TÜBİTAK

Turkish Journal of Agriculture and Forestry

http://journals.tubitak.gov.tr/agriculture/

Research Article

Turk J Agric For
(2014) 38: 603-614

doi:10.3906/tar-1312-104

Development of microsatellite markers in sesame (*Sesamum indicum* L.)

Jyothi BADRI[1,*], Vijay YEPURI[2], Anuradha GHANTA[2], Sivaramakrishnan SIVA[2], Ebrahimali Abubacker SIDDIQ[2]

[1]Directorate of Rice Research, Rajendranagar, Hyderabad, India

[2]Institute of Biotechnology, Acharya N. G. Ranga Agricultural University, Rajendranagar, Hyderabad, India

Received: 23.12.2013 • **Accepted:** 22.05.2014 • **Published Online:** 15.08.2014 • **Printed:** 12.09.2014

Abstract: Sesame (*Sesamum indicum* L.) is an important oilseed crop with high-quality seed oil and many antioxidant properties. Owing to its commercial and medicinal values, there is a renewed interest among agricultural scientists in this ancient crop. Efforts to strengthen the sesame-specific marker base have been initiated in the recent past; however, the available number of microsatellite markers is still not sufficient for the development of high-resolution genetic linkage maps for important agronomic traits and there is a need to increase the number of informative DNA markers in sesame. In the present study, we developed 25 microsatellite markers by employing the selective hybridization strategy and 95 mining expressed sequence tags of the NCBI database. This new set of microsatellite markers was characterized and screened for genetic diversity in an array of 16 sesame germplasms. Of the 120 SSRs, 92 were polymorphic, consisting of 18 SSRs from selective hybridization and 74 from the EST data set. The number of alleles per microsatellite locus ranged from 2 to 5, with an average of 3.11 alleles. The allele size ranged widely (100–510 bp) among the primer pairs. Polymorphic information content estimates ranged from 0.2982 to 0.912. Jaccard's similarity coefficient ranged from 0.21 to 0.82. The potential of the markers was assessed by diversity analysis using the sequential hierarchical agglomerative nonoverlapping clustering technique of the unweighted pair group method of arithmetic means on a set of 16 genotypes of sesame, including 2 wild species. Results supported the hypothesis that *S. malabaricum* could be the immediate progenitor of the cultivar species and that *S. mulayanum* is distinct from *S. malabaricum* and *S. indicum*, while suggesting hardly any diversity among the cultivars.

Key words: EST-SSRs, microsatellite markers, sesame, variability

1. Introduction

Sesame (*Sesamum indicum* L.) is one of the oldest oilseed crops known to man and it is valued for its high-quality seed oil. Sesame seeds are an important source of oil (44%–58%), protein (18%–25%), and carbohydrates (13.5%) (Bedigian et al., 1985). The greater stability of the oil is due to the presence of natural antioxidants such as sesamin, sesamol (Brar and Ahuja, 1979; Ashri, 1998), sesaminol, sesamolinol, and squalene (Mohamed and Awatif, 1998). Among the primary edible oils, sesame oil has the highest antioxidant content (Cheung et al., 2007) and contains abundant fatty acids such as oleic acid (43%), linoleic acid (35%), palmitic acid (11%), and stearic acid (7%) (Bedigian et al., 1985). The confectionary value of the seed and the exceptionally superior quality of the oil has enabled sesame to emerge as an important commodity in international trade. As for its medicinal value, sesame reduces plasma cholesterol and thereby lowers blood pressure (Sankar et al., 2005). Such features and benefits of sesame have recently renewed the interest of agricultural scientists in this ancient crop (Laurentin and Karlovsky, 2006), previously considered an orphan crop as far as genomic tools are concerned.

Molecular marker technology employment is still at its infancy when it comes to sesame. The major limiting factor in many studies, excluding those of Wei et al. (2009) and Zhang et al. (2012), is the lack of crop-specific marker-based high-density linkage maps. Compared to other oil seed crops, there are very few efforts in the development of sesame-specific microsatellite markers. The majority of the molecular marker-based genetic diversity studies in sesame employed DNA markers like random amplified polymorphic DNA (RAPD) (Bhat et al., 1999; Nathan Kumar et al., 2000; Davila et al., 2003; Ercan et al., 2004; Abdellatef et al., 2008), intersimple sequence repeats (ISSRs) (Kim et al., 2002), both RAPD and ISSRs (Sharma, 2009), amplified fragment length polymorphisms (AFLPs) (Laurentin and Karlovsky, 2006; Ali et al., 2007), and restriction fragment length polymorphisms (Yamada et al., 1993).

Dixit et al. (2005) were the first to report polymorphic SSRs in sesame; that study only named 10, compared to

* Correspondence: jyothi_rishik@yahoo.com

the more than 1000 SSR loci mapped in other oilseed crops like soybean (Song et al., 2004; Xin et al., 2012) and groundnut (Ferguson et al., 2004; Wang et al., 2012). However, with the establishment of expressed sequence tag (EST) sequencing projects for gene discovery programs in several plant species, a wealth of DNA sequence information has been generated and deposited in online databases. Wei et al. (2008) assembled a total of 1785 nonredundant EST sets among 3328 identified sesame ESTs, and 50 primer pairs were designed in the flanking regions of repeat-containing ESTs. Transcriptomic studies using Illumina paired-end sequencing revealed a wealth of unigenes, from which 40 polymorphic EST-SSRs were obtained (Wei et al., 2011). Wang et al. (2012) detected 59 cDNA-SSRs among 60,960 unigenes deposited in GenBank. Spandana et al. (2012) reported 111 SSRs in sesame and Vijay et al. (2013) reported 156 EST-SSRs, along with primer sequence information, from 16,619 ESTs mined from GenBank. Zhang et al. (2012) used RNA sequencing to develop a huge set of 2164 genic SSRs, of which 276 have shown successful amplification. The recent advances made in the development of molecular tools in sesame are encouraging; however, to make use of a high-density molecular linkage map, the number of available SSR markers is still very meager. Further research efforts are needed to develop sesame-specific markers in abundance to make use of the variability present in the sesame germplasm.

Keeping in view the need for informative DNA markers in sesame improvement, an attempt was made to develop sesame-specific SSR markers using a selective hybridization approach and mining of EST-SSRs from the NCBI database. The SSRs identified in the present study were characterized and used in the diversity analysis of sesame germplasm consisting of varieties from cultivated species and accessions from wild sesame.

2. Materials and methods

Selective hybridization and isolation from the EST-database were the 2 strategies employed for the isolation of microsatellite markers. In the selective hybridization approach, the varieties Swetha, RT-54, TMV-3, and MKN-6 of the cultivated species (*Sesamum indicum* L.) were used as a source of DNA. DNA was isolated from the leaves of 1-month-old plants using the protocol of Laurentin and Karlovsky (2006), with modifications to account for mucilage interfering with the DNA isolation procedure. The basic protocol of the selective hybridization method was used here, though with a different combination of restriction enzymes and oligo repeats. The genomic DNA (2 µg) was digested with 4 base pair cutter restriction enzymes (R*sa*I, B*st*UI, N*he*I, and B*fu*I) in 4 separate reactions in a reaction volume of 50 µL. The 4 restriction digestions were incubated at 37 °C for 4 h. The digestion was confirmed by identifying a dense smear within the 100–1000 bp range.

