Mol Biol Rep (2014) 41:1189–1200
DOI 10.1007/s11033-013-2825-7

Development and molecular characterization of genic molecular markers for grain protein and calcium content in finger millet (*Eleusine coracana* (L.) Gaertn.)

M. Nirgude · B. Kalyana Babu · Y. Shambhavi · U. M. Singh · H. D. Upadhyaya · Anil Kumar

Received: 5 January 2013 / Accepted: 25 October 2013 / Published online: 30 January 2014

Abstract Finger millet (*Eleusine coracana* (L.) Gaertn), holds immense agricultural and economic importance for its high nutraceuticals quality. Finger millets seeds are rich source of calcium and its proteins are good source of essential amino acids. In the present study, we developed 36 EST-SSR primers for the *opaque2 modifiers* and 20 anchored-SSR primers for calcium transporters and calmodulin for analysis of the genetic diversity of 103 finger millet genotypes for grain protein and calcium contents. Out of the 36 *opaque2 modifiers* primers, 15 were found polymorphic and were used for the diversity analysis. The highest PIC value was observed with the primer FMO2E33 (0.26), while the lowest was observed FMO2E27 (0.023) with an average value of 0.17. The gene diversity was highest for the primer FMO2E33 (0.33), however it was lowest for FMO2E27 (0.024) at average value of 0.29. The percentage polymorphism shown by *opaque2 modifiers* primers was 68.23 %. The diversity analysis by calcium transporters and calmodulin based anchored SSR loci revealed that the highest PIC was observed with the primer FMCA8 (0.30) and the lowest was observed for FMCA5 (0.023) with an average value of 0.18. The highest gene diversity was observed for primer FMCA8 (0.37), while lowest for FMCA5 (0.024) at an average of 0.21. The *opaque2 modifiers* specific EST-SSRs could able to differentiate the finger millet genotypes into high, medium and low protein containing genotypes. However, calcium dependent candidate gene based EST-SSRs could broadly differentiate the genotypes based on the calcium content with a few exceptions. A significant negative correlation between calcium and protein content was observed. The present study resulted in identification of highly polymorphic primers (FMO2E30, FMO2E33, FMO2-18 and FMO2-14) based on the parameters such as percentage of polymorphism, PIC values, gene diversity and number of alleles.

Keywords Finger millet · Protein · Calcium transporter and calmodulin genes · Polymorphism information content · Gene diversity · Allele frequency · Mini core collection

M. Nirgude · B. K. Babu · Y. Shambhavi · U. M. Singh · A. Kumar (✉)
Department of Molecular Biology and Genetic Engineering, College of Basic Sciences & Humanities, G.B. Pant University of Agriculture and Technology, Pantnagar 263 145, India
e-mail: anilkumar.mbge@gmail.com

B. K. Babu
Vivekananda Parvateeya Krishi Anusandhan Sansthan (VPKAS), Almora 263 601, India

H. D. Upadhyaya
International Crops Research Institute for Semi-Arid Tropics (ICRISAT), Patancheru 502 324, India

Introduction

In many developing countries of Latin America, Africa and Asia, cereals are the major staple food and often the only source of protein. Cereals typically provide 50 % of the dietary protein for humans and can comprise 70 % of the protein intake for people in developing countries [1]. The amount of seed protein varies from ~10 % (in cereals) to ~40 % (in certain legumes and oilseeds) of the dry weight, forming a major source of dietary protein [2]. The demand for cereal grains will continue to increase as a consequence of the expanding human population. However, it is well known that cereals do not provide a nutritionally balanced source of protein. The deficiency of protein leads to protein

energy malnutrition (PEM) and several other disorders which affect the normal biological functions. Among the minor-cereals, finger millet (*Eleusine coracana* L. Gaertn.) is known for its nutritional quality. Finger millet, is a tetraploid crop ($2n = 4x = 36$; genome constitution AABB) belongs to the grass family Poaceae, subfamily Chloridoideae. The calcium content of finger millet is 5–30 times higher than the other cereals [4]. The second major component of finger millet is its grain protein content (GPC) which constitutes on an average 7 % protein, however it varies from 4.88 to 15.58 % [5], and contains 44.7 % of the essential amino acids [6]. It is also recognized for its health benefits such as hypoglycemic, hypocholesterolemic and anti-ulcerative characteristics [3]. The crop is adapted to a wide range of environments, can withstand significant levels of salinity, drought and is relatively resistant to water logging, and has few serious diseases like blast. Finger millet is grown mainly by subsistence farmers and serves as a food security crop because of its high-nutritional value and excellent storage qualities.

The improvement of grain quality, either for food or for feed, is an important objective in the crop improvement programs encompassing molecular breeding. It is known that transcription factors play an important role in controlling expression during seed development. Genetic differences in the synthesis of storage proteins can be observed at the transcriptional level [7]. *Opaque2* (*o2*) is a basic leucine zipper (bZIP) transcription factor (TF) which binds to the GCN4 like motif and regulating the *zein* genes, ribosome-inactivating protein *b32* and cytosolic pyruvate orthophosphate dikinase *cyPPDK1* [8]. In *o2* mutants of maize, lysine-poor α-*zein* decreased, but lysine-rich non-*zein* increased in compensation, resulting in higher total lysine content [9]. In addition to protein-bound lysine, free lysine could be increased by altering lysine catabolism. In *o2* kernels, the expression and accumulation levels of *Zea mays* lysine keto gluterate (*ZmLKR*) are lower than those of the wild type, resulting in reduced LKR activity and a higher free lysine content [10].

There have been initiatives for finger millet improvement using plant breeding approach for different traits. The prerequisite for this involves screening of diverse germplasm for desired trait by using morphological, biochemical, and molecular markers. Molecular techniques using DNA polymorphism have been increasingly used to characterize and identify a novel germplasm for uses in the crop breeding process [11]. Several DNA marker systems are now commonly used in diversity studies of plants. The most commonly used marker systems are random amplified polymorphic DNA (RAPD) [12, 13], restriction fragment length polymorphism (RFLP) [14], amplified fragment length polymorphism (AFLP) [15], inter simple sequence repeats (ISSRs) [16] and microsatellites or simple sequence

Table 1 List of finger millet accessions used in present investigation with their calcium and protein contents

