Development of Microsatellite Markers for Carallia brachiata (Rhizophoraceae)

Authors: Qiang, Yinmeng, Xie, Hongxian, Qiao, Sitan, Yuan, Yang, Liu, Ying, et al.

Source: Applications in Plant Sciences, 3(3)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1400125

Applications in Plant Sciences 2015 3(3): 1400125

PRIMER NOTE

DEVELOPMENT OF MICROSATELLITE MARKERS FOR *CARALLIA BRACHIATA* (RHIZOPHORACEAE)[1]

YINMENG QIANG[2,4], HONGXIAN XIE[2,4], SITAN QIAO[2,4], YANG YUAN[2], YING LIU[2], XIANGGANG SHI[2], MI SHU[2], JIANHUA JIN[2], SUHUA SHI[2], FENGXIAO TAN[3], AND YELIN HUANG[2,5]

[2]Guangdong Provincial Key Laboratory of Plant Resources and State Key Laboratory of Biocontrol, Sun Yat-sen University, Guangzhou 510275, Guangdong, People's Republic of China; and [3]South China Agricultural University, Guangzhou 510642, Guangdong, People's Republic of China

- *Premise of the study:* Microsatellite markers were developed for *Carallia brachiata* to assess the genetic diversity and structure of this terrestrial species of the Rhizophoraceae.
- *Methods and Results:* Based on transcriptome data for *C. brachiata*, 40 primer pairs were initially designed and tested, of which 18 were successfully amplified and 11 were polymorphic. For these microsatellites, one to three alleles per locus were identified. The observed and expected heterozygosities ranged from 0 to 0.727 and 0 to 0.520, respectively. In addition, all primers were successfully amplified in two congeners: *C. pectinifolia* and *C. garciniifolia*.
- *Conclusions:* The microsatellite markers described here will be useful in population genetic studies of *C. brachiata* and related species, suggesting that developing microsatellite markers from next-generation sequencing data can be efficient for genetic studies across this genus.

Key words: *Carallia brachiata*; genetic diversity; microsatellite marker; Rhizophoraceae; transcriptome.

The Rhizophoraceae comprise three tribes, 15 genera, and ca. 140 species (Schwarzbach and Ricklefs, 2000). Although often described as a mangrove family, only members of tribe Rhizophoreae, which includes four genera and 16 species, live exclusively in intertidal habitats. Mangroves differ from their terrestrial relatives in that the former are characterized by their coastal habitat and peculiar adaptive traits, such as their unusual adaptive viviparous fruits and their pneumatophores and knee roots (Zhong et al., 2002). Population genetic markers such as microsatellite loci have been developed in many Rhizophoraceae mangroves (Triest, 2008). However, less attention has been paid to the population genetics of closely related terrestrial groups (Zhong et al., 2002).

Carallia Roxb., known as the freshwater mangrove, is in the tribe Gynotrocheae, which is the closest inland relative to Rhizophoreae. *Carallia* is fairly common in evergreen forests of the tropical Old World, especially along rivers (Schwarzbach and Ricklefs, 2000; Shi et al., 2002). This genus comprises approximately 10 species, four of which are found in China, namely *C. brachiata* (Lour.) Merr., *C. pectinifolia* W. C. Ko, *C. garciniifolia* F. C. How & C. N. Ho, and *C. diplopetala* Hand.-Mazz. (Qin and Boufford, 2007). *Carallia brachiata* is the most common species in this genus and is geographically widespread from Madagascar throughout tropical Asia to Australia (Queensland), Melanesia, and Micronesia (Hou, 1958). Recently, we have sequenced the leaf transcriptome of *C. brachiata* to investigate genetic characteristics of this species (Huang et al., unpublished data). Here we developed and characterized 18 expressed sequence tags–derived simple sequence repeat (EST-SSR) markers in *C. brachiata*, and tested the potential utility of those markers across species within *Carallia*. As Ellis and Burke (2007) proposed, EST-SSRs are easier and less expensive to develop, as well as more transferable across taxonomic boundaries, in comparison to genomic SSR markers.

