I J T A
© Serials Publications
National Academy of Agricultural Science (NAAS) Rating : 3.03

Development and validation of SSR markers in Finger millet (*Eleusine coracana* (L.) Gaertn)

Jagadeesh Selvam. N., Muthukumar. M., Hifzur Rahman, Senthil. N and M. Raveendran*

Abstract: *Finger millet (*Eleusine coracana *(L.) Gaertn), a nutritionally rich millet is widely adapted to harsh environments like drought, salinity etc. Since, the crop is neglected over years, the genomic resources are limited. The present study aims at supplementing the genomic resources in the form of eSSR markers. We have identified 12 novel functional eSSR markers from RNA Seq assembled contigs of finger millet through ion torrent sequencing platform. The twelve flanking eSSR primers designed were used for validation in 23 finger millet genotypes/varieties collected from different eco-geographical adaptations by PCR amplification and allele sizing through capillary electrophoresis. Twelve eSSR primers generated a total of 92 alleles with an average of 7.6 alleles per locus and allele sizes ranging between 33 and 268 bp and an average major allele frequency of 0.4398. The discriminatory power and efficiency of the markers was determined by the parameters like average band informativeness (Av.Ib= 0.3816), average resolving power (Av.Rp=2.6956), average gene diversity (He=0.49), observed heterozygousity (Ho=0.19) and average polymorphism information content (PIC=0.67). The cross amplification of the eSSR markers in the related family of millets were affirmed in three other millets such as proso millet, barnyard millet and fox tail millet. The phylogenetic analysis revealed that the 23 genotypes clustered mainly into two distinct groups and similar pattern was observed with principal component analysis. The allelic pattern observed in all the primers together were pictorially depicted as eSSR barcodes or DNA fingerprints for the 23 genotypes of finger millet. The eSSR markers developed from this study could be used for molecular characterization, genetic diversity and hybridity analysis of finger millet.*

Key words: *Finger millet; eSSRs; genetic diversity, barcodes, finger printing*

INTRODUCTION

Finger millet (*Eleusine coracana* L. Gaertn, family Poaceae (Gramineae), subfamily Chloridoideae and tribe Eragrostideae), native to the Uganda and Ethiopian regions of Africa (Harlan 1971; deWet 1995), is one of the most important small millets cultivated in the tropics, especially South Asia and African continents. It occupies 12% of the global millet area with the major producers being Uganda, India, Nepal, and China having a yield potential of >10 t/ha under optimum irrigated conditions (Goron, and Raizada 2015). In India, it is cultivated in area of 1.8 m ha with the average yields of 1.3t/ha (http://www.icrisat.org/crop-fingermillet.htm). Finger millet tops the list among the potential crops to be exploited for area expansion programmes for millets globally, because of its' wider adaptability to harsh environments, besides its nutritional and therapeutic properties (NAAS 2013). Being grown predominantly under rainfed conditions, it is a considered as gene resource for resistance against biotic and abiotic stresses (Bhatt, *et al*., 2011; Babu, *et al*., 2014). In terms of nutritive value, it is rich in amino acids (methionine & tryptophan), minerals (iron & calcium), dietary fibre and polyphenols (Goron, and Raizada 2015). Likewise, its health care benefits are correlated to the grain's polyphenol content with anti-oxidant property and dietary fiber promoting slow digestion as well as stability of blood sugar levels, which are advocated by dieticians for cancer and diabetics patients (Chandrasekara, and Shahidi 2011; Devi, *et al*., 2014). Despite these potentials of adaptability, nutritive and therapeutic values, it has been a neglected crop over several years and was regarded as poor person's crop, famine food or birdseed. Only in the recent past, the need for the development of genomic resources, understanding the mechanism of stress resistance and

* Department of Plant Biotechnology, Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore 641 003, India, *E-mail: raveendrantnau@gmail.com*

exploitation of the genes for cereal crop improvement has been realized.

The present day cultivated species *E. coracana* is evolved by domestication of the wild species, *Eleusine coracana* subsp. *africana*. The cultivated species is divided into 5 races on the basis of inflorescence morphology, among which the race *coracana* is widely distributed and is present in the archaeological record of early African agriculture dating 4000 years back (Purugganan, and Fuller 2009). It is reported that the racial evolution took place further in Africa: races *vulgaris*, *elongate*, *plana*, and *compacta* evolved from the race *coracana*, and were introduced into India some 3000 years ago (Prasad, and Staggenborg 2009). Very little is known on the genetics and genomics of the crop. It is an allotetraploid with chromosome number 9 (2n=4x=36) and composed of AABB genome (Bisht, and Mukai 2001). The progenitor of A-genome is *E. indica* and B-genome is yet unknown although some claim it to be *E. tristachya* (Dida, and Devos 2006; Liu, *et al.,* 2014). The genome size based on C-value (2C= 3.34-3.87 pg), approximated to 1593 Mbp (Bharathi 2011; Liu, *et al.,* 2014; Goron, and Raizada 2015). Till date, the genomic resources of finger millet available in the public domain are very limited, for instance; only 1780 EST sequences and 440 genomic sequences are retrievable from NCBI database. At present, the SSR (simple sequence repeats) markers available for molecular characterization and genetic analysis in finger millet are very less. Global collaborative efforts have been initiated recently in 2014, for the whole genome sequencing and transcriptome profiling in finger millet by Bio-innovate and ICRISAT, Hyderabad (Goron, and Raizada 2015); which would enrich the genomic resources by discovery of new markers.

The SSRs marker system is considered to be the ideal molecular system for genetic diversity analysis, establishing evolutionary relationships, marker trait associations because of the advantages, *viz*., high level of polymorphism, locus-specificity, multiallelic and co-dominant nature, wide genome coverage, high abundance, and reproducibility (Powell, *et al*., 1996). Genetic diversity analysis is the first step for understanding the genetic similarity among the genotypes and is a tool for identification of parents for crossing. So far, limited work on genetic diversity analysis of finger millet has been reported globally (Babu, *et al*., 2007; Dida, and Devos 2006; Dida, *et al*., 2007; Babu, *et al.,* 2014). With respect to genetic linkage map, only one genetic linkage map has been reported in finger millet. Discovery of genomic SSRs is a time consuming and laborious process as it involves construction of microsatellite enriched genomic libraries, SSR capture, sequencing and identification of SSR markers. Alternatively, genic SSRs could be mined *in silico* from the expressed sequence tags (EST sequences) available in public databases and are commonly known as EST-SSRs. It was reported that 39 to 49% of the genome of finger millet consists of repetitive DNA sequences (Gupta, and Ranjekar 1981). In an earlier study by our group, RNA sequencing of two varieties of finger millet under salinity stress conditions (control vs stress) was performed for understanding the transcriptome (Rahman, *et al.,* 2014). This sequence information was used to mine the EST-SSRs (eSSRs) and validate them in few local and exotic varieties. We also present here, the genetic diversity and phylogenetic relationships established by the EST-SSRs and a first report of fingerprints for 23 finger millet varieties as SSR barcodes based on the allele sizes generated by 12 eSSR markers.

