Korean J. Plant Res. 30(3):287-297(2017)
https://doi.org/10.7732/kjpr.2017.30.3.287

Print ISSN 1226-3591
Online ISSN 2287-8203

Original Research Article

Development of EST-SSRs and Assessment of Genetic Diversity in Little Millet (*Panicum sumatrense*) Germplasm

Asjad Ali, Yu-Mi Choi, Do-Yoon Hyun, Sukyeung Lee, Jin-Hee Kim, Sejong Oh and Myung Chul Lee*

National Agrobiodiversity Center, National Institute of Agricultural Sciences, RDA, Jeonju 54874, Korea

Abstract - Little millet (*Panicum sumatrense*) is well known for its salt and drought stress tolerance and high nutritional value, but very limited knowledge of genetic variation and genomic information is available. In this study, a total of 779 primer pairs were designed from the 22,961 EST sequences of switchgrass (*Pancium virgatum*), of which 48 EST-SSR markers were developed based on the trials of transferability of these primers in little millet. The EST-SSR amplicons showed reproducible single band polymorphism and produced a total of 160 alleles with an average of 3.3 alleles per locus in 37 accessions of little millet. The average values of expected and observed heterozygosities were 0.266 and 0.123, respectively. The polymorphic information content (PIC) values were observed in range of 0.026 to 0.549 with an average of 0.240. The genetic relatedness among the little millet accessions was evaluated by neighbor-joining dendrogram, which grouped all accessions into two distinct groups. The validation thus demonstrated the utility of the switchgrass EST-SSR markers in assessing genomic relationships in little millet. The findings from this study could be useful for designing strategies for the identification of diverse germplasm for conservation and future molecular breeding programs for little millet.

Key words - EST-SSRs, Genetic diversity, Little millet

Introduction

Small grain millets represent eight small grain species of millet which are cultivated as food crops in different countries of the world. These include finger millet [*Eleusine coracana* (L.) Gaertn.], foxtail millet [*Setaria italica* (L.) Beauv.], proso millet (*Panicum miliaceum* L.), little millet (*Panicum sumatrense* Roth. ex Roem. & Schult.), barnyard millet [*Echinochloa crusgalli* (L.) Beauv. & Echinochloa colona (L.)], kodo millet (*Paspalum scrobiculatum* L.), teff [*Eragrostis tef* (Zucc.)] and fonio millet (*Digitaria exilis* Stapf. & Digitaria iburua Stapf.) (Goron *et al*., 2015). The foxtail millet and proso millet are cultivated as a miner crop in Korea due to their ability to grow in barren soil and tolerate drought stress. However, little millet (*Panicum sumatrense*) is cultivated in semi-arid area in India, SriLanka, Pakistan, Myanmar, and other South East Asian countries (Hiremath *et al*., 1990). The maturing time of little millet is ~90 days (deWet *et al*., 1983) and it is tolerant against biotic and abiotic stresses such as pest, drought and salt (Bhaskaran and Panneerselvam, 2013; Ajithkumar and Panneerselvam, 2014). The nutritional value of little millet is comparable to other cereals and rich in phytochemicals including phenolic acids, flavonoids, tannins, and phytates (Pradeep and Guha, 2011). For this reason, little millet is one of the putative crops to introduce and cultivate on reclaimed land and prepare the global climate exchange in Korea.

The collection and conservation of plant genetic resources is important for proper utilization and improvement of existing crop resources along with the introduction of new crops. However, diversity of conserved plant genetic resources is also emphasized to provide an opportunity for plant breeders to develop new and improved cultivars with desirable characteristics. The assessment of genetic diversity within and between plant populations is routinely performed using morphological and DNA marker analysis (Govindaraj *et al*., 2015). In plant genetics, SSR markers were widely used for genetic diversity studies, molecular mapping, molecular

*Corresponding author. E-mail : mcleekor@korea.kr
Tel. +82-63-238-4900

fingerprinting and conservation strategies due to their high variability, abundance, multiallelic nature, reproducibility, polymorphism, transferability as well as their codominant inheritance, chromosome-specific location and wide genomic distribution (Jia *et al*., 2009; Gupta *et al*., 2012; Im *et al*., 2014). Recently, the expressed sequence tags (ESTs)-SSRs were suggested more efficient than genomic SSR markers, because the EST-SSRs are physically associated to coding regions and can enhance the evaluation of plant populations by enabling the variation assay in expressed genes with known function (Nicot *et al*., 2004; Seo *et al*., 2013). Therefore, identification of EST-SSRs is important to study different species of the same genus. Furthermore, the transferability of EST-SSRs between species may support the idea of exploring genes with similar function and contribute to comparative genomics and diversity analysis.

The genomic information and molecular marker related studies in little millet are very limited; however, little millet belongs to the same genus as that of switchgrass (*Panicum virgatum*) that has well defined molecular information (Wang *et al*., 2011; Casler, 2012). It is now possible to easily obtain thousands of switchgrass ESTs that could be the main source of identification for *in silico* EST-SSRs.

The objective of this study was, to develop highly polymorphic EST-SSR markers based on cross-species transferability of derived SSRs from switchgrass EST databases and characterize newly developed EST-SSRs to better understand the genetic diversity of collected germplasm of little millet.

Materials and Methods

Plant material and DNA extraction

Thirty seven accessions of little millet, obtained from ICRISAT, India were used in this study (Table 1). All the accessions were maintained at the National Agrobiodiversity Center (NAS, RDA, Republic of Korea). The seeds were germinated and grown in the greenhouse. The leaves were harvested at two-leaf stage after 15 days of planting and genomic DNA was extracted according to Qiagen DNeasy Plant Mini kit protocol (QIAGEN, Germany). The concentrations of DNA were estimated using NanoDrop ND-1000 (NanoDrop Technologies Inc., USA) and final adjustment was made at 100 ng/μL. DNA samples were stored at a temperature of -20℃ for further use.

Database mining for EST-SSRs and primer design

Publicly available database of the National Center for Biotechnology Information (NCBI) was used to deduce the sequences of switchgrass as on March 17, 2016. A total of 22,961 EST sequences were retrieved from NCBI to search simple sequence repeats. Those ESTs were assembled into unigenes using SeqMan DNA Star lasergene version 7.1 (DNASTAR Inc, Madison, WI) and the parameters for clustering were set at a minimum of 98% identity in 30-bp overlap. The unigenes were used for identification of the microsatellite primer pairs via simple sequence repeat identification tool (SSRIT) (http://www.gramene.org/gramene/searches/ssrtool) and SSR locator V.1 software (da Maia *et al*., 2008). The criteria for selection of markers were repeat units of di-, tri-, tetra-, penta-, and higher nucleotides and expected product sizes in the range of 100 to 400 bp.