Restriction fragments were ligated overnight with freshly prepared double-stranded super SNX linkers (MWG Biotech) at a ligation temperature of 16 °C. A PCR check was performed with a linker sequence as the primer to ensure successful ligation. Linker-ligated DNA was hybridized with 3 biotinylated oligo repeats [$(AT)_{12}$, $(CT)_{10}$, and $(TCG)_{10}$] at a hybridization temperature of 60 °C using 2X hybridization solution in 3 separate hybridization reactions. Prior to incubation, hybridization mixtures were heated at 95 °C for 5 min, followed by quick chilling on ice for 2 min. Hybridization mixtures were conjugated with streptavidin-coated magnetic beads at room temperature for 30 min with constant gentle agitation. For this, 50 µL of beads for each hybridization reaction were washed twice in $T_{10}E_2$, followed by 1X hybridization solution using a magnetic particle concentrator, and were finally suspended in 150 µL of 1X hybridization solution.

After the bead-hybridized fragment complex was washed twice in 2X and 4 times in 1X SSC along with 0.1% SDS, repeat-enriched DNA was extracted in TLE, NaOAc/EDTA solution, and 95% ethanol. DNA fragments enriched with oligonucleotide repeats were separated from the beads after this process. This DNA mixture was incubated on ice for more than 15 min and centrifuged at full speed for 10 min. The supernatant was pipetted out and the enriched DNA was air-dried to form a pellet known as 'pure-gold DNA'. The pellet was resuspended in 25 µL of TLE. PCR was performed to increase the quantity of the gold DNA. PCR-amplified gold DNA was ligated into the pGEM T-easy cloning vector using T_4 DNA ligase and T_4 DNA ligase buffer at 14 °C overnight. Repeat-containing DNA fragments ligated into plasmid are also known as recombinant DNA (rDNA).

The DH10B strain of *E. coli* was used as a host for transformation with rDNA. Competent cells of DH10B were prepared as described by Sambrook and Russell (2001). Competent *E. coli* cells were transformed with rDNA using an electroporation apparatus and the transformed mixture was plated on LB amp+ agar plates coated with IPTG and X-gal. The plates were incubated at 37 °C for 16 h. Only white colonies were picked up for further colony PCR and primary culture inoculation. Plasmids containing inserts larger than 300 bp, according to insert PCR results, were selected and diluted to 100 ng/µL and sequenced using an automated sequencer. Raw sequences were extracted into FASTA format using Chromas Lite software (http://www.technelysium.com.au/chromas_lite.html). Microsatellite Analysis Server (MICAS) software (http://sunserver.cdfd.org.in:8080/MIC/index.html) and SSRIT (http://www.gramene.org/)

were used for screening the sequenced colonies. SSR-containing clones were submitted to VecScreen software (http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html) and genomic regions having strong match with vector sequences were eliminated. After deleting vector-contaminated regions, primers were designed for clones that were found to possess flanking sequences for primer design. Primers were designed using Primer 3 software (http://bioinfo.ut.ee/primer3-0.4.0/primer3/) for primers of lengths of 18–22 bp and product sizes of 100–450 bp.

A total of 3328 EST sequences were downloaded from the NCBI database and screened for repeat-containing sequences using SSRIT. Primers were designed using Primer 3 software for repeat-containing sequences possessing sufficient flanking sequences on either side of the repeat region.

For characterization of newly developed SSRs, DNA amplification was performed in a reaction volume of 10 µL containing 50 ng of DNA template (2 µL), 1X PCR reaction buffer (15 mM Tris-HCl) (1 µL), 2 mM dNTPs (1 µL), 1 U of Taq DNA polymerase, 10 µM of forward and reverse primers (0.5 µL), and sterile distilled water (4.8 µL). PCR was carried out in a Veriti 96-well thermal cycler (Applied Biosystems). The PCR conditions were programmed for an initial denaturation step of 94 °C for 5 min, followed by 35 cycles of denaturation at 94 °C (annealing temperature) for 45 s, extension carried out at 72 °C for 1 min, and then a final extension of 72 °C for 10 min. The PCR products were fractionated on a 3% metaphor agarose (Lonza) gel with 0.05 µg μL^{-1} ethidium bromide. Samples were loaded with a reference 50-bp DNA ladder (NEB). Gels were electrophoresed at 120 V. After separation, gels were documented using Molecular Imager Gel Doc (BIO-RAD).

After characterization, SSR primer pairs were used to assess genetic diversity in 16 accessions including 13 cultivars (TKG-22, TMV-3, RT-54, Swetha, Hima, Chandana, NSKMS-128, Uma, VRI-1, Rajeshwari, TAC-89-309, Paiyur-1, and CO-1), 2 accessions of *S. mulayanum* (BBL-46-2K and BB-3-8), and 1 accession of *S. malabaricum* (IC-204492), the last 2 species being wild. Amplification products were visualized on 3% MetaPhor agarose gel and were scored using a binary code for presence ("1") or absence ("0") of bands (alleles) for every SSR locus.

Only DNA bands in the range of best resolution in the agarose gel (approximately 100 to 500 bp) were counted. For each genotype, the presence or absence of each band was determined and designated as "1" if present and "0" if absent. The genetic distance between individuals was estimated by using the markers that produced the expected size (100–500 bp) of amplification product. Polymorphic information content (PIC) was calculated as described by Botstein et al. (1980) using the below formula:

$$PIC = 1 - [\Sigma P_i^2] - [\Sigma\Sigma 2P_i^2 P_j^2] \quad \begin{matrix} n\ n-1\ n \\ I=1\ I=1j=I+1 \end{matrix}$$

where p_i equals the frequency of the ith allele and p_j is the frequency of the $(I+1)$th allele. This computation was done using Genstat 7.10. For diversity analysis, only data from polymorphic SSR loci were used. Genetic diversity was estimated by computing the mean number of pair-wise differences over each locus among SSR binary phenotypes using Genstat 7.10 software. Similarities between any 2 genotypes were estimated according to Nei and Li (1979):

$S_{ij} = 2\ N_{ij}\ (N_i + N_j)$,

where N_{ij} is the number of bands in common accessions i and j, and N_i and N_j are the total number of bands in common between any 2 accessions and may range from 0 (no common bands) to 1 (identical band profile for the 2 accessions).

A dendrogram was constructed based on the S_{ij} values by adopting the sequential hierarchical agglomerative nonoverlapping clustering technique of the unweighted pair group method of arithmetic means (UPGMA), which is a variant of the average linkage clustering algorithm (Sneath and Sokal, 1973). These computations were performed using the statistical analysis package NTSYS-pc v2.10t (Rohlf, 1994).

3. Results

In the selective hybridization approach, digestion with R*sa*I and B*fu*I enzymes was complete, resulting in a continuous smear of fragments ranging from 100 to 1000 bp, while it was only a partial digestion resulting in a discontinuous smear throughout and a thick smear of high molecular weight DNA close to the well. Hence, only restriction fragments from R*sa*I and B*fu*I digestions were used for ligation with super SNX linkers.

A total of 350 white colonies (positive for the insert) were obtained after transformation; of them, 172 colonies were randomly picked for colony PCR. All of them were harboring plasmids with inserts of the expected size. Plasmids were extracted from the colonies and inserts were sequenced. Figure 1 shows the raw sequence of clone M82 containing $(TC)_{17}$ and $(AC)_{17}$ repeats. Out of 172 sequenced colonies, 52 were found to have unique SSR regions. SSR-containing clones were submitted to VecScreen software and genomic regions that strongly matched vector sequences were eliminated. After deleting vector-contaminated regions, primers were designed for clones that were found to possess flanking sequences for primer design. Primers were designed for 25 clones, which contained 38 repeat sequences (Table 1). The sesame genomic library containing these 25 clones obtained in the

