J. Plant Biochem. Biotechnol. (July–August 2017) 26(3):356–365
DOI 10.1007/s13562-017-0398-6

ORIGINAL ARTICLE

Development and validation of EST derived SSR markers with relevance to downy mildew (*Sclerospora graminicola* Sacc.) resistance in pearl millet [*Pennisetum glaucum* (L.) R. Br.]

Harshvardhan N. Zala[1] · Kalyani S. Kulkarni[1,2] · Tejas C. Bosamia[3] · Yogesh M. Shukla[4] · Sushil Kumar[1] · Ranbir S. Fougat[1] · Ankita Patel[1]

Received: 6 September 2016 / Accepted: 16 January 2017 / Published online: 30 January 2017

Abstract Molecular markers associated or linked with a particular trait are immensely essential for marker-assisted selection. Downy mildew is a severe disease in pearl millet causing huge production losses annually. There is a need to develop reliable functional markers associated with downy mildew resistance. In this study, expressed sequence tag-SSR (EST-SSRs) markers were developed from transcriptome data of downy mildew inoculated resistant and susceptible pearl millet genotypes. The high quality pearl millet reads were assembled by Trinity into 26,690 transcripts and 3586 (13.4%) transcripts harbored 4510 SSR motifs with the frequency of one SSR per 5.73 kb. The distribution of SSR motifs in different repeat type classes highlighted tri-nucleotide repeats as the highest (54.6% with CCG/CGG) followed by hexa-nucleotide (16.1% with CCGGCG/CCGGCG), penta-nucleotide (14.5% with CCGCG/CGCGG), tetra-nucleotides (8.0% with AGGG/CCCT) and di-nucleotide (6.8% with AG/CT) repeats. Total of 2419 EST-SSRs markers were developed and 230 markers based on their function in downy mildew-pearl millet interaction were validated in 12 pearl millet genotypes which are parents of mapping population. A total of 33 (20.62%) EST-SSR markers displayed polymorphism with a polymorphism information content (PIC) value ranging from 0.188 to 0.375 with an average of 0.354 and 2.2 alleles. The highest polymorphism was displayed by tri-nucleotide repeat EST-SSR markers followed by hexa-nucleotide, penta-nucleotide and di-nucleotide. The EST-SSR primers designed from eight transcripts coding for peroxidase 16-like, transcription factor myc7e, wound stress protein precursor, peroxidase 16-like, zinc finger ccch domain-containing, wrky transcription factor, beta purothionin were found to be more polymorphic than rest of the transcripts. The newly developed EST-SSR markers have added to the repository of molecular markers for pearl millet and can be employed for downy mildew resistance breeding.

Keywords Downy mildew · Pearl millet · EST-SSR · Disease resistance transcripts · PIC

Abbreviations

EST	Expressed sequence tag
GBS	Genotyping by sequencing
PIC	Polymorphism information content
SSR	Simple sequence repeat

Introduction

Downy mildew disease caused by the obligate oomycetes pathogen, *Sclerospora graminicola* Schott is a serious threat to pearl millet [*Pennisetum glaucum* (L.) R. Br.]

Harshvardhan N. Zala and Kalyani S. Kulkarni have contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s13562-017-0398-6) contains supplementary material, which is available to authorized users.

✉ Harshvardhan N. Zala
zala_harsh@yahoo.co.in

✉ Sushil Kumar
sushil254386@yahoo.com

[1] Department of Agricultural Biotechnology, Anand Agricultural University, Anand 388 110, India

[2] ICAR-Indian Institute of Rice Research, Rajendranagar, Hyderabad 500 030, India

[3] Department of Biotechnology, Junagadh Agriculture University, Junagadh 362 001, India

[4] Department of Biochemistry, Anand Agricultural University, Anand 388 110, India

production reported to cause losses of up to 40% in favorable environments. Pearl millet is a principal crop for food security in the arid and semi-arid regions of India and Africa and now-a-days gaining much more attention on the fact of its nutritive index (Kodkany et al. 2013; Kumar et al. 2016). It is one of the most widely cultivated drought and high temperature tolerant C_4 cereals for forage, grain and stover. Development of downy mildew resistant parental lines to develop durable hybrids through marker-assisted selection is a viable and feasible strategy of disease control. Molecular markers are the most important genetic resources for crop improvement.

Molecular markers linked with a particular trait like disease resistance can be employed to track that trait in the segregating population and be effectively utilized for marker-assisted selection (MAS). Even though, molecular marker studies in pearl millet date back to the early 1990s with the first molecular map of pearl millet reported by Liu et al. (1994), efforts are still underway for the development and application of molecular-markers in pearl millet to make it stand with other cereals like rice, sorghum, maize, wheat and barley (Sehgal et al. 2012; Punnuri et al. 2016). The development of downy mildew resistant hybrid, HHB-67 improved by RFLP marker-assisted selection is a landmark event in the history of pearl millet improvement which not only ensued the discovery of QTLs linked to the downy mildew disease resistance but also led to successful adoption of this variety in north western India (Hash et al. 2003). The availability of pearl millet genome sequence in the near future (http://ceg.icrisat.org/ipmgsc/) is expected to append and enhance the genetic resources. Nevertheless, efforts must continue in the direction of developing functional molecular markers for specific biotic or abiotic resistance/tolerance.

The genetic maps in pearl millet based on markers such as restriction fragment length polymorphism (RFLP) and amplified fragment length polymorphism (AFLP) with the simple sequence repeat (SSR) and diversity array technology (DArT)-based maps have proven useful in the identification of quantitative trait loci (QTLs) and breeding for drought tolerance, disease resistance, stover quality and grain mineral (Hash et al. 2003; Supriya et al. 2012; Kannan et al. 2014; Kumar et al. 2016). A high density linkage map and Genotyping by sequencing (GBS) markers has been recently developed by Punnuri et al. (2016) for tagging leaf spot resistance.

