002; Barker et al., 2003], *S. arbutifolia* Pall. [Hoshikawa et al., 2009], *P. nigra* L. [Van der Schoot et al., 2000; Smulders et al., 2001], and *S. hukaoana* Kimura [Kikuchi et al., 2005]). Fluorescent labeling using M13 primer tails was performed according to Schuelke (2000) to test this high number of loci in a cost-efficient way. We used a subset of eight individuals to prescreen the quality of the amplified SSRs on a MegaBACE 1000 automated capillary sequencer (GE Healthcare, Freiburg, Germany). Scorable polymorphic bands were revealed by 23 SSR loci (Table 2), while no amplification, pronounced stutter bands, multibanding patterns, or monomorphic bands were shown by the remaining 67 loci. To further characterize the 23 selected loci, the number of analyzed samples was increased to at least 22 and up to 60 individuals. Only 14 loci were screened using 60 individuals belonging to two populations (Appendix 1), while the remaining loci were evaluated with a panel of individuals sampled at different locations along the river. Fluorescence-labeled primers were ordered for those loci with a high level of polymorphism and good scorability. PCRs were performed in a 16.6-μL mix containing 1.2 ng/μL of template DNA, 1× PCR reaction buffer (Molegene, Butzbach, Germany), 0.3 mM of each dNTP (Bioline, Luckenwalde, Germany), 0.04 U/μL of *Taq* polymerase (Molegene), 0.2 μM of each primer (Metabion), 0.16 mg/mL of bovine serum albumin (BSA; Thermo Scientific, St. Leon-Rot, Germany), and 2.4–3.0 mM of $MgCl_2$ (Molegene), depending on the locus to be amplified (Table 3). PCR amplification was conducted using a T1 Thermocycler (Biometra, Göttingen, Germany) performed with 5 min of initial denaturation at 94°C, followed by 30–40 cycles of 30–45 s of denaturation at 94°C, 45 s of annealing at 53–60°C, 30–45 s of elongation at 72°C, and 10 min of final elongation at 72°C. For some primer pairs, a touchdown PCR was conducted. The cycling process of the touchdown PCR was performed with 30–45 s of denaturation at 94°C, 45 s of annealing with temperatures decreasing 1°C per cycle from 65–60°C to 56–51°C during the first 10 cycles and temperatures of 57–60°C for the last 20–25 cycles, and 30–45 s of elongation at 72°C (Table 3). Primer pairs are reported in Table 1.

A MegaBace 1000 automated capillary sequencer (GE Healthcare) was used to separate the SSR amplicons by capillary electrophoresis. For allele sizing, the internal size standard MegaBACE ET400-R (GE Healthcare) and MegaBACE Genetic Profiler software (version 1.2; GE Healthcare) were used.

Genetic diversity parameters (Table 2) and deviations from Hardy–Weinberg equilibrium (HWE) were estimated using GenAlEx version 6.5 (Peakall and Smouse, 2012). The number of alleles ranged from two to eight with an average of 4.43 alleles per locus, while the effective number of alleles ranged from 1.15 to 3.09 per locus. Allelic richness standardized by rarefaction was estimated according to El Mousadik and Petit (1996), and values ranged from 2.0 to 7.73 alleles per locus. The observed and expected heterozygosities ranged from 0.036 to 0.786 and 0.035 to 0.721, respectively. Significant deviation from expected heterozygote frequencies was observed for Shum_070 ($P < 0.05$); Shum_032 and Shum_076 ($P < 0.01$); and Shum_033 ($P < 0.001$). The software MICRO-CHECKER version 2.2.3 (van Oosterhout et al., 2004) was used to detect null alleles, and evidence for null alleles was detected at only one locus (Shum_032) in one population. Deviation from HWE can be explained by the presence of null alleles in the case of Shum_032. Inbreeding due to small effective population size can explain deviations from HWE in the case of Shum_033, Shum_070, and Shum_076. Fisher's exact test analysis to detect linkage disequilibrium was performed using GENEPOP version 4.3 (Rousset, 2008). Linkage disequilibrium ($P < 0.05$) was detected between two loci: Shum_002 and Shum_070 (Table 2).

TABLE 2. Characteristics of 23 nuclear microsatellite loci developed in this and other studies for *Salix humboldtiana*.

