DNA Research **20**, 197–207, (2013)
Advance Access publication on 3 February 2013
doi:10.1093/dnares/dst002

Genome-Wide Development and Use of Microsatellite Markers for Large-Scale Genotyping Applications in Foxtail Millet [*Setaria italica* (L.)]

Garima Pandey†, Gopal Misra†, Kajal Kumari†, Sarika Gupta, Swarup Kumar Parida, Debasis Chattopadhyay, and Manoj Prasad*

National Institute of Plant Genome Research (NIPGR), Aruna Asaf Ali Marg, New Delhi 110 067, India

*To whom correspondence should be addressed. Tel. +0091 11-2673-5160. Fax. +0091 11-2674-1658.
Email: manoj_prasad@nipgr.ac.in

Edited by Dr Satoshi Tabata
(Received 7 November 2012; accepted 9 January 2013)

Abstract

The availability of well-validated informative co-dominant microsatellite markers and saturated genetic linkage map has been limited in foxtail millet (*Setaria italica* L.). In view of this, we conducted a genome-wide analysis and identified 28 342 microsatellite repeat-motifs spanning 405.3 Mb of foxtail millet genome. The trinucleotide repeats (~48%) was prevalent when compared with dinucleotide repeats (~46%). Of the 28 342 microsatellites, 21 294 (~75%) primer pairs were successfully designed, and a total of 15 573 markers were physically mapped on 9 chromosomes of foxtail millet. About 159 markers were validated successfully in 8 accessions of *Setaria* sp. with ~67% polymorphic potential. The high percentage (89.3%) of cross-genera transferability across millet and non-millet species with higher transferability percentage in bioenergy grasses (~79%, Switchgrass and ~93%, Pearl millet) signifies their importance in studying the bioenergy grasses. *In silico* comparative mapping of 15 573 foxtail millet microsatellite markers against the mapping data of sorghum (16.9%), maize (14.5%) and rice (6.4%) indicated syntenic relationships among the chromosomes of foxtail millet and target species. The results, thus, demonstrate the immense applicability of developed microsatellite markers in germplasm characterization, phylogenetics, construction of genetic linkage map for gene/quantitative trait loci discovery, comparative mapping in foxtail millet, including other millets and bioenergy grass species.

Key words: foxtail millet (*Setaria italica* L.); microsatellite; comparative mapping; physical mapping; transferability

1. Introduction

Foxtail millet [(*Setaria italica* (L.) P. Beauv.; $2n = 18$)], a member of the Poaceae family, is the second most-widely planted species of millets, next to pearl millet (*Pennisetum glaucum*) (FAOSTAT 2005; http://faostat.fao/org/). It is mainly cultivated in semi-arid regions of India, China and other parts of Asia, North Africa and America. It is a potential model system for biofuel crops due to its small genome size (~515 Mb), less repetitive DNA and self-pollinated nature[1–3] and the most important are its close phylogenetic relationships with several biofuel crops such as switchgrass (*Panicum virgatum*), napier grass (*Pennisetum purpureum*) and pearl millet (*P. glaucum*). Furthermore, foxtail millet is an excellent genetic system to study evolution and physiology of C4 photosynthesis and abiotic stress tolerance mechanism particularly for salinity and dehydration stresses.[2–4] Recently, the Joint Genome Institute of the Department of Energy, USA and Beijing Genomics Institute, China have sequenced the genome of two foxtail

† Authors contributed equally.

millet accessions.[5,6] The sequence-based comparative genome mapping has revealed a high degree of syntenic relationship of foxtail millet with rice and sorghum, in spite of their divergence more than 50 million years ago.[7] With the availability of genome-wide sequence resources in foxtail millet, it is now required to develop and validate useful sequence-based genetic markers and genes in a larger genotype sets and utilize them in genetic improvement of target millet and its related crop species.

Among the available different genetic marker system, the microsatellite markers are considered important in various genetic studies because of their high reproducibility, co-dominant inheritance, multi-allelic variation and abundance in the genome.[8,9] These desirable characteristics of microsatellite markers had made them advantageous in various applications of structural, functional and comparative genomics, including variety identification, studying genetic diversity and phylogenetic relationships, construction of high-density genome maps, mapping of useful genes, comparative genome mapping and marker-assisted selection.[10–15] Till now, a total of 218 genic and genomic microsatellite markers have been developed from Expressed Sequence Tag sequences (26)[16], random amplified polymorphic DNA (RAPD)-enriched genomic sequences from GenBank (45)[14] and microsatellite-enriched genomic libraries (147)[15] in foxtail millet. The availability of limited EST sequence resources and involvement of huge cost, time and labor in mining microsatellite markers from enriched genomic libraries have impeded the development of large-scale microsatellite markers in foxtail millet. However, foxtail millet with large genome size would require a large number of microsatellite markers for construction of high-density genetic linkage map and identification and mapping of genes/quantitative trait loci (QTLs) for useful agronomic traits.

A large number of microsatellite markers have been developed by utilizing the pseudomolecule genomic sequences from the completely sequenced plant genomes, including rice (IRGSP 2005, http://rgp.dna.affrc.go.jp/IRGSP/), *Populus* (genome.jgi.doe.gov/poplar/), sorghum (genome.jgi-psf.org/Sorbi1/Sorbi1.home.htm), maize (mips.helmholtz-muenchen.de/plant/maize/) and *Arabidopsis thaliana* (*Arabidopsis* Genome Initiative 2001, http://www.arabidopsis.org/), and these markers were further utilized in genome mapping and genotyping applications in target crop species. A similar approach has been implemented in rice, sorghum, *Brachypodium*, *Brassica* and *Arabidopsis* for the development of thousands of microsatellite markers from whole genome sequences.[12,13,17–19] Unfortunately, no such marker resource is currently available for recently sequenced foxtail millet genome. With the availability of foxtail millet whole genome sequence in public domain, it is now possible to mine and develop large-scale *in silico* microsatellite markers that could be utilized for various applications of structural and comparative genomics in foxtail millet. The development and large-scale validation of such genomic microsatellite markers in a genome-wide scale could be useful for not only genetic improvement of foxtail millet, but also for other under-utilized/orphan crop species for which very little or no genomic information is available.

Considering the utility of whole genome sequence-based microsatellite markers, we made an attempt to develop and validate microsatellite markers from foxtail millet genome sequences and determine their genomic distribution on foxtail millet genome. Furthermore, a selected set of validated and physically mapped markers were utilized in evaluating polymorphic and diversity potential in foxtail millet accessions, studying cross-transferability across millet and non-millet species and *in silico* comparative genome mapping among foxtail millet and three members of grass family.