Of late, RNA-seq or transcriptome sequencing through next-generation sequencing platforms has been able to provide reliably large amount of information on the functions and expression of genes at a given time point (Devran et al. 2015; Varshney 2016). The SSRs markers developed from such data ply insight on functions and provides association with that particular trait in study and aptly known as EST-SSRs. The EST-SSRs have intrinsic advantage over genomic SSRs in terms of its origin from coding regions of the genome and transferability to other related species (Kalia et al. 2011; Fougat et al. 2014). The open accessibility of huge amount of data emanated from transcriptome sequencing in various crops concerning various stages or time points have greatly levered and favored the development of SSRs (Varshney et al. 2005a, b, 2016).

There are reports on development of EST-SSRs for drought tolerance (Rajaram et al. 2013), abiotic stress tolerance (Bertin et al. 2005; Senthilvel et al. 2008; Thudi et al. 2010; Sehgal et al. 2012) in pearl millet. The EST-SSRs associate with particular function and hence can be efficiently employed in QTL tagging and molecular breeding of a particular trait. In the present study, transcriptome data of downy mildew resistant and susceptible genotypes was employed to mine and develop functional EST-SSR derived SSRs markers on the basis of annotated functional genes involved in plant disease resistance.

Materials and methods

Plant materials and DNA extraction

A total of 12 downy mildew (DM) resistant/susceptible genotypes of pearl millet were used for the validation of newly designed EST-SSR primers (Supplementary Table I). These resistant and susceptible genotypes can be used as potential parental lines for development of mapping population(s) to identify DM resistance QTL/genes. Genotypes were obtained from the International Crop Research Institute for Semi-Arid Tropics (ICRISAT), Patancheru, Telangana and Main Pearl Millet Research Station, Junagadh Agricultural University, Jamnagar, Gujarat. Genomic DNA was extracted from fresh tissue of 10 days old plant by CTAB method described by Zidani et al. (2005). The quality and integrity of DNA was checked on 0.8% agarose gel and quantified using NanoDrop ND-1000 (NanoDrop products, DE, USA).

Screening of trinity assembly for EST-SSRs mining and primer designing

In the earlier study, we performed transcriptome analysis of downy mildew resistant genotype (P310-17) and susceptible genotype (7042S) of pearl millet upon inoculation and control. We have employed the transcripts emanated from assembling the pre-processed 1,295,196 high-quality (HQ) reads of four multiplexed pearl millet samples through Trinity assembler yielded 26,690 transcripts (Kulkarni et al. 2016). These assembled transcripts were used to

mining and development of EST-SSR markers in present study. The transcripts were screened for SSR motifs using Perl script, MIcroSAtellite identification tool (MISA) with criteria of unit size or minimum number of repeats (2-7, 3-5, 4-4, 5-3 and 6-3) and maximum number of bases interrupting two SSRs in a compound microsatellite = 100 (Thiel et al. 2003). The fasta sequences of transcripts harboring SSR motifs were extracted by in-house Perl script and subjected to design primers flanking the SSR motif using BatchPrimer3 (You et al. 2008). The parameters used for designing primers as follows primer length 18–23 bp, with optimum value 20 bp; Tm 57–62 °C, with optimum value 60 °C; GC content 30–70%, with the optimum value 55%; and product size range 120–300 bp optimum value 150 bp, Max 5′ self-complementarity 5.0 and Max 3′ self-complementarity 0. The EST-SSRs primers were synthesized from Sigma-Aldrich, USA.

Functional annotation of SSRs harboring transcripts

The assembled transcripts harboring SSR motifs were functionally annotated using Blast2GO, online tool for functional annotation of (novel) sequences (Conesa et al. 2005). Blastx of contigs was carried out with non-redundant (nr) database using e value cut-off of 1E-6. The sequences showing blast hits were mapped for retrieving the gene ontology (GO) terms associated to the hits. Annotation was carried out on the mapped sequences for selecting GO terms and InterPro annotation was executed to retrieve domain/motif information in a sequence-wise manner. The GOSlim terms, cut-down versions of the GO ontologies containing a subset of the terms in the whole GO for molecular function, biological process, and cellular component categories associated with the best BLASTX hit of the *Arabidopsis thaliana* protein were assigned to the corresponding pearl millet transcripts (http://www.geneontology.org/GO.slims.shtml).

Polymorphism survey of newly designed EST-SSR markers

On the basis of functional annotation of the transcripts harboring SSR motifs, transcripts possessing functional relevance to downy mildew resistance were selected for validation on a panel of 12 genotypes. The PCR mixture (10 µL) contained 5 µL of 2X PCR master mix (Fermentas), 1 µL of genomic DNA (50 ng), 0.8 µL primers (10p moles of each forward and reverse primers), and 3.2 µL of nuclease free water. A touchdown PCR thermal profile was performed for 94 °C for 5 min of initial denaturation, followed by first five cycles of 94 °C for 30 s, 65 °C to 60 °C for 30 s and 72 °C for 1 min, with 1 °C decrement in annealing temperature per cycle, then 30 cycles of 94 °C for 30 s with constant annealing temperature of 60 °C for 30 s and 72 °C for 1 min followed by a final extension for 7 min at 72 °C. Amplified PCR product were separated on 3% Metaphor agarose gel (0.5 µg/ml Et Br) along with 50 bp DNA ladder (Fermentas, USA) in 1X TBE buffer at constant power 120 V for about 2.5–3 h. The gels were visualized and documented by gel documentation system (Bio-Rad, USA). The amplified bands were scored as presence (1) or absence (0). Polymorphism information content (PIC) was calculated using PowerMarker V3.25 (Liu and Muse 2005).