METHODS AND RESULTS

Using an improved cetyltrimethylammonium bromide (CTAB) method (Fu et al., 2005), total RNA was extracted from the leaf of one individual of *C. brachiata* collected from Baiyun Mountain (23°10′47″N, 113°17′50″E) in Guangzhou, Guangdong, China. A voucher specimen (*L. Hu 110154*) was deposited in the herbarium of Sun Yat-sen University (SYSU). RNA integrity was checked using 1.0% agarose gel electrophoresis and a 2100 Bioanalyzer (Agilent, Santa Clara, California, USA) analysis. 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) at BGI-Shenzhen, China (Huang et al., 2012). A total of 13.66 million 90-nucleotide paired-end reads were obtained and assembled using Trinity (release 20110519) with the default parameters (Grabherr et al., 2011). Redundant sequences (minimum identity = 99%) were removed using CAP3 (Huang and Madan, 1999). The analysis yielded 47,788 contigs with an average length of 695 bp, an N50 length of 1077, and an average depth of coverage of 29.4×.

[1]Manuscript received 30 December 2014; revision accepted 3 February 2015.

The authors thank Qiang Fan and Yu Liu for their assistance in collecting plant materials. This study was supported by grants from the National Natural Science Foundation of China (41276107, 31200466, J1210074, J1310025), Chang Hungta Science Foundation of Sun Yat-Sen University, and the Ministry of Education of China for Sun Yat-sen University–Heishiding Nature Reserve Biological Field Practice Project.

[4]These authors contributed equally to this work.

[5]Author for correspondence: lsshyl@mail.sysu.edu.cn

doi:10.3732/apps.1400125

To detect SSRs in the contigs, the software MISA (Thiel et al., 2003) was used with the following search criteria: six, five, five, five, and five repeat units for di-, tri-, tetra-, penta-, and hexanucleotide motifs, respectively. A total of 6114 SSRs were identified from those contigs, with 798 contigs containing more than one SSR and 375 contigs containing compound SSRs. The frequency of EST-SSRs observed in the *C. brachiata* transcriptome was 1.07%, and the distribution density was 184.06 per mega base pair. The most abundant repeat type was trinucleotide (3326 [54.40%]), followed by dinucleotide (2612 [42.72%]), tetranucleotide (151 [2.47%]), pentanucleotide (13 [0.21%]), and hexanucleotide (12 [0.20%]) repeat units. Using Primer3 (Rozen and Skaletsky, 1999), primer pairs were designed for 40 SSR loci whose PRIMER_PAIR_PENALTY was <0.2, with the following parameters: PRIMER_PRODUCT_SIZE_RANGE = 100–300 bp, PRIMER_MAX_END_STABILITY = 250, the start base of the TARGET = three bases before the start position of SSR, and the length of the TARGET = six bases longer than the size of SSR. Default values were used for all other parameters.

To assess variability among these 40 loci, 35 individuals of *C. brachiata* were sampled from three populations: Zhuhai, Guangdong (22°15′17″N, 113°16′13″E; *Q. Fan 090108*); Yangchun, Guangdong (21°53′17″N, 111°22′59″E; *Y. Liu 140318*); and Wenchang, Hainan (19°31′56″N, 110°44′40″E; *Q. Fan 121202*). In addition, individuals from one population of *C. pectinifolia* (Heishiding Natural Reserve, Fengkai, Guangdong; 23°31′12″N, 111°52′13″E; *Y. Liu 090718*) and one population of *C. garciniifolia* (Xishuangbanna, Yunnan; 21°55′50″N, 101°15′09″E; *Y. Huang 140211*) were also sampled to detect the efficiency of these markers in cross-species amplification. All voucher specimens were deposited at SYSU.

