Microsatellite markers analysis for evaluation of genetic variation in mango genotypes

Kamlesh Kumar, Manish Srivastav, Sanjay Kumar Singh, Vinod and Chet Ram

Abstract

The study was under taken to evaluate the genetic variation in seven mango genotypes, namely, Amrapali, Dushehari, Jazarmand Pasand, Neelam, Pusa Arumma, Sensation and Tommy Atkins using SSR markers. At first, forty-two microsatellite markers were screened in mango genotypes for getting polymorphic SSR markers then polymorphic SSR markers were utilized for further analysis. Out of 42 primer pairs screened, 13 primer pairs namely LMMA2, LMMA3, LMMA4, LMMA7, LMMA8, LMMA9, LMMA11, LMMA12, LMMA16 and ESTD1, ESTD2, ESTD9 and ESTD10) were found polymorphic and rest of primers got monomorphic status. These 13 polymorphic markers were generated 43 alleles varied from 2 to 5 with an average of 3.30 alleles per locus among the mango genotype tested. The major allele frequency was varied from 0.36 to 0.86 for the locus ESTD9 and LMMA4, respectively. Mean major allele frequency was 0.60 for all the polymorphic loci. The observed heterozygosity (Ho) for individual loci ranged from 0.14 (ESTD2) to 0.86 (LMMA16), with an average of 0.44 over all loci. Whereas, gene diversity varied from 0.24 (LMMA4) to 0.74 (ESTD9), with an average of 0.53. The PIC value ranged from 0.21 (LMMA4) to 0.70 (ESTD9) with an average of 0.47. The microsatellite markers study helped in identification of genetic distance among mango genotypes. Dissimilarity matrix suggested a clear separation between Indigenous and exotic genotypes.

Keywords: Microsatellite markers, genetic diversity, mango, PCR, PIC

Introduction

Microsatellites are tandemly arranged repeats of mono-, di-, tri-, tetra-, and penta-nucleotides with different lengths of repeat motifs. The term microsatellite was coined by Litt and Lutty (1989) [9] which is also known as simple sequence repeats (SSRs). The SSRs markers are characterized by their abundance in the genome, highly polymorphic, reproducible and simple to use (Trojanowska and Bolibok, 2004) [18]. Mango (Mangifera indica L.) is the most popular fruit in the tropical and subtropical regions of the world. The mango belongs to the subgenus Mangifera. Owing to its religious and social importance, delicious taste, capitative flavour, attractive aroma and diverse end uses, it acknowledged as ‘King of fruits’ in India. The genus Mangifera contains about 70 species mostly restricted to tropical Asia and can be divided into two subgenera (Limus and Mangifera) [Bompard, 2009] [5]. Although some authors have considered India as the centre of origin due to the high degree of mango diversity observed in the country (Ravishankar et al., 2000) [14]. However, taxonomic and molecular level evidences supported that mango was evolved at north-western Myanmar, Bangladesh and North-Eastern India (Bompard, 2009 [5]; Mukherjee and Litz, 2009) [12]. Currently, mango is produced in more than 90 countries with the global production of 50.65 million metric tons. India’s share in the world’s mango production is around 56 per cent. India is the largest producer of mango in the world with an annual production of 19.68 MT from an area of 2.26 mha. Mango contributes 21.20 per cent of the total fruit production in the country with the productivity of 8.70 t/ha (NM database, 2016-17) [2]. In India, about 25 commercial cultivars are being grown at large areas; however, only few of them are preferred in the international market. Despite the recognized high quality of few well known mango cultivars, considerable cultivar improvement is needed in most regions of mango culture. The highly heterozygous and allopolyploid nature of the mango with complex genetic nature and lack of knowledge on inheritance pattern of several agronomic traits and pre-selection criteria makes the mango improvement work very cumbersome (Litz, 1997) [10]. Precise information on the genetic relationships within the mango germplasm is always
needed for carrying out efficient breeding programmes. In order to assess the genetic variability in mango, PCR based DNA markers are the most preferred genomic tools. A large number of DNA-based techniques were developed to identify genetic variability within species. The implication of DNA markers for mango improvement viz., estimation of genetic variation in existing populations, characterization of rootstocks, variety identification and validation of parentages has been reported (González et al., 2002) [7]. Among the molecular markers analyzed, SSR markers are the most potent for screening the genotypes for genetic variation because they are polymorphic, co-dominant, high reproducible, widely distributed throughout the genomes, amenable to automation, shows high levels of genetic variation in the number of tandemly repeating units at a locus and required low quantities of template DNA (Kumar et al., 2016) [8]. SSRs are actually considered the most efficient markers, but their use is still limited in mango due to long and laborious process to generate them (Trojanowska and Bolibok, 2004) [18]. Therefore, in the present study, a set of 42 SSR markers have been used for detecting polymorphism and evaluation of genetic variation among seven mango genotypes growing at ICAR-IARI, New Delhi.