Locus[a]	N	A	A_R	A_e	RNo03[b] H_o	RNo03[b] H_e[c]	RNo13[b] H_o	RNo13[b] H_e[c]
Shum_002	60	4	3.98	2.42	0.517	0.563[ns]	0.677	0.598[ns]
Shum_006	36	3	2.61	1.73	—	—	—	—
Shum_029	31	2	2.00	1.29	—	—	—	—
Shum_032	60	4[‡]	3.85	1.40	0.103	0.267**	0.355	0.301[ns]
Shum_033	53	3	3.00	1.92	0.680	0.473[ns]	0.786	0.477***
Shum_047	60	7	6.83	2.33	0.414	0.555[ns]	0.516	0.571[ns]
Shum_049	32	3	2.69	1.21	—	—	—	—
Shum_060	53	6	5.96	2.10	0.542	0.531[ns]	0.586	0.505[ns]
Shum_061	56	5	5.00	3.09	0.654	0.637[ns]	0.667	0.701[ns]
Shum_062	31	4	3.42	2.13	—	—	—	—
Shum_064	53	3	2.98	2.00	0.458	0.510[ns]	0.517	0.490[ns]
Shum_066	60	8	7.73	2.46	0.448	0.488[ns]	0.613	0.668[ns]
Shum_067	32	4	3.37	1.29	—	—	—	—
Shum_070	59	4	3.76	1.17	0.071	0.070[ns]	0.194	0.203*
Shum_071	59	3	2.88	1.44	0.321	0.275[ns]	0.226	0.331[ns]
Shum_074	59	3	3.00	1.15	0.036	0.035[ns]	0.226	0.207[ns]
Shum_076	58	5	4.78	1.19	0.071	0.135**	0.200	0.186[ns]
Shum_077	52	5	5.00	1.30	0.083	0.081[ns]	0.393	0.340[ns]
Sa54A	60	7	6.86	2.74	0.414	0.508[ns]	0.710	0.721[ns]
gSIMCT24	43	7	5.24	1.49	—	—	—	—
ORPM446	22	3	3.00	1.50	—	—	—	—
SB196	29	3	3.00	1.64	—	—	—	—
WPMS18	32	6	5.71	1.73	—	—	—	—

Note: A = number of alleles; A_e = effective number of alleles; A_R = allelic richness standardized by rarefaction; H_e = expected heterozygosity; H_o = observed heterozygosity; N = number of individuals analyzed.

[a] Previously developed loci: Sa54A (King et al., 2010); gSIMCT24 (Stamati et al., 2003); ORPM446 (Tuskan et al., 2004); SB196 (Barker et al., 2003); WPMS18 (Smulders et al., 2001).

[b] Values of H_e and H_o are not shown for loci evaluated with a panel of individuals that were not part of populations RNo03 or RNo13.

[c] Significant deviations from Hardy–Weinberg equilibrium: $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; ns = not significant.

[‡] Locus showing null allele at one of the analyzed populations.

CONCLUSIONS

In the near future, the described microsatellite markers will be used to analyze the genetic structure and diversity of *S. humboldtiana* along river stretches in northern Patagonia, Argentina, and to unravel dispersal processes as well as effects of landscape fragmentation and biological invasions. To our knowledge, no SSR markers had been previously developed for *S. humboldtiana*. The new marker set can be used for future studies of genetic diversity and differentiation as well as for estimating dispersal distances and determining spatial genetic structures. Beyond population genetic applications, these markers may also be useful for clone identification, genome mapping, and breeding purposes. Furthermore, they may be useful in testing for cross-amplification in related species and developing PCR multiplexes for fast and economic genotyping.

LITERATURE CITED

BARKER, J. H. A., A. PAHLICH, S. TRYBUSH, K. J. EDWARDS, AND A. KARP. 2003. Microsatellite markers for diverse *Salix* species. *Molecular Ecology Notes* 3: 4–6.

BECERRA, A. G., E. R. NOUHRA, M. P. SILVA, AND D. MCKAY. 2009. Ectomycorrhizae, arbuscular mycorrhizae, and dark-septate fungi on *Salix humboldtiana* in two riparian populations from central Argentina. *Mycoscience* 50: 343–352.

TABLE 3. Summary list of PCR cocktail $MgCl_2$ content and cycle profiles for 23 polymorphic nuclear microsatellite markers amplified in *Salix humboldtiana*.

Locus[a]	$MgCl_2$ (mM)[b]	Cycle profile[c] Denaturation (s)	Cycle profile[c] T_a (°C)	Cycle profile[c] Elongation (s)	No. of cycles
Shum_002	2.6	35	60	35	30
Shum_032	3.0	40	55	40	35
Shum_033	3.0	40	54	40	35
Shum_047	3.0	30	55	30	35
Shum_060	3.0	30	60	30	30
Shum_061	3.0	30	58	30	40
Shum_064	3.0	30	TD (60/51) 57	30	10–25
Shum_066	3.0	35	TD (63/54) 59	35	10–25
Shum_070	3.0	40	TD (61/52) 58	40	10–20
Shum_071	2.7	40	56	40	34
Shum_074	2.6	30	57	30	30
Shum_076	2.7	45	TD (65/56) 60	45	10–23
Shum_077	3.0	40	54	40	40
Shum_006	2.6	30	53	30	30
Shum_029	3.0	45	57	45	35
Shum_049	3.0	45	57	45	35
Shum_062	3.0	45	55	45	35
Shum_067	3.0	45	TD (60/51) 57	45	10–25
Sa54A	2.6	40	59	40	35
gSIMCT24	2.4	45	54.5	45	33
ORPM446	3.0	45	55	45	30
SB196	3.0	45	54	45	30
WPMS18	3.0	45	59	45	35

Note: T_a = annealing temperature.

