Applications in Plant Sciences 2016 4(1): 1500091

Primer Note

Development and characterization of polymorphic microRNA-based microsatellite markers in *Nelumbo nucifera* (Nelumbonaceae)[1]

Xiaolei Wang[2], Songtao Gui[2], Lei Pan[3], Jihong Hu[2,4], and Yi Ding[2,4]

[2]State Key Laboratory of Hybrid Rice, College of Life Sciences, Wuhan University, Wuhan 430072, People's Republic of China; and [3]Hubei Province Engineering Research Center of Legume Plants, School of Life Sciences, Jianghan University, Wuhan 430056, People's Republic of China

- *Premise of the study:* Polymorphic microRNA (miRNA)–based microsatellite markers were developed to investigate the genetic diversity and population structure of *Nelumbo nucifera* (Nelumbonaceae).
- *Methods and Results:* A total of 485 miRNA-based microsatellites were found from the genomic DNA sequences of *N. nucifera*. After several rounds of screening, 21 primer pairs flanking di-, tri-, or pentanucleotide repeats were identified that revealed high levels of genetic diversity in four populations with two to five alleles per locus. The observed and expected heterozygosity per locus ranged from 0.000 to 1.000 and from 0.000 to 0.803, respectively.
- *Conclusions:* The polymorphic microsatellite markers will be useful for studying the genetic diversity and population structure of *N. nucifera*.

Key words: genetic diversity; microRNA (miRNA); microsatellites; *Nelumbo nucifera*; Nelumbonaceae; polymorphism.

Sacred lotus (*Nelumbo nucifera* Gaertn.) ($2x = 2n = 16$), an aquatic perennial plant in the family Nelumbonaceae, has been cultivated as an ornamental or vegetable plant for more than 7000 yr throughout Asia (Hu et al., 2012; Yang et al., 2015). Microsatellite (simple sequence repeat [SSR]) markers are sensitive tools for evaluating genetic diversity, population genetic structure, and intraspecific variation. Because microsatellites can be either intergenic or intragenic (Tóth et al., 2000), the variable length of repeat motifs at the SSR may be related to differential function or activity of the segments of chromosomes in which they reside. MicroRNAs (miRNAs) are ca. 21-nucleotide, noncoding, small RNAs that play an important role in gene expression under diverse stress conditions including various biotic as well as abiotic stresses (Bartel, 2004). miRNA-based SSR (miRNA-SSRs) markers are a novel type of functional marker exploited predominantly in animal sciences, but little reported in plants. In this study, we performed a genome-wide analysis of miRNA-SSRs in *N. nucifera* and validated 45 SSRs among the 36 genotypes. This is the first report of genome-wide identification and characterization of miRNA-SSRs in *N. nucifera*.

[1]Manuscript received 2 August 2015; revision accepted 10 September 2015.

The authors thank Weidong Ke (Institute of Wuhan Vegetable Scientific Research) for providing the sacred lotus materials. This work was supported by the National Natural Science Foundation of China (31271310).

[4]Authors for correspondence: hujh2010@163.com, yiding@whu.edu.cn

doi:10.3732/apps.1500091

METHODS AND RESULTS

The 106 *N. nucifera* pre-miRNA sequences identified in our previous study (Pan et al., 2015) were used for the present investigation. The 1000-bp (500 bp upstream and 500 bp downstream of mature miRNA sequence) sequences were obtained from the sacred lotus reference genome (Ming et al., 2013). The miRNA-SSR loci distributed throughout the *N. nucifera* genome were screened using MISA (Thiel et al., 2003) with default parameters. SSRs were selected based on the length of the core repeat motif ≥10 nucleotides (e.g., five units of dinucleotide repeat motifs, four units of trinucleotide repeat motifs, or three units of tetranucleotide repeat motifs). A total of 485 miRNA-based SSRs were present in the genome of *N. nucifera*. Using the MISA output, primers of each of the SSR-containing sequences were designed using the program BatchPrimer3 (http://probes.pw.usda.gov/batchprimer3) (You et al., 2008). The parameters of each primer were set using the following criteria: (1) primer size of 18–22 nucleotides in length; (2) GC content of 40–60%; (3) annealing temperature between 50°C and 60°C (55°C optimum); and (4) expected amplicon size of 100–300 bp. In total, 138 miRNA-SSR primer pairs of *N. nucifera* were designed, and 45 primer pairs were synthesized for further analysis (GenScript, Nanjing, China).