Materials and methods

Plant material and DNA extraction

In the present investigation, seven mango genotypes were taken as experimental materials. Leaf samples (Young and tender) were collected from well-established 7 genotypes of mango which were maintained in the Division of Fruits and Horticultural Technology, ICAR-Indian Agricultural Research Institute (ICAR-IARI), New Delhi. The samples were immediately frozen in to liquid nitrogen (-196 °C) and stored till DNA isolation. DNA was isolated from frozen leaf samples using CTAB method (Murray and Thompson, 1980)[13] with some minor modifications. The composition of extraction buffer was 4% w/v CTAB, 1% PVP w/v, 20 mM EDTA, 100 mM Tris-HCl (pH-8), 1.4 M NaCl, and 0.2% β-mercaptoethanol. Leaf samples were ground to fine powder using pre-chilled pestle and mortar in the presence of liquid nitrogen. Powdered material was quickly transferred to centrifuge tubes containing CTAB extraction buffer (6 ml pre-heated) and vortexed it. The tubes were incubated at 65 °C for 1 h with intermittent shaking (after each 15 min.). After incubation, tubes were cooled to room temperature and 6 ml chloroform: isooamyl alcohol (24:1) was added. The contents were mixed by inversion for about 5-10 minutes. Then samples were centrifuged at 12,000 rpm for 10 min at 25 °C. The supernatant was transferred into new centrifuge tubes. Then DNA was precipitated by adding 1.4 M NaCl (0.5 vol.) and chilled isopropanol (1 vol.) and left it for over-night at 4 °C. Thereafter, the tubes were centrifuged at 10,000 rpm for 5 min. at 10 °C. Supernatant was discarded and the DNA pellet was washed twice with 70% ethanol. The DNA pellet was dried for complete removal of ethanol and was dissolved in 2ml TE buffer.

DNA purification and quantification

Two μl RNase A (10 mg/ml) was added to 200 μl DNA solution and incubated for 1 hour at 37 °C. The DNA was treated with an equal volume of Phenol: Chloroform: Isooamyl alcohol (25:24:1) and mixed the content gently by swirling for 5 minute. Then tubes were centrifuged at 10,000 rpm for 5 min. and supernatant was taken in a fresh tube. Followed by 2 extractions were given with Chloroform: Isooamyl alcohol (24:1). RNA free purified DNA was precipitated by adding Sodium acetate (0.1 vol.) and chilled Isopropanol (2.5 vol.) to the aqueous phase and then collected by spinning at 12,000 rpm for 10 minute. Then precipitate was washed twice with 70% ethanol. The pellet was air dried and dissolved in TE buffer. DNA was run in 0.8% agarose gel to check quality and quantity. The quantity of DNA was measured by comparing with the control λ phage DNA indicated height and weight of DNA. Each DNA sample was diluted with sterilized double distilled water to yield a working concentration of 25 ng/μl.

PCR amplification

PCR amplification was carried out in 20 μl containing 50 ng of genomic DNA, 2·0 μl MgCl₂ (25 mM), 1U Taq DNA polymerase (Fermetas, USA), 2·0 μl of 10X PCR buffer (without MgCl₂), 0.5 μM of each of primers and 200 μM of dNTPs. The final volume was made up with sterile double-distilled water. The PCR was performed in Perkin Elmer 9,600 thermal cycler using the following thermocycling conditions. PCR products were analysed electrophoretically on 3.5% high resolution agarose (Metaphor) gels containing ethidium bromide (10mg/ml). Three μl of 100 bp standard DNA ladder (Fermentas, USA) was loaded in the first well of each gel to determine sizes of identified bands. Electrophoresis was carried out at 5 V/cm for 3 hour. The agarose gel was documented using gel documentation system.