Results and discussion

Identification, type and frequency of EST-SSRs

The selection of assembled transcripts is critical for mining and development of EST-SSRs from RNA-seq data. In the present study, *de novo* assembled transcripts emanated by Trinity programme were selected on the basis of the largest N50 value (1219 bp), the maximum transcript length (1.74 Kb) and average contig length of 969.02 bp, was employed for development and mining SSRs (Kulkarni et al. 2016). Higher length of transcripts provides better opportunity in terms of longest open reading frame for mining SSRs and for effectively picking of primers from flanking sequence of SSRs (Bosamia et al. 2015). Out of the 26,690 transcripts, 3586 (13.4%) transcripts harbored SSR motifs. A total of 4510 SSRs were identified and 695 transcripts harbored more than one SSR (Table 1). The SSRs have been found in a range of 6–11% of the total EST sequences from several cereal species viz., barley, maize, rye, rice, sorghum and wheat (Varshney et al. 2002, 2009). In the present study, we found comparatively higher percentage of SSRs (13.4%). The frequency of SSR was one per 5.73 kb which is comparatively higher than the previous studies in pearl millet (Senthilvel et al. 2008) and corroborates with the reports stating frequency of SSRs every 4–8 kb in cereals (Kantety et al. 2002). The

Table 1 Statistics of EST-SSRs identified in pearl millet transcripts assembled by Trinity

Features	Contigs
Total number of transcripts examined	26,690
Total size of examined sequences (bp)	25,863,011
Total number of identified SSRs	4510
Number of SSR containing sequences	3586 (13.4%)
Number of sequences containing more than one SSR	695
Number of SSRs present in compound formation	336
Average frequency of SSRs	1/5.73 kb

Table 2 Distribution and frequencies of EST-SSRs in pearl millet

Motif length	Repeats number										
	3	4	5	6	7	8	9	10	>10	Total	%
Di-nucleotide	–	–	–	–	133	70	54	19	31	307	6.8
Tri-nucleotide	–	–	1490	670	262	33	5	–	3	2463	54.6
Tetra-nucleotide	–	253	87	15	2	1	–	–	1	359	8.0
Penta-nucleotide	505	122	24	1	1	–	–	–	1	654	14.5
Hexa-nucleotide	611	102	6	–	3	2	3	–	–	727	16.1
Total	1116	477	1607	686	401	106	62	19	36	4510	–

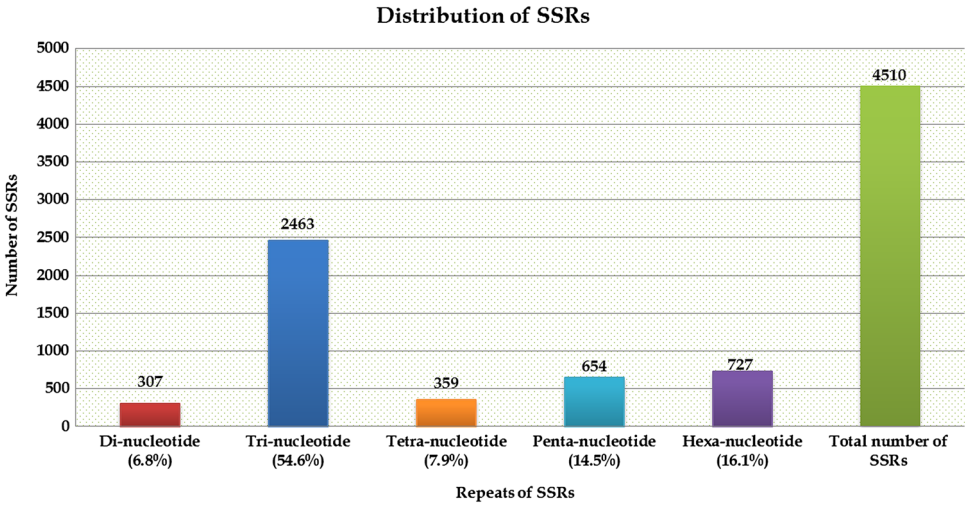

Fig. 1 Distribution of EST-SSR repeat motifs

frequency of SSR depends on amount of data and the parameters employed for mining SSRs (Varshney et al. 2007; Ding et al. 2015). Higher frequency of SSRs in the present study can be attributed to the RNA-seq generated transcripts.

The EST-SSRs can be classified on the basis of their size, type of repeat unit, motif length (Bosamia et al. 2015). The distribution of SSR motifs in different repeat type classes highlighted tri-nucleotide repeats as the highest (54.6%) followed by hexa-nucleotide (16.1%), penta-nucleotide (14.5%), tetra-nucleotides (8.0%) and di-nucleotide (6.8%) repeats (Table 2; Fig. 1). However, the results contrasted with the earlier studies of Senthilvel et al. (2008) in pearl millet where di-nucleotide repeat was the most common repeat followed by tri-, penta-, tetra- and hexa-nucleotide repeats. Earlier studies in cereals also demonstrated tri-nucleotides as the most abundant type of repeat class (Varshney et al. 2002, 2005a, b). Tri-nucleotide repeats are generally abundant in SSRs (Parekh et al. 2016). This abundance is dependent on factors like SSR search criteria, size of the dataset, database mining tools and the EST sequence redundancy (Varshney et al. 2005a, b; wang et al. 2014). The prominence of tri-nucleotide repeats in the pearl millet transcriptome also corroborated with the highest percentage of tri-nucleotide repeats (57.9%) in chickpea transcriptome sequenced through 454 pyrosequencing (Garg et al. 2011) and with studies of (Kantety et al. 2002; Varshney et al. 2002; Thiel et al. 2003; Gupta and Varshney 2000) in sorghum, barely and bread wheat. The tri-nucleotide repeat motifs were followed by hexa-nucleotide repeat motifs contributing 16.1% of the total SSR motifs. It has also been observed in many species that the transcripts or exons have more tri-nucleotide SSRs than other repeats (Bosamia et al. 2015; Kumar et al. 2016; Cardle et al. 2002; Morgante et al. 2002; Li et al. 2004). Nevertheless, hexa-nucleotide repeat has not yet been reported as the second highest repeat motif in cereals, however, such a presence could be explained on the basis of findings that hexa-nucleotide repeats along with tri-nucleotide repeat do not perturb reading frames and tolerate expansions or deletions in coding regions (Asp et al. 2007; Metzgar et al. 2000; Morgante et al. 2002).

