Published October 18, 2018

RESEARCH

Genetic Diversity Patterns and Heterosis Prediction Based on SSRs and SNPs in Hybrid Parents of Pearl Millet

Satbeer Singh,* S. K. Gupta, Mahendar Thudi, Roma Rani Das, Anilkumar Vemula, Vanika Garg, R. K. Varshney, Abhishek Rathore, S. K. Pahuja, and Dev Vart Yadav

ABSTRACT

The present investigation was performed to assess genetic divergence and heterosis prediction in hybrid parents of pearl millet [*Pennisetum glaucum* (L.) R Br.] using simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers. Using 56 SSR loci, 412 alleles were detected in 147 lines with an average of 7.36 alleles per locus, and 75,007 SNP loci were detected in 117 lines. Both SSR- and SNP-based clustering and structure analysis partitioned all maintainer (B) and restorer (R) lines into two clear-cut separate groups, indicating the existence of two diverse gene pools, each representing the seed and restorer parents in pearl millet. Results of analysis of molecular variance and principal coordinate analysis also showed significant diversity between B and R lines. The correlation between parental genetic distances estimated based on SSRs and SNPs was high and significant ($r = 0.58$, $p < .01$). Similar clustering pattern of hybrid parents was observed with both marker systems, although the cost of genotyping was 41% less with SNPs than with SSRs, and the ratio of loci detected with SNPs was much higher (1:364 SSR/SNP), hence the use of SNPs is indicated over SSRs for germplasm characterization. A set of 136 hybrids (including all B × B, R × R, and B × R crosses) generated crossing 17 hybrid parents (nine B lines and eight R lines) in half diallel (without reciprocal) fashion, and evaluation at two locations revealed that the correlation between genetic distance and better parent heterosis for grain yield was moderate, positive, and significant (with SSR, 0.33, $p < .01$; with SNP, 0.35, $p < .01$), hence both SSRs and SNPs were found comparable in results for heterosis prediction.

S. Singh, S.K. Gupta, M. Thudi, R.R. Das, A. Vemula, V. Garg, R.K. Varshney, and A. Rathore, International Crops Research Institute for the Semi-Arid Tropics, Patancheru 502 324, Telangana, India; S.K. Pahuja and D.V. Yadav, Chaudhary Charan Singh Haryana Agricultural Univ., Hisar 125 004, Haryana, India. Received 12 Mar. 2018. Accepted 30 Aug. 2018. *Corresponding author (satbeeryadav87@gmail.com). Assigned to Associate Editor Jeffrey Endelman.

Abbreviations: AMOVA, analysis of molecular variance; B line, maintainer line; BPH, better parent heterosis; PCoA, principal coordinate analysis; PCR, polymerase chain reaction; PIC, polymorphic information content; R line, restorer line; SNP, single nucleotide polymorphism; SSR, simple sequence repeat.

Pearl millet [*Pennisetum glaucum* (L.) R. Br], a diploid species ($2n = 14$), is cultivated on ~30 m ha in the dry areas of the world. It has the ability to grow in environments of low and erratic rainfall, high temperature, and low soil fertility and is the main source of food and fodder for the subsistence farmers in the arid and semiarid tropics of Asia and Africa (Yadav and Rai, 2013).

After the discovery of stable A_1 cytoplasmic male sterility in 1962 in pearl millet (Burton, 1965), followed by its successful transformation to hybrid technology with 30 to 40% heterosis for grain yield, there has been phenomenal increase in area under hybrids, especially in India (Yadav et al., 2012). At present, hybrids occupy ~70% (~5 m ha^{-1}) of the area under pearl millet cultivation in India, and this percentage is further increasing (Satyavathi, 2017).With the global mandate of pearl millet crop improvement, ICRISAT is continuously enhancing genetic diversity of hybrid parents by involving new germplasm of Asian and African origin to improve the productivity and adaptability of this crop in diverse agroecologies. These efforts to enhance parental genetic diversity by ICRISAT and other breeding programs in the public and private sector led to significant hybrid cultivar diversity in India (Yadav et al., 2016). In any year, ~80 to 100 pearl millet

Published in Crop Sci. 58:2379–2390 (2018).
doi: 10.2135/cropsci2018.03.0163

hybrids (by name) have been cultivated in last two decades in India, and ~60 to 70% of them are directly or indirectly based on ICRISAT-bred hybrid parental lines (Mula et al., 2007; Rao et al., 2018). Hence, considering such high-order germplasm impact, it is important to monitor levels of genetic diversity continuously in such pearl millet breeding programs.

Often, breeders use morphological traits and pedigree information for the purpose of germplasm characterization. However, these methods are time consuming, environmentally influenced, and less accurate across spatial and temporal conditions. Recently, some of the molecular-marker-based investigations in pearl millet were able to well characterize germplasm into genetically distinct groups (Kapila et al., 2008; Stich et al., 2010; Singh et al., 2013a; Gupta et al., 2015; Ramya et al., 2018). Different types of markers have proven to successfully unfold genetic diversity patterns in many crops, although simple sequence repeats (SSRs) have been judged as more reliable DNA markers for genetic diversity studies. More recently, single nucleotide polymorphisms (SNPs) have also attracted more attention due to their higher abundance throughout the genome than SSRs. He et al. (2012) investigated hybrid parents in rice (*Oryza sativa* L.) using SSRs and SNPs and concluded that both SSR and SNP markers have distinguishable power to detect polymorphism, whereas Hamblin et al. (2007) found that SSRs give higher grouping probabilities than SNPs in maize (*Zea mays* L.). Moreover, the real test of marker technology in breeding could be its precision to predict heterosis for grain yield. However, the relationship between marker based genetic distance and heterosis when investigated in several crops using different marker systems revealed results in both positive and negative directions (Dias et al., 2004; Kiula et al., 2008; Barbara, 2011; Gupta et al., 2017; Ramya et al., 2018).

To date, most of the investigations on genetic diversity in pearl millet have been based on morphological traits and pedigree analyses, while few are based on SSR markers (Chowdari et al., 1998; Mariac et al., 2006; Stich et al., 2010; Nepolean et al., 2012; Singh et al., 2013a; Bashir et al., 2014; Gupta et al., 2015). Thus, the present study was undertaken to assess the genetic diversity in hybrid parents of pearl millet using two different marker systems, SSRs and SNPs, to understand the patterns of genetic variability and its association with heterosis to help pearl millet breeders strategize productivity enhancement in this crop.

MATERIALS AND METHODS

Plant Material and DNA Extraction

A set of advanced generation (>$F_{7/8}$) 150 pearl millet hybrid parents involving 75 seed parents (B lines) and 75 restorer parents (R lines) was used in present study, including 120 hybrid parents (60 B and 60 R lines) from ICRISAT, Patancheru, India, and 30 (15 B and 15 R lines) from Chaudhary Charan Singh Haryana Agricultural University, Hisar, India. This set of parents was selected based on diverse parentage in pedigrees considering that it represents existing diversity in hybrid parents of pearl millet. The ICRISAT pearl millet hybrid breeding program has contributed significantly to the parentage of hybrids released to date from both public and private sector organizations in India, hence a higher number of 120 diverse parental lines were taken from ICRISAT pearl millet breeding program. The B lines were coded as B01 to B75, and R lines were coded as R01 to R75 (pedigrees are given in Supplemental Table S1). Tift $23D_2B_1$, a maintainer of the A_1 cytoplasmic male sterility system (Burton, 1969), was used as a reference genotype, which was bred by introducing the d_2 dwarfing gene in the genetic background of Tift $23B_1$ at Tifton, GA.

Thirty to thirty-five seeds from each of 150 hybrid parents and Tift $23D_2B_1$ were planted in pots in darkroom at 18 to 25°C. DNA was isolated from ~100 mg of fresh leaf tissue from 8-d-old seedlings using the NucleoSpin 96 Plant II kit. The first and second elutions of DNA were used for SNP and SSR markers, respectively. The quality, quantity, and integrity of DNA was checked by using 0.8% (w/v) agarose gel electrophoresis in Tris-acetate-EDTA (TAE) buffer for 1 h at 90 V with visualization under ultraviolet light after staining with ethidium bromide and then diluted up to 5 ng μL^{-1}.

Genotyping Using SSR Markers

A set of 56 SSR markers, identified as highly polymorphic by earlier studies involving diverse pearl millet germplasm (Yadav et al., 2007; Senthilvel et al., 2008; Rajaram et al., 2013; Moumouni et al., 2015), were selected to establish SSR profiles for the 150 hybrid parents. These 56 SSRs were reported to be distributed across all seven linkage groups in pearl millet genome (Supplemental Table S2).