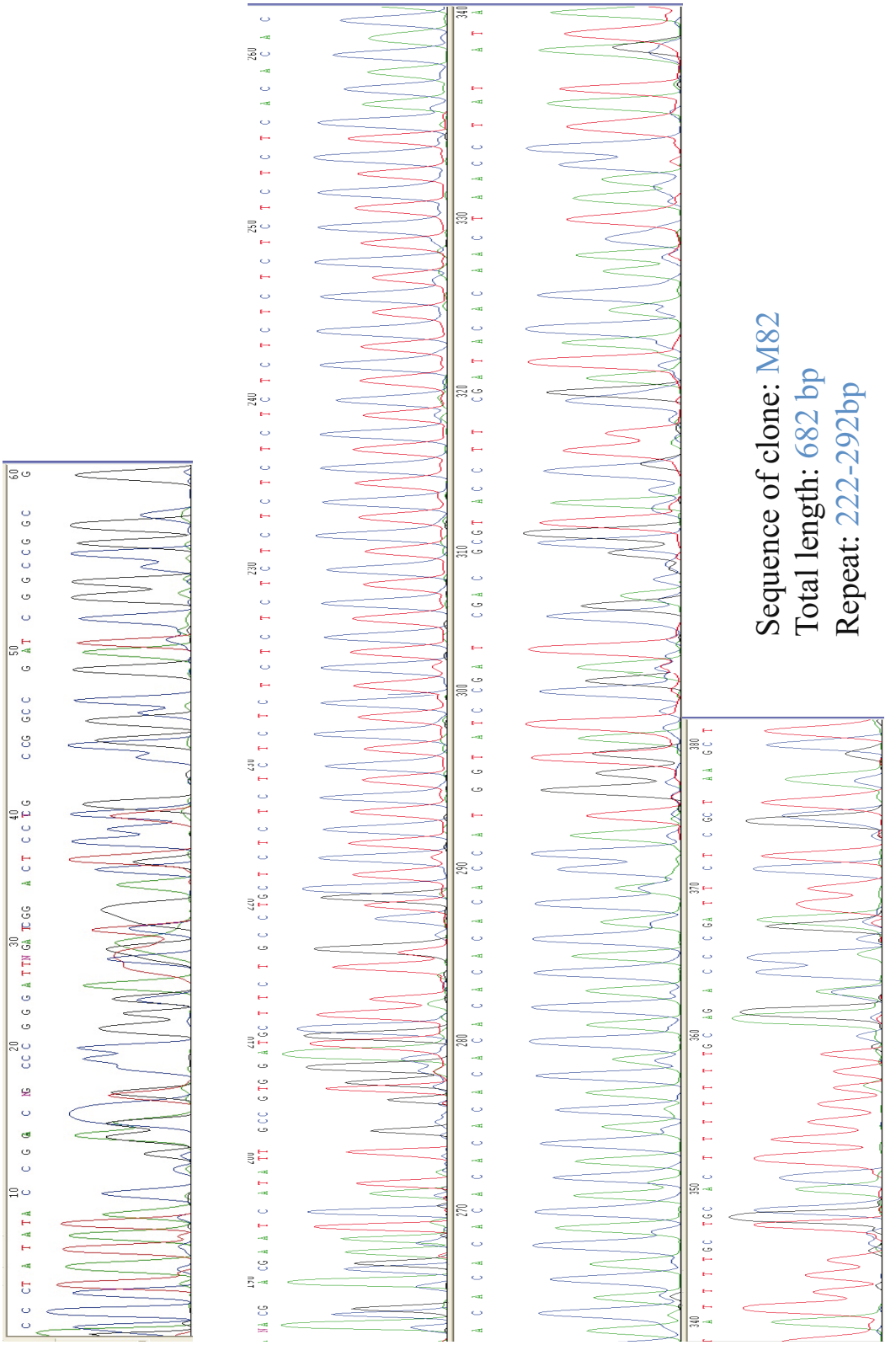

Figure 1. Raw sequence of clone number M82 containing (TC)17(AC)17 repeats.

Table 1. Characteristics of SSR markers isolated using the selective hybridization approach.

No.	GenBank accession ID	Clone	Repeat motif	Primer sequence	Primer ID	Ta (°C)	Allele size (bp)	No. of alleles	PIC
1	HQ224869	M9	(GA)23	GCGGGAAATTCGATTGTTTA (F) CGAGGCAGATCATGAGGTTT (R)	SM1	52	190–240	3	0.432
2	HQ224870	M24	(TCG)7	TAGCAGAATCGTCCCAGTCC (F) AGCAGAATCCGTGGCTTAGA (R)	SM2	57	385–400	4	0.766
3	HQ224872	M82	(TC)17(AC)17	ATCATATTGCCGTGGATGCT (F) CTTAGCGAGAATCGGGTCTG (R)	SM4	50	190–210	4	0.652
4	HQ224873	M93	(TC)13	GCAGAATCACTGCAGAAGGA (F) AACCAACAACCGCTTTTACG (R)	SM5	57	294–328	2	0.431
5	HQ224874	M99	(CT)10, (AGC)5	AGAATCGGTTCTCTGCTGCT (F) ATGACGAGAACGGAAAGAGC (R)	SM6	53	440–480	3	0.436
6	HQ224875	M104	(TC)32	CCGGCTATTCCTCAGTTGTC (F) GTAAGCGTTGCCAGAAGGAC (R)	SM7	58	190–210	3	0.81
7	HQ224876	M106	(AC)10, (TC)7	CACGCTCGAACTCTCTCCTT (F) GACTTGTCCGACCATCCATC (R)	SM8	58	420–480	4	0.83
8	HQ236477	M112	(TC)6,(TC)7, (TC)7,(CT)6	AGGCTGCTCTGGGACTTCTT (F) CATCAAACGCCTTTTTACGG (R)	SM9	57	450–480	3	0.753
9	HQ236478	M68	(GTAATG)4, (ATGGCA)4, (ATGGCA)3	TAGGCGAATGGGGTTTAAGG (F) CGCCTTTGAGTGTGACTTGA (R)	SM10	53	380–400	3	0.843
10	HQ236480	M12	(CT)3(CA)2(TA)2	CACATAGAGTTTGCGGCTCA (F) ATCCCTCGGGCTCAATCTAT (R)	SM12	57	190–210	2	0.322
11	HQ236482	M21	(GTC)5	TCTTCTTGGCACGACACTTG (F) CAAAAGAGCCGGAAAGACAG (R)	SM14	53	100–130	3	0.82
12	HQ236483	M32	(GCC)2(TG)2(AGA)2, (GA)3(GTT)2(CA)3(CGT)4	GCGGGAATCGATTGTTTAAG (F) ACGAAGAGGATGGTGACGAC (R)	SM15	58	250–280	3	0.75
13	HQ236486	M85	(GA)9,(CTA)3 (CTT)3(GTT)2	GCTAGCAGAATCACGATTTAATCTC (F) TTGGTGTTGGTGTTGCTGTT (R)	SM18	55	330–360	4	0.845
14	HQ236489	M105	(TC)8	TTTTGTAGCCGTTTTTGGATG (F) CGGTTATCCCCCTGATTTCT (R)	SM21	52	200–227	4	0.784
15	HQ236490	M18	(TCG)4(TCC)2,(TCG)5	GATCTCGTGGTTGTCGATGA (F) CGGTCACGTAGCCTATTCGT (R)	SM22	53	290–310	3	0.764
16	HQ236491	SI5	(AG)4(GAA)2,(AG)7	AGGAAGAACAACGGTGGAGA (F) CGCCCTTTTACGTTTCTCTG (R)	SM23	53	200–230	4	0.82
17	HQ236492	SI2	(TTC)2(CCTTT)2	TGGGAAATAGGATTGCCACT (F) GGGTTTCAATAAGGGGGAGA (R)	SM24	52	250–270	4	0.81
18	HQ236493	SI18	(TC)12	TAGCAGAATCGCTAGCAGCA (F) CCCATCTAACCTTCCCCCTA (R)	SM25	53	250–270	4	0.81

present study was submitted to the nucleotide sequence NCBI database with GenBank accession IDs ranging from HQ 224869.1 to HQ 224876.1 and from HQ 236477.1 to HQ 236492.1. From the 3328 EST data set from the NCBI database, 95 SSR primer pairs were designed (Table 2). An example of designing a primer using Primer 3.0 software in clone M93 is given in Figure 2.

