Mining of EST-SSR Markers of *Musa* and Their Transferability Studies Among the Members of Order the Zingiberales

**S. Backiyarani · S. Uma · P. Varatharj ·
M. S. Saraswathi**

Abstract Expressed sequence tags (ESTs) databases of 11 *Musa* complementary DNA libraries were retrieved from National Center of Biotechnology Information and used for mining simple sequence repeats (SSRs). Out of 21,056 unique ESTs, SSR regions were found only in 5,158 ESTs. Among these SSR containing ESTs, the occurrence of trinucleotide repeats are the most abundant followed by mono-, di-, tetra-, hexa-, and pentanucleotides. Moreover, this study showed that the rate of class II SSRs (<20 nucleotides) was higher than the class I SSRs (<20 nucleotides), and proportion of class I and II SSRs as abundant for tri-repeats. As a representative sample, primers were synthesized for 24 ESTs, carrying >12 nucleotides of SSR region, and tested among the various genomic group of *Musa* accessions. The result showed that 88 % of primers were functional primers, and 43 % are showing polymorphism among the *Musa* accessions. Transferability studies of *Musa* EST-SSRs among the genera of the order Zingiberales exhibited 100 and 58 % transferability in Musaceae and Zingiberaceae, respectively. The sequence comparison of SSR regions among the different *Musa* accessions confirmed that polymorphism is mainly due to the variation in repeat length. High percentage of cross-species, cross-genera, and cross-family transferability also suggested that these *Musa* EST-SSR markers will be a valuable resource for the comparative mapping by developing COS markers, in evolutionary studies and in improvement of the members of Zingiberaceae and Musaceae.

Keywords *Musa* sp. · EST-SSR markers · Transferability · Zingiberaceae

Introduction

Genetic improvement of plantain and banana has been restricted by the complex genetic structure and ploidy nature of the crop, which often leads to low levels of fertility and seed viability. It is further hindered by the large amount of space required for growth,

S. Backiyarani (✉) · S. Uma · P. Varatharj · M. S. Saraswathi
National Research Centre for Banana (ICAR), Thogamalai Road, Thayanur Post, Trichy 620 102 Tamil Nadu, India
e-mail: backiyarani@gmail.com

maintenance, and evaluation of plant populations. Molecular marker assisted breeding could be potentially exploited to enhance efficiency of genetic improvement in *Musa*. One of the major setbacks in the application of molecular breeding for *Musa* improvement is the paucity of informative molecular markers. Although simple sequence repeats (SSRs) are considered to be the best marker in many crop species because they are codominant, multiallelic, and randomly dispersed throughout the genome [1], their development through the conventional approach of genomic library construction is time consuming and resource intensive [2]. Harvesting of publicly available DNA sequence databases for the development of SSRs provides an attractive, time and cost-effective alternative. Expressed sequence tags (ESTs) are particularly attractive for marker development since they represent coding regions of the genome and are also being developed at an extremely faster pace for many genomes [3–5]. Moreover, the frequency of microsatellites is significantly higher in ESTs than in genomic DNA [6, 7]. In addition, they can lead to the development of gene-based maps, which helps in the identification of candidate genes and increases the efficiency of marker assisted selection [8]. Similar advantages of EST-SSRs have also been reported in other horticultural crops such as grapes [9], *Prunus* [10], and cocoa [11]. To date, nearly 35,000 banana ESTs are currently available in the GenBank of National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/), and their number is increasing every year. Therefore, the present study has been initiated (1) to mine the SSR markers from *Musa* ESTs, (2) to analyze the frequency and distribution of SSRs, (3) to validate the EST-SSR markers for detection of polymorphism in *Musa* genotypes, and (4) to assess the transferability of markers among *Musa* species, related genera Ensete, and family Zingiberaceae of the order Zingiberales.

Materials and Methods

Data Mining for SSR Marker

ESTs of 11 *Musa* complementary DNA (cDNA) libraries developed from various organs and different stress conditions were retrieved from NCBI. ESTs having poly-A tails were trimmed using TrimEST program (http://emboss.bioinformatics.nl/cgi-bin/emboss/trimest). Then, all the ESTs were reassembled using CAP3 program (http://pbil.univ-lyon1.fr/cap3.php). These assembled sequences were run in WEBSAT for finding the SSR motifs (http://purl.oclc.org/NET/websat/). EST-SSR primers were designed for at least six repeat units in case of dinucleotide, four repeat units for tri- and tetranucleotide and three repeat units for penta- and hexanucleotide SSRs using the same program.