S. no.	Genotype	Origin	Ca (mg/100 g)	Protein (%)
1	IE2312	India	345	6.77
2	IE3945	Uganda	255	7.26
3	IE4671	India	207	6.95
4	IE5870	Nepal	201	7.67
5	IE2296	India	221	6.25
6	IE6294	Zimbabwe	262	8.05
7	IE5201	India	318	7.36
8	IE6326	Zimbabwe	232	7.71
9	IE3721	Uganda	343	7.11
10	IE2457	Kenya	239	6.85
11	IE6337	Zimbabwe	277	7.5
12	IE2043	India	271	6.87
13	IE5537	Nepal	281	7.4
14	VR708	Unknown	294	7.91
15	IE4570	Zimbabwe	302	7.08
16	IE518	India	287	7.64
17	IE3391	Zimbabwe	333	7.31
18	IE3317	Zimbabwe	334	7.1
19	IE2034	India	319	7.34
20	IE2589	USA	309	7.04
21	IE3470	India	279	7.44
22	IE3475	India	348	7.21
23	IE3614	Unknown	326	7.62
24	IE4057	Uganda	283	7.53
25	IE7079	Kenya	261	6.92
26	IE2957	Germany	447	7.72
27	IE6473	Uganda	228	7.32
28	IE7018	Kenya	229	6.48
29	IE3045	India	386	9.96
30	IE6350	Zimbabwe	224	7.33
31	IE2790	Malawi	345	7.91
32	IE4797	Maldives	226	7.35
33	IE3077	India	356	7.92
34	IE4121	Uganda	246	7.31
35	IE4073	Uganda	249	6.87
36	IE2710	Malawi	351	8.18
37	IE2872	Zambia	313	7.0
38	IE5066	Senegal	231	7.43
39	IE7320	Kenya	292	7.67
40	IE4491	Zimbabwe	256	8.01
41	IE2606	Malawi	332	7.08
42	IE2572	Kenya	421	7.69
43	IE2619	Malawi	291	7.58
44	IE2911	Zambia	351	7.31
45	IE501	India	324	7.47
46	IE2437	Kenya	273	6.58
47	IE6082	Nepal	239	7.69
48	IE6154	Nepal	248	7.28

Table 1 continued

S. no.	Genotype	Origin	Ca (mg/100 g)	Protein (%)
49	IE3618	India	308	7.27
50	IE4757	India	307	7.4
51	IE6514	Zimbabwe	217	7.29
52	IE2042	India	265	7.52
53	IE5817	Nepal	234	7.9
54	IE1055	Unknown	293	7.22
55	IE5091	Zimbabwe	237	6.97
56	IE3973	Uganda	196	6.8
57	IE2217	India	231	6.92
58	IE3952	Uganda	235	7.06
59	IE2430	Kenya	294	7.31
60	IE4734	India	303	8.01
61	IE4028	Uganda	237	7.36
62	IE6165	Nepal	292	7.0
63	IE6221	Nepal	251	7.55
64	IE4709	Burundi	386	9.95
65	IE6240	Zimbabwe	258	7.13
66	IE6421	Uganda	232	7.29
67	IE4497	Zimbabwe	298	6.98
68	IE4816	India	304	7.49
69	IE5306	Zimbabwe	225	7.47
70	IE2871	Zambia	351	7.22
71	IE3392	Zimbabwe	367	6.54
72	IE4646	Zimbabwe	295	7.25
73	IE5106	Zimbabwe	313	7.5
74	IE3104	India	242	7.87
75	IE6537	Nigeria	439	11.09
76	IE5367	Kenya	223	6.92
77	IE6059	Nepal	255	7.21
78	IE2821	Nepal	233	7.32
79	IE4329	Zimbabwe	224	7.59
80	IE4673	India	271	6.81
81	IE4565	Zimbabwe	244	7.95
82	IE4795	Zimbabwe	268	7.43
83	IE4622	Zimbabwe	273	7.42
84	IE4545	Zimbabwe	312	7.4
85	GPHCPB1	Uttarakhand	117.56	14.0
86	GPHCPB2	Uttarakhand	136.11	11.5
87	GPHCPB13	Uttarakhand	146.16	10.83
88	GPHCPB27	Uttarakhand	150.50	11.3
89	GPHCPB50	Uttarakhand	192.33	11.3
90	GPHCPB10	Uttarakhand	177.77	11.2
91	GPHCPB28	Uttarakhand	165.45	10.9
92	GPHCPB4	Uttarakhand	184.13	11.0
93	GPHCPB9	Uttarakhand	163.94	10.3
94	GPHCPB16	Uttarakhand	175.22	10.0
95	GPHCPB26	Uttarakhand	384.88	10.6
96	GPHCPB31	Uttarakhand	417.46	11.3
97	GPHCPB35	Uttarakhand	392.60	11.3
98	GPHCPB37	Uttarakhand	326.05	10.9
99	GPHCPB40	Uttarakhand	415.43	11.8
100	GPHCPB44	Uttarakhand	439.76	10.7
101	GE1680	Bangalore	355.63	13.95
102	GE3885	Bangalore	344.73	13.76
103	GE1583	Bangalore	380.50	11.36

repeats (SSRs) [17]. For research involving finger millet the most widely used molecular method has been RAPD [18–20], although isozymes [21], RFLP [22] and SSRs [23] have also been used successfully in genetic diversity analyses. Recently, Kumar et al. [24] used RAPD, SSR and protein profile based analysis to differentiate finger millet genotypes differing in their protein content and they found significant negative correlation ($r = -0.69$*) between the protein and calcium content of finger millet genotypes. Till, now there are no reports on the evaluation of finger millet genotypes based on their protein content using gene specific SSR markers and few reports available on calcium content based differentiation. Thus, the present study aimed at identification of genic microsatellites markers for the candidate genes involved in calcium transport pathway and *opaque2 modifiers* specific markers for differentiating the mini-core collection of 103 finger millet genotypes belongs to several parts of the world based on the protein and calcium content.

Materials and methods

Plant materials and DNA extraction

A total of 103 finger millet genotypes were used in the present study and the details are given in Table 1. Out of the 103 genotypes, 80 are mini-core collection [25] accessions and 4 control cultivars obtained from International Crop Research Institute for Semi Arid and Tropics (ICRISAT), Patancheru, India and the remaining 19 are from Ranichauri Hill Campus, G. B. Pant University of Agriculture and Technology (GBPUA&T), Pantnagar, India and University of Agricultural Sciences (UAS) Bangalore. The genotypes are categorized as follows; the genotypes having Ca content of 300–450 mg per 100 g of sample denotes as high, 200–300 mg per 100 g medium and 100–200 mg per 100 g low level of Ca. Protein content of 10 % and above considered as high protein containing genotypes, 8–10 % denotes medium, 6–8 % denotes low level of protein content. The genomic DNA of different accessions of finger millets were

isolated by standard methods [26], and subsequently quantified and analysed via agarose gel electrophoresis [27].

Development of genic microsatellite markers

Maize *opaque2* modifier sequences deposited in maize database (www.maizegdb.org) were downloaded through ftp and reported *bZIP* genes of sorghum and rice were searched in NCBI and TIGR database respectively. Using BLAST tool, the *bZIP* genes from sorghum and rice were searched for homology against maize *opaque2 modifiers* and developed the genic microsatellite markers. The selected *opaque2 modifiers* sequences were aligned by multiple sequence alignment of ClustalW software for identification of conserved regions and were further used for the identification of microsatellites by using SSRIT tool available at Gramene (www.gramene.org). Similarly, the nucleotide sequences of different candidate genes viz. calcium exchangers, channels and ATPase of cereals (finger millet, rice, maize, wheat and barley) were downloaded from NCBI and were used for SSR identification using online available SSRIT tool. Dimers, trimers, tetramers and pentamers of $\geq$5 repeats and hexamers of $\geq$3 repeats were chosen for primer designing. Primers flanking the SSRs were designed using WebSat (http://wsmartins.net/websat/) software [28] where all parameters were kept default. The SSR primers were designed for the calcium transporters using primer3 software.

