Theoretical and Applied Genetics
https://doi.org/10.1007/s00122-019-03512-z

ORIGINAL ARTICLE

Identification of heterotic groups in South-Asian-bred hybrid parents of pearl millet

S. K. Gupta[1] · K. Sudarshan Patil[1] · Abhishek Rathore[1] · Dev Vart Yadav[2] · L. D. Sharma[3] · K. D. Mungra[4] · H. T. Patil[5] · Suresh K. Gupta[6] · Ramesh Kumar[2] · Vaibhav Chaudhary[1] · Roma R. Das[1] · Anil Kumar[1] · Vikas Singh[1,7] · Rakesh K. Srivastava[1] · Rajeev Gupta[1] · M. Boratkar[1] · Rajeev K. Varshney[1] · K. N. Rai[1] · O. P. Yadav[8]

Received: 7 December 2018 / Accepted: 13 December 2019

Abstract

Key message **Pearl millet breeding programs can use this heterotic group information on seed and restorer parents to generate new series of pearl millet hybrids having higher yields than the existing hybrids.**

Abstract Five hundred and eighty hybrid parents, 320 R- and 260 B-lines, derived from 6 pearl millet breeding programs in India, genotyped following RAD-GBS (about 0.9 million SNPs) clustered into 12 R- and 7 B-line groups. With few exceptions, hybrid parents of all the breeding programs were found distributed across all the marker-based groups suggesting good diversity in these programs. Three hundred and twenty hybrids generated using 37 (22 R and 15 B) representative parents, evaluated for grain yield at four locations in India, showed significant differences in yield, heterosis, and combining ability. Across all the hybrids, mean mid- and better-parent heterosis for grain yield was 84.0% and 60.5%, respectively. Groups G12 B × G12 R and G10 B × G12 R had highest heterosis of about 10% over best check hybrid Pioneer 86M86. The parents involved in heterotic hybrids were mainly from the groups G4R, G10B, G12B, G12R, and G13B. Based on the heterotic performance and combining ability of groups, 2 B-line (HGB-1 and HGB-2) and 2 R-line (HGR-1 and HGR-2) heterotic groups were identified. Hybrids from HGB-1 × HGR-1 and HGB-2 × HGR-1 showed grain yield heterosis of 10.6 and 9.3%, respectively, over best hybrid check. Results indicated that parental groups can be formed first by molecular markers, which may not predict the best hybrid combination, but it can reveal a practical value of assigning existing and new hybrid pearl millet parental lines into heterotic groups to develop high-yielding hybrids from the different heterotic groups.

Introduction

Pearl millet (*Pennisetum glaucum* (L.) R. Br., syn. *Cenchrus americanus* (L.) Morrone, is cultivated in environments of low and erratic rainfall, high temperatures and low soil fertility and is the main source of food and fodder for the farming communities in arid and semiarid tropics of sub-Saharan Africa and South Asia. It occupies about 27 m ha area worldwide and is staple food for more than 90 million people globally. In South Asia, it occupies an area of about 8.0 m ha in India, with about 60 to 70% grown under hybrids (about 5.0 m ha) (Satyavathi 2017) and about 0.5 m ha (mostly under hybrids) in Pakistan (Ullah et al. 2017). Successful deployment of hybrids in India led to phenomenal increase in average productivity of pearl millet from 305 kg ha^{-1} in 1950s to present yield of 1132 kg ha^{-1} (Yadav and Rai 2013). To make this possible, breeding programs in public and private sector in India have worked closely with pearl millet breeding program of ICRISAT-Asia and are continuously engaged in enhancing genetic diversity of hybrid parents utilizing significant breeding material of African and Asian origin.

ICRISAT continues to be one of the major developers of advanced breeding lines and hybrid parents which have been widely disseminated worldwide since 1980s. The long-term goal of ICRISAT's hybrid pearl millet breeding program is to develop and disseminate broad-based germplasm, generate parents to develop hybrids with diverse phenotypic traits, adaptation to multiple environments and having high

Communicated by Alain Charcosset.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-019-03512-z) contains supplementary material, which is available to authorized users.

✉ S. K. Gupta
s.gupta@cgiar.org

Extended author information available on the last page of the article

Published online: 02 January 2020

yield, multiple resistance to diseases and tolerance to abiotic stresses. ICRISAT-bred hybrid pearl millet germplasm has been playing a major role in hybrid pearl millet programs in public and private sectors globally, and especially in India, where 60 to 70% hybrids cultivated at farmer's field derives one or both of their parents (directly or indirectly) from this program (Mula et al. 2007; ICRISAT 2012; Rao et al. 2018).