MATERIALS AND METHODS

De novo assembly of RNA sequence information

The RNA sequence information in an earlier study (Rahman, *et al.,* 2014) on transcriptome analysis of the finger millet under salinity stress and control conditions reported for understanding the differentially expressed genes was used for the present investigation. *De novo* assembly for the high quality RNA-Seq reads with the default parameters (work flow shown in Fig.1) through CLC Genomics workbench (http://www.clcbio.com/products/clc-genomics-workbench) and the contigs were generated.

Identification of SSR regions and designing of flanking eSSR primers

The presence of repeat regions in the contigs and their positions were determined through phobos plugin (http://www.ruhr-uni-bochum.de/ecoevo/cm/cm_phobos.htm) available in the Geneious software. Primers were synthesized for the ESTs containing at least 8 di-, 5 tri-, 4 tetra-, 3 penta- or 2 hexa-nucleotide repeats by Primer3 (*in- built* in Geneious software) based on the parameters of minimum 50% GC content, melting temperature range between 50°C and 63°C, 18–27 bp long primers, amplicon sizes of 100 to 400 bp. The quality of the designed primers were checked through *in silico* analysis in the web version of Oligoanalyzer (Integrated DNA technologies,

Iowa) for the absence of secondary structures (hairpins, homodimers, heterodimers).

Source of finger millet varieties and relative millets for primer validation

Twenty three finger millet genotypes (Table 1) collected from various locations across India were used for validation of these designed primers by PCR amplification. Cross amplification of these primers in the closely relatives of the family Poaceae, such as Proso millet (*Panicum miliaceum,* var. CO 5), Barnyard Millet (*Echinochloa frumentacea* var. CO 5) and Foxtail millet (*Setaria italica* var. CO 7) was also carried out.

DNA isolation, quantification and quality checking

Genomic DNA of the 23 finger millet genotypes (Table 1) and 3 other minor millets (Table 2) were isolated by mini-prep CTAB method as described by Doyle (1987) with minor modifications, for instance, tissue lysis performed in a tissue lyzer (Qiagen, USA). The leaf tissues from 10-day-old seedlings were sliced into small bits of 1-2 cm long and transferred to 2 ml eppendorf tubes along with two beads per tube. About 600 µl of pre-heated extraction buffer (2% CTAB, 100 mM Tris–HCL pH 8.0, 50 mM EDTA pH 8.0, 1.5 M NaCl) was added to the leaf tissues and mixed well by vortexing and inverting. The tubes were placed in a 65°C water bath in a tube holder for 30 min with occasional mixing. Equal volume of Chloroform: Isoamyl alcohol mix (24:1) was added to the tubes with the samples and mixed by repeated inversion for 3 min. The tubes were centrifuged at 10,000 rpm at 4°C for 10 min. After centrifugation, the upper aqueous phase was collected and two-thirds volume of pre-chilled ethanol was added to precipitate DNA. The tubes were incubated at -70°C for 10 min. The DNA pellets were washed with 70 % ethanol and dissolved in sterile water. After dissolving the pellet, the concentrated DNA was stored at -20°C. DNA concentration and purity based on A260/A280 ratios was quantified using a Nano-Drop spectrophotometer (ND-1000 Spectrophotometer, NanoDrop Technologies, USA) and the DNA concentration was normalized to be 25 ng/µl for PCR reaction.

Validation of eSSR primers

Validation of the designed EST-SSR primers was done by PCR amplification and separation of amplicons by capillary electrophoresis**.** A standard PCR reaction of total reaction volume 15ìL containing 25 ng of DNA template, 1X Taq buffer, 2 mM $MgCl_2$, 0.2 mM dNTPs,

Table 1
List of finger millet varieties used for this study, geographical origin and location from which the seeds were obtained.

S.No	*Genotypes*	*ORIGIN*	*Pedigree*	*Collected From*
1	CO7	INDIA	Selection from Local Fingermillet	TNAU, Coimbatore, Tamil Nadu
2	CO12	INDIA	Selection from variety PR722	TNAU, Coimbatore, Tamil Nadu
3	CO13	INDIA	CO7 x TAH 107	TNAU, Coimbatore, Tamil Nadu
4	CO14	INDIA	Malawi 1305 x CO 13	TNAU, Coimbatore, Tamil Nadu
5	DPI 009-04	INDIA	CO12 X TNAU 946	TNAU, Paiyur, Tamil Nadu
6	GPU26	INDIA	(Ind5 x Ind9)x IE1012	GKVK, Bangalore, Karnataka
7	GPU28	INDIA	Indof5 x IE1012	GKVK, Bangalore, Karnataka
8	GPU66	INDIA	PR202xGP28	GKVK, Bangalore, Karnataka
9	HR374	INDIA	EC-4840 x IE-927	GKVK, Bangalore, Karnataka
10	HR911	INDIA	UAS1 x 1E927	GKVK, Bangalore, Karnataka
11	IE3317	ZIMBABWE	-	ICRISAT, Hyderabad, Andhra Pradesh
12	IE3952	UGANDA	-	ICRISAT, Hyderabad, Andhra Pradesh
13	IE4121	UGANDA	-	ICRISAT, Hyderabad, Andhra Pradesh
14	IE4329	ZIMBABWE	-	ICRISAT, Hyderabad, Andhra Pradesh
15	IE4671	INDIA	-	ICRISAT, Hyderabad, Andhra Pradesh
16	Indof5	INDIA	Cauvery x IE 927	GKVK, Bangalore, Karnataka
17	KM252	INDIA	-	TNAU, Coimbatore, Tamil Nadu
18	Paiyur2	INDIA	VL 145 x Selection 10	TNAU, Paiyur, Tamil Nadu
19	PES110	INDIA	Selection from germplasm	GKVK, Bangalore, Karnataka
20	PR202	INDIA	Pureline selection from peddapuram	GKVK, Bangalore, Karnataka
21	RAU8	INDIA	BR-407x Ranchi local	TNAU, Coimbatore, Tamil Nadu
22	Trichy1	INDIA	Selection from HR 374	TNAU, Coimbatore, Tamil Nadu
23	KMR301	INDIA	MR-1xGE-1409	TNAU, Coimbatore, Tamil Nadu