SSR and anchored-SSR marker analysis

A total of 36 SSR primers for *opaque2 modifiers* and 20 anchored SSR primers for calcium candidate genes were designed for the analysis of molecular diversity of 103 finger millet accessions. PCR amplification was performed as per the standard protocol using 25–50 ng of template DNA, 30 ng of primer (Life Tech), 0.1 mM dNTPS, 0.2 U Taq DNA polymerase (Bangalore Genei pvt. Bangalore, India), 1X PCR buffer (10 mM Tris pH 8.0, 50 mM KCl and 1.8 mM MgCl2) in a volume of 25 µl. Amplification was performed with thermal cycler (Eppendorf Germany). The standardized amplification was: Initial denaturation 95 °C for 5 min followed by 40 cycles of denaturation 94 °C for 1 min; Primer annealing based on *Tm* value for 1 min; primer extension at 72 °C for 2 min; and final primer extension at 72 °C for 7 min. The annealing temperatures of the cycling parameter were readjusted for each microsatellite primers according to their calculated melting temperature (*Tm*). PCR amplified products of all the primers were subjected to gel electrophoresis using 3.5 % SFR (super fine resolution) agarose gel in 1X TAE buffer at 100 V. The fragment sizes were determined by comparing with a 100 bp DNA ladder (Genei Pvt., Bangalore, India) and the ethidium bromide stained gels were documented using Alpha Imager 1200 TM (Alpha Innotech Corporation, USA).

Statistical analysis

DNA fingerprints were scored for the presence (1) or absence (0) of bands of various molecular weight sizes in the form of binary matrix. Data were analyzed to obtain Jaccard's coefficients [29] among the genotypes by using NTSYS-pc (version 2.11 W; Exeter Biological Software, Setauket, NY, [30]. The SIMQUAL program was used to calculate the Jaccard's coefficient, a common estimator of genetic identity. Similarity matrices were utilized to construct the UPGMA (unweighted pair- group method with arithmetic average) dendrograms. The matrix comparison was carried out using the MAXCOMP function in the NTSYSpc version 2.02*i*. The polymorphic percentage was calculated by using the formula as polymorphism percentage = (no. of polymorphic bands/total bands) $\times$ 100. The polymorphism information content (PIC), gene diversity, and allele frequency were calculated using Power Marker V3.0 software [31].

Results

Development of SSR markers

The DNA sequences retrieved from the different database sources were screened for the presence of the SSRs in the UTR, CDS, Introns and ESTs. A total of 36 and 20 SSRs were designed for *opaque2 modifiers*, and calmodulin candidate genes respectively. Out of 36 *opaque2 modifier* primers, 30 showed amplification with finger millet genomic DNA, however seven out of fifteen primers were amplified for calcium transporter and calmodulin genes. The primers which shown amplification with finger millet genomic DNA were used to study the molecular diversity in the 103 finger millet genotypes with respect to the protein and calcium content.

SSR analysis and genetic diversity of finger millet genotypes based on *Opaque2 modifier* SSRs

Out of the 30 amplified SSR loci, 15 primers were polymorphic which generated 59 reproducible alleles. The number of alleles produced was in the range of 2 to 8 with an average of 4.8 alleles per primer. Out of the 74 alleles, 59 were polymorphic and the remaining 15 were monomorphic with a size ranged from 100 to 900 bp. The molecular profiling of the 103 finger millet genotypes with the *opaque2 modifiers* specific SSR loci FMO2E29 showed in Fig. 1. The average polymorphism percentage revealed by *opaque2*

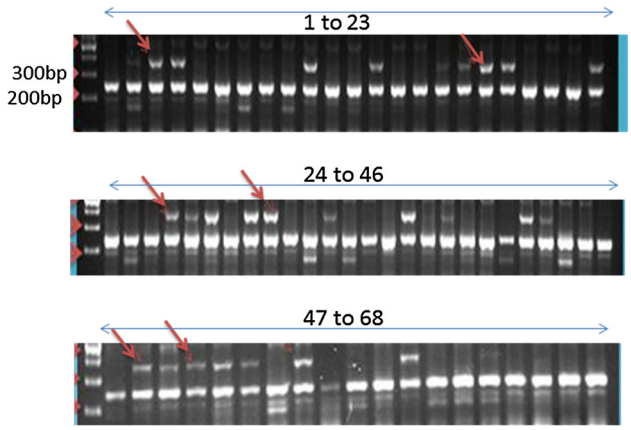

Fig. 1 SSR profile of 68 finger millet genotypes generated by primer FMO2E29

modifier SSR loci was 68.23 %. The PIC for each *o2* SSR marker was calculated to know its capability of making distinctions, assess the quality of markers and to compare the effectiveness of each primer combination in rendering genetic information. The PIC value and gene diversity was calculated by using the software Power Marker v 3.25. The highest PIC value was observed with the primer FMO2E33 (0.256) and the lowest was observed for FMO2E27 (0.023) at an average PIC value of 0.171. The highest major allele frequency is 0.98 for the primer FMO2E27, while it was lowest for primer FMO2E33 (0.757) with an average value of 0.852. The highest gene diversity was observed for primer FMO2E33 (0.327), however lowest was for primer FMO2E27 (0.024) at an average gene diversity value of 0.290. The details of the designed SSRs with theirs expected size of alleles, PIC, gene diversity, allele frequency were given in Table 2. The sequences of the polymorphic loci along with GenBank accession numbers was given in Table 4.The distribution of *opaque2 modifier* alleles among the high protein, medium protein and low protein content genotypes indicates that the alleles O2A13, O2A30, O2A31 and O2A33 are distributed 75, 75, 100 and 95 % in high protein containing genotypes respectively which is comparatively higher than in medium and low protein containing genotypes. So it can be hypothesized that these alleles may be associated with the high protein content. The structural variation of *o2* might affect in the transcriptional efficiency of seed storage protein genes.