A total of 120 primer pairs were designed for the repeat sequences obtained from selective hybridization and the EST database. Out of a diverse array of 16 sesame accessions (Figure 3), 104 (86.67%) SSRs that generated clearly scorable amplification products were evaluated for polymorphism. Of the 104 SSRs employed, 92 primer pairs [18 by selective hybridization (Table 1) and 74 from the EST-database (Table 2)] were polymorphic. The profiles generated using these polymorphic markers were able to differentiate among the sesame accessions studied.

A total of 286 amplification products were obtained

Table 2. Characteristics of EST-derived SSRs.

No.	Accession ID	Repeat motif	Primer sequence	Primer ID	Ta (°C)	Allele size	No. of alleles	PIC
1	BU670690	(CT)10	TCCCCCAAATTTCACAAAAA (F) AGAGTAGGTTGCGCTCCTCA (R)	SEM1	52	220–240	3	0.77
2	BU670685	(TC)9	CCCAGCCAAGAAACAAGAAA (F) AACCCCACTAGGCGAAGAAT (R)	SEM2	51	170–192	2	0.37
3	BU670338	(AG)12	GAGAAGCAGAAGCTCAAAGAGT (F) ACTTCTCCACTCCCATGACG (R)	SEM3	53	380–410	2	0.43
4	BU670669	(AGAAGA)3	CTTTGAAGGAGTGGGTGAGC (F) TTTCCAGCAATACCATACATCA (R)	SEM4	52	160–186	2	0.37
5	BU670662	(AGC)5	CCAGGGAAGTAAGAGGAGGTG (F) ACGGCTCCAAATGTGTTTGT (R)	SEM5	54	180–220	2	0.42
6	BU670541	(AGA)5	AGGACAAGATCCACGGTGAG (F) TCAGCATCACATGATTCAAGC (R)	SEM8	52	280–300	2	0.352
7	BU670534	(GCA)5	TTCCCGGAACATTCTGATTC (F) GCTTACCTCCCCCAAAAGTC (R)	SEM9	52	480–510	3	0.861
8	BU670516	(GA)8	GGACCATGTAATCCCAGCAC (F) GGGGCACAGAGTGGATGTAG (R)	SEM10	55	210–260	4	0.912
9	BU670476	(TA)7 and (TTC)2(TC)3	TTCCAGTACCGATCCTCACC (F) AAAATCTGCCAAATAAACCAAAA (R)	SEM 32	50	240–280	3	0.56
10	BU670450	(TA)6	GTCCGCCAGCTCAATACCTA (F) CGGAAACCGTACATTCATCA (R)	SEM 37	52	185–232	4	0.78
11	BU670434	(CT)5 and (ACACC)4	ACAGCACTTACCCCAAAGGA (F) TGGGAGGCAACTTTCATTCT (R)	SEM 38	51	400–450	4	0.644
12	BU670397	(TG)7	GTGCAGGAGGGGACTTTGTA (F) AGCACCAGCACCAGCACT (R)	SEM 42	53	210–230	2	0.32
13	BU670348	(AATGCT)3	TGCCTTTACAAATGGCTTCA (F) CCCATGAACCCATATCCTTG (R)	SEM 44	50	260–310	3	0.53
14	BU670327	(AGA)5 and (CAG)4	AGGACAAGATCCACGGTGAG (F) TCCCTTATTTGCAAGGCAAC (R)	SEM 48	52	285–320	2	0.39
15	BU670310	(AC)12	GCTGCATGCACAACCTATACA (F) GGTTTGAAGGGAGAGGAAGG (R)	SEM 51	53	196–224	2	0.43
16	BU670264	(CTCTCTCTC)3(CAC)3	AATTGACGCGAGGAGTCTTG (F) AAGCCTTTTGCACCTTCTGA (R)	SEM 57	51	350–400	3	0.79
17	BU670253	(GAGTGAG)3	CGAAAGAAGAGGCAGAGGTG (F) TCTCCGACCATCAAAACCAT (R)	SEM 62	52	260–280	2	0.45
18	BU670238	(CT)10	TCCATTCCTCTCATCCTCAA (F) CTGTGTCCGATCACCAAAAA (R)	SEM 64	50	340–400	2	0.48
19	BU670137	(GAGTGAG)3	CGAAAGAAGAGGCAGAGGTG (F) AGCAGTCTCCGACCATCAAG	SEM 73	54	180–230	3	0.49
20	BU670128	(AAGAAC)3	CTAGGAATGTCGGAGGCGTA (F) AATCCGAAACGTTGGCACT (R)	SEM 74	52	120–160	4	0.834
21	BU670118	(CA)4(TG)5,(AC)4(AG)2	GCTTCTGCGCTTTTACATCC (F) TTCTTACCCGCTGCCCTAAT (R)	SEM 76	52	400–450	4	0.43
22	BU670238	(CT)10	CAAACCTCACTGGTCTTCGAT (F) CCCGGATTGTCAAAGTCATT (R)	SEM 80	51	260–280	4	0.72
23	BU670068	(TC)7	TTTTCACGCTATCATCAAACC (F) CCTCCTCACCCTTGAACTGA (R)	SEM440	52	200–220	4	0.872
24	BU670030	(TA)7	CCATCAGGGAGTGAATTGCT (F) TCTCCGTCTGAACTGCCTCT (R)	SEM 82	53	100–130	3	0.765

Table 2. (Continued).

