Journal of Genetics, Vol. 97, Suppl 1, December 2018, Online Resources, pp. e39–e47
https://doi.org/10.1007/s12041-018-0893-0

ONLINE RESOURCES

Identification of novel microsatellite markers in okra (*Abelmoschus esculentus* (L.) Moench) through next-generation sequencing and their utilization in analysis of genetic relatedness studies and cross-species transferability

KUNDAPURA V. RAVISHANKAR*, GAYATHRI MUTHAIAH, PITCHAIMUTHU MOTTAIYAN and SANTHOSH-KUMAR GUNDALE

Division of Biotechnology and Division of Vegetable Crops, ICAR-Indian Institute of Horticultural Research, Bengaluru 560 089, India
*For correspondence. E-mail: kv_ravishankar@yahoo.co.in, kvravi@iihr.res.in.

Received 10 January 2017; revised 13 June 2017; accepted 16 June 2017; published online 11 April 2018

Keywords. microsatellite markers; transferability; polymorphic information content; *Abelmaschus esculentus*.

Introduction

Abelmoschus esculentus L. (Malvaceae), populary known as okra, is one of the important pod vegetable crops in tropics and subtropics. To develop the microsatellite markers, we have partially sequenced genomic DNA using Roche 454 Titanium pyrosequencing technology. A total of 979,806 bp data from 61,722 reads were obtained. These reads were assembled into 3735 contigs of which 2708 had microsatellites. Primers were designed for 402 microsatellites, from which 50 randomly selected SSR primers were standardized for amplification of okra DNA. Genetic analysis was done by employing 29 genotypes. Polymorphic information content (PIC) across all 50 loci ranged from 0.000 to 0.865 with a mean value of 0.519. The observed and expected heterozygosity ranged from 0.000 to 0.750 and 0.000 to 0.972, respectively. Alleles per locus ranged from 1 to 27. This is the first report on the development of genomic SSR markers in okra using next-generation sequencing technology. We have also assessed cross species transferability of microsatellite markers with related wild species *A. tetraphyllus* (84%), *A. ficulneus* (78%), *A. moschatus* (92%), *A. angulosus grandiflorus* (90%) and *A. caillei* (86%), and demonstrated their use in genetic relatedness study and molecular characterization.

Okra (*Abelmoschus esculentus* L.) belongs to the family Malvaceae and is considered as one of the important pod vegetable crops in India, Southeast Asia, West Africa, USA, Brazil and Turkey. Okra is rich in calcium, mineral salt, and vitamins (Aladele *et al.* 2008). The total production of okra in the world is around 4.8 million ton pods with highly mucilage, where India contributes 70% Nigeria 15%, Pakistan 2%, Ghana 2%, Egypt 1.7%, and Iraq 1.7% (Gulsen *et al.* 2007).The genus *Abelmoschus* includes eight species (Ramya and Bhat 2012). The genome size of okra is 1.6 Gb, and the chromosome number is generally found to be $2n = 130$ (Joshi and Hardas 1956). There are two kinds of *A. esculentus* L. genotypes $2n = 60-70$ as diploid and $2n = 120-130$ as tetraploids, this variation in chromosome number could be due to irregularities in chromosome movement during the mitotic phase of cell division (Nwangburuka *et al.* 2011). The polyploidy complex was further assessed through chloroplast DNA (cpDNA) intronic spacer, which revealed that *A. ficulneus* and *A. moschatus* are the closest wild relatives of *A. esculentus* (Ramya and Bhat 2012).

Molecular markers have proved as a powerful tool for the assessment of genetic variation and elucidation of genetic relationships within and among the species (Chakravarthi and Naravaneni 2006; Yuan *et al.* 2014, 2015). Molecular marker techniques like RAPD, RFLP, simple sequence repeat (SSR) and AFLP are the most widely used for the genetic characterization of species and also for the crop improvement (Sawadogo *et al.* 2009).

However, okra lacks information on molecular markers (Sunday *et al.* 2008). There are only two studies where markers were used to assess the genetic diversity in okra using general DNA markers (Gulsen *et al.* 2007; Prakash *et al.* 2011). SSR markers are an important marker tool in the application of plant genetics and breeding because of their high reproducibility, multiallelic nature, codominant inheritance and a good genome coverage (Bertini *et al.* 2006). The next-generation sequencing (NGS) technology is revolutionizing the field of molecular biology, especially in the development of microsatellite markers. The advantage of NGS is its cost effectiveness, time saving and large number of SSR can be identified (Ravishankar *et al.* 2015). Recently Roche/454 Solexa, Illumina and SOLID are the technologies that are predominantly used in sequencing genetics and breeding applications. The objective of this study was to isolate and characterize the microsatellite markers for okra through NGS of okra genome, and to examine their transferability to the related species.