Plant Materials, Amplification, and Confirmation of SSRs

Eighteen *Musa* accessions representing different genomic groups of the core collection of *Musa* germplasm, maintained at NRCB, Trichy, were used to validate the EST-SSR markers, which are designed from the *Musa* EST database. The genomic DNA of these accessions was extracted as described by Gawel and Jarnet [12] with minor modification.

Twenty-four primer pairs for *Musa* EST-SSRs were synthesized, which are representing all possible combinations of mono- [1], di- [11], tri- [8], tetra- [1], penta- [2], and hexa-repeats [1]. PCR reaction was carried out in a 12.5-μl volume containing template DNA (20 ng), dNTPs (0.2 mM), primers (each 0.5 μM), 1.25 U of *Taq* polymerase, and 1× PCR buffer. Amplification was carried out using an Ependorff thermocycler with the following

program: initial denaturation at 94 °C for 5 min; 25 cycles of 94 °C for 1 min; specific primer annealing temperature for 1 min; and extension at 72 °C for 1 min with an extra elongation period of 7 min at 72 °C. The PCR products were run on 3 % Agarose in 1× Tris–borate–EDTA buffer and visualized by ethidium bromide staining. A 50-bp molecular ladder was used to estimate the allele sizes.

The putative functions of identified functional markers were annotated by BLASTX against the NCBI Non-redundant Protein http://www.ncbi.nlm.nih.gov/ RefSeq/. Those EST-SSRs, which were successfully amplified in various genomic groups of *Musa*, were used to test the transferability across genera and family using transferability of *Musa* EST-SSR markers all primer pairs producing successful PCR bands using *Musa* genomic DNA were tested using *Ensete superbum*, *Ensete glaucum* of family Musaceae and *Zingiber officinale*, *Elettaria cardamomum*, and *Curcuma longa* of family Zingiberaceae).

To confirm the presence of microsatellite repeats, MESSR 17 was used for the amplification of genomic genomic DNA of six cultivars, representing different genomic group such as AA, BB, AAB, ABB, and AAAB, and they were electrophoresed on a 3 % agarose gel. The bands were eluted from the gel and purified using the Qiagen gel extraction kit. The purified fragment was cloned using the pGEMT TA cloning kit for sequencing. The positive clones identified by PCR were sequenced by an automated sequencer (ABI PRISM 3700 DNA analyzer). The final edited sequences belonging to different genotypes were compared with the original SSR containing EST sequence through multiple sequence alignment using Lasergene 7.1v software.

Results and Discussion

Abundance, Frequency, and Distribution of SSRs in *Musa* ESTs

A total of 32,708 EST sequences were retrieved from 11 cDNA libraries available in the public domain. They have been derived from different organs and/or stress conditions of two different species (*Musa acuminata* and *Musa balbisiana*) and bispecific hybrids, and they were used to evaluate the presence of SSR motifs. These ESTs were trimmed with the help of TrimEST. Elimination of redundant sequences and improvement of sequence quality using CAP3 analysis resulted in 21,050 potential unique ESTs (Table 1). These unique sequences were subjected to WEBSAT analysis resulting in 5,158 (24.5 %) SSRs containing ESTs. This included 1,349 contigs and 3,809 singletons. Among the 11 libraries screened, all libraries except IV and VII had 20 % of SSR containing ESTs. This study revealed that banana exhibited higher percentage of SSR containing *Musa* ESTs (24.5 %) compared to other horticulture crops like cassava (9.7 %) [13], pepper (10.2 %) [14], and coffee (6.8 %) [15]. Varshney et al. [16] reported that the frequency of EST-SSRs may be attributed to the parameters used for mining SSRs. High frequency of SSRs in banana ESTs might be attributed to the higher mutation rate in clonally propagated crops [17]. This is supported by Sreenu et al. [18], suggesting that the distribution and frequency are resultant of factors like mutation and DNA repair mechanisms.

