Development of EST-SSR Markers in *Barringtonia racemosa* (Lecythidaceae) and Cross-Amplification in Related Species

HONGXIAN XIE², YANG YUAN², XIAOTING FANG², YING LIU², CHAO YANG⁴, JIANHUA JIN², FENGXIAO TAN⁵, AND YELIN HUANG²

²Guangdong Key Laboratory of Plant Resources and State Key Laboratory of Biocontrol, Sun Yat-sen University, Guangzhou 510275, Guangdong, People’s Republic of China; ³Division of Life Science, The Hong Kong University of Science and Technology, Hong Kong, People’s Republic of China; ⁴Department of Neurosurgery, First Affiliated Hospital of Sun Yat-sen University, Guangzhou 510080, Guangdong, People’s Republic of China; and ⁵College of Natural Resources and Environment, South China Agricultural University, Guangzhou 510642, Guangdong, People’s Republic of China

**Premise of the study:** Microsatellite markers were identified and characterized to study the genetic diversity and structure of *Barringtonia racemosa* (Lecythidaceae).

**Methods and Results:** Based on the transcriptome data of *B. racemosa*, 30 primer pairs were initially designed and tested, of which 15 were successfully amplified and displayed clear polymorphisms across the 43 individuals from three distant populations tested in the study. The results showed that the number of alleles per locus ranged from two to seven and the expected heterozygosity and observed heterozygosity per locus varied from 0 to 0.772 and from 0 to 0.933, respectively.

**Conclusions:** The expressed sequence tag–simple sequence repeat (EST-SSR) markers described here will be useful for studying genetic diversity and structure of *B. racemosa*. Furthermore, all loci were successfully cross-amplified in *B. asiatica* and *B. acutangula* and will be of great value for genetic studies across this genus.

**Key words:** Barringtonia racemosa; cross-amplification; EST-SSR; Lecythidaceae; transcriptome.

*Barringtonia racemosa* (L.) Spreng. (Lecythidaceae), the most widespread species of *Barringtonia*, is characteristic of coastal communities in the Paleotropics (Tomlinson, 1986; Prance, 2012). The flowers open at night and are pollinated by night-flying animals. The bark and crushed fruit contain sapo-nins and are used as a fish poison (Payens, 1967). This tree has a distribution from the Pacific to East Africa (Prance, 2012); it is variously recorded as a mangrove associate and is abundant along tidal rivers and in areas subject to tide and salinity. It also occurs inland on the edge of peat swamp forests and on hillsides to altitudes of 200 m (Tomlinson, 1986). The fruit are buoyant because of the spongy fibrous pericarp, and the wide distribution is on account of its dispersal by water (Prance, 2012). However, the actual level of interpopulation migration has never been studied. Recent advances in molecular techniques, such as microsatellite markers, have created new opportunities for population genetic research that can be used to estimate the genetic variation and clarify the level of migration among populations of *B. racemosa* across its range. Transcriptome sequencing is an efficient method for acquiring expressed sequence tag–simple sequence repeat (EST-SSR) markers (Bouck and Vision, 2007) based on its low cost and high reliability. Here, we describe the development of 15 nuclear microsatellite markers for ongoing population genetic research in *B. racemosa*. Additionally, we tested the transferability of these markers for studying two related species, *B. asiatica* (L.) Kurz and *B. acutangula* (L.) Gaertn.; these three species comprise all of the wide-ranging, water-dispersed species in *Barringtonia* (Prance, 2012).

**Methods and Results**

In this study, a three-month-old seedling of *B. racemosa* (from Dongzhai harbor, Hainan, China) was immediately frozen in liquid nitrogen and stored at −80°C for RNA extraction. The total RNA of *B. racemosa* was extracted using an improved cetyltrimethylammonium bromide (CTAB) method (Fu et al., 2005). The RNA quality and quantity were measured using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, California, USA). The mRNA was isolated using Oligotex-dT30 (TaKaRa Biotechnology Co., Dalian, China) after extraction and then fragmented ultrasonically. The cDNA libraries were prepared for sequencing following the Illumina protocol, and paired-end short read sequencing was done using the Illumina Genome Analyzer II system (Illumina, San Diego, California, USA; sequencing performed by BGI, Shenzhen, Guangdong Province, China) (Huang et al., 2012; Qiang et al., 2015). A total of 13,510,000 90-nucleotide paired-end reads were obtained and assembled using Trinity (release 20140413) with the default parameters (Grabherr et al., 2011). CAP3 (Huang and Madan, 1999) was further used to cluster similar contigs and obtain a high-quality reference genome with nonredundant unigenes (minimum identity = 99%). The analysis yielded 63,932 contigs with an average length of 589 bp, an N50 length of 808, and an average depth of coverage of 18.14×. The software MISA (Thiel et al., 2003) was used to detect
Table 1. Characteristics of 15 microsatellite loci isolated from *Barringtonia racemosa*.