Materials and methods

Total genomic DNA isolation

Twenty-eight okra cultivars and five related wild species of *A. manihot* var. *tetraphyllus, A. ficulneus, A. moschatus, A. angulosus* var. *grandiflorus* and *A. caillei* were used in this study (table 1). Young leaves were collected from plants that were maintained at the ICAR-Indian Institute of Horticultural Research, Bengaluru. Total genomic DNA was isolated using modified CTAB method (Ravishankar *et al.* 2000) with the repetition of washing step using chloroform: isoamyl alcohol (24 : 1) for five times until its mucilage was removed. The DNA concentration was determined using a spectrophotometer (Gene Quanta, Amersham Biosciences, Holliston, USA) by taking absorption at 260 and 280 nm.

Sequencing and assembly

Total genomic DNA of cultivar Arka Anamika was used for shotgun sequencing using Roche 454 GS FLX Titanium sequencing technology (at M/S Eurofins, Bengaluru facility) following manufacture's instructions. Raw reads were used for the assembly of contigs using CAP3 Software (Huang and Anup 1999).

Mining SSRs and design of primers

The assembled contigs were further analysed for the presence of microsatellites using MISA software (Suping *et al.* 2013). For the identified microsatellites, primers were designed by using Primer3 software (Untergasser *et al.* 2012). A total of 402 SSR primers were designed for the SSRs (see electronic supplementary material at http://www.ias.ac.in/jgenet/). From these, a set of randomly selected 50 primers were used for amplification of DNA from 28 okra genotypes and five related species.

Table 1. Genotypes used in the study.

	Genotypes/accession number
1	Kashi Vibhuti
2	Komal 2486
3	Kashi Lalima
4	Kashi Pragathi
5	Acc. 315
6	Parbhani Kranti
7	SB 2
8	SB 8
9	Arka Anamika
10	JOKH 7315
11	Hisar Unnat
12	Varsha Upahar
13	IIHR-G10
14	IIHR-1685
15	IIHR 31
16	Acc. 444
17	Acc. 344
18	Acc. 340
19	Salkirthi
20	Acc. 170
21	Acc. VRO 6
22	Pusa A4
23	CO 1
24	Acc. 294
25	Acc. 281
26	Acc. M64
27	Acc. F1
28	IIHR 1
29	Arka Abay
30	*A. tetraphyllus*
31	*A. ficulneus*
32	*A. moschatus*
33	*A. angulosus grandifloras*
34	*A. caillei*

PCR conditions and genotyping

We used a fluorescence-based M13 tailed PCR assay to amplify the microsatellites (Oetting *et al.* 1995). The primers had the addition of a standard M13 tail, at the 5′ end of the primer sequence (Schuelke 2000). A total of 100 SSR markers were initially screened with pooled DNA, and later 50 SSRs were employed for 28 okra genotypes and five related species (table 2). The PCR amplification conditions were as follows: initial denaturation at 94°C for 2 min, followed by 35 cycles of denaturation, annealing and polymerization steps (94°C for 30 s, 50–60°C and 72°C for 1 min respectively. A final extension of polymerization was done at 72°C for 5 min. PCR amplification

Table 2. Genetic analysis of microsatellite loci using 29 okra genotypes.