The di-nucleotide repeat class AG/CT (53.7%; Table 3) was the most abundant in pearl millet transcriptome as was evident in sorghum, barley, wheat and in tall fescue grass (Gupta and Varshney 2000; Thiel et al. 2003; Saha et al. 2004). The presence of AG/CT in the highest percentage was also supported by the studies of Kantety et al. (2002), Varshney et al. (2002) and

Table 3 Frequencies of different repeat motifs of di- and tri-nucleotide repeats in pearl millet EST-SSRs

Repeat motif	Repeats number								
	5	6	7	8	9	10	>10	Total	%
AC/GT	–	–	30	15	13	3	10	71	23.1
AG/CT	–	–	65	39	30	13	18	165	53.7
AT/AT	–	–	14	9	8	3	3	37	12.1
CG/CG	–	–	24	7	3	–	–	34	11.1
Total	–	–	133	70	54	19	31	307	–
AAC/GTT	16	13	7	1	2	–	1	40	1.6
AAG/CTT	61	26	11	2	–	–	1	101	4.1
AAT/ATT	12	4	5	1	–	–	–	22	0.9
ACC/GGT	110	42	23	2	–	–	–	177	7.2
ACG/CGT	109	34	13	2	–	–	1	159	6.5
ACT/AGT	13	5	–	–	–	–	–	18	0.7
AGC/CTG	215	140	78	10	–	–	–	443	18.0
AGG/CCT	219	71	27	7	–	–	–	324	13.2
ATC/ATG	30	29	8	5	–	–	–	72	2.9
CCG/CGG	705	306	90	3	3	–	–	1107	44.9
Total	1490	670	262	33	5	-	3	2463	–

Jayashree et al. (2006) where SSRs were mined from the nucleotide databases. The studies of Yu et al. (2004) and Jayashree et al. (2006) also indicated that AG and CCG repeats were the most abundant class of repeats in cereals. Among the tri-nucleotides, CCG/CGG class was the most abundant (44.9%) motif as was observed in closely related species viz., sorghum and foxtail millet (Jia et al. 2007). It was also in conformance with most other studies (Eujayl et al. 2002; Gupta and Varshney 2000; Kantety et al. 2002; Saha et al. 2004; Varshney et al. 2002). Moreover, the abundance of CCG repeat is found to be specific feature of monocots and could be attributed to the increased GC content or high frequencies of amino acids with GC nucleotide codons (Morgante et al. 2002). In case of the tetra-, penta- and hexa-nucleotide repeats, AGGG/CCCT (8.6%), CCGCG/CGCGG (9.5%) and CCGGCG/CCGGCG (8.9%) were the most abundant repeat classes respectively (Supplementary Table S2). As with the case of tri-nucleotides, the dominance of GC repeats in penta- and hexa-nucleotide could be attributed to the GC content in monocots.

The SSR motifs were further classified into class I (≥20 bp) and class II (≥10 but <20 bp) repeats on the basis of the motif length. The class II repeats (3786) were present in more number than class I repeats (724) (Table 4). Similar observations were recorded by Rajaram et al. (2013) in pearl millet drought transcriptome. The average frequency (Kb/SSR) was found to be the highest for di-nucleotide motif followed by tetra-nucleotide and penta-nucleotide repeats. The least average frequency of tri-nucleotide repeat motif colligated with its highest number in the pearl millet transcripts.

Functional annotation of transcripts harboring EST-SSRs

Functional annotation of transcripts harboring SSRs was carried with respect to identification of the function and significance in disease resistance so as to correlate function with the molecular markers (Supplementary Table 3). Therefore, functional annotation constituted the prerequisite for selection of SSR primer designing and validation with respect to downy mildew resistance.

Out of the total 3586 transcripts subjected to BlastX with non-redundant database, 2566 transcripts were annotated (Supplementary Figure S1). The annotated transcripts showed the maximum homology was seen with *Setaria italica* (2510) followed by *Zea mays* (184) and *Sorghum bicolor* (167) (Supplementary Figure S2). Pearl millet and foxtail millet belong to the same millet family and share common ancestry which could have allotted pearl millet top blast hits to foxtail millet (Devos et al. 2000).

The functionally annotated SSR harbored transcripts were categorized into the cellular components, molecular functions and biological processes (Fig. 2). In cellular component ontology, the maximum number of transcripts were associated with plastid, GO:0009536 (450) followed by mitochondria, GO:0005739 (275), plasma membrane, GO:0005886 (160) and cytosol, GO:0005829 (100). The maximum number of transcripts in molecular function were attributed to nucleotide binding, GO:0000166 (501)

Table 4 Classification of EST-SSR according to the length of motif

Motif length	Class I (≥20 bp)	Class II (≥10 but <20 bp)	Total number of SSRs	Avg. frequency (Kb/SSR)
Di-nucleotide	50	257	307	84.24
Tri-nucleotide	303	2160	2463	10.50
Tetra-nucleotide	106	253	359	72.04
Penta-nucleotide	149	505	654	39.55
Hexa-nucleotide	116	611	727	35.57
Total	724 (16.05%)	3786 (83.94%)	4510	5.73

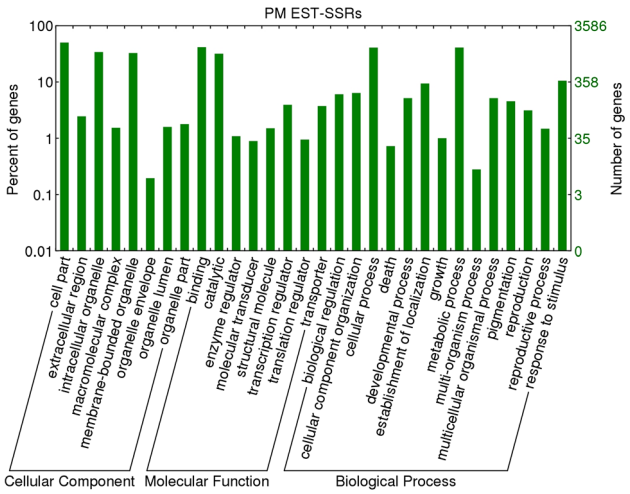

Fig. 2 Gene ontology distribution of EST-SSR harboring pearl millet transcripts into cellular component, molecular function and biological process

followed by protein binding, GO: (380), DNA binding, GO:0003677 (300) and kinase activity, GO:0016301 (200). In biological function, the maximum number of transcripts was represented by transport, GO:0006810 (337) followed by transcription, DNA-templated GO:0009056 (289), cellular protein modification, GO:0006464 (287) and response to stress, GO:0016043 (285).