Out of 5,158 SSR containing ESTs, 4,691 contained only one SSR, 234 had more than one SSR, and 505 had compound SSR, and the detailed data for individual library are depicted in Table 2. The *Musa* EST-SSRs contained diverse types of repeat motifs. Among different repeat units, the trinucleotide repeat motifs were maximum (37.6 %) followed by mono- (32.9 %), di- (15.9 %), tetra- (8.6 %), hexa- (3.07 %), and pentanucleotide (1.7 %). The occurrences of different repeat units were tri-repeats (37.6 %). Similar trend has been

Table 1 Details of *Musa* EST libraries retrieved from NCBI

Library	Genotype	Genome	Tissue	Stress	Number of ESTs	Assembled ESTs	Number of EST-SSRs
I	Grand Naine	AAA	Leaf	Unstress	4,030	3,161	930
II	Grand Naine	AAA	Fruit	Unstress	2,528	2,182	536
III	Grand Naine	AAA	Fruit	Crown disease	6,604	5,066	1,506
IV	Nabusa	AAA	Root	*Fusarium*	50	47	2
V	Calcutta 4	AA	Leaf	Hot stress	1,144	524	146
VI	Calcutta 4	AA	Leaf	Cold stress	1,143	574	140
VII	Pisang Klutuk Wulung	BB	Leaf	Unstress	2,645	2,523	229
VIII	Pisang Klutuk Wulung	BB	Root	Drought Stress	2,644	2,480	759
IX	Cachaco	ABB	Shoot meristem	Osmotic Stressed	8,568	2,608	504
X	Cachaco	ABB	Shoot meristem	Un stress	2,502	1,532	332
XI	Manoranjitham	ABB	Leaf	*Mycosphaerela*	850	353	74

observed in cereals [19] and in legumes [20]. Most EST sequences consist of exonic regions, which are under strong selection against frameshift mutations as they translate into proteins. As codons are functional units of three nucleotides, indel mutations causing a shift in three nucleotides do not perturb the current reading frame of a given gene [21].Therefore, trinucleotide repeats are expected to be the most abundant SSR class found in ESTs.

The size of the SSR varied from 10 bp (5 repeats for dinucleotide motifs) to 60 bp (20 repeats for trinucleotide motif), and the average number of repeats was 8.3 for dimeric, 4.8 for trimeric (58.3 % with 4 repeats), 3.2 for tetrameric (85.8 % with 3 repeats), 3.42 for pentameric (72.4 % with 3 repeats), and 3.2 for hexameric (85 % with 3 repeats) SSRs. Maximum length of SSR region was observed for tri-repeats (CGA),[20] di-repeats (TA),[25] tetra-repeats (GATA),[9] penta-repeats (ATGCC),[7] and hexa-repeats (GAACCG)[6] and (GCCGTC).[6] Among 11 libraries, library III had maximum repeat length for di-, tri-, and hexa-repeats followed by library IX for tetra- and penta-repeats. In general trinucleotides with repeat motifs CCG/CGG were dominant followed by AAG/CTT when compared with other repeats (Table 3). The SSRs in *Musa* ESTs showed a more or less equal contribution of A+T (52 %) rich repeat motifs and G+C rich repeat motifs (48 %). This is accordance with the findings of Lescot et al. [22] and Cheung and Town [23] who also registered 47 % GC content in the protein matching regions of cDNAs and *Musa* bacterial artificial chromosome (BAC) end sequences, respectively.

Class I and II Group of *Musa* EST-SSRs

Based on the nucleotide length of the SSR region, simple perfect SSRs were classified into two groups, namely, class I (>20 nucleotides) and class II (10–20 nucleotides) SSRs. The number of class I SSR ranged from 0 to 234, and the number of class II SSRs ranged from 2 to 1,272 was observed in libraries IV and III, respectively. In general, the most abundant repeat types in class I SSRs were mono-repeats followed by tri-repeats, while trinucleotide repeats followed by mono-repeats were more abundant in class II SSRs. The pentanucleotide

