Genetic Diversity in Maintainer and Restorer Lines of Pearl Millet

T. Nepolean, S. K. Gupta,* S. L. Dwivedi, R. Bhattacharjee, K. N. Rai, and C. T. Hash

ABSTRACT

Molecular markers facilitate rapid and environment-neutral characterization of the pattern of genetic diversity. The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) at Patancheru, India, has developed a large number and diverse range of maintainer lines (B-lines) and restorer lines (R-lines) of pearl millet [*Pennisetum glaucum* (L.) R. Br.] in last three decades. In the present study, 213 lines (98 B-lines and 115 R-lines) were genotyped using high throughput assay and 38 polymorphic simple sequence repeat (SSR) markers, which detected 308 alleles averaging 8.1 alleles per locus. Restorer lines were genetically more diverse than B-lines, as higher average gene diversity was detected among R-lines (0.62) than among B-lines (0.50). A neighbor-joining tree based on simple matching dissimilarity distance matrix clearly differentiated B-lines from R-lines into two clusters, with further subdivision of B-lines into four subclusters and R-lines into five subclusters. Seven B-lines clustered with R-lines, and only two R-lines clustered with B-lines. Few unique alleles (1 to 3) were detected in 14 B- and 30 R-lines, which could be used to distinguish them from each other as well as from the remaining lines used in this study. The SSR alleles of each of the parental lines along with distinctness, uniformity, and stability (DUS) characterization may enable ICRISAT to protect these lines from any kind of infringement by the multiple users who have received these lines for hybrid development.

T. Nepolean, current address, Indian Agricultural Research Institute, New Delhi, 110 012, India; S.K. Gupta, S.L. Dwivedi, K.N. Rai, and C.T. Hash, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, 502 324, Andhra Pradesh, India; R. Bhattacharjee, current address, International Institute of Tropical Agriculture (IITA), PMB 5320, Ibadan, Nigeria. Received 16 Nov. 2011. *Corresponding author (s.gupta@cgiar.org).

Abbreviations: AMOVA, analysis of molecular variance; B-line, maintainer line; CMS, cytoplasmic-nuclear male sterility; DUS, distinctness, uniformity, and stability; Fst, fixation index; ICMB, ICRISAT millet B-line; ICMP, ICRISAT millet pollinator; IPC, ICRISAT pollinator collection; LG, linkage group; PCR, polymerase chain reaction; PIC, polymorphic information content; R-line, restorer line; SSR, simple sequence repeat; WCA, West and Central Africa.

Pearl millet ranks sixth globally among cereals in terms of area cultivated after wheat (*Triticum aestivum* L.), rice (*Oryza sativa* L.), maize (*Zea mays* L.), barley (*Hordeum vulgare* L.), and sorghum [*Sorghum bicolor* (L.) Moench] (FAO, 2010). It is a staple food crop for about 90 million people living in the semi-arid tropical regions of Africa and the Indian subcontinent (Gulia et al., 2007). India is the largest producer of this crop with 9.6 million ha cultivated area and 10 million tons of grain production (Indiastat, 2010). It is also valued as an important fodder crop. Until the late 1980s, population improvement had been the major breeding strategy at ICRISAT targeted for developing improved open-pollinated varieties. However, single-cross hybrids in pearl millet are reported to confer 20 to 30% grain yield advantages over open-pollinated varieties (Rai et al., 2006). Thus, ICRISAT's research strategy at Patancheru was aligned with the regional priority to breed parental lines of potential hybrids. Currently, there are over

Published in Crop Sci. 52:2555–2563 (2012).
doi: 10.2135/cropsci2011.11.0597

90 hybrids (by name) grown on 4.5 million ha in India of which about 70 hybrids, mostly from the private seed companies, are based on ICRISAT-bred parental lines or on proprietary parental lines developed from ICRISAT-bred improved germplasm (Mula et al., 2007). Therefore, ICRISAT is a major producer and supplier of the parental lines of potential hybrids, mostly to public and private sector hybrid breeding programs in India.

Maintainer lines (B-lines) bred at ICRISAT have been characterized for morphological diversity using limited field evaluation (Rai et al., 2009). Similar characterization of restorer lines (R-lines) has also been done (Gupta et al., 2011). In comparison to morphological descriptors, molecular markers have many advantages, as their expression is independent of environmental conditions. The objective of the present study was to characterize genetic diversity among 98 B-lines and 115 R-lines developed by the pearl millet breeding program at ICRISAT and identify alleles that could be used as genetic tags for finger printing these lines and their derivatives.

MATERIALS AND METHODS

Experimental Material

A set of 98 B-lines and 115 R-lines, primarily developed by the pearl millet breeding program of ICRISAT-Patancheru, were used in this study. All the 115 R-lines were restorers of A_1 cytoplasmic-nuclear male sterility (CMS) system while 60, 33, 4, and 1 B-lines were maintainers of A_1, A_4, A_5, and A_{egp} CMS systems, respectively. Both B- and R-lines were derived using a range of diverse germplasm and/or improved breeding lines and these differ for numerous agronomic traits such as days to flowering, plant height, panicle exsertion, panicle length, panicle diameter, panicle shape, panicle density, seed size and shape, and seed color (Talukdar et al., 1995; Rai et al., 2009; Gupta et al., 2011). Tift $23D_2B_1$, used as a reference genotype, was bred at the Coastal Plain Experiment Station, Tifton, GA, by introducing the d_2 dwarfing gene into the genetic background of an elite line Tift $23B_1$, which is a maintainer of the A_1 CMS system (Burton, 1967, 1969).

DNA Extraction

The B- and R-lines along with Tift $23D_2B_1$ were grown in small plastic pots in a greenhouse. Approximately 30 mg of leaf tissue from 20 to 25 seedlings (14 d old) were collected and bulked after removing the leaf tips and midribs. The harvested leaf samples were immediately collected in 96-well plate that consisted of 95 lines and one control (Tift $23D_2B_1$). Genomic DNA was extracted following the protocol described in Sharp et al. (1988). Quantification of DNA, quality check, and normalization up to 5 ng μl^{-1} were done on agarose gel (1.2%, containing ethidium bromide).

Simple Sequence Repeat Markers

A set of 32 genomic- and six expressed sequence tag-simple sequence repeat (SSR) primer pairs were selected to genotype the B-lines and R-lines using an ABI3700 DNA fragment analyzer (Applied Biosystems). Of these, 27 SSRs had dinucleotide repeats, four had trinucleotide repeats, and the remaining seven had compound repeats. Twenty-nine of the 38 SSRs had been mapped on pearl millet genome: five to seven SSR loci on linkage groups (LGs) 1, 2, 3, and 5 and one to two SSR loci each on LGs 4, 5, and 6 (Qi et al., 2004; Senthilvel et al., 2008).

