MICROSATELLITE MARKERS FOR THE MOLECULAR CHARACTERISATION OF POTENTIALLY COMMERCIAL MANGO (Mangifera indica) PROGENIES

SHAHRIL AB RAZAK1*, MUHAMMAD AFIQ TAJOL ARIFFIN2, SHARIFAH MUNIRAH SYD MOHAMAD1, NOR HELWA EZZAH NOR AZMAN1, MOHD AZHAR HASSAN3 and JOHARI SARIP3

1Biotechnology and Nanotechnology Research Centre, MARDI Headquarters, 43400 Serdang, Selangor, Malaysia
2Horticulture Research Centre, MARDI Sintok, 06050, Bukit Kayu Hitam, Kedah, Malaysia
3Horticulture Research Centre, MARDI Headquarters, 43400 Serdang, Selangor, Malaysia
*E-mail: shahrilf@mardi.gov.my or shahrilfirdaus87@gmail.com

Accepted 2 September 2020, Published online 25 October 2020

ABSTRACT

Mango, or Mangifera indica as it is scientifically known, is a fruit crop that has significant importance to the global economy. This study characterised 10 commercially potential mango progenies using a total of 10 SSR markers. The analysis of the 10 SSR markers revealed the presence of 34 alleles. Each locus had 2–6 alleles equal to an average of 3.40 alleles per locus. The data also showed 0.115 (MiIIHR21) to 0.692 (MiIIHR28) polymorphic information content (PIC), equal to an average of 0.431. Each locus had heterozygosity and gene diversity values ranging from 0.044 (M1) to 0.957 (MiIIHR29) and from 0.122 (MiIIHR21) to 0.730 (MiIIHR29), respectively, averaging 0.480 and 0.476 each. The range of the pairwise genetic distance was 0.000 (ND014 ↔ ND019) to 0.611 (ND020 ↔ SB003). The dendrogram analysis showed that each mango progeny was successfully differentiated using the 10 SSR markers except for ND014 and ND019. The allele combination using these 10 SSR markers could be used to develop a DNA profile for fingerprinting purposes for each progeny except for ND014 and ND019, which can only be differentiated with the addition of an extra marker. The developed DNA fingerprinting profile could be used for cultivar identification and authentication. Additionally, breeders could also use the developed DNA profiles for plant variety protection (PVP) purposes.

Key words: DNA fingerprinting, genetic relationship Mangifera indica, SSR marker

INTRODUCTION

Anacardiaceae is a family of predominantly tropical species, with a few temperate region representatives. It is also the family to which mango (Mangifera indica L. 2n=40) belongs. There are 600 species in this family. These species are further classified into 70 genera. Of these genera, pistachio (Pistacia vera L.) and cashew (Anacardium occidentale L.) are some of the cultivated species that have emerged. There are 70 species in the genus Mangifera and a further two subgenera under it (Limus and Mangifera). These species are mostly restricted to tropical Asia (Kostermans & Bomgaard, 1993). According to Singh et al. (2016), mango has an estimated genome size of 439Mb. It is also a diploid (2n=40). In 2010, the world’s largest producer of mango was India (42.2%), which produced 16.38 million metric tons in that period. China came in second at 4.35 million metric tons (11.3%) followed by Thailand at 2.55 million metric tons (6.6%) (Mitra, 2014).

Mango is a perennial crop that is hard to characterise. Phenotypic-based cultivar identification is used to characterise mangoes. In this method, similar to other fruit tree species, the morphological traits of mango based on descriptors are identified from its cultivars (IPGRI, 2006). This method is, however, limited to the number of discriminating traits and environmental influences, making it difficult to precisely assess the varietal variability of the fruits, thus rendering the method inaccurate. Further adding to the confusion is the nature of the mango, which is a cross-pollinated crop, so different
sources of the mother tree could be used for seed and clonal propagation; therefore leading to intra- and inter-varietal variability. Nevertheless, a powerful tool that could be used for crop species in crop breeding research has recently emerged, known as molecular technology (Surapaneni et al., 2018).