PCR amplifications

EST-SSRs were amplified in a 20 μl total volume containing 50 ng of genomic DNA, 2 μl of each EST-SSR primer (10 pmol), 4 μl of 5x Green GoTaq reaction Buffer (Promega Co, USA), 1 U of *Taq* DNA polymerase (Promega Co., USA), 1.6 μl of dNTP (2.5 mM), and 11 μl nuclease-

Table 1. List of little millet accessions used in this study

Origin country	Accession Number								
India	IT153625	IT261890	IT261894	IT261898	IT261902	IT261906	IT261910	IT284250	IT284254
	IT153626	IT261891	IT261895	IT261899	IT261903	IT261907	IT261911	IT284251	IT284255
	IT153627	IT261892	IT261896	IT261900	IT261904	IT261908	IT261912	IT284252	IT297286
	IT153628	IT261893	IT261897	IT261901	IT261905	IT261909	IT261913	IT284253	IT297288
	IT153629								

free water. DNA amplifications were performed in PTC-100 thermal controller (MJ Research Watertown, MA, USA). The PCR profile was: initial denaturation of 3 min at 94℃, followed by 35 cycles of 45s at 50-55℃ and 45s at 72℃, and a final extension of 10 min at 72℃. All the amplifications were separated by capillary electrophoresis using Fragment Analyzer™ 12-capillary Automated CE System using DNF-900 double stranded DNA Reagent Kit (Advanced analytical Technologies, USA).

Data Analysis

Blastn was used for the annotation of EST-SSRs with default parameters at NCBI database. Analysis for different parameters of variability such as number of alleles (N_A), expected heterozygosity (H_E), observed heterozygosity (H_o), and polymorphism information content (PIC) was performed by employing PowerMarker V3.25 (Liu and Muse, 2005). The neighbor joining method was used to construct a dendrogram with the help of DARwin 6.0. (Perrier and Jacquemoud-Collet, 2006). A Principal Coordinate Analysis (PCoA) was used to investigate the genetic relationships between garlic in each country based on EST-SSR data (Peakall and Smouse, 2006).

Results

Frequency distribution of EST-SSR markers

A total of 22, 961 non-redundant ESTs of switchgrass were used to find the EST-SSR domain. The mining of SSRs from the pool of EST sequences demonstrated the presence of 779 microsatellite repeats with a minimum SSR length set to 20 bp and an amplicon length ranging from 100 to 280 nucleotides. Thus 3.39% ESTs contained at least one SSR. A variation of repeat motifs in terms of di-, tri-, tetra-, penta- or higher nucleotides (hexa- and above) was observed among those 779 EST-SSRs. The vast majority of EST-SSRs were di-nucleotide repeats 289 (37%). The second most dominant type of EST-SSRs was tri-nucleotides 242 (31%), which are comparatively more enriched in protein coding regions, followed by penta- and tetra-nucleotides with frequencies of 60 (8%) and 40 (5%), respectively. The remaining proportion of 19% was covered by hexa- and higher nucleotides (Fig. 1).

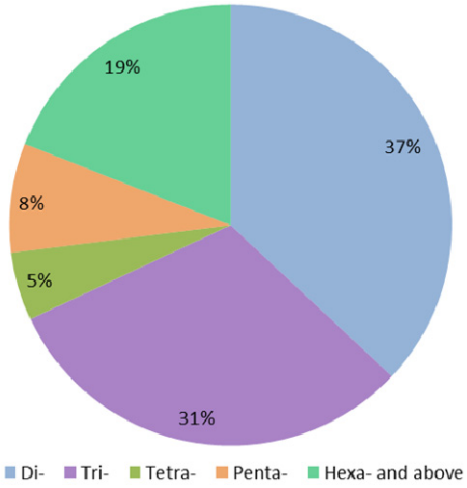

Fig. 1. Frequency distribution of EST-derived SSR markers.

Genetic diversity analysis of EST-SSR markers

Forty eight (6.2%) of 779 EST-SSRs were shown highly polymorphic among 37 accessions of little millet, for instance, Fig. 2 shows the amplification of PS-E-SSR299. The primer sequence information, product sizes, and repeat motifs for the 48 EST-SSR markers developed in this study have been mentioned in Table 2, while the summary of genetic diversity analysis of 48 EST-SSR markers by PowerMarker has been shown in Table 3. A total of 160 alleles were identified and the number of alleles per locus ranged from 2 to 7 for polymorphic markers. A mean of 3.3 alleles per marker was observed across the 48 EST-SSR markers. Generally, higher genetic variability is uncovered by higher values of H_E and H_o. As listed in Table 3, the average value for the H_E was 0.266 ranging from 0.027 (PS-E-SSR121) to 0.615 (PS-E-SSR362). However, the value for H_o ranged from 0.000 to 1. The PIC values ranged from 0.026 to 0.549 with an average of 0.240 across all the EST-SSRs. Out of 48 EST-SSRs, 60.4% markers (PS-E-SSR47, PS-E-SSR71, PS-E-SSR167, PS-E-SSR178, PS-E-SSR182, PS-E-SSR185, PS-E-SSR195, PS-E-SSR201, PS-E-SSR220, PS-E-SSR258, PS-E-SSR264, PS-E-SSR271, PS-E-SSR284, PS-E-SSR303, PS-E-SSR304, PS-E-SSR308, PS-E-SSR320, PS-E-SSR328, PS-E-SSR357, PS-E-SSR362, PS-E-SSR363, PS-E-SSR379, PS-E-SSR383, PS-E-SSR384, PS-E-SSR404, PS-E-SSR411, PS-E-SSR415, PS-E-SSR433, PS-E-SSR448) were highly polymorphic with PIC values in a range of 0.174 to 0.549, whereas 39.6% of the markers (PS-E-SSR31, PS-E-SSR121, PS-E-SSR125, PS-E-