Table 2
Description of eSSR markers developed from RNA contigs of finger millet

Primer No	*Primer Name*	*Primer Sequence*	*Length*	*Repeat*	*Annotation*	*E-Value*
1	C_15674_F	CTTAGCCTCCAGCCTCAA	18	(CAG)6	MYB family transcription factor	8e-26
	C_15674_R	CTTCACTTCATCGTTCGTTAG	21			
2	C_1574_F	AGGTAGTAGCGATAAAGATGG	21	(CCG)7	RNA recognition motif containing protein	5e-17
	C_1574_R	GTGTGTGGTAAGAAAGAGATG	21			
3	C_2042_F	AGGACGACGACACCATTG	18	(GCG)7	Respiratory burst oxidase	2e-25
	C_2042_R	GCTGTTGCTGCTGATGAA	18			
4	C_22006_F	GCGAATTACTCACGAATCC	19	(ATCA)4A	Expressed protein	0.70
	C_22006_R	CCATGCTCCTCGTTACAG	18			
5	C_32131_F	GGAAGAAGAAGAAGTCATCAAG	22	(GCA)6G	Expressed protein	2e-28
	C_32131_R	CCTTATCATCACCATAGAACTG	22			
6	C_32605_F	CCTTCGTATGTGCTCATATTAG	22	(TCA)6	Cig3, putative	5.7e-10
	C_32605_R	CTTCTGGAACCAATGTATCTC	21			
7	C_32654_F	CGCTCAAGTATCTCACAAG	19	(CGA)6	Acidic leucine-rich nuclear phosphoprotein 32-related protein 1	1e-31
	C_32654_R	TCGTCATCATCGTCTTCC	18			
8	C_38341_F	GGGTTCCTCATCATCTTCTT	20	(GATG)4	Zinc-binding protein	0.24
	C_38341_R	CTAGGCAGTAGGCACAAC	18			
9	C_41565_F	CATGCAGCTCCCAAGATC	18	(GATC)4	OsSAUR29 - Auxin-responsive SAUR gene family member	0.089
	C_41565_R	ACAACAACAACCACTTCCA	19			
10	C_47540_F	CCGACTTCTACTACTACTACAT	22	(CTG)6	Expressed protein	1e-25
	C_47540_R	CGTTACTCAAGCCGACAT	18			
11	C_4972_F	CCGAAGCCGAAGAAGAAG	18	(GTTGC)4	Ras-related protein	1e-18
	C_4972_R	GCCAGAACACAAGCATCA	18			
12	C_5279_F	ACCACCGCTACAACAACATT	20	(CCTG)5	Ulp1 protease family protein	8e-06
	C_5279_R	GCATATTATTATTCATGGTGGCAGC 25				

Table 3
Summary statistics of 12 eSSRs used to assess the genetic variation in finger millet

PrimerName	*Amplicon Length (bp)*		*No of Alleles*	*PIC*	*MAF*	*Heterozygosity*		*Av.Ib*	*Rp*
	Expected	*Observed*				*He*	*Ho*		
C_15674	207	254-268	8	0.8582	0.2174	0.7895	0.0000	0.2500	2.0000
C_1574	125	260-266	7	0.8318	0.2174	0.6343	0.0000	0.2857	2.0000
C_2042	125	62-87	7	0.6459	0.4667	0.6697	0.9688	0.5590	3.9130
C_22006	116	88-97	6	0.7211	0.3810	0.3536	0.0000	0.3043	1.8261
C_32131	225	81-92	4	0.4339	0.7273	0.1556	0.0000	0.4783	1.9130
C_32605	111	72-92	12	0.8809	0.2174	0.7672	0.0357	0.1667	2.0000
C_32654	117	83-95	7	0.5351	0.6667	0.5753	0.0000	0.2609	1.8261
C_38341	110	168-194	9	0.8224	0.2800	0.5906	0.1190	0.2415	2.1739
C_41565	104	84-111	14	0.8284	0.3077	0.5578	0.5563	0.3230	4.5217
C_47540	143	33-43*	5	0.4941	0.6923	0.2403	0.1979	0.4522	2.2609
C_4972	125	116-136	10	0.8469	0.1912	0.4031	0.5000	0.5913	5.9130
C_5279	125	82-98	3	0.1626	0.9130	0.1472	0.0000	0.6667	2.0000
Max			14	0.8809	0.9130	0.7895	0.9688	0.6667	5.9130
Min			3	0.1626	0.1912	0.1472	0.0000	0.1667	1.8261
Mean			7.6	0.6718	0.4398	0.4903	0.1981	0.3816	2.6956

* Not Significant; PIC- Polymorphism Information Content; MAF- Major allele frequency; H_e- Heterozygosity Expected; H_o- Heterozygosity observed; Av.Ib- Average band informativeness; Rp- Resolving power of the primer

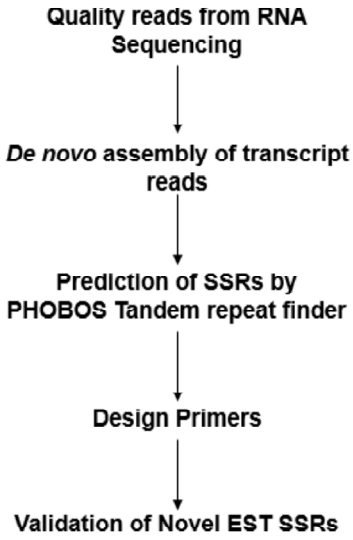

Figure 1: Bioinformatics work flow showing the steps involved in identification of EST SSRs from RNA-Seq data

1 U Taq DNA polymerase (GeNei) and 0.5ìM each of forward and reverse primers was performed in MyCycler (BioRAD, USA) with the following PCR conditions: DNA denaturation at 95°C for 5 min followed by 30 cycles of 95°C for 45 s, primer annealing at 55°C for 45 s and a primer extension step of 72°C for 45 s and a final extension step at 72°C for 10 min. The PCR products were separated on 3% agarose gel. There was little polymorphism observed between the varieties on agarose gels. Therefore, allele sizing was done by capillary electrophoresis of the amplified products in a fragment analyzer (ABI, Bioanalyzer, USA) using 20, 100, 500 and 1000 bp DNA standards.

Data scoring and marker statistics

The data scoring was done for each specific allele of each primer in the format of 0, 1 scoring, where ''1'' indicated the presence of a specific allele (band) and ''0'' indicated its absence. Polymorphism Information Content (PIC), which measures the discriminatory power of each SSR locus (Anderson, *et al.*, 1993), number of alleles per locus, frequency of the major allele, observed heterozygosity and expected heterozygosity for the 12 polymorphic markers were calculated using the formula as described by (Liu, and Muse 2005). A PIC value of each locus was calculated as: $PIC_j = 1 - \sum_{l=1 \text{ to } L} P^2_{lj}$, where P_{lj} is the relative frequency of the l^{th} allele for the locus j and is summed across all the alleles (L) over all lines. The expected heterozygousity and observed heterozygousity estimates were determined by data analysis using Arlequin software (Excoffier, *et al.*, 2005). The other parameters such as resolving power, band informativeness (Ib), major allele frequency, allelic range were also estimated for confirming the marker efficiency as described by (Guasmi, *et al.*, 2012). Band informativeness was estimated using the formula, $Ib = 1-(2 \times | 0.5 - p |)$, where p is the proportion of the total genotypes containing the band, and resolving power (Rp) is the sum of Ib values of all the bands amplified by a primer, $Rp = \Sigma Ib$.

