**DEVELOPMENT AND CHARACTERIZATION OF POLYMORPHIC MICRORNA-BASED MICROSATellite MARKERS IN *NELUMBO NUCIFERA* (NELUMBONACEAE)**

**XIAOLEI WANG**, **SONGTAO GUI**, **LEI PAN**, **JIHONG HU**, and **YI DING**

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- to 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*.

**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 MgCl₂, 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).


| Locus     | mRNA     | Primer sequences (5′ – 3′)       | Repeat motif  | Allele size range (bp) | $T_a$ (°C) | A | GenBank accession no. |
|-----------|----------|---------------------------------|---------------|------------------------|------------|---|----------------------|
| NnmiR-SSR1| Nnu-miR156a | F: GCCATGCAATGATGAATGAC         | (CT)$_7$      | 196–220                | 59         | 3 | KT344795             |
|           |          | R: GCAAACGAGCTGGGATATGGA       |               |                        |            |   |                      |
| NnmiR-SSR2| Nnu-miR156b | F: TGGGTGCTGGCCTGCTCTTA         | (TGCTT)$_3$   | 176–182                | 60         | 3 | KT344796             |
|           |          | R: GCAAATATGAGCTGGTGGAAA        |               |                        |            |   |                      |
| NnmiR-SSR3| Nnu-miR157a | F: TGGCAATAGATCCCTTTGGT         | (AAT)$_7$     | 179–200                | 56         | 4 | KT344797             |
|           |          | R: GTGGATTGTGGAGGTGGTTTTT       |               |                        |            |   |                      |
| NnmiR-SSR4| Nnu-miR160a | F: TGCTTATGCAGGTAGTTGGA         | (TC)$_8$      | 175–180                | 59         | 2 | KT344798             |
|           |          | R: CGAGCGCCAGCTGATTGGA          |               |                        |            |   |                      |
| NnmiR-SSR5| Nnu-miR160a | F: CGAGCGCCAGCTGATTGGA          | (TC)$_7$      | 172–178                | 58         | 2 | KT344799             |
|           |          | R: GACGGTGCTGCTGCTTTAGG         |               |                        |            |   |                      |
| NnmiR-SSR6| Nnu-miR160d | F: CAGCGATCATACATCCGACGA       | (TA)$_9$      | 160–166                | 58         | 4 | KT344800             |
|           |          | R: GTCCCAGCAGCGATGAGGATGG        |               |                        |            |   |                      |
| NnmiR-SSR7| Nnu-miR165a | F: CCTATGACCTCGGACAGACG         | (TC)$_9$      | 180–186                | 59         | 2 | KT344801             |
|           |          | R: CTTGAAAGCGAAACATCA           |               |                        |            |   |                      |
| NnmiR-SSR8| Nnu-miR165b | F: TCATGGCTTACCTCACTCA          | (TC)$_7$      | 136–173                | 58         | 2 | KT344802             |
|           |          | R: ACCCTGAGCGAGCAAGCTAT          |               |                        |            |   |                      |
| NnmiR-SSR9| Nnu-miR171 | F: CGTACTGTTGTTGGAGGTA          | (CT)$_12$     | 200–208                | 60         | 2 | KT344803             |
|           |          | R: CCGCCATATTATCTCTGAC          |               |                        |            |   |                      |
| NnmiR-SSR10| Nnu-miR172a | F: CCTCAGCTTCTCTTTCTC          | (CT)$_7$      | 128–138                | 60         | 3 | KT344804             |
|           |          | R: CCGATCTTCACTCTTCCG           |               |                        |            |   |                      |
| NnmiR-SSR11| Nnu-miR396a | F: GCAGATCGCTATCCCTTCT         | (CT)$_7$      | 193–210                | 58         | 5 | KT344805             |
|           |          | R: AGCTTGAGGAAGGGCGTA           |               |                        |            |   |                      |
| NnmiR-SSR12| Nnu-miR828 | F: TCTCTATGGTGAGAGGACGGA       | (CT)$_11$     | 162–183                | 59         | 4 | KT344806             |
|           |          | R: AAAAGCTGGCTCTCTCTC          |               |                        |            |   |                      |
| NnmiR-SSR13| Nnu-miR441a | F: TGCAATGCGCAAAAGGGAGA         | (GA)$_10$    | 130–140                | 59         | 3 | KT344807             |
|           |          | R: AGCTATGGAGCCAGAGGCGA         |               |                        |            |   |                      |
| NnmiR-SSR14| Nnu-miR441c | F: TATACGTACGCTCTTCC          | (TC)$_12$     | 145–152                | 60         | 2 | KT344808             |
|           |          | R: GTTCCTTGGTGCTCTGATC         |               |                        |            |   |                      |
| NnmiR-SSR15| Nnu-miR5227 | F: ATGGCATGAGAGGCTCTAT         | (GA)$_4$      | 128–140                | 60         | 2 | KT344809             |
|           |          | R: TGGGTGCTGGGGAAATCAT          |               |                        |            |   |                      |
| NnmiR-SSR16| Nnu-miR157d | F: GAGCTGCTGCTGCTCTTGG         | (CT)$_5$      | 136–170                | 58         | 3 | KT344810             |
|           |          | R: AGGCTCTCTCTCTTCCTT          |               |                        |            |   |                      |
| NnmiR-SSR17| Nnu-miR157d | F: TGGGTGCTGGGGGGTAGGGA       | (TA)$_12$     | 150–170                | 59         | 3 | KT344811             |
|           |          | R: GAAATGGGACTTTTTCTCC          |               |                        |            |   |                      |
| NnmiR-SSR18| Nnu-miR165a | F: TTTTATGGGCTGGCTCTTCTT       | (TC)$_10$     | 135–145                | 58         | 3 | KT344812             |
|           |          | R: CCAACAGCTAGAACTCA           |               |                        |            |   |                      |
| NnmiR-SSR19| Nnu-miR169b | F: TGAGCTCAGCAAGGTGGTTG         | (AAT)$_12$    | 252–260                | 60         | 5 | KT344813             |
|           |          | R: TGGATGCTCAGGGGTCTCCCTG       |               |                        |            |   |                      |
| NnmiR-SSR20| Nnu-miR172b | F: TCCAGCATCAGACAGTTCC         | (TCCCT)$_4$   | 120–140                | 59         | 2 | KT344814             |
|           |          | R: TGGATGCTCAGGGGTCTCCCTT       |               |                        |            |   |                      |
| NnmiR-SSR21| Nnu-miR319b | F: TGATGTCAGGTGGGCTCTG         | (TC)$_7$      | 170–190                | 60         | 3 | KT344815             |
|           |          | R: GCCCTCTTGGCTGCAAAACG         |               |                        |            |   |                      |