The dendrogram was constructed by using both NTSYSpc2.11 and Power Marker v 3.25 software, and both were similar. The similarity coefficients were used as input data for the cluster analysis using NTSYSpc 2.11 program and the unweighted pair group method with arithmetic mean (UPGMA) cluster analysis method was followed for construction of phylogenetic tree (Fig. 2). The dendrogram generated using SSR similarity matrix data resulted in four major clusters (A, B, C and D) at a similarity of 0.64. Grouping of genotypes were broadly consistent with the protein content of the genotypes. The high protein containing genotypes, which belonged to Ranichauri Hill Campus, GBPUA & T, Pantnagar were grouped into three clusters B, C and D with the similarity coefficient ranged from 0.49 to 0.87. However, four genotypes (GPHCPB 13, 27, 44 and 50) were exceptionally clustered together with

Table 2 Details of the polymorphic *opaque2* SSRs with their genetic parameter values

S. no.	Primer name	Product range (bp)	Number of polymorphic alleles	Percentage of polymorphism	Major allele frequency	Gene diversity	PIC value	Unique allele (bp)
1	FMO2E25	150–350	7	100	0.897	0.177	0.158	125
2	FMO2E26	125–1,000	4	80.0	0.904	0.165	0.146	–
3	FMO2E27	225–350	1	25.0	0.987	0.0248	0.023	–
4	FMO2E28	100–700	4	100	0.906	0.144	0.125	180
5	FMO2E29	100–450	4	66.66	0.898	0.1432	0.116	350
6	FMO2E30	100–650	5	83.33	0.812	0.2862	0.235	500
7	FMO2E31	100–700	2	100	0.810	0.244	0.190	–
8	FMO2E32	175–700	4	80.0	0.895	0.175	00.153	–
9	FMO2E33	150–500	3	75.0	0.757	0.327	0.256	–
10	FMO2E34	100–350	1	25.0	0.858	0.163	0.123	–
11	FMO2E35	250–400	3	75.0	0.951	0.083	0.074	–
12	FMO2E36	100–400	4	80.0	0.853	0.204	0.164	–
13	FMO2-18	150–900	8	100	0.758	0.313	0.245	–
14	FMO2-3	400–1,000	4	80.0	0.797	0.259	0.206	–
15	FMO2-14	100–450	5	83.33	0.776	0.290	0.230	–
	Average		3.9	63.23	0.852	0.209	0.171	

Table 3 Details of the polymorphic anchored SSRs of calcium transporters with their genetic parameter values

S. no.	Primer name	Product range (bp)	Number of polymorphic alleles	Perecentage of polymorphism	Major allele frequency	Gene diversity	PIC value	Unique allele (bp)
1	FMCA8	25–400	–	100	0.751	0.368	0.299	350, 250
2	FMCA5	150–500	3	25	0.987	0.0249	0.300	–
3	FMCA14	100–250	–	100	0.805	0.285	0.235	–
4	FMCA13		2	50	0.976	0.046	0.046	–
5	FMCA11	100–500	–	100	0.785	0.301	0.243	650
	Average		1.0	75	0.855	0.213	0.175	

Table 4 List of *Opaque2* primers used in present study along with sequences and Tm value

S. no.	GenBank accession number	Primer source	Primer name	Primer sequence	Tm (°C)
1	AB053474.1	*Zea mays*	FMO2E – 25	**F** CAGAAACCTTCCCAAACAAATC	47
				R CAGCACCAAAACACCAGAAATA	
2	AB053475.1	*Zea mays*	FMO2E- 26	**F** CTTTCCCACCATTTCCTTCTCT	47
				R GCTCCTCCAGAAACCTCTCG	
3	AB053475.1	*Zea mays*	FMO2E – 27	**F** TGAGAGTGTGTGTTTAGCTCCG	47
				R TGCACTTCTGTGAACTCAGGAT	
4	AB053475.1	*Zea mays*	FMO2E – 28	**F** GTCAACTCTGATGCTTCTGTGC	47
				R GACGATCCCCATGCTGTATATT	
5	AB053473.1	*Zea mays*	FMO2E – 29	**F** TTCTCTTTCCCTTCTCCTCCTC	47
				R ATTGGTAGCTGATGTGGGACTT	
6	AB053473.1	*Zea mays*	FMO2E – 30	**F** ACGGACAACAGAATCCTCAAAT	47
				R GTTCACTTGGACACATCACGTT	
7	AB053474.1	*Zea mays*	FMO2E – 31	**F** TAGTAAATAACGGGGCAATTCG	47
				R TGCAACAACACAACAGATGCTA	
8	AF395819.1	*Zea mays*	FMO2E – 32	**F** CCTCTCTCCTCTCCTCTTCTCC	47
				R ACCGCCTCCTCCAGAAAC	
9	AF395819.1	*Zea mays*	FMO2E – 33	**F** AAGTGATACATTGGCAGGGAAC	47
				R CCCATTGCCGTATTATTGTCTT	
10	AB021736.1	*Zea mays*	FMO2E – 34	**F** GAGGAAAGTAAGCCCAGATTCA	47
				R CGAGAGGGACACACACAGATT	
11	AB021736.1	*Zea mays*	FMO2E – 35	**F** TTCGCAGGTTTTATAGGATTGC	47
				R TTTAGCATTGTCCACACACACA	
12	AB021736.1	*Zea mays*	FMO2E – 36	**F** TGGGTGATTTTGTATGGAGATG	47
				R TGTCAACAGCAGCATCATTGTA	
13	SB03G004290.1	*Sorghum bicolor*	FMO2-18	**F** GACTCTTTTCTCTCTCCTCCCC	55
				R CTCCTCTCGACCTCTCCTCC	
14	NM_001112217.1	*Zea mays*	FMO2-3	**F** GCTGTGGGTTCTGGAAAAGTAG	55
				R AAGAAGGTGTTGACGAGAGAGG	
15	SB02G018870.1	*Sorghum bicolor*	FMO2-14	**F** ATATGGACTGACGACGCAAATA	55
				R TGGAGAGATCAGAAGTAGACAAGG	

genotypes of low protein containing genotypes under cluster A. The major cluster A further divided into two sub-clusters. The sub-cluster AI has two genotypes IE5367 and IE3392 with a similarity coefficient of 0.85. The sub-cluster AII forms several mini-clusters in which genotypes IE3104 and VR708 show maximum similarity with similarity coefficient 1.00. The major cluster A comprised of largely low protein containing genotypes. The principal component analysis (PCA) analysis has clearly differentiated the selected finger millet genotypes based on the

Table 5 List of primers of calmodulin and calcium transporters used in present study