Genetic diversity analysis

Genetic diversity analysis was performed using NTSYSpc ver.2.02i (Rohlf, 2000) for generating dendrogram along with principal component analysis. Genetic similarities (GS) between pairs of accessions were measured by the Jaccard's similarity coefficient based on the proportion of shared alleles with SIMQUAL module. The clustering of accessions was done based on a similarity matrix using an unweighted pair group method with arithmetic average (UPGMA) algorithm following SHAN (Sequential, hierarchial, agglomerative and nested) module. The clustering result was used to construct a dendrogram following TREE module (Ali, *et al.*, 2008). The ordinates were developed for the similarity matrix and based on the eigen values, the grouping of the varieties were plotted in the 3-D plot in the principal component analysis in NTSYS (Rohlf, 2000).

eSSR barcoding

Microsatellite allele based barcodes are best pictorial representations of molecular blue prints of varieties. The eSSR based barcodes for the alleles generated by the 12 primers for the 23 finger millet varieties were developed using MS EXCEL 2013 where binary data were prepared as bar(s) indicating the presence of a specific allele (Heszky, and Kiss 2009; Kanupriya, *et al.*, 2011; Chinnappareddy, *et al.*, 2012).

RESULT AND DISCUSSION

Finger millet (*E. coracana* (L.) Gaertn) are rich source of nutritive cum therapeutic values, mostly neglected over years. Most recently, the need for the exploitation of the crop has been realized. Since, the genomic resources are very limited for finger millet, the present study was focused on development of eSSR (EST-SSRs) using the transcriptome data. As early as 1981 itself, through agarose gel filtration and electrophoresis experiments, presence of 20% long interspersed repeated sequences (LINES) of 4000-4200 bp, 60% short interspersed repeated sequences

(SINES) of 150-200 bp, 18% single copy genes of 1900 bp was reported in the genome of fingermillet (Gupta, and Ranjekar 1981). These results suggests that the finger millet genome has considerable number of repetitive DNA elements. This gave the cue to explore the transcriptome data set for identification of SSRs which resulted in identification of 12 eSSRs and were validated through genetic diversity analysis and development of varietal DNA fingerprints in the form of eSSR barcodes.

Identification of frequency and distribution of eSSRs:

RNA sequence information of the finger millet reported by Rahman *et al.* (2014) in a previous study were assembled to form contigs. A total of 58,522 contigs were used for the prediction of the repeat regions of di-, tri-, tetra- nucleotide repeats and the repeating units for each type of repeats varied as shown in Fig. 2. Maximum repeat regions comprised of predominantly tri nucleotide repeats (5,980). The frequency of tri-nucleotide repeats was the highest followed by hexa-, di- and tetra- nucleotide repeats, respectively. The type of repeats was predominantly composed of tri nucleotide repeats as observed in other species (Dillon *et al*., 2014).

Development of eSSR markers

The contigs having di, tri-, tetra-, and penta-nucleotide repeats were annotated by BLASTx search against rice (*Oryza sativa*) with RGAP v7.0. The repeat regions were found to fall within the genes encoding for MYB family transcription factor, RNA recognition motif containing protein, Respiratory burst oxidase, cig3, acidic leucine-rich nuclear phosphoprotein 32-related protein Auxin-responsive SAUR gene family member, ras-related protein, ulp1 protease family protein, expressed protein (Table 2). Thus, these repeats can be regarded as functional eSSRs. The flanking primers for the identified eSSRs were designed and their quality were checked *in silico* analysis by using web version of Oligo Analyser (Integrated DNA technologies, Coralville, Iowa) which showed that the primers were devoid of hairpins, homodimers and heterodimers.

Validation of eSSR primers

The validation of the designed eSSR primers were done by PCR amplification and separation of the PCR products on a 3% agarose gel electrophoresis. The resolution of the eSSR primers was poor as the size of the bands observed on the agarose gel was uniform