Polymerase Chain Reaction and Genotyping

Polymerase chain reactions (PCRs) were performed in 5 μL volumes in 384-well PCR plates (ABGene Rochester). Each PCR reaction mixture contained 5 ng of genomic DNA, 2 pmol μL^{-1} of each primer, 25 mM $MgCl_2$, 2 mM of each deoxyribonucleotide triphosphate, 10x reaction buffer, and 0.2 U Amplitaq Gold Polymerase (Applied Biosystems India). After one denaturing step of 15 min at 94°C, a touchdown amplification program was performed on a GeneAmp 9700 thermal cycler (Applied Biosystems). This profile consisted of a denaturing step of 25 sec at 94°C and an extension step of 30 sec at 72°C. The initial annealing step was 20 sec at 64°C for one cycle and subsequently the temperature was reduced by 1°C for every cycle until a final temperature of 55°C was reached. The annealing temperature of 55°C was maintained for the last 35 cycles of the amplification followed by the final extension of 72°C for 7 min. Polymerase chain reaction products were pooled after PCR, which contained 1 μL of each of dye-labeled (FAM, VIC, NED, and PET) product, 7 μL of formamide, 0.3 μL of the LIZ-labeled (500 [−250]) size standard, and 4.2 μL of distilled water. The DNA fragments were size separated on an ABI 3700 automatic DNA sequencer (Perkin-Elmer/Applied Biosystems). GeneScan 3.1 (Applied Biosystems, 1998a) was used to size the peak patterns by using the internal LIZ-labeled (500 [−250]) size standard and Genotyper 3.1 (Applied Biosystems, 1998b) was used for allele calling. Final bins for individual markers were assigned with the help of marker repeat length using the AlleloBin 2.0 program (Prasanth et al., 2006.) and the binned data was used for further data analysis.

Data Analysis

Polymorphic information content (PIC); allelic richness as determined by total number of alleles, alleles per locus, and occurrence of common, most frequent, rare, and unique alleles; gene diversity; and heterozygosity were estimated using the PowerMarker V3.0 software (Liu and Muse, 2005). Unique alleles are those that are present in one line but absent in other lines. Rare alleles are defined as those whose frequency is ≤1% in the investigated materials. Common alleles are defined as those whose frequency is 1 to 20% while most frequent alleles are those whose frequency is >20%. This classification of alleles was done following Li et al. (2008) and Upadhyaya et al. (2008).

Analysis of molecular variance (AMOVA) was performed (Excoffier et al., 1992) to estimate the variance components among and within B- and R-line groups. The *F* value—the fixation index (or Wright's F statistic) (Fst)—from the AMOVA analysis provided genetic differentiation of the subgroups. Simple matching allele frequency-based distance matrix was used in the DARwin-5.0 program (Perrier et al., 2003) to construct a tree diagram (38 SSR loci data on 98 B- and 115 R-lines) to examine the genetic structure and diversity among B- and R-lines. The grouping of B- and R-lines into clusters and subclusters was done at 5% dissimilarity level.

Table 1. Allelic composition, polymorphic information content (PIC), gene diversity, and heterozygosity of the 38 simple sequence repeat (SSR) loci in 98 maintainer lines (B-lines) and 115 restorer lines (R-lines) of pearl millet.

SSR loci	Allelic richness	Allele size range (bp)	Rare allele (≤1%)	Common allele (1 to ≤20%)	Most frequent allele (>20%)	PIC†	Gene diversity†	Heterozygosity†
Xctm08	4	6 (249–255)‡	0	2	2	0.47 (0.44, 0.42)	0.57 (0.51, 0.48)	0.03 (0.01, 0.04)
Xctm10	16	32 (167–199)	6	9	1	0.77 (0.41, 0.89)	0.78 (0.43, 0.90)	0.04 (0.01, 0.06)
Xctm12	9	20 (319–339)	2	4	3	0.73 (0.60, 0.73)	0.76 (0.65, 0.76)	0.04 (0.05, 0.03)
Xicmp3002	5	12 (197–209)	1	2	2	0.54 (0.52, 0.56)	0.61 (0.58, 0.62)	0.04 (0.01, 0.06)
Xicmp3032	6	15 (180–195)	2	2	2	0.66 (0.62, 0.55)	0.71 (0.67, 0.61)	0.08 (0.07, 0.10)
Xicmp3043	4	15 (229–244)	1	2	1	0.42 (0.29, 0.50)	0.46 (0.34, 0.55)	0.00 (0.00, 0.00)
Xicmp3048	2	5 (244–249)	0	1	1	0.10 (0.08, 0.12)	0.11 (0.08, 0.13)	0.01 (0.02, 0.00)
Xicmp3080	6	15 (211–226)	2	2	2	0.65 (0.61, 0.55)	0.70 (0.67, 0.61)	0.07 (0.03, 0.10)
Xicmp3088	6	15 (152–167)	1	3	2	0.64 (0.54, 0.58)	0.70 (0.59, 0.61)	0.01 (0.01, 0.01)
Xpsmp2045	6	12 (195–207)	3	1	2	0.48 (0.28, 0.49)	0.56 (0.30, 0.56)	0.01 (0.01, 0.02)
Xpsmp2068	17	38 (99–137)	5	11	1	0.86 (0.71, 0.87)	0.87 (0.73, 0.88)	0.02 (0.03, 0.02)
Xpsmp2070	24	74 (191–265)	11	12	1	0.88 (0.66, 0.89)	0.88 (0.68, 0.89)	0.02 (0.00, 0.05)
Xpsmp2077	7	38 (138–176)	3	2	2	0.53 (0.47, 0.46)	0.06 (0.55, 0.49)	0.05 (0.03, 0.06)
Xpsmp2079.2	18	46 (213–259)	6	11	1	0.85 (0.73, 0.88)	0.86 (0.75, 0.89)	0.03 (0.00, 0.07)
Xpsmp2086	8	20 (110–130)	3	3	2	0.66 (0.68, 0.47)	0.70 (0.72, 0.51)	0.01 (0.01, 0.02)
Xpsmp2089	11	30 (103–133)	2	8	1	0.83 (0.74, 0.86)	0.85 (0.77, 0.87)	0.02 (0.00, 0.04)
Xpsmp2090	9	20 (171–191)	4	3	2	0.68 (0.56, 0.74)	0.72 (0.59, 0.77)	0.00 (0.00, 0.01)
Xpsmp2201	6	4 (332–366)	2	3	1	0.42 (0.17, 0.56)	0.45 (0.18, 0.61)	0.03 (0.01, 0.04)
Xpsmp2202	2	16 (145–161)	0	0	2	0.29 (0.30, 0.28)	0.35 (0.37, 0.33)	0.04 (0.03, 0.05)
Xpsmp2203	10	26 (335–361)	3	6	1	0.72 (0.34, 0.81)	0.75 (0.38, 0.83)	0.04 (0.00, 0.07)
Xpsmp2204	10	94 (173–267)	3	5	2	0.73 (0.36, 0.82)	0.76 (0.44, 0.84)	0.02 (0.02, 0.03)
Xpsmp2207	7	14 (296–310)	0	4	3	0.73 (0.55, 0.70)	0.77 (0.60, 0.74)	0.05 (0.02, 0.07)
Xpsmp2209	6	20 (337–357)	2	2	2	0.64 (0.62, 0.55)	0.70 (0.68, 0.62)	0.10 (0.08, 0.11)
Xpsmp2211	5	10 (110–120)	2	1	2	0.47 (0.42, 0.49)	0.56 (0.53, 0.58)	0.06 (0.07, 0.04)
Xpsmp2212	4	6 (188–194)	2	0	2	0.32 (0.37, 0.14)	0.38 (0.50, 0.14)	0.03 (0.06, 0.01)
Xpsmp2214	4	6 (242–248)	1	0	3	0.56 (0.37, 0.60)	0.63 (0.41, 0.67)	0.04 (0.01, 0.07)
Xpsmp2218	21	50 (217–267)	8	12	1	0.86 (0.75, 0.87)	0.87 (0.77, 0.88)	0.05 (0.02, 0.07)
Xpsmp2220	12	24 (108–132)	4	7	1	0.81 (0.56, 0.82)	0.83 (0.59, 0.84)	0.02 (0.00, 0.04)
Xpsmp2222	4	6 (152–158)	0	3	1	0.25 (0.24, 0.25)	0.26 (0.26, 0.26)	0.01 (0.02, 0.01)
Xpsmp2227	3	4 (194–198)	1	1	1	0.22 (0.18, 0.68)	0.25 (0.19, 0.72)	0.02 (0.02, 0.06)
Xpsmp2232	10	24 (226–250)	5	2	3	0.71 (0.59, 0.64)	0.75 (0.64, 0.69)	0.00 (0.00, 0.00)
Xpsmp2237	6	40 (214–258)	1	3	2	0.62 (0.30, 0.36)	0.66 (0.36, 0.47)	0.04 (0.01, 0.04)
Xpsmp2246	4	6 (258–264)	0	2	2	0.57 (0.44, 0.58)	0.63 (0.47, 0.65)	0.01 (0.01, 0.01)
Xpsmp2248	5	8 (162–170)	1	2	2	0.46 (0.55, 0.36)	0.51 (0.61, 0.39)	0.01 (0.01, 0.01)
Xpsmp2249	5	10 (150–160)	2	2	1	0.28 (0.15, 0.37)	0.30 (0.17, 0.41)	0.03 (0.01, 0.04)
Xpsmp2251	8	14 (154–168)	4	2	2	0.45 (0.38, 0.48)	0.53 (0.45, 0.57)	0.00 (0.00, 0.00)
Xpsmp2267	2	1 (256–257)	0	0	2	0.37 (0.25, 0.19)	0.49 (0.29, 0.21)	0.03 (0.01, 0.03)
Xpsmp2273	16	62 (158–220)	7	8	1	0.75 (0.68, 0.78)	0.78 (0.70, 0.80)	0.04 (0.01, 0.06)
Total	308		100	143	65			
Mean	8.1					0.58 (0.46, 0.58)	0.62 (0.50, 0.62)	0.03 (0.02, 0.04)