Band scoring and data analysis

The clear and scorable loci of SSR were considered for scoring as 0 for the absence and 1 for the presence of alleles. Based on 0 and 1 allele scoring, binary matrix was prepared and binary data was analysed using NTSYS-PC program (Numerical Taxonomy and multivariate analysis SYSTAT program) and Power Marker version 3.23 software (Liu and Muse, 2004) [11].

The specific bands useful for identifying species and cultivars were named with a primer number followed by the approximate size of the amplified fragment in base pairs. Polymorphism was calculated based on the presence or absence of bands. The 0 or 1 data matrix was created and used to calculate the genetic distance and similarity using “Simqual” a subprogram of the NTSYS-PC program (Numerical Taxonomy and multivariate analysis SYSTAT program). The genetic associations between accessions were evaluated by calculating the Jaccard’s similarity coefficient for pair wise comparisons based on the proportions of shared bands produced by the primers. The dendrogram was constructed by using a distance matrix using the un-weighted pair group method with arithmetic average.

Result and Discussion

Identification of polymorphic SSR markers

In order to get polymorphic SSR markers, forty-two SSR primers were tested on seven mango genotypes representing exotic as well as indigenous cultivars (Fig. 1). Out of 42 primer pairs screened, thirteen primer pairs viz.LMMA2, LMMA3, LMMA4, LMMA7, LMMA8, LMMA9, LMMA11, LMMA12, LMMA16 and ESTD1, ESTD2, ESTD9 and ESTD10) were polymorphic. Rest of the primers found to be monomorphic in nature, thus could not included in the further study.
Forty-two SSR markers were screened for polymorphism among seven mango genotypes (Table 1 and 2). Out of 42 primer pairs screened, thirteen primer pairs (LMMA2, 3, 4, 7, 8, 9, 11, 12, 16 and ESTD1, 2, 9 and 10) were polymorphic and 19 primers (LMMA1, 4, 5, 6, 10, 13, 14, 15 and ESTD 3, 4, 5, 6, 7, 8 and 11, MiSHRS23, 26, 29, 30 and 32) were monomorphic. Ten primer pairs (MiSHRS1, 4, 18, 33, 34, 36, 37, 39, 44 and 48) did not amplify the DNA of any genotype. Analysis of SSR data obtained for polymorphism studies among seven mango genotypes was performed by Power Marker software version 3.23 software to generate genetic similarity matrix based on possible pair-wise combinations of accessions using the “proportion of shared alleles”. UPGMA cluster analysis was used to construct a dendrogram from the genetic similarity matrix using Power Marker version 3.23 software (Liu and Muse, 2004). Observed heterozygosity (H0) is the proportion of individuals in a population that is heterozygous at a given number of loci. On the other hand, the expected heterozygosity (He) is defined as the proportion of individuals which are prospective heterozygotes based on the allele frequencies and assuming Hardy-Weinberg equilibrium.