25	BU670027	(GCACCT)4	TTGCATCAGGAGATCCAACA (F) CACTCAAAGCAAACCAGCAA (R)	SEM 83	50	360–380	2	0.42
26	BU670003	(TCT)5	GACGACGCTAAGTCCGAATC (F) AGGGGTTAAGTGAGGCTGGT (R)	SEM 88	54	170–240	5	0.81
27	BU669994	(AGA)5 and (CAGCGA)3	AGGACAAGATCCACGGTGAG (F) TCCCTTATTTGCAAGCAACC (R)	SEM 90	52	140–170	3	0.643
28	BU669957	(AG)9(AAAAG)2	AACCATCCCATTTGTTTTGC (F) TCCTCAGAGCTGCACATTTTT (R)	SEM 94	49	220–240	4	0.783
29	BU669908	(TTGT)4	CCAACTTTTCTGGGTTGGAA (F) ATGGGCGTATCAGTTTCGAC (R)	SEM 100	51	181–199	4	0.814
30	BU669848	(CT)10 and (CT)7	CAAACCTCACTGGTCTTCGAT (F) CCCGGATTGTCAAAGTCATT (R)	SEM 104	52	250	3	0.542
31	BU669811	(TCT)6	GACGACGCTAAGTCCGAATC (F) AGGGGTTAAGTGAGGCTGGT (R)	SEM 108	54	162–208	3	0.674
32	BU669782	(CCA)5(CGG)3	CGGTCACCTGAATTTCCATC (F) GTACTCTTCCTCCGCCTCCT (R)	SEM112	54	260–320	2	0.45
33	BU669703	(TCC)6	CTCTCCCCTTCCCAATCAAT (F) GTGATGCAGCTGAAGTGGAA (R)	SEM120	52	150–190	3	0.61
34	BU669462	(CT)10	TCCCCCAAATTTCACAAAAA (F) AGAGTAGGTTGCGCTCCTCA (R)	SEM138	51	182–228	3	0.77
35	BU669409	(AT)7 (GTAT)6	AGGCTGGAGTCCATTGAGAA (F) TTACTTGGACCACCACAAAAA (R)	SEM 146	51	160–190	4	0.833
36	BU669221	(TCTCA)5	TGAAGCTGCCTTACGTGAAA (F) GCTTGATAGAGAAGTTACGACAAAAA (R)	SEM 170	52	110–130	4	0.72
37	BU669189	(TC)10	CCAAGAAACCGCTCACTAGC (F) CCAGCTCGTACTTCCCATGT (R)	SEM 176	54	190–240	2	0.39
38	BU669103	(CGG)7	GGTGGAGGTGGTGGAAGATA (F) ACCCAGCCGATAAACATCAC (R)	SEM 188	53	300–340	5	0.81
39	BU669001	(TC)9 (TG)6(TTTG)2	TTGACAATACCGCAATTAGCC (F) CATTGCGTCAGTTGCATTCT (R)	SEM 201	50	232–242	3	0.75
40	BU668961	(CT)10	TCCCCCAAATTTCACAAAAA (F) AGAGTAGGTTGCGCTCCTCA (R)	SEM 206	51	181–223	3	0.79
41	BU668814	(GAT)5	TCCTCTCTTTTCCTCCACCA (F) GGCTCTGCTTTGACCTACCA (R)	SEM219	53	200–210	2	0.5
42	BU668777	(TC)8	AATCCCTTTTCTCACTGCTCA (F) TGCACCACTAGGAACAGCAG (R)	SEM226	52	480–500	2	0.47
43	BU668643	(AG)10	CAGAATTCATTCTTCAACAACTCTTC (F) CGTGTTCCATCCCGTAACTT (R)	SEM 248	52	100–120	4	0.68
44	BU668626	(GCCACC)3	CATTAGGCCTTGTCCATGCT (F) CAATAACCCGTGAGGTGGAG (R)	SEM249	53	290–310	5	0.83
45	BU668561	(AC)5 (TG)4	ACTTGACAGCCATGGGAAAG (F) GAATAGCCTTCACGCTCCAG (R)	SEM 253	53	220–250	3	0.693
46	BU668543	(TC)4(TTCTCT)4 TCAA)4	AACGACATCACTTCGATCCAT (F) TGCTGACTTTCTTCCCGTTA (R)	SEM260	50	220–250	3	0.693
47	BU668467	(CCA)6 and (CAT)4	GCTCATGGACTACCCTCACG (F) AATTCGTCGACACTGTGGTG (R)	SEM265	53	200	3	0.41
48	BU668438	(CAGCCA)4	TGAGAGGAATTGGATTGGAAA (F) GTGGGGAATGAGGAAATGTG (R)	SEM270	51	187–224	4	0.77
49	BU669684	(TC)10	ACATTTCCTCATTCCCCACA (F) TGAAAGGAGGGAAAAACCAG (R)	SEM445	50	280–320	4	0.81

Table 2. (Continued).

50	BU668405	(CCTG)4	CTCCATTCCTCCACTTCCAA (F) CGCAATAGCTTGCATCTGAA (R)	SEM278	51	480–510	3	0.72
51	BU668493	(CATTCA)3	GAATTGAGAAAAAGAAAATGTTTGAA (F) AACGTTGAAGGTCCAACCAG (R)	SEM279	55	180–200	2	0.32
52	BU668385	(GAT)7	CTGGGGAAGGAAGTGGTGTA (F) TTGCAGAAGCCTTAACAGCA (R)	SEM282	52	185–240	5	0.79
53	BU668365	(AG)11	GGATTCCGACTGTTTCCAGA (F) ATTCACGCAACTCTCCCTCT (R)	SEM285	52	150	3	0.36
54	BU668318	(TC)6(CT)7(AT)6	GCAAACCTAAATGCCCTTGA (F) CAGTGCCTGTGTGCCTGTAT (R)	SEM292	52	140–170	5	0.84
55	BU668208	(AG)10	CCCTCGTTCGAAATCTCTGT (F) GTTTGGCCTTAGTTGCCTTG (R)	SEM312	52	195–250	2	0.47
56	BU668159	(AG)7(TTGA)2(GA)2	TCACACAATTACACACACACACC (F) GTTGATGGCTTGGAGGGTTA (R)	SEM314	53	170–210	3	0.57
57	BU668125	(AG)10	CACTTACAGGGCTCCTTGAATC (F) GGAGAGAACAAAGACAGACACG (R)	SEM315	55	111–173	5	0.72
58	BU668121	(AT)10	CACGGAAGCAGCTCATCAT (F) CCTGCCGACATGACTACAAC (R)	SEM316	54	150	2	0.32
59	BU668088	(CT)6&(AC)6	CACTCCCATCCCACCATACT (F) AACCCCATTTCTTCGCTGTA (R)	SEM324	53	200	2	0.34
60	BU668080	(GA)6(GTGA)2(GA)4	CCACAGGAATTCCGACACTT (F) CCTTTCCCTCGAAGATCACA (R)	SEM326	52	300–350	2	0.37
61	BU667806	(TC)6 & (TC)7	CAAACTTGAACCACGACAGC (F) CTCCATGTTCCTCAGCTTCC (R)	SEM361	53	250–270	3	0.72
62	BU667772	(AT)5 & (ATGTAT)3	TTTTCTTCCCCTCCTCAACA (F) GCCCTGAGGGATTTGAGTTT (R)	SEM364	51	280–300	2	0.42
63	BU667711	(CT)6 & (AC)7	CACTCCCATCCCACCATACT (F) AACCCCATTTCTTCGCTGTA (R)	SEM371	52	175–210	2	0.48
64	BU667689	(TG)7	ATTCTTTGCGCCTCTTTGTG (F) TTCCTCACATCGAACAACCA (R)	SEM378	50	198–218	2	0.482
65	BU667627	(AG)10	CCCTCGTTCGAAATCTCTGT (F) TCTGAGTTGCCACATGCTTC (R)	SEM384	52	200–215	2	0.48
66	BU667555	(CT)8	TTCTGTGGCACTCGTAGTCG (F) TAGGCATTGCCAATTTGTGA (R)	SEM396	52	195–220	2	0.47
67	BU667547	(CT)9	TTTCTCCTCTCACTCTGCAATC (F) TGCACCACTAGGAACAGCAG (R)	SEM399	54	390–430	3	0.2982
68	BU667505	(CCA)5	TAGCTCTCGCCGTTCTGTTT (F) CTCCTCCTCGAACCTTCCTT (R)	SEM406	53	230–250	3	0.79
69	BU667448	(CT)8	TTCTGTGGCACTCGTAGTCG (F) CTGTCGCCTTTGCTTTTACC (R)	SEM417	52	300	3	0 .64
70	BU667447	(GAGTGAG)3	CAAGAAAAGGCCACAGAGGA (F) CAACACAAACTCGACAGCACA (R)	SEM428	52	160–200	5	0.842
71	BU667391	(AT)10	CACGGAAGCAGCTCATCAT (F) TCTGGCTGCTCAACAAGAAA (R)	SEM430	51	260–310	3	0.73
72	BU667382	(CT)8	TATCGGCGATTTCTCCAAAC (F) CAAATGCACCGTGAATCAAC (R)	SEM434	50	184–196	3	0.83
73	BU667375	(TCCC)3	CAACCAAATCAACACCAACG (F) CGTCGCTTGCACATACAAAT (R)	SEM435	50	222–236	3	0.753
74	BU667372	(TGGA)3	GACCCAAGATCACCACCATC (F) TATGGAATGGGACCAAATCA (R)	SEM436	52	174–176	2	0.49