Table 2 Summary of SSRs detected after sequences assembled and categorized

Library	ESTs containing SSR	ESTs having more than 1 SSR	Compound SSRs	Mononucleotide repeats	Dinucleotide repeats	Trinucleotide repeats	Tetranucleotide repeats	Pentanucleotide repeats	Hexanucleotide repeats
I	828	50	133	184	128	519	78	16	35
II	518	9	52	148	98	221	67	11	20
III	1,317	94	173	670	234	628	103	24	47
IV	2	0	0	0	1	1	0	0	0
V	138	3	20	14	35	69	20	4	4
VI	133	3	7	18	39	59	12	6	6
VII	209	10	9	185	24	20	0	0	0
VIII	661	58	2	382	71	27	0	0	0
IX	498	2	41	18	80	159	46	11	16
X	332	3	64	29	128	252	31	14	24
XI	70	2	4	62	0	6	2	4	0
Total	4,706	234	505	1,710	838	1,961	359	90	150

Table 3 Frequency and distribution of different SSR types in the 21,050 nonredundant *Musa* ESTs redundant

	3	4	5	6	7	8	9	10	11	12	13	14	15	>15	Total
A/T								169	121	84	66	53	146	1,029	1,668
C/G								18	14	1	5	2	1	1	42
AC/CA			9	5	6	6	6	7	4	0	2				45
AG/GA			104	42	31	36	10	28	22	16	20	5	10	26	350
AT/TA			46	23	11	6	6	7	6	0	2	4	0	7	118
CG/GC			7	1	1	3									12
CT/TC			62	35	32	32	14	11	20	6	8	4	3	37	264
GT/TG			19	11	5	3	0	6	2	0	2	0	0	1	49
AGA/ACA		57	15	14	0	4	10	4	-	2				1	107
AAG/CCT		94	24	14	14	-	2	1							149
ATC/GAT		31	13	3	10	1	3								61
CCG/CGG		120	42	5	5										172
CGT/GGT		22	9	0	1										32
ACC/AGC		36	13	7	8	4									68
GAA/TCC		102	31	17	12	5	1								168
GAC/GGC		40	8	17	10	1	1	1							78
GAG/GCT		65	26	11	5						1		1		109
Tri (Other)	550	160	138	56	23	3	5	5	1	3	0	1		945	
Tetra	310	33	9	5	2										359
PENTA	61	22	3	3	1										90
HEXA	131	11	5	3											150
N								187	135	85	71	55	147	1,030	1,710
NN			247	117	86	83	39	59	54	22	34	13	13	71	838
NNN		1,117	341	226	121	38	20	11	5	4	3	1	1	3	1,891
NNNN	310	33	9	5	2										359
NNNNN	61	22	3	3	1										90
NNNNNN	131	11	5	3											150
	502	1,183	605	354	210	121	59	257	194	111	108	69	161	1,104	5,038

repeats were very meager in both classes of SSRs. A number of ESTs having class I SSRs types were observed for different repeat types such as tri- (307), di- (187), hexa- [21], tetra- [10], and pentanucleotides [7]. Similarly, a number of class II SSRs for tri- (1,700), di- (648), hexa- (131), penta- (83), and tetranucleotides (70) were observed. It is understood that the number of class II SSRs is maximum than that of class I SSRs, and both class I and II SSRs were abundant for tri-repeats. Among the SSRs, a maximum of 38 % of class I SSRs when compared to the total SSRs were observed in library II, whereas library III had maximum number of class II SSRs.

Transferability of EST-Derived SSR Markers

A total of 24 EST-SSR primer pairs were synthesized and evaluated with DNA extracted from 18 *Musa* germplasm accessions representing diverse genomic groups. Out of these, 21 EST-SSR primers amplified the genomic DNA of all the accessions. The high amplification success suggested that the flanking regions of these loci are sufficiently conserved, and they

could be used for comparative analyses of genetic diversity in the genus *Musa*. Failure in amplification of the three primers is mainly due to the deletion of microsatellite at a specified locus [24], or nonamplification is due to mutations (in-dels) in the primer binding sites. [25]

Nine out of 24 (43 %) showed polymorphism among the *Musa* accessions. The high polymorphism observed in this study is contradictory to the results of Wang et al. [26] who reported lower rate of polymorphism for EST-SSRs than the genomic SSRs. This suggested that these EST-SSR markers could be exploited for improving banana through MAS breeding and developing linkage maps. The higher frequency of polymorphism observed is possibly due to the inclusion of diverse indigenous land races, exotic introduction, and wild species. The number of alleles detected ranged from 1 to 3 with an average of 1.8 alleles per locus. The amplification results suggested that di- and tri-repeat SSR motifs are highly polymorphic than tetra-, penta- and hexa-repeats.