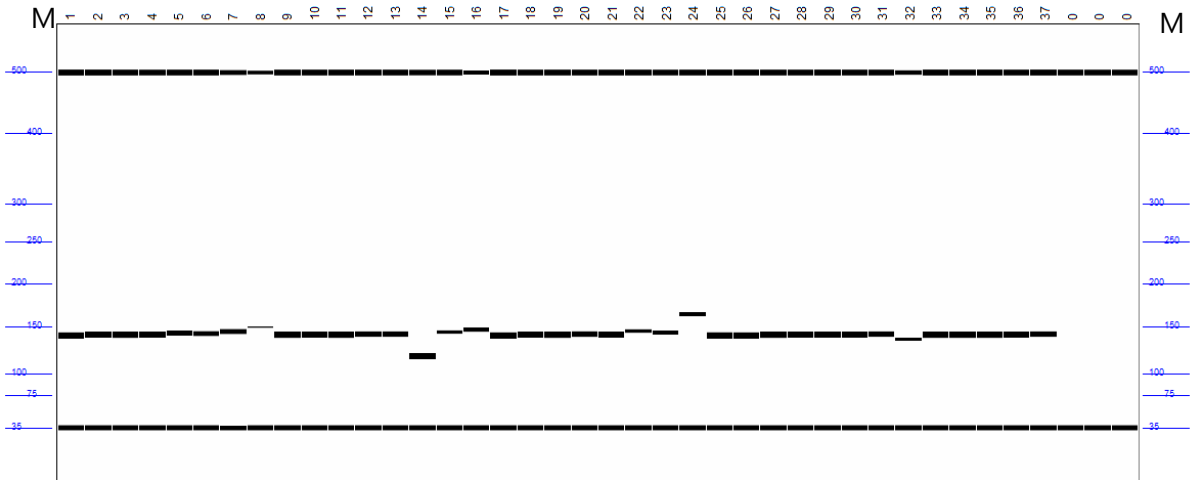

Fig. 2. Amplicons of 37 little millet accessions using PS-E-SSR299 primer. M: size markers of 500bp and 35bp, Lanes 1-37: The name of accessions as listed in Table 1.

Table 2. Description of EST-SSR markers developed in this study

Marker Name	Forward Primer	Product size	Repeat motif	E-value	Putative identities
PS-E-SSR31	F-ACTTCCCTAGAGTTCCAGT R-TTCTGAAACTGTTCTATTGG	217	(CCACGG)3	0.0	Predicted: *Setaria italica* la-related protein 1B
PS-E-SSR47	F-ACTTTGATCACATTAGCATT R-GATTATGGTACGGACTTGTA	278	(AG)44	0.0	Predicted: *Setaria italica* 1-aminocyclopropane-1-carboxylate oxidase
PS-E-SSR71	F-ACTCATCTGACAAACTATGG R-ATAGAACTGTGTGTTGGTGT	137	(GA)34	0.0	Predicted: *Setaria italica* flowering time control protein FPA
PS-E-SSR121	F-ACCTGATCGCCTGAG R-GATTCTGGACAAGCTGTAT	273	(AT)28	0.0	Predicted: *Setaria italica* uncharacterized LOC101762880
PS-E-SSR125	F-ACCTTAAGGATTGGAATATC R-GTTGAGTAAGTTTCTCCTCA	254	(TC)28	5e-111	*Panicum virgatum* microsatellite PVSSR008 sequence
PS-E-SSR142	F-ACGAGACGACAGAACTACTA R-ACCTCCTCTAGTACATACACC	127	(CCG)9	1e-152	Predicted: *Setaria italica* gibberellin receptor GID1
PS-E-SSR144	F-ACGAGATATTCTACCTGGAG R-CTCCACTTCATCTCCTTT	269	(AG)27	0.0	Predicted: *Setaria italica* uncharacterized LOC101778542
PS-E-SSR167	F-ACGTACGTGTTTTGTACTTT R-TACTATTTGAAGAGCTGAGG	125	(CAG)9	0.0	Predicted: *Setaria italica* aquaporin PIP1-5
PS-E-SSR178	F-ACTACAAGTCAAGTCTGCAA R-AAGGAGGAGGTGGTGT	196	(TA)25	0.0	*Sorghum bicolor* hypothetical protein
PS-E-SSR182	F-ACTACCTCACCAACATCTG R-GCCCCTGCTGCTTAG	270	(GA)25	4e-126	Predicted: *Setaria italica* B-box zinc finger protein 20-like
PS-E-SSR185	F-ACTACGACGTCCAAGAC R-GTTGGACGAATCGTACTC	221	(CA)24	0.0	Predicted: *Setaria italica* decapping nuclease DXO homolog, chloroplastic
PS-E-SSR195	F-ACATCCATGATCTCCCT R-ACAGGAAAACAAACAAACTA	167	(CGC)16	1e-137	Predicted: *Setaria italica* DEAD-box ATP-dependent RNA helicase 2
PS-E-SSR199	F-ACATGACGATGATGAAGA R-GAACTGGCAGAAGCAC	197	(AG)14	2e-80	*Sorghum bicolor* hypothetical protein
PS-E-SSR201	F-ACATGTTCTGGACAATTTAC R-AACTACTGTCTAGAGTCCCC	186	(TAA)16	0.0	Predicted: *Setaria italica* probable protein phosphatase 2C 28
PS-E-SSR220	F-ACAACAAGAAGAGCAACC R-GAGATGTAGAGTTTGGTGC	280	(TC)23	0.0	Predicted: *Setaria italica* THO complex subunit 4A-like (LOC101768390)

Table 2. Description of EST-SSR markers developed in this study (Continued)