The polymerase chain reactions (PCR) were performed on a GeneAmp 9700 thermal cycler (Applied Biosystems) in 10-μL reaction mixtures containing 10 ng of template DNA, 10× KAPA *Taq* polymerase buffer with $MgCl_2$, 1 mM of each deoxyribonucleotide triphosphate, and 0.25 U of *Taq* DNA polymerase in 384-well PCR plates. We added 2 pM μL^{-1} of forward and 4 pM μL^{-1} of reverse primers to each PCR reaction mixture. Polymerase chain reaction program cycle parameters included an initial denaturation step of 5 min at 94°C, followed by 10 cycles of 25 s denaturation at 94°C, 20 s primer annealing at 64°C, in which the temperature was decreased by 1°C with each cycle down to 54°C, and an extension step of 30 s at 72°C. This was followed by 37 cycles of 25 s at 94°C, 20 s at 56°C, and 30 s at 72°C, with a final extension step of 72°C for 20 min. Polymerase chain reactions of different loci were pooled according to the size of the amplicon and the usage of fluorophore dye. One microliter of the resulting mixture of the PCR product pool was combined with 7 μL of HiDi loading buffer (containing formamide), 0.1 μL of the LIZ-labeled (500 [−250]) internal size standard, and 3.9 μL of distilled water. Samples were denatured for 5 min at 95°C, quickly cooled on ice, and genotyped on an Applied Biosystems 3700 automatic DNA sequencer. Fragment scoring was done manually using Gene Mapper 4.0 (Applied Biosystems, 1998).

Genotyping Using SNPs

Genotyping by sequencing was used to identify SNPs from hybrid parents as described by Elshire et al. (2011). The DNA extracted from each line was digested using ApeKI endonuclease (recognition site: G/CWCG) for 2 h at 75°C, then ligated with adapters having unique multiplex sequence indices (barcodes) using ligase buffer with ATP and T4 ligase, and a small aliquot of such digested ligated fragments were purified to remove excess adapters. Finally, the PCR amplification of pooled amplicons was performed before sequencing on the Illumina Hiseq2500 platform. The pearl millet genome reference (Varshney et al., 2017) was used to identify SNPs in TASSEL version 4 software (Bradbury et al., 2007). The software takes the raw sequencing reads and barcode sequence information as input, wherein only barcodes containing reads were used for SNP calling. Such barcodes containing reads were trimmed to 64 bp from the barcode side, aligned against each other, and used for SNP identification. The SNPs identified were assigned to each hybrid parent according to the barcode sequence information.

Evaluating Association between Genetic Distance and Heterosis for Grain Yield

Seventeen parents (one to three parents representing each cluster) were identified considering the number of lines in the cluster, allelic variation, and genetic distance within the cluster (Supplemental Table S3). The range of genetic distances of 17 sampled lines was from 0.36 to 0.85, with an average of 0.69, which was representative of 147 lines (average of 0.68 and ranging from 0.17 to 0.90). One hundred and thirty-six hybrids were generated using these identified 17 hybrid parents (nine B lines and eight R lines) and crossing them in diallel fashion (without reciprocals). These hybrids were evaluated along with parents in a trial planted in an α-lattice design with two replications at each of two locations, Chaudhary Charan Singh Haryana Agricultural University, Hisar (in north India; 29°15′ N, 75°7′ E longitude, 215.2 m asl), and at ICRISAT, Patancheru (in south India; 17°3′ N, 78°27′ E, 545 m asl), during the rainy season of 2015. At Patancheru, each entry was planted in two rows of 4-m length with rows spaced 75 cm apart, and plants were maintained 15 cm apart. At Hisar, each entry was planted in four rows of 4-m length with rows spaced 50 cm apart, and plants were maintained 12 cm apart. All the recommended agronomic practices were followed at both the locations for good crop growth. All the panicles in a plot were harvested at physiological maturity for each entry separately. The harvested material was sundried for 10 to 15 d, threshed, and recorded for grain yield in kilograms.

Statistical Analysis

Out of the 150 hybrid parents, SSR data on 147 lines (73 B and 74 R lines) were analyzed, whereas two B lines and one R line could not be analyzed due to >20% missing data. The number of alleles per locus, major allele frequency, gene diversity, and polymorphism information content (PIC) for each locus were calculated using Power Marker 3.5 (Liu and Muse, 2005). The occurrence of common, most frequent, and rare alleles were determined following Li et al. (2008) and Upadhyaya et al. (2008). In the case of SNPs, 11 B and 22 R lines had significant missing data, hence 117 lines (64 B and 53 R lines) were analyzed with a set of 444,075 SNPs. Further, SNPs having a minor allele frequency <5% and missing data >30% were eliminated, resulting in 75,007 high-quality SNPs.

A simple matching allele frequency-based distance matrix was calculated in the DARwin 5.0 program (Perrier and Collet, 2006) using SSR markers, and the Rogers dissimilarity matrix (Rogers, 1972) was calculated using SNP markers in R (R Core Team, 2017). The DARwin program was used to construct unweighted neighbor-joining cluster diagrams for both the markers (SSRs and SNPs) to examine the genetic structure and relationship of hybrid parents. Principle coordinate analysis (PCoA) was performed for both the markers (SSRs and SNPs) using software GenAlEx version 6.5 (Peakall and Smouse, 2012). Analysis of molecular variance (AMOVA) was performed to estimate the variance components among and within B- and R-line groups using Arlequin software version 3.5.2.2 (Excoffier and Lischer, 2010), wherein the significance of fixation index value was tested with 10,100 permutations. Using the software GenAlEx 6.5, a Mantel test was performed to assess the association between genetic distance matrices generated from SSR and SNP markers. To infer population stratification using SNPs, ADMIXTURE software version 1.3.0 (Alexander et al., 2009) was used to estimate admixture proportions. STRUCTURE software version 2.3.1 was used to identify actual subpopulations (K). A range of 1 to 10 K values were run for 20 times each with a burn-in period of 10,000 steps, followed by 10,000 Markov chain Monte Carlo replicates. The mean posterior probability [ln(PD)] derived for each K was plotted to find the plateau of the ΔK values using online available program Structure Harvester (http://taylor0.biology.ucla.edu).

Analysis of variance was performed using PROC MIXED in SAS 9.4 (SAS Institute, 2017), considering location, genotype, and replication as fixed and block as random. To pool the data across two locations and to make the error variance homogeneous, individual location variances were estimated and modeled to error distribution using the residual maximum likelihood (REML) procedure. Genotype effect was further portioned into hybrids, parents, and hybrids vs. parents using the SAS CONTRAST statement. The phenotypic observations (Z_{ijkl}) on Accession l in replicate j of block k of location i was modeled as

$$Z_{ijkl} = \mu + e_i + (e/r)_{ij} + (e/r/b)_{ijk} + g_l + \text{eg}_{il} + \varepsilon_{ijkl}$$

where μ is the grand mean, e_i is the fixed effect of location I, g_l is the fixed effect of genotype l, $(e/r)_{ij}$ is the fixed effect of replication j nested with in location i, $(e/r/b)_{ijk}$ is the random effect of block k nested with in replication j and location i and is ~NID(0, σ^2_b), where σ^2_b is block variance; eg_{il} is the fixed effect of the interaction between genotype l and location i, and ε_{ijkl} is the random residual effect and is ~NID(0, σ^2_ε), where σ^2_ε is environmental variance.

Better parent heterosis (BPH) for grain yield was estimated as BPH = $100(F_1 - \text{BP})/\text{BP}$, where F_1 is hybrid yield and BP is the yield of the better-yielding parent. The significance of PH was tested using a t test [$t = (F_1 - \text{BP})/\sqrt{(2\text{MSE}/r)}$; where MSE is mean square error and r is the number of replications]. The association between genetic distance and heterosis was estimated using PROC CORR in SAS.