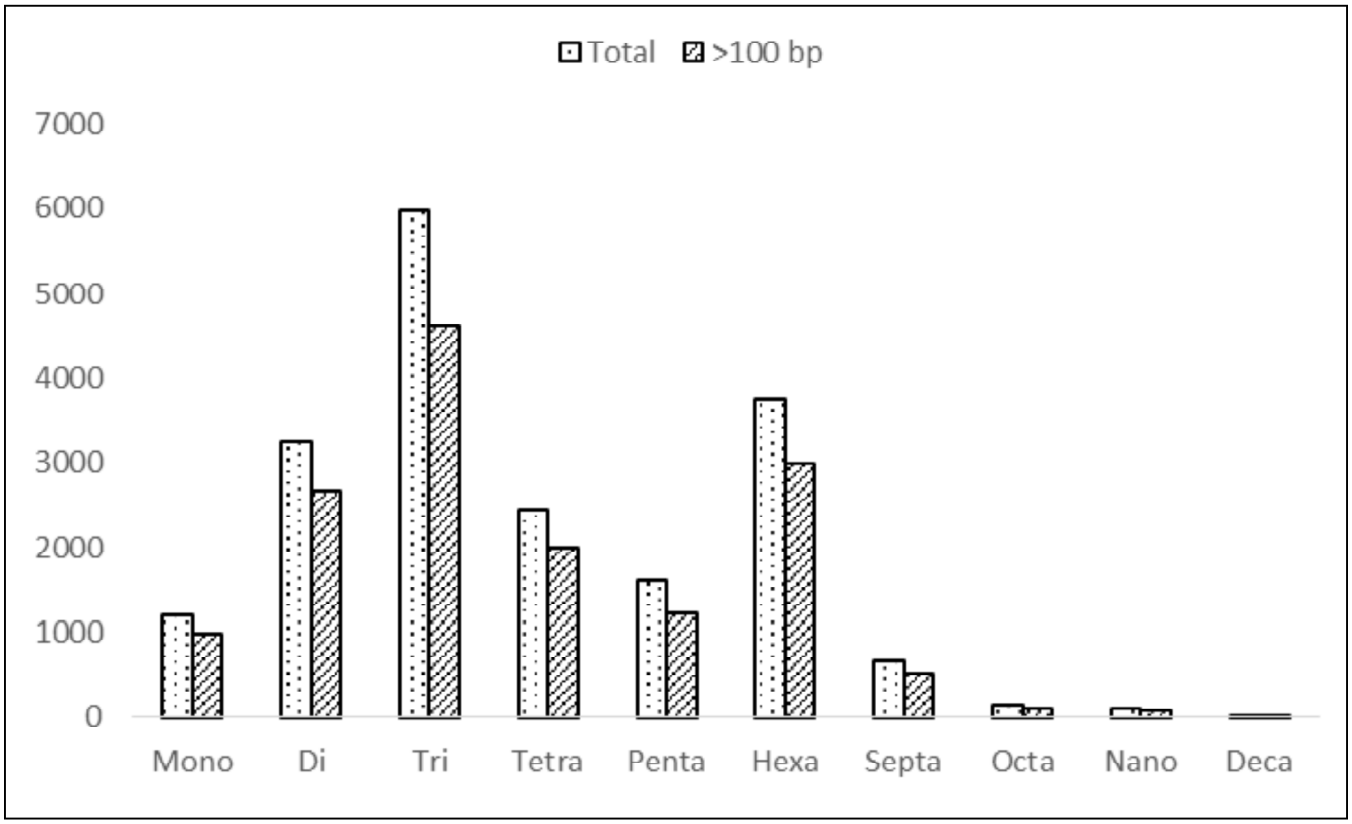

Figure 2: Distribution of eSSRs in finger millet

for all the 23 genotypes of finger millet. The SSR markers or microsatellites are 1-6 bp small tandem repeats of DNA vary between different varieties within the same species with a difference of few bp based on the number of times the repeating units are added during slippages in recombination. In order to determine the minor variations in few bp, the best method for resolving the PCR products is capillary electrophoresis. The same products were run on capillary electrophoresis system which determined the allele sizes. Already, capillary electrophoresis based allele size determination for SSR amplicons have been reported in several crops (Dillon, *et al*., 2014; Mathithumilan, *et al*., 2013) which has been useful in establishing the genetic variation. Among the 12 eSSR primer pairs, 2 primers produced amplicons exactly of the expected sizes, 3 primer pairs resulted in larger size amplicons than expected, indicating that there may be an insertions or intron within the genomic sequences. The amplicons of the remaining 7 primer pairs were smaller than expected, indicating that a deletion within the genomic sequence might have occurred. The other probable reasons that could be attributed are either a lack of primer specificity or the possibility of assembly errors (Zhou, *et al*., 2014). Cross species amplification is one of the important property of SSR markers. The eSSR markers developed for finger millet were used for confirmation of cross species amplification in proso millet (*Panicum miliaceum*), barnyard millet (*Echinochloa frumentacea*) and fox tail millet (*Setaria italica*). All the primers produced higher allelic range (~1 kb) invariably in all the 3 millets, which suggests that the genes must have undergone major deletions during evolution of finger millet.

Statistical analysis for establishing the eSSR's efficiency

In total, 92 alleles were identified with average of 7.6 alleles per locus varying from 3 to 14. PIC provides an estimate of the discriminatory power of a locus by taking into account, not only the number of alleles that are expressed, but also the relative frequencies of those alleles. PIC values usually range from 0 (monomorphic) to 1 (very highly discriminative), with many alleles in equal frequencies. It depends on the number of detectable alleles and the distribution of their frequency (Botstein, *et al*., 1980; Anderson, *et al*., 1993). In this study, PIC values of 12 eSSRs ranged from 0.1626 to 0.8809 with the average of 0.6718. The lowest PIC value was observed in C_5279 primer with 3 alleles while the highest PIC value was obtained for C_32605 with 12 alleles. Earlier, Ramdoss (2014) has reported PIC values ranging between 0.024 and 0.853 with a mean of 0.495 in finger millet genotypes. The average PIC values obtained in the present study is much better in comparison with the earlier report proving that the eSSRs developed are efficient in terms of discriminating power. The observed heterozygosity (H_O) for individual loci ranged from 0.0000 to 0.9688 with a mean of 0.1981. The expected heterozygosity also referred as gene diversity ranged from 0.1472 to 0.7895 with a mean of 0.4903. However, gene diversity was earlier reported in the range of 0.024 to 0.327 with an average of 0.290 (Nirgude *et al*., 2014). Average band informativeness (AvIb) is a measure of closeness of a band to be present in 50% of the genotypes. The average band informativeness varied from 0.1667 to 0.6667 with the average of 0.3816. The resolving power of the primers was observed from 1.8261 to 5.9130 with the average of 2.6956. Resolving power is a characteristic of a primer which reflects overall suitability of a marker system for the purpose of identification, as it is related to the number of specimens distinguished by that primer. Since cumulative band informativeness is a part of the resolving power of the primers, both these estimates are indicators of the efficiency of the primers to produce informative alleles (Table-3). The primers with AvIb and Rp values ranging above 0.25 and 1.0, respectively, are considered as efficient primers (Srivastav *et al*., 2012). Out of the 12 eSSRs, 9 primers were found to be efficient (Table 4). The major allele frequency ranged from 0.1912 to 0.9130 with the average of 0.4398 which were in accordance with the report on major allele frequency in the range of 0.167 to 0.982 with average of 0.485 by Manyasa *et al*. (2015). Earlier, it was reported as 0.757 to 0.98 with an average major allele frequency of 0.757 (Nirgude *et al*., 2014). Based on the statistical analysis, it is evident that the eSSR primers are more efficient in terms of PIC, discriminary power, major allele frequency, AvIB and Rp over the SSR markers used for finger millet genetic diversity analysis by different research groups.

Genetic diversity analysis

The Jaccard's similarity coefficients ranged between 0.029 and 0.650. The least similarity coefficient of 0.029 was observed between RAU8-GPU28 and GPU28-IE4671 genotypes, while the highest similarity coefficient 0.650was observed between GPU26-GPU86. (Table 4). This wide range of similarity coefficients indicates that there exists a considerable genetic variation among the finger millet genotypes.