Marker Name	Forward Primer	Product size	Repeat motif	E-value	Putative identities
PS-E-SSR232	F-ACAAGAAGGATAGCAGGA R-TATTTCTTTGAAGACTGCAT	197	(CGC)15	0.0	Predicted: *Setaria italica* splicing regulatory glutamine/lysine-rich protein 1
PS-E-SSR258	F-ACACCACACCACCATC R-GAGATGATGTGGCCC	277	(AT)21	0.0	Predicted: *Setaria italica* shaggy-related protein kinase eta-like
PS-E-SSR264	F-ACACCATTGATCTTACACAT RGTATGCTCCGTGTGCT	117	(AG)21	7e-08	Predicted: *Setaria italica* chlorophyll a-b binding protein
PS-E-SSR269	F-ACACTACAGACCTCACCTCT R-AAACTCTTCATTTTGGTGAT	271	(AT)21	0.0	Predicted: *Setaria italica* histone deacetylase complex subunit SAP18
PS-E-SSR270	F-ACACTAGTACACCGGTCAT R-CGAAGTAGACGCTCTTG	238	(GAA)10	0.0	Predicted: *Setaria italica* photosystem I reaction center subunit VI
PS-E-SSR271	F-ACACTCCCACCATCTCT R-CAAGACCAATAAAACAAAAG	192	(AT)21	2e-61	Predicted: *Setaria italica* ethylene-responsive transcription factor RAP2-13-like
PS-E-SSR284	F-ACAGCATCTACACGTTTATT R-CGGCGAGATATAGACAA	174	(GA)21	0.0	Predicted: *Setaria italica* probable NAD(P)H-dependent oxidoreductase 2
PS-E-SSR299	F-ACATAAATGACAAAGGAAAA R-AATCAGTATGAGCAGAACAC	162	(TGT)13		No significant similarity found
PS-E-SSR303	F-ACATACACAATCAAGGAAAG R-TGGCTATACTAGAAAGGTTG	139	(CAG)10	0.0	Predicted: *Setaria italica* peroxidase 16-like (LOC101774598)
PS-E-SSR304	F-ACATACACATACCGTCTCAT R-GGAACTGAAGTGTACCAATA	201	(TC)19	3e-137	Predicted: *Setaria italica* glucose-induced degradation protein 4 homolog
PS-E-SSR308	F-ACATCAAACTTGAAGAGAAA R-CATCATTTACAACAGGGAC	274	(TCGTCT)5	0.0	Predicted: *Setaria italica* aquaporin PIP1-1-like
PS-E-SSR309	F-ACATCAAACTTGAAGAGAAA R-CATCTTCTACAACAGGGAC	274	(TC)19	0.0	*Sorghum bicolor* hypothetical protein
PS-E-SSR315	F-AATTTCTCCTTGATCTTCTC R-AGGAGCACAAGACCG	121	(GA)16	0.0	*Panicum miliaceum* dehydrin mRNA
PS-E-SSR320	F-AATTTCTCCTTGATCTTCTC R-CATGAGAGTACGGTGGTA	220	(AG)19	2e-170	Predicted: *Setaria italica* dehydrin DHN1-like
PS-E-SSR328	F-ACAAAAAGTTTATTGGTTGA R-TAAACTGAGAAGAAGGATGA	268	(GA)19	3e-88	Predicted: *Setaria italica* non-specific lipid-transfer protein 3-like
PS-E-SSR357	F-AATGTTATTGCTTGAGAATC R-TATACAAGTCATAAGGGGTG	127	(GA)18	1e-147	Predicted: *Setaria italica* serine carboxypeptidase-like 19
PS-E-SSR362	F-AATTACATTTGGATCGTTAC R-TATGTTACCCTGTCCATTAC	218	(AAT)12	2e-64	Predicted: *Setaria italica* trihelix transcription factor ASR3-like
PS-E-SSR363	F-AATTACCCATCTGATCTTTT R-GTATAATGCGGTGCTATAAT	131	(GGGGAT)6	5e-155	Predicted: *Setaria italica* SPX domain-containing protein 5-like
PS-E-SSR364	F-AATTAGCCAAAGCAATTT R-CTACTCGCTCAGCTCCT	160	(GGC)12	4e-52	Predicted: *Setaria italica* mitochondrial import inner membrane translocase subunit TIM17-2-like
PS-E-SSR379	F-AATATTGACTACAACCGATG R-AGATTTGTACAGTTGTGGTG	193	(CCG)12	3e-118	Predicted: *Setaria italica* uncharacterized LOC101776482
PS-E-SSR383	F-AATCAATGTTTAATTCCGTA R-CAGTCGACGTAGTTGTTC	232	(TAG)12	6e-26	Predicted: *Setaria italica* uncharacterized LOC101776882
PS-E-SSR384	F-AATCAGCAATAAAGATAACG R-ACTTGCTTGGGATTAAAA	210	(GA)18	2e-149	Predicted: *Setaria italica* sucrose nonfermenting 4-like protein
PS-E-SSR404	F-AAGCAGCTGAGGATAAAG R-GTACACTCCGAACTCAAAG	213	(CGATT)7	0.0	Predicted: *Setaria italica* photosynthetic NDH subunit of subcomplex B 5
PS-E-SSR405	F-AAGCATATGAATACATCTTGA R-TCAAGGAGTACATATCCAAG	274	(AATCG)7	0.0	Predicted: *Setaria italica* probable 4-coumarate--CoA ligase 2

Table 2. Description of EST-SSR markers developed in this study (Continued)

Marker Name	Forward Primer	Product size	Repeat motif	E-value	Putative identities
PS-E-SSR411	F-AAGCTTCACTGCAAGG R-CTCCTTGTTCTTTTTCTCTT	269	(CT)17	0.0	Predicted: *Setaria italica* uncharacterized LOC101765121
PS-E-SSR415	F-AAGGAGCATAGAAACATACA R-TGGAAGAATGGAACATATAG	114	(AG)17	6e-99	Predicted: *Setaria italica* uncharacterized LOC101764180
PS-E-SSR421	F-AAGGCCTTATAGATGGTG R-TGGTCTTCTTCGTATCATAA	102	(AC)17	3e-91	Predicted: *Setaria italica* nucleosome assembly protein 1;2
PS-E-SSR423	F-AAGGGTGTACCAACTCC R-ATCCTCCTGTCGTCG	226	(GA)17	8e-168	Predicted: *Setaria italica* CRIB domain-containing protein RIC4-like transcript variant X2
PS-E-SSR427	F-AAGGTCAGCTACTTCCAG R-ACGAACTCCGCATCT	200	(CT)17	0.0	Predicted: *Setaria italica* plant UBX domain-containing protein 10
PS-E-SSR433	F-AAGTAAACAAACTGTTGGAC R-CCTCCTCGAGCTTTATC	239	(GCC)11	0.0	Predicted: *Setaria italica* uncharacterized LOC101758960
PS-E-SSR438	F-AAGTACGAGAACCTGATTG R-0AGTTTCTTACCCTTTTCAAC	239	(GCC)11	0.0	Predicted: *Setaria italica* ruBisCO large subunit-binding protein subunit alpha
PS-E-SSR448	F-AAGTATATGCTTTGCTCAAG R-GAATGTTCCATATGATGAAG	242	(CGC)11	0.0	Predicted: *Setaria italica* protein root UVB sensitive 6-like
PS-E-SSR458	F-AAGTTCGCCTCTGGA R-GAACGGTGGGCTCAG	264	(GGC)11	0.0	Predicted: *Setaria italica* PLASMODESMATA CALLOSE-BINDING PROTEIN 2-like

Table 3. Genetic parameters of 48 EST-SSR markers among 37 accessions of little millet