Table 1: A list of SSR primers (forward and reverse) and their base sequence

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Primer ID</th>
<th>Sequence (5′-3′)</th>
<th>Tm (°C)</th>
<th>Ta (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LMMA1</td>
<td>F: ATGGAGACTGAATGTACAGAG R: ATTAAATCTCGTCCACAAGT</td>
<td>56.5</td>
<td>51.2</td>
</tr>
<tr>
<td>2</td>
<td>LMMA2</td>
<td>F: AAATAGATGAAGCACTAAAG R: TTAGTGATTTTTGATGTTGTC</td>
<td>50.9</td>
<td>50.9</td>
</tr>
<tr>
<td>3</td>
<td>LMMA3</td>
<td>F: AAAAAACCTTCATAAAGTGAATC R: CAGTAAAACCTGGATACCTT</td>
<td>50.9</td>
<td>52.0</td>
</tr>
<tr>
<td>4</td>
<td>LMMA4</td>
<td>F: AGATTTAAAAGCTCAAGAAAAA R: AAAGCTTAAATGTTGTTTCTC</td>
<td>48.1</td>
<td>52.0</td>
</tr>
<tr>
<td>5</td>
<td>LMMA5</td>
<td>F: AGAATAAGCTGATACCTACAC R: TAACAAATATCTAATTGACAGG</td>
<td>54.0</td>
<td>50.9</td>
</tr>
<tr>
<td>6</td>
<td>LMMA6</td>
<td>F: ATATCTCAGGGTCTGAAATGA R: TTATTAATTTTCACAGACTATGTTCA</td>
<td>53.2</td>
<td>53.1</td>
</tr>
<tr>
<td>7</td>
<td>LMMA7</td>
<td>F: ATTTAATCTTCATACTTCAAC R: AGATTTAGTTTTGATTGAG</td>
<td>50.9</td>
<td>50.9</td>
</tr>
<tr>
<td>8</td>
<td>LMMA8</td>
<td>F: CATGGAGTTTGTGATACCTAC R: CAGATGTTAAGCCATATAGGTG</td>
<td>53.3</td>
<td>55.9</td>
</tr>
<tr>
<td>9</td>
<td>LMMA9</td>
<td>F: TTGCAACTGATAAACAAATATAG R: TTCACATGACAGATATACCTT</td>
<td>50.9</td>
<td>52.8</td>
</tr>
<tr>
<td>10</td>
<td>LMMA10</td>
<td>F: TTCTTATTAGCTAAAGAGACACATT R: AGTTACAGATCTTCTCACAAT</td>
<td>52.8</td>
<td>52.0</td>
</tr>
<tr>
<td>11</td>
<td>LMMA11</td>
<td>F: ATATTTAGACCTACAGAGTTAC R: GTATTATCGGAATGTGCTTTC</td>
<td>54.0</td>
<td>52.8</td>
</tr>
<tr>
<td>12</td>
<td>LMMA12</td>
<td>F: AAGAGATAGCATTATTAATATGAAGA R: GTAAAGATCGCTGGTGTGTTATT</td>
<td>49.1</td>
<td>52.8</td>
</tr>
<tr>
<td>13</td>
<td>LMMA13</td>
<td>F: CACAGCTAATAAACAATCTATG R: CATTATATCCCTAATCTACATC</td>
<td>54.0</td>
<td>52.8</td>
</tr>
<tr>
<td>14</td>
<td>LMMA14</td>
<td>F: ATTACTTCTATAGTCCCTATAT R: CTCGGTAAAACTTTTGGACTAT</td>
<td>52.0</td>
<td>53.3</td>
</tr>
<tr>
<td>15</td>
<td>LMMA15</td>
<td>F: AACACTGTGGCTGACATAT R: CTGGTTAACATATGACCATCTT</td>
<td>53.2</td>
<td>52.8</td>
</tr>
<tr>
<td>16</td>
<td>LMMA16</td>
<td>F: ATAGATTTACATATCTCTTCTG CAT R: TATAATTATATCGTCTCTC</td>
<td>50.9</td>
<td>50.9</td>
</tr>
<tr>
<td>17</td>
<td>MiSHRS1</td>
<td>F: TAACAGCTTTTGCTGCTGCTC R: TCCGCCGATAAACATCACAC</td>
<td>57.3</td>
<td>57.3</td>
</tr>
<tr>
<td>18</td>
<td>MiSHRS4</td>
<td>F: CCACGAATATCACACTGCTGC R: TCTGACACTGCTCTTCCACC</td>
<td>59.8</td>
<td>59.4</td>
</tr>
<tr>
<td>19</td>
<td>MiSHRS18</td>
<td>F: AAACGGAGAAACAGAGAC R: CAAGTACCTGCTGCAACTAG</td>
<td>56.0</td>
<td>57.3</td>
</tr>
</tbody>
</table>

Fig 1: Banding pattern of LMMA16 (a) and LMMA2 (b) primers on mango parental genotypes. M indicates 100 bp DNA ladder; A=Amrapali; D= Dushehari; JP= Janardan Pasand; N=Neelum; PA= Pusa Arunima; S= Sensation; TA= Tommy Atkins.
Genetic diversity in mango cultivars based on SSR markers