S. no.	Accession no.	Primer source	Primer name	Primer sequence	*Tm*
1	EB739736	*Eleusine coracana*	FMCA1	**F** ACACCCCGCCGTAGGCCA	55
				R CGCGTCAGATGAGAGAGA	
2	CX265535	*Eleusine coracana*	FMCA2	**F** CGCGTCTCCTTGCCGCAG	57
				R ACTGACTCCGAGGAGGAG	
3	CX265535	*Eleusine coracana*	FMCA3	**F** CGCGGAGTTCAAGGAGGC	55
				R TTATTAGGCNTACCAANT	
4	CX264726	*Eleusine coracana*	FMCA4	**F** GAGAGAAGCCTCGAGGAC	55
				R CTGCTGCTGAACATTTGA	
5	EB739736	*Eleusine coracana*	FMCA5	**F** GTGTAGAGCTGGCAAGAG	51
				R GAGAGAGCGCGGCGAGAG	
6	AF489111	*Oryza sativa*	FMCA6	**F** CCTCCTCCTCGTCCTCGG	58
				R TGAAGGGGATTCTCCTCCTC	
7	AF489111	*Oryza sativa*	FMCA7	**F** CGCGACGTGCGGGAACGCG	58
				R TCGGGTTCGCCAGCGCGC	
8	EF446604	*Hordeum vulgare*	FMCA8	**F** CATGGCATGTCCTCCTCC	58
				R AGGGAGTCCTCCTCCTCC	
9	AF256229	*Zea mays*	FMCA9	**F** GAGATATGCTGTGAGCTC	50
				R GAGACTGGTTGTTTGTTT	
10	NM_001051880	*Oryza sativa*	FMCA10	**F** CCGCCGCCGCCGCTTCCG	60
				R CGTCGACGGCGGCGGCGG	
11	NM_001051880	*Oryza sativa*	FMCA11	**F** CCGCCGCCGCCGCTTCCG	60
				R GCGCGCGCTCGTGCGCCA	
12	NM_001051880	*Oryza sativa*	FMCA12	**F** GCGCGCCGGGCGGGCTGC	60
				R CGTCGACGGCGGCGGCGG	
13	AF096871	*Zea mays*	FMCA13	**F** GGAGGAGCGACGACACAC	60
				R GCGGTCCAGGAAGAAGAA	
14	AF096871	*Zea mays*	FMCA14	**F** TCGCAGGAGGACGACGAC	60
				R GGAGGAGCGACGACACAC	
15	AF096871	*Zea mays*	FMCA15	**F** CGCCGCCGCCGCCGTCTC	60
				R GCTCTGCATCGTCGCGCG	
16	AB071014	*Oryza sativa*	FMCA16	**F** TGTGTGCTGTTACACTGG	51
				R ATCATTCAGCACTGTGTG	
17	AB071014	*Oryza sativa*	FMCA17	**F** TGTGTGCTGTTACACTGG	55
				R GAGTTTGTCTTTGGATGGA	
18	AB071014	*Oryza sativa*	FMCA18	**F** CTCTCTGGTGCCGGGGAT	57
				R GAAGGAAGAAACAGGCGGCG	
19	EU954499	*Zea mays*	FMCA19	**F** TGTGTGCTGTTACGCTGG	57
				R CATTAGGATTTGGCGC	
20	EU954499	*Zea mays*	FMCA20	**R** GGACTATTTTTATTGGAGAGA	56
				R GAGATGGAACTGGAGAAAG	

protein content into three clusters (I, II and III). The cluster I contained all the genotypes of low and medium protein content belonged to mini-core collection. However, cluster II and III contained the high protein containing finger millet genotypes. The two high protein containing genotypes GPHCPB1 and GPHCPB37 are not grouped with any other genotypes. The overall grouping pattern of PCA is similar to the clustering pattern of the dendrogram. The grouping of the finger millet cluster indicates the high level of similarity within clustered genotypes. The high levels of genetic similarity indicated that accessions were related and the variation is limited.

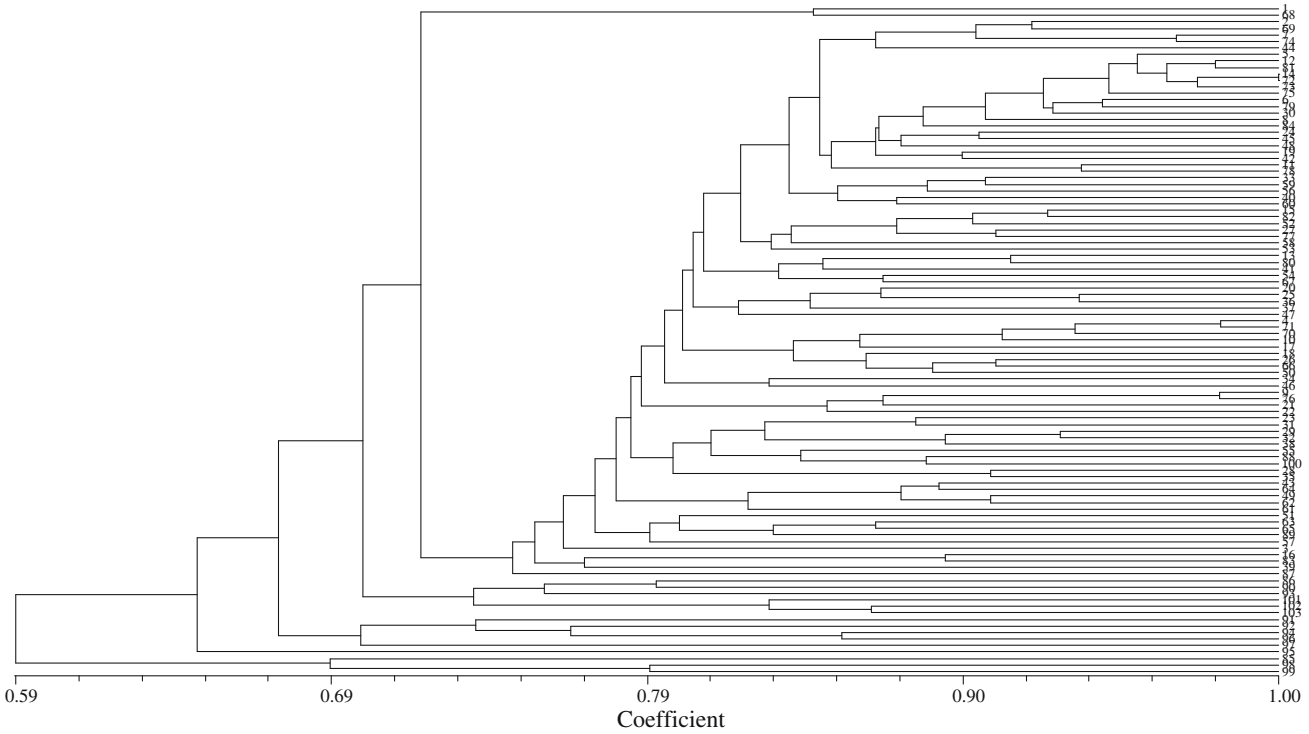

Fig. 2 The dendrogram of 103 finger millet genotypes based on UPGMA analysis using *opaque2* gene based SSR loci (for labels, please refer to Table 1)

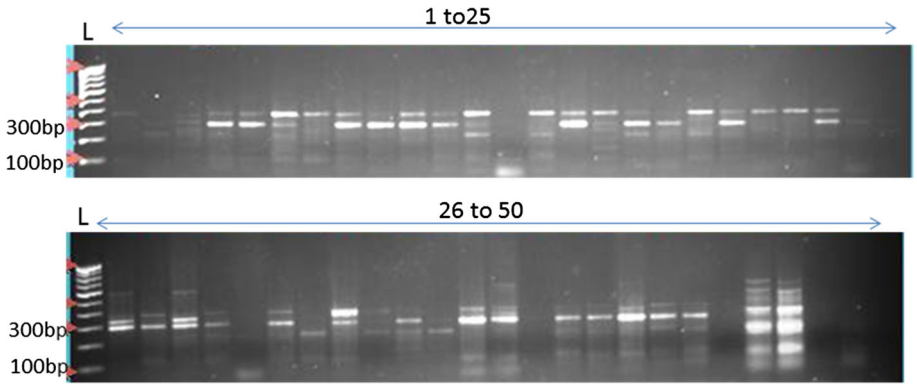

Fig. 3 SSR profile of FMCA8 on finger millet genotypes

SSR analysis and genetic diversity using calcium transporters and candidate genes based anchored SSRs