[a] Previously developed loci: Sa54A (King et al., 2010); gSIMCT24 (Stamati et al., 2003); ORPM446 (Tuskan et al., 2004); SB196 (Barker et al., 2003); WPMS18 (Smulders et al., 2001).

[b] $MgCl_2$ content in the PCR cocktails.

[c] Denaturing temperature 94°C; annealing cycle run for 45 s; TD = touchdown PCR, with the range of annealing temperatures for the first 10 cycles in parentheses; elongation temperature 72°C.

DUMOLIN, S., B. DEMESURE, AND R. J. PETIT. 1995. Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. *Theoretical and Applied Genetics* 91: 1253–1256.

EL MOUSADIK, A., AND R. J. PETIT. 1996. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics* 92: 832–839.

HANLEY, S., A. BARKER, W. VAN OOIJEN, C. ALDAM, L. HARRIS, I. AHMAN, S. LARSSON, AND A. KARP. 2002. A genetic linkage map of willow (*Salix viminalis*) based on AFLP and microsatellite markers. *Theoretical and Applied Genetics* 105: 1087–1096.

HOSHIKAWA, T., S. KIKUCHI, T. NAGAMITSU, AND N. TOMARU. 2009. Eighteen microsatellite loci in S*alix arbutifolia* (Salicaceae) and cross-species amplification in *Salix* and *Populus* species. *Molecular Ecology Resources* 9: 1202–1205.

KIKUCHI, S., W. SUZUKI, N. BAN, A. KANAZASHI, AND H. YOSHIMARU. 2005. Characterization of eight polymorphic microsatellites in endangered willow *Salix hukaoana*. *Molecular Ecology Notes* 5: 869–870.

KING, R. A., S. L. HARRIS, A. KARP, AND J. H. A. BARKER. 2010. Characterisation and inheritance of nuclear microsatellite loci for use in population studies of the allotetraploid *Salix alba–Salix fragilis* complex. *Tree Genetics & Genomes* 6: 247–258.

MEGLÉCZ, E., C. COSTEDOAT, V. DUBUT, A. GILLES, T. MALAUSA, N. PECH, AND J.-F. MARTIN. 2010. QDD: A user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics* 26: 403–404.

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.

Rousset, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.

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.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.

Smulders, M. J. M., J. Van Der Schoot, P. Arens, and B. Vosman. 2001. Trinucleotide repeat microsatellite markers for black poplar (*Populus nigra* L.). *Molecular Ecology Notes* 1: 188–190.

Stamati, K., S. Blackie, J. W. S. Brown, and J. Russell. 2003. A set of polymorphic SSR loci for subarctic willow (*Salix lanata*, *S. lapponum* and *S. herbacea*). *Molecular Ecology Notes* 3: 280–282.

Thomas, L. K., and I. Leyer. 2014. Age structure, growth performance and composition of native and invasive Salicaceae in Patagonia. *Plant Ecology* 215: 1047–1056.

Tortorelli, L. A. 2009. Maderas y bosques Argentinos, 2nd ed. Orientación Gráfica Editora, Buenos Aires, Argentina.

Tuskan, G. A., L. E. Gunter, Z. K. Yang, T. Yin, M. M. Sewell, and S. P. DiFazio. 2004. Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. *Canadian Journal of Forest Research* 34: 85–93.

Van der Schoot, J., M. Pospíšková, B. Vosman, and M. J. M. Smulders. 2000. Development and characterization of microsatellite markers in black poplar (*Populus nigra* L.). *Theoretical and Applied Genetics* 101: 317–322.

van Oosterhout, C., W. F. Hutchison, P. Shipley, and D. P. M. Wills. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.

Appendix 1. Voucher and location information for the *Salix humboldtiana* populations used in this study. The vouchers are deposited in Herbarium Marburgense, University of Marburg, Marburg, Germany.

Population	Locality	Geographic coordinates	N	Voucher no.	Herbarium ID	Collector
RNo03	Allen, Río Negro, Argentina	39°2′43.50″S, 67°47′44.34″W	29	RNo03-26	MB-001506	Bozzi, J.
RNo13	Beltrán, Río Negro, Argentina	39°16′13.96″S, 65°49′26.86″W	31	RNo13-06	MB-001538	Bozzi, J.

Note: N = number of individuals sampled.