Marker Name	N_A[z]	H_E[y]	H_o[x]	PIC[w]	Marker Name	N_A	H_E	H_o	PIC
PS-E-SSR31	2	0.053	0.000	0.051	PS-E-SSR304	4	0.200	0.000	0.193
PS-E-SSR47	6	0.539	0.784	0.444	PS-E-SSR308	5	0.576	0.838	0.494
PS-E-SSR71	5	0.291	0.000	0.280	PS-E-SSR309	2	0.149	0.000	0.138
PS-E-SSR121	2	0.027	0.027	0.026	PS-E-SSR315	3	0.152	0.000	0.146
PS-E-SSR125	2	0.102	0.000	0.097	PS-E-SSR320	5	0.612	0.108	0.538
PS-E-SSR142	2	0.053	0.000	0.051	PS-E-SSR328	7	0.559	0.514	0.517
PS-E-SSR144	2	0.102	0.000	0.097	PS-E-SSR357	2	0.193	0.000	0.174
PS-E-SSR167	7	0.577	0.027	0.549	PS-E-SSR362	5	0.615	1.000	0.539
PS-E-SSR178	4	0.586	1.000	0.498	PS-E-SSR363	2	0.289	0.081	0.248
PS-E-SSR182	5	0.405	0.000	0.379	PS-E-SSR364	2	0.078	0.0811	0.075
PS-E-SSR185	2	0.193	0.000	0.174	PS-E-SSR379	4	0.586	1.000	0.498
PS-E-SSR195	4	0.244	0.000	0.232	PS-E-SSR383	3	0.283	0.000	0.263
PS-E-SSR199	2	0.149	0.000	0.138	PS-E-SSR384	3	0.324	0.000	0.302
PS-E-SSR201	3	0.239	0.0541	0.221	PS-E-SSR404	3	0.197	0.000	0.186
PS-E-SSR220	2	0.234	0.000	0.206	PS-E-SSR405	3	0.103	0.027	0.100
PS-E-SSR232	2	0.149	0.000	0.138	PS-E-SSR411	5	0.551	1.000	0.450
PS-E-SSR258	3	0.279	0.000	0.252	PS-E-SSR415	4	0.329	0.000	0.310
PS-E-SSR264	6	0.291	0.027	0.281	PS-E-SSR421	3	0.104	0.000	0.101
PS-E-SSR269	2	0.102	0.000	0.097	PS-E-SSR423	2	0.102	0.000	0.097
PS-E-SSR270	2	0.102	0.000	0.097	PS-E-SSR427	3	0.128	0.027	0.122
PS-E-SSR271	4	0.439	0.108	0.409	PS-E-SSR433	3	0.197	0.000	0.186
PS-E-SSR284	4	0.244	0.000	0.232	PS-E-SSR438	2	0.149	0.000	0.138
PS-E-SSR299	2	0.053	0.000	0.051	PS-E-SSR448	5	0.574	0.946	0.483
PS-E-SSR303	3	0.199	0.000	0.189	PS-E-SSR458	2	0.053	0.000	0.051
Mean						**3.3**	**0.266**	**0.123**	**0.240**

[z]N_A: number of alleles per locus, [y]H_E: expected heterozygosity, [x]H_o: observed heterozygosity, [w]PIC: polymorphic information content.

SSR142, PS-E-SSR144, PS-E-SSR199, PS-E-SSR232, PS-E-SSR269, PS-E-SSR270, PS-E-SSR299, PS-E-SSR309, PS-E-SSR315, PS-E-SSR364, PS-E-SSR405, PS-E-SSR421, PS-E-SSR423, PS-E-SSR427, PS-E-SSR438, PS-E-SSR458) were only informative with PIC values in a range of 0.026 to 0.146.

Cluster analysis

Neighbor-joining dendrogram based on genetic dissimilarity matrix data of EST-SSR alleles demonstrated the structure of the genetic diversity among little millet germplasm collections (Fig. 3). The germplasm was clustered into two main groups (Group-I and II). Group-I included 12 accessions and represented majority of accessions (11) from USDA (USA), including one accession (No. 7: IT261891) from ICRISAT (India). Group-II included 25 accessions, of which 21 accessions belonged to USDA (USA) and 4 accessions from ICRISAT (India) including number 3, 4, 5 and 6 (IT153625, IT153626, IT153627, IT153628 (Fig. 3, Table 1).

Possible association of EST-SSR markers

Sequence similarity search was conducted using blastn function of NCBI database (www.ncbi.nlm.nih.gov) for the annotation of newly developed markers and the majority of sequences carried high homology hits (Table 2). Among 48 marker sequences, 30 (62.5%) markers showed high similarity with various important genes in the database (E value <1E-157). The match percentage in the database was higher in case of *Setaria italica* (87.5%) as compare to other plant species (12.5%) (Fig. 4). Among the *Setaria italica* matched sequences, seven genes were annotated as uncharacterized loci, while others were linked to multiple predicted genes such as La-related protein 1B (PS-E-SSR31), flowering time