Twenty anchored-SSR primers of calcium transporters were used to characterize the genetic diversity among the 103 finger millet genotypes. The sequences of the SSR loci along with GenBank accession numbers was given in Table 5.Five, out of twenty primers generated 19 reproducible alleles, and each of these primers varied in their ability to resolve the variability among the genotypes. The molecular profiling of finger millet genotypes with the SSR loci FMCA8 was given in Fig. 3. The number of alleles by each individual primer was in the range of 2–7 with an average of 3.8 alleles per primer. Out of the 19 alleles, 15 were polymorphic and 4 were monomorphic. The molecular weight of the amplified alleles was ranged from 100 to 900 bp with the polymorphism percentage of 75 %. The highest PIC value was observed with the primer FMCA8 (0.299), while lowest was for primer FMCA13 (0.04) at an average PIC value of 0.175. The highest major allele frequency was observed for the primer FMCA5 (0.987), however it was lowest for FMCA8 (0.751) with a mean frequency of 0.854. The gene diversity was highest for primer FMCA8 (0.368) and lowest for FMCA5 (0.024) at an average value of 0.213 (Table 3).

The dendrogram was constructed by using both NTSYSpc2.11 and POWERMARKER software, and both were similar. The similarity coefficients were used as input data for the cluster analysis using NTSYSpc 2.11 program

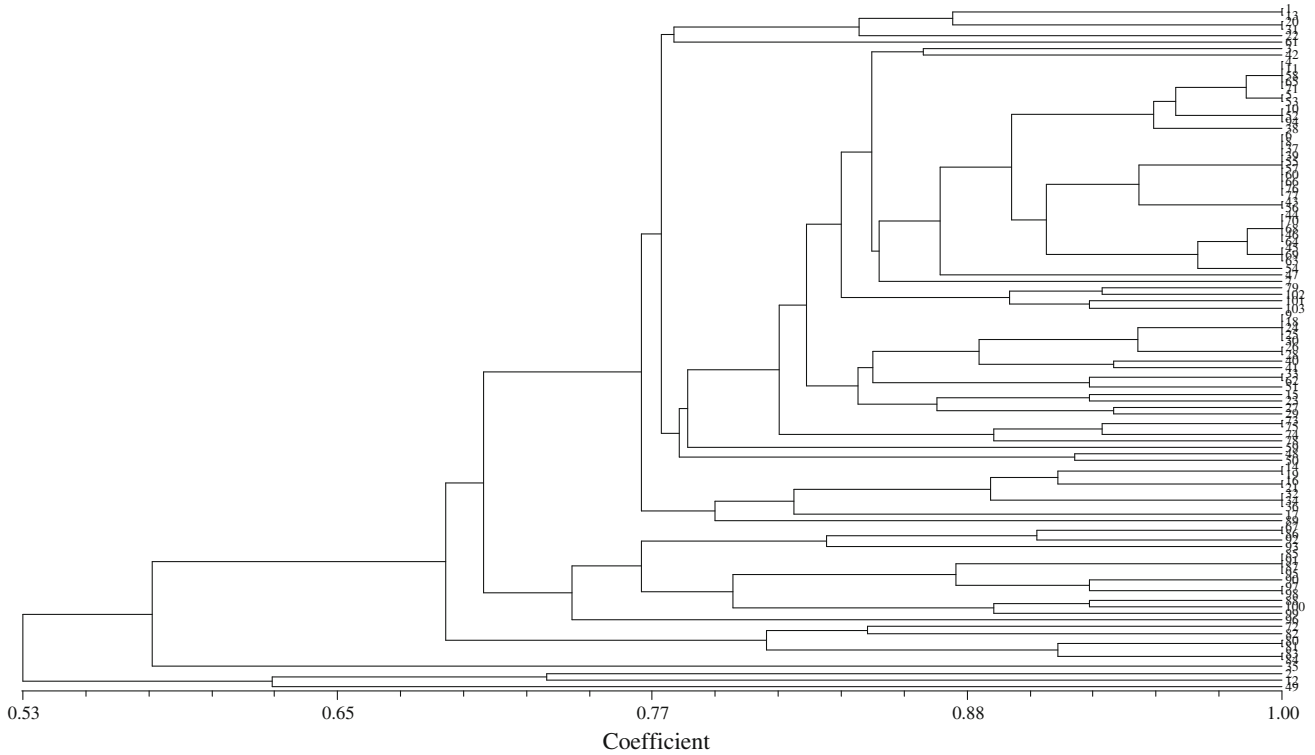

Fig. 4 The dendrogram based on UPGMA analysis of 103 finger millet genotypes using anchored SSRs of calcium transporters and candidate genes (for labels, please refer to Table 1)

and the unweighted pair group method with arithmetic mean (UPGMA) cluster analysis method was followed for construction of phylogenetic tree (Fig. 4). The dendrogram generated using SSR similarity matrix data resulted in four major clusters (A, B, C and D) at a similarity of 0.54. Grouping of genotypes were largely in congruence with the calcium content of the genotypes. The cluster A comprised of medium and high calcium genotypes mostly belonged to mini-core collection of finger millet genotypes and along with GPHCPB 40 and GPHCPB 50. These two are high and medium calcium containing genotypes, contained 415 and 200 mg respectively. The genotype GPHCPB 40 was clustered together with high calcium containing genotypes viz., IE4757, and IE4497. Similarly, cluster B mostly consisted of high and medium ca containing genotypes with few exceptions. These results indicated that the SSR loci used in the present study could able to broadly differentiate the finger millet genotypes into high and medium calcium containing and low ca containing genotypes. Hence there is a need to develop more genic based polymorphic SSR loci to explore the marker systems with high genome coverage so as to identify the markers linked to high calcium content. Similar results also have been obtained with the 2D pattern of principal component analysis. Group A contained all the high and medium ca containing genotypes. In group B, high and medium ca containing genotypes were close to each other with few exceptions.

Discussion

Cereals in-general and millets in particular are the staple food for the people living in the most of the countries of under-developed and developing parts of the world [32]. Cereals typically provide 50 % of the dietary protein for humans and can comprise 70 % of the protein intake for people in developing countries [33]. But unfortunately, the nutritional profile of cereals is poor as it is deficient in essential amino acids such as lysine, tryptophan and methionine due to a relatively higher proportion of prolamines as storage proteins which are essentially devoid of lysine and tryptophan. The reason concerning is that lysine, tryptophan and methionine are the limiting amino acids in human beings [33]. Finger millet is the excellent source of the calcium and also the minerals like chromium, zinc, copper, magnesium. Recently, Upadhyaya et al. [34] studied the genetic diversity analysis of core collection [35] of finger millet genotypes for their mineral composition like iron zinc, calcium and protein content and found a substantial genetic variability. The proteins of the finger millet are rich source of the amino acids like methionine, and lysine which limiting the human diet. Hence keeping in mind the importance of this crop, research efforts are necessary to explore the genetic information of this crop and understand the basis of the potential of this crop with high mineral and protein quality content.