Validation of EST-SSRs and polymorphism

The SSR markers were selected on the basis of functionally annotated transcript sequences harboring SSR motifs associated with downy mildew resistance genes as previously reported in pearl millet-downy mildew interaction and similar plant-pathogen interactions (Kini et al. 2000; Shivakumar et al. 2000; Geetha and Shetty, 2002; Shivakumar et al. 2003; Chandrashekhara et al. 2010). Thus, EST-SSR primers were selected for disease resistance genes, pathogenesis related proteins, transcription factors, genes involved in disease resistance. Out of the total 3586 transcripts subjected for primer designing, 2419 successful primer pairs were designed from 2321 transcripts with stringent criteria, out of which 230 primers pairs were selected on the basis of transcript function in

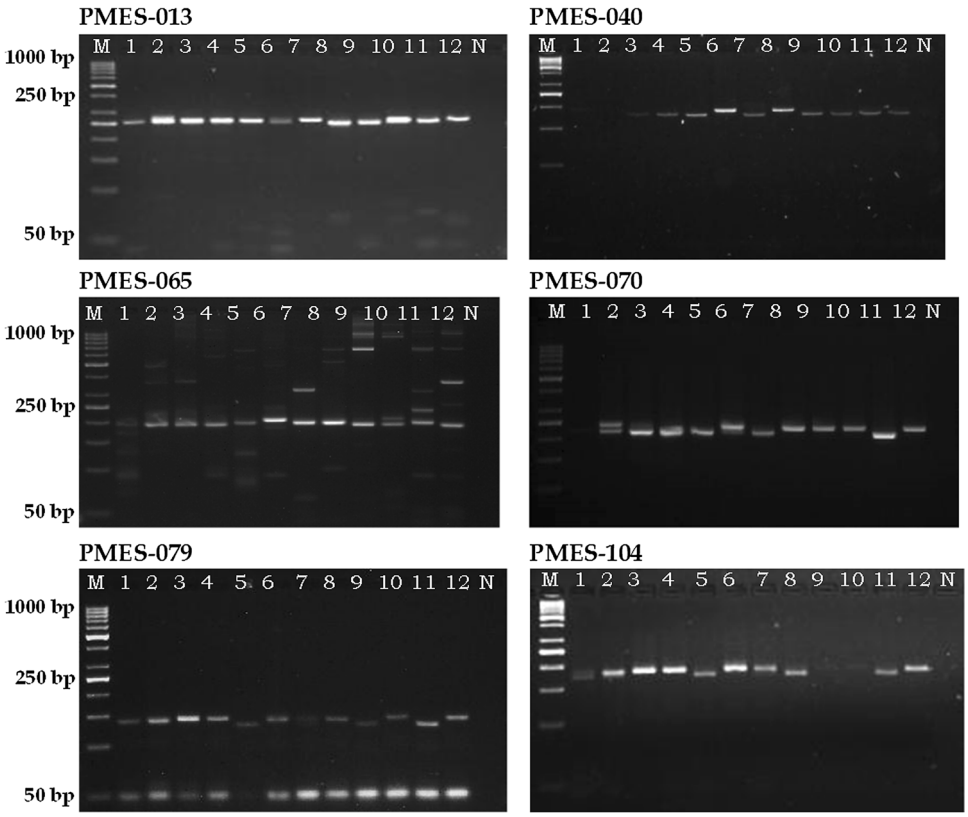

Fig. 3 Electrophoretic profile of polymorphic EST-SSR markers run on 3% metaphor agarose

Table 5 Details of 33 EST-SSR primers showing polymorphism among 12 pearl millet genotypes

Primer	SSR motif	Function	No. of alleles	Product size (bp)	PIC value
PMES-001	(AGCCGG)3	Serine arginine-rich splicing factor rs2z33-like	2	105–121	0.375
PMES-002	(CCGCCT)3	Mitogen-activated protein kinase 2	2	113–127	0.344
PMES-005	(AGC)5	Nac domain-containing protein 21 22-like	2	115–138	0.375
PMES-013	(GCGCCT)4	Ankyrin repeat domain-containing protein 2-like	2	250–285	0.375
PMES-015	(CCCGT)4	Basic leucine zipper 9-like isoform ×4	2	238–255	0.375
PMES-016	(ATGGAC)3	Peroxidase 16-like	3	162–182	0.346
PMES-040	(TTGGT)3	Mlo protein homolog 1-like	2	185–197	0.375
PMES-050	(CAG)5	Probable serine threonine-protein kinase wnk7-like	3	148–172	0.346
PMES-052	(GAAGAG)6	Probable serine threonine-protein kinase wnk7-like	2	228–263	0.305
PMES-053	(CAG)3	Transcription factor myc7e	3	143–154	0.346
PMES-055	(TCGC)4	Cinnamyl alcohol dehydrogenase 2	2	185–212	0.375
PMES-061	(CTG)6	Disease resistance protein rpm1	2	224–238	0.375
PMES-065	(CGG)5	Wound stress protein precursor	3	179–223	0.346
PMES-070	(CGA)5	Protein kinase superfamily protein	2	222–248	0.344
PMES-072	(CCA)6	Protein kinase superfamily protein	2	127–146	0.375
PMES-075	(CCG)5	c2h2 zn finger protein	2	142–171	0.344
PMES-079	(GTGGCT)6	Cadmium zinc-transporting atpase hma2-like	2	136–148	0.375
PMES-082	(CCA)5	Peroxidase 16-like	3	151–159	0.346
PMES-083	(CAC)6	Chitinase 6-like	2	125–149	0.375
PMES-086	(GGA)5	Zinc finger ccch domain-containing protein 11-like	3	208–230	0.346
PMES-087	(GCGAG)3	e3 ubiquitin-protein ligase rglg2-like	2	210–226	0.375
PMES-101	(TCG)5	Map kinase family protein	2	183–199	0.375
PMES-104	(CTT)6	Probable lrr receptor-like serine threonine-protein kinase mrh1-like isoform x1	2	172–189	0.344
PMES-108	(TCGAA)3	Ras-related protein rab-2-a	2	195–217	0.188
PMES-109	(CCGAT)5	Ras-related protein rab-2-a	2	314–324	0.375
PMES-110	(GCT)6	Wrky transcription factor	3	101–119	0.346
PMES-125	(AT)10	4-coumarate–ligase-like 1-like	2	173–187	0.313
PMES-158	(CGG)5	Aldehyde dehydrogenase family 2 member mitochondrial-like	2	139–152	0.313
PMES-159	(GA)8	Beta purothionin	3	152–165	0.3457
PMES-160	(GCT)5	Stress responsive protein	2	149–163	0.375
PMES-161	(GCT)5	Stress responsive protein	2	148–158	0.375
PMES-224	(GAG)6	Phenylalanine ammonia-lyase	2	153–167	0.375
PMES-226	(CGC)5	Cinnamoyl-reductase 1-like	2	173–184	0.375