†PIC, gene diversity, and heterozygosity for the B- and R-lines are given in parentheses, respectively.

‡Numbers in parentheses indicate range.

RESULTS

The 38 SSR loci detected a total of 308 alleles in 98 B- and 115 R-lines, with an average of 8.1 alleles per locus. The number of alleles per locus varied from 2 (*Xicmp3048*, *Xpsmp2202*, and *Xpsmp2267*) to 24 (*Xpsmp2070*), with four to six alleles at 17 SSR loci (Table 1). Markers *Xpsmp2070* and *Xpsmp2218* had more than 20 alleles while *Xctm10*, *Xpsmp2068*, *Xpsmp2079*, *Xpsmp2089*, *Xpsmp2203*, *Xpsmp2204*, *Xpsmp2220*, *Xpsmp2232*, and *Xpsmp2273* amplified 10 to 18 alleles per locus. The allele size across the loci and lines varied from 1 (*Xpsmp2267*) to 94 bp (*Xpsmp2204*). The PIC ranged from 0.10 (*Xicmp3048*) to 0.88 (*Xpsmp2070*) with an average of 0.58. Twenty of the 38 SSRs were highly polymorphic, with PIC values ranging from 0.64 to 0.88 and averaging 0.74. Furthermore, the average PIC values for B- and R-lines were 0.46 and 0.58, respectively. Gene diversity, defined as the probability that two randomly chosen alleles from the population are different, varied from 0.11 (*Xicmp3048*) to 0.88 (*Xpsmp2070*) with an average of 0.62. Restorer lines had higher average gene diversity (0.62) than

Table 2. Genotype-specific alleles present in 14 maintainer lines (B-lines) and 30 restorer lines (R-lines) of pearl millet.

Accession identity†	SSR‡ allele	Accession identity	SSR allele
	Maintainer (B) lines		
ICMB 04444	*Xicmp3032*$_{180}$ and *Xicmp3080*$_{211}$	ICMB 91444	*Xpsmp2209*$_{353}$
ICMB 96222	*Xpsmp2068*$_{137}$	ICMB 92444	*Xpsmp2218*$_{217}$ and *Xpsmp2273*$_{186}$
842B	*Xpsmp2070*$_{235}$ and *Xpsmp2086*$_{114}$	ICMB 92333	*Xpsmp2227*$_{198}$
ICMB 01333	*Xpsmp2070*$_{251}$	ICMB 95333	*Xpsmp2232*$_{228}$
ICMB 91666	*Xpsmp2079.2*$_{259}$	ICMB 98333	*Xpsmp2232*$_{226}$
ICMB 94333	*Xpsmp2086*$_{110}$	ICMB 98111	*Xpsmp2251*$_{154}$
ICMB 04111	*Xpsmp2086*$_{130}$	ICMB 03999	*Xpsmp2251*$_{166}$
	Restorer (R) lines		
IPC 689	*Xctm12*$_{335}$	IPC 821	*Xpsmp2203*$_{361}$
IPC 1536	*Xctm10*$_{199}$	IPC 244	*Xpsmp2204*$_{191}$
IPC 645	*Xpsmp2045*$_{207}$	IPC 487	*Xpsmp2204*$_{175}$
IPC 1307	*Xpsmp2045*$_{205}$ and *Xpsmp2079.2*$_{235}$	IPC 795	*Xpsmp2214*$_{248}$ and *Xctm10*$_{167}$
IPC 417	*Xpsmp2068*$_{135}$	IPC 1027	*Xpsmp2218*$_{247}$
IPC 1351	*Xpsmp2068*$_{129}$	IPC 1078	*Xpsmp2218*$_{219}$
IPC 338	*Xpsmp2070*$_{191}$ and *Xctm10*$_{169}$	IPC 637	*Xpsmp2220*$_{132}$
IPC 1329	*Xpsmp2070*$_{233}$ and *Xpsmp2273*$_{176}$	IPC 1114	*Xpsmp2220*$_{114}$
IPC 1047	*Xpsmp2077*$_{142}$ and *Xpsmp2079.2*$_{223}$	IPC 687	*Xpsmp2232*$_{242}$
IPC 404	*Xpsmp2079.2*$_{245}$	IPC 774	*Xpsmp2232*$_{250}$
IPC 1445	*Xpsmp2079.2*$_{219}$	IPC 1178	*Xpsmp2249*$_{150}$
IPC 367	*Xpsmp2089*$_{133}$, *Xpsmp2220*$_{124}$, and *Xpsmp2203*$_{343}$	IPC 419	*Xpsmp2251*$_{168}$
IPC 896	*Xpsmp2089*$_{105}$	IPC 487	*Xpsmp2251*$_{167}$
IPC 976	*Xpsmp2090*$_{171}$	IPC 1254	*Xpsmp2273*$_{176}$
IPC 954	*Xpsmp2201*$_{362}$	IPC 1536	*Xpsmp2273*$_{199}$

†ICMB, ICRISAT millet B-line; IPC, ICRISAT pollinator collection.