In the present study, the observed size of the amplicon was larger than the expected size for the primer MESSRm17. Similar result was also reported by Thiel et al. [27]. The difference in the amplicon size is due to the presence of introns and insertions–deletions (in–dels) in the corresponding genomic sequences, which was substantiated by Saha et al. [28] through sequence analysis.

Sequence Comparison of SSR Region Across *Musa* Genomes

EST-SSR polymorphism was confirmed at the nucleotide level by cloning and sequencing the amplified products of the primer MESSR17, for six out of 18 accessions exhibiting polymorphism were cloned and sequenced. The BLAST analysis (Table 4) showed that all the sequences hit with the original EST sequence from which MESSR17 primer was designed. The sequence alignment showed that there was variation in the repeat length of GGA and CTT motif among the *Musa* accessions (Fig. 1). This suggested that allelic diversity for the MESSR17 primer must be due to the differences in repeat lengths, which could be due to insertion or deletion of repeats in the microsatellite regions.

The comparison of sequence sizes and nucleotides of the different genomic groups suggested that BB genome (Attikole and Bhimkol) are having more nucleotide base pairs compared to other genomic groups, namely, AA AAB, ABB, and AAAB. This is mainly due to variation in the length of CTT and GGA motifs and also the point mutation in GGA motif in both BB accessions. Apart from this, variation in motif length, a few point mutations were observed in the SSR motif flanking region. This phenomenon is supposed to be the innate evolving nature of the genome, which is indicative of the evolutionary relationships among the taxa [29].

To confirm the specificity of this MESSR 17 marker across the BB genomes, 10 accessions of BB genome were compared with accessions belonging to different genomic group through mobility differences of the amplified product in 3 % Metaphore agarose gel. Eight out of ten accessions (BB genome) exhibited monophorphic bands of same size of the sequenced BB accessions genomes, and the remaining two accessions exhibited polymorphic bands similar to other accessions. This result suggested that MESSR 17 marker could not be used as BB-specific markers as SSRs are highly mutable loci [30].

Transferability of *Musa* EST-SSR Markers Across Genera and Families

To test the transferability of *Musa* EST-SSR markers among the Zingiberale order, eight *Musa* species (*M. acuminata*, *M. balbisisana*, *M. ornata*, *M. laterita*, *M. nagensium*, *M. itinerans*, and *M. sikkinensis*) related genus (*E. superbum* and *E. glaucum*) and three

Table 4 Details of EST-SSR primers used for studying the transferability across Musaceae and Zingiberaceae