plant-pathogen interaction. (Supplementary Table S4) and rest of the 2586 EST-SSRs markers have been listed in Supplementary Table S5. The newly developed EST-markers were prefixed PMES1-PMES230 and were validated on twelve parental genotypes for amplification. Out of the 230 EST-SSRs markers, 160 displayed desired amplification of defined product size referring to that of the designing parameters and 33 markers (20.62%) displayed polymorphism (Fig. 3).

The PIC of newly developed EST-SSRs was found in the range of 0.188 to 0.375 with an average value of 0.354 (Table 5). A total of 69 polymorphic loci with 2.2 average number of alleles were detected across the surveyed panel of twelve genotypes Eight primer pairs (PMES-016, PMES-050, PMES-053, PMES-061, PMES-065, PMES-082, PMES-086, PMES-110, PMES-158 and PMES-159) produced three alleles and rest of the 25 polymorphic primer pairs displayed two alleles across the panel of twelve genotypes. The PIC values of markers provided an estimate of their discriminating power in a set of accessions by taking not only the number of alleles but also the relative frequencies of each allele (Smith et al. 2000). Although the PIC values were slightly less than the polymorphism percentage of the pearl millet SSR markers reported in earlier studies, it has been well documented that EST-derived SSRs are less polymorphic than those derived from

genomic libraries (Senthilvel et al. 2008; Rajaram et al. 2013). The previous studies using EST-SSR markers for genetic diversity analysis in other crops exhibited PIC value of 0.44 in bread wheat, 0.45 in barley (Gupta and Varshney 2000; Thiel et al. 2003). The average PIC value was lower compared to the PIC values of genomic SSR (0.4–0.9) markers in pearl millet, sorghum, foxtail millet (Caniato et al. 2007). The EST-SSR markers are generally less polymorphic and provide less alleles and lower PIC values as compared to the genomic SSR (Varshney et al. 2009). In the present study, tri-nucleotide repeat markers were more polymorphic than the di-nucleotide, tetra-nucleotide and penta-nucleotide repeat-based markers, as observed previously in pearl millet (Gerard et al. 2006; Senthilvel et al. 2008; Rajaram et al. 2013).

On the basis of the functions assigned in Blast2GO annotation, eight transcripts harboring the SSR motifs and coding for peroxidase 16-like, transcription factor myc7e, wound stress protein precursor, peroxidase 16-like, zinc finger ccch domain-containing protein 11-like, wrky transcription factor, beta purothionin were found to be more polymorphic than rest of the SSR markers. Functional or gene-based SSR and SNP primers underlying drought tolerance have been reported in pearl millet (Rajaram et al. 2013; Sehgal et al. 2012) and in this study, we preliminary reported functional markers associated with downy mildew resistance. The polymorphic EST-SSR markers could be further employed in validation of mapping population segregating for downy mildew resistance in pearl millet. Such information can be further deployed for QTL mapping for disease resistance.

With reference to the economic losses caused by downy mildew disease in pearl millet and the fact that hitherto little information on gene/EST-SSRs, the primers developed in the present study have added to the existing repertoire of data for pearl millet. The polymorphic EST-SSR markers out of total of 2419 EST-SSRs developed in present study can be used in mapping population segregating for downy mildew resistance for advancing marker-assisted breeding for downy mildew resistance in pearl millet and for association mapping, genetic diversity assessment, genome mapping and saturation of pearl millet genetic map. The EST-SSRs markers can be employed for preparing a "transcript map" and can also be utilized for transferability in other minor millets for which less genetic resources are available.

Acknowledgement Authors are thankful to Rajan Sharma, Principal Scientist, ICRISAT, Patancheru and K. D. Mundra, Research Scientist, Main Pearl Millet Research Station, Jamnagar for kindly providing pearl millet seeds. HNZ and KSK express gratitude to Rajiv Gandhi National Fellowship, Government of India and INSPIRE, DST respectively for furnishing fellowship for the doctoral programme.

Author's contribution YMS, HNZ conceived and planned the study. HNZ, KSK, TCB, SK analyzed, interpreted the data and drafted manuscript. AP provided technical help. YMS, SK, RSF critically reviewed and edited the manuscript.

Compliance with ethical standards

Conflict of interest We declare that we have no conflict of interest.