All sampled leaves were dried by silica gel and then total genomic DNA from each individual was isolated using the CTAB method (Doyle, 1991). PCR amplifications were performed in a final volume of 30 µL, containing 60 ng of genomic DNA, 1× PCR buffer (10 mM Tris-HCl [pH 8.4] and 1.5 mM $MgCl_2$; TransGen Biotech Co., Beijing, China), 0.2 mM dNTPs (Bocai Biotech Co., Shanghai, China), 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) with the following cycling conditions: initial denaturation at 94°C for 4 min, followed by 32 cycles of 94°C for 40 s, 56–60°C for 30 s, and 72°C for 30 s, with a final extension of 10 min at 72°C. The annealing temperatures ranged from 56°C to 60°C for different primer pairs. Amplification products were electrophoresed through 8% polyacrylamide denaturing gels and visualized by silver staining. The band size was calculated by comparison with a 20-bp DNA ladder (TaKaRa Biotechnology Co., Dalian, China).

Of the 40 primer pairs, 22 failed to amplify the expected products. Of the remaining 18 loci, 11 displayed clear polymorphisms in *C. brachiata* (Tables 1 and 2). The presence of null alleles was detected by MICRO-CHECKER version 2.2.3 (van Oosterhout et al., 2004) for all loci. Genetic diversity indices were calculated using the software POPGENE (Yeh and Boyle, 1997), including the number of alleles, observed heterozygosity, and expected heterozygosity (Table 2). The number of alleles per locus ranged from one to three in each population. The observed heterozygosity ranged from 0 to 0.727, and the expected heterozygosity ranged from 0 to 0.520 (Table 2). All 18 primer sets also successfully amplified SSR loci in *C. pectinifolia* and *C. garciniifolia*. All 18 SSR loci were at Hardy–Weinberg equilibrium and had no indication of null alleles. There was also no significant linkage equilibrium ($P < 0.05$) between locus pairs (Table 2).

CONCLUSIONS

Our results have shown that transcriptome sequencing is a valuable source of microsatellite markers in *C. brachiata* as

Table 1. Characteristics of 18 microsatellite loci developed in *Carallia brachiata*.

Locus	Primer sequences (5′–3′)	Repeat motif	Allele size (bp)	T_a (°C)	GenBank accession no.	Putative function
CBSSR03	F: TCCTCCTCCAAGAAAGGGAT	$(CTT)_5$	182	60	KM921762	Hypothetical protein
	R: TAACAAACAGTCGCTCACCG					
CBSSR05	F: ATCAACCACTTGGAGATGCC	$(AG)_6$	219	60	KM921764	Serine-threonine protein kinase, plant-type
	R: GCGATACATGTAAACGGCCT					
CBSSR09	F: AAGCGAAAGCGAGGTAACAA	$(AAT)_5$	198	60	KM921763	Bromodomain-containing protein
	R: CATTCGGGAAGCGTATTCAT					
CBSSR11	F: CCTGTAGGCATCTTACCCCA	$(CAA)_5$	212	60	KM921747	Hydrolase, hydrolyzing O-glycosylcompounds, putative
	R: TGCTGGGAGGGATTAACAAC					
CBSSR13	F: TTTCAGCGAGCTCAGGTTTT	$(GA)_7$	196	60	KM921748	—
	R: ACCCAGCATTACACGAGTCC					
CBSSR14	F: CAGGTTGTGGAAGTGGGACT	$(GTG)_5$	223	60	KM921749	Ninja-family protein mc410-like
	R: AACCGCAAAAATCGACATTC					
CBSSR15	F: ATCAACCACTTGGAGATGCC	$(AG)_6$	219	60	KM921750	Probable LRR receptor-like serine/threonine-protein kinase At3g47570-like
	R: GCGATACATGTAAACGGCCT					
CBSSR18	F: TCGTGGCCTTAGCTTCTTGT	$(GGT)_5$	161	58	KM921751	Programmed cell death 2 C-terminal domain-containing family protein
	R: CAGTCAGGGAGACCACCAAT					
CBSSR19	F: GCTGGAGTTCTTCCTCAACG	$(GAT)_6$	155	58	KM921752	Preprotein translocase secA subunit
	R: AAATCGGCCAAGATTGTGAC					
CBSSR20	F: GTTGCTCACCCAGCAATTTT	$(AG)_6$	133	58	KM921753	Basic helix-loop-helix DNA-binding superfamily protein
	R: TTCCTTCCCCAAACTGAGTG					
CBSSR22	F: TAATGCTATGCTGTGCCTGC	$(AAT)_5$	244	58	KM921754	Hypothetical protein
	R: TTGAAGGCGGTGAGACTTTT					
CBSSR23	F: AGAAGAGGCTGGGAATGGAT	$(TTC)_5$	267	58	KM921755	RNA-binding protein 38-like
	R: AACCCACAAGTTCAACAGCC					
CBSSR25	F: CTTCCCAAAGCTTCTCGTTG	$(TCC)_5$	247	58	KM921756	Glucan endo-1,3-beta-glucosidase precursor, putative
	R: AAGTTTGCAACTTGGGATGG					
CBSSR27	F: GGGTTGTGATTTCTGATGGG	$(AGG)_5$	222	56	KM921757	ATP-binding protein, putative
	R: TCACTTTCACATCCTGCTGC					
CBSSR28	F: AAGTTGCACTCATCCCGAAC	$(TC)_7$	216	56	KM921758	Ring finger protein, putative
	R: TCCGTTTGAAGGGACATAGG					
CBSSR30	F: CCAGAACTTGTAGCGCATGA	$(TGG)_6$	214	56	KM921759	Zinc finger family protein
	R: GAATTGCAACTGTAACGCGA					
CBSSR32	F: CGCTAACCGCTCTCTAATCG	$(TC)_7$	183	56	KM921760	Ubiquitin-activating enzyme E1, putative
	R: GGTTAAGGCTAGGTTTCGGG					
CBSSR37	F: GTCCGTCTCCGAAATCAAAA	$(CAG)_6$	223	56	KM921761	Deoxyuridine 5′-triphosphate nucleotidohydrolase-like
	R: CATGGGGCATTGAGAGACTT					