RESULTS

Genetic Divergence Statistics

SSRs

Fifty-six SSR loci detected 412 alleles in 147 hybrid parents (73 B and 74 R lines) of pearl millet with an average of 7.36 alleles per locus. The R lines had a higher number of alleles (384) with an average of 6.80 alleles per locus, in comparison with 332 alleles for B lines with an average of 5.90 alleles per locus. The allelic richness varied from two alleles on loci *Xpsmp2202* and *Xpsmp2267* to 16 alleles on loci *Xicmp3032* and *Xicmp3088*. Twelve SSR loci had >10 alleles, whereas 27 loci amplified 5 to 10 alleles per locus (Table 1). Furthermore, of the total 412 alleles detected, 35 were rare alleles (frequency $\leq$ 1%), 283 were common alleles (frequency = 1–20%), and 94 were the most frequent (frequency > 20%) alleles. The most frequent alleles ranged from one (15 loci) to three (*Xipes0220.1*, *Xpsmp2222*, and *Xpsmp2248*) with an average of 1.68. The rare alleles ranged from one (18 loci) to four (*Xpsmp2070*), averaging 0.63, whereas the common alleles ranged from 1 (nine loci) to 15 (*Xpsmp3032* and *Xipes0203*), averaging 5.05 alleles per locus (Table 1). Major allele frequency for all parental lines ranged from 0.14 (*Xipes0203*) to 0.89 (*Xpsmp2202*), averaging 0.42, whereas for B and R lines, major allele frequency ranged from 0.14 (*Xipes0203*) to 0.99 (*Xpsmp2246*) and 0.14 (*Xicmp3088*) to 0.85 (*Xipes0213*) with averages of 0.50 and 0.40, respectively. The PIC values for all parental lines ranged from 0.17 for the *Xpsmp2202* locus to 0.90 for the locus *Xipes0203*, averaging 0.65 (Table 1). Out of 56 SSR loci, 35 were found highly polymorphic with PIC value ranging from 0.68 to 0.90. Moreover, the PIC values ranged from 0.03 (*Xpsmp2246*) to 0.90 (*Xipes0203*) for B lines and from 0.25 (*Xipes0213*) to 0.91 (*Xipes0203*) for R lines, averaging 0.57 and 0.70, respectively. Gene diversity for all parental lines varied from 0.19 (*Xpsmp2202*) to 0.91 (*Xipes0203*), averaging 0.69 (Table 1). The R lines had higher average gene diversity (0.73, ranging from 0.27 to 0.92) than B lines (0.61, ranging from 0.03 to 0.91).

SNPs

A total of 75,007 SNPs were detected on 117 hybrid parents (Table 2). The PIC value for all parental lines ranged from 0.00 to 0.38 with an average of 0.27. Furthermore, the PIC value ranged from 0.02 to 0.38 (average of 0.26) and 0.02 to 0.38 (average of 0.27) for B and R lines, respectively. Gene diversity for all parental lines varied from 0.00 to 0.50, averaging 0.33. Among B and R lines, average gene diversity was 0.30 (ranging from 0.01 to 0.50) and 0.31 (ranging from 0.02 to 0.50), respectively. The average major allele frequency for all the hybrid parents ranged from 0.50 to 1.00 with a mean of 0.75. The results pertaining to allelic variation, gene diversity, and PIC values indicated sufficient polymorphism for all the SNPs loci detected in the plant material in this study.

Genetic Distance Estimates based on SSRs and SNPs

The pairwise SSR-based simple matching genetic distances for all the hybrid parents varied from 0.17 to 0.90 with an average of 0.68. The average genetic distance for B $\times$ R pairs was higher (0.72, ranging from 0.39 to 0.90) than for B $\times$ B (0.56, ranging from 0.17 to 0.81) and R $\times$ R pairs (0.69, ranging from 0.28 to 0.90) (Table 3, Fig. 1a) for SSRs. Genetic distance estimates using SNPs varied from 0.11 to 0.75 with an average of 0.62. Similarly, the SNP-based average genetic distance of B $\times$ R pairs (0.65, ranging from 0.34 to 0.75) was higher than for B $\times$ B (0.54, ranging from 0.11 to 0.75) and R $\times$ R pairs (0.63, ranging from 0.12 to 0.72). The maximum numbers of pairs had genetic distance of >0.60 for both SSRs and SNPs (Table 3, Fig. 1b). The above results revealed a wide range of genetic distance between any two parents in the material for both SSR and SNP markers. A Mantel test was performed to assess the relationship between genetic distances estimated based on SSRs and SNPs and revealed a significant positive correlation (0.58, $p < 0.01$). This correlation was also significant for B and R lines individually, although the value was higher for B lines (0.53, $p < 0.01$) than for R lines (0.23, $p < 0.01$).

Genetic Relatedness and Clustering Patterns

SSRs

Simple sequence repeat markers differentiated majority of the B (81%) and R lines (72%) into their separate respective groups (Fig. 2a). Fourteen B lines (19%) and 21 R lines (28%) were found in common cluster, although they formed separate groups within this mixed cluster. Further, SSRs grouped all B lines into five clusters, with 26 B lines in B-I, 14 in B-II, 16 in B-III, five in B-IV, and 12 in B-V (Fig. 2b). Hybrid parents sharing a common parent were found in the same group, as the majority of the B lines (67%) having 843B in their parentage were found clustered in B-I. In addition, 10 of the ICRISAT-bred B lines having ICMB 01222 (thick and long panicle line) in their parentage were found in cluster B-III. The B lines (i.e., B63 and B62) having thick and long panicles but from a different breeding center (Hisar) also grouped in the same cluster, B-III. Moreover, many of the trait-specific lines also pooled in the same cluster, as most of the B lines (60%) in the B-II cluster were bold seeded (having >10 g of 1000-seed weight); this cluster had four lines with ICMB 99111 (very bold-seeded B line) in their parentage. Five hybrid parents bred at ICRISAT and known for arid-type plants (early maturity, thin stem, small seeded, and small panicle) grouped in cluster B-I, along with hybrid parents developed at the Hisar center (primarily breeding plant types adapted to arid zone).In the case of restorer lines, SSRs grouped six R lines in R-I, 15 in R-II, 18 in R-III, eight in R-IV, six in R-V and 21 in R-VI (Fig. 2c).

Table 1. Allelic richness, allele size, polymorphic information content (PIC), and gene diversity of the 56 simple sequence repeat (SSR) loci in 147 hybrid parental lines (73 maintainer [B] lines and 74 restorer [R] lines) of pearl millet.