Table 4
Similarity matrix of Finger millet Accessions tested with eSSR markers

	KMR301	IE4121	IE3952	GPU66	IE4329	PR202	DPI 00904	GPU26	CO13	CO12	IE4671	CO(RA)14	IE3317	KM252	HR314	Trichy1	CO7	RAU8	Paiyur2	Indof5	PES110	HR911	GPU28
KMR301	1.000																						
IE4121	0.348	1.000																					
IE3952	0.348	0.455	1.000																				
GPU66	0.292	0.280	0.524	1.000																			
IE4329	0.54	0.192	0.240	0.292	1.000																		
FR202	0.48	0.333	0.455	0.39	0.48	1.000																	
DPI 00904	0.54	0.192	0.409	0.192	0.250	0.348	1.000																
GPU26	0.333	0.222	0.435	0.650	0.231	0.320	0.185	1.000															
CO13	0.348	0.231	0.231	0.23	0. 48	0.143	0.107	0.320	1.000														
CO12	0.348	0.524	0.455	0.280	0. 48	0.280	0.240	0.320	0.455	1.000													
IE4671	0.280	0.222	0.179	0.100	0. 03	0.100	0.103	0.133	0.320	0.269	1.000												
CO(RA)14	0.269	0.350	0.417	0.308	0. 38	0.308	0.179	0.400	0.308	0.478	0.250	1.000											
IE3317	0.222	0.214	0.308	0.478	0.222	0.308	0.138	0.458	0.214	0.259	0.207	0.286	1.000										
KM252	0. 92	0.231	0.391	0.524	0.248	0.600	0.192	0.500	0.185	0.280	0.100	0.360	0.417	1.000									
HR314	0. 92	0.185	0.333	0.39	0.292	0.455	0.192	0.320	0.143	0.185	0.100	0.214	0.308	0.524	1.000								
Trichy1	0. 54	0.292	0.409	0.292	0. 54	0.409	0.250	0.333	0.192	0.348	0.143	0.269	0.222	0.348	0.292	1.000							
CO7	0.240	0.333	0.455	0.280	0.240	0.333	0.292	0.320	0.231	0.391	0.222	0.308	0.360	0.280	0.333	0.476	1.000						
RAU8	0.320	0.259	0.214	0.172	0. 38	0.133	0.100	0.250	0.308	0.417	0.400	0.333	0.241	0.214	0.214	0.222	0.360	1.000					
Paiyur2	0.292	0.391	0.333	0.185	0. 48	0.231	0.192	0.222	0.280	0.524	0.320	0.417	0.172	0.185	0.143	0.348	0.455	0.308	1.000				
Indof5	0.348	0.333	0.280	0.185	0. 48	0.143	0.107	0.222	0.231	0.391	0.320	0.259	0.133	0.185	0.143	0.409	0.280	0.417	0.524	1.000			
PES110	0. 92	0.333	0.231	0.280	0. 07	0.185	0.069	0.269	0.280	0.280	0.138	0.214	0.214	0.185	0.143	0.292	0.280	0.214	0.280	0.231	1.000		
HR911	0.036	0.154	0.250	0.200	0. 60	0.304	0.208	0.192	0.034	0.154	0.069	0.143	0.231	0.250	0.250	0.381	0.304	0.103	0.304	0.250	0.111	1.000	
GPU28	0. 00	0.063	0.133	0.214	0.222	0.172	0.100	0.167	0.063	0.063	0.029	0.059	0.200	0.214	0.172	0.100	0.172	0.029	0.133	0.063	0.063	0.143	1.000

The highest frequency of 15 was observed for similarity coefficient 0.280, followed by 11 for 5 similarity coefficients, *viz*., 0.333, 0.348, 0.231, 0.214 and 0.192. Single frequency or unique value was obtained in 17 similarity coefficients. The frequency of the similarity coefficients indirectly explains the probability of recombination events between the two genotypes at a specific genetic lineage. However, the sites of chromosomal cross overs and number of chiasmata may not be fixed at a specific location but to distinct locations between two genotypes surveyed by the marker.

The dendrogram constructed by UPGMA (Unweighted pair group method with arithmetic mean) involving SHAN module grouped the 23 finger millet genotypes into major two clusters with two out-groups comprising of GPU28 and HR911 in the peripheral clades. Interestingly, both these genotypes belong to the same geographical location, Bangalore of Karnataka, India (Fig. 4). The major cluster I comprised of the genotypes collected from Coimbatore, Tamil Nadu (CO12, CO13, CO14, CO7, Trichy1, Paiyur2, RAU8, KMR301); Bangalore, Karnataka (PES110, Indof 5); and other areas of India (IE4671, IE4121). The major cluster II clustered the genotypes of Uganda (IE3952), Zimbabwe (IE4329, IE3317), certain genotypes from Coimbatore, TamilNadu (DPI 009-04, KM252) and Bangalore, Karnataka (PR202, GPU26, GPU66). Even in the principal component analysis, similar grouping pattern was observed as two major groups (Fig. 5). This clustering pattern though did not follow any specific pattern of grouping of the genotypes. The genotypes from same eco-geographical adaptations was clustered in different clusters. This suggests that the genotypes probably have wider genetic variation which is evident from the sub-clustering pattern as well as their low similarity coefficients. This is also supported by the statistical estimates of PIC, AvIb and Heterozygosity values.

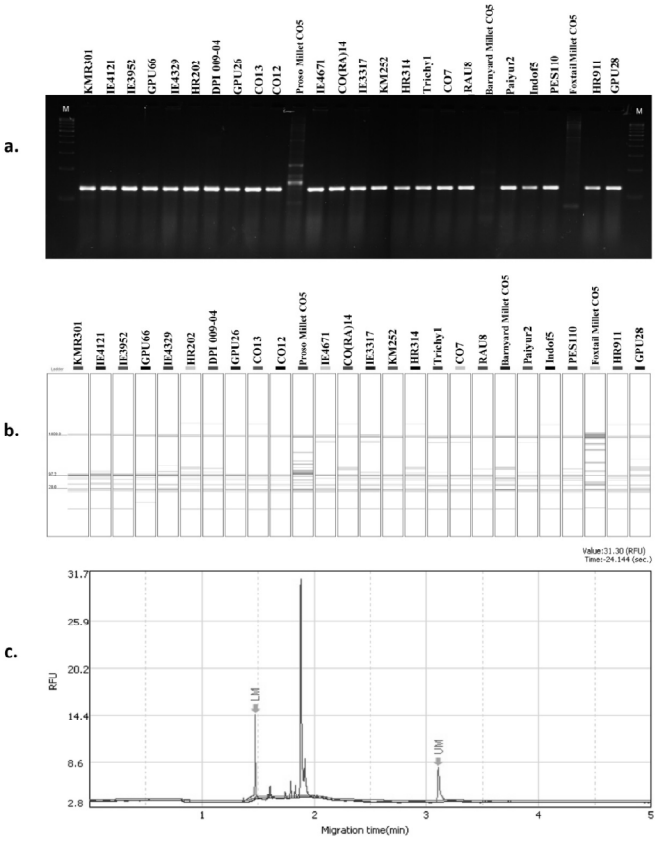

Figure 3: Representative gel showing amplification profiles of eSSR primer pair C_5279 a. PCR products were separated through 3% agarose gel for finger millet and other related millets b. Capillary electrophoresis of C_5279 c. Electrophoretogram depicting allele for KMR301 genotype amplified through C_5279 primer

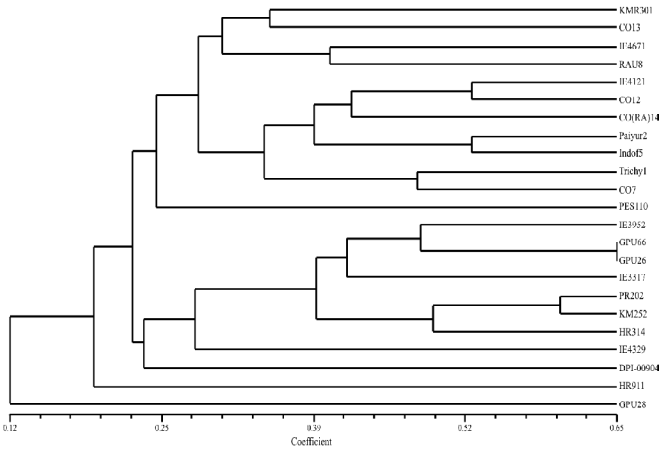

Figure 4: Dendrogram of 23 finger millet (*Eleusine coracana*) genotypes revealed by cluster analysis of genetic similarity estimates generated by Jaccard coefficient based on 12 EST-SSR markers