In order to continuously enhance the genetic gains achieved through pearl millet hybrids, there is need to evolve new strategies to increase magnitude of heterosis further. One of them is to increase heterozygosity through hybridization of genetically distinct parental materials belonging to distinct heterotic pools (Melchinger and Gumber 1998). To achieve this, an ideal approach would be to cross all potential germplasm sources in diallel crosses and select the most promising heterotic patterns, which is not practically feasible due to large number of materials and resources involved. Alternatively, available germplasm needs to be organized into heterotic groups to increase the efficiency of hybrid breeding program (Reif et al. 2005). Importance of formation of heterotic groups has been suggested in maize (*Zea mays*) (Menkir et al. 2004; Reif et al. 2003; Akinwale et al. 2014; Suwarno et al. 2014); rye (*Secale*) (Fischer et al. 2010); sunflower (*Helianthus annuus*) (Reif et al. 2013); sorghum (Menz et al. 2004) and in rice (Xie et al. 2013; Wang et al. 2015). Of the two recently concluded studies on formation of heterotic groups in pearl millet, one investigated combining ability patterns in West African population hybrids but could not come out with clear heterotic groups (Pucher et al. 2016); while the other assessed hybrid parents from one breeding program only (Ramya et al. 2018) and did not estimate heterosis in identified groups over currently popular high-yielding commercial hybrids.

To overcome the limitation of evaluating all possible crosses in available germplasm for identification of heterotic pools, Melchinger (1999) suggested (i) clustering germplasm based on genetic similarities, (ii) selecting representative genotypes from each subgroup, (iii) evaluating crosses among representative genotypes in field trials, and (iv) finally identifying heterotic pools based on per se performance, combining ability and heterosis. Hence, clustering of available germplasm, through precise phenotyping or genotyping, into groups is a prerequisite to formulate working heterotic pools in any crop. Interestingly, recent investigations assessing molecular genetic diversity in pearl millet classified breeding lines into genetically distinct groups (Kapila et al. 2008; Stich et al. 2010; Nepolean et al. 2012; Singh et al. 2013; Gupta et al. 2015; Singh et al. 2018; Ramya et al. 2018) and further indicated the existence of two broad-based pools in hybrid parents, each representing B- (seed parents) and R-lines (restorers parents). However, the precise information on heterotic pattern in two groups of material is lacking. This study was, therefore, conducted to identify the patterns of heterotic groups among hybrid parents available across ICRISAT and national (both public and private sector) pearl millet breeding programs of India, which has broader adaptation in South Asia. We also assessed whether there is any association between molecular diversity and magnitude of heterosis realized in terms of higher productivity. Such information can provide a guideline reference to maximize germplasm potential useful for increasing heterosis in pearl millet hybrid breeding.

Materials and methods

Genetic material

Five hundred and eighty hybrid parents (320 R-lines and 260 B-lines) were used in this study. They included 391 parental lines having diverse pedigrees from ICRISAT pearl millet breeding program (195 R-lines coded between R-1 and R-200; 196 B-lines, coded between B-1 and B-200); and 189 lines from 5 different breeding programs representing both public and private sectors (125 R-lines coded between R-201 and R-342; and 64 B-lines coded between B-201 and B-264) (Supplementary Table 1). These breeding programs were: Chaudhary Charan Singh Haryana Agricultural University (CCSHAU), Hisar, Haryana; Junagadh Agricultural University (JAU), Jamnagar, Gujarat; Mahatma Phule Krishi Vidyapeeth (MPKV), Dhule, Maharashtra; Sri Karan Narendra Agriculture University (SKNAU), Durgapura, Rajasthan, and one private seed company (name undisclosed). The four public funded research programs at Hisar, Jamnagar, Dhule and Durgapura represented states which collectively cultivate about 90% of pearl millet in India.

DNA extraction and genotyping

Thirty to thirty-five seeds from each parental line were sown in a dark room maintained at a temperature between 18 and 25 °C at ICRISAT-Patancheru. Etiolated leaf tissues were harvested eight days after sowing. Pooled leaf tissue from 20 to 25 seedlings per line was collected for DNA extraction using a modified DNA extraction method described by Mace et al. (2003). The DNA was stained by 5 ng/µl of ethidium bromide and checked using 0.8% (w/v) agarose gel electrophoresis in tris-acetate-EDTA (TAE) buffer for 1 h at 90 V with visualization under ultraviolet (UV) light.