2. Materials and methods

2.1. *Identification of microsatellites and primer design*

The whole genome sequences of *S. italica* were retrieved from Phytozome (http://www.phytozome.net) and searched for microsatellites using the default parameter of MIcroSAtellite (MISA) identification tool (http://pgrc.ipk-gatersleben.de/misa). The search criteria were: six repeat units for dinucleotide repeats (DNRs), five repeat units for trinucleotide repeats (TNRs), tetranucleotide repeats (TeNRs), pentanucleotide repeats (PNRs) and hexanucleotide repeats (HNRs). The forward and reverse primers flanking the identified microsatellite repeat motifs were designed in batches using the default parameters of integrated Perl 5 interface module of MISA-Primer3 software (http://pgrc.ipk-gatersleben.de/misa/primer3.html).

2.2. *Physical mapping of foxtail millet genomic microsatellite markers*

For determining the specific physical location (bp) of designed microsatellite markers on the nine foxtail millet chromosomes, the flanking genomic sequences of identified microsatellite repeat motifs were BLAST searched against the whole genome sequence of foxtail millet at Phytozome (http://www.phytozome.net). Optimized BLASTN search parameters with E-value = 0 and low-complexity filter options were used for this analysis. The microsatellite

markers were plotted separately for each of the nine foxtail millet chromosomes according to their ascending order of physical position (bp), from the short arm telomere to the long arm telomere and finally visualized in MapChart software.[20]

2.3. Validation of genomic microsatellite markers

For validation of selected set of physically mapped microsatellite markers distributed over nine chromosomes of foxtail millet, the genomic DNA was isolated from the leaves of eight foxtail millet accessions (Table 1) using CTAB method. The isolated genomic DNA was resolved on 0.8% agarose gel (Cambrex, USA) in Tris-borate EDTA (TBE) buffer (pH 8.0), quantified using different concentration of λ DNA standard and analysed in the GelDoc-It imaging system (UVP, Cambridge, UK). The polymerase chain reaction (PCR) was carried out in MyCycler thermal controller (Bio-Rad) in a 25-μl total volume containing one unit of Taq DNA polymerase (Bangalore Genei), 50 ng of genomic DNA, 10 μmol/l of each primer, 0.5 mmol/l of each dNTPs and 2.5 μl of 10× PCR buffer (Bangalore Genei). DNA amplification was done following the PCR program; a preliminary denaturation of 3 min at 94°C, followed by 40 cycles of 1 min at 94°C, 1.5 min at 48–55°C and 2 min at 72°C and a final extension for 10 min at 72°C. The amplicons were resolved on 2% agarose gel (Cambrex, USA) in TBE buffer (pH 8.0), and analysed in the GelDoc-It imaging system. Standard size marker of 100 bp (NEB) was used to determine the fragment sizes for each markers. One of the markers showing amplification in all the accessions used in the study was cloned and sequenced. For cloning, the amplicons amplified by microsatellite marker were eluted using *AccuPrep* DNA Gel Purification Kit (Bioneer, Korea) and cloned into pGEM-T Easy vector (Promega, USA) following the manufacturer's instructions. The plasmids were transformed into competent cells of *Escherichia coli* DH5α and grown on LB media supplemented with ampicillin (100 μg/ml), X-Gal (20 μg/ml) and IPTG (0.1 M). The transformants were selected based on blue–white screening, and recombinant plasmids were isolated using *AccuPrep* Plasmid MiniPrep DNA Extraction Kit following the manufacturer's protocol. The plasmids were sequenced in automated sequencer (3730xl DNA Analyzer, Applied Biosystem) using M13 forward and reverse primers. Multiple sequence alignment was performed with the obtained sequences along with reference *S. italica* sequence using ClustalW2 program (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

Table 1. Details of plant materials used in the present study

Serial number	Common name	Genus/species	Accessions
1.	Foxtail millet	*S. italica*	cv. Prasad
2.	Foxtail millet	*S. italica*	cv. Lepakshi
3.	Foxtail millet	*S. italica*	IC403579
4.	Foxtail millet	*S. italica*	IC480117
5.	Foxtail millet	*S. italica*	GS2038
6.	Green millet	*S. italica* subsp.viridis	EC539251
7.	African bristlegrass	*S. sphacelata*	EC539291
8.	Bristly foxtail	*S. verticillata*	EC539293
9.	Barnyard millet	*Echinochloa frumentacea*	CO-LV2
10.	Finger millet	*Eleusine coracana*	CO-RA114
11.	Kodo millet	*Paspalum scrobiculatum*	C0-3
12.	Little millet	*Panicum sumatrense*	C0-4
13.	Pearl millet	*Pennisetum glaucum*	CO-CU 9
14.	Proso millet	*Panicum miliaceum*	CO-5
15.	Switchgrass	*P. virgatum*	PI421521
16.	Guinea grass	*Panicum maximum*	SPM92
17.	Sorghum	*Sorghum bicolor*	CO30
18.	Wheat	*Triticum aestivum*	PH132
19.	Rice	*Oryza sativa*	cv. Pusa Basmati
20.	Maize	*Zea mays*	B73

2.4. Assessment of genetic diversity pattern and phylogenetic relationships

The microsatellite alleles amplified in millets and non-millets were scored in a binary format to infer the molecular diversity and genetic relationship among them. The genetic similarity coefficient was calculated according to Jaccard's coefficient[21] using the NTSYS-pc software package (version 2.02e, Exeter Software, Setauket, NY, USA).[22] The pairwise similarity matrix of the Jaccard coefficient was used to construct a phylogenetic tree using the neighbor-joining module of NTSYS-pc software.

2.5. Comparative genome mapping between foxtail and other grass species

The flanking sequences of microsatellite marker loci that were physically mapped on the nine chromosomes of foxtail millet were BLAST searched against genome sequences of sorghum, maize and rice (http://gramene.org/; www.phytozome.net) to derive marker-based syntenic relationships among the chromosomes of foxtail millet and three other grass species. A cutoff bit score of 54.7 and E-value of $<1e-05$ were considered significant for BLAST analysis. The marker-based syntenic relationships

among foxtail millet, rice, sorghum and maize were finally visualized with visualization blocks using Circos 0.55 (http://circos.ca).