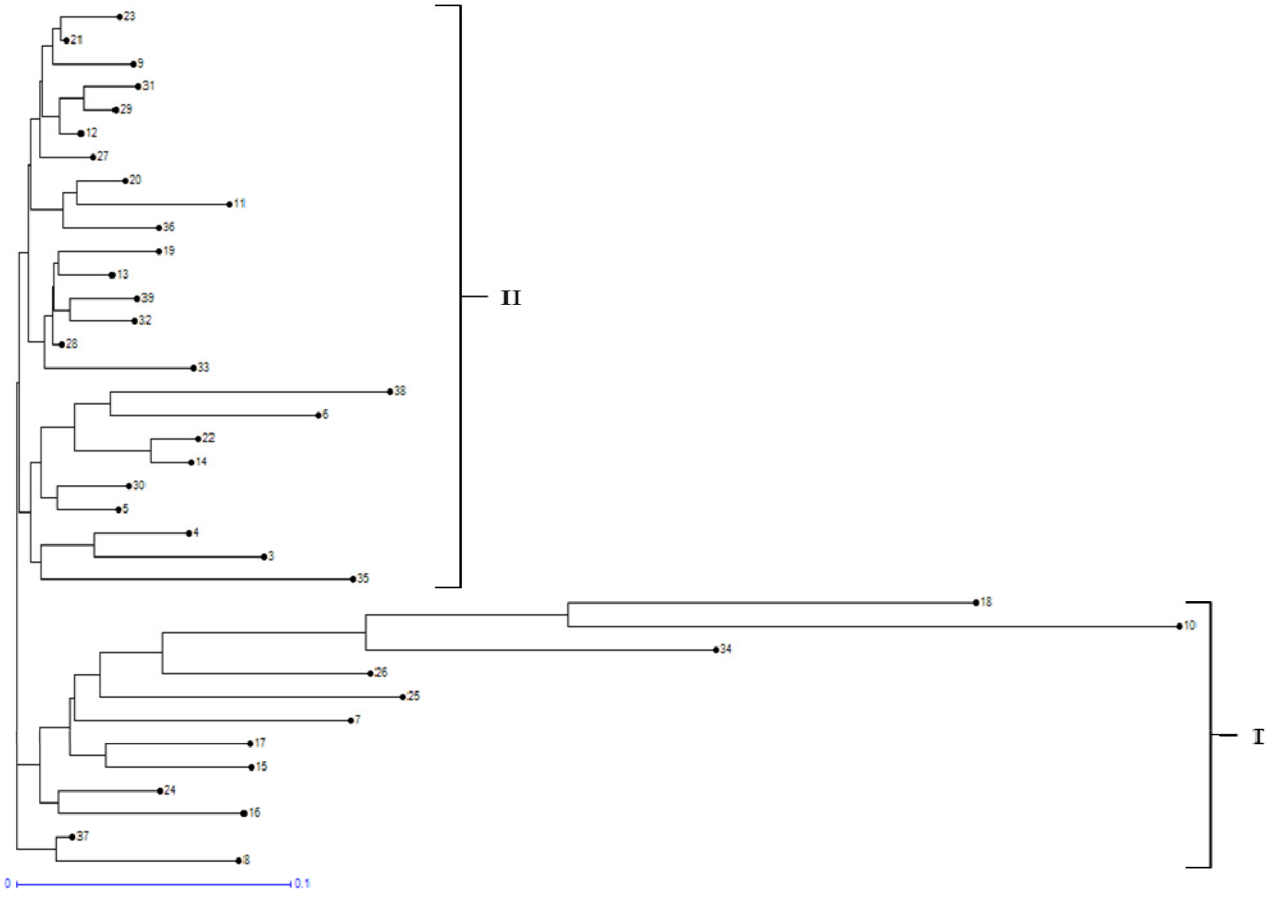

Fig. 3. Dendrogram representing the structure of genetic diversity observed among the 37 accessions of little millet by using 48 EST-SSR markers.

control protein FPA (PS-E-SSR71), gibberellin receptor GID1 (PS-E-SSR142), ethylene-responsive transcription factor RAP2-13-like (PS-E-SSR271) controlling different functions (Table 2). Similarly, four EST-SSRs (PS-E-SSR264, PS-E-SSR270, PS-E-SSR284 and PS-E-SSR404) expressed sequence similarity with light harvesting proteins. Three markers (PS-E-SSR178, PS-E-SSR199, and PS-E-SSR309) showed sequence similarity with *Sorghum bicolor* hypothetical proteins, whereas one EST-SSR (PS-E-SSR315) depicted similarity with a stress responsive *Panicum miliaceum* dehydrin mRNA. However, PS-E-SSR299 could not show any significant match in the nucleotide database.

Principal coordinate (PCoA) analysis

Principal coordinate analysis (PCoA) was performed based on EST-SSR profiles. Coordinates were calculated for two first axes with positive Eigen values. The first axis (Coord. 1) accounted for 22.38% and the second axis (Coord. 2) accounted for 9.09%. Among the 37 little millet accessions used in this experiment, except for the No. 8, 16, and 32 (IT261892, IT261900, IT284252) belonging to USDA (USA), most of accessions were biased toward one side (Fig. 5).

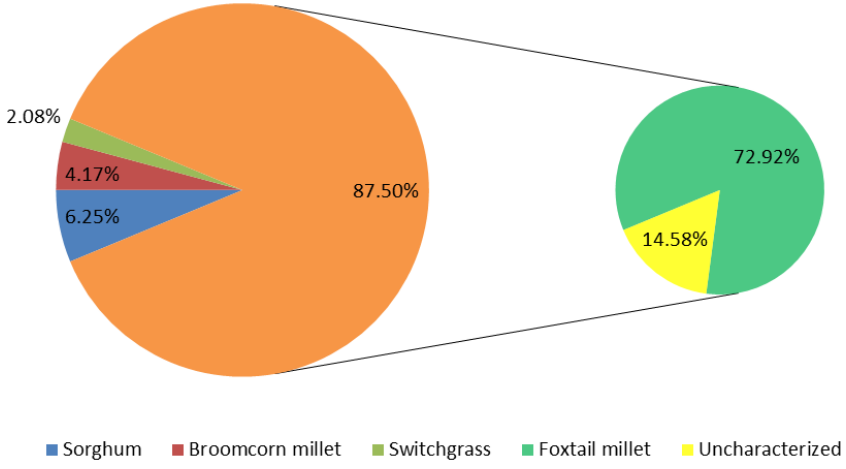

Fig. 4. Annotation based similarity distribution of EST-SSR sequences across species.

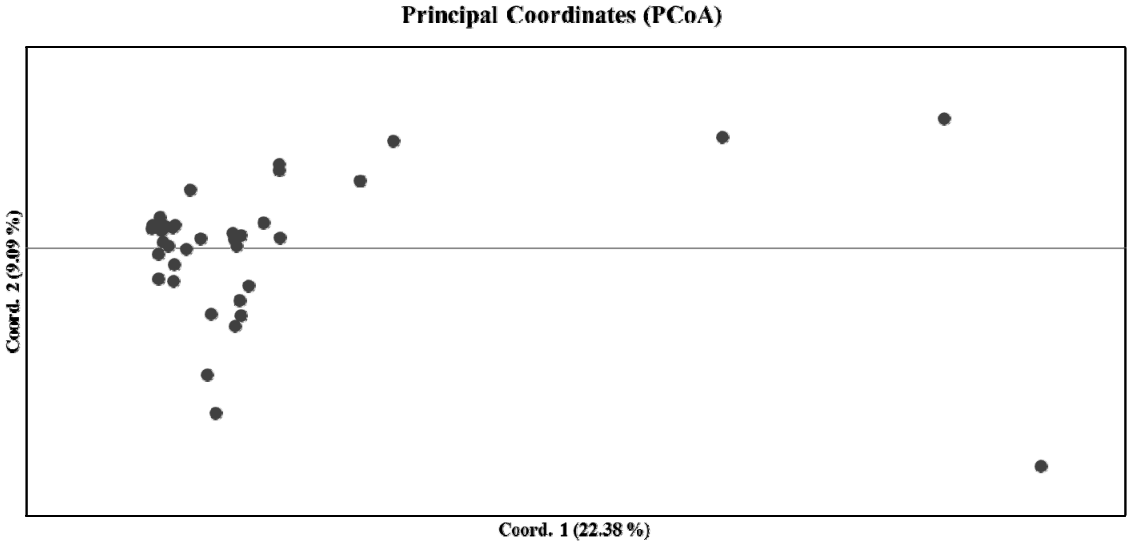

Fig. 5. Principal coordinates analysis (PCoA) of the first and second coordinates estimated based on the presence or absence of alleles using EST-SSR markers for 37 little millet varieties. The first two coordinates explained 22.38% and 9.09% of the total variation, respectively.