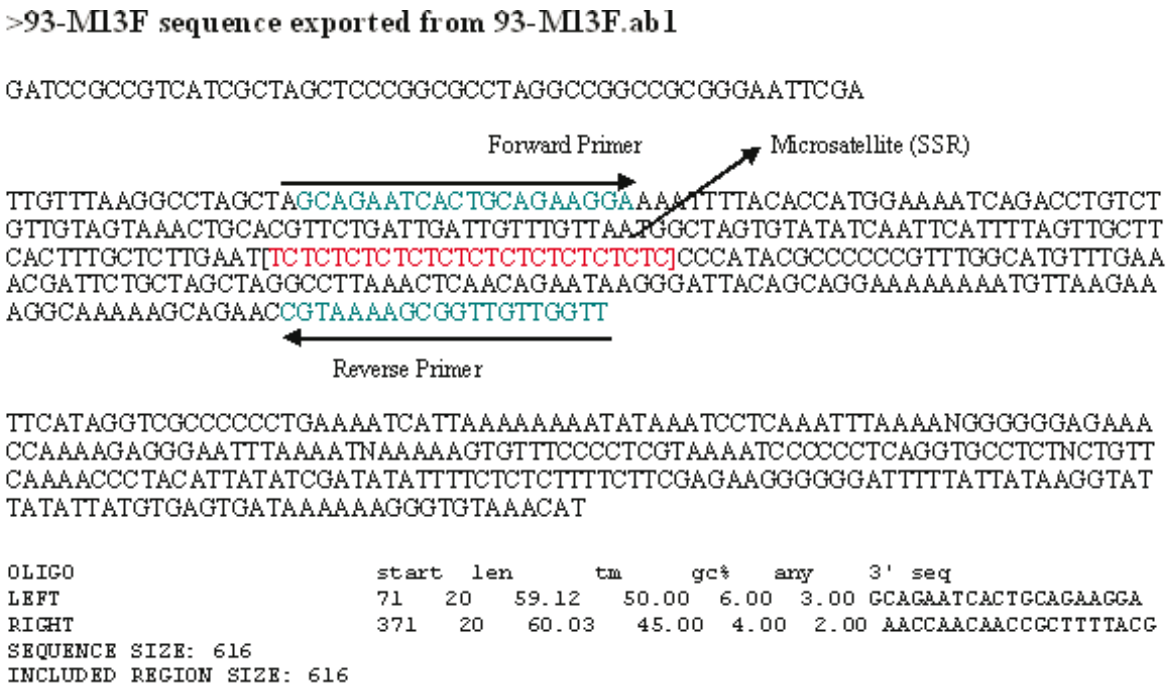

Figure 2. Primer design in clone number M93 with Primer 3.0 software.

and the number of alleles per microsatellite locus ranged from 2 to 5 with an average of 3.11 alleles. Allele size ranged widely (100–510 bp) among the primer pairs. Estimated PIC ranged from 0.2982 to 0.912. Jaccard's similarity coefficient ranged from 0.21 to 0.82 among the 16 accessions evaluated for polymorphic value of the isolated SSRs. The UPGMA-based dendrogram obtained using these data (Figure 4) revealed that they existed in 5 clusters each, with a varied number of accessions. Cluster I included 2 accessions, TAC-89-309 and VRI-1; cluster II included 5 accessions, NSKMS-128, CO-1, Paiyur-1, Uma, and Madhavi; and cluster III included 5 accessions, Chandana, Hima, Rajeshwari, Swetha, and TKG-22. TMV-3 and IC-204492 (of the wild species *S. malabaricum*) made up cluster IV, and cluster V consisted of 2 accessions, BB-3-8 and BBL-46-2K (of the wild species *S. mulayanum*).

4. Discussion

In the present study, sesame-specific microsatellite markers were isolated following a microsatellite-enriched genomic library approach as well as mining from the EST database. Dixit et al. (2005) reported 10 polymorphic SSR markers in sesame. After that initial report, there was a considerable gap (until 2011) in developing genomic tools in sesame, except for scattered reports on the use of anonymous markers in assessment of genetic diversity. Use of advanced genomic tools like Illumina paired-end sequencing (Wei et al., 2011) and RNA sequencing (Zhang et al., 2012), along with the traditional selective hybridization approach (Spandana et al., 2012) and mining from GenBank depositions (Wei et al., 2008; Wang et al., 2012; Vijay et al., 2013), added a number of SSRs to the genomic library of sesame. Renewed interest in developing genomic wealth in this ancient crop is encouraging; however, the number of microsatellites that have been reported in sesame is still minimal. To cover the entire genome and to construct a reasonable genetic linkage map of sesame, research should be directed towards increasing the sesame-specific marker database.

The enriched library prepared in the present study indeed resulted in a higher percentage (45%) of microsatellites, but redundancy problems further reduced the actual percentage to 30%. Problems of redundancy in enriched libraries have been reported in other crops (Rallo et al., 2000; Mba et al., 2001) as well as sesame (Dixit et al., 2005). The AG/TC class of dinucleotide repeats was the most abundant in the enriched library as well as in the EST-derived SSRs, as has been the case in other crops (Ferguson et al., 2004; Lichtenzveig et al., 2005). In a study analyzing wheat (*Triticum aestivum* L.), rice (*Oryza sativa* L.), maize (*Zea mays*), and soybean (*Glycine max*) ESTs, TC repeats have been found to be the most frequent, suggesting that dinucleotide repeats could be the most abundant in coding regions of most plant genomes (Gao et al., 2003).

In the present study, the selection of a restriction enzyme was as important as the selection of repeat probe. An enzyme that has the recognition site within the repeat region would yield only fragments without any

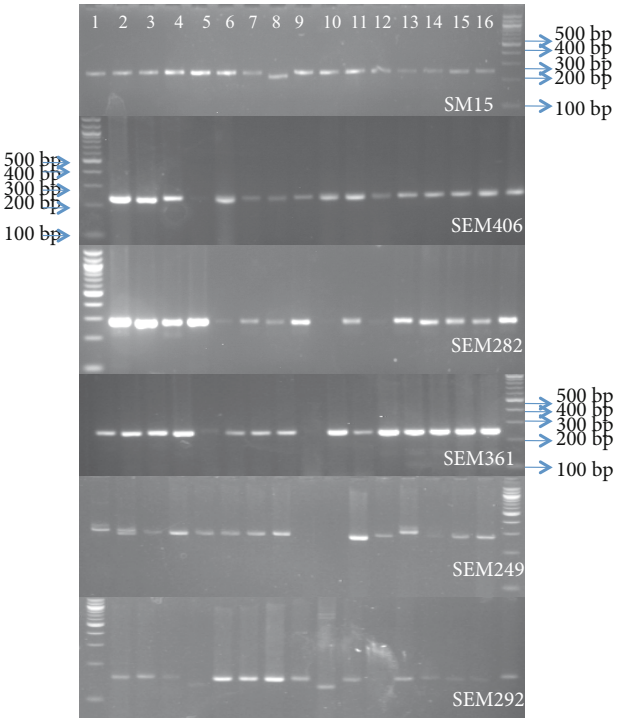

Figure 3. Evaluation of SSR primer pairs for polymorphism in 16 sesame accessions. 1. TKG-22, 2. TMV-3, 3. NSKMS-28, 4. Chandana, 5. Swetha, 6. BB-3-8, 7. Madhavi, 8. BBL-46-2K, 9. VRI-1, 10. C0-1, 11. PAIYUR-1, 12. Hima, 13. IC-204492, 14. Rajeshwari, 15. Uma, 16. TAC-89-309.

repeats for which it is being probed for. For example, the recognition site of the restriction enzyme R*sa*I is GT'AC/CA'TG; a digestion product of this enzyme therefore cannot be used for enriching with any such probe like GTAC/CATG/TAC/GTA and so on. Digestion fragments from different enzymatic digestions after ligation with linkers were pooled for fragments with linkers. Use of more than one restriction enzyme has been suggested for library construction (Kölliker et al., 2001) to result in a maximum yield of potentially useful SSRs and ensure even distribution of SSRs across the genome. Though 4 different enzymes (R*sa*I, B*fu*I, B*st*UI, and N*he*I) were chosen in the present study for digestion, only the first 2 enzymes were used as they resulted in complete digestion; the latter 2 only resulted in partial digestion of the genome. About 13.3% of all the primers designed failed to amplify despite optimization efforts. Poor amplification could be due to divergence in the sequences flanking SSRs, thus creating null alleles (Smulders et al., 1997). Null alleles are presumably caused by DNA polymorphisms in primer sites (Tang et al., 2002), especially in noncoding regions (Mogg et al., 2002).