3. Results and discussion

3.1. *Frequency, distribution and characterization of microsatellites in foxtail millet genome*

The 405.73 Mb available genome sequences of foxtail millet were searched for microsatellites, and a total of 28 342 microsatellites comprising different kinds of desirable repeat-motifs (from DNRs to HNRs) were identified with average frequency of about 69 microsatellites per megabase of genomic sequences. The frequency and distribution of different microsatellite repeat-motif types in the genomic sequences of foxtail millet are shown in Fig. 1A and Supplementary Table S1. The frequency of microsatellites (69/Mb) estimated in the foxtail millet genomic sequences at genome-wide level is comparable to that documented in *Brachypodium* (78.5/Mb) genome, but lower than that reported earlier in other related monocot plant genomes such as rice (189.4/Mb) and sorghum (99.8/Mb) and dicot *Arabidopsis* (127.5/Mb).[19] These differences could be due to the variation in search criteria, size of the database and bioinformatics software tools used in different studies for identification of microsatellites.[10,11,15,23]

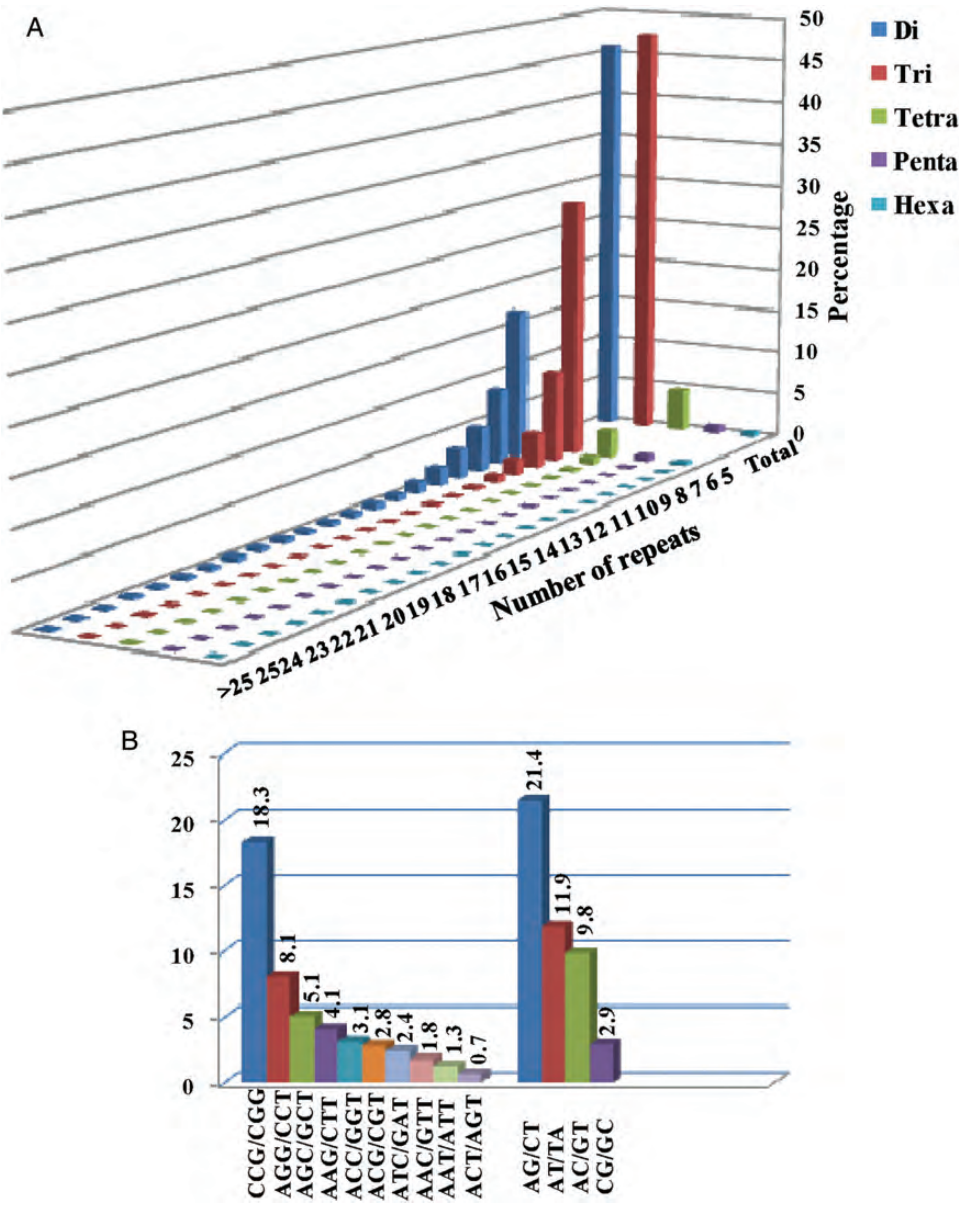

Figure 1. Analysis of simple sequence repeats from foxtail millet genome. (A) Relative frequency, proportion (%) and number of selected microsatellite repeat-motif types. (B) Different classes of selected microsatellite repeats.

Among the 28 342 microsatellites obtained, the TNRs were the most abundant (13 521) with a proportion of 47.7%, followed by DNRs (13072, 46.1%), TeNRs (1354, 4.8%), PNRs (236, 0.8%) and HNRs (159, 0.56%) (Fig. 1A). The distribution of different microsatellite motifs is not uniform, particularly as in case of TNRs and DNRs. Among DNRs, AG/CT motifs (21.4%) were more frequent, followed by AT/AT (11.9%), AC/GT (9.8%) and GC/CG (2.9%) motifs. Among the TNRs, CCG/CGG (18.3%) motifs were most abundant followed by AGG/CCT (8.1%) and AGC/GCT (5.1%), whereas ACT/AGT (0.7%) motifs were less abundant (Fig. 1B). Based on length of the repeat motifs, a total of 7387 (26%) microsatellites were classified as long and hypervariable class I ($\geq$20 bp) types and remaining 20 995 (74%) microsatellites as variable class II (12–19 bp) types. Interestingly, as in case of class I microsatellites, the proportion of DNRs (47%) was higher when compared with TNRs (30%) (Fig. 2). In contrast, in the class II microsatellite types, the proportion of TNRs (54%) was more than that of DNRs (46%). This converse distribution pattern of microsatellite repeat motifs in the foxtail millet genome overall resulted in slightly higher frequency of TNRs when compared with DNRs in this study. Generally, DNRs are reported to be abundant in sequenced plant genomes such as rice, sorghum and *Arabidopsis*.[19,24,25] Conversely, in our study, we observed an abundance of TNRs (47.7%) in foxtail millet genome that agreed well with the earlier reports in *Brachypodium*,[19] bamboo[26] and switchgrass.[27] The earlier reports of higher frequency of TNRs in the transcribed genic sequence components of the plant genomes suggested that the genomic sequences used in this study for mining microsatellites possibly representing more of expressed sequence components of the foxtail millet genome.[10–13] Furthermore, the prevalence of AG/CT (21.4%) dinucleotide and CCG/CGG (18.3%), AGG/CCT (8.1%) and AGC/GCT (5.1%) trinucleotide repeat motifs in the foxtail millet genome is comparable to earlier similar genome-wide microsatellite identification studies in monocot and dicot plant species such as *Brachypodium*, rice, sorghum and bamboo.[19,26]