	Primer name	Primers	T_m	No. of allele	H_o	H_e	PIC value	PI
1	AeKVR 102	F:TTCCGATGAGGAACTAGATGGAT R:CCCAGTTGGACCTTGAGAGA	55	18	0.148	0.886	0.861	0.024
2	AeKVR-103	F:TACGACACGTATATGCCTACTTGAG R:ACGTCCTAAACGAACGTCCTTA	55	4	0.160	0.490	0.450	0.227
3	AeKVR-110	F:TCTCGAAAACACTTGAGAGCAAA R:TCGTGTGCCCTTTGAGAGA	57	1	0.000	0.000	0.000	0.609
4	AeKVR-111	F:TGCTATTATCCCTGCTCCTTCTG R:GTGATGGTTTCCTTGGAGTCA	50	5	0.214	0.449	0.417	0.537
5	AeKVR-132	F:TAACTTTGCACGAAACTGTGAAC R:TAGGTTGTCGATTTGGATGTTG	57	5	0.214	0.471	0.434	0.314
6	AeKVR-125	F:TCGTATCGTTGAAGAAGGAAACA R:TGAATCGTCCGTGATATAAACC	57	3	0.038	0.298	0.259	0.333
7	AeKVR-116	F:TAGCAAAAGCGATGATTGTCTG R:CCCCTAAACCCTAATCCTGACT	55	4	0.115	0.310	0.287	0.308
8	AeKVR-137	F:TAAAGGGTATAATGCAGCCATGT R:TCGTGGTTAGTTTCGTTTTCCT	60	18	0.148	0.886	0.861	0.681
9	AeKVR-134	F:TCTTATCGACTGATGTTGAGGCA R:TGTGGGAAACTATAACCGGAAC	60	6	0.536	0.734	0.675	0.113
10	AeKVR-156	F:TTAGAGGTGCTCATCCTTCCCTA R:TGGTTTCATGCTCCCACTTAAT	60	5	0.214	0.449	0.417	0.298
11	AeKVR-131	F:TATTCATAAACAGGGATGAGCAG R:TCTTCAGTCCGATACAGCACTT	50	3	0.000	0.222	0.205	0.415
12	AeKVR-173	F:TCCGTGACTATCTACTACGTTCGTG R:GTTCGACACGCTAAACGCTAA	50	3	0.160	0.344	0.307	0.334
13	AeKVR-153	F:TGGAGTGTGCTGCCAAGTTTAT R:TTGTTTATGCTGTGATGCTGAC	55	18	0.148	0.886	0.861	0.024
14	AeKVR-114	F:TGAGAAGCGATGTTCTAGCGATT R:ATGAGGAAATAACTGACCCAGC	57	11	0.000	0.875	0.843	0.269
15	AeKVR-192	F:TCGTGACCGTGGACTCGTAGGTA R:ACGACCGACCGACCGAAC	57	9	0.308	0.744	0.695	0.743
16	AeKVR-193	F:TAACGCAAACTGAACTCTCGTTA R:ACCTAACCCTAACCCTAAACCG	57	8	0.750	0.617	0.544	0.175
17	AeKVR-113	F:TAACTGAGCTATCCCGACCATTA R:CTCGTTCATCCTATCTTTTGCC	55	9	0.036	0.710	0.677	0.095
18	AeKVR-118	F:TAGCAAAAGCGATGATTGTCTG R:CCCCTTAAAACCCTAATCCTGA	55	5	0.000	0.633	0.555	0.145
19	AeKVR-182	F:TAAAGCGAGGTGGTCTACATGAC R:ATTGGGTGAAAGACAGAAAGGA	55	27	0.407	0.972	0.952	0.003
20	AeKVR-150	F:TATCTCCCCATCGACTAACAAGA R:AACGAGCGGACACCTACG	55	2	0.000	0.509	0.375	0.290
21	AeKVR-176	F:TCCGTTTTATTTCGACCGTTACC R:TAACCGAACCCGAACCGTA	55	27	0.407	0.972	0.952	0.003
22	AeKVR-195	F:TCGTAACCCGTATAATGCAACAG R:AACGTAACCTAACCCTAAACCCT	55	12	0.148	0.720	0.678	0.091
23	AeKVR-117	F:TACGTTCCGTACCTTACTTCGG R:GTTACGACGAGGTTTACCAAGG	57	16	0.593	0.890	0.865	0.012
24	AeKVR-119	F:TAACTGAGCTATCCCGACCATTA R:CTCGTTCATCCTATCTTTTGCC	57	9	0.037	0.764	0.727	0.068
25	AeKVR-161	F:TAGAGGGATAAATTGGGAGATGC R:CGTTGGGACTACTGTGACTCTG	57	7	0.037	0.535	0.497	0.178
26	AeKVR-141	F:TATGTCAAGGATGAATGCTCTCT R:ACGTACAACGAACGACGAC	57	6	0.083	0.306	0.291	0.324
27	AeKVR-177	F:TCCGTAAACCGTAAAACCGTA R:AAGGTAAGGTAGGTAGGTAAGGTAAGT	50	4	0.160	0.490	0.450	0.227
28	AeKVR-199	F:TCGTAAACCGTAAACCGTAAACC R:CGTCGTCGTAGTAGTGTGTGTAGTAG	50	3	0.708	0.519	0.415	0.217
29	AeKVR-172	F:TCGTCCTATTAAGCGTATGGTCTC R:GTAACGAACGAACGAACGAAC	55	18	0.148	0.886	0.861	0.024
30	AeKVR-160	F:TACGACTTTCTTCTCCCTTTTCC R:ATCAACCCGGACATCAATAAAC	57	4	0.160	0.490	0.450	0.227

Table 2 (*contd*)