Note: F = forward primer; R = reverse primer; T_a = annealing temperature.

TABLE 2. Results of initial primer screening in populations of *Carallia brachiata*, *C. pectinifolia*, and *C. garciniifolia.*

	C. brachiata												*C. pectinifolia*				*C. garciniifolia*			
	Zhuhai (N = 12)				Yangchun (N = 11)				Wenchang (N = 12)				Heishiding (N = 12)				Xishuangbanna (N = 11)			
Locus	A	H_o	H_e	P	A	H_o	H_e	P	A	H_o	H_e	P	A	H_o	H_e	P	A	H_o	H_e	P
CBSSR03	2	0.273	0.247	0.675	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—
CBSSR05	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—
CBSSR09	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—
CBSSR11	2	0.250	0.228	0.692	2	0.273	0.368	0.346	2	0.083	0.083	1.000	1	0.000	0.000	—	2	0.091	0.091	1.000
CBSSR13	2	0.583	0.431	0.192	2	0.546	0.416	0.264	2	0.417	0.489	0.590	1	0.000	0.000	—	1	0.000	0.000	—
CBSSR14	1	0.000	0.000	—	2	0.091	0.091	1.000	1	0.000	0.000	—	2	0.583	0.431	0.192	2	0.182	0.173	0.819
CBSSR15	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—
CBSSR18	2	0.250	0.228	0.692	3	0.455	0.385	0.865	2	0.083	0.083	1.000	2	0.250	0.228	0.692	2	0.455	0.456	1.000
CBSSR19	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—
CBSSR20	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—
CBSSR22	2	0.083	0.083	1.000	2	0.727	0.520	0.164	2	0.250	0.228	0.692	2	0.250	0.344	0.297	2	0.364	0.312	0.531
CBSSR23	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—
CBSSR25	2	0.500	0.464	0.775	2	0.091	0.091	1.000	2	0.250	0.344	0.297	1	0.000	0.000	—	1	0.000	0.000	—
CBSSR27	1	0.000	0.000	—	2	0.182	0.173	0.819	2	0.667	0.464	0.109	2	0.500	0.391	0.299	1	0.000	0.000	—
CBSSR28	1	0.000	0.000	—	2	0.091	0.091	1.000	2	0.167	0.159	0.827	1	0.000	0.000	—	1	0.000	0.000	—
CBSSR30	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—	1	0.000	0.000	—
CBSSR32	1	0.000	0.000	—	2	0.546	0.416	0.264	2	0.333	0.290	0.556	2	0.250	0.431	0.121	2	0.636	0.507	0.371
CBSSR37	2	0.500	0.464	0.775	2	0.727	0.485	0.079	2	0.250	0.228	0.692	2	0.250	0.228	0.692	1	0.000	0.000	—