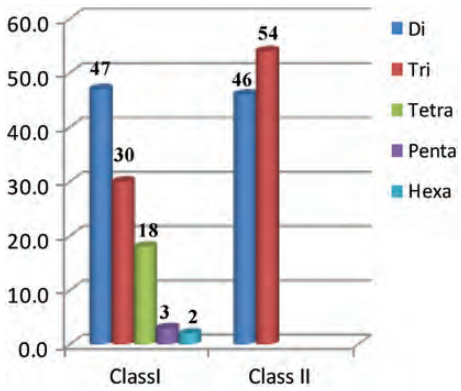

Figure 2. Frequency and relative distribution of long and hypervariable class I and variable class II microsatellite repeats in the foxtail millet genome.

3.2. *Development of microsatellite markers and their physical mapping in foxtail millet genome*

From 28 342 identified microsatellites, forward and reverse primer pairs could be designed from either side of the flanking genomic sequences of 21 294 microsatellites with successful primer designing potential of 75.1%. The primers could not be designed for remaining 7048 (24.9%) microsatellites due to constraint of obtaining enough flanking sequences from either side of identified microsatellites. Similar observations have also been observed in other genome-wide microsatellite mining studies in monocot and dicot crop plants.[12,13,19]

The determination of genomic distribution of 21 294 microsatellite markers on the foxtail millet genome revealed physical localization of 15 573 markers on the 9 chromosomes of foxtail millet with average marker density of 42 markers/Mb (Supplementary Table S2, Supplementary Fig. S1). All the physically mapped 15 573 genomic microsatellite markers were reported in publicly available NCBI Probe Database with Accession numbers from PUID16589829 to PUID 16605401. The average marker density was maximum 46.2/Mb in chromosome 9, followed by 42.8/Mb in chromosome 5 and 41.6/Mb in chromosome 1 and minimum 30/Mb in chromosome 8. The comprehensive analysis of chromosome-wise distribution and frequency of these physically mapped microsatellite markers showed higher frequency of physically mapped markers on chromosome 9 (2719 markers, 17.5%) and minimum in chromosome 8 (1222, 7.8%) with average physical gap size of ~24 kb between corresponding mapped markers (Table 2).

Table 2. Summary of chromosomal distribution and average density of microsatellite markers mapped on the nine chromosomes of foxtail millet

Chromosome	Markers mapped (%)	Density (markers/Mb)
Chr. 1	1751 (11.3)	41.59
Chr. 2	1929 (12.4)	39.21
Chr. 3	1969 (12.6)	38.91
Chr. 4	1352 (8.8)	33.55
Chr. 5	2019 (12.9)	42.78
Chr. 6	1235 (7.9)	34.31
Chr. 7	1377 (8.8)	38.36
Chr. 8	1222 (7.8)	30.02
Chr. 9	2719 (17.5)	46.16

These observations suggested that the high-density microsatellite marker-based physical map constructed in this study for foxtail millet genome could be useful for rapid selection of genome-wide microsatellite markers well distributed over nine chromosomes for various large-scale genotyping applications, including mapping of genes and genomes and comparative genome mapping involving foxtail millet and other related crop plants.

3.3. *Amplification and polymorphic potential of microsatellite markers*

A selected set of 159 genomic microsatellite markers (148 from class I and 11 from class II), distributed over 9 chromosomes of foxtail millet, gave clear, successful and reproducible amplification with expected product size (bp) in *S. italica* cv. Prasad. It indicated about 100% amplification potential of designed genomic microsatellite markers, and, thus, the developed 15 573 microsatellite markers for foxtail millet genome are a useful resource for foxtail millet genomics and molecular breeding. One hundred fifty-two (95.6%) of the 159 microsatellite markers amplified unique single allele, whereas 7 markers amplified more than 1 allele/multiple alleles. Thus, a total of 167 alleles were amplified by 159 genomic microsatellite markers in *S. italica* cv. Prasad (Supplementary Table S3). Furthermore, 132 out of 159 microsatellite markers produced fragments of desired sizes, whereas the amplification products for 27 markers differed from the expected sizes. This inconsistency may be either due to the use of different accessions of foxtail millet in this study and sequenced genome or due to sequencing error.

To evaluate polymorphism and molecular diversity potential of developed microsatellite markers, the validated and physically mapped 159 microsatellite markers were amplified in a set of 7 accessions of *Setaria*, including 4 cultivated and 3 wild species. One hundred seven (67.2%) of the 159 genomic microsatellite markers showed polymorphism in 8 cultivated and wild accessions of foxtail millet. A total of 342 alleles ranging from 1 to 6 alleles were amplified by 159 microsatellite markers in 8 foxtail accessions with an average 2.15 alleles per marker locus (Supplementary Table S3; Fig. 3). The polymorphic potentials of these 159 microsatellite markers were further evaluated in the genomic DNA of 2 parental foxtail millet accessions, Prasad and Lepakshi of an F_2 mapping population (cv. Prasad x cv. Lepakshi). Thirty (~19%) of the 159 markers showed polymorphism between parental accessions of F_2 mapping population. It includes 20 (21.5%), 8 (14.8%) and 2 (16.6%) DNR, TNR and TeNR motifs containing markers, respectively (Supplementary Table S3). Remarkably, all the 30 markers that showed parental polymorphism belonged to class I microsatellite types. This clearly suggests the added advantage of long hypervariable class I microsatellites over variable class II microsatellites for revealing higher degree of polymorphism in the foxtail millet accessions and, thus, could be utilized for large-scale genotyping applications. Such observations have also been reported in other genome-wide studies in plant species using a larger set of class I and class II microsatellite markers.[13,15,28] The polymorphic potential (67.2%) of genomic microsatellite markers estimated among foxtail millet cultivated and wild accessions was higher than that reported using the microsatellite markers derived from foxtail millet genomic sequences.[14,15] The higher polymorphic potential of genomic microsatellite markers is expected because of their derivation from non-coding sequence components of the foxtail genome.[10,12,15] Besides, the validated physically mapped markers could enable to discriminate all the eight cultivated and wild foxtail millet accessions from each other with a level diversity from 32 to 59% (Supplementary Fig. S2). The distant clustering patterns of *S. italica* subsp. viridis and other related *Setaria* wild species such as *Setaria sphacelata* and *Setaria verticillata* from the cultivated *Setaria* species were expected. The clustering patterns derived among cultivated and wild *Setaria* species corresponded well with their phylogenetic relationships and taxonomic classifications. Moreover, this accords well with the earlier studies conducted using ISSR and transposon-based markers.[29,30] The microsatellite markers, thus, showed differentiation between cultivated and wild *Setaria* species that could be useful for introgression breeding for transferring genes/alleles of agronomic importance like stress tolerance from the wild