Primer name	Primers	T_m	No. of allele	H_o	H_e	PIC value	PI
31AeKVR-149	F:TCACCAGGCTCGACCACTC R:GTACGTCGGGTACGACCG	57	4	1.000	0.581	0.478	0.218
32AeKVR-162	F:TTGGCATTAGCGTGTGTATGAAT R:ACGACTTTAGCCGACAACTCTG	57	3	0.036	0.227	0.205	0.537
33AeKVR-147	F:TGTACGTTACCGACGTTCGAC R:GTTAGACCGGAGTACCGAGC	57	5	0.214	0.471	0.434	0.314
34AeKVR-183	F:TGGTTTAGGGTTTACCGACTACG R:TAAGTTCGGGTTTAGGGTACGA	55	3	0.038	0.298	0.259	0.333
35AeKVR-101	F:TAAAACAACCATAAGCAGCACCT R:TGTTCTAATTCTTGCCTGACGA	55	4	0.115	0.310	0.287	0.308
36AeKVR-104	F:TAGCGAAGAAATCACAGTTCACA R:TAGGACCCAACGGAAGATAGAG	55	4	1.000	0.581	0.478	0.218
37AeKVR-108	F:TAGCGAAGAAATCACAGTTCACA R:CGGGGAAATAAAGTAGAAAGGC	55	6	0.536	0.734	0.675	0.113
38AeKVR-145	F:TGGGCTGGCACTGGAGGTC R:GACGACGACGACCGACGA	55	5	0.214	0.449	0.417	0.298
39AeKVR-187	F:TCCGAGATTCAAGCGGATTATAG R:ACGACCACGCAACCGTAT	55	3	0.000	0.222	0.205	0.415
40AeKVR-107	F:TCTGGGTGTCTTCTGCTTGTATCT R:CAGATTTTGCTAGGGGAGAATG	55	3	0.160	0.344	0.307	0.334
41AeKVR-200	F:TCGTTTACCGTTTACCGTTTACCT R:AACCGAACCGAACCTAACG	55	5	0.214	0.449	0.417	0.298
42AeKVR-106	F:TAGCGAAGAAATCACAGTTCACA R:CCCAAAACGGGTAACGTAAAG	60	3	0.000	0.222	0.205	0.415
43AeKVR-143	F:TCAATCACTCCCAATCACCTTTA R:TGGAGAAGTAAAACCCAACACA	50	3	0.160	0.344	0.307	0.334
44AeKVR-196	F:TGTTTTACCCCTAACCGTCCC R:TACCCGTACCCGTCCGTA	60	4	0.080	0.345	0.311	0.367
45AeKVR-126	F:GTACGGATACTCAAACGAAGGC R:GTAGGTGCAGTGTTGTTACCGA	57	11	0.000	0.875	0.843	0.026
46AeKVR-194	F:TCGAACCCCTGAACTTGGTATT R:CACCACCGTAATAACCTAACCC	55	9	0.308	0.744	0.695	0.074
47AeKVR-165	F:TAGCAAAAGCGATGATTGTCTG R:CCCCTAAACCCTAATCCTGACT	55	8	0.750	0.617	0.544	0.175
48AeKVR-163	F:TCAAGACAAGGATAAAGTCGCAC R:TAATGCAGGGTGATTTAGATGG	55	9	0.036	0.710	0.677	0.095
49AeKVR-167	F:TGGATTTAGGCTTTTAGGGATGG R:AACACACACACACACACACACA	55	5	0.000	0.633	0.555	0.145
50AeKVR-169	F:TGGTTTCAATATGCACCATCCA R:ACTAACTAACGAACGAACGAACG	55	2	0.000	0.444	0.341	0.260
Mean			8.370	0.218	0.561	0.519	Total PI: 3.9490×10^{-43}

was performed in a 15 μL volume containing 75–100 ng okra DNA, 1.5 μL Taq Buffer 10x (Tris pH with 15 mm $MgCl_2$). dNTPs 1.5 μL of (10 mM), 0.5 μL of forward primer M13 tail (5 pM), 0.5 μL of reverse primer M13 tail (5 pM), 1 μL of fluorescent probes FAM, VIC, PET and NED (5 pM) M13 primers, 0.5 μL of *Taq* polymerase (3 u per μL), and nuclease free water (8.7 μL). All the PCR reactions were performed using a Bioer Life Pro Thermal cycler (BioerTechnology, China). The amplified PCR products labelled with FAM, VIC, PET or NED were pooled before separation on the ABI 3730 Genetic Analyzer (Applied Biosystem, USA). The data obtained was further analysed using the Peak Scanner software (Applied Biosystems, USA) to determine exact fragment size of the PCR product.

Statistical analysis

The data (fragment size) were used for calculating the number of alleles per locus, observed heterozygosity (H_o), expected heterozygosity (H_e) and PIC, probability of identity (PI) using Cervus 3.0 software (Kalinowski *et al.* 2007). The dendrogram analysis for 29 genotypes and wild species was done separately following neighbour-joining method (NJ) using Darwin software (Perrier *et al.* 2003; DARwin software. http://darwin.cirad.fr/darwin).

Table 3. Distribution of microsatellite types in okra genome.

	Repeats	Percentage
Monorepeats	572	21.10
Direpeats	607	22.40
Trirepeats	345	12.70
Tetrarepeats	462	17.06
Pentarepeats	294	10.80
Hexarepeats	149	05.50
Complexrepeats	279	10.30

Results

Assembly statistics and designing of primers

The total sequences obtained using Roche 454 Titanium platform was 979,806 bp from 61,722 reads. These reads were assembled into 3735 contigs. Among these, 2708 contigs had microsatellites and were identified using the software MISA. They had 1303 perfect microsatellites (including 572 monorepeats, 607 direpeats, 345 trirepeats, 462 tetrarepeats, 294 pentarepeats and 149 hexarepeats) and 279 complexrepeats (table 3). However, we were able to design primers for 402 microsatellite loci (ESM1) using Primer3 software (Untergasser *et al.* 2012). Direpeats were found to be more abundant class of microsatellite than the tri, tetra, penta and hexarepeats (table 3).