‡SSR, simple sequence repeat.

Table 3. Analysis of molecular variance of 98 maintainer lines (B-lines) and 115 restorer lines (R-lines).

Source of variation	df	Sum of squares	Variance components	Percentage of variation	*P* (10,000 permutations)
Between B- and R-line group	1	355.562	1.64439	17.91	0.0000
Within B- and R-line group	424	3196.175	7.53815	82.09	
Total	425	3551.737	9.18254		
Fixation index: 0.17908					0.0000

B-lines (0.50). The level of heterozygosity in SSRs across B- and R-lines ranged from 0.01 to 0.10 and averaging 0.03, with >0.05 in four SSRs. Heterozygosity was not detected at four SSR loci (*Xicmp3043*, *Xpsmp2090*, *Xpsmp2232*, and *Xpsmp2251*). The R-lines had greater average heterozygosity (0.04) than B-lines (0.02). Moreover, ICRISAT millet B-line ICMB 96666, ICMB 89111, ICMB 00333, and ICMB 04333 among B-lines; and ICRISAT millet pollinator ICMP 451, ICRISAT pollinator collection IPC 492, IPC 962, IPC 1018, IPC 1078, IPC 1307, IPC 1503, IPC 1617, and IPC 2089 among R-lines were genetically more variable than other lines, with heterozygosity at three to six SSR loci detected in the former and at 6 to 16 SSR loci in the latter. Allele sizes for the internal control (Tift $23D_2B_1$) were uniform and reproducible for each of the markers indicating the accuracy of the protocol and reproducibility of allelic data for a given primer across assays.

Of the 308 alleles detected in B- and R-lines, 100 were rare (frequency ≤ 1%), 143 common (frequency between 1 and 20%), and 65 most frequent (frequency > 20%) (Table 1), with more alleles observed in R-lines (284 alleles in 115 lines averaging 2.47 alleles per line) than the B-lines (214 alleles in 98 lines averaging 2.18 alleles per line). The most frequent alleles were detected at all the SSR loci, ranging from one to three and averaging 1.7 alleles per locus. The rare alleles ranged from 1 (*Xicmp3002*, *Xicmp3043*, *Xicmp3088*, *Xpsmp2214*, *Xpsmp2227*, *Xpsmp2237*, and *Xpsmp2248*) to 11 (*Xpsmp2070*), averaging 2.6, while the common alleles ranged from 1 (*Xicmp3048*, *Xpsmp2045*, *Xpsmp2211*, and *Xpsmp2227*) to 12 (*Xpsmp2070* and *Xpsmp2218*), averaging 3.7. Twenty-nine SSR loci detected the rare, common, and most frequent alleles. For five SSR loci in this study, both common and most frequent alleles were detected but no rare alleles were detected whereas for two SSR loci, only rare and most frequent alleles were detected. Simple sequence repeat locus *Xpsmp2070* detected 11 rare and 12 common alleles but only one most frequent allele while *Xpsmp2202* and *Xpsmp2267* each detected only two most frequent alleles and

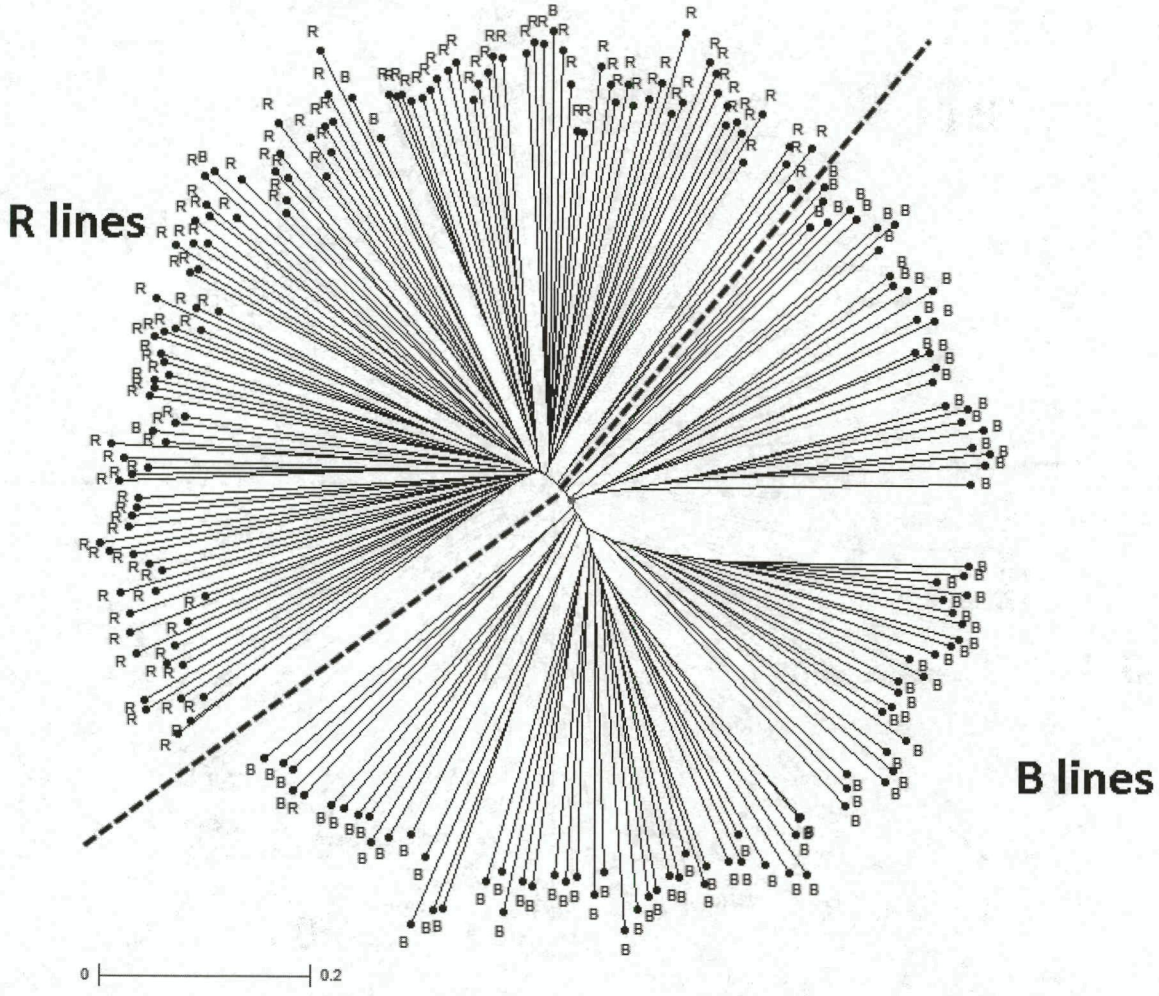

Figure 1. Unweighted neighbor-joining tree based on a simple matching dissimilarity matrix for allele sizes detected by 38 simple sequence repeat primer pairs across 98 maintainer lines (B-lines) and 115 restorer lines (R-lines) of pearl millet. Accessions are identified as "B" for maintainer lines and "R" for restorer lines.

no rare or common alleles. Genotype-specific alleles were detected in 14 B- and 30 R-lines (Table 2).