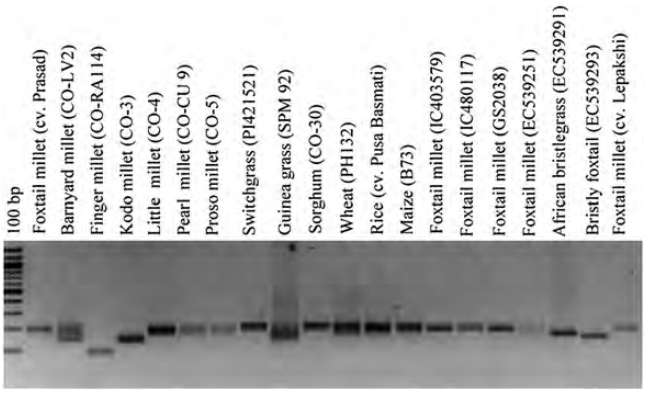

Figure 3. Representative gel showing amplification profiles of one microsatellite marker SiGMS 3261 and its fragment length polymorphism among foxtail millet and related species. The amplicons are resolved in 2% agarose gel along with 100 bp DNA size standard.

species to the cultivated genetic backgrounds of foxtail millet. Therefore, the genomic microsatellite markers with high amplification and polymorphic potential distributed over nine chromosomes of foxtail millet genome designed in the study could be useful for many large-scale genotyping applications in foxtail millet.

3.4. *Cross-genera transferability of microsatellite markers*

Of the 159 genomic microsatellite markers validated, 58 markers were randomly selected to assess their utility for cross-genera transferability among millet and non-millet species. Of the 58 microsatellite markers assayed, 53 (91.3%) gave consistent amplification in Barnyard millet, 55 (94.8%) in finger millet, 51 (87.9%) in kodo millet, 55 (94.9%) in little millet, 54 (93.1%) in pearl millet, 51 (87.9%) in proso millet, 46 (79.3%) in switchgrass, 57 (98.2%) in guinea grass, 52 (89.6%) in sorghum, 42 (71.2%) in wheat, 51 (87.9%) in rice and 53 (91.4%) in maize with an average cross-genera transferability potential of ~89% (Fig. 3). All the 58 microsatellite markers have ability to distinguish the investigated 13 millet and non-millet species into 2 distinct groups (Fig. 4). The obtained high average cross-genera transferability potential (~89%) of genomic microsatellite markers is consistent with our previous study, where also a higher level of transferability (73%) of foxtail millet genomic microsatellite markers across various millet and non-millet species was observed.[15] Apart from this, our results are in close agreement to several previous studies reporting high cross-species transferability of genomic microsatellite markers from *Lolium* to *Festuca* complex (~71%),[31] rice to bamboo (~68.3%)[32] and groundnut to other legumes (~71%).[33] The high degree of cross-transferability of genomic microsatellite markers across related foxtail millet species could be due to more representation of microsatellite markers that were derived from the conserved expressed component of the foxtail millet genome sequences and, thus, resulting in higher transferability of microsatellite markers like genic microsatellite markers across closely related and even distantly related genera.[10–12] The sequencing of PCR products of microsatellite markers amplifying multiple alleles in millet and non-millet species revealed the presence of expected microsatellite repeat motifs in foxtail millet as predicted *in silico* (Supplementary Fig. S3). The fragment length polymorphism among the accessions explained the variation in the length of the microsatellite repeats and the presence of point mutations, including insertion/deletions (Supplementary Fig. S3). Therefore, based on amplicon sequencing, the microsatellite markers showing cross-transferability across millet and non-millet species were valid and, thus, have utility in comparative genome mapping studies in both millets and non-millets. Besides, the markers showing high-cross transferability across millet and non-millet species can be utilized for genetic enhancement of orphan crops using the marker-related genetic information of foxtail millet genome.

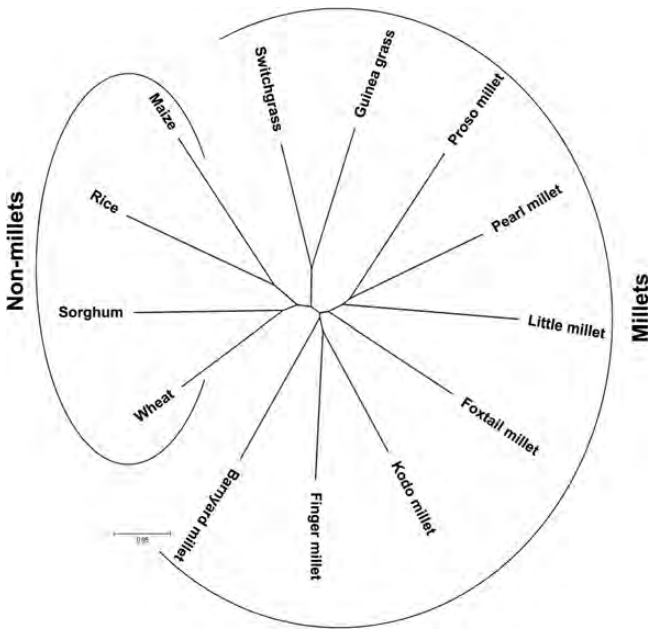

Figure 4. Genetic relationships among 9 millet and 4 non-millet grass species based on 58 foxtail millet microsatellite markers using neighbor-joining clustering. Nine millet species, including foxtail millet, were clearly differentiated from the four non-millet grass species, and expected genetic relationships among species under study were also evident.