The parental lines were resequenced using restriction-site-associated DNA (RAD) sequencing, followed by SNP calling and filtering as described by Varshney et al. (2017). Resequencing data of these pearl millet lines are available at https://www.ncbi.nlm.nih.gov//sra/?term=SRP063925. Also, the world reference germplasm line Tift 23 D_2B_1–P_1–P_5 was used as a control. This line is a single plant

selection done at ICRISAT from Tift 23 D_2B_1 line which was bred at the Coastal Plain Experiment Station, Tifton, Georgia, USA.

Selection of parents for development of hybrids

Diversity analysis was carried on RAD-GBS data of 580 hybrid parents based on Roger's distance (Rogers 1972) and following unweighted neighbor joining clustering method using DARwin software version 6.0.015 (Perrier et al. 2003). The grouping of B- and R-lines into clusters was done at 5% dissimilarity level. All the 580 hybrid parents were found clustered into 13 clear-cut groups (9 R-line groups) and (4 B-line groups). Three groups which had mix of both B- and R-lines were separated into 3B- and 3R-line groups (Table 1, Fig. 1). Hence, total of 7 B-line groups and 12 R-line groups were identified. Groups dominated by R-lines were designated as G1R to G9R; groups dominated by B-lines were designated as G13B to G16B, while mixed groups were designated as G10B, G11B and G12B for B-lines while G10R, G11R and G12R for R-lines. Representative lines for each of these groups were identified on the basis of total number of parents falling in that respective group and also considering genetic distance (GD) of the parent in the respective group. Based on the number of lines a group had, the groups were characterized as: large (> 10% of the total lines), medium (5–10% lines) and small (< 5% lines). As a rule, 3, 2 and 1 representative lines were identified from these large, medium and small groups, respectively (Table 1). Representative lines were identified which can represent the entire range of GD of the respective group. The average GD of the groups and the average GD of representative lines in relation to all other lines in the respective group are presented in Table 1. In the case of small groups, one line having average GD with respect to all other lines of the specific group was identified. In the case of medium groups, one line having higher and the other having lower GD with respect to all other lines than the

Table 1 Distribution of 580 pearl millet hybrid parents in different groups based on SNP genotyping, size of groups, and number and name of identified representative parents

Group name	No. of R-lines	No. of B-lines	Total no. of parents	Group category	No. of parents selected	Average genetic distance of group	Representative hybrid parents (average genetic distance from other lines in respective group)[a]
G1R	47	1	48	L	3	0.31	R83 (0.28), R110 (0.31), R171 (0.34)
G2R	23	0	23	M	2	0.33	R9 (0.31), R11 (0.26)
G3R	21	6	27	M	2	0.34	R203 (0.30), R243 (0.38)
G4R	16	1	17	S	1	0.32	R2 (0.32)
G5R	12	1	13	S	1	0.34	R177 (0.34)
G6R	36	3	39	L	3	0.27	R22 (0.22), R70 (0.27), R183 (0.30)
G7R	26	2	28	M	2	0.27	R75 (0.28), R151 (0.23)
G8R	19	4	23	S	1	0.36	R3 (0.35)
G9R	33	3	36	L	3	0.36	R69 (0.40), R80 (0.34), R157 (0.37)
G10R	27	0	27	M	2	0.34	R21 (0.38), R115 (0.29)
G11R	17	0	17	S	1	0.32	R187 (0.32)
G12R	23	0	23	M	1	0.32	R167 (0.32)
G10B	0	32	32	L	2	0.28	B130 (0.22), B159 (0.38)
G11B	0	12	12	S	1	0.25	B234 (0.25)
G12B	0	15	15	S	1	0.34	B110 (0.35)
G13B	2	32	34	L	3	0.24	B37 (0.18), B169 (0.31), B191 (0.25)
G14B	10	50	60	L	3	0.29	B86 (0.29), B96 (0.34), B194 (0.24)
G15B	2	80	82	L	3	0.23	B24 (0.18), B65 (0.22), B108 (0.31)
G16B	6	18	24	M	2	0.24	B3 (0.31), B64 (0.20)
Total	320	260	580	7 L, 6 M, 6 S	22 R and 15 B-lines	0.28	37 lines