Various molecular systems such as microsatellite markers (Viruel et al., 2005), isozymes (Degani et al., 1990), ISSRs (Pandit et al., 2007) and AFLPs (Kashkush et al., 2001) have been used to identify mango cultivars at the molecular level. Of these systems, microsatellites, also known as simple sequence repeats (SSRs) are codominant, allow reproducibility, and have a high level of polymorphism, making them the markers of choice for fingerprinting most plant species (Gupta & Varshney, 2000). Besides these advantages, SSRs perform particularly well in marker-assisted breeding and in the mapping of traits of interest of perennial crops. Therefore, this study aims to characterise and measure the genetic variability of ten potentially commercial mango progenies using microsatellite markers.

**MATERIALS AND METHODS**

**Genomic DNA isolation and plant material**

Ten potentially commercial mango progenies were selected for characterisation using microsatellite markers. The details of these progenies are summarised in Table 1. Five individuals of each progeny were collected before DNA extraction purposes. The Malaysian Agriculture Research and Development Institute (MARDI), Sintok, Malaysia (6° 28' 53"N, 100° 29' 00"E), containing the specific mango plots, supplied the progenies for this study. Young fresh leaves of each were collected and stored with silica gel in a zip-locked plastic bag for air-drying. A 96-well plate containing stainless steel beads (2.3 mm diameter) was prepared. Then, small fragments of the young leaves were punched onto the plate and were immediately frozen at –80°C for overnight. The frozen tissue was ground using Tissue Lyser (Qiagen, Germany). After the tissue was ground, an extraction buffer (2% CTAB, pH 8 of 100 mM Tris-HCl, 20 mM EDTA, 1.4 M NaCl, 0.05% β-mercaptoethanol) was immediately added. The protocol in Mace et al. (2003) was used to extract the total genomic DNA. Finally, 0.8% agarose gels were used to measure the DNA quality while Fluoraskan Ascent (Thermo Fisher Scientific, United States) was used to measure the DNA quantity.

**SSR genotyping**

The 10 SSR markers used in this study were taken from Ravishankar et al. (2011). These are listed in Table 2. The PCR was conducted as described by Schuelke (2000). First, a non-fluorescent labelled M13 sequence tail (TGTTAAACGCCTGCGGAT) and an M13 adapter labelled with a fluorescent dye (FAM, PET, NED or VIC) were used for ligating the primers (either forward or reverse). This was followed by preparing the PCR reaction mixture and making it up to a final volume of 10 µL, containing 1× buffer (Invitrogen, United States), 10 µM each of forward and reverse primer, 5 µM fluorescence-labeled M13 adaptor, 2 µM of each dNTP (Invitrogen, United States), 1U of Taq polymerase (Invitrogen, United States), and 0.1 µL of bovine serum albumin (BSA) as a PCR enhancer. A GeneAmp® PCR System 9700 (Applied Biosystems, United States) was used to perform amplification. In the amplification process, initial denaturation was set to 94°C for 2 min. Then, 35 cycles were run according to the following setting: i) 4°C for 30 s ii) 41–65°C for 45 s, and ii) 72°C for 45 s. The final extension was set to 72°C for 5 min. Upon the completion of amplification, the PCR product was multiplexed using up to four primers with a combination of four different fluorescent dyes. Then, Hi-Di formamide and GeneScan 500 LIZ, standard molecular weight ladders (Applied Biosystems, United States), were mixed with the products. Finally, an ABI 3130xl Genetic Analyzer (Applied Biosystems, United States) was used to resolve the products.