SSR loci	Total allele†	Allele size range	Rare alleles (≤1%)	Common alleles (1–20%)	Most frequent alleles (>20%)	Major allele frequency†	PIC†	Gene diversity†
		bp						
Xctm08	7	127–145	0	5	2	0.33 (0.30, 0.51)	0.75 (0.74, 0.60)	0.78 (0.78, 0.65)
Xctm10	10	162–177	0	8	2	0.33 (0.42, 0.27)	0.77 (0.61, 0.84)	0.79 (0.67, 0.86)
Xctm12	6	210–234	0	4	2	0.45 (0.45, 0.55)	0.62 (0.64, 0.60)	0.68 (0.69, 0.64)
Xicmp3002	11	200–206	2	8	1	0.26 (0.32, 0.19)	0.82 (0.79, 0.85)	0.84 (0.81, 0.87)
Xicmp3032	16	254–266	0	15	1	0.22 (0.40, 0.30)	0.88 (0.76, 0.85)	0.89 (0.78, 0.86)
Xicmp3043	7	236–254	0	5	2	0.53 (0.84, 0.47)	0.59 (0.25, 0.69)	0.64 (0.28, 0.72)
Xicmp3048	4	240–253	0	2	2	0.53 (0.88, 0.61)	0.54 (0.19, 0.53)	0.61 (0.22, 0.57)
Xicmp3080	6	150–154	0	4	2	0.37 (0.44, 0.31)	0.68 (0.58, 0.78)	0.72 (0.65, 0.81)
Xicmp3088	16	172–184	1	14	1	0.29 (0.42, 0.14)	0.85 (0.72, 0.91)	0.86 (0.75, 0.92)
Xipes0004	6	140–170	1	3	2	0.45 (0.78, 0.46)	0.66 (0.34, 0.70)	0.70 (0.37, 0.73)
Xipes0082	8	150–168	0	7	1	0.22 (0.27, 0.27)	0.84 (0.81, 0.79)	0.86 (0.83, 0.81)
Xipes0105	6	212–214	0	4	2	0.43 (0.49, 0.42)	0.58 (0.61, 0.59)	0.65 (0.66, 0.65)
Xipes0152.2	10	176–202	0	8	2	0.29 (0.41, 0.27)	0.79 (0.69, 0.80)	0.81 (0.73, 0.83)
Xipes0174	13	193–215	0	12	1	0.23 (0.32, 0.35)	0.87 (0.81, 0.80)	0.88 (0.83, 0.82)
Xipes0176	4	352–388	1	1	2	0.46 (0.67, 0.62)	0.56 (0.39, 0.49)	0.64 (0.47, 0.54)
Xipes0186	3	172–204	0	1	2	0.51 (0.64, 0.68)	0.44 (0.37, 0.42)	0.54 (0.47, 0.48)
Xipes0198	9	226–244	1	7	1	0.61 (0.82, 0.35)	0.57 (0.29, 0.80)	0.60 (0.31, 0.82)
Xipes0200	15	229–279	3	11	1	0.28 (0.27, 0.28)	0.85 (0.85, 0.84)	0.86 (0.86, 0.85)
Xipes0203	15	155–164	0	15	0	0.14 (0.14, 0.16)	0.90 (0.90, 0.90)	0.91 (0.91, 0.91)
Xipes0213	4	172–192	1	2	1	0.82 (0.77, 0.85)	0.29 (0.36, 0.25)	0.31 (0.39, 0.27)
Xipes0220.1	8	338–352	1	4	3	0.30 (0.32, 0.36)	0.73 (0.72, 0.75)	0.77 (0.76, 0.78)
Xipes0223	7	114–126	0	6	1	0.43 (0.71, 0.27)	0.73 (0.42, 0.82)	0.75 (0.46, 0.84)
Xipes0227	4	240–250	0	2	2	0.64 (0.53, 0.74)	0.48 (0.49, 0.40)	0.53 (0.57, 0.43)
Xipes0236	10	225–262	2	6	2	0.27 (0.26, 0.31)	0.77 (0.80, 0.77)	0.80 (0.82, 0.80)
Xpsmp2045	7	115–213	0	5	2	0.27 (0.23, 0.32)	0.77 (0.80, 0.77)	0.80 (0.82, 0.80)
Xpsmp2059	5	120–128	2	1	2	0.47 (0.41, 0.49)	0.53 (0.62, 0.61)	0.61 (0.68, 0.66)
Xpsmp2068	12	211–233	1	10	1	0.61 (0.90, 0.26)	0.59 (0.16, 0.86)	0.61 (0.18, 0.87)
Xpsmp2070	14	264–274	4	10	0	0.16 (0.22, 0.19)	0.88 (0.85, 0.87)	0.89 (0.86, 0.88)
Xpsmp2077	7	100–140	1	5	1	0.39 (0.32, 0.43)	0.73 (0.77, 0.71)	0.76 (0.80, 0.74)
Xpsmp2079.2	13	229–273	1	12	0	0.18 (0.23, 0.19)	0.87 (0.85, 0.88)	0.88 (0.87, 0.89)
Xpsmp2081.1	11	148–183	0	10	1	0.21 (0.4, 0.28)	0.86 (0.73, 0.85)	0.87 (0.76, 0.86)
Xpsmp2086	9	126–170	1	6	2	0.27 (0.29, 0.18)	0.80 (0.77, 0.86)	0.83 (0.80, 0.88)
Xpsmp2089	7	174–217	0	5	2	0.25 (0.30, 0.22)	0.78 (0.74, 0.79)	0.80 (0.78, 0.82)
Xpsmp2090	7	125–141	0	5	2	0.38 (0.48, 0.26)	0.69 (0.45, 0.77)	0.73 (0.55, 0.80)
Xpsmp2201	3	216–241	0	1	2	0.71 (0.77, 0.61)	0.36 (0.29, 0.49)	0.43 (0.36, 0.55)
Xpsmp2202	2	120–144	0	1	1	0.89 (0.82, 0.78)	0.17 (0.29, 0.33)	0.19 (0.31, 0.36)
Xpsmp2203	8	152–188	0	6	2	0.37 (0.41, 0.22)	0.71 (0.64, 0.84)	0.75 (0.69, 0.86)
Xpsmp2204	6	151–168	1	3	2	0.57 (0.56, 0.45)	0.56 (0.58, 0.69)	0.61 (0.62, 0.72)
Xpsmp2207	4	150–154	1	1	2	0.48 (0.60, 0.62)	0.53 (0.51, 0.53)	0.61 (0.56, 0.57)
Xpsmp2208	7	112–122	0	5	2	0.39 (0.48, 0.22)	0.72 (0.61, 0.82)	0.75 (0.66, 0.84)
Xpsmp2209	3	164–166	0	1	2	0.51 (0.41, 0.47)	0.45 (0.61, 0.61)	0.55 (0.67, 0.67)
Xpsmp2211	4	352–389	0	2	2	0.59 (0.44, 0.43)	0.46 (0.64, 0.60)	0.54 (0.69, 0.67)
Xpsmp2212	6	264–274	1	3	2	0.41 (0.37, 0.49)	0.70 (0.73, 0.66)	0.74 (0.76, 0.69)
Xpsmp2214	4	170–184	0	2	2	0.45 (0.47, 0.43)	0.52 (0.52, 0.60)	0.60 (0.60, 0.67)
Xpsmp2218	5	167–182	1	2	2	0.60 (0.62, 0.51)	0.49 (0.40, 0.63)	0.55 (0.49, 0.67)
Xpsmp2220	8	320–328	2	4	2	0.35 (0.48, 0.22)	0.75 (0.64, 0.81)	0.78 (0.69, 0.84)
Xpsmp2222	6	156–182	1	2	3	0.38 (0.66, 0.47)	0.66 (0.47, 0.64)	0.71 (0.52, 0.69)
Xpsmp2227	4	260–276	0	2	2	0.55 (0.71, 0.49)	0.42 (0.34, 0.58)	0.53 (0.42, 0.64)
Xpsmp2232	8	207–253	0	6	2	0.32 (0.60, 0.27)	0.75 (0.42, 0.80)	0.78 (0.51, 0.82)
Xpsmp2237	6	108–115	1	4	1	0.47 (0.63, 0.22)	0.67 (0.52, 0.80)	0.71 (0.55, 0.83)
Xpsmp2246	3	102–105	0	1	2	0.67 (0.99, 0.57)	0.39 (0.03, 0.46)	0.47 (0.03, 0.55)
Xpsmp2248	6	320–326	1	2	3	0.37 (0.34, 0.36)	0.67 (0.68, 0.72)	0.72 (0.73, 0.76)
Xpsmp2249	6	257–263	2	2	2	0.47 (0.51, 0.36)	0.62 (0.61, 0.71)	0.67 (0.66, 0.75)
Xpsmp2251	6	196–206	0	4	2	0.47 (0.47, 0.39)	0.63 (0.65, 0.73)	0.68 (0.69, 0.76)
Xpsmp2267	2	261–265	0	1	1	0.86 (0.92, 0.76)	0.21 (0.15, 0.31)	0.24 (0.15, 0.37)
Xpsmp2273	8	146–155	1	5	2	0.28 (0.29, 0.23)	0.77 (0.77, 0.81)	0.80 (0.80, 0.83)
Total	412 (332, 384)		35	283	94			
Mean	7.36 (5.90, 6.80)		0.63	5.05	1.68	0.42 (0.50, 0.40)	0.65 (0.57, 0.70)	0.69 (0.61, 0.73)

† Total alleles, major allele frequency, PIC, and gene diversity for B and R lines, respectively, are given in parentheses.

Ten out of fourteen restorers bred at the Hisar center grouped in clusters R-I and R-II. The R-III cluster had 60% of the R lines derived from MRC (Mandore restorer composite), whereas eight R lines sharing SDMV 90031 as a common parent grouped in cluster R-VI.

SNPs

Similarly, the SNPs partitioned the majority of B (78%) and R lines (64%) into two clear-cut separate groups. Fourteen B lines (22%) and 17 R lines (32%) were found grouped in a common cluster (Fig. 3a). Further, SNPs grouped all the B lines into five clusters: 17 in B-I, nine in B-II, 27 in B-III, seven in B-IV, and five in B-V (Fig. 3b). Similar to the SSR pattern, most of the lines (12) having ICMB 01222 in their parentage grouped together in cluster B-III. Also, two B lines with thick and long panicles (B62, B63, B65, B67, and B68) bred at the Hisar center were found grouped in same cluster, B-III. Eight B-lines having the common parent 843B in there parentage were found in group B-I. The ICRISAT-bred arid-type lines (thin stem, small seeded, and small panicles) and most of the B lines from the Hisar center (bred for arid ecology) were found in common clusters B-III and B-I. Furthermore, B lines known for bold seededness grouped in cluster B-II,

Table 2. Average polymorphism in hybrid parents of pearl millet based on simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs).

