Development of microsatellite markers in *Hagenia abyssinica* (Bruce) J.F. Gmel, an endangered tropical tree of eastern Africa, using next-generation sequencing

A.W. Gichira¹,², Z.C. Long¹, G.W. Hu¹, R.W. Gituru³, Q.F. Wang¹ and J.M. Chen¹

¹Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China
²University of Chinese Academy of Sciences, Beijing, China
³Botany Department, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Corresponding authors: Q.F. Wang / J.M. Chen
E-mail: qfwang@wbgcas.cn / jmchen@wbgcas.cn

Received November 27, 2015
Accepted January 18, 2016
Published June 21, 2016
DOI http://dx.doi.org/10.4238/gmr.15028168

**ABSTRACT.** *Hagenia abyssinica* (Bruce) J.F. Gmel is an endangered tree species endemic to the high mountains of tropical Africa. We used Illumina paired-end technology to sequence its nuclear genome, aiming at creating the first genomic data library and developing the first set of genomic microsatellites. Seventeen microsatellite markers were validated in 24 individuals. The average number of alleles per locus was 7.6, while the observed and expected heterozygosities ranged from 0.000 to 0.958 and from 0.354 to 0.883, respectively. These polymorphic markers will be used as tools for further molecular studies to facilitate formulation of appropriate conservation strategies for this species.

**Key words:** Afrotomante, *Hagenia abyssinica*; Microsatellites; Rosaceae; SSR; Next-generation sequencing
INTRODUCTION

*Hagenia abyssinica* (Bruce) J.F. Gmel (Rosaceae) is a dioecious tree species endemic to the isolated and scattered high mountains of tropical Africa at the ecological zone range of 2300-3400 m above sea level (Hedberg, 1970). Over-exploitation for its medicinal benefits, habitat destruction, and selective logging has led to drastic reduction in population sizes of this species and subsequent classification of this species as highly endangered in Ethiopia and tropical East Africa (Negash, 1995; Feyissa et al., 2007; Assefa et al., 2010; Seburanga et al., 2014).

Previously conducted studies on genetic diversity of *H. abyssinica* have used random-amplified polymorphic DNA (Kumilign, 2005), inter-simple sequence repeat markers (Feyissa et al., 2007), consensus chloroplast microsatellite primers (Ayele et al., 2009), and amplified fragment length polymorphism markers (Ayele et al., 2011). However, there is a need for more extensive surveys and studies using more informative markers, e.g., the co-dominant simple-sequence repeat (SSR) markers. SSR loci generate a large number of alleles per population and display a relatively higher genetic diversity (Powell et al., 1996). Here, we report the development of genomic microsatellite markers for use in outlining the genetic diversity and evolutionary significant units for conservation.

MATERIAL AND METHODS

Plant samples were collected from natural populations in eastern Africa. Species identification was done based on data collected from the National Museums of Kenya. Voucher specimens have been deposited at the Herbarium of Wuhan Botanical Garden, Chinese Academy of Sciences. Genomic DNA was extracted using the plant genomic DNA isolation kit (Tiangen, Beijing, China) following the manufacturer’s protocol. The quality and concentration of the DNA was checked using a Qubit DNA assay kit in Qubit 2.0 fluorometer (Life Technologies, San Diego, CA, USA). High-quality DNA samples were used for library construction, and sequenced using the Illumina HiSeq 2000 Platform at Novogene Bioinformatics Technology Co. Ltd. (Beijing, China). De novo assembly was conducted using the Velvet 2.0 (Zerbino and Birney, 2008) software. A total of 5 G clean DNA data were generated. Using the MIcroSAtellite identification tool (http://pgrc.ipk-gatersleben.de/misa/), 132,557 SSR motifs were identified and 38,819 primer pairs were designed using Primer3 (http://biotools.umassmed.edu/bioapps/primer3_www.cgi).

Firstly, we randomly selected 25 SSR primers and tested them for amplification and specificity, using genomic DNA extracted from four individuals obtained from Mt. Meru (03°13’S/036°46’E). Polymerase chain reactions (PCR) were carried out in a 25-µL reaction mixture composed of 2 µL (50 ng) genomic DNA, 2.5 µL 10X Taq buffer, 1 µL dNTPs (each 0.25 mM), 1 µL (0.25 µM) forward and reverse primer each, 0.2 U Taq polymerase (TaKaRa Bio, Dalian, China) and 17.3 µL ddH₂O. The PCR was performed in a T100™ Thermal Cycler (Bio-Rad, USA) as follows: an initial denaturation step of 4 min at 95°C was followed by 35 cycles of denaturation for 30 s at 95°C, annealing for 30 s at 50°-55°C, and extension of 30 s at 72°C. A final extension step was run for 10 min at 72°C. To quantify the PCR products, a 1.5% agarose gel was used. Secondly, we selected 17 primer pairs that produced clear unambiguous bands and labeled the forward sequence, with 6-FAM fluorescent dye at the 5’-end. These were then used for genotyping 24 samples obtained from natural populations of *H. abyssinica* in East Africa (Mt. Elgon, Mt. Kenya, Mt. Meru).
short-tandem repeat sequences were separated on an ABI 3730 XL automated sequencer (TsingKe Biotech, Beijing, China) and visualized using the GeneMarker software (Soft Genetics). The diversity analyses of the genetic markers, including the number of alleles ($N_A$), expected heterozygosity ($H_O$), and observed heterozygosity ($H_E$), were conducted using GenAlex 6.5 (Peakall and Smouse 2012).