Genetic analysis

The allelic data such as number of alleles, observed heterozygosity, expected heterozygosity and the allele per locus were analysed using Cervus 3.0 software. The values estimated for observed heterozygosity ranged from 0.000–0750 with a mean of 0.218; the values for expected heterozygosity ranged from 0.000–0.972 with a mean of 0.561; alleles per locus ranged from 1 to 27. The mean value of PIC was 0.519. The PI values ranged from 0.003 to 0.681 with a combined PI for all the loci is 3.949×10^{-43} (table 2). Cross species amplification showed transferability of 78 to 92% among the wild species, for *A. manihot* var.*tetraphyllus* amplified 84%; *A. ficulneus* amplified 78%; *A. moschatus* amplified 92%; *A. angulosus* var.*grandiflorus* amplified 90% and *A. caillei* amplified 86% of SSR primers standardized (table 4). Dendrogram analysis of data showed that genotypes used here are classified into three groups (figure 1). The genetic relatedness among the wild related species with *A. esculentus* also showed three groups among them (figure 2).

Discussion

Several vegetable crops, e.g. okra lack genomic resources such as microsatellite markers. This has resulted in a very few studies on the genetic diversity and genome mapping (Gulsen *et al.* 2007; Sunday *et al.* 2008). Hence, molecular markers like microsatellite needs to be generated which are cost effective in genotyping. They also have a wide application in rapid identification of species, cultivars and in crop improvement programme. At present, the discovery of SSR or microsatellite loci can be done by sequencing whole or partial genome through NGS technology (Seo *et al.* 2012; Ravishankar *et al.* 2015).

NGS approaches generate highly accurate and quality sequence data as it covers the same genomic region multiple times (Gilles *et al*. 2011). In this study, the 454 technology GS FLX Titanium platform was used to identify microsatellites sequences in okra genomic sequences and further to evaluate the genetic relationships among the okra cultivars and species.

In a good number of species belonging to Malavceae family, the microsatellite markers were developed through the microsatellite-enriched library method (Phuekvilai and Wolff 2013; Rymer *et al.* 2014). This method is cumbersome, time consuming and only a very few microsatellite markers can be identified. Recently there are few reports on developing microsatellite using transcriptome data (Fan *et al.* 2013; Ju *et al.* 2015; Schafleitner *et al.* 2013). The major disadvantage of transcriptome-derived microsatellite is that they suffer from low polymorphism. However, until now there are no reports of microsatellite development using genomic DNA sequences in okra.

In this study, we employed NGS approach to develop SSR markers for okra and reported SSRs using GS FLX Titanium pyrosequencing technology with a total of 979,806 bp sequences from a partial genomic DNA sequence of cv. Arka Anamika. The frequency analysis of repeats revealed that direpeats were in maximum frequency followed by mono, tetra, tri, penta and hexarepeats. Generally, in plants, mononucleotide repeats were present in high number (rice, sorghum, Arabidopsis, Medicago and Populus) species (Sonah *et al.* 2011). Apart from direpeats, in this study, monorepeats were the most prevalent SSR type okra genomic DNA, accounting 20.1% of all SSRs identified, followed by tetranucleotide repeats (17.06%; table 3). While the mono, di, tri and tetranucleotide repeats contribute to the major proportion of microsatellites (73.26%). However, trirepeats were in high frequency for microsatellites derived from transcriptome data in okra (Schafleitner *et al.* 2013). This is expected because trirepeats result in extension or shortening of amino acid repeats motifs and other type of repeat cause frame-shift mutations. The mechanism and selection pressure for the origin and evolution of microsatellite repeats is not clearly understood. The relative dominant occurrence of repeat motif of a particular sequence type and its length in plant genome may be the outcome of certain selection pressure applied on that specific motif during evolution. The most common mutation mechanisms assumed to be operating are replication slippage,

Table 4. Cross species transferability of orka genomic micrsatellite markers.