Note: A = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; N = number of individuals in the population sampled; P = P values for deviation from Hardy–Weinberg equilibrium.

well as across the genus. These newly developed EST-SSRs will be valuable resources for the investigation of population genetic variation and structure in terrestrial relatives of Rhizophoraceae mangroves.

LITERATURE CITED

DOYLE, J. J. 1991. DNA protocols for plants—CTAB total DNA isolation. *In* G. M. Hewitt and A. Johnston [eds.], Molecular techniques in taxonomy, 283–293. Springer-Verlag, Berlin, Germany.

ELLIS, J. R., AND J. M. BURKE. 2007. EST-SSRs as a resource for population genetic analyses. *Heredity* 99: 125–132.

FU, X. H., Y. L. HUANG, S. L. DENG, R. C. ZHOU, G. L. YANG, X. W. NI, W. J. LI, AND S. H. SHI. 2005. Construction of a SSH library of *Aegiceras corniculatum* under salt stress and expression analysis of four transcripts. *Plant Science* 169: 147–154.

GRABHERR, M. G., B. J. HAAS, M. YASSOUR, J. Z. LEVIN, D. A. THOMPSON, I. AMIT, X. ADICONIS, ET AL. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644–652.

HOU, D. 1958. Rhizophoraceae. *In* C. G. G. J. van Steenis [ed.], Flora Malesiana, Series 1, vol. 5, 429–493. Noordhoff-Kolff, Djakarta, Indonesia.

HUANG, X. Q., AND A. MADAN. 1999. CAP3: A DNA sequence assembly program. *Genome Research* 9: 868–877.

HUANG, Y. L., X. T. FANG, L. LU, Y. B. YAN, S. F. CHEN, C. C. ZHU, L. HU, ET AL. 2012. Transcriptome analysis of an invasive weed *Mikania micrantha. Biologia Plantarum* 56: 111–116.

QIN, H. N., AND D. E. BOUFFORD. 2007. Rhizophoraceae. *In* Z. Y. Wu and P. H. Raven [eds.], Flora of China, vol. 13, 295–299. Science Press, Beijing, China, and Missouri Botanical Garden Press, St. Louis, Missouri, USA.

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.

SCHWARZBACH, A. E., AND R. E. RICKLEFS. 2000. Systematic affinities of Rhizophoraceae and Anisophylleaceae, and intergeneric relationships within Rhizophoraceae, based on chloroplast DNA, nuclear ribosomal DNA, and morphology. *American Journal of Botany* 87: 547–564.

SHI, S. H., Y. ZHONG, Y. L. HUANG, X. Z. QIU, AND H. T. CHANG. 2002. Phylogenetic relationships of the Rhizophoraceae in China based on sequences of the chloroplast gene *matK* and ITS regions of nuclear ribosomal DNA and combined data set. *Biochemical Systematics and Ecology* 30: 309–319.

THIEL, T., W. MICHALEK, R. K. VARSHNEY, AND A. GRANER. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 10: 411–422.

TRIEST, L. 2008. Molecular ecology and biogeography of mangrove trees towards conceptual insights on gene flow and barriers: A review. *Aquatic Botany* 89: 138–154.

VAN OOSTERHOUT, C., W. F. HUTCHINSON, D. P. M. WILLS, AND P. SHIPLEY. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.

YEH, F. C., AND T. J. B. BOYLE. 1997. Population genetic analysis of codominant and dominant markers and quantitative traits. *Belgian Journal of Botany* 129: 157.

ZHONG, Y., Q. ZHAO, S. H. SHI, Y. L. HUANG, AND M. HASEGAWA. 2002. Detecting evolutionary rate heterogeneity among mangroves and their close terrestrial relatives. *Ecology Letters* 5: 427–432.