The results of AMOVA showed significant difference between B- and R-line groups, which accounted for 17.9% of the total genetic variation detected by 38 marker loci. The within-group variation, both in the B- and R-lines, was much larger and accounted for 82.1% of the total variation (Table 3). The Fst between B- and R-lines for the individual markers varied from 0.00138 to 0.43697 and was significant for 33 of the 38 marker loci (data not shown). Some of the markers, *Xpsmp2237*, *Xpsmp2045*, *Xpsmp2212*, *Xpsmp2203*, *Xpsmp2267*, and *Xpsmp2204*, had high Fst values of 0.436, 0.354, 0.340, 0.268, 0.249, and 0.246, respectively, and contributed maximum to the variation between B- and R-line groups. The neighbor-joining tree based on simple matching dissimilarity matrix of 213 lines clearly differentiated B-lines from R-lines with majority of B- and R-lines grouping separately into two clusters and with a further four subclusters in B-lines and five subclusters in R-lines. However, seven B-lines grouped with R-lines and two R-lines grouped with B-lines (Fig. 1). The tree diagram was also constructed separately for B- and R-lines, which provided clear differentiation of lines among clusters. The B-lines grouped into four subclusters with 39 lines in subcluster B I, 29 in B III, 18 in B IV, and 12 in B II (Fig. 2). Line 81B and 11 other B-lines that have 81B in their parentage grouped in the B I subcluster as did five B-lines with large panicle volume. Furthermore, 81B and 841B, which share Tift $23B_1$ as a common ancestor, grouped relatively closely within this subcluster. ICMB 89111 is another promising B-line that was found in subcluster B II, which also has another five lines having ICMB 89111 in their parentage. The B III subcluster had highest number of lines (26 out of 29 B-lines) with 843B in their parentage although 843B itself was found in subcluster B IV. Likewise, the 115 R-lines formed five subclusters (R I, R II, R III, R IV, and R V) with 23, 21, 30, 13, and 28 lines, respectively (Fig. 3). Nine lines in subcluster R I had B 282 (a d_2 dwarf restorer obtained from Rockefeller Foundation Collection, Bangkok) in their parentage. Subcluster R II (13 lines) was dominated mostly by lines

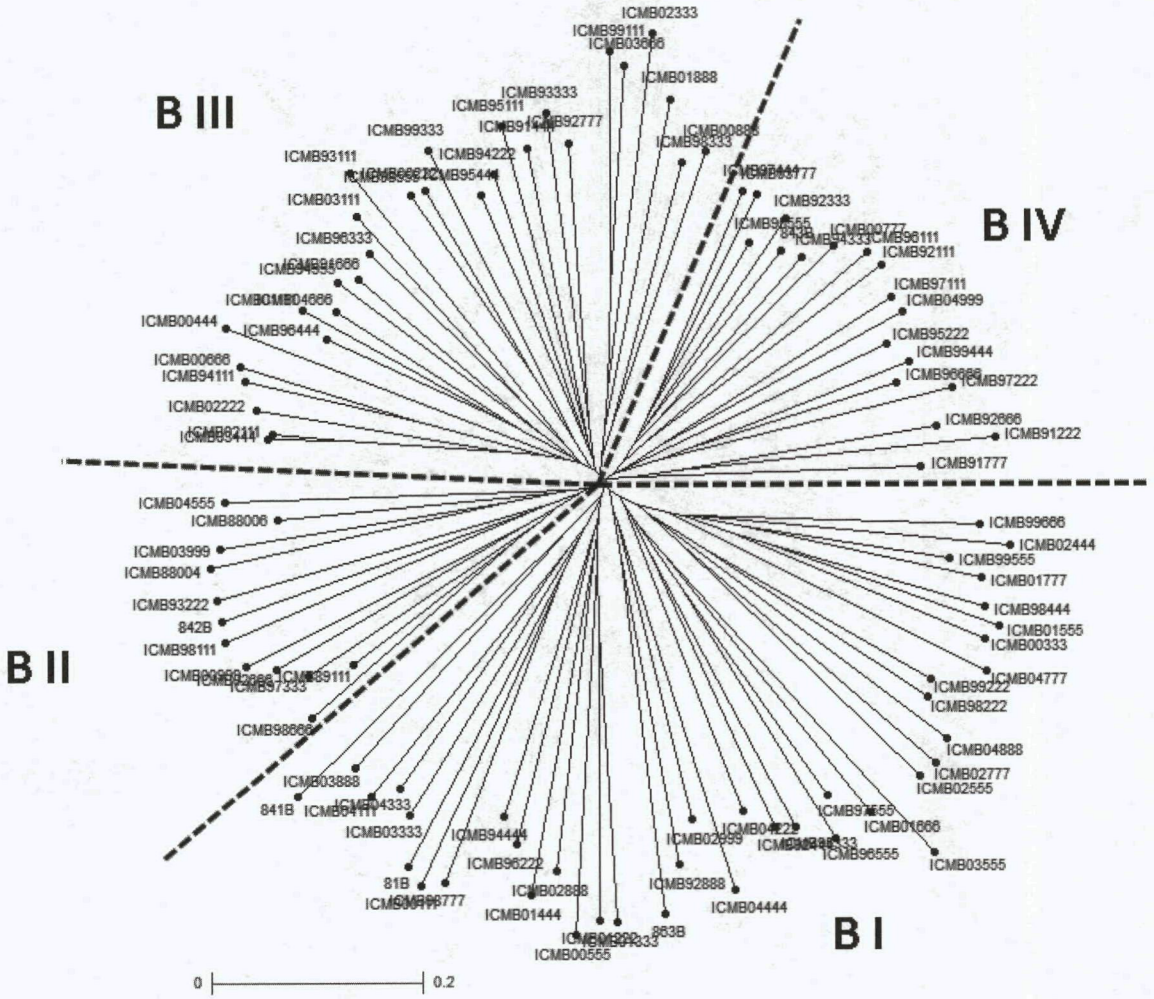

Figure 2. Unweighted neighbor-joining tree based on a simple matching dissimilarity matrix for allele sizes detected by 38 simple sequence repeat primer pairs across 98 maintainer lines (B-lines) of pearl millet. Identity of these lines: ICMB, ICRISAT millet B-line.

developed at Gujarat Agricultural University, Jamnagar, or those derived from crosses involving Jamnagar-bred lines in their parentage. In contrast, subclusters R III, R IV, and R V had lines with mixed parentage.