In the present study, a total of 43 alleles were amplified using 13 polymorphic SSR markers. Number of allele varied from 2 to 5 with an average of 3.30. The major allele frequency varied from 0.36 to 0.86 for the locus ESTD9 and LMMA4, respectively. Mean major allele frequency was 0.60 for all the polymorphic loci. The observed heterozygosity (Ho) for individual loci ranged from 0.14 (ESTD2) to 0.86 (LMM16), with an average of 0.44 over all loci. Whereas, gene diversity varied from 0.24 (LMM4) to 0.74 (ESTD9), with an average of 0.53. The PIC values ranged from 0.21 to 0.70 with an average of 0.47. The most polymorphic locus was ESTD9 and least polymorphic locus was LMMA4 (Table 3). Nine out of 13 SSR loci showed a positive estimated frequency of null alleles (r). Locus ESTD2 had the highest positive r value of 0.214, which indicates a likely occurrence of null alleles (Table 3). Begum et al. (2013) [4] also screened 109 SSRs on Panchadarakalasa mangoes for polymorphism and variability analysis. Out of them, they reported only 4 SSRs as polymorphic which produced 11 bands. They concluded that polymorphism level of the employed SSRs was variable ranging from 2-4 alleles per SSR and average was 2.75 per SSR. The PIC value ranged from 0.25-0.56 with an average of 0.42.
Dissimilarity matrix was used to construct the dendrogram which revealed the clones into two major clusters. Cluster I comprised only 2 genotypes namely Sensation and Tommy Atkins. Both the genotypes are exotic. Whereas, cluster II, grouped the remaining 5 genotypes. Dushehari and Janardan Pasand grouped together and Neelum, Amrapali and Pusa Arunima were separated in one group. Maximum genetic distance was found between Tommy Atkins and Dushehari (0.67), whereas, minimum genetic distance was noted between Janardan Pasand and Dushehari (0.25). Average genetic distance was 0.48 among the seven genotypes studied (Fig. 2). Similar study was done by Surapaneni et al. (2013)[17] using microsatellite markers in mango. They characterized 90 genotypes and detected 301 alleles from 106 polymorphic loci with an average of 2.87 alleles per locus and polymorphism information content was 0.67. They did UPGMA cluster analysis and categorized all the genotypes into two major groups with a genetic similarity range of 47-88 percent.

Fig 2: Dendrogram based on dissimilarity matrix of mango genotypes

Genetic and genomic complexity of the cultivars has hindered identification of markers in mango. However, remarkable progress has been witnessed in molecular marker technology in mango. Different molecular markers types as isozymes (Aron et al., 1997) [3], RAPD (Ravishankar et al., 2000) [14], minisatellite and microsatellites (Schnell et al., 2005) [16] and Amplified Fragment Length Polymorphism (Eiadthong et al., 2000) [6] were used to study polymorphism in mango. Recently, SSR primers have been developed and reported by many laboratories across the globe, these being choicest markers for their co-dominant nature. Schnell et al. (2005) [16] reported on the development and characterization of 15 microsatellite loci isolated from Mangifera indica L. Out of which, nine loci exhibited significant linkage disequilibrium. The genotypes used in the present study differed significantly at SSR loci, which is evident from estimated parameters such as the number of detected alleles and per cent polymorphism. Wahdan et al. (2011) [19] screened 42 primers in two Egyptian mango genotypes ‘Aml’ and ‘Hania’ using SSR markers, in which, 36 were polymorphic. By obtaining banding patterns from 36 primers, each genotype could be distinguished from the other, indicating that SSR primers were effective for identification of genotypes. Ravishankar et al. (2011) [15] developed new microsatellite markers to characterize and assess the genetic diversity among mango cultivars. They isolated 36 SSRs loci by a microsatellite-enriched partial genomic library method and designed 86 primers pairs for putative microsatellite loci and characterized 30 mango cultivars. They found expected heterozygosity (He) values ranged from 0.195 to 0.941 with average of 0.728 and PIC values varied from 0.185 to 0.920 with average of 0.687. The majority of them represented high PIC values (>0.700). For all the 36 SSRs loci the total value of PI was 2.42x10-31. Grouping of seven mango genotypes resulted in two major clusters. From the perusal of the geographical locality of each genotypes and their clustering pattern it could be inferred that the grouping of the genotypes is associated with their geographical location as well as parent used in the hybridization programme. Similar results were also revealed by Abirami et al. (2008) [1]. In this study, the markers ESTD9 with high PIC value provide an opportunity for direct comparison and identification of different genotypes independent of any influences. Specific bands/loci observed in the Sensation and Tommy Atkins genotypes studied may be useful for cultivar identification. Further, the presence of specific loci indicates the genetic distinctness of the genotype under this study.

Acknowledgements

Authors are highly thankful to Dr. A K Singh and Dr. K Usha, Heads, Division of Fruits and Horticultural Technology, ICAR-IARI, Pusa, New Delhi, India for financial support and their valuable suggestions for conducting the experiment.

References