Discussion

EST-SSR markers are believed to be an important tool for genotyping to study genetic relationships and cross transferability in crop species due to its high level of conservation (Cordeiro *et al*., 2001). Because of limited knowledge about the genome of little millet, no progress has been made in marker development and to the best of our knowledge, no report has been published on EST-SSR markers in this crop. This scenario motivated us to develop EST-SSR markers in little millet by using the publically available genomic resources of switchgrass, which shares the same family and genus with little millet. Switchgrass has high percentage of cross-species transferability of molecular markers (Pandey *et al*., 2013) that supports cross species EST-SSRs as a useful resource for developing markers in minor crops. Previously, Wang *et al*. (2005) also succeeded to transfer SSR markers from major cereal crops to minor grass species

The average frequency (3.39%) of EST-SSRs in the selected transcribed region was comparable to barley (3.4%), wheat (3.2%), sorghum (3.6%) and rice (3.57%) (Kentety *et al*., 2002). It demonstrates that different motif frequency distributions occur in different species. Among 779 designed EST-SSRs, di-nucleotide repeat motifs were found in higher proportion (37%) followed by tri-nucleotides (31%). The higher ratio of dinucleotides in our results varied from the outcomes of Pandey *et al*. (2013) but comparable to that documented in lettuce (Rauscher and Simko, 2013). However, the abundance of tri-nucleotides in the coding regions of plant genomes can be explained by the fact that expansions or deletions in these regions can be endured for tri-nucleotides, which don't perturb the reading frames (Biswas *et al*., 2014).

The newly developed EST-SSR markers were polymorphic exhibiting 160 alleles with an average of 3.3 alleles per locus (Table 3) which is comparable to 2.4 and 2.5 alleles per locus described by Lin *et al*. (2012) and Jia *et al*. (2007) in foxtail millet. However, H_E mean value (0.266) is less than that was reported in proso millet (0.37) (Cho *et al*., 2010). The reason of variation might be due to artificial outcrossing or natural heterosis. A lower value of average PIC (0.240) was calculated as compared to earlier reports of Lin *et al*. (2012) and Ali *et al*. (2016) showing PIC values of 0.381 and 0.518 in foxtail, respectively. Low PIC value with EST-SSR markers observed in the present study indicates lower genetic diversity among the little millet accessions considered for analysis. Furthermore, various factors such as breeding behavior of species, size of the collection and primer locations in the genome can contribute to affect PIC values.

Moreover, efficiency of EST-SSR markers was assessed by genetic relatedness among the accessions based on their grouping pattern in the neighbor-joining dendrogram (Fig. 3). It revealed two distinct groups among the little millet accessions. The genotypes scattered into two groups irrespective of their geographical origin due to less polymorphism of EST-SSRs and narrow genetic base. It also indicates that the origin of all accessions is common as the whole germplasm collection was from India rather it was maintained at USDA or ICRISAT. A previous report from the MSSR (1999) researchers described that no genetic variation was observed during molecular analysis of eleven landraces of little millet. Hence it could be deduced that the little millet has narrow genetic base that needs to be studied in detail.

The annotation of the marker sequences revealed a diverse functional category of the genes and almost all were found to have homology to other plant genes (Table 2). For instance, PS-E-SSR31 has been annotated for La-related protein 1B which can be characterized with specific experiments in little millet. La-related protein showed its involvement in the stability and efficiency of mRNAs required for cell division, development, migration and cell death (Burrows *et al*., 2010). In addition, little millet is an arid area crop and tolerant to drought stress. Analysis of stress responsive gene *DHN1*, which showed sequence similarity with PS-E-SSR320, could open the secrets of stress tolerance in little millet. A previous study revealed that *DHN1* showed positive effect on plant growth under abiotic stress conditions (Beck *et al*., 2007). Many other genes such as flowering time control protein FPA, gibberellin receptor GID1, Photosystem I reaction center subunit VI, sucrose non-fermenting protein, photosynthetic NDH and plasmodesmata callose binding protein have shown their involvement in phenotypic and physiological functions. We also found a gene 'aquaporin PIP1' which plays its role in stem development and storage of sugar in *setaria viridis* (McGarghey *et al*., 2016). Our annotations provide the basis

and reference for designing strategies and experiments to explore the diverse nature and genetic variability of little millet.

As a result of Principle coordinate analysis (PCoA), among the 37 little millet accessions except some little millet belonging to USDA, most of accessions were distributed to one side (Fig. 5). It couldn't be divided into clusters at a clear population level. Little millet showed a distribution irrespective of geographical origins as the same result of dendrogram (Fig. 3). Also, Ali *et al*. (2016) study was twenty-two EST-SSR markers were used to analyze the genetic diversity and related genes of 621 fox millets in Korea landraces. PCoA explained variation of 13.88% and 10.99% by the first and second coordinates, respectively, but no clear cluster level clusters were found.

In summary, we developed first time EST-SSR markers for little millet utilizing cross genomic resources. Our findings from EST-SSR markers showed the presence of low genetic differentiation among the germplasm collections of little millet. The EST-SSRs developed here would serve as an important source for the assessment of genetic diversity in little millet and related species. However, the results indicated the need to introduce new germplasm to broaden the genetic base of the crop. It also reflected the need to develop more microsatellite and/or single nucleotide polymorphism (SNP) markers for effective implications in marker assisted breeding.

Acknowledgment

This study was carried out with the support of “Research Program for Agricultural Science & Technology Development (Project No. PJ010871022017)”, National Institute of Agricultural Sciences, Rural Development Administration, Republic of Korea.

References

Ali, A., Y.M. Choi, D.Y. Hyun, S. Lee, S. Oh, H.J. Park, Y.H. Cho and M.C. Lee. 2016. EST-SSR based genetic diversity and population structure among Korean landraces of foxtail millet (*Setaria italica*). Korean J. Plant Res. 29(3):322-330.

Ajithkumar, I.P. and Panneerselvam. 2014. ROS scavenging system, osmotic maintenance, pigment and growth status of panicum sumatrense roth under drought stress. Cell Bio-Chem. Biophys. 68:587-595.

Biswas, M.K., Q. Xu, C. Mayer and X. Deng. 2014. Genome wide characterization of short tandem repeat markers in sweet orange (*Citrus sinensis*). PloS One. 9:e104182. Doi: 10.1371/journal.pone.0104182.

Beck, E.H., S. Fettig, C. Knake, K. Hartig and T. Bhattarai. 2007. Specific and unspecific responses of plants to cold and drought stress. J. Biosci. 32:501-510.