Thirty-six *N. nucifera* accessions were used in the current study (Appendix 1). Total genomic DNA was isolated from frozen young leaves using the modified cetyltrimethylammonium bromide (CTAB) method as described in Doyle and Doyle (1987). A preliminary study using 12 *N. nucifera* individuals from a population from Hubei (Appendix 1) resulted in the selection of 21 microsatellite loci (Table 1) that were polymorphic. The sequences of polymorphic microsatellite loci were deposited into GenBank (accession no. KT344795–KT344815; Table 1). PCR amplifications were performed in a 15-μL reaction containing 50–100 ng genomic DNA, 1.5 μL 10× PCR buffer, 0.4 μM for each primer, 1.5 mM $MgCl_2$, 250 μM each dNTP, and 0.5 units *Taq* DNA polymerase (TianGen, Beijing, China). The thermocycling conditions were: 95°C for 3 min; 35 cycles of 94°C for 30 s, annealing temperature optimized for each primer for 30 s (Table 1), and 72°C for 40 s; and a final extension step at 72°C for 7 min. The amplified products were separated on 6% denaturing polyacrylamide sequencing gels in 0.5× TBE buffer and visualized by silver nitrate staining. The size of fragments was determined using a 20-bp marker of 20–500 bp (TaKaRa Biotechnology Co., Dalian, China).

TABLE 1. Characteristics of 21 miRNA microsatellite loci and primer pairs developed in *Nelumbo nucifera*.

Locus	miRNA	Primer sequences (5′–3′)	Repeat motif	Allele size range (bp)	T_a (°C)	*A*	GenBank accession no.
NnmiR-SSR1	Nnu-miR156a	F: GCGATGCATGATGAAATGAC R: CCAACCAAGATAACGCATCA	$(CT)_7$	196–220	59	3	KT344795
NnmiR-SSR2	Nnu-miR156b	F: TCCACCACTCCGGCTATCTA R: GCAACGTTAAGTGCTGCAAA	$(TGCTT)_3$	176–182	60	3	KT344796
NnmiR-SSR3	Nnu-miR157a	F: TGCAAATAGATCCCCTTTGT R: GTGGAATGTTGGAGGTTTTT	$(AAT)_7$	179–200	56	4	KT344797
NnmiR-SSR4	Nnu-miR160a	F: TGGCTTATGCAAGAGTAGGTGA R: ACTGCCTGCCGTATATGTGA	$(TC)_8$	175–180	59	2	KT344798
NnmiR-SSR5	Nnu-miR160a	F: CGAGGAGCCATGCATATTG R: GACGATGCTGCTGCTTTATG	$(TTC)_7$	172–178	58	2	KT344799
NnmiR-SSR6	Nnu-miR160d	F: CAAGCAGCTAACATACCACGA R: GTCCCACACACCATGTGAAG	$(TA)_9$	160–166	58	4	KT344800
NnmiR-SSR7	Nnu-miR165a	F: CCTAAGTGACCTCGGACCAG R: CTGCAAGCCAGAATCAAACA	$(TC)_{10}$	180–186	59	2	KT344801
NnmiR-SSR8	Nnu-miR165b	F: TCATTCCCCTCAACCATGA R: ACCTCGAGCCAGACAACATT	$(TC)_7$	136–173	58	2	KT344802
NnmiR-SSR9	Nnu-miR171	F: CGGTACTGTTTTGCAGGTGA R: CCCGCCATTAATTCTCATCA	$(CT)_{12}$	200–208	60	2	KT344803
NnmiR-SSR10	Nnu-miR172a	F: CCCTCAGCTTCTCCTTTTCC R: CCCATCTTCTCAACCTTCCA	$(CT)_{17}$	128–138	60	3	KT344804
NnmiR-SSR11	Nnu-miR396a	F: GCAAAGCTCCATTTCACCTT R: AGCTGTGGAAAAGCATGACA	$(CT)_{17}$	193–210	58	5	KT344805
NnmiR-SSR12	Nnu-miR828	F: TCTCTATGGATGAAGCACCAGA R: AAGCAGAGCTCCCCAACATA	$(CT)_{11}$	162–183	59	4	KT344806
NnmiR-SSR13	Nnu-miR4414a	F: TGCAAAGTCAGCAAAGAGGA R: GGATTGGACAAAGAGGGAAGA	$(GA)_{10}$	130–140	59	3	KT344807
NnmiR-SSR14	Nnu-miR4414c	F: TATTCTACGGCCCCTTACCC R: GGTCCTCTTGCTCTTGCATC	$(TC)_{12}$	145–152	60	2	KT344808
NnmiR-SSR15	Nnu-miR5227	F: ATGGCGAAACAGGGTTCATA R: TGTTTGCCTGGGGAATACAT	$(GAC)_4$	128–140	60	2	KT344809
NnmiR-SSR16	Nnu-miR157d	F: GAGGTGTCTGGGGACTCTCTT R: AGTGCCTTCTCTGTCCCTTG	$(CT)_{15}$	136–170	58	3	KT344810
NnmiR-SSR17	Nnu-miR157d	F: TGTGGTCTTGGCTGAATGAA R: GAAATGGAACTTTTCCCCACT	$(TA)_{13}$	150–170	59	3	KT344811
NnmiR-SSR18	Nnu-miR165a	F: TTTTATGGGCTTGCTCGTTT R: CACAGCAAGTCAGAATCAAACA	$(TC)_{16}$	135–145	58	3	KT344812
NnmiR-SSR19	Nnu-miR169b	F: CCAAAGTTCTTCCCTTGAAACA R: TGAGTTCTGCAAGGGCTTCT	$(AAT)_{12}$	252–260	60	5	KT344813
NnmiR-SSR20	Nnu-miR172b	F: TCTCAAGGCACCAGTCAGTG R: TGCAGCATCATCAAGATTCC	$(TCCCT)_4$	120–140	59	2	KT344814
NnmiR-SSR21	Nnu-miR319b	F: TTGTAGATGCATGGGTTCTGTC R: GCTCCCTTCAGTCCAAAACA	$(TC)_{21}$	170–190	60	3	KT344815