3.5. In silico *comparative genome mapping between foxtail millet and other grass species*

For comparative mapping, the physically mapped 15 573 microsatellite markers on the 9 chromosomes of foxtail millet were compared with their physical location on the chromosomes of other related grass genomes namely sorghum, maize and rice (Table 3; Fig. 5A–C; Supplementary Tables S4–S6). *In silico* comparative genome mapping revealed maximum proportion of sequence-based orthology and, thus, syntenic relationship of microsatellite markers distributed over 9 foxtail millet chromosomes with 10 chromosomes of sorghum (16.9%, 2631) and maize (14.5%, 2252) and 12 rice (6.4%, 1005) chromosomes. Many of the physically mapped microsatellite markers of foxtail millet also showed syntenic relationships with more than one chromosome of sorghum, maize and rice, reflecting the occurrence of segmental duplication events

Table 3. A summary of microsatellite marker-based comparative mapping showing maximum syntenic relationships of foxtail millet chromosomes with sorghum, maize and rice chromosomes

Foxtail millet chromosomes	Sorghum chromosomes	Maize chromosomes	Rice chromosomes
Chr. 1	Chr. 4 (309, 92.5%)	Chr. 5 (184, 62.5%) and Chr. 4 (85, 28.9%)	Chr. 2 (131, 86.1%)
Chr. 2	Chr. 2 (271, 86.9%)	Chr. 7(155, 54.1%) and Chr. 2 (90, 31.4%)	Chr. 9 (45, 46.8%) and Chr. 7 (31, 32.3%)
Chr. 3	Chr. 9 (157, 49.8%) and Chr. 8 (85, 26.9%)	Chr. 6 (83, 33.1%), Chr. 8 (42, 16.7%) and Chr. 10 (28, 11.1%)	Chr. 5 (55, 50%) and Chr. 12 (37, 33.6%)
Chr. 4	Chr. 10 (169, 84.07%)	Chr. 9 (75, 42.1%), Chr. 6 (51, 28.6%) and Chr. 5 (28, 15.7%)	Chr. 6 (37, 69.8%)
Chr. 5	Chr. 3 (332, 88.06%)	Chr. 3 (206, 60.2%) and Chr. 8 (93, 27.2%)	Chr. 1 (163, 89.5%)
Chr. 6	Chr. 7 (140, 76.5%)	Chr. 1 (56, 34.3%) and Chr. 4 (51, 31.3%)	Chr. 8 (35, 67.3%)
Chr. 7	Chr. 6 (185, 71.7%)	Chr. 2 (99, 43.8%) and Chr. 10 (74, 32.7%)	Chr. 4 (68, 64.1%)
Chr. 8	Chr. 5 (36, 48.6%)	Chr. 4 (50, 48.1%)	Chr. 1 (39, 79.6%) and Chr. 11 (29, 59.2%)
Chr. 9	Chr. 1 (495, 86.9%)	Chr. 1 (262, 64.2%), Chr. 9 (92, 22.5%) and Chr. 5 (64, 15.6%)	Chr. 3 (151, 73.7%), Chr. 1 (41, 20%) and Chr. 10 (31, 15.1%)

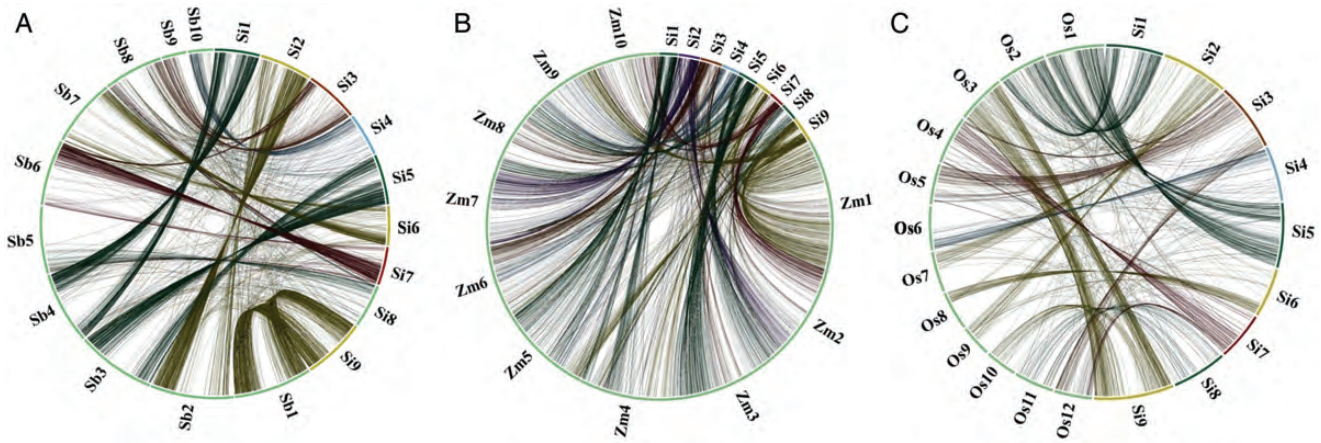

Figure 5. Genome relationships of foxtail millet with other grass species. Syntenic relationship of foxtail millet genome with (A) sorghum, (B) maize and (C) rice chromosomes using 15 573 physically mapped foxtail millet microsatellite markers. Maximum syntenic relationships of foxtail millet chromosomes with sorghum chromosomes based on microsatellite markers were apparent.

between the chromosomal stretches of respective genomes. This is expected considering the presence of higher number (10–12) of chromosomes in rice, maize and sorghum when compared with 9 chromosomes of foxtail millet. A number of foxtail millet microsatellite markers showed significant matches with different chromosomes of sorghum, maize and rice.

Foxtail millet–sorghum synteny: between foxtail millet and sorghum genomes, 2631 microsatellite marker loci distributed over 9 chromosomes of foxtail millet showed significant matches with 2635 genomic regions of 10 sorghum chromosomes (Table 3; Fig. 5A). All the nine foxtail millet chromosomes showed considerable and higher average frequency (16.9%) of microsatellite marker-based syntenic relationship with specific sorghum chromosomes. The physically mapped microsatellite markers on the foxtail millet chromosome 1 (309) showed maximum synteny (92.5%) with sorghum chromosome 4, followed between foxtail millet chromosome 5 and sorghum chromosome 3 (88.1%), between foxtail millet chromosome 2 and sorghum chromosome 2 (86.9%), between foxtail millet chromosome 9 and sorghum chromosome 1 (86.9%) and minimum between foxtail millet chromosome 8 and

sorghum chromosome 5 (48.6%) (Supplementary Table S4).