References

Asp T, Frei UK, Didion T, Nielsen KK, Lübberstedt T (2007) Frequency type and distribution of EST-SSRs from three genotypes of *Lolium perenne* and their conservation across orthologous sequences of *Festuca arundinacea*, *Brachypodium distachyon* and *Oryza sativa*. BMC Plant Biol 127:36

Bertin I, Zhu JH, Gale MD (2005) SSCP-SNP in pearl millet-a new marker system for comparative genetics. Theor Appl Genet 110:1467–1472

Bosamia TC, Mishra GP, Thankappan R, Dobaria JR (2015) Novel and stress relevant EST derived SSR markers developed and validated in peanut. PLoS ONE. doi:10.1371/journal.pone.0129127

Caniato FF, Guimaraes CT, Schaffert RE, Alves VM, Kochian LV, Borem A, Klein PE, Magalhaes JV (2007) Genetic diversity for aluminum tolerance in sorghum. Theor Appl Genet 114:863–876

Cardle L, Ramsay L, Milbourne D, Macaulay M, Marshall D, Waugh R (2002) Computational and experimental characterization of physically clustered simple sequence repeats in plants. Genetics 156:847–854

Chandrashekhara S, Niranjan S, Deepaka G, Manjunatha H, Shettya S (2010) Thionins (PR protein-13) mediate pearl millet downy mildew disease resistance. Arch Phytopathol Plant Prot 43:1356–1366

Conesa A, Gotz S, Miguel J, Terol J, Talón M, Robles M (2005) Blast2GO: a universal tool for annotation visualization and analysis in functional genomics research. Bioinformatics 21(18):3674–3676

Devos KM, Pittaway TS, Reynolds A, Gale MD (2000) Comparative mapping reveals a complex relationship between the pearl millet genome and those of foxtail millet and rice. Theor Appl Genet 100:190–198

Devran Z, Kahveci E, Özkaynak E, Studholme DJ, Tör M (2015) Development of molecular markers tightly linked to *Pvr4* gene in pepper using next-generation sequencing. Mol Breed 35:101

Ding Q, Li J, Wang F, Zhang Y, Li H, Zhang J, Gao J (2015) Characterization and development of EST-SSRs by deep transcriptome sequencing in Chinese cabbage (*Brassica rapa* L. ssp. pekinensis). Int J Genomics 2015:473028

Eujayl I, Sorrells ME, Baum M, Wolters P, Powell W (2002) Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor Appl Genet 104:399–407

Fougat R, Joshi C, Kulkarni K, Kumar S, Patel A, Sakure A, Mistry J (2014) Rapid development of microsatellite markers for *Plantago ovata* Forsk.: using next generation sequencing and their cross-species transferability. Agriculture 4:199

Garg R, Patel RK, Jhanwar S, Priya P, Bhattacharjee A, Yadav G, Bhatia S, Chattopadhyay D, Tyagi AK, Jain M (2011) Gene discovery and tissue-specific transcriptome analysis in chickpea with massively parallel pyrosequencing and web resource development. Plant Physiol 156:1661–1678

Geetha HM, Shetty HS (2002) Expression of oxidative burst in cultured cells of pearl millet cultivars against *Sclerospora*

graminicola inoculation and elicitor treatment. Plant Sci 163(3):653–660

Gerard PR, Fernandez JF, Bertolino P, Dufour J, Raquin C, Frascaria F (2006) New insights in the recognition of the European ash species *Fraxinus excelsior* L. and *Fraxinus augustifolia* Vahl. as useful tools for forest management. Ann For Sci 63:733–738

Gupta PK, Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica 113:163–165

Hash CT, Yadav RS, Sharma A, Bidinger FR, Devos KM, Gale MD, Howarth CJ, Chandra S, Cavan GP, Serraj R, Kumar PS, Breese WA, Witcombe JA (2003) Pearl millet molecular marker research. PSP Ann Rep 35:23–56

Jayashree B, Punna R, Prasad P, Bantte K, Hash CT, Chandra S, Hoisington DA, Varshney RK (2006) A database of simple sequence repeats from cereal and legume expressed sequence tags mined in silico: survey and evaluation. In Silico Biol 6(6):607–620

Jia XP, Yun-Su S, Yan-Cun S, Guo-Yin Y, Tian-Yu W, Li Y (2007) Development of EST-SSR in foxtail millet (*Setaria italica*). Genet Res Crop Evol 54(2):233–236

Kalia RK, Rai MK, Kalia S, Singh R, Dhawan AK (2011) Microsatellite markers: an overview of the recent progress in plants. Euphytica 177(3):309–334

Kannan B, Senthilvel S, Gajaraj A, Raj B, Chandra S, Prabhu A, Dhanapal D, Hash CT (2014) Association analysis of SSR markers with phenology grain and stover-yield related traits in pearl millet (*Pennisetum glaucum* (L.) R. Br.). Sci World J 2014:14

Kantety RV, Mauricio LA, Matthews DE, Sorrells ME (2002) Data mining for simple sequence repeats in expressed sequence tags from barley maize rice sorghum and wheat. Plant Mol Biol 48:501–510

Kini KR, Vasanthi NS, Shetty HS (2000) Induction of β-1, 3-glucanase in seedling of pearl millet in response to infection by *Sclerospora graminicola*. Eur J Plant Pathol 106:267–274

Kodkany BS, Bellad RM, Mahantshetti NS, Westcott JE, Krebs NF, Kemp JF, Hambidge KM (2013) Biofortification of pearl millet with iron and zinc in a randomized controlled trial increases absorption of these minerals above physiologic requirements in young children. J Nutr 143:1489–1493

Kulkarni KS, Zala HN, Bosamia TC, Shukla YM, Kumar S, Fougat RS, Narayanan S, Patel MS, Joshi CG (2016) *De novo* transcriptome sequencing to dissect candidate genes associated with pearl millet-downy mildew (*Sclerospora graminicola* Sacc.) interaction. Front Plant Sci 7:847. doi:10.3389/fpls.2016.00847