Bhaskaran, J. and R. Panneerselvam. 2013. Accelerated reactive oxygen scavenging system and membrane integrity of two panicum species varying in salt tolerance. Cell BioChem. Biophys. 67:885-892.

Burrows, C., N. Abd Latip, S.J. Lam, L. Carpenter, K. Sawicka, G. Tzolovsky, H. Gabra, M. Bushell, D.M. Glover, A.E. Willis and S.P. Blagden. 2010. The RNA binding protein Larp1 regulates cell division, apoptosis and cell migration. Nucleic Acids Res. 38:5542-5553.

Casler, M.D. 2012. Switchgrass breeding, genetics and genomics. *In* Monti, A. (ed.), Switchgrass, Green Energy and Technology, Springer London, UK. pp. 29-53.

Cho, Y.I, J.W. Chung, G.A. Lee, K.H. Ma, A. Dixit, J.G. Gwag and Y.J. Park. 2010. Development and characterization of twenty-five new polymorphic microsatellite markers in proso millet (*Panicum miliaceum* L.). Genes Genom. 32: 267-273.

Cordeiro, G.M., R. Casu, C.L. McIntyre, J.M. Manners and R.J. Henry. 2001. Microsatellite markers from sugarcane (*Saccharum* spp) EST cross transferable to erianthus and sorghum. Plant Sci. 160:1115-1123.

De Wet, J.M.J., K.E. Prasada Rao and D.E. Brink. 1983. Systematics and domestication of *Panicum sumatrense* (Graminae). J. d'agriculture Tradit. Bot. appliquée. 30:159-168.

Da Maia, L.C., D.A. Palmieri, V.Q. de Souza, M.M. Kopp, F.I.F. de Carvalho and A.C. de Oliveira. 2008. SSR locator: tool for simple sequence repeat discovery integrated with primer design and PCR simulation. Int. J. Plant Genom. Doi: 10.1155/2008/412696.

Goron, T.L. and M.N. Raizada. 2015. Genetic diversity and genomic resources available for the small millet crops to accelerate a new green revolution improvement. Front. Plant Sci. Doi:10.3389/fpls.2015.00157.

Govindaraj, M., M. Vetriventhan and M. Srinivasan. 2015.

Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its Analytical perspectives. Genet. Res. Int. Doi.org/10.1155/2015/431487.

Gupta, S.K. Kumari, P.P. Sahu, S. Vidapu and M. Prasad. 2012. Sequence-based novel genomic microsatellite markers for robust genotyping purposes in foxtail millet [*Setaria italica* L. Beauv]. Plant Cell Rep. 31:323-337.

Hiremath, S.C., G.N.V. Patil and S.S. Salimath. 1990. Genome homology and origin of *Panicum sumatrense* (Gramineae). Cytologia 55:315-319. Doi:10.1508/cytologia.55.315.

Im, C.H., K.H. Kim, H.J. Je, A. Ali, M.K. Kim, W.K. Joung, S.D. Lee, H.Y. Shin and J.S. Ryu. 2014. Multiplex simple sequence repeat (SSR) markers discriminating *Pleurotus eryngii* cultivar. Korean J. Mycol. 42(2):159-164.

Jia, X.P., Y.S. Shi, Y.C. Song, G.Y. Wang, T.Y. Wang and Y. Li. 2007. Development of EST-SSR in foxtail millet (*Setaria italica*). Genet. Resour. Crop Ev. 54:233-236.

Jia, X.P., Z. Zhang, Y. Liu, C. Zhang, Y. Shi, Y. Song, T. Wang and Y. Li. 2009. Development and genetic mapping of SSR markers in foxtail millet [*Setaria italica* L. Beauv]. Theor. Appl. Genet. 118:821-829.

Kantety, R.V., M. La Rota, D.E. Matthews and M.E. Sorrells. 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. Plant Mol. Biol. 48:501-510. Doi: 10.1023/A:1014875206165.

Lin, H.S., G.I. Liao, C.Y. Chian, C.S. Kuoh and S.B. Chang. 2012. Genetic diversity in the foxtail millet (*Setaria italica*) germplasm as determined by agronomic traits and microsatellite markers. Aus. J. crop Sci. 6(2):342-349.

Liu, K. and M. Muse. 2005. PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128-2129.

McGaughey, S.A., H.L. Osborn, L. Chen, J.L. Pegler, S.D. Tyerman, R.T. Furbank, C.S. Byrt and C.P.L. Grof. 2016. Roles of aquaporins in *Setaria viridis* stem development and sugar storage. Front. Plant Sci. 7:1815. Doi:10.3389/fpls. 2016.01815.

Nicot, N., V. Chiquet, B. Gandon, L. Amilhat, F. Legeai, P Leroy, M. Bernar and P. Sourdille. 2004. Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). Theor. Appl. Genet. 109(4):800-805.

MSSR. 1999. Ninth annual report: 1998-99 M.S. Swaminathan Research Foundation, Chennai.

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 Res. 20:197-207.

Peakall, R.O.D. and Smouse, P.E. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6(1):288-295.

Perrier, X. and J.P. Jacquemoud-Collet. 2006. Darwin software, http://darwin.cirad.fr/darwin

Pradeep, S.R. and M. Guha. 2011. Effect of processing methods on the nutraceutical and antioxidant properties of little millet (*Panicum sumatrense*) extracts. Food Chem. 126:1643-1647.

Rauscher, G. and I. Simko. 2013. Development of genomic SSR markers for fingerprinting lettuce (*Lactuca sativa* L.) cultivars and mapping genes. BMC Plant Biol. 13:11. **Doi:** 10.1186/1471-2229-13-11.

Seo, D.H., K.M. Jung, S.J. Kim and K.M. Kim. 2013. Development of EST-SSR markers and analysis of genetic diversity using persimmon (*Diospyros kaki* Thunb) cultivars collecting from domestic. Korean J. Plant Res. 26(4):491-502 (in Korean).

Wang, Y.W., T.D. Samuels and Y.Q. Wu. 2011. Development of 1,030 Genomic SSR Markers in Switchgrass. Theor. Appl. Genet. 122(4):677-686.

Wang, M.L., N.A. Barkley, J.K. Yu, R.E. Dean, M.L. Newman, M.E. Sorrells and G.A. Pederson. 2005. Transfer of simple sequence repeat (SSR) markers from major cereal crops to minor grass species for germplasm characterization and evaluation. Plant Genet. Resour. 3(1):45-57.

(Received 22 March 2017 ; Revised 20 April 2017 ; Accepted 4 May 2017)