Groups†	No. of alleles‡	Major allele frequency‡	Gene diversity‡	PIC value‡
SSRs				
B and R lines	7.4 (2–16)	0.42 (0.14–0.89)	0.69 (0.19–0.91)	0.65 (0.17–0.90)
B lines	5.9 (1–14)	0.50 (0.14–0.99)	0.61 (0.03–0.91)	0.57 (0.03–0.90)
R lines	6.8 (2–15)	0.40 (0.14–0.85)	0.73 (0.27–0.92)	0.70 (0.25–0.91)
SNPs				
B and R lines	2.0 (1–2)	0.75 (0.50–1.00)	0.33 (0.00–0.50)	0.27 (0.00–0.38)
B lines	2.0 (1–2)	0.77 (0.50–0.99)	0.30 (0.01–0.50)	0.26 (0.02–0.38)
R lines	2.0 (1–2)	0.75 (0.50–0.99)	0.31(0.02–0.50)	0.27 (0.02–0.38)

† B line, maintainer line; R line, restorer line.

‡ Range of alleles, major allele frequency, gene diversity, and polymorphic information content (PIC) is given in parentheses.

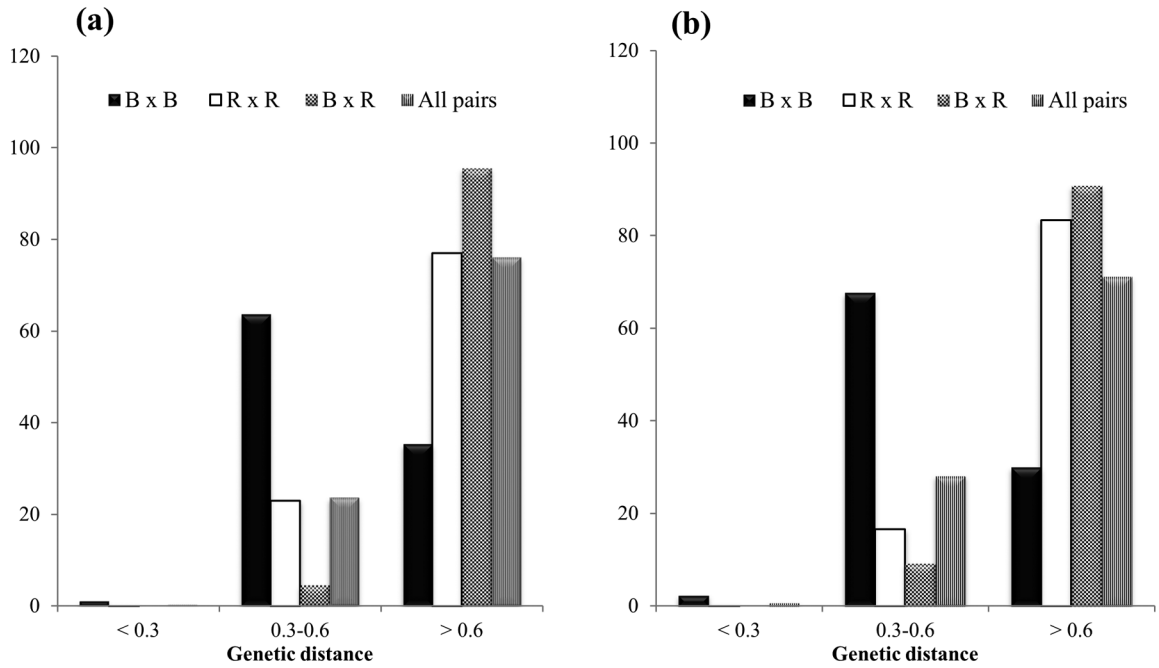

Fig. 1. Frequency distribution of genetic distance between hybrid parents using (a) simple sequence repeats and (b) single nucleotide polymorphisms. B, maintainer lines; R, restorer lines.

Table 3. Average (minimum–maximum) of simple matching genetic distance of pearl millet hybrid parental lines based on simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs).

	Group†			
Marker type	All line pairs	B × B pairs	R × R pairs	B × R pairs
SSR	0.68 (0.17–0.90)	0.56 (0.17–0.81)	0.69 (0.28–0.90)	0.72 (0.39–0.90)
SNP	0.62 (0.11–0.75)	0.54 (0.11–0.75)	0.63 (0.12–0.72)	0.65 (0.34–0.75)

† B line, maintainer line; R line, restorer line.

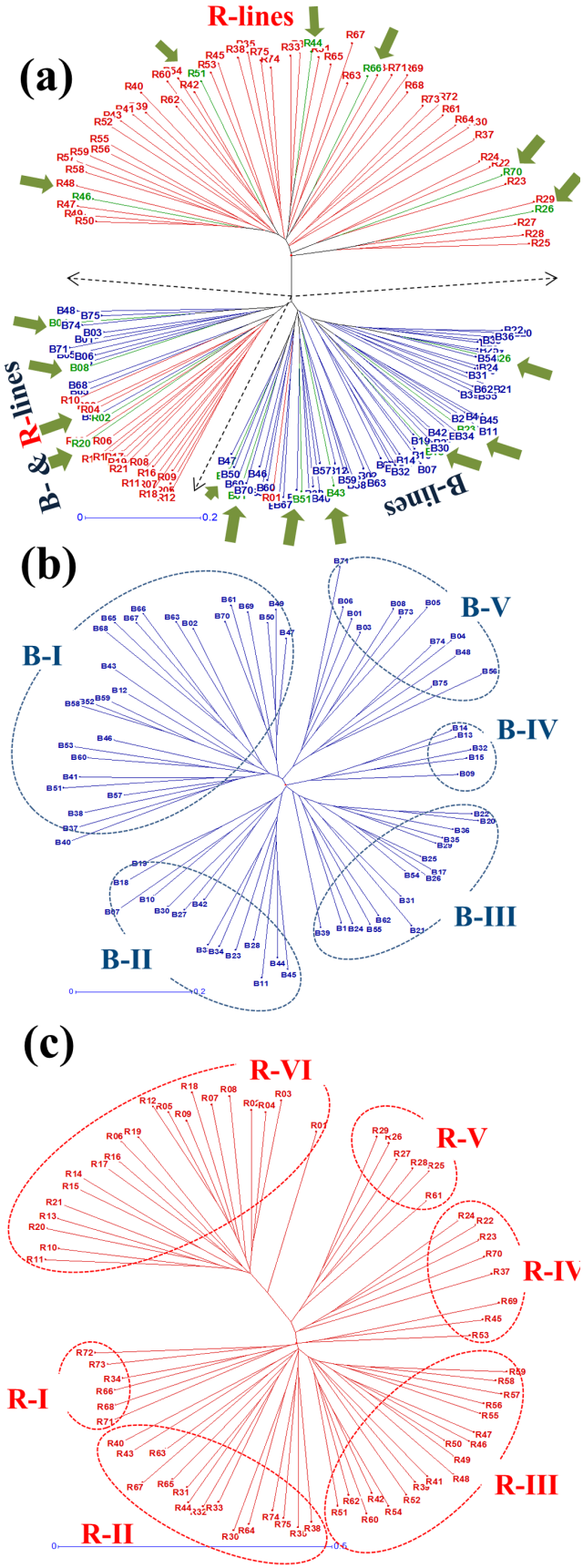

Fig. 2. Unweighted neighbor-joining tree of (a) all 147 hybrid parents using (b) 73 maintainer (B) lines and (c) 74 restorer (R) lines using simple sequence repeats. Red and blue coloring indicates R and B lines, respectively. Green arrows indicate the selected representative parental lines.