[a]Details of parental lines are available in Supplementary Table 1

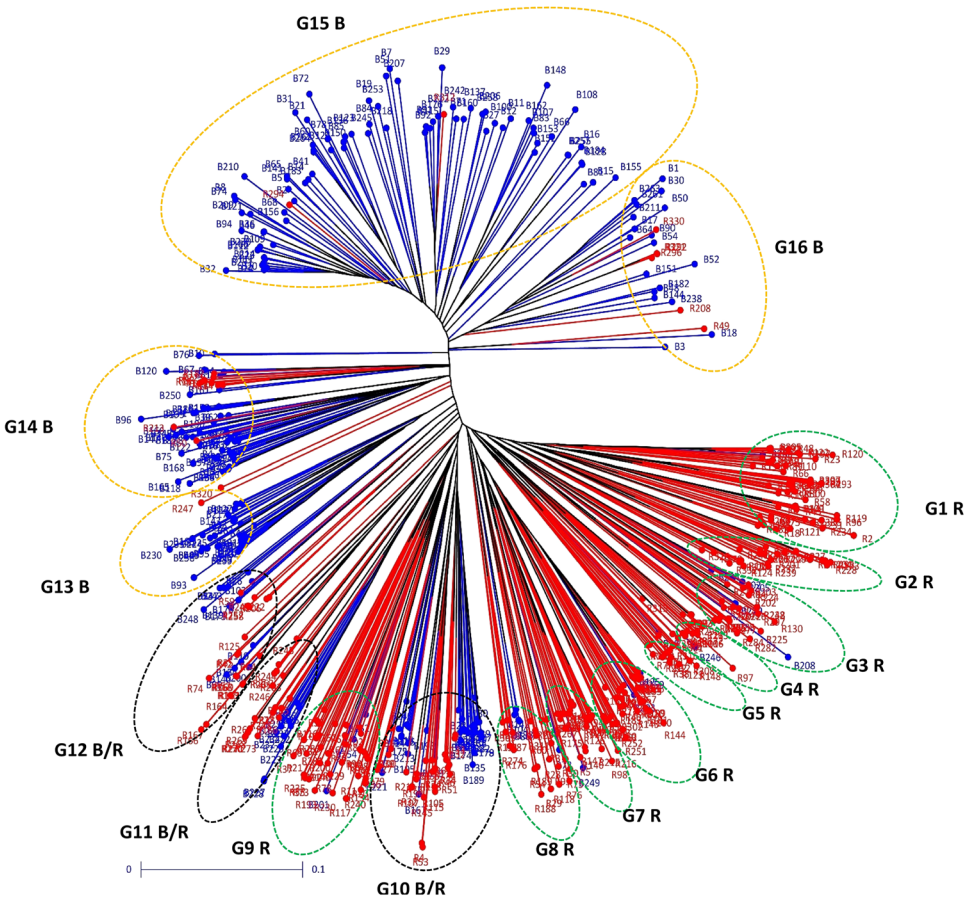

Fig. 1 Nineteen marker-based groups in the clustering pattern of 580 B- and R-lines of pearl millet. Red and blue colors indicate restorer and seed parents, respectively (high resolution of this figure is also available as Supplementary Figure S1) (color figure online)

average GD of the specific group were identified. In the case of large group, one line having GD equivalent to the average GD with respect to all other lines of the specific group, and two other with one having higher and the other having lower GD with respect to all other lines than the average GD of the specific group were identified. The genetic structure of the representative lines consisted of 13 groups, each having 1 to 3 representative lines (Fig. 1). The set of 37 lines represented the allelic variation presented in the entire set of breeding lines (Fig. 3). The GD values for the 37 parents ranged from 0.136 to 0.349 with an average of 0.285, which was 0.286 in the original set of 580 parental lines. In total, 37 (15 B- and 22 R-lines) representative lines were selected across these groups and crossed in line × tester mating design to generate 320 test cross hybrids (10 hybrids could not be generated due to flowering synchronization problems in parental lines) during summer season of 2015 at ICRISAT. These 320 hybrids represented 84 marker-based group crosses (7 B-line groups × 12 R-line groups).

Field evaluation of hybrids and parents

A trial comprising 320 hybrids, 37 parents and 4 hybrid check entries was planted in alpha lattice design with two replications during rainy season (June–September) of 2015 at four locations in India: (Dhule: N 20.90°, E 74.77°, Patancheru: N 17.53°, E 78.27°