DNA fingerprinting of finger millet genotypes through EST-SSR barcoding:

DNA barcode is a powerful and deployable tool for assessing the identity of a genotype, variety or cultivar. For developing DNA barcodes, the unique nucleotide sequence patterns of small DNA fragments (400–800 bp) are taken as specific reference collections to identify the plant species using candidate genes, *viz*., *cox*, *rbc*L, *psb*A, *trn*L, *mat*K, etc (Ratnasingham and Hebert, 2007 and Vijayan and Tsou, 2010). The best alternative to develop barcode is to use the

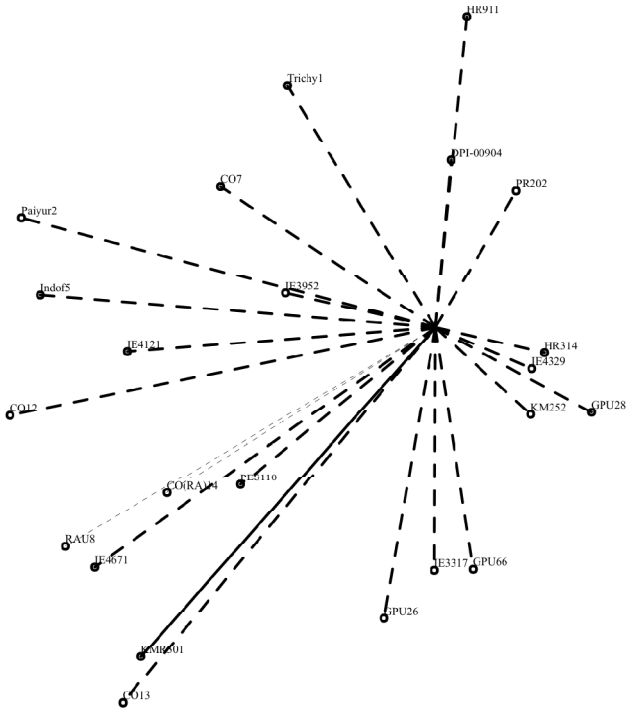

Figure 5: Principle component analysis of 23 finger millet (*Eleusine coracana*) genotypes based on 12 polymorphic EST SSR markers

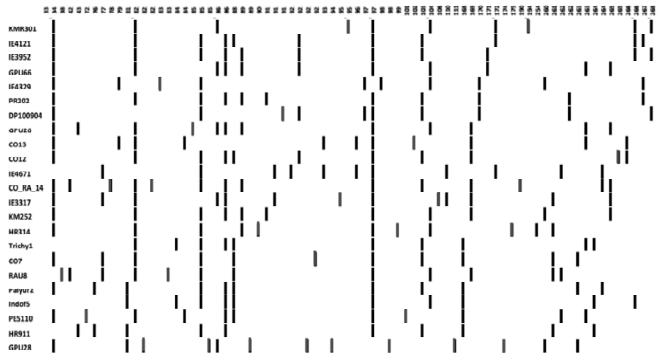

Figure 6: eSSR barcodes of 23 finger millet (*Eleusine coracana*) genotypes. Each bar represent an allele. Each red bar represents unique allele specific to the variety

molecular marker (allelic) data of microsatellite or simple sequence repeat (SSR) markers. Microsatellite regions are usually biallelic or multi-allelic forms generally found at each locus which make them highly informative. Hence, eSSR barcoding was attempted in finger millet genotypes. The allelic data of 23 finger millet genotypes generated by 12 eSSRs was developed into barcode (Fig. 6). The unique alleles are distinct alleles for the identification of the genotypes (marked in red in Fig.6). GPU28 variety produced 9 alleles which are specific to the variety followed by PES110 and HR314. GPU28 has distinct unique alleles which has out-grouped it from other genotypes even in the clustering. HR314 produced 4 specific alleles by four different primers (C_38341; C_32605; C_41565; C_15674) indicating that this genotype is distinct from other genotypes. DPI00904 produced 3 unique alleles by C_32654 (1 allele) and C_2042 (2 alleles). Each barcode represents a specific allele in that genotype irrespective of the primer (Heszky and Kiss. 2009). Earlier this kind of microsatellite barcodes have been reported in grapes, brinjal, guava (Heszky and Kiss. 2009; Chinnappareddy *et al*., 2012; Kanupriya *et al*., 2011). This system of microsatellite barcoding solves the problem of reproducibility and reliability in comparison to the organellar gene sequence based barcoding and would also serve as molecular signatures of documented fingerprints of the finger millet genotypes.

ACKNOWLEDGMENTS

The financial support of Department of Bio-technology, Government of India, New Delhi (Grant Number BT/PR-10482/AGR/02/564/2008) is greatly acknowledged.

REFERENCES

Ali, M., Rajewski, J., Baenziger, P., Gill, K., Eskridge, K., and Dweikat, I. (2008), Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. *Molecular breeding,* **21**(4): 497-509.

Anderson, J., G. Churchill, J. Autrique, S. Tanksley and M. Sorrells. (1993), Optimizing parental selection for genetic linkage maps. **Genome. 36:** 181-186.

Babu, B. K., Dinesh, P., Agrawal, P. K., Sood, S., Chandrashekara, C., Bhatt, J. C., and Kumar, A. (2014), Comparative Genomics and Association Mapping Approaches for Blast Resistant Genes in Finger Millet Using SSRs. *PLoS ONE,* **9**(6): e99182.

Babu, B. K., Senthil, N., Gomez, S. M., Biji, K., Rajendraprasad, N., Kumar, S. S., and Babu, R. C. (2007), Assessment of genetic diversity among finger millet (*Eleusine coracana* (L.) Gaertn.) accessions using molecular markers. *Genetic Resources and Crop Evolution,* **54**(2): 399-404.

Bharathi (2011), Phenotypic and genotypic diversity of global finger millet (*Eleusine coracana* (L.) Gaertn.) composite collection. Ph. D. Thesis submitted to Tamil Nadu Agricultural University.

Bhatt, D., Negi, M., Sharma, P., Saxena, S. C., Dobriyal, A. K., and Arora, S. (2011)., Responses to drought induced oxidative stress in five finger millet varieties differing in their geographical distribution. *Physiol. Mol. Biol. Plants.* 17, 347–353.

Bisht MS, Mukai Y (2001), Genomic in situ hybridization identifies genome donor of finger millet (*Eleusine coracana*). *Theor. Appl. Genet.* 102: 825–832.

Botstein, D., R.L. White, M. Skolnick and R.W. Davis. (1980), Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics.* **32:**314.