*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*.

| Locus | Hubei population (N = 6) | Fujian population (N = 5) | Jiangxi population (N = 5) |
|-------|--------------------------|---------------------------|---------------------------|
|       | $H_o$ | $H_e$ | $H_o$ | $H_e$ | $H_o$ | $H_e$ | $H_o$ | $H_e$ |
| NnmiR-SSR1 | 0.500 | 0.429 | 0.500 | 0.429 | 0.500 | 0.429 |
| NnmiR-SSR5 | 0.380 | 0.294 | 0.380 | 0.294 | 0.380 | 0.294 |
| NnmiR-SSR7 | 0.000 | 0.356 | 0.000 | 0.356 | 0.000 | 0.356 |
| NnmiR-SSR10 | 0.000 | 0.356 | 0.000 | 0.356 | 0.000 | 0.356 |
| NnmiR-SSR12 | 0.000 | 0.356 | 0.000 | 0.356 | 0.000 | 0.356 |
| NnmiR-SSR14 | 0.600 | 0.467 | 0.600 | 0.467 | 0.600 | 0.467 |
| NnmiR-SSR16 | 0.000 | 0.533 | 0.000 | 0.533 | 0.000 | 0.533 |
| NnmiR-SSR17 | 0.000 | 0.533 | 0.000 | 0.533 | 0.000 | 0.533 |
| Mean | 0.429 | 0.294 | 0.429 | 0.294 | 0.429 | 0.294 |

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

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. I., 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. Deng. 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. Michael, 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.

Töth, G., Z. Gaspári, 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.*