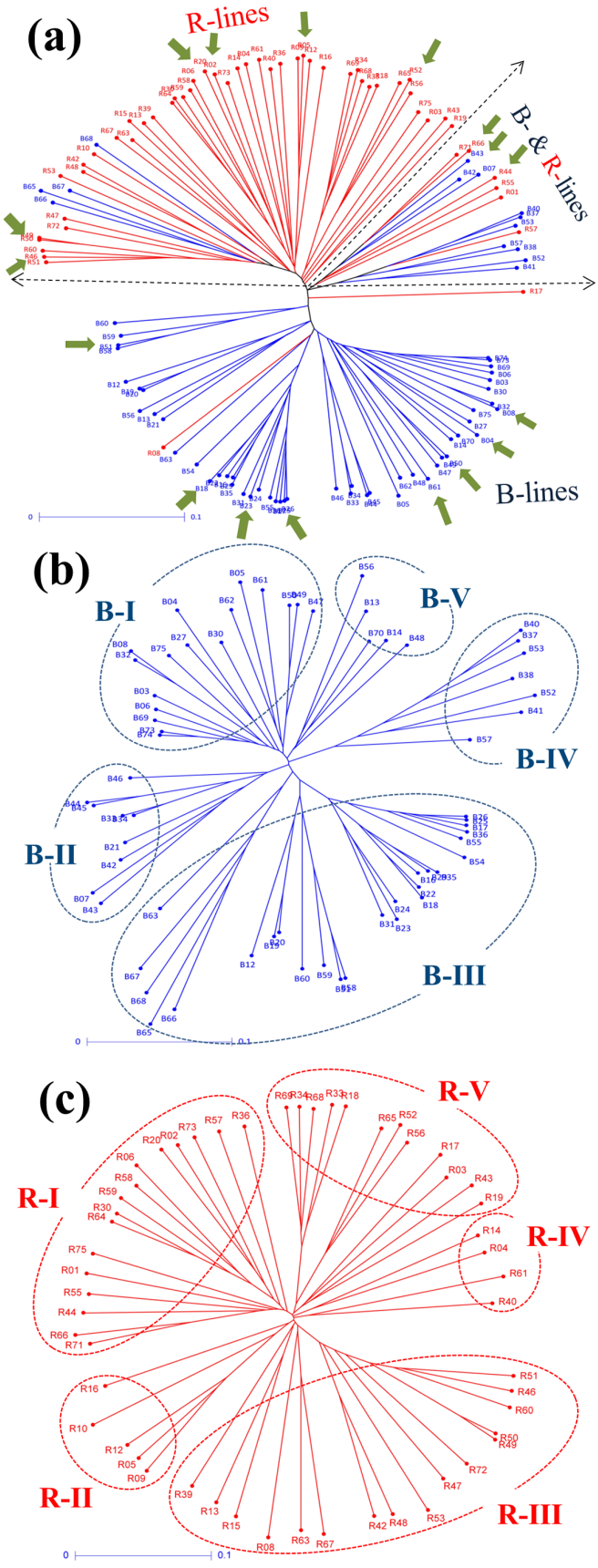

Fig. 3. Unweighted neighbor-joining tree of (a) all 117 hybrid parents using (b) 64 maintainer (B)lines and (c) 53 restorer (R) lines using single nucleotide polymorphisms. Red and blue coloring indicates R and B lines, respectively. Green arrows indicate the selected representative parental lines.

with most of the lines in this group having ICMB 99555 (bold-seeded B line) in their parentage. Likewise, SNPs grouped 16 R lines in R-I, five lines in R-II, 16 lines in R-III, four lines in R-IV, and 12 lines in R-V (Fig. 3c). Again, 5 out of 12 R lines bred at Hisar for arid adaptation grouped in common cluster R-I, whereas seven R lines bred at ICRISAT, with a composite well known for its high drought tolerance (MRC) in their parentage, were grouped in R-III

Analysis of Molecular Variance and Structure Analysis

Results of AMOVA showed significant difference between B- and R-line groups, which accounted for 13 and 8% of the total genetic variation detected by SSRs and SNPs, respectively (Table 4). The within-group variation, both in B and R lines, was much larger and accounted for 87 and 92% of the total variation detected by SSRs and SNPs, respectively. The fixation index value varied from 0.001 (*Xpsmp* 2211) to 0.571 (*Xpsmp* 2246) with an average of 0.15 per locus for SSRs, while it was 0.07 for SNPs, indicating that significant genetic variation was detected at molecular level using SSRs and SNPs.

In the case of SSRs, log likelihood had two as the optimum value of K. Also, the maximum ΔK was found at $K = 2$, indicating that the entire population can be partitioned into two groups. The accessions with the probability of $\geq$0.80 membership fractions were assigned to corresponding groups, whereas others were categorized as admixture. The first group consisted of 94 parental lines with the majority of the B lines (73) along with 21 R lines, whereas the second group consisted of only R lines (53) (Supplemental Fig. S2a). Similarly, in the case of SNPs, the first group consisted of only B lines (36), whereas the second group consisted of 55 parental lines with most of the R lines (49) along with six B lines, and 26 parental lines were retained as admixture (Supplemental Fig. S2b). Results pertaining to PCoA further support that B and R lines form two separate broad-based groups in pearl millet (Supplemental Fig. S3). The intermixing of B and R lines across Axis 1 for both SSRs and SNPs showed the relatedness of some of seed and restorer parents.

Grain Yield and Better Parent Heterosis

Grain yield of the hybrids ranged from 1719 (B26 × B61) to 5386 kg ha^{-1} (B49 × R66) at Hisar, whereas it varied from 2982 (R46 × R51) to 7012 kg ha^{-1} (B49 × R70) at Patancheru. Across the locations, better parent heterosis varied from −13.1 to 175.9% with a mean of 32.6%, and 109 hybrids out of 136 had significant positive BPH (Supplemental Table S4). Also, hybrids from the B × R group had higher grain yield and BPH (4185 kg ha^{-1} and 74.4%, respectively) than B × B (3597 kg ha^{-1} and 50.5%) and R × R groups (3851 kg ha^{-1} and 60.8%). Analysis of variance for grain yield showed large and highly significant variation due to location, indicating that the materials were evaluated under diverse locations (Table 5). Large and highly significant variation observed due to parents and hybrids indicated wide genetic differences among parental lines, as well as among hybrids. Highly significant mean squares due to "hybrids vs. parents" indicated significant heterosis in crosses. Effect of genotype × environment interaction was significant; however, crossover interaction (Ponnuswamy et al., 2018) was negligible (2.56%), suggesting that ranks of hybrids and parents for yield were similar in both the locations. The correlation of genetic distance and BPH was significant for both SSR- (0.33, $p < 0.01$) and SNP-based (0.35, $p < 0.01$) marker systems.

DISCUSSION

Genetic Divergence

The SSR marker-based (56 loci) analysis of 147 pearl millet hybrid parents indicated presence of 7.36 alleles per locus. This was more than the previously reported values of 6.26 alleles per locus (Kapila et al., 2008), 3.0 alleles per locus (Singh et al., 2013a), and 2.76 alleles per locus (Sumathi et al., 2013) in pearl millet, which might be because of the greater number of hybrid parents involved in present study (147) in comparison with the 72, 20, and 42 hybrid parents in the studies mentioned above, respectively. However, the present set of hybrid parents had fewer alleles per locus than reported by Gupta et al. (2015) (12.68 alleles per locus) and Nepolean et al. (2012) (8.10 alleles per locus), which could be because of the involvement of more lines (379 and 213, respectively, in these studies). Stich et al. (2010)

Table 4. Analysis of molecular variance using simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) for maintainer (B) and restorer (R) lines of pearl millet.

	SSR					SNP				
Source	df	SS†	MS‡	Estimated variance	Percentage	df	SS	MS	Estimated variance	Percentage
Between B and R lines	1	3.86	3.86	0.05	13	1	4,871	4,871	38	8
Within B and R lines	145	45.42	0.31	0.31	87	232	109,076	470	470	92
Total	146	49.28				233	113,948			
P				0.001					0.0001	
Fixation index				0.15					0.07	

† SS, sum of squares.

‡ MS, means square.

Table 5. Analysis of variance for grain yield.

Source of variation	Numerator df	Denominator df	*F* value	Pr > *F*
Environment	1	23.5	1342.57	<0.0001
Replication (environment)	2	22.2	17.45	<0.0001
Genotype	156	248	18.05	<0.0001
Hybrid (H)	139	260	14.10	<0.0001
Parent (P)	16	273	5.50	<0.0001
H vs. P	1	24.5	759.54	<0.0001
Environment × genotype	156	248	3.73	<0.0001
Environment × H	139	260	3.91	<0.0001
Environment × P	16	273	1.46	0.1162
Environment× (H vs. P)	1	24.5	15.07	0.0007

also reported more alleles per locus (16.40) than the present study, with similar number of inbred lines (145), which seems to be due to the involvement of quite a diverse set of inbreds derived from landraces of Western and Central African origin. In the present study, SSR markers detected more alleles among R lines (384 alleles, with an average of 6.80 alleles per locus) than among B lines (332 alleles, with an average of 5.90 alleles per locus). In addition, the average PIC and gene diversity values were higher for R lines (0.70 and 0.73, respectively) than for B lines (0.57 and 0.61, respectively). This indicated that R lines were more genetically diverse than B lines, which was as expected due to the involvement of broader genetic base germplasm in the development of restorer lines. Similar findings on R lines being more diverse than B lines were earlier reported in pearl millet (Nepolean et al., 2012; Gupta et al., 2015). This suggests that there is a need to broaden the genetic base of B-line development programs.