| Locus | Primer sequences (5′–3′) | Repeat motif | Allele size range (bp) | \( T_a \) (°C) | GenBank accession no. | BLAST top hit description [organism] | BLAST top hit accession no. | E-value |
|-------|--------------------------|--------------|------------------------|---------------|-----------------------|-------------------------------------|-------------------------------|---------|
| br-39616 &br-60198 &br-6303 &br-39818 &br-39641 &br-7338 &br-6800 &br-7130 &br-38256 &br-39436 &br-32408 &br-38577 &br-11107 &br-54330 &br-23885 | F: TTGGCTCTGCTCTGGTATCC R: AGTTGGCGCTATCTCTCCTC | (TG)\(_8\) | 112–126 | 56 | KR822410 | CASP-like protein 4 [Zea mays] | NM_001137348 | 5e-07 |
| F: ATCTACATGGTGGCTGCACA R: CGAAGCTCAGCTCCCATCAA | (TC)\(_8\) | 106–131 | 58 | KR822411 | Hypothetical protein partial [Morus notabilis] | XM_010103213 | 3e-20 |
| F: GTTTTTGAGAGACAGAGGCG R: ACTCGATATCCTGGAGCCTAG | (GA)\(_5\) | 233–266 | 58 | KR822412 | DNA photolyase, putative [Ricinus communis] | XM_002523046 | 0.000 |
| F: TCGGATATCTTTGGTGTTCC R: AAAAGCACCATTTGAGCCTCAT | (ATC)\(_5\) | 193–200 | 60 | KR822413 | Uncharacterized protein [Arabidopsis thalina] | NM_121752 | 2e-68 |
| F: TAGTCCTCAAACAGGAGGCTT R: AAGGAGGAATGGAGAATCA | (GCT)\(_5\) | 113–128 | 56 | KR822414 | Transcription factor E2F isoform 1 [Theobroma cacao] | XM_007015654 | 0.000 |
| F: GTGAAGTGGCTCTGCTGTTA R: AGCACATATTACAGCAATCC | (TGG)\(_8\) | 244–259 | 58 | KR822415 | E2F transcription factor-I family protein [Populus trichocarpa] | XM_006384725 | 0.000 |
| F: ATCTGAGCTGGGCGGGCGCAA R: AGCAAGCTCAGCTCCCATCAA | (AG)\(_8\) | 184–206 | 60 | KR822409 | Major facilitator protein isoform 1 [Theobroma cacao] | XM_007031842 | 0.000 |
| F: AGTGACGCGATTTACCTGCGT | (CCG)\(_5\) | 154–170 | 60 | KR822416 | Major facilitator superfamily protein isoform 1 [Theobroma cacao] | XM_007031841 | 0.000 |
| F: ATTACTCCTGCGCACATTTG R: TGAGGTTGACATCGCATCAGTAG | (AGG)\(_8\) | 124–133 | 60 | KR822418 | Hypothetical protein [Morus notabilis] | XM_010114075 | 0.000 |
| F: GACGGAGTTGAGAGAACTTTG C: GTGTGAATCAGTCTCGTAG | (TA)\(_6\) | 154–178 | 58 | KR822417 | Class III peroxidase [Populus trichocarpa] | KJ_201853 | 4e-50 |
| F: GGCTCCTGGCTGCTGCTGTTA R: GGGGCTTTGCTGCTGTGAAT | (TG)\(_8\) | 218–248 | 58 | KR822404 | Rhodanese/cell cycle control phosphatase superfamily protein [Francoa sonchifolia] | KM_461132 | 2e-142 |
| F: CTCTCTCCCTCCCTCTGCTCAA R: GAACTCGAGATCGCTGGGCTT | (GAA)\(_7\) | 244–268 | 58 | KR822405 | Leucine-rich repeat extension-like protein 4 [Morus notabilis] | XM10090551 | 2e-110 |
| F: AGACAAACGCGGTCTCTCATT R: GCCGGCATGAGGGAAGGAACAGTT | (AGG)\(_9\) | 262–274 | 60 | KR822406 | Hypothetical protein [Phaseolus vulgaris] | XM_007161158 | 1e-165 |
| F: ACCGTATCTGGTGGGCTGTTA R: AACGCTCCGAGAAACACCC | (AT)\(_9\) | 240–270 | 56 | KR822407 | Hypothetical protein [Morus notabilis] | XM_010902747 | 1e-180 |
| F: GTTTGSGCAGTGTTGGCTGG T: TGCTGGCCATTTGATTCCTT | (TGC)\(_7\) | 262–280 | 58 | KR822408 | Auxin response factor 5 [Morus notabilis] | XM_010108646 | 0.000 |