Chandrasekara, A., and Shahidi, F. (2011), Determination of antioxidant activity in free and hydrolyzed fractions of millet grains and characterization of their phenolic profiles by HPLC-DAD-ESI-MS. *J. Funct. Foods* **3,** 144–158.

Chase, M.W., R.S. Cowan, P.M. Hollingsworth, C. van den Berg, S. Madriñán, G. Petersen, O. Seberg, T. Jørgsensen, K.M. Cameron and M. Carine. (2007), A proposal for a standardised protocol to barcode all land plants. *Taxon.* **56:** 295-299.

Chinnappareddy, L.R.D., K. Khandagale, S.H. Srinivas Reddy, C. Kanupriya, A. Chennareddy and T. Singh. 2012. SSR-Based DNA Barcodes as a Tool for Identification of Eggplant Genotypes. *International Journal of Vegetable Science.* **18:** 260-271.

Devi, P. B., Vijayabharathi, R., Sathyabama, S., Malleshi, N. G., and Priyadarisini, V. B. (2014), Health benefits of finger millet (*Eleusine coracana* L.) polyphenols and dietary fiber: a review. *Journal of food science and technology,* **51**(6):1021-1040.

deWet, J. (1995), Finger millet. In *Evolution of Crop Plants*, eds. S. J and S. NW, 137 – 140. Harlow: Longman Scientific and Technical.

Dida, M. M., and Devos, K. M. (2006), "Finger millet," in *Cereals and Millets*, ed C. Kole (New York, NY: Springer), 333–343.

Dida, M. M., Ramakrishnan, S., Bennetzen, J. L., Gale, M. D., and Devos, K. M. (2007), The genetic map of finger millet, Eleusine coracana. *Theoretical and applied genetics,* **114**(2):321-332.

Dillon, N. L., Innes, D. J., Bally, I. S., Wright, C. L., Devitt, L. C., and Dietzgen, R. G. (2014), Expressed Sequence Tag-Simple Sequence Repeat (EST-SSR) Marker Resources for Diversity Analysis of Mango (*Mangifera indica* L.). *Diversity,* **6**(1):72-87.

Doyle, JJ. (1987), A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19, 11-15

Excoffier, L., G. Laval and S. Schneider. (2005), Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics online.* 1: 47.

Ganapathy, S., A. Nirmalakumari and A.R. Muthiah, (2011), Genetic variability and interrelationship analyses for economic traits in finger millet germplasm. *World Journal of Agricultural Sciences*, **7:** 185-188.

Goron, T.L. and M. Raizada. (2015), Genetic diversity and genomic resources available for the small millet crops to accelerate a New Green Revolution. *Frontiers in Plant Science.* 6.

Guasmi, F., Elfalleh, W., Hannachi, H., Fères, K., Touil, L., Marzougui, N., Triki, T., and Ferchichi, A. (2012), The use of ISSR and RAPD markers for genetic diversity among South Tunisian Barley. *ISRN Agronomy,* **2012**.

Gupta, V. S., and Ranjekar, P. (1981), DNA sequence organization in finger millet (Eleusine coracana). *Journal of Biosciences,* **3**(4):417-430.

Harlan, J. R. (1971), Agricultural origins: centers and noncenters. *Science,* **174**(4008): 468-474.

Heszky, L. and E. Kiss. (2009), Identification of grapevine cultivars using microsatellite-based DNA barcodes. *Vitis.* **48:** 17-24.

Kanupriya, Latha, P. M., Aswath, C., Reddy, L., Padmakar, B., Vasugi, C., and Dinesh, M. (2011), Cultivar identification and genetic fingerprinting of guava (*Psidium guajava*) using microsatellite markers. *International Journal of Fruit Science,* 11(2):184-196.

Liu, K., and Muse, S. V. (2005), PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics,* 21(9): 2128-2129.

Liu, Q., B. Jiang, J. Wen and P.M. Peterson. (2014), Low-copy nuclear gene and McGISH resolves polyploid history of *Eleusine coracana* and morphological character evolution in Eleusine. *Turkish Journal of Botany.* 38: 1-12.

Mathithumilan, B., Kadam, N., Biradar, J., Reddy, S., Ankaiah, M., Narayanan, M., Makarla, U., Khurana, P., and Sreeman, S. (2013), Development and characterization of microsatellite markers for Morus spp. and assessment of their transferability to other closely related species. *BMC Plant Biology,* 13(1): 194.

NAAS. (2013), Role of Millets in Nutritional Security of India. Policy Paper No. 66, National Academy of Agricultural Sciences, New Delhi.16 p.

Pandey, G., G. Misra, K. Kumari, S. Gupta, S.K. Parida, D. Chattopadhyay and M. Prasad. (2013), Genome-wide development and use of microsatellite markers for large-scale genotyping applications in foxtail millet (*Setaria italica* (L.)). *DNA research.* 20: 197-207.

Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey and A. Rafalski. (1996), The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular breeding.* 2: 225-238.

Prasad, P. V., and Staggenborg, S. A. (2009), Growth and production of sorghum and millets. *Soils, Plant Growth and Crop Production,* **2**.

Prevost, A. and M. Wilkinson. (1999), A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theoretical and Applied Genetics.* 98: 107-112.

Purugganan, M. D., and Fuller, D. Q. (2009), The nature of selection during plant domestication. *Nature,* **457** (7231): 843-848.

Rahman, H., Jagadeeshselvam, N., Valarmathi, R., Sachin, B., Sasikala, R., Senthil, N., Sudhakar, D., Robin, S., and Muthurajan, R. (2014), Transcriptome analysis of salinity responsiveness in contrasting genotypes of finger millet (*Eleusine coracana* L.) through RNA-sequencing. *Plant Molecular Biology,* **85**(4-5): 485-503.

Ratnasingham, S. and Hebert, P. D. N. (2007), BOLD: The Barcode of Life Data System (*www.barcodinglife.org*). *Mol. Ecol. Notes,* 7: 355–364.

Rohlf, F. J. (2000), NTSYS-pc: Numerical taxonomy and multivariate analysis system. Exeter Publishing, Setauket. pp 119.

Srivastava, N., Bajpai, A., Chandra, R., Rajan, S., Muthukumar, M., and Manoj Kumar Srivastava. (2012), Comparison of PCR based marker systems for genetic analysis in different cultivars of mango. *J. Environ. Biol.,* 33: 159-166.

Vanaja, R. and Green Foundation. (2013), Millets conservation in Southern India. In: Conservation and Sustainable use of agricultural biodiversity. pp. 362-367.

Vijayan, K. and C.H. Tsou. (2010), DNA barcoding in plants: taxonomy in a new perspective. *Current Science*, 99(11): 1530-1541.

Zhou, Q., Chen, T., Wang, Y., and Liu, Z. (2014), The development of 204 novel EST-SSRs and their use for genetic diversity analyses in cultivated alfalfa. *Biochemical Systematics and Ecology,* **57**: 227-230.