Kumar S, Hash CT, Nepolean T, Singh G, Rajaram V, Rathore A, Mahendrakar MD, Senthilvel S, Yadav R, Srivastava R (2016) Mapping quantitative trait loci controlling high Fe and Zn density in self and open pollinated grains of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. Front. Plant Sci 7:1636

Li YC, Korol AB, Fahima T, Nevo E (2004) Microsatellites within genes: structure function and evolution. Mol Biol Evol 21:991–1007

Liu K, Muse S (2005) Power marker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129

Liu CJ, Witcombe JR, Pittaway TS, Nash M, Busso CS, Hash CT, Gale MD (1994) An RFLP-based genetic map of pearl millet (*Pennisetum glaucum*). Theor Appl Genet 89:481–487

Metzgar D, Bytof J, Wills C (2000) Selection against frameshift mutations limits microsatellite expansion in coding DNA. Genome Res 10:72–80

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with non-repetitive DNA in plant genomes. Nat Genet 30:194–200

Parekh MJ, Kumar S, Zala HN, Fougat RS, Patel CB, Bosamia TC, Kulkarni KS, Parihar A (2016) Development and validation of novel fiber relevant dbEST–SSR markers and their utility in revealing genetic diversity in diploid cotton (*Gossypium herbaceum* and *G. arboreum*). Ind Crops Prod 83:620–629

Punnuri SM, Wallace JG, Knoll JE, Hyma KE, Mitchell SE, Buckler ES, Varshney RK, Singh BP (2016) Development of a high-density linkage map and tagging leaf spot resistance in pearl millet using genotyping-by-sequencing markers. Plant Genome 09:01–13

Rajaram V, Nepolean T, Senthilvel S, Varshney RK, Vadez V, Srivastava RK, Shah TM, Supriya A, Kumar S, Ramana B, Bhanuprakash A, Narasu ML, Riera-Lizarazu O, Hash CT (2013) Pearl millet [*Pennisetum glaucum* (L.) R. Br.] consensus linkage map constructed using four RIL mapping populations and newly developed EST-SSRs. BMC Genomics 14:159

Saha MC, Mian MA, Eujayl I, Zwonitzer JC, Wang L, May GD (2004) Tall fescue EST-SSR markers with transferability across several grass species. Theor Appl Genet 109(4):783–791

Sehgal D, Rajaram V, Armstead IP, Vadez V, Yadav YP, Hash CT, Yada RS (2012) Integration of gene-based markers in a pearl millet genetic map for identification of candidate genes underlying drought tolerance quantitative trait loci. BMC Plant Biol 12:9

Senthilvel S, Jayashree B, Mahalakshmi V, Sathish KP, Nakka S, Nepolean T, Hash CT (2008) Development and mapping of simple sequence repeat markers for pearl millet from data mining of expressed sequence tags. BMC Plant Biol 8:119–127

Shivakumar PD, Vasanthi NS, Shetty HS, Smedegaard-Petersen V (2000) Ribonucleases in the seedlings of pearl millet and their involvement in resistance against downy mildew disease. Eur J Plant Pathol 106:825–836

Shivakumar PD, Geetha HM, Shetty HS (2003) Peroxidase activity and isozyme analysis of pearl millet seedlings and their implications in downy mildew disease resistance. Plant Sci 164:85–93

Smith JSC, Kresovich S, Hopkins MS, Mitchell SE, Dean RE, Woodman WL, Lee M, Porter K (2000) Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. Crop Sci 40:226–232

Supriya A, Senthilvel S, Nepolean T, Eshwar K, Rajaram V, Shaw R, Hash CT, Kilian A, Yadav RC, Lakshmi NM (2012) Development of a molecular linkage map of pearl millet integrating DArT and SSR markers. Theor Appl Genet 123:239–250

Thiel T, Michalek W, Varshney RK, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor App Gen 106(3):411–421

Thudi M, Senthilvel S, Bottley A, Hash TC, Reddy AR, Feltus AF, Varshney RK (2010) A comparative assessment of the utility of PCR-based marker systems in pearl millet. Euphytica 174(2):253–260

Varshney RK (2016) Exciting journey of 10 years from genomes to fields and markets: some success stories of genomics-assisted breeding in chickpea pigeonpea and groundnut. Plant Sci 242:98–107

Varshney RK, Thiel T, Nils S, Langridge P, Graner A (2002) *In silico* analysis on frequency and distribution of microsatellites in ESTs of some cereal species. Cell Mol Biol Lett 7:537–546

Varshney RK, Graner A, Sorrells ME (2005a) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23:48–55

Varshney RK, Sigmund R, Borner A, Korzun V, Stein N, Graner A (2005b) Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat rye and rice. Plant Sci 168:195–202

Varshney RK, Mahender T, Aggrawal RK, Borner A (2007) Genic molecular markers in plants: development and applications. In: Varshney RK, Tuberosa R (eds) Genomics assisted crop improvement vol I: genomics approaches and platforms. Springer, Dordrecht, pp 13–30

Varshney RK, Close TJ, Singh NK, Hoisington DA, Cook DR (2009) Orphan legume crops enter the genomics era. Curr Opin Plant Biol 12:202–210

Wang Z, Yu GH, Shi BB, Wang XM, Qiang HP, Gao HW (2014) Development and characterization of simple sequence repeat (SSR) markers based on RNA sequencing of *Medicago sativa* and in silico mapping onto the *M. truncatula* genome. PLoS ONE 9:e92029

You FM, Huo N, Gu YQ, Luo M, Ma Y, Hane D, Lazo GR, Dvorak J, Anderson OD (2008) BatchPrimer3: a high throughput web application for PCR and sequencing primer designing. BMC Bioinform 9:253

Yu JK, Dake TM, Singh S, Benscher D, Li W, Gill B, Sorrells ME (2004) Development and mapping of EST-derived simple sequence repeat markers for hexaploid wheat. Genome 47:805–818

Zidani S, Ferchichi A, Chaieb M (2005) Genomic DNA extraction method from pearl millet (*Pennisetum glaucum* L.) leaves. Afr J Biotechnol 4(8):862–866