**Table 1.** Characterization of ten progenies used in this study

<table>
<thead>
<tr>
<th>Progeny ID</th>
<th>Mother Plant</th>
<th>Fruit Shape</th>
<th>Fruit Colour (ripe)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MK002</td>
<td>Mun Dian Kou</td>
<td>Oblong</td>
<td>Yellow</td>
</tr>
<tr>
<td>MM007</td>
<td>Masmuda</td>
<td>Elliptic</td>
<td>Yellow</td>
</tr>
<tr>
<td>ND001</td>
<td>Nam Dok Mai</td>
<td>Obovoid</td>
<td>Green</td>
</tr>
<tr>
<td>ND004</td>
<td>Nam Dok Mai</td>
<td>Oblong</td>
<td>Yellow</td>
</tr>
<tr>
<td>ND010</td>
<td>Nam Dok Mai</td>
<td>Elliptic</td>
<td>Yellow</td>
</tr>
<tr>
<td>ND014</td>
<td>Nam Dok Mai</td>
<td>Elliptic</td>
<td>Yellow</td>
</tr>
<tr>
<td>ND019</td>
<td>Nam Dok Mai</td>
<td>Elliptic</td>
<td>Yellow</td>
</tr>
<tr>
<td>ND020</td>
<td>Nam Dok Mai</td>
<td>Elliptic</td>
<td>Green</td>
</tr>
<tr>
<td>SB003</td>
<td>Strawberry</td>
<td>Obovoid</td>
<td>Red</td>
</tr>
<tr>
<td>SB019</td>
<td>Strawberry</td>
<td>Obovoid</td>
<td>Red</td>
</tr>
</tbody>
</table>
Statistical analysis

GeneMapper Version 5 (Thermo Fisher Scientific, United States) software was used to score the allele size. The methods described in Arif et al. (2010) were used as a basis for scoring and analysing the generated electropherograms. The number of alleles, the Major Allele Frequency, the Gene Diversity, the extent of heterozygosity, and the Polymorphism Information Content (PIC) of each microsatellite marker were calculated using PowerMarker software. The same program was also used to calculate the shared-allele based genetic-distance, as per Liu and Muse (2005). From the calculated genetic distance matrix of the shared microsatellite alleles, an Unweighted Pair Group Method with Arithmetic Mean (UPGMA) was constructed. Next, the UPGMA was used in the MEGA7 software (Kumar et al., 2016) to generate the dendrogram for this study.

RESULTS AND DISCUSSION

Characterisation of microsatellite marker

After analysing the 10 microsatellite markers, the author observed 34 alleles in each locus, equal to 3.4 alleles per locus on average. There were two (MiIIHR21) to six (MiIIHR29) alleles and 0.115 (MiIIHR21) to 0.692 (MiIIHR28) PIC values, equal to an average of 0.431. The informativeness of the molecular markers can be assessed from the PIC value. As per Hildebrand et al. (1994), only five microsatellite markers (25%) were considered to be highly informative (PIC value > 0.7). Also, the value of heterozygosity per locus ranged from 0.000 (MiIIHR10) to 0.8443 (MiIIHR29). The details of the characterisation are summarised in Table 2.

For the past 20 years or more, SSRs have remained the most trusted markers for genotyping plants, owing to their characteristics of informativeness, high polymorphic, multi-allele, co-dominant, experimentally reproducible, and transferability across species (Vieira et al., 2015). Moreover, microsatellites are greatly useful for diversity assessment measured in terms of genetic distance, cultivar identification through DNA fingerprinting, linkage and QTL analysis, and even in evolutionary studies (Kalia et al., 2011). MiIIHR29, which showed the presence of the highest number of alleles, is a simple dinucleotide perfect repeat motif microsatellite type with a high number of repeat motifs (n=10). Merrit et al. (2015) also found results consistent with this study, wherein the frequency of repeats were positively correlated with the number of alleles detected, suggesting that longer repetitive sequences tended to have higher chances and a greater extent of polymorphism irrespective of the mutation mechanism (Ellegren, 2004).