Note: *A* = number of alleles per locus; T_a = optimal annealing temperature.

TABLE 2. Genetic properties of 14 polymorphic miRNA-SSR markers in four populations of *Nelumbo nucifera*.[a]

	Jiangxi population (N = 5)			Hunan population (N = 6)			Fujian population (N = 3)			Hubei population (N = 22)		
Locus	A	H_o	H_e[b]	A	H_o	H_e[b]	A	H_o	H_e	A	H_o	H_e[b]
NnmiR-SSR1	2	0.500	0.429	3	0.000	0.545**	1	0.000	0.000	3	0.048	0.675**
NnmiR-SSR5	2	0.600	0.467	4	0.667	0.591	1	0.000	0.000	3	0.300	0.600**
NnmiR-SSR7	2	0.000	0.356	2	0.000	0.485*	2	0.000	0.667	2	0.000	0.406**
NnmiR-SSR8	3	0.600	0.511	3	0.600	0.467	2	0.500	0.500	4	0.364	0.732**
NnmiR-SSR9	3	0.600	0.511	3	0.500	0.440	2	0.667	0.533	3	0.150	0.591**
NnmiR-SSR10	2	0.000	0.356	2	0.000	0.485*	2	0.000	0.667	2	0.046	0.460**
NnmiR-SSR11	2	0.600	0.467	3	0.667	0.712	2	1.000	0.600	3	0.227	0.606**
NnmiR-SSR12	2	0.750	0.536	2	0.000	0.356	2	0.333	0.333	4	0.526	0.668**
NnmiR-SSR13	2	0.000	0.533*	2	0.000	0.485	2	0.000	0.533	2	0.048	0.512**
NnmiR-SSR14	2	0.600	0.467	3	0.167	0.621*	2	0.333	0.333	4	0.300	0.413*
NnmiR-SSR15	2	0.600	0.467	2	0.500	0.530	3	0.667	0.733	3	0.524	0.605
NnmiR-SSR16	2	0.800	0.533	4	0.667	0.803*	2	0.000	0.400	4	0.105	0.694**
NnmiR-SSR17	3	0.000	0.711**	4	0.000	0.800**	2	0.000	0.533	5	0.350	0.744**
NnmiR-SSR19	2	0.200	0.200	3	0.167	0.621*	2	1.000	0.600	2	0.045	0.333**
Mean	2.21	0.418	0.467	2.86	0.281	0.567	1.93	0.321	0.459	3.14	0.217	0.574

Note: A = total number of alleles per locus; H_e = expected heterozygosity; H_o = observed heterozygosity; N = sample size for each population.