Foxtail millet–maize synteny: the comparative mapping between foxtail millet and maize genomes revealed syntenic relationship of 2252 microsatellite marker loci distributed over 9 chromosomes of foxtail millet with 2337 genomic regions on 10 chromosomes of maize (Table 3; Fig. 5B). On an average, 14.5% syntenic relationship of microsatellite marker loci between foxtail millet and maize chromosomes was observed. The syntenic relationships of microsatellite marker loci were maximum between foxtail millet chromosome 9 with maize chromosome 1 (64.2%), followed between foxtail millet chromosome 1 and maize chromosome 5 (62.5%), between foxtail millet chromosome 5 and maize chromosome 3 (60.2%) and minimum between foxtail millet chromosome 3 and maize chromosome 6 (33.1%) (Supplementary Table S5).

Foxtail millet–rice synteny: the microsatellite markers physically mapped on the foxtail millet chromosomes showed least synteny with 12 rice chromosomes with average frequency of 6.4% that is much lower than that with the sorghum and maize chromosomes (Table 3; Fig. 5C). Only 1005 foxtail microsatellite marker loci showed significant matches with 1130 genomic regions spanning over 12 chromosomes of rice genome (Supplementary Table S6). Maximum synteny of microsatellite marker loci between foxtail millet chromosome 5 and rice chromosome 1 (89.5%) and minimum between foxtail millet chromosome 2 and rice chromosome 9 (46.8%) was observed.

Previous reports showed the mapping of similar microsatellite markers either genetically or physically on orthologous or syntenic chromosomes of different related plant genomes.[1,6,34–36] Comparative microsatellite marker-based genome mapping revealed higher degree of synteny between foxtail and sorghum genome, followed by maize and rice that is expected due to their taxonomic relationship in which foxtail millet, sorghum and maize belong to the same subfamily Panicoideae, whereas rice belongs to Ehrhartoideae. This result also indicates declination of synteny with increasing phylogenetic distance among plant species. The syntenic relationships inferred among the chromosomes of foxtail millet and three cereal species (sorghum, maize and rice) in the present study agreed well with the earlier study of Devos.[35] However, using the genome-wide microsatellite markers, the present study generated few previously unobserved interesting information like syntenic relationships between foxtail chromosome 9 and rice chromosome 3 (73.7%) and foxtail millet chromosome 8 with rice chromosome 1 (79.5%) that was not reported earlier.[35]

Our results are in accordance with our earlier study, where we reported ~18% sequence homology of microsatellite markers in the foxtail millet with sorghum, ~16% with maize and ~5% with rice.[15] In the synteny analysis, the observed high level of similarity provides an idea on genome conservation among grasses. Chromosomal rearrangements such as duplication, inversion and translocation were also evidenced as many of foxtail millet marker loci that were mapped to more than one chromosome of analysed species. Although our results do not provide a whole genome view due to the limitations of regions studied and biasness of markers used, our comparative studies have been insightful in understanding the evolutionary process among grasses involving the foxtail millet genome. Therefore, further investigation is required to evaluate the detailed level of collinearity between sorghum, maize and rice with foxtail millet genomes. The microsatellite marker-based comparative genome mapping between foxtail millet and other diploid grass species such as sorghum, maize and rice would be useful in map-based isolation of genes of agronomic importance from foxtail millet using the marker-based genotyping information of other related small and diploid grass members.

4. Conclusions

In the present study, we conducted a genome-wide analysis and identified 28 342 microsatellite motifs in the foxtail millet. A total of 21 294 primer pairs were designed, of which 15 573 microsatellites were physically mapped on 9 chromosomes of foxtail millet genome, and 159 markers were validated successfully in foxtail millet accessions. The microsatellite markers developed here showed their utility in high amplification and polymorphic potential, cross-genera transferability and *in silico* comparative physical mapping. These genomic microsatellite markers reported here would be of enormous and immense use for various large-scale genotyping applications, including germplasm characterization, cultivar identification, construction of high-density microsatellite marker-based physical linkage map for gene/QTL discovery and comparative genome mapping involving foxtail millet and other millets, cereals and bioenergy grass species. More importantly, the constructed high-density physical map in this study could act as reference for future studies of researchers for rapid selection of microsatellite markers either across genome-wide/chromosomal level or at targeted chromosomal location. Thus, it would expedite the construction of high-density genetic linkage map and fine mapping of genes/QTLs for important agronomic traits in foxtail millet.

Acknowledgements: Grateful thanks are due to the Director, National Institute of Plant Genome Research (NIPGR), New Delhi, India for providing facilities. Ms Garima Pandey acknowledges the award of Junior Research Fellowship from University Grants Commission, New Delhi. We are also thankful to the National Bureau of Plant Genetic Resources, New Delhi/Hyderabad/Akola, India and Tamil Nadu Agricultural University, India for providing the seed material. We thank Mr Mehanathan Muthamilarasan, NIPGR, New Delhi for critically reading the manuscript. Computational resources under BTISNET program (DISC facility) are gratefully acknowledged.

Conflict of Interest statement

None declared.

Supplementary Data: Supplementary data are available at www.dnaresearch.oxfordjournals.org.

Funding

This work was financially supported by the core grant of National Institute of Plant Genome Research (NIPGR), New Delhi, India.