Molecular characterisation and genetic relationship among selected mango progenies

The UPGMA dendrogram (Figure 1), generated from shared allele matrix, indicate that there are two major clusters (Cluster A and Cluster B) for the studied progenies. There are two progenies, namely SB003 and SB009, in Cluster A. Both progenies are open-pollinated progenies from the same mother plant, namely the Strawberry cultivar. Meanwhile, the remaining eight progenies were clustered in Cluster B. Based on the results of the pairwise genetic distance (Table 3), the highest degree of variability between two progenies was observed for ND020 with SB003, as these had the highest genetic distance of 0.611. Meanwhile, the set of microsatellite markers used in this study was unable to differentiate the two progenies, ND014 and ND019, obtaining a value of 0.000 for the pairwise genetic distance between the two, which was also the lowest value. Both progenies were derived from the open-pollinated seed of the Nam Dok Mai variety. The limited number of microsatellite markers applied in this study resulted in its inability to differentiate the two mango progenies (ND014 and ND019). Further characterisation via the

<table>
<thead>
<tr>
<th>Marker</th>
<th>Major Allele Frequency</th>
<th>Allele No</th>
<th>Gene Diversity</th>
<th>Heterozygosity</th>
<th>PIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MiIIHR01</td>
<td>0.413</td>
<td>3.000</td>
<td>0.651</td>
<td>0.044</td>
<td>0.577</td>
</tr>
<tr>
<td>MiIIHR02</td>
<td>0.804</td>
<td>3.000</td>
<td>0.332</td>
<td>0.391</td>
<td>0.304</td>
</tr>
<tr>
<td>MiIIHR12</td>
<td>0.818</td>
<td>3.000</td>
<td>0.313</td>
<td>0.364</td>
<td>0.289</td>
</tr>
<tr>
<td>MiIIHR16</td>
<td>0.826</td>
<td>3.000</td>
<td>0.302</td>
<td>0.348</td>
<td>0.280</td>
</tr>
<tr>
<td>MiIIHR20</td>
<td>0.685</td>
<td>3.000</td>
<td>0.464</td>
<td>0.630</td>
<td>0.401</td>
</tr>
<tr>
<td>MiIIHR21</td>
<td>0.935</td>
<td>2.000</td>
<td>0.122</td>
<td>0.130</td>
<td>0.115</td>
</tr>
<tr>
<td>MiIIHR26</td>
<td>0.620</td>
<td>3.000</td>
<td>0.544</td>
<td>0.500</td>
<td>0.486</td>
</tr>
<tr>
<td>MiIIHR28</td>
<td>0.413</td>
<td>5.000</td>
<td>0.732</td>
<td>0.696</td>
<td>0.692</td>
</tr>
<tr>
<td>MiIIHR29</td>
<td>0.402</td>
<td>6.000</td>
<td>0.730</td>
<td>0.957</td>
<td>0.690</td>
</tr>
<tr>
<td>MiIIHR32</td>
<td>0.478</td>
<td>3.000</td>
<td>0.575</td>
<td>0.739</td>
<td>0.482</td>
</tr>
<tr>
<td>Mean</td>
<td>0.639</td>
<td>3.400</td>
<td>0.476</td>
<td>0.480</td>
<td>0.431</td>
</tr>
</tbody>
</table>

Table 2. Characterization of ten microsatellite marker used in this study
addition of additional microsatellite markers is needed to differentiate both progenies. Smýkal et al. (2008) demonstrated that if an adequate number of microsatellite markers were used, the method could prove to have powerful discriminatory power and polymorphism information content, and could even perfectly identify all varieties, more so than the majority of biochemical, morphological, or DNA markers out there.

**CONCLUSION**

There are a few factors that contribute to the high degree of divergence observed among mango hybrids. These include the high discriminatory power of the SSR markers, the different parental combinations of the mango hybrids, and the highly cross-pollinated nature of the mango crop. A cluster consists of the hybrids related to each other by descent. This study proved that SSR markers are not only useful for identifying the variety of mangoes, but also to estimate the heterosis for future mango breeding programmes.

**ACKNOWLEDGEMENT**

This study was funded by the Malaysian Government under Projek Pembangunan MARDI.
REFERENCES