Primers	*A. tetraphyllus*	*A. ficulneus*	*A. moschatus*	*A. angulosus grandiflorus*	*A. caillei*
AeKVR-102	A	NA	NA	A	A
AeKVR-103	NA	A	A	NA	A
AeKVR-110	A	A	NA	A	A
AeKVR-111	A	A	A	A	A
AeKVR-132	A	A	A	A	A
AeKVR-125	A	A	A	A	A
AeKVR-116	A	NA	A	A	A
AeKVR-137	A	A	A	NA	A
AeKVR-134	A	A	A	A	A
AeKVR-156	A	A	A	A	A
AeKVR-131	A	A	A	A	A
AeKVR-173	A	NA	A	A	A
AeKVR-153	A	A	A	A	A
AeKVR-114	A	A	A	A	A
AeKVR-192	A	NA	A	A	A
AeKVR-193	A	A	A	A	NA
AeKVR-113	A	NA	A	NA	A
AeKVR-118	NA	A	A	NA	A
AeKVR-182	A	A	A	A	A
AeKVR-154	A	A	A	A	NA
AeKVR-176	A	NA	A	A	A
AeKVR-195	NA	A	A	A	NA
AeKVR-117	A	NA	A	A	A
AeKVR-119	NA	A	A	A	A
AeKVR-161	NA	A	A	A	NA
AeKVR-141	A	A	A	A	A
AeKVR-177	A	A	A	A	A
AeKVR-199	A	A	A	A	A
AeKVR-172	A	A	A	A	A
AeKVR-160	A	A	A	A	A
AeKVR-149	A	A	A	A	A
AeKVR-162	NA	A	A	A	A
AeKVR-147	A	A	A	A	A
AeKVR-183	A	A	NA	A	A
AeKVR-101	A	A	A	A	A
AeKVR-104	A	A	A	A	A
AeKVR-108	A	A	A	A	A
AeKVR-145	A	A	A	A	NA
AeKVR-187	A	NA	A	A	A
AeKVR-107	A	A	A	A	A
AeKVR-200	A	NA	NA	A	A
AeKVR-106	A	A	A	A	NA
AeKVR-143	A	A	A	A	A
AeKVR-196	NA	NA	A	A	A
AeKVR-126	A	A	A	A	A
AeKVR-194	A	A	A	A	A
AeKVR-165	NA	A	A	A	A
AeKVR-163	A	NA	A	A	A
AeKVR-167	A	A	A	NA	NA
AeKVR-169	A	A	A	A	A
Transferability	84%	78%	92%	90%	86%

A, amplified; NA, not amplified.

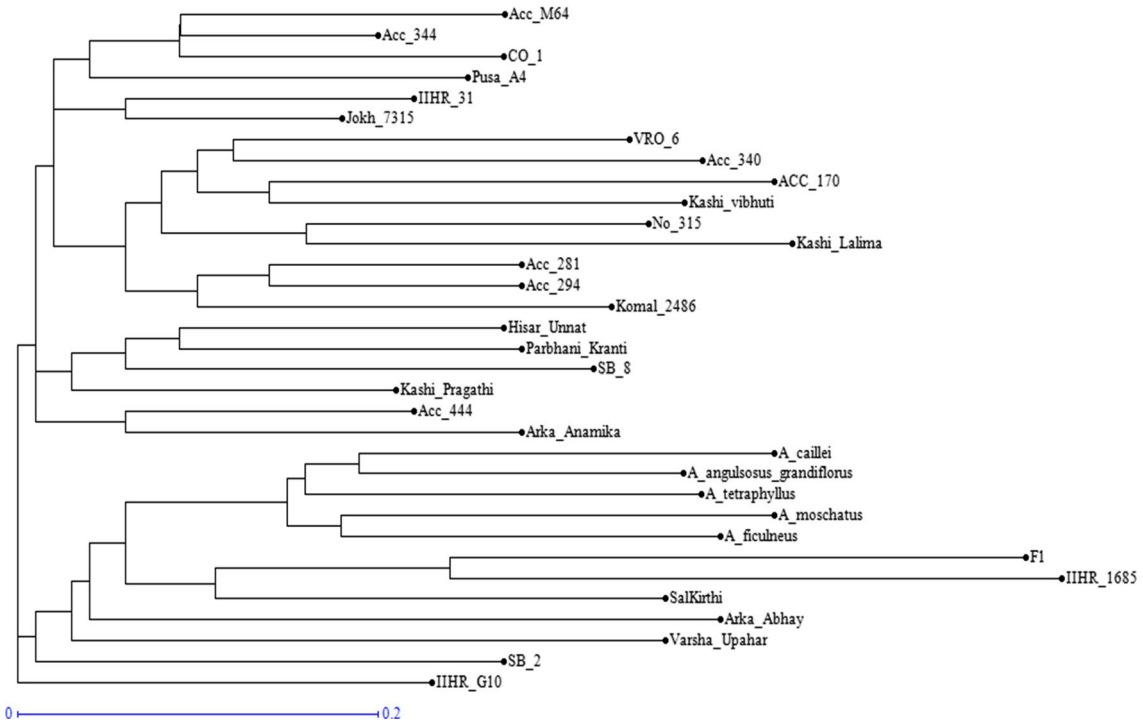

Figure 1. Dendrogram analysis showing the genetic relationship among *A. esculentus* L. accession using SSR marker data.

Figure 2. Dendrogram analysis showing the genetic relationship among the *A. esculentus* L. genotypes and wild related species using SSR marker data.

and unequal crossing over which might contribute to the addition or removal of one or more motifs, and variation in the length (Buschiazzo and Gemmell 2006; Sonah *et al.* 2011).