Evaluating the polymorphic SSR markers isolated by grouping a set of sesame genotypes discriminated cultivars from the wild species by placing them in separate clusters. The sesame germplasm consisted of 13 genotypes of cultivated species and 3 accessions from 2 wild species. Wild accessions were included in the diversity analysis to test for conservation of the SSRs isolated. All the cultivars except TMV-3 grouped into the first 3 clusters. TMV-3 clustered with the wild species *S. malabaricum*, indicating its involvement in the evolution of the cultivar. Close proximity of *S. malabaricum* with the wild cultivars supports the assertion that it could be the immediate progenitor of the cultivar species *S. indicum* (Bhat et al., 1999; Hiremath and Patil, 1999; Nathankumar et al., 2000; Bedigian, 2003). The 2 accessions of the species *S. mulayanum* clustered together into a separate group, suggesting that it had no role in the evolution of cultivated sesame and it was distinctly different from *S. malabaricum*.

Though a large number of molecular markers have been employed in sesame genetic diversity studies, the majority of them were not based on the sesame genome sequence information. Robust crop-specific SSR markers employed for the first time in sesame revealed their highly informative nature, but the number of SSR markers used was quite low (Dixit et al., 2005). Zhang et al. (2007), utilizing the sequence information from the EST data set, revealed high genetic variability in sesame germplasm with EST-SSR and SRAP markers. In the present study, the sesame-specific SSR markers consisting of 25 genomic SSRs and 90 EST-derived SSRs proved to be highly informative by providing further evidence to support the hypothesis that *Sesamum malabaricum* could be the immediate progenitor of the cultivated species *Sesamum indicum*. A similar inference was drawn by Bhat et al. (1999) using RAPD markers.

A detailed review of the various investigations on sesame genetic diversity revealed that different methods give different results regarding the level of genetic diversity, depending upon marker type, composition, and size of germplasm. The differences in genetic diversity reported by employing SSRs are not only due to the aforementioned factors, but could also be due to the type of SSRs used. Evaluation of the level of the distribution of different classes of SSRs in the genomes of wheat, *Arabidopsis*, maize, and rice revealed that the frequency of microsatellites was significantly higher in ESTs than in genomic DNA across all species (Morgante et al., 2002). Contrary to this, the frequency of genomic SSRs was higher than EST-SSRs, though the percent of genomic SSRs used was quite low (21%) when compared to EST-SSRs (79%) in the present study. When exploring genetic relationships, though it is advantageous to use quite a large number of molecular markers, it is equally important that these markers evenly cover the entire genome. In the majority of the crop species (rice, maize, sorghum, tomato, *Arabidopsis*, etc.), saturated genetic maps are available and

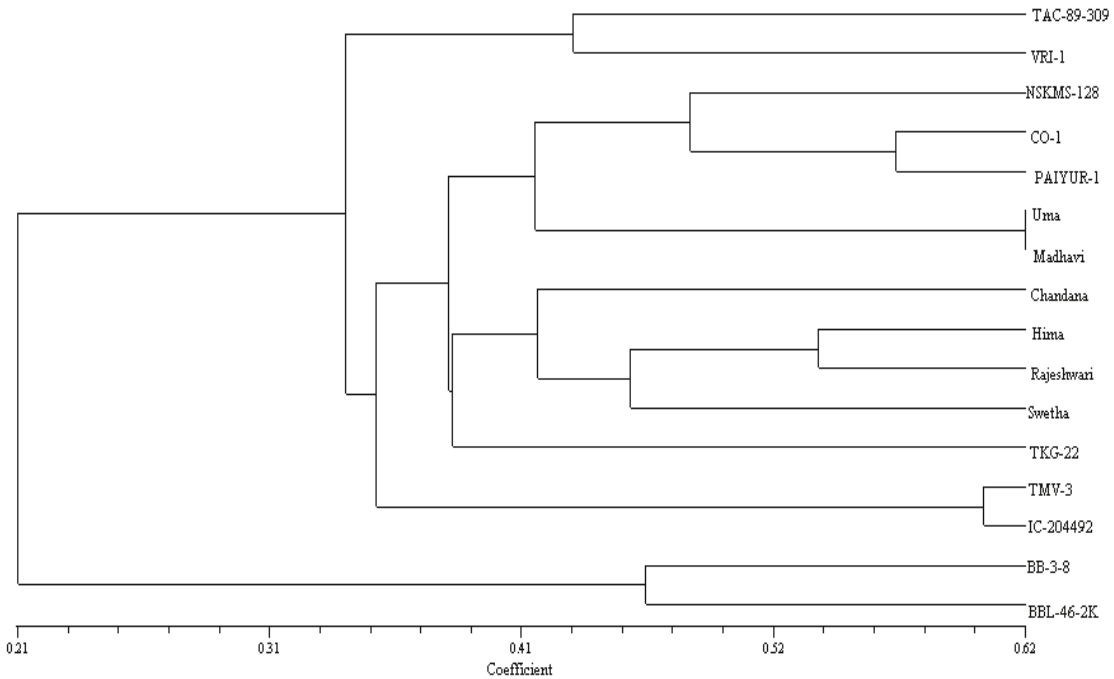

Figure 4. Dendrogram showing genetic diversity among 16 sesame accessions using the newly isolated SSRs.

selection of molecular markers that adequately represent the entire genome can be done without any difficulty. However, in this study, there was no scope for the selection of SSR markers, as there is neither a saturated genetic map based on them nor are there abundant SSRs in sesame. The first genetic linkage map using 284 PCR-based markers of 3 different types (10 EST-SSR markers, 30 AFLP markers, and 124 RSAMPL markers) covering 76% of the genome was recently reported by Wei et al. (2009). Owing to the pace of technological advancement, development of sesame-specific molecular markers and the creation of a saturated genetic map based on the same are not far from reality.

Considering the highly informative nature of SSR markers compared to all other marker types used in sesame, considerable emphasis should be placed on generating species-specific SSR/SNP markers that would help understand the population structure and genetic diversity and detect and exploit genes relating to both qualitatively and quantitatively inherited traits.

Acknowledgments

Financial support from the National Agricultural Innovation Project, Indian Council of Agricultural Research, and provision of seed material from the National Bureau of Plant Genetic Resources in New Delhi are greatly appreciated.

References

Abdellatef E, Sirelkhatem R, Ahmed MMM, Radwan KH, Khalafalla MM (2008). Study of genetic diversity in Sudanese sesame (*Sesamum indicum* L.) germplasm using random amplified polymorphic DNA (RAPD) markers. Afr J Biotechnol **7**: 4423–4427.

Ali GM, Yasumoto S, Katsuta MS (2007). Assessment of genetic diversity in sesame (*Sesamum indicum* L.) detected by amplified fragment length polymorphism. Electron J Biotechn 10: 12–23.

Bedigian D (2003). Evolution of sesame revisited: domestication, diversity and prospects. Genet Resour Crop Ev 50: 779–787.

Bedigian D, Seigler DS, Harlan JR (1985). Sesamin, sesamolin and the origin of sesame. Biochem Syst Ecol 13: 133–139.