References

1. Devos, K.M. and Gale, M.D. 2000, Genome relationships: the grass model in current research, *Plant Cell*, **12**, 637–46.
2. Doust, A.N., Kellogg, E.A., Devos, K.M. and Bennetzen, J.L. 2009, Foxtail millet: a sequence-driven grass model system, *Plant Physiol.*, **149**, 137–41.
3. Lata, C., Gupta, S. and Prasad, M. 2012, Foxtail millet: a model crop for genetic and genomic studies in bioenergy grasses, *Crit. Rev. Biotechnol.* DOI: 10.3109/07388551.2012.716809. (September 18, 2012).
4. Li, P. and Brutnell, T.P. 2011, *Setaria viridis* and *Setaria italica*, model genetic systems for Panicoid grasses, *J. Exp. Bot.*, **62**, 3031–37.
5. Bennetzen, J.L., Wang, J.S.H., Percifield, R., et al. 2012, Reference genome sequence of the model plant *Setaria*, *Nature Biotech.*, **30**, 555–61.
6. Zhang, G., Liu, X., Quan, Z., et al. 2012, Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential, *Nat. Biotechnol.*, **30**, 549–54.
7. Mitros, T.K., Bennetzen, J.L., Schmutz, J. and Rokhsar, D.S. 2010, The foxtail millet genome and its relationship amongst the grasses. In: *Plant and Animal Genomes XVIII Conference*, January 9–13, San Diego, CA. W328 (abstract).
8. Delseny, M., Laroche, M. and Penon, P. 1983, Detection of sequences with Z-DNA forming potential in higher plants, *Biochem. Biophys. Res. Commun.*, **116**, 113–20.
9. Tautz, D. and Renz, M. 1984, Simple sequences are ubiquitous repetitive components of eukaryotic genomes, *Nucleic Acids Res.*, **12**, 4127–38.
10. Varshney, R.K., Graver, A. and Sorrells, M.E. 2005, Genic microsatellite markers in plants: features and applications, *Trends Biotechnol.*, **23**, 48–55.
11. Parida, S.K., Rajkumar, K.A., Dalal, V., Singh, N.K. and Mohapatra, T. 2006, Unigene derived microsatellite markers for the cereal genomes, *Theor. Appl. Genet.*, **112**, 808–17.
12. Parida, S.K., Kalia, S.K., Sunita, K., et al. 2009, Informative genomic microsatellite markers for efficient genotyping applications in sugarcane, *Theor. Appl. Genet.*, **118**, 327–38.
13. Parida, S.K., Yadav, D.K. and Mohapatra, T. 2010, Microsatellites in *Brassica* unigenes: relative abundance, marker design and use in comparative physical mapping and genome analysis, *Genome*, **53**, 55–67.
14. Lin, H.S., Chiang, C.Y., Chang, S.B. and Kuoh, C.S. 2011, Development of simple sequence repeats (SSR) markers in *Setaria italica* (Poaceae) and cross-amplification in related species, *Int. J. Mol. Sci.*, **12**, 7835–45.
15. Gupta, S., Kumari, K., Sahu, P.P., Vidapu, S. and Prasad, M. 2012, Sequence based novel genomic microsatellite markers for robust genotyping purposes in foxtail millet [*Setaria italica* (L.) P. Beauv], *Plant Cell Rep.*, **31**, 323–37.
16. Jia, X.P., Shi, Y.S., Song, Y.C., Wang, G.Y., Wang, T.Y. and Li, Y. 2007, Development of EST-SSR in foxtail millet (*Setaria italica*), *J. Genet. Res. Crop Evol.*, **54**, 233–6.
17. Singh, H., Deshmukh, R.K., Singh, A., Singh, A.K. and Gaikwad, K. 2010, Highly variable SSR markers suitable for rice genotyping using agarose gels, *Mol. Breed.*, **25**, 359–64.
18. Sayama, T., Hwang, T.-Y., Komatsu, K., et al. 2011, Development and application of a whole-genome simple sequence repeat panel for high-throughput genotyping in soybean, *DNA Res.*, **18**, 107–15.
19. Sonah, H., Deshmukh, R.K., Sharma, A., et al. 2011, Genome-wide distribution and organization of microsatellites in plants: an insight into marker development in *Brachypodium*, *PLoS One*, **6**, e21298. doi:10.1371/journal.pone.0021298 ma1.
20. Voorrips, R.E. 2002, MapChart: software for the graphical presentation of linkage maps and QTLs, *J. Hered.*, **93**, 77–8.
21. Jaccard, P. 1908, Nouvelle recherches sur la distribution florale, *Bull Soc. Vaud. Sci. Nat.*, **44**, 223–70.
22. Rohlf, F.J. 1997, NTSYS-pc numerical taxonomy and multivariance analysis system. Version 2.02e, Exeter Software, Setauket, NY.
23. Kalia, R.K., Rai, M.K., Kalia, S., Singh, R. and Dhawan, A.K. 2011, Microsatellites markers: an overview of the recent progress in Plants, *Euphytica*, **177**, 309–34.
24. Morgante, M. and Olivieri, A.M. 1993, PCR-amplified microsatellites as markers in plant genetics, *Plant J.*, **3**, 175–82.

25. Sharma, P.C., Grover, A. and Kahl, G. 2007, Mining microsatellites in eukaryotic genomes, *Trends Biotechnol.*, **25**, 490–8.
26. Yasodha, R. 2011, Characterization of microsatellites in the tribe bambuseae, *Gene Conserv.*, **10**, 51–64.
27. Sharma, M.K., Sharma, R., Cao, P., et al. 2012, A genome-wide survey of switchgrass genome structure and organization, *PLoS One*, **7**, e33892. doi:10.1371/journal.pone.0033892.
28. Temnykh, S., Declerk, G., Lukashover, A., Lipovich, L., Cartinhour, S. and McCouch, S. 2001, Computational and experimental analysis of microsatellites in rice (Oryza sativa L.): frequency, lengthvariation, transposon associations and genetic marker potential, *Genome Res.*, **11**, 1441–52.
29. Wei, L., Hui, Z., Yong-fang, W., Hai-quan, L. and Xian-min, D. 2012, Assesment of genetic relationship of foxtail millet with its wild ancestor and close relatives by ISSR markers, *J. Int. Agri.*, **11**, 556–66.
30. Hirano, R., Naito, K., Fukunaga, K., Watanabe, K.N., Ohsawa, R. and Kawase, M. 2011, Genetic structure of landraces in foxtail millet (Setaria italica (L.) P. Beauv.) revealed with transposon display and interpretation to crop evolution of foxtail millet, *Genome*, **54**, 1–9.
31. Jones, E.S., Dupal, M.P., Kollier, R., Drayton, M.C. and Forster, J.W. 2001, Development and characterization of simple sequence repeat (SSR) markers for perennial ryegrass (*Lolium perenne*. L.), *Theor. Appl. Genet.*, **102**, 405–15.
32. Chen, S.Y., Lin, Y.T., Lin, C.W., Chen, W.Y., Yang, C.H. and Ku, H.M. 2010, Transferability of rice SSR markers to bamboo, *Euphytica*, **175**, 23–33.
33. Gautami, B., Ravi, K., Narasu, M.L., Hoisington, D.A. and Varshney, R.K. 2009, Novel set of groundnut SSRs for genetic diversity and inter-specific transferability, *Int. J. Integr. Biol.*, **7**, 100–6.
34. Devos, K.M., Wang, Z.M., Beales, J., Sasaki, T. and Gale, M.D. 1998, Comparative genetic maps of foxtail millet (*Setaria italica*) and rice (*Oryza sativa*), *Theor. Appl. Genet.*, **96**, 63–8.
35. Devos, K.M. 2005, Updating the crop circle, *Curr. Opin. Plant Biol.*, **8**, 155–62.
36. Srinivasachary, D.M.M., Gale, M.D. and Devos, K.M. 2007, Comparative analyses reveal high levels of conserved colinearity between the finger millet and rice genomes, *Theor. Appl. Genet.*, **115**, 489–99.