| EST-SSR marker	Sequence	Primer name	SSR motif	Blast	E value	Product size
MESSR1	DN239954	Str4	(T)16	*Musa acuminata* AAA Group mRNA for putative high mobility group protein	$5E-162$	105
MESSR2	DN238160	Str1	(AG)15	*Nelumbo nucifera* catalase mRNA, complete cds	$2E-148$	262
MESSR3	CONTIG46	Str10	(AG)15	*Phyllostachys edulis* cDNA clone full insert sequence	$9E-64$	194
MESSR4	DN238418	Str2	(AG)9	*Zea mays* oxygen-evolving enhancer protein 3-1 mRNA, complete cds	$8E-91$	342
MESSR5	CONTIG91	Str9	(CT)13	*Zea mays* clone 1062364 nonspecific lipid-transfer protein 2 precursor, mRNA	$6E-44$	179
MESSR6	DN238348	Str3	(CT)16	*Zea mays* clone 285841 homocysteine S-methyltransferase 3 mRNA, complete cds	$3E-95$	340
MESSR7	DN239829	Str7	(GA)10	*Hordeum vulgare* subsp. Vulgare mRNA for predicted protein	$6E-162$	201
MESSR8	CONTIG88	Str11	(GA)8	*Zea mays* chlorophyll a-b binding protein 6A	0	313
MESSR9	FF559253	SSR2	(AG)19	*Musa acuminate* clone 1-59 response regulator 6 mRNA partial cds	$1.00E-11$	116
MESSR10	ES437536	SSR8	(CG)9	*Mus musculus* mitochondrial ribosomal protein L50,mRNA complete cds	0.004	150
MESSR11	DN239594	Str6	(GAT)7	PREDICTED: *Vitis vinifera* similar to nascent polypeptide associated complex alpha	$6E-149$	213
MESSR12	DN239930	Str8	(GGA)5	*Musa* ABB Group dehydrin (Dhn12) mRNA, complete cds	0	371
MESSR13	DN239403	Str5	(TGC)4	*Musa acuminata* AAA Group mannose-binding lectin mRNA	$1E-38$	179
MESSR14	FF559627	SSR6	(TGG)5	*Musa acuminata* sub sp.*burmannicoides* var.*calcutta* 4 mRNA containing microsatellite,	$0.00E+00$	111
MESSR15	ES437610	SSR7	(AGA)4	*Ricinus communis* WRKY transcription factor, putative, mRNA	$1E-12$	121
MESSR16	S38	SSR9	(CTT)4	*Musa acuminata* clone BAC MA4-72 F16, complete sequence	0	240
MESSR17	FF878775	SSR4	(GAA)11	*Musa acuminate* sub sp. *burmannicoides* var.*calcutta* 4 mRNA containing microsatellite,	0	274
MESSR18	Contig132	SSR11	(TCT)12	*Sorghum bicolor* hypothetical protein, mRNA	$1.00E-77$	124
MESSR19	DN238292	SSR10	(ATAA)4	No significant similarity	–	135
MESSR20	Contig89	SSR12	(CACTG)4	No significance similarity	–	210
MESSR21	FF559235	SSR1	(CTTGAA)2	*Arabidopsis thaliana* NAD+ADP-ribosyltransferase/zinc ion binding (PARP2) mRNA,complete cds	$2.00E-101$	297
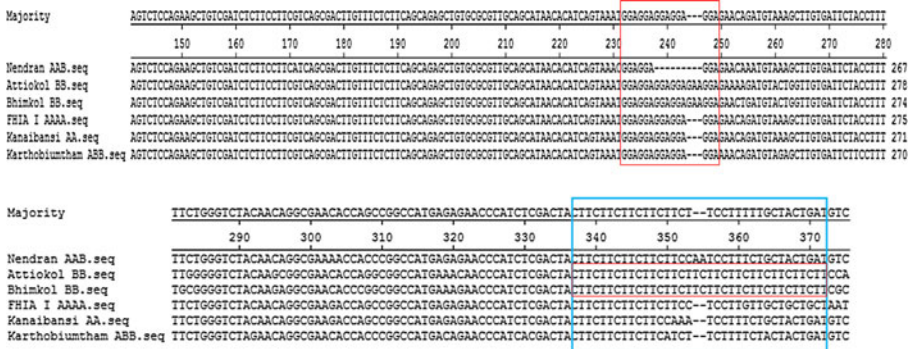
Fig. 1 Alignment of six different genomic group of *Musa* EST sequences obtained by amplification using MESSR17 primer pairs. Repetitive sequences are indicated in *boxes*

different genera of the family Zingiberaceae (*E. cardamomum*, *C. longa*, and *Z. officinale*) were amplified with the functional *Musa* EST-SSR primers (Fig. 2). The amplification results revealed that all the 21 functional EST-SSR primers produced amplicons in all *Musa* and *Ensete* species of *Musa*ceae. Higher levels of transferability of EST-derived SSRs reflect the conserved nature of coding sequences in the SSR flanking region. This result suggested that these transferable EST-SSR markers could be used for identification of markers associated with specific traits in other *Musa* species and related genera. Similar results were observed among the related species of grapes [31] and pines [32] and among related genera of sugarcane [33] and sunflower [34]

