Development and characterization of genic-SSR markers from different Asia lotus (Nelumbo nucifera) types by RNA-seq


1State Key Laboratory of Hybrid Rice, College of Life Sciences, Wuhan University, Wuhan, Hubei, China
2College of Forestry and Life Sciences, Chongqing University of Arts and Sciences, Yongchuan, Chongqing, China
3Guangchang White Lotus Research Institute, Fuzhou, Jiangxi, China
4Lotus Center, Wuhan University, Wuhan, Hubei, China
5College of Pharmaceutical Sciences, Wuhan University, Hubei, China

*Theses authors contributed equally to this study.
Corresponding authors: Z.L. Hu
E-mail: huzhongli@whu.edu.cn

Received January 6, 2015
Accepted May 25, 2015
Published September 22, 2015
DOI http://dx.doi.org/10.4238/2015.September.22.11

ABSTRACT. Nelumbo nucifera is an important economic vegetable and traditional medicine, but available genetic resources remain limited. Next generation sequencing has proven to be a rapid and effective means of identifying genic simple sequence repeat (genic-SSR) markers. This study developed genic-SSRs for N. nucifera using Illumina sequencing technology to assess diversity across cultivated and wild lotus. A total of 105,834 uni-contigs were produced with an average read length of 722 bp. Exactly 11,178 genic-SSR loci were identified in 9523 uni-contigs. Di-nucleotide (64.5%) was the most abundant SSR, followed by tri-nucleotide (23%), tetra-nucleotide (8.9%), penta-nucleotide (2.5%), and hexa-nucleotide (1%) repeat types. The most common di- and
tri-nucleotide repeat motifs were AG/CT (51%) and AAG/CTT (8%), respectively. Based on these SSRs sequences, 6568 primer pairs were designed, of which 72 primers were randomly selected for synthesis and validation, and 38 in-silico polymorphic primers were obtained using in-house perl scripts. A total of 110 primers were screened in the lotus samples and the results showed that 101 primers yielded amplification products, of which 80 were polymorphs. The number of alleles ranged from 2 to 17 and the PIC (polymorphism information content) ranged from 0.19 to 0.87 with a mean value of 0.55. An Unweighted Pair Group Method with Arithmetic Mean (UPGMA) dendrogram based on Jaccard’s similarity coefficients showed that the correlation between geographical source and genotype was low. This study describes the distribution of genic-SSRs in the expressed portion of the lotus genome. These genic-SSRs have an important role to play in molecular mapping, diversity analysis, and marker-assisted selection strategies in *Nelumbo*.

Key words: Transcriptome; Genic-SSRs; *Nelumbo nucifera*; Genetic diversity