Genetic Relatedness and Clustering Patterns

The genetic diversity analysis based on both SSRs and SNPs partitioned B and R lines into two clear-cut groups, indicating the existence of two diverse broad-based groups, each one representing the seed and restorer parents in pearl millet. The results are consistent with the previous studies on clustering of B and R lines as separate groups based on molecular markers in pearl millet (Nepolean et al., 2012; Gupta et al., 2015). Similarly, He et al. (2012) found B and R lines to exist as two broad groups in a study on 168 hybrid parents of rice. Two separate diversity studies in maize also partitioned diverse set of populations into two broad groups (i.e., Red Yellow Dent and Lancaster Sure Crop) groups (Dudley et al., 1991), and Flint and Dent groups (Dhillon et al., 1993). Results of AMOVA, PCoA, and Structure analysis in our study also confirmed significant diversity between B and R lines. Although the majority of B and R lines were found together in clusters dominated by B and R lines, respectively, few B lines grouped with clusters dominated by R lines and vice versa. This might be due to the occasional involvement of some trait-specific donor B lines in the R-line breeding program and vice versa (Gupta et al., 2015). For instance, four R lines that grouped with B lines had involvement of a progeny from thick-panicle seed parent composite HHVBC (high head volume B line composite), which was used as a donor parent for improving panicle thickness in the R-line breeding program. Hence, the utilization of a common donor line should be avoided in both B- and R-line development programs to maintain significant genetic distance between these two gene pools.

Hybrid parents having common parents in their pedigrees and trait similarities were found in the same clusters, indicating that molecular markers were able to identify and group genetically related breeding lines. For instance, 10 B lines sharing ICMB 01222 in their parentage grouped in cluster B-III (of both SSR and SNP), whereas six lines with 843B in their parentage grouped in cluster B-I. However, B lines sharing 843B in their parentage were scattered in all clusters of B lines, which might be due to this line being used as a parent in a diverse array of crosses in the B-line development program (Kapila et al., 2008). Also, many of R lines sharing a common parent derived from MRC, the composite with arid-type breeding material, were found grouped in R-III (of both SSR and SNP), and many of the R lines sharing SDMV 90031 as a common parent grouped in a common cluster.

Many of the hybrid parents bred at a particular breeding program grouped in same cluster, which again seems to be due to presence of same set of breeding lines in their parentage that have been frequently used in crossing program at that location. For instance, 8 out of 13 B lines bred at the Hisar center grouped in the same cluster, B-I. Also, 11 out of 15 R-lines from Hisar grouped in clusters R-I (five) and R-II (six). Nepolean et al. (2012) and Gupta et al. (2015) reported that pearl millet parental lines bred at Gujrat Agricultural University, Jamnagar, India, grouped in the same cluster in an investigation on pearl millet hybrid parents using SSRs.

Correlation between Genetic Distances and Heterosis

The positive significant correlation (r = 0.58, < 0.001) between SSR- and SNP-based genetic distances among hybrid parents must be responsible for the similar genetic

diversity patterns observed in this study based on both SSRs and SNPs. However, the molecular-marker-based diversity patterns in this study revealed that the number of alleles per locus, gene diversity, and PIC values based on SSR markers were higher than those analyzed from SNP markers. Similar findings on SSRs being better indicators of genetic diversity were reported by previous studies (Van Inghelandt et al., 2010; Varshney et al., 2010; He et al., 2012; Singh et al., 2013b). This might be because SNP markers are usually biallelic and the maximum gene diversity observable with biallelic markers is 0.5, vs. a maximum of 1.0 for SSR-type multiallelic markers. So SNP markers are less polymorphic than SSRs on an individual marker basis, but this limitation can be compensated for by the abundance of SNPs in the genome (Jones et al., 2007). It was also noted that genotyping cost using SNPs was 41% less than for SSRs, and the ratio of detected loci was 1:364 (SSR/SNP) in this set of hybrid parents. Under such a scenario of high correlation between genetic distances estimated using SSRs and SNPs, finding almost the same levels of genetic diversity patterns, and considering factors such as cost and the number of total detected loci, we suggest that SNPs are a better marker system than SSRs for capturing genetic relatedness in hybrid parents. Similar findings about SNP applicability over SSRs were reported by Hamblin et al. (2007), Jones et al. (2007), and Yang et al. (2009).

The results from 136 hybrids developed using parents involved in this study reported moderate and similar values of positive significant correlation between SSR- and SNP-based genetic distance, and better parent heterosis for grain yield. Considering the comparable results using SSRs or SNPs but the lower cost associated with SNPs, this indicated that SNPs should be a better marker system for heterotic prediction and further for formulating heterotic groups in pearl millet. Results from earlier studies in pearl millet and other crops have also indicated that markers can be a powerful tool for heterosis prediction and heterotic grouping (Reif et al., 2003; Wang et al., 2015; Ramya et al., 2018).

CONCLUSION

Overall, estimates of genetic distance between pearl millet hybrid parents using SSRs and SNPs were found positively correlated, leading to similar clustering patterns of hybrid parents. Both types of markers were well able to capture the genetic relatedness between genotypes based on pedigrees and traits. Considering the advantage of SNPs over SSRs for having lower cost, more loci detected, and comparable heterotic prediction values, we suggest that SNPs are preferable over SSR markers for genetic diversity and heterosis prediction investigations.

Conflict of Interest

The authors declare that there is no conflict of interest.

Supplemental Material Available

Supplemental material for this article is available online.

Acknowledgments

The research reported here was financially supported by the ICRISAT Pearl Millet Hybrid Parent Research Consortium (PMHPRC) and was conducted under the CGIAR Research Program on Dryland Cereals.

References

Alexander, D.H., J. Novembre, and K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19:1655–1664. doi:10.1101/gr.094052.109

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

Barbara, J. 2011. The relationship between heterosis and genetic distances based on RAPD and AFLP markers in carrot. Plant Breed. 130:574–579.

Bashir, E.M.A., A.M. Ali, E.T.I. Mohamed, A.E. Melchinger, H.K. Parzies, and B.I.G. Haussmann. 2014. Genetic diversity of Sudanese pearl millet [*Pennisetum glaucum* (L.) R. Br.] landraces as revealed by SSR markers, and relationship between genetic and agro-morphological diversity. Genet. Resour. Crop Evol. 62:579–591. doi:10.1007/s10722-014-0183-5

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635. doi:10.1093/bioinformatics/btm308

Burton, G.W. 1965. Pearl millet Tift 23A released. Crops Soils 17:19.

Burton, G.W. 1969. Registration of pearl millet inbreds Tift $23B_1$, Tift $23A_1$, Tift $23D_2B_1$, and Tift $23D_2A_1$. Crop Sci. 9:397. doi:10.2135/cropsci1969.0011183X000900030064x

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

Dhillon, B.S., J. Boppenmaier, W.G. Pollmer, R.G. Herrmann, and A.E. Melchinger. 1993. Relationship of restriction fragment length polymorphisms among European maize inbreds with ear dry matter yield of their hybrids. Maydica 38:245–248.

Dias, L.A., E.A. Picoli, R.B. Rocha, and A.C. Alfenas. 2004. A priori choice of hybrid parents in plants. Genet. Mol. Res. 3:356–368.

Dudley, J.W., M.A. Saghai-Maroof, and G.K. Rufener. 1991. Molecular markers and grouping of parents in maize breeding programs. Crop Sci. 31:718–723. doi:10.2135/cropsci1991.0011183X003100030036x

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A robust, simple genotyping- by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379. doi:10.1371/journal.pone.0019379

Excoffier, L., and H.E.L. Lischer. 2010. Arlequin Suite ver 3.5, a new series of programs to perform population genetics Analyses under Linux and Windows. Mol. Ecol. Resour. 10:564–567. doi:10.1111/j.1755-0998.2010.02847.x

Gupta, S.K., T. Nepolean, S.M. Sankar, A. Rathore, R.R. Das, K.N. Rai, and C.T. Hash. 2015. Patterns of molecular diversity in current and previously developed hybrid parents of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. Am. J. Plant Sci. 6:1697–1712. doi:10.4236/ajps.2015.611169

Gupta, S.K., T. Nepolean, C.G. Shaikh, K.N. Rai, C.T. Hash, R.R. Das, and A. Rathore. 2017. Phenotypic and molecular diversity-based prediction of heterosis in pearl millet (*Pennisetum glaucum* L. (R.) Br.). The Crop J. 6:271–281. doi:10.1016/j.cj.2017.09.008

Hamblin, M.T., M.L. Warburton, and E.S. Buckler. 2007. Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. PLoS One 2:e1367. doi:10.1371/journal.pone.0001367

He, Z.Z., F.M. Xie, L.Y. Chen, and M.A. Dela-Paz. 2012. Genetic diversity of tropical hybrid rice germplasm measured by molecular markers. Rice Sci. 19:193–201. doi:10.1016/S1672-6308(12)60040-7

Jones, E.S., H. Sullivan, D. Bhattramakki, and J.S.C. Smith. 2007. A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (*Zea mays* L.). Theor. Appl. Genet. 115:361–371. doi:10.1007/s00122-007-0570-9

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.