Bhat KV, Babrekar PP, Lakhanpaul S (1999). Study of genetic diversity in Indian and exotic sesame (*Sesamum indicum* L.) germplasm using random amplified polymorphic DNA (RAPD) markers. Euphytica 110: 21–33.

Botstein D, White LR, Sholnick M, Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am J Hum Genet 32: 314–331.

Brar GS, Ahuja KL (1979). Sesame, its culture, genetics, breeding and biochemistry. Annu Rev Plant Sci 1: 245–313.

Cheung SC, Szeto YT, Benzie IF (2007). Antioxidant protection of edible oils. Plant Foods Hum Nutr 62: 39–42.

Dixit A, Jin MH, Chung JW, Yu JW, KichungH, Ma KH, Park YP, Chi EG (2005). Development of polymorphic microsatellite markers in sesame (*Sesamum indicum* L.). Mol Ecol Notes 5: 736–738.

Ercan A, Taskin M, Turgut K (2004). Analysis of genetic diversity in Turkish sesame (*Sesamum indicum* L.) populations using RAPD markers. Genet Resour Crop Ev 51: 599–607.

Ferguson ME, Burrow MD, Schulze SR, Bramel PJ, Paterson AH, Kresovich S, Mitchell S (2004). Microsatellite identification in peanut (*Arachis hypogeal* L.). Theor Appl Genet 108: 1064–1070.

Gao L, Tang J, Hongwei L, Jia J (2003). Analysis of microsatellites in major crops assessed by computational and experimental approaches. Mol Breeding 12: 245–261.

Hiremath SC, Patil CG (1999). Genome homology and the putative progenitor of sesame. J Cytol Genet 34: 69–74.

Kim DH, Zur G, Danin PY, Lee SW, Shim KB, Kang CW, Kashi Y (2002). Genetic relationship of sesame germplasm collection as revealed by inter-simple sequence repeats. Plant Breeding 121: 259–262.

Kölliker R, Jones ES, Drayton MC, Dupal MP, Forsters JW (2001). Development and characterization of simple sequence repeat (SSR) markers for white clover (*Trifolium repens* L.). Theor Appl Genet 102: 416–424.

Laurentin HE, Karlovsky P (2006). Genetic relationship and diversity in a sesame (*Sesamum indicum* L.) germplasm collection using amplified fragment length polymorphism. BMC Genet 7: 1–10.

Lichtenzveig, J, Scheuring C, Dodge J, Abbo S (2005). Construction of BAC and BIBAC libraries and their application for generation of SSR markers for analysis of chick pea *Cicer arietinum* L. Theor Appl Genet 110: 492–510.

Mba REC, Stephenson P, Edwards K, Melzer S, Nkumbira J, Gullberg U, Apel K, Gale M, Tohme J, Fregene M (2001). Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta carantz*) genome towards and SSR based molecular genetic map of cassava. Theor Appl Genet 102: 21–31.

Mogg R, Batley J, Hanley S, Edwards D, O'Sullivan H, Edwards K (2002). Characterization of the flanking regions of *Zea mays* microsatellite reveals a large number of useful sequence polymorphisms. Theor Appl Genet 105: 532–543.

Mohamed HMA, Awatif II (1998). The use of sesame oil un-saponifiable matter as a natural antioxidant. Food Chem 62: 269–276.

Morgante HMA, Hanafey M, Powell W (2002). Microsatellites are preferentially associated with no repetitive DNA in plant genomes. Nat Genet 30: 194–200.

Nathan Kumar G, Singh KN, Vaidyanatahan P (2000). Relationships between cultivated sesame (*Sesamum* sp.) and the wild relatives based on morphological characters, isozymes and RAPD markers. J Genet Breed 54: 5–12.

Nei M, Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Nat Acad Sci USA 76: 5269–5273.

Rallo P, Dorado G, Martin A (2000). Development of simple sequence repeats (SSRs) in olive tree (*Olea europaea* L.). Theor Appl Genet 101: 984–989.

Rohlf FJ (1994). NTSYS-PC*:* Numerical Taxonomy and Multivariate Analysis System. New York, NY, USA: Exeter Software.

Sambrook JF, Russell DW (2001). Molecular Cloning: A Laboratory Manual. Vol. 1, 2, and 3. 3rd ed. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press.

Sankar D, Sambandam G, Ramakrishna RM, Pugalendi KV (2005). Modulation of blood pressure, lipid profiles and redox status in hypertensive patients taking different edible oils. Clin Chim Acta 355: 97–104.

Sharma SN, Vinod K, Mathur S (2009). Comparative analysis of RAPD and ISSR markers for characterization of sesame (*Sesamum indicum* L.) genotypes. J Plant Biochem Biot 18: 37–43.

Smulders MJM, Bredemeijer G, Rus-Kortekaas W, Arens P, Vosman B (1987). Use of short microsatellites from database sequence to generate polymorphism among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. Theor Appl Genet 97: 264–272.

Sneath PHA, Sokal RR (1973). Numerical Taxonomy: The Principles and Practice of Numerical Classification. San Francisco, CA, USA: Freeman.

Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004). A new integrated genetic linkage map of the soybean. Theor Appl Genet 109: 122–128.

Spandana B, Prathap Reddy V, John Prasanna G, Anuradha G, Sivaramakrishnan S (2012). Development and characterization of microsatellite markers (SSR) in sesamum (*Sesamum indicum* L.) species. Appl Biochem Biotechnol 168: 1594–1607.

Tang S, Yu JK, Slabaugh MB, Shintani DK, Knaps SJ (2002). Simple sequence repeat map of the sunflower genome. Theor Appl Genet 105: 1124–1136.

Vijay Y, Malathi S, Rao KVS, Vemireddy LR, Jyothi B, Dineshkumar V, Anuradha G, Siddiq EA (2013). Assessment of genetic diversity in sesame (*Sesamum indicum* L.) genotypes using EST derived SSR markers. J Crop Sci Biotechnol 16: 93–103.

Wang H, Varma RP, Yuan M, Gong L, Zhao Y, Guo B, Farmer AD, Rosen BD, Gao J, Isobe S et al. (2012). Development and characterization of BAC-end sequence derived SSRs, and their incorporation into a new higher density genetic map for cultivated peanut (*Arachis hypogaea* L.). BMC Plant Biology 12: 10.

Wang L, Zhang Y, Qi X, Gao Y, Zhang X (2012). Development and characterization of 59 polymorphic cDNA-SSR markers for the edible crop *Sesamum indicum* L. (Pedaliaceae). American J Bot 99: e394–e398.

Wei W, Qi X, Wang L, Zhang Y, Hua W, Li D, Lv H, Zhang X (2011). Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using illumine paired end sequencing and development of EST-SSR markers. BMC Genomics 12: 451.

Wei YB, Zhang HY, Zheng YZ, Miao HM, Zhang TZ, Guo WZ (2009). A genetic linkage map construction for sesame (*Sesamum indicum* L.). Genes and Genomics 31: 207–216.

Xin D, Sun J, Wang J, Jiang H, Hu G, Liu C, Chen Q (2012). Identification and characterization of SSRs from soybean (*Glycine max*) ESTs. Mol Biol Rep 39: 9047–9057.

Yamada K, Morita T, Masuda K Sugai M (1993). Chloroplast DNA variation on the genus *Sesamum*. J Plant Res 106: 81–87.

Zhang H, Wei L, Miao H, Zhang T and Wang C (2012). Development and validation of genic-SSR markers in sesame by RNA-seq. BMC Genomics 13: 316.

Zhang P, Yang ZH, Zhen GW, Zhan ZY, Bin WL, Zhen ZT (2007). Genetic diversity analysis of *Sesamum indicum* L. germplasm using SRAP and EST-SSR markers. Acta Agronomica Sinica 33: 1696–1702.