Amplification of *Musa* EST-SSR markers in other genera of Zingiberaceae revealed that these markers could be used for a variety of studies. Out of 21 polymorphic EST-SSR markers, transferability was observed for *Z. officinale* [12], *C. longa* [10], and *E. cardamomum* [12] of Zingiberaceae family. Gao et al. [35] have also experimentally proven the transferability of EST-SSR markers from monocot (wheat) to dicot (soybean) species. These results indicated that EST-SSRs can be transferred across relatively large taxonomic distances, spanning not just species within a genus and multiple genera within a family but in some instances families with in the same order. Currently, no ESTs are available for the *E. cardamomum* in the public domain unlike *Zingiber* spp. (38115) and *Curcuma* spp. (12593), which have sufficient number of ESTs (NCBI). It is suggested that these transferable *Musa* EST-SSR markers could also be used in the cardamom improvement program as the availability of cardamom specific primers are very meager. Fatokun et al. [36] reported that different plant species often share orthologous genes for very similar functions. Similarly,
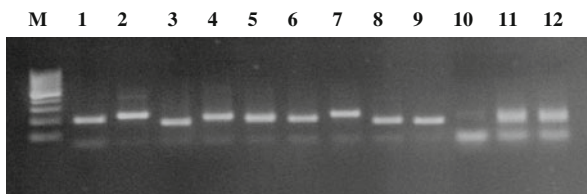

Fig. 2 Amplification patterns obtained with primer MESSR17 in 3 % agarose gel electrophoresis of various genera of order Zingiberales. *M* 100 bp Molecular weight marker, 1 *Musa acuminate*, 2 *M. balbisiana*, 3 *M. ornata*, 4 *M. laterita*, 5 *M. nagensium*, 6 *M. itinerans*, 7 *M. sikkimenis*, 8 *Ensete superbum*, 9 *Ensete glacum*, 10 *Zingiber officinale*, 11 *Elettaria cardamomum*, 12 *Curcuma longa*

Bennetzen and Freeling [37] and Gale and Devos [38] also stated that gene content and gene order among different plant species are highly conserved. This suggested that *Musa* EST-SSR markers have potential to develop conserved orthologous set (COS) markers, which facilitate comparative studies between species and members of the order Zingiberales. Hence, it is suggested that these *Musa* EST-SSR markers are potential markers for comparative mapping among the Zingiberaceae and Musaceae families, which could be used in the development of COS markers for the order Zingiberales.

In summary, the *Musa* EST databases have enormous amount of functional SSR markers, which could be exploited not only for studying the genetic diversity but also used for identifying candidate gene for specific trait, as these EST-SSRs are part of or adjacent to functional genes. High percentage of cross-species, cross-genera and cross-family transferability also suggested that these *Musa* EST-SSR markers will be a valuable resource for the comparative mapping by developing COS markers in evolutionary studies and in the improvement of the members of Zingiberaceae and Musaceae.

References

1. Yu, J., & Kohel, R.J. (1999). Proceedings of the Belt Wise Cotton Research Conferences, National Cotton Council, Memphis, p. 439.
2. Zane, L., Bargelloni, L., & Patarnello, T. (2002). *Molecular Ecology, 11*, 1–16.
3. Gu, Z., Hillier, L., & Kwok, P. Y. (1998). *Human Mutation, 12*, 221–225.
4. Kantety, R. V., La Rota, M., Matthews, D. E., & Sorrells, M. E. (2002). *Plant Molecular Biology, 48*, 501–510.
5. Picoult-Newberg, L., et al. (1999). *Genome Research, 9*, 167–174.
6. Morgante, M., Hanafey, M., & Powell, W. (2002). *Nature Genetics, 30*, 194–200.
7. Toth, G., Gaspari, Z., & Jurka, J. (2000). *Genome Research, 10*, 967–981.
8. Gupta, P. K., Rustgi, S., Sharma, S., Singh, R., Kumar, N., & Balyan, H. S. (2003). *Molecular Genetics Genomics, 270*, 315–323.
9. Scott, K. D., Eggler, P., Seaton, G., Rossetto, M., Ablett, E. M., Lee, L. S., et al. (2000). *Theoretical and Applied Genetics, 100*, 723–726.
10. Ogundiwin, E. A., Martí, C., Forment, J., Pons, C., Granell, A., Gradziel, T. M., et al. (2009). *Plant Molecular Biology, 68*, 379–397.
11. Fouet, O., Allègre, M., Argout, X., Jeanneau, M., Lemainque, A., Pavek, S., et al. (2011). *Tree Genetics and Genomics, 74*, 799–817.
12. Gawel, N. J., & Jarret, R. L. (1991). *Plant Molecular Biology Reporter, 9*(3), 262–266.
13. Raji, A., Anderson, J., Kolade, O., Ugwu, C., Dixon, A., & Ingelbrecht, I. (2009). *BMC Plant Biology, 9*, 118.
14. Yi, G., Lee, J., Lee, S., Choi, D., & Kim, B. D. (2006). *Theoretical and Applied Genetics, 114*(1), 113–130.
15. Poncet, V., Rondeau, M., Tranchant, C., Cayrel, A., Hamon, S., de Kochko, A., et al. (2006). *Molecular Genetics and Genomics, 276*(5), 436–449.
16. Varshney, R. K., Graner, A., & Sorrells, M. E. (2005). *Trends in Biotechnology, 23*(1), 48–55.
17. Gaudeul, M., Stenøien, H. K., & Agren, J. (2007). *American Journal of Botany, 94*, 1146–1155.
18. Sreenu, V. B., Kumar, P., Nagaraju, J., & Nagarajaram, H. A. (2007). *Journal of Biosciences, 32*, 3–15.
19. Varshney, R. K., Thiel, T., Stein, N., Langridge, P., & Graner, A. (2002). *Cellular and Molecular Biology Letters, 7*(2A), 537–546.
20. Eujayl, I., Sledge, M. K., Wang, L., May, G. D., Chekhovskiy, K., Zwonitzer, J. C., et al. (2004). *Theoretical and Applied Genetics, 108*, 414–422.
21. Metzgar, D., Bytof, J., & Wills, C. (2000). *Genome Research, 10*, 72–80.
22. Lescot, M., Piffanelli, P., Ciampi, A. Y., Ruiz, M., et al. (2008). *BMC Genomics, 9*, 58. doi:10.1186/1471-2164-9-58.