In the present study, we report the successful development of 402 SSRs markers (ESM1) of which 50 SSR primers were screened over 28 okra genotypes, 22 SSR primers were found to be highly polymorphic markers with the PIC values more than 0.5. Overall, the statistical analysis revealed that the observed heterozygosity ranged from 0.000–0.750 with a mean value of 0.218; the values for expected heterozygosity ranged from 0.000–0.972 with a mean value of 0.561; allele per locus ranged from 1 to 27 with a mean value of 8.37, and the mean PIC was found to be 0.519 (table 4). Similar kind of work was done for the development and characterization of SSR marker in Gossypium with the mean PIC value observed as 0.65 (John *et al.* 2012) and in another study, genetic diversity of cotton with the PIC values ranged from 0.34 to 0.86 and the mean value PIC was 0.80 (Muhammad *et al.* 2013). Forty-nine markers of 50 showed polyrmorphism in this study indicating presence of diversity. Here, the PI was calculated and values ranged from 0.003 to 0.681 and combined PI value is 3.949×10^{-43}. These low values indicate that these microsatellite markers can be employed for DNA finger printing.

The dendrogram analysis classified okra genotypes into three major groups (figure 1). The genotypes IIHR-31, Jokh7315 and Pusa A4 are closely related and grouped together. Similarly, genetic relatedness study using SSR marker grouped species into three clusters. *A. tetraphyllus* grouped seperately and even *A. esculentus* genotypes also formed a separate group. Cross species amplification showed 78 to 92% of transferability among the wild-related species (*A. manihot* var. *tetraphyllus* amplified 84%; *A. ficulneus* amplified 78%; *A. moschatus* amplified 92%; *A. angulosus* var. *grandiflorus* amplified 90% and *A. caillei* amplified 86%). However, the extent of transferablility has no relationship with their genetic relatedness. This may be due to distinct alleles present in each species (table 4).

Based on the results of the analysis, we conclude that, NGS is an efficient method for the development of microsatellite markers in okra. We can develop a large number of microsatellite markers using NGS sequence data from okra. The microsatellite markers developed here are a useful tool for the genetic diversity studies, crop improvement programmes and also in gene discovery.

Acknowledgements

Authors acknowledge the financial support through institute project from ICAR-Indian Institute of Horticultural Research, Bengaluru.

References

Aladele S. E., Ariyo O. J. and Lapena R. 2008 Genetic relationships among West African okra (*Abelmoschuscaillei*) and Asian genotypes (*Abelmoschu sesculentus*) using RAPD. *Afr. J. Biotechnol.* **7**, 1426–1431.

Andre G., Emese M., Nicolas P., Stephanie F., Thibaut M. and Jean F. M. 2011 Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics* **12**, 245.

Bertini C. D., Schuster I., Sediyama T., de Barros E. G. and Moreira M. A. 2006 Characterization and genetic diversity analysis of cotton cultivars using microsatellites. *Genet. Mol. Biol.* **29**, 321–329.

Buschiazzo E. and Gemmell N. 2006 The rise fall and renaissance of microsatellites in eukaryotic genomes. *BioEssays*. **28**, 1040–1050.

Chakravarthi B. K. and Naravaneni R. 2006 SSR Marker based DNA fingerprinting and diversity study in rice (*Oriza sativa* L.). *Afr. J. Biotechnol.* **5**, 684–688.

Fan Q., Chen S., Li M., He S., Zhou R. and Liao W. 2013 Development and characterization of microsatellite markers from the transcriptome of *Firmiana danxiaensis* (Malvaceaes.l.). *Appl. Plant Sci*. **28**, 12–16.

Gilles A., Meglé Cz E., Pech N., Ferreira S., Malausa T., and Martin J. F. 2011 Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics* **12**, 245.

Gulsen O., Karagul S. and Abak K. 2007 Diversity and relationships among Turkish okra germplasm by SRAP and phenotypic marker polymorphism. *Biologia* **62**, 41–45.

Huang and Anup M. 1999 CAP3: A DNA sequence assembly programe. *Genome Res*. **9**, 868–877.

John Z., Yu D. D., Fang, Russell J., Mauricio U. L. L., Hinze R. G. *et al.* 2012 Development of a core set of SSR markers for the characterization of *Gossypium* germplasm. *Euphytica* **187**, 203–213.

Joshi A. B. and Hardas M. W. 1956 Alloploid nature of okra, *Abelmoschu sesculentus* (L.) Moench. *Nature* **178**, 1190.

Ju M. M., Ma H. C., Xin P. Y., Zhou Z. L. and Tian B. 2015 Development and characterization of EST-SSR markers in *Bomba xceiba* (Malvaceae). *Appl. Plant Sci*. **3**, 1–4.

Kalinowski S. T., Taper M. L. and Marshall T. C. 2007 Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol*. **16**, 1099–1006.

Muhammad S., Bahaeldeen B. M., Muhammad A. A., Agung N. P., Kamran S. B., Ahmad A. S. and Tayyab H. 2013 Genetic diversity assessment of cotton (*Gossypium hirsutum* L.) genotypes from pakistan using simple sequence repeat markers. *Aust. J. Crop Sci*. **7**, 261–267.