In the present study, a mini core collection of finger millet (80 accessions), along with accessions from different parts of India and four control cultivars were used to analyse the genic molecular marker based genetic diversity with respect to the protein and calcium content. Understanding the diversity of finger millet germplasm collections is important for effective exploitation of their genetic potential as well as for selection of genotypes as breeding lines, maintenance and for conservation. The genic microsatellites are often located in introns and un-translated regions of sequenced cDNAs but can also be located in exons and translated regions of ORFs [36]. These genic SSRs tend to be more readily transferable between related species or genera than genomic ones, since coding sequence is better conserved than non-coding sequence. They also provide a powerful means to link the genetic maps of related species, and since many of them are located within genes of known or at least putative function, any allelic variation present can be exploited to generate perfect markers [37]. Alleles of genic microsatellites can be associated with structural mutations that lead to novel proteins that are larger or smaller than those of the original alleles and which can have substituted or repeated amino acids.

The analysis of the molecular diversity by the 15 *opaque2 modifier* SSRs resulted in 72 alleles, of which 59 were polymorphic. The percentage of polymorphism observed among the selected germpalsm was 68.23 % which is higher than the polymorphism showed by RAPD (49.43 %), SSR (50.2 %) and cytochrome P450 gene based markers (58.75 %) [38]. Salimath et al. [39] detected 10 % and 26 % of polymorphism respectively in 17 accessions of *E. coracana* from Africa and Asia with RAPDs which is very low than the present study. Our results showed that the designed *opaque2 modifiers* SSRs were able to differentiate the selected 103 finger millet genotypes into high, medium and low protein containing genotypes (Fig 2). Similar results were also obtained by Kumar et al. [24], where they used RAPD, SSR and protein profiling for differentiating 52 finger millet genotypes of India based on protein content. But in the present study, a large collection of mini core finger millet genotypes including Indian accessions were taken and the primers were designed from the *opaque2* specific bZIP related protein sequences. Hence, it gave better understanding of the genetic variation for protein content among the finger millet genotypes and can be enhanced through molecular breeding programmes. The four genotypes (GPHCPB 13, 27, 44 and 50) were exceptionally clustered together with genotypes of low protein containing genotypes under cluster A though they contain high protein content, which may be due to more genome identity among the accessions.

The higher level of polymorphism revealed by the *o2* SSRs showed that there may be structural variation in the *o2* gene which may affect the total grain protein content of finger millet. There have been reports indicating the association of microsatellite repeats with the regulation and/or functioning of genes in plant specific pathways [40]. The SSR variations in the transcription units could affect transcription, translation, and cause transcription slippage thereby leading to disruption in splicing and possibly disturbing cellular functions [41]. The present results indicate a possible role of genic SSRs in governing trait variations amongst a collection of germplasm of finger millet. The SSRs designed from *o2* ESTs played a major role in discriminating finger millet genotypes with respect to protein content followed by primers designed from introns and UTRs. Hence it can be hypothesised that there may be allelic variation in *o2* transcription factor genes which is responsible for the differential regulation of prolamin genes and this may ultimately affect the total seed protein content in different finger millet genotypes. The identification of three important *o2* alleles showed variation among the coding regions and associated with high grain protein content (Fig 2). The structural variation of *o2* might affect in the transcriptional efficiency of seed storage protein genes. The association of 103 finger millet genotypes by PCA analysis also confirmed in differentiating the genotypes into three groups based on the protein content.

The polymorphism percentage observed was 75 % for the calcium candidate genes, which is significantly higher than earlier studies [38]. Panwar et al. [38] analysed the genetic diversity with respect to calcium content using RAPD, SSR and CytochromeP_{450} based markers which revealed respectively 49.4, 50.2 and 58.7 % polymorphism in 52 genotypes of *E. coracana*. The average PIC value obtained in the present study was 0.171; however Panwar et al. [38] found average PIC as 0.35 for RAPD, 0.50 for SSR and 0.40 for cytP450 gene based markers which are higher than our results. The polymorphic anchored SSR markers designed from coding regions of selected calcium sensor and transporters could broadly resolve finger millet genotypes with respect to their calcium content. However, Panwar et al. [38] differentiated the 52 finger millet genotypes of India were analyzed using RAPD, SSR and cytochromeP450 based marker and found wide genetic base and clustered according to calcium contents. This low amount of variation obtained in our results may be due to less number of polymorphic loci and there is a more scope to explore the marker systems The distribution of calcium transporters and signaling alleles among the high calcium, medium calcium and low calcium content genotypes indicates that the single allele is distributed 64.1 % present in high calcium containing genotypes which is higher than medium and low calcium containing genotypes, so this allele may contribute to high calcium content in these genotypes. In the present study, we found negative

correlation between the results of protein and calcium containing genotypes both at molecular level and phenotypic level. Similarly, Kumar et al. [24] used RAPD, SSR and protein profile based analysis to differentiate finger millet genotypes differing in their protein content and they found significant negative correlation ($r = -0.69$*) between the protein and calcium content of finger millet genotypes. These results indicated that there may be a negative correlation between protein and calcium content in the finger millet crop.

Thus the present study resulted in characterizing the world wide mini-core collection and Indian finger millet genotypes by using the genic molecular markers for protein and calcium content. The present results indicated a possible role of genic SSRs in governing trait variations amongst a collection of germplasm of finger millet. The *opaque2 modifier* specific EST-SSRs could able to differentiate the finger millet genotypes into high, medium and low protein containing genotypes. However, calcium dependent candidate gene based EST-SSRs could broadly differentiate the genotypes based on the calcium content with a few exceptions. A significant negative correlation between calcium and protein content was observed. The present study resulted in identification of highly polymorphic primers (FMO2E30, FMO2E33, FMO2-18 and FMO2-14) based on the parameters such as percentage of polymorphism, PIC values, gene diversity and number of alleles.