Kiula, B.A., N.G. Lyimo, and A.M. Botha. 2008. Association between AFLP-based genetic distance and hybrid performance in tropical maize. Plant Breed. 127:140–144. doi:10.1111/j.1439-0523.2007.01434.x

Li, Y., R. Guan, L. Zhangxiong, M. Yansong, W. Lixia, L. Linhai, et al. 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., and S.V. Muse. 2005. Power Marker: An integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129. doi:10.1093/bioinformatics/bti282

Mariac, C., V. Luong, I. Kapran, A. Mamadou, F. Sagnard, M. Deu, et al. 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

Moumouni, K.H., B.A. Kountche, M. Jean, C.T. Hash, Y. Vigouroux, B.I.G. Haussmann, and F. Belzile. 2015. Construction of genetic map for pearl millet, *Pennisetum glaucum* (L.) R. Br., using a genotyping by sequencing approach. Mol. Breed. 35:5. doi:10.1007/s11032-015-0212-x

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

Nepolean, T., S.K. Gupta, S.L. Dwivedi, R. Bhattacharjee, K.N. Rai, and C.T. Hash. 2012. Genetic diversity in maintainer and restorer lines of pearl millet. Crop Sci. 52:2555–2563. doi:10.2135/cropsci2011.11.0597

Peakall, R., and P. Smouse. 2012. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539. doi:10.1093/bioinformatics/bts460

Perrier, X., and J.P. Collet. 2006. DARwin software. Ctr. Coop. Int. Rech. Agron. Dév., Paris. http://darwin.cirad.fr/darwin (accessed 8 Aug. 2016).

Ponnuswamy, R., A. Rathore, A. Vemula, R.R. Das, A.K. Singh, D. Balakrishnan, et al. 2018. Analysis of multi-location data of hybrid rice trials reveals complex genotype by environment interaction. Cereal Res. Commun. 46:146–157. doi:10.1556/0806.45.2017.065

R Core Team. 2017. R: A language and environment for statistical computing. R Found. Stat. Comput., Vienna.

Rajaram, V., T. Nepolean, S. Senthilvel, R.K. Varshney, V. Vadez, R.K. Srivastava, et al. 2013. Pearl millet [*Pennisetum glaucum* (L.) R. Br.] consensus linkage map constructed using four RIL mapping populations and newly developed EST-SSRs. BMC Genom. 14:159. doi:10.1186/1471-2164-14-159

Ramya, K., M.L. Ahamed, C.T. Satyavathi, A. Rathore, P. Katiyar, A.G.B. Raj, et al. 2018. Towards defining heterotic gene pools in pearl millet [*Pennisetum glaucum* (L) R Br]. Front. Plant Sci. 8:1934. doi:10.3389/fpls.2017.01934

Rao, V.N., K.P.C. Rao, S.K. Gupta, M. Kizito, K.D. Charyulu, N. Nagraj, et al. 2018. Impact of ICRISAT Pearl Millet Hybrid Parents Research Consortium (PMHPRC) on the livelihoods of Farmers in India. Res. Rep. 75. ICRISAT, Telangana, India.

Reif, J.C., A.E. Melchinger, X.C. Xia, M.L. Warburton, D.A. Hoisington, S.K. Vasal, et al. 2003. Use of SSRs for establishing heterotic groups in subtropical maize. Theor. Appl. Genet. 107:947–957. doi:10.1007/s00122-003-1333-x

Rogers, J.S. 1972. Measures of genetic similarity and genetic distance: Studies in genetics. Univ. of Texas, Austin.

SAS Institute. 2017. SAS/STAT 14.3 user's guide. SAS Inst., Cary, NC.

Satyavathi, C.T., editor. 2017. Project coordinator review: 52nd Annual Group Meeting of the All India Coordinated Pearl Millet Improvement Project, Ludhiana, India. 28–30 Apr. 2017. All India Coord. Res. Proj. Pearl Millet, Jodhpur, India. http://www.aicpmip.res.in/pcr2017.pdf (accessed 16 June 2018).

Senthilvel, S., B. Jayashree, V. Mahalakshmi, P.S. Kumar, S. Nakka, T. Nepolean, and C. 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

Singh, A.M., M.K. Rana, S. Singh, S. Kumar, D. Kumar, and L. Arya. 2013a. Assessment of genetic diversity among pearl millet [*Pennisetum glaucum* (L.) R Br.] cultivars using SSR markers. Range Manag. Agrofor. 34:77–81.

Singh, N., D.R. Choudhury, A.K. Singh, S. Kumar, K. Srinivasan, R.K. Tyagi, et al. 2013b. Comparison of SSR and SNP markers in estimation of genetic diversity and population structure of Indian rice varieties. PLoS One 8:e84136. doi:10.1371/journal.pone.0084136

Stich, B., B.I.G. Haussmann, P. Raj, 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 environmental parameters. BMC Plant Biol. 10:216. doi:10.1186/1471-2229-10-216

Sumathi, P., M. Sumathi, N.K. Vinodhana, and M. Sathya. 2013. Assessment of genetic distance among the inbred lines of pearl millet [*Pennisetum glaucum* (L.) R. Br.] using SSR markers. Indian J. Biotechnol. Allied Fields 1:153–162.

Upadhyaya, H.D., S.L. Dwivedi, M. Baum, R.K. Varshney, S.M. Udupa, C.L.L. Gowda, et al. 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

Van Inghelandt, D., A.E. Melchinger, C. Lebreton, and B. Stich. 2010. Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. Theor. Appl. Genet. 120:1289–1299. doi:10.1007/s00122-009-1256-2

Varshney, R.K., M. Baum, P. Guo, S. Grando, S. Ceccarelli, and A. Graner. 2010. Features of SNP and SSR diversity in a set of ICARDA barley germplasm collection. Mol. Breed. 26:229–242. doi:10.1007/s11032-009-9373-9

Varshney, R.K., C. Shi, M. Thudi, C. Mariac, J. Wallace, P. Qi, et al. 2017. Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. Nat. Biotechnol. 35:969–976. doi:10.1038/nbt.3943

Wang, K., F.L. Qiu, W. Larazo, M.A. Dela Paz, and F.M. Xie. 2015. Heterotic groups of tropical *indica* rice germplasm. Theor. Appl. Genet. 128:421–430. doi:10.1007/s00122-014-2441-5

Yadav, S., S. Jain, V. Jain, and R.K. Jain. 2007. Genetic analysis of CMS, restorer, hybrid and open pollinated genotypes of Indian pearl millet [*Pennisetum glaucum* (L.) R. Br.] using ISSR markers. Indian J. Biotechnol. 6:340–348.

Yadav, O.P., and K.N. Rai. 2013. Genetic improvement of pearl millet in India. Agric. Res. 2:275–292. doi:10.1007/s40003-013-0089-z

Yadav, O.P., K.N. Rai, B.S. Rajpurohit, C.T. Hash, R.S. Mahala, S.K. Gupta, et al. 2012. Twenty-five Years of Pearl millet improvement in India. All India Coord. Pearl Millet Improv. Proj., Jodhpur, India.

Yadav, O.P., K.N. Rai, H.P. Yadav, B.S. Rajpurohit, S.K. Gupta, A. Rathore, and C.G. Karjagi. 2016. Assessment of diversity in commercial hybrids of pearl millet in India. Indian J. Plant. Genet. Resour. 29:130–136. doi:10.5958/0976-1926.2016.00018.8

Yang, X., Y. Xu, T. Shah, H. Li, Z. Han, J. Li, and J. Yan. 2009. Comparison of SSRs and SNPs in assessment of genetic relatedness in maize. Genetica (The Hague) 139:1045. doi:10.1007/s10709-011-9606-9