DISCUSSION

Assessment of genetic variation among lines may help identify genetically diverse parental lines to exploit heterosis in cross-pollinated crops such as pearl millet. Molecular markers have proven to be a powerful tool for fingerprinting and assessing genetic variation and relatedness among germplasm, elite cultivars, and genetic stocks. Simple sequence repeat markers in the present study detected high average gene diversity (0.62), indicating sufficient polymorphisms, which can be used to characterize genetic diversity. Thirty-eight SSR loci detected a total of 308 alleles in 98 B-lines and 115 R-lines, with more alleles detected in R-lines (284) than B-lines (214). The higher gene diversity and greater number of alleles detected in R-lines may be due to the broader genetic base used in the development of these lines and the differences in sample size (115 R-lines vs. 98 B-lines). The average number of alleles per locus obtained in the present study (8.1) was higher than those previously reported in pearl millet (Budak et al., 2003; Mariac et al., 2006; Kapila et al., 2008; Oumar et al., 2008). However, Stich et al. (2010) found even higher number of alleles per locus (16.4) in a study involving 145 pearl millet inbreds that had been derived from diverse landraces of West and Central Africa (WCA). The lower number of alleles per SSR locus detected in the present study, in comparison to Stich et al. (2010), was probably due to much diverse inbreds that were derived from populations originating from WCA countries, the center of diversity for pearl millet, and this set of inbreds well represented the available diversity of pearl millet in this region while the inbreds used in the present investigation were from ICRISAT-Patancheru breeding program and were developed with the objective of their adaptation to pearl millet agro-ecological zones in India, which obviously narrowed down the diversity in comparison to the available pearl millet global diversity. The pattern of these results are similar to those reported in maize, where Liu et al. (2003) found abundant allelic variation with an average of 21.7 alleles per locus in a study of 260 diverse inbreds (representing tropical, subtropical, and temperate

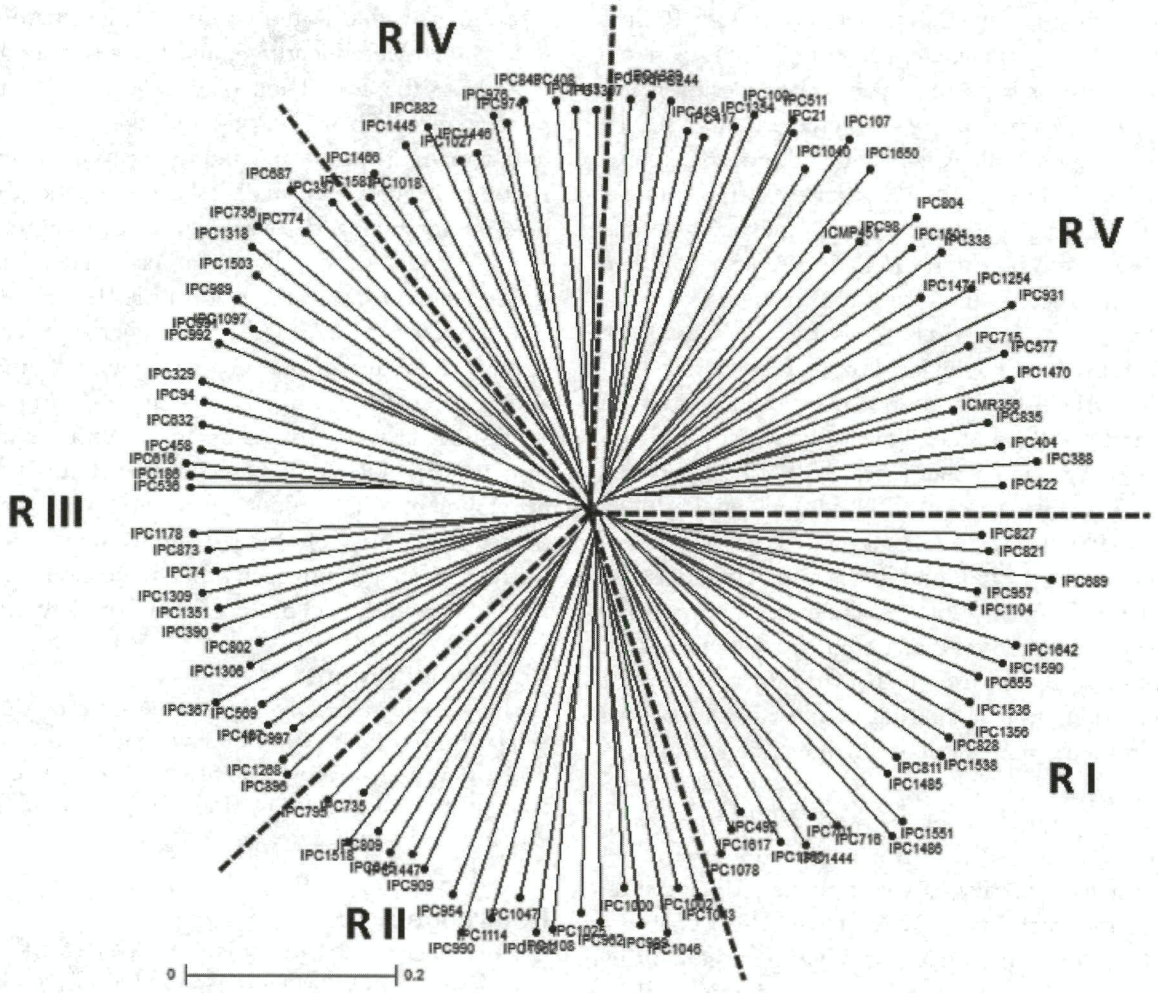

Figure 3. Unweighted neighbor-joining tree based on a simple matching dissimilarity matrix for allele sizes detected by 38 simple sequence repeat primer pairs across 115 restorer lines (R-lines) of pearl millet. Identity of these lines: ICP/ICMP, ICRISAT millet pollinator; IPC, ICRISAT pollinator collection.

regions) while in other studies in maize that used less diverse or single region-specific materials, the number of alleles per locus were much lower, ranging from 4.9 to 6.9 (Taramino and Tingley, 1996; Senior et al., 1998; Lu and Bernardo, 2001; Matsouka et al., 2002).

The levels of heterozygosity detected in B- and R-lines were low (<0.050) for most of the markers, which is acceptable, while in some SSRs it was up to 0.096. A few SSR primer pairs detected heterozygosity in greater numbers of B- and R-lines than others. For example, SSR locus *Xicmp3080* detected heterozygosity in three B-lines and 10 R-lines. Likewise, some of the B- and R-lines were more heterozygous than others. For example, a B-line, ICMB 96666, showed heterozygosity at six SSR loci while ICRISAT millet pollinator ICP 492, IPC 1617, and ICMP 451 among R-lines were found heterozygous at 10 to 16 SSR loci. The high level of heterozygosity detected in elite pollinator ICMP 451, the male parent of formerly popular hybrid ICMH 451, was not unexpected as this partial inbred pollinator is known to be variable for several morphological traits. The probable reasons for high heterozygosity in some lines than others could be either due to (i) residual heterozygosity, (ii) contamination, or (iii) mutation and mutational bias at specific SSR loci or the amplification of similar sequences in two separate genomic regions (Senior et al., 1998; Udupa and Baum, 2001; Matsouka et al., 2002).