Nwangburuka C. C., Kehinde O. B., Ojo D. K., Denton O. A. and Popoola A. R. 2011 Morphological classification of genetic diversity in cultivated okra, *Abelmoschu sesculentus* (L.) Moench using principal component analysis (PCA) and single linkage cluster analysis (SLCA). *Afr. J. Biotechnol*. **10**, 54.

Oetting W. S., Lee H. K., Flanders D. J. and Wiesner G. L. 1995 Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. *Genomics* **30**, 450–458.

Perrier X., Flori A. and Bonnot F. 2003 Data analysis methods. *Enfield. Sci. Publish. Montpellier*. 43–76.

Phuekvilai P. and Wolff K. 2013 Characterization of microsatellite loci in *Tiliap latyphyllos* (Malvaceae) and cross-amplification in related species. *Appl. Plant Sci*. **1**. https://doi.org/10.3732/apps.1200386.

Prakash K., Pitchaimuthu M. and Ravishankar K. V. 2011 Assessment of genetic relatedness among okra genotypes (*Abelmoschu sesculentus* L. Moench) using RAPD markers. *Electron. J. Plant Breed*. **2**, 80–86.

Ramya P. and Bhat K. V. 2012 Analysis of phylogenetic relationships in *Abelmoschus* species (Malvaceae) using ribosomal and chloroplast intergenic spacers. *Indian J. Genet. Plant Breed*. **72**, 445–453.

Ravishankar K. V., Dinesh M. R., Nischita P. and Sandya B. S. 2015 Development and characterization of microsatellite markers in mango *(Mangifera indica)* using next-generation sequencing technology and their transferability across species. *Mol. Breed*. **35**, 1–13.

Ravishankar K. V., Lalitha A. and Dinesh M. R. 2000 Assesment of genetic relatedness among a few Indian mango varieties using RAPD marker. *J. Hortic. Sci. Biotechnol*. **75**, 198–201.

Rymer P. D., Harris S. and Boshier D. 2014 Isolation and characterization of microsatellite markers for the neotropical tree *Pachir aquinata* (Malvaceae). *Conserv. Genet. Resour*. **6**, 375–382.

Sawadogo M., Ouedraogojt B. D., Ouedraogo M., Gowdabs B. C. and Timko M. 2009 The use of cross species SSR primers to study genetic diversity of okra from Burkina Faso. *Afr. J. Biotechnol*. **8**, 2476–2482.

Schafleitner R., Kumar S., Lin C. Y., Hegde S. G. and Ebert A. 2013 The okra (*Abelmoschu sesculentus*) transcriptome as a source for gene sequence information and molecular markers for diversity analysis. *Gene* **517**, 27–36.

Schuelke M. 2000 An economic method for the fluorescent labeled of PCR fragment. *Nat. Biotechnol*. **18**, 233–234.

Seo K. I., Lee G., Park S. K., Yoon M. S., Ma K. H., Lee J. R. et al. 2012 Genome shotgun sequencing and development of microsatellite markers for gerbera (*Gerbera hybrid* H.) by 454 GS-FLX. *Afr. J. Biotechnol*. **11**, 7388–7396.

Sonah H., Deshmukh R. K., Sharma A., Singh V. P., Gupta D. K., Gacche R. N. et al. 2011 Genome wide distribution and organization of microsatellites in plants: an insight into marker development in *Brachypodium*. *PLoS One* **10**, 137.

Sunday E., Aladele O. J. Ariyo L. and Robert D. L. 2008 Genetic relationship among West African Okra (*Abelmo schuscaillei*) and Asian genotypes (*Abelmo schusesculentus*) using RAPD. *Afri. J. Biotechnol*. **7**, 1426–1431.

Suping F., Helin T., You C., Jingyi W., Yeyuan C., Guangming S. et al. 2013 Development of pineapple microsatellite markers and germplasm genetic diversity analysis. *Biomed. Res. Int.* 1–11. https://doi.org/10.1155/2013/317912.

Untergasser A., Cutcutache J., Koressaar T., Ye J., Faircloth B. C., Remm M. and Rozen S. G 2012 Rozen.Primer3 - new capabilities and interfaces. *Nucleic. Acids. Res.* **40**, 15.

Wang S., Xiaokun L., Jianwei L., Juan H., Gang C., Xinxin X. *et al.* 2013 Effects of controlled-release urea application on the growth, yield and nitrogen recovery efficiency of cotton. *Agric. Sci*. **4**, 3–38.

Yuan C. Y., Wang P., Chen P. P., Xiao W. J., Zhang C., Hu S. *et al.* 2015 Genetic diversity revealed by morphological traits and ISSR markers in 48 okras (*Abelmoschus escullentus* L.). *Physiol. Mol. Biol. Plant*. **21**, 359–364.

Yuan C. Y., Zhang C., Wang P., Hu S., Chang H. P., Xiao W. J. et al. 2014 Genetic diversity analysis of okra (*Abelmoschus esculentus* L.) by inter-simple sequence repeat (ISSR) markers. *Genet. Mol. Res.***13**, 3165–3175.

Corresponding editor: MANOJ PRASAD