23. Cheung, F., & Town, C. D. (2007). *BMC Plant Biology, 7*, 29. doi:10.1186/1471-2229-7-29.
24. Lothe, R. A., Andersen, S. N., Hofstad, B., Meling, G. I., Peltomäki, P., Heim, S., et al. (1995). *Genes Chromosome Cancer, 14*(3), 182–188.
25. Callen, D. F., Thompson, A. D., Shen, Y., Phillips, H. A., Richards, R. I., Mulley, J. C., et al. (1993). *American Journal of Genetics, 52*, 922–927.
26. Wang, J. Y., Zheng, L. S., Huang, B. Z., Liu, W. L., & WU, Y. T. (2010). *Genetic Resources and Crops Evolution, 57*(4), 553–563.
27. Thiel, T., et al. (2003). *Theoretical and Applied Genetics, 106*, 411–422.
28. Saha, M. C., Mian, M. A. R., Eujayl, I., Zwonitzer, J. C., Wang, L., & May, G. D. (2004). *Theoretical and Applied Genetics, 109*, 783–791.
29. Agarwal, R. K., Hendre, P. S., Varshney, R. K., Bhat, P. R., Krishnakumar, V., & Singh, L. (2007). *Theoretical and Applied Genetics, 114*(2), 359–372.
30. Gow, C., Noble, J. L., Rollinson, D., & Jones, C. (2005). *Genetica, 124*(1), 77–83.
31. Decroocq, V., Fave, M. G., Hagen, L., Bordenave, L., & Decroocq, S. (2003). *Theoretical and Applied Genetics, 106*, 912–922.
32. Chagne, D., Chaumeil, P., Ramboer, A., Collada, C., Guevara, A., Cervera, M. T., et al. (2004). *Theoretical and Applied Genetics, 109*, 1204–1214.
33. Cordeiro, G. M., Casu, R., McIntyre, C. L., Manners, J. M., & Henry, R. J. (2001). *Plant Science, 160*, 1115–1123.
34. Ellis, J. R., & Burke, J. M. (2007). *Heredity, 99*, 125–132.
35. Gao, L., Tang, J., Li, H., & Jia, J. (2003). *Molecular Breeding, 12*, 245–261.
36. Fatokun, C. A., Menacio-Hautea, D. I., Danesh, D., & Young, N. D. (1992). *Genetics, 132*, 841–884.
37. Bennetzen, J. L., & Freeling, M. (1993). *Trends in Genetics, 9*, 259–261.
38. Gale, M. D., & Devos, K. M. (1998). *Proceedings of the National Academy, 95*, 1971–1974.