Besides the diverse parentage, B- and R-lines in pearl millet have distinct plant characteristics. Maintainer lines are generally bred for short height (<100 cm) and larger seed size while R-lines are generally bred for taller height (150–180 cm), more tillers, relatively small seed size, and profuse pollen production (Rai et al., 2006). Therefore, parents involved in respective B- and R-line development are morphologically quite diverse from each other. This distinctness in parentage and ideotypes was also reflected in the marker-based clustering pattern, which clearly separated B-lines from R-lines in two clusters (Fig. 1). This was also evidenced from AMOVA, which had shown significant variation between B- and R-line groups. The grouping of B- and R-lines into two separate clusters was detected by markers under study, as 33 of the 38 markers had significant Fst. Furthermore, the B-lines and R- lines grouped into four and five subclusters,

respectively, with very few B-lines clustering with R-lines and vice-versa. For example, an R-line, IPC 1551, clustered with B-lines, but this is not unexpected considering it had a B-line (843B) in its parentage. Likewise, a B-line, ICMB 98777, clustered with R-lines, which could probably be due to the fact that an R-line (J 104) was involved in its parentage. Furthermore, the clustering pattern in B-lines revealed that B-lines sharing common parents in their parentage clustered together. For instance, B-lines with 81B in their parentage clustered along with 81B in cluster B I while those with ICMB 89111 in their parentage grouped in cluster B II along with ICMB 8911. The two downy mildew resistant lines, 81B and 841B, clustered together, which confirms the previous reports about the genetic relatedness between 81B and 841B (Chowdari et al., 1998; Kapila et al., 2008). The ICMB 00555, ICMB 01222, and ICMB 01333, known for large panicle size, were found in the same cluster, as also reported by Kapila et al. (2008). Maintainer lines sharing 843B in their parentage were scattered in all four clusters, with maximum being in cluster B III, which further supports earlier observation about clustering of most of the B-lines with 843B in their parentage in one large cluster (Kapila et al., 2008).

This study identified 14 B- and 30 R-lines with unique alleles (1 to 3 alleles) not present in other lines that can be used as genetic tags to supplement distinctness, uniformity, and stability (DUS) tests, thus enabling ICRISAT to protect these from possible infringement by multiple users who have received these lines for use in their hybrid program. Furthermore, presence of line-specific alleles may have association with some distinct trait(s) of the particular line, which merits further investigation. Moreover, the higher number of rare and unique alleles in both B- and R-lines suggests that these lines have been derived from diverse genetic base. Six to eight rare alleles were detected in each of ICMB 03555, ICMB 04444, ICMB 92444, and ICMB 00111 among B-lines and IPC 367, IPC 404, IPC 990, IPC 1486, IPC 1583, and IPC 1650 among the R-lines.

The identification of parental combination with strong yield heterosis is the most important step in development of hybrids (Becker et al., 1999; Melchinger, 1999).The information available on genetic distances between the parental lines can be used to predict hybrid performance, as is done in several other crops such as maize (Smith et al., 1990; Ajmone Marsan et al., 1998; Mohammadi et al., 2008), rape (*Brassica napus* L.) seed (Diers et al., 1996), and rice (Zhang et al., 1995). However, an earlier study conducted on hybrid parental lines of pearl millet indicated molecular-marker based genetic-distance as a nonreliable tool for predicting heterotic combinations (Chowdari et al., 1998). This could be due to use of limited number of parental lines (only 12 lines) studied that also had lesser range of similarity index values (0.81 to 0.50). However, in the present study, 23 pairs of B- and R-lines with wide genetic distance from 0.19 to 0.90 were identified, and hybrids were generated to investigate the correlation of genetic distance with hybrid performance. These hybrids and their parents are being tested in multiseason and multiyear trials to examine the relationship between marker-based genetic distance and heterosis for grain and stover yield and other agronomic traits.

In conclusion, this marker-based study clearly differentiated most of the pearl millet B- and R-lines bred or assembled at ICRISAT-Patancheru in two separate clusters. Substantial diversity was detected among the B- and R-lines, indicating that B-line × B-line and R-line × R-line crosses involving lines with greater genetic distance can generate new recombinants for B- and R-line development. The unique alleles and allele combinations detected in some of the B- and R-lines will enable to distinguish them from each other as well as from other lines used in this study and protect them from any infringement.

Acknowledgment

Authors acknowledge the support provided by the ICRISAT-Sehgal Family Foundation Endowment Fund to conduct this study, as a part of project (YSFF06) "Molecular diversity in relation to morphological diversity and yield heterosis in sorghum and pearl millet."

References

Ajmone Marsan, P., P. Castiglioni, F. Fusari, M. Kuiper, and M. Motto. 1998. Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. Theor. Appl. Genet. 96:219–227. doi:10.1007/s001220050730

Applied Biosystems. 1998a. GeneScan analysis software user manual. Release 3.1. Applied Biosystems, Foster City, CA.

Applied Biosystems. 1998b. Genotyper 3.1. Applied Biosystems, Foster City, CA.

Becker, H.C., H. Loptien, and Robbelen G. 1999. Breeding: An overview. In: C. Gomez-Campo, editor, Biology of *Brassica* coenospecies. Elsevier, Amsterdam, Netherlands. p. 413–460.

Budak, H., F. Pedraza, P.B. Cregan, P.S. Baenziger, and I. Dweikat. 2003. Development and utilization of SSRs to estimate the degree of genetic relationships in a collection of pearl millet germplasm. Crop Sci. 43:2284–2290. doi:10.2135/cropsci2003.2284

Burton, G.W. 1967. Pearl millets Tift 23DA and Tift 23DB released. Georgia Agric. Res. 9:6.

Burton, G.W. 1969. Registration of pearl millet inbreds Tift $23B_1$, Tift $23A_1$, Tift $23DB_1$, and Tift $23D2A_1$. Crop Sci. 9:397.

Chowdari, K.V., S.R. Venkatachalam, A.P. Davierwala, V.S. Gupta, P.K. Ranjekar, and O.P. Govila. 1998. Hybrid performance and genetic distance as revealed by the (GATA)4 microsatellite and RAPD markers in pearl millet. Theor. Appl. Genet. 97:163–169. doi:10.1007/s001220050881

Diers, B.W., P.B.E. McVetty, and T.C. Osborn. 1996. Relationship between heterosis and genetic distance based on restriction fragment length polymorphism markers in oilseed rape (*Brassica napus* L.). Crop Sci. 36:79–83. doi:10.2135/cropsci1996.0011183X003600010014x

Excoffier, L., P. Smouse, and J. Quattro. 1992. Analysis of molecular variance inferred for metric distances among DNA

haplophytes: Applications to human mitochondrial DNA restriction data. Genetics. 131: 479–491.

FAO. 2010. FAOSTAT. FAO, Rome, Italy. http://faostat.fao.org/site/567/default.aspx#ancor (accessed 29 Apr. 2010).

Gulia, S.K., J.P. Wilson, J. Carter, and B.P. Singh. 2007. Progress in grain pearl millet research and market development. In: J. Janick and A. Whipkey, editors, Issues in new crops and new uses. ASHS Press, Alexandria, VA. p. 196–203.

Gupta, S.K., R. Bhattacharjee, K.N. Rai, and M. Suresh Kumar. 2011. Characterization of ICRISAT-bred restorer parents of pearl millet. J. SAT Agric. Res. 9:1–5.

Indiastat. 2010. Agricultural production. Indiastat, Datanet India Pvt. Ltd., New Delhi, India. http://www.indiastat.com/agriculture/2/stats.aspx (accessed 1 July 2010).

Kapila, R.K., R.S. Yadav, P. Plaha, K.N. Rai, O.P. Yadav, C.T. Hash, and C.J. Howarth. 2008. Genetic diversity among pearl millet maintainers using microsatellite markers. Plant Breed. 127:33–37.