[a] See Appendix 1 for population locality information.

[b] Deviations from Hardy–Weinberg equilibrium: *$P < 0.05$, **$P < 0.01$.

Fourteen polymorphic SSR primers were used to genotype 36 individuals of *N. nucifera* collected from Jiangxi Province (N = 5; 1°17′N, 103°50′E), Hunan Province (N = 6; 26°54′N, 112°36′E), Fujian Province (N = 3; 26°15′N, 117°37′E), and Hubei Province (N = 22; 30°34′N, 116°16′E). Voucher specimens were deposited in the Wuhan National Field Observation and Research Station for Aquatic Vegetables (Appendix 1). Parameters of genetic diversity including number of alleles (A), observed heterozygosity (H_o), expected heterozygosity (H_e), and Hardy–Weinberg equilibrium (HWE) were determined by Arlequin version 3.5.1.2 (Excoffier et al., 2005). Each of the 14 loci exhibited two to five alleles among the 36 *N. nucifera* individuals, with H_o and H_e ranging from 0.000 to 1.000 and from 0.000 to 0.803, respectively (Table 2). A relatively high level of genetic diversity was found in the Hubei population (H_o = 0.217, A = 3.14) compared with the other three populations. This may be due to the fact that we sampled more individuals from the Hubei population. Some loci showed significant deviation from HWE (Table 2) due to heterozygote deficiency.

CONCLUSIONS

We developed a novel set of 21 miRNA-based SSR markers for *N. nucifera*. These markers will enable researchers to estimate the genetic diversity and genetic structure of populations of *N. nucifera*. They may also be used as a novel genotyping tool for plant molecular breeding.

LITERATURE CITED

BARTEL, D. P. 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116: 281–297.

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

EXCOFFIER, L., G. LAVAL, AND S. SCHNEIDER. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47–50.

HU, J. H., L. PAN, H. G. LIU, S. Z. WANG, Z. H. WU, W. D. KE, AND Y. DING. 2012. Comparative analysis of genetic diversity in sacred lotus (*Nelumbo nucifera* Gaertn.) using AFLP and SSR markers. *Molecular Biology Reports* 39: 3637–3647.

MING, R., V. B. ROBERT, Y. L. LIU, M. YANG, Y. P. HAN, L. T. LI, Q. ZHANG, ET AL. 2013. Genome of the long-living sacred lotus (*Nelumbo nucifera* Gaertn.). *Genome Biology* 14: R41.

PAN, L., X. L. WANG, J. JIN, X. L. YU, AND J. H. HU. 2015. Bioinformatic identification and expression analysis of *Nelumbo nucifera* microRNA and their targets. *Applications in Plant Sciences* 3: 1500046.

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.

TÓTH, G., Z. GÁSPÁRI, AND J. JURKA. 2000. Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Research* 10: 967–981.

YANG, M., L. M. XU, Y. L. LIU, AND P. F. YANG. 2015. RNA-Seq uncovers SNPs and alternative splicing events in Asian lotus (*Nelumbo nucifera*). *PLoS One* 10: e0125702.

YOU, F. M., N. HUO, Y. Q. GU, M. C. LUO, Y. MA, D. HANE, G. R. LAZO, ET AL. 2008. BatchPrimer3: A high throughput web application for PCR and sequencing primer design. *BMC Bioinformatics* 9: 253.

Appendix 1. Voucher and location information for populations of *Nelumbo nucifera* used in the study. The voucher specimens are deposited in the Wuhan National Field Observation and Research Station for Aquatic Vegetables Herbarium (NOH).

Population code	Population locality	Voucher no.	*n*	Geographic coordinates
JX1	Fuzhou, Jiangxi Province, China	NOH-JX6	5	1°17′N, 103°50′E
HN2	Hengyang, Hunan Province, China	NOH-HN8	6	26°54′N, 112°36′E
FJ3	Sanming, Fujian Province, China	NOH-FJ4	3	26°15′N, 117°37′E
HB4	Wuhan, Hubei Province, China	NOH-HB50	22	30°34′N, 116°16′E

Note: *n* = number of individuals.