References

1. Deutscher D (1978) The current status of breeding for protein quality in corn. In: Friedman M (ed) Nutritional improvement of food and feed grains. Plenum, New York, pp 281–300
2. Shewry PR, Napier JA, Tatham AS (1995) Seed storage proteins: structures and biosynthesis. Plant Cell 7:945–956
3. Chethan S, Malleshi NG (2007) Finger millet polyphenols: characterization and their nutraceutical potential. American J Food Technol 2(7):582–592
4. National Research Council (1996) Lost crops of Africa vol. I: grains. Board on Science and technology for International Development. National Academy Press, Washington, DC
5. Singh P, Srivastava S (2006) Nutritional composition of sixteen new varieties of finger millet. J Community Mobil Sustain Dev 12:81–84
6. Mbithi-Mwikya S, Ooghe W, Van Camp J, Nagundi D, Huyghebaert A (2000) Amino acid profile after sprouting, autoclaving and lactic acid fermentation of finger millet (*Eleusine coracana*) and kidney beans (*Phaseolus vulgaris* L.). J Agric Food Chem 48(8):3081–3085
7. Sørensen MB, Cameron-Mills V, Brandt A (1989) Transcriptional and post-transcriptional regulation of gene expression in developing barley endosperm. Mol Gen Genet 217(2–3):195–201
8. Maddaloni M, Donini G, Balconi C, Rizzi E, Gallusci P, Forlani F, Lohmer S, Thompson R, Salamini F, Motto M (1996) The transcriptional activator*Opaque-2*controls the expression of a cytosolic form of pyruvate orthophosphate dikinase-1 in maize endosperms. Mol Gen Genet 250(5):647–654
9. Mertz ET, Bates LS, Nelson OE (1964) Mutant gene that changes protein composition and increases lysine content of maize endosperm. Science 145(3629):279–280
10. Kemper EL, Neto GC, Papes F, Martinez Moraes KC, Leite A, Arrudaa P (1999) The role of *Opaque2* in the control of lysine-degrading activities in developing maize endosperm. Plant Cell 11:1981–1993
11. O'Neill R, Snowdon R, Köhler W (2003) Population genetics: aspects of biodiversity. Prog Bot 64:115–137
12. Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535
13. Jacobson A, Hedrén M (2007) Phylogenetic relationships in *Alisma*(Alismataceae) based on RAPDs, and sequence data from ITS and *trn*L. Plant Syst Evol 265(1–2):27–44
14. Soller M, Beckmann JS (1982) Restriction fragment length polymorphisms and genetic improvement. In: World Congress on genetics applied to livestock production. Madrid, vol. 6, pp. 396–404
15. Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407
16. Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20(2):176–183
17. Becker J, Heun M (1995) Barley microsatellites: allele variation and mapping. Plant Mol Biol 27(4):835–845
18. Salimath SS, Olivera ACD, Godwin ID, Bennetzen JL (1995) Assesment of genome origins and diversity in the genus *Eleusine* with DNA markers. Genome 38:757–763
19. Fakrudin B, Shashidhar HE, Kulkarni RS, Hittalmani S (2004) Genetic diversity assessment of finger millet, *Eleusine coracana* (Gaertn.), germplasm through RAPD analysis. PGR Newslett 138:50–54
20. Hilu KW, Johnson JL (1992) Ribosomal DNA variation in finger millet and wild species of *Eleusine* (Poaceae). Theor Appl Genet 83:895–902
21. Babu BK, Senthil N, Gomez SM, Biji KR, Rajendraprasad NS, Kumar SS, Babu RC (2007) Assessment of genetic diversity among finger millet (*Eleusine coracana* (L.) Gaertn.) accessions using molecular markers. Genet Resour Crop Evol 54:399–404
22. Parani M, Rajesh K, Lakshmi M, Parducci L, Szmidt AE, Parida A (2001) Species identification in seven small millet species using polymerase chain reaction-restriction fragment length polymorphism of trn S-psb C gene region. Genome 44(3):495–499
23. Panwar P, Saini RK, Sharma N, Yadav D, Kumar A (2010) Efficiency of RAPD, SSR and Cytochrome P450 gene based markers in accessing genetic variability amongst finger millet (*Eleusine coracana*) accessions. Mol Biol Rep 37(8):4075–4082
24. Kumar A, Sharma N, Panwar P, Gupta AK (2012) Use of SSR, RAPD markers and protein profiles based analysis to differentiate *Eleusine coracana* genotypes differing in their protein content. Mol Biol Rep 39(4):4949–4960
25. Upadhyaya HD, Sarma NDRK, Ravishankar CR, Albrecht T, Narsimhudu Y, Singh SK, Varshney SK, Reddy VG, Singh S, Dwivedi SL, Wanyera N, Oduori COA, Mgonja MA, Kissandu DB, Parzies HK, Gowda CLL (2010) Developing a mini-core collection in finger millet using multilocation data. Crop Sci 50:1924–1931
26. Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4326
27. Maniatis T, Sambrook J, Fritsch EF (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

28. Martins WS, Lucas DCS, Neves KFS, Bertioli DJ (2009) WebSat: a web software for microsatellite marker development. Bioinformation 3:282–283
29. Jaccard P (1908) Nouvelles recherches sur la distribuition florale. Bull Soc Vaudoise Sci Nat 44:223–270
30. Rohlf FJ (1997) NTSYS-pc numerical taxonomy and multivariate analysis system, version 20. Exeter Publications, New York
31. Liu K, Muse M (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129
32. Deutscher D (1978) The current status of breeding for protein quality in corn. In: Friedman M (ed) Nutritional improvement of food and feed grains. Plenum, New York, pp 281–300
33. Sofi PA, Shafiq AW, Rather AG, Shabir HW (2009) Quality protein maize (QPM): genetic manipulation for the nutritional fortification of maize. J Plant Breed Crop Sci 1(6):244–253
34. Upadhyaya HD, Ramesh S, Shivali Sharma, Singh SK, Varshney SK, Sarma NDRK (2011) Genetic diversity for grain nutrients contents in a core collection of finger millet (*Eleusine coracana* (L.) Gaertn.) germplasm. Field Crops Res 121:42–52
35. Upadhyaya HD, Gowda CLL, Pundir RPS, Reddy VG, Singh S (2006) Development of core subset of finger millet germplasm using geographical origin and data on 14 quantitative traits. Genet Resour Crop Evol 53:679–685
36. Fujimori S, Washio T, Higo K, Ohtomo Y, Murakami K, Matsubara K, Kawai J, Carninci P, Hayashizaki Y, Kikuchi S, Tomita M (2003) A novel feature of microsatellites in plants: a distribution gradient along the direction of transcription. FEBS Lett 554:17–22
37. Andersen JR, Lübberstedt T (2003) Functional markers in plants. Trends Plant Sci 8(11):554–560
38. Panwar P, Nath M, Kumar V, Kumar A (2010) Comparative evaluation of genetic diversity using RAPD, SSR and cytochrome P450 gene based markers with respect to calcium content in finger millet (*Eleusine coracana* L. Gaertn.). J Genet 89(2):121–133
39. Salimath SS, de Oliveira AC, Bennetzen JL, Godwin ID (1995) Assessment of genome origins and genetic diversity in the genus Eleusine with DNA markers. Genome 38(4):757–763
40. Kashi Y, Soller M (1999) Functional roles of microsatellites and minisatellites. In: Goldstein DB, Schlötterer C (eds) Microsatellites: evolution and applications. Oxford University Press, New York, pp 10–23
41. Lawson MJ, Zhang L (2006) Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. Genome Biol 7:R14