Li, Y., R. Guan, L. Zhangxiong, M. Yansong, W. Lixia, L. Linhai, L. Fanyun, L. Weijiang, C. Pengyin, Y. Zhe, G. Yuan, Z. Li, N. Xuecheng, M.J.M. Smulders, L. Wei, P. Rihua, C. Yanhua, Y. Zhongmei, G. Min, C. Ruzhen, H. Anfu, S. Ainong, Z. Bo, Z. Shenlong, and Q. Lijuan. 2008. Genetic structure and diversity of cultivated soybean (*Glycine max* (L.) Merr.) landraces in China. Theor. Appl. Genet. 117:857–871. doi:10.1007/s00122-008-0825-0

Liu, K., M. Goodman, S. Muse, J.S. Smith, E. Buckler, and J. Doebley. 2003. Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. Genetics 165:2117–2128.

Liu, K., and S.V. Muse. 2005. PowerMarker: Integrated analysis environment for genetic marker data. Bioinformatics 21:2128–2129. doi:10.1093/bioinformatics/bti282

Lu, H., and R. Bernardo. 2001. Molecular marker diversity among current and historical maize inbreds. Theor. Appl. Genet. 103:613–617. doi:10.1007/PL00002917

Mariac, C., V. Luong, I. Kapran, A. Mamadou, F. Sagnard, M. Deu, J. Chantereau, B. Gerard, J. Ndjeunga, G. Bezançon, J.-L. Pham, and Y. Vigouroux. 2006. Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers. Theor. Appl. Genet. 114:49–58. doi:10.1007/s00122-006-0409-9

Matsouka, Y., S.E. Mitchell, S. Kresovich, M. Goodman, and J. Dobeley. 2002. Microsatellite in *Zea* – Variability, patterns of mutations and use for evolutionary studies. Theor. Appl. Genet. 104:436–450. doi:10.1007/s001220100694

Melchinger, A.E. 1999. Genetic diversity and heterosis. In: C.G. Coors and S. Pandey, editors, Genetics and exploration of heterosis in crops. American Society of Agronomy, Madison, WI. p. 99–118.

Mohammadi, S.A., B.M. Prasanna, C. Sudan, and N.N. Singh. 2008. SSR heterogenic patterns of maize parental lines and prediction of hybrid performance. Biotechnol. and Biotechnol. Eq. 22:541–547.

Mula, R.P., K.N. Rai, and V.N. Kulkarni. 2007. Public-private partnership and impact of ICRISAT's pearl millet hybrid parents research. J. SAT Agric. Res. 5: 1–5.

Oumar, I., C. Mariac, J.L. Pham, and Y. Vigouroux. 2008. Phylogeny and origin of pearl millet (*Pennisetum glaucum* [L.] R. Br) as revealed by microsatellite loci. Theor. Appl. Genet. 117:489–497. doi:10.1007/s00122-008-0793-4

Perrier, X., A. Flori, and F. Bonnot. 2003. Data analysis methods. In: P. Hamon, M. Seguin, X. Perrier, and J.C. Glaszmann, editors, Genetic diversity of cultivated tropical plants. Enfield, Science Publishers, Montpellier, France. p. 43–76.

Prasanth, V.P., S. Chandra, B. Jayashree, and D. Hoisington. 2006. AlleloBin – A program for allele binning of microsatellite markers based on the alogirithm of Idury and Cardon (1997). ICRISAT, Patancheru, Andhra Pradesh, India. http://www.icrisat.org/bt-software-d-allelobin.htm (accessed 23 Oct. 2009).

Qi, X., T.S. Pittaway, S. Lindup, H. Liu, E. Waterman, F.K. Padi, C.T. Hash, J. Zhu, M.D. Gale, and K.M. Devos. 2004. An integrated genetic map and a new set of simple sequence repeat markers for pearl millet, *Pennisetum glaucum*. Theor. Appl. Genet. 109:1485–1493. doi:10.1007/s00122-004-1765-y

Rai, K.N., S.K. Gupta, R. Bhattacharjee, V.N. Kulkarni, A.K. Singh, and A.S. Rao. 2009. Morphological characteristics of ICRISAT-bred pearl millet hybrid seed parents. J. SAT Agric. Res. 7.

Rai, K.N., V.N. Kulkarni, R.P. Thakur, B.I.G. Haussmann, and M.A. Mgonja. 2006. Pearl millet hybrid parents research: Approaches and achievements. In: C.L.L. Gowda, K.N. Rai, B.V.S. Reddy, and K.B. Saxena, editors, Hybrid parents research at ICRISAT. ICRISAT, Patancheru, Andhra Pradesh, India. p. 11–74.

Senior, M.L., J.P. Murthy, M.M. Goodman, and C.W. Stubber. 1998. Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. Crop Sci. 38:1088–1098. doi:10.2135/cropsci1998.0011183X0038000400034x

Senthilvel, S., B. Jayashree, V. Mahalakshmi, P.S. Kumar, S. Nakka, T. Nepolean, and C.T. Hash. 2008. Development and mapping of simple sequence repeat markers for pearl millet from data mining of expressed sequence tags. BMC Plant Biol. 8:119. doi:10.1186/1471-2229-8-119

Sharp, P.J., M. Kries, P.R. Sherry, and M.D. Gale. 1988. Location of β-amylase sequences in wheat and its relatives. Theor. Appl. Genet. 75:286–290. doi:10.1007/BF00303966

Smith, O.S., J.S.C. Smith, S.L. Bowen, R.A. Tenborg, and S.J. Wall. 1990. Similarities among a group of elite inbreds as measured by pedigree, F_1 grain yield heterosis, and RFLPs. Theor. Appl. Genet. 80:833–840. doi:10.1007/BF00224201

Stich, B., B.I.G. Haussmann, R. Pasam, S. Bhosale, C.T. Hash, A.E. Melchinger, and H.K. Parzies. 2010. Patterns of molecular and phenotypic diversity in pearl millet [*Pennisetum glaucum* (L.) R. Br.] from West and Central Africa and their relation to geographical and environmental parameters. BMC Plant Biol. 10:216. doi:10.1186/1471-2229-10-216

Talukdar, B.S., P.P. Prakash Babu, C.T. Hash, and J.R. Witcombe. 1995. Pearl millet pollinators: Descriptions, pedigree and distribution. ICRISAT, Patancheru, AP, India.

Taramino, G., and S. Tingley. 1996. Simple sequence repeats for germplasm analysis and mapping in maize. Genome 39:277–287. doi:10.1139/g96-038

Udupa, S.M., and M. Baum. 2001. High mutation rate and mutational bias at (TAA)n microsatellite loci in chickpea (*Cicer arietinum* L.). Mol. Genet. Genom. 265: 1097–1103. doi:10.1007/s004380100508

Upadhyaya H.D., S.L. Dwivedi, M. Baum, R.K. Varshney, S.M. Udupa, C.L.L. Gowda, D. Hoisington, and S, Singh. 2008. Genetic structure, diversity, and allelic richness in composite collection and reference set in chickpea (*Cicer arietinum* L.). BMC Plant Biol. 8:106. doi:10.1186/1471-2229-8-106

Zhang, Q., Y.J. Gao, M.A. Saghai Maroof, S.H. Yang, and J.X. Li. 1995. Molecular divergence and hybrid performance in rice. Mol. Breed. 1:133–142. doi:10.1007/BF01249698