Note: \( T_a \) = annealing temperature.
microsatellites from all unigenes obtained in the study. We screened for SSR motifs containing two to six nucleotides with the minimum number of repeats as follows: six for dinucleotide and five for trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide. Altogether, 6524 SSR motifs were found, and 30 of them were selected at random to design primers using Primer3 (Rozen and Skaletsky, 1999), with the optimum conditions set at a length of 20 bp (18–22 bp), a temperature of 60.0°C (56–62°C), and a product size range of 120–250 bp.

Forty-three individuals of B. racemosa from three natural populations (Appendix 1) were used to evaluate the polymorphisms of the target microsatellite loci. In addition, individuals from one population of B. asiatica and one population of B. acutangula were also sampled to detect the efficiency of these markers in cross-species amplification (Appendix 1). Genomic DNA was extracted from silica gel–dried leaves using the DNA Extraction Kit (Magen, Guangzhou, China) following the manufacturer’s protocol. PCR amplifications were performed in a final volume of 50 μL, containing 60 ng of genomic DNA, 1× PCR buffer (10 mM Tris-HCl pH 8.4 and 1.5 mM MgCl2; TransGen Biotech Co., Beijing, China), 0.2 mM dNTPs (TransGen Biotech Co.), 0.5 μM of each primer (Life Technologies, Shanghai, China), and 1 unit EasyTaq DNA polymerase (TransGen Biotech Co.). The PCR reactions were carried out under standard conditions for all primers in a Bio-Rad PTC-200 thermocycler (Bio-Rad Laboratories, Hercules, California, USA). Raw data were analyzed and allele sizes and number of alleles per locus were called using PROSize version 2.0 software (AATI). GenAlEx version 6.5 (Peakall and Smouse, 2012) was used to calculate the average number of alleles per locus (A), the observed heterozygosity (Ho), and the expected heterozygosity (He) of each microsatellite locus. Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium were tested using GENEPOP 4.3 (Rousset, 2008). MICRO-CHECKER (van Oosterhout et al., 2004) was employed for testing scoring errors and null alleles.

In total, 15 out of 30 primer pairs were successfully amplified. The sequences of microsatellite loci were deposited into GenBank (accession no. KR822404–KR822418; Table 1). The SSR genotyping was performed using a Fragment Analyzer Automated CE System (Advanced Analytical Technologies [AAATT], Ames, Iowa, USA) under the following conditions: initial denaturation was at 94°C for 4 min, followed by 35 cycles at 94°C for 40 s, then annealing for 45 s at the optimal temperature for each primer pair (from 56–60°C, see Table 1), and 72°C for 1 min, with a final extension of 10 min at 72°C. To test the utility of the primers, PCR products were detected on 1% agarose gels. Finally, a total of 15 of the 30 primer pairs were successfully amplified.

CONCLUSIONS

The 15 microsatellite markers developed here have been proven to be useful in the evaluation of the genetic diversity of B. racemosa. Cross-amplification of these microsatellite loci in B. asiatica and B. acutangula suggests that they will also be useful in studies of other species within Barringtonia. We are currently using these markers for inferring genetic diversity and spatial genetic structure in these species. In particular, we are using these markers to assess the impact of seed dispersal by water and to estimate seed dispersal distances for these species in coastal tropical landscapes of the Indo-West Pacific.
APPENDIX 1. Voucher and location information for the species and populations used in this study. All voucher specimens are deposited at the herbarium of Sun Yat-sen University (SYSU), Guangzhou, China.

| Species         | Population code | Voucher no. | Collection locality              | Geographic coordinates          | N |
|-----------------|-----------------|-------------|----------------------------------|---------------------------------|---|
| *Barringtonia racemosa* | BRSH            | ZCR2013081  | Hainan, China                    | 19°32'58.2"N, 110°34'22"E       | 19 |
|                 | BRKL            | HYL2014154  | Sabah, Malaysia                  | 5°27'1.44"N, 115°36'47.88"E     | 15 |
|                 | BRDR            | SSH2012112  | Daintree River, Australia        | 12°23'0.6"S, 130°51'58.68"E     | 9  |
| *B. asiatica*   | BASB            | HSN2014026  | Palawan, Philippines             | 10°11'48.48"N, 118°54'16.92"E   | 7  |
| *B. acutangula* | BACC            | LY2013216   | Kampong Pluck, Cambodia          | 13°12'33.12"N, 103°58'25.68"E   | 7  |

Note: N = number of individuals.