

Discovery of novel genic-SSR markers from transcriptome dataset of an important non-human primate, *Macaca fascicularis*

ABSTRACT

Macaca fascicularis, also known as the cynomolgus macaque, is an important non-human primate animal model used in biomedical research. It is an Old-World primate widely distributed in Southeast Asia and is one of the most abundant macaque species in Malaysia. However, the genetic structure of wild cynomolgus macaque populations in Malaysia has not been thoroughly elucidated. In this study, we developed genic-simple sequence repeat (genic-SSR) markers from an in-house transcriptome dataset generated from the Malaysian cynomolgus macaque via RNA sequencing, and applied these markers on 26 cynomolgus macaque individuals. A collection of 14,751 genic-SSRs were identified, where 13,709 were perfect SSRs. Dinucleotide repeats were the most common repeat motifs with a frequency of 65.05%, followed by trinucleotide repeats (20.55%). Subsequently, we designed 300 pairs of primers based on perfect di- and trinucleotide SSRs, in which 105 SSRs were associated with functional genes. A subset of 30 SSR markers were randomly selected and validated, yielding 19 polymorphic markers with an average polymorphism information content value of 0.431. The development of genic-SSR markers in this study is indeed timely to provide useful markers for functional and population genetic studies of the cynomolgus macaque and other related non-human primate species.

Keyword: *Macaca fascicularis*; Cynomolgus macaque; Genetic variation; Population genetics; Transcriptomics