RESULTS AND DISCUSSION

The $N_A$ varied from 4 to 13 (average of 7.6). The $H_O$ and $H_E$ values ranged from 0.000 to 0.958 and from 0.354 to 0.883, respectively (Table 1). These markers may be used to analyze the genetic distinctiveness and phylogenetic relationships among natural populations of *H. abyssinica* laying a foundation on which proper conservation measures will be conceived.

### Conflicts of interest

The authors declare no conflict of interest.

---

**Table 1.** Characterization of 17 microsatellite markers for *Hagenia abyssinica*.

<table>
<thead>
<tr>
<th>Name</th>
<th>Repeat motif</th>
<th>Primers (5’-3’)</th>
<th>Exp</th>
<th>Allele range (bp)</th>
<th>$T_m$ (°C)</th>
<th>$N_A$</th>
<th>$H_O$</th>
<th>$H_E$</th>
<th>GenBank accession No.</th>
</tr>
</thead>
</table>
| Hagssr1 | (TAT)$_6$    | F: AAATCTTATATCGGACGCA  
            R: AACCCTATGCTATGCTACAGA | 253 | 204 - 257 | 53 | 12 | 0.292 | 0.838 | KU128498 |
| Hagssr2 | (AG)$_9$     | F: TCCAAAAGCCAATTTGAA  
            R: AAAGCAGCAACCAATGACGA | 259 | 247 - 261 | 53 | 5 | 0.667 | 0.549 | KU128499 |
| Hagssr3 | (TA)$_6$     | F: AAGAGGTGAAAAATCGAAAGGA  
            R: CACCCGAGCTACTAATTTCT | 192 | 173 - 214 | 53 | 13 | 0.958 | 0.883 | KU128500 |
| Hagssr12|(AGAGAA)$_5$  | F: TTGGCTTTGTTGATATCTGCT  
                       R: ATCAATCGGTCATCTCACCC | 198 | 192 - 198 | 53 | 6 | 0.042 | 0.631 | KU128501 |
| Hagssr13|(AG)$_9$      | F: GAAATAGACGATGTCTGTTG  
                       R: CACTGACGTTGACTCTCTCTT | 185 | 167 - 186 | 53 | 5 | 0.958 | 0.685 | KU128502 |
| Hagssr15|(GA)$_7$      | R: TTATGGTGGGCTTTGGTTTC  
                       R: TGGTTTACTTTTGAGGACTGCT | 225 | 224 - 231 | 53 | 4 | 0.422 | 0.659 | KU128503 |
| Hagssr19|(GAA)$_5$     | F: GCCATTGAGGAGGCAAGCA  
                       R: ACCCTAGGCTATGGGACTGA | 221 | 207 - 255 | 54 | 13 | 0.292 | 0.857 | KU128504 |
| Hagssr20|(TTG)$_7$     | F: AAGAAAAGTTGTGTTGCTGCT  
                       R: TGTTGGTGTCTTGCTGCTCTT | 200 | 196 - 200 | 53 | 4 | 0.000 | 0.681 | KU128505 |
| Hagssr21| (CT)$_6$     | F: TCGGACAGCTAGACCTCACC  
                       R: TGGTGGTTGCTGGGATGAAA | 235 | 226 - 236 | 53 | 6 | 0.042 | 0.471 | KU128506 |
| Hagssr26|(AG)$_9$      | F: GACGAAGACAATCTGTGGCA  
                       R: TCTCATTCATTCCGTCGTCA | 223 | 217 - 223 | 53 | 4 | 0.000 | 0.354 | KU128507 |
| Hagssr27|(AG)$_6$      | F: AGTTTCACGCTTTGGAGGAA  
                       R: AACCTCCAACCCCTCCTAGA | 250 | 205 - 255 | 53 | 8 | 0.125 | 0.806 | KU128508 |
| Hagssr28|(TTC)$_6$     | F: TCGTCGATACGAGTTTGTGG  
                       R: AAGCTTGCTCTGTTTGAGG | 232 | 228 - 259 | 53 | 8 | 0.708 | 0.740 | KU128509 |
| Hagssr32|(TG)$_12$    | F: GGCTGAGCGCCCTGTGGATG  
                       R: TGGCAAGCATTAAATGACCACAG | 239 | 205 - 261 | 53 | 10 | 0.250 | 0.814 | KU128510 |
| Hagssr33|(CT)$_9$     | F: TTCAACTTGGTTGCTGTTC  
                       R: CACTGGAAAGAATCCCAAT | 188 | 181 - 188 | 54 | 4 | 0.083 | 0.480 | KU128511 |
| Hagssr35|(AAC)$_5$     | F: TCTAATTGGTCTACCCAGCA  
                       R: TCAATTTGCTGGTGTGGTGA | 279 | 274 - 285 | 53 | 5 | 0.417 | 0.711 | KU128512 |
| Hagssr38|(CT)$_13$    | F: AAGATTTTTTGCCGAGAGGCTG  
                       R: CACAGGAGAAAGGATGGG | 270 | 282 - 297 | 53 | 10 | 0.875 | 0.865 | KU128513 |
| Hagssr40|(AAG)$_5$     | F: GAATCTGACGCTTCAAGGACA  
                       R: CCCAGAAGAACCAACAAAA | 264 | 262 - 282 | 53 | 13 | 0.292 | 0.870 | KU128514 |

Exp = expected product range; $T_m$ = annealing temperature; $N_A$ = average number of alleles; $H_O$ = observed heterozygosity; $H_E$ = expected heterozygosity.
ACKNOWLEDGMENTS

We acknowledge the assistance offered by Edwin L. Kadiori and Josphat Mbogo in the field and John K. Muchuku and Elizabeth M. Kamande in the laboratory. Research supported by the Sino-Africa Joint Research Center (#Y323771W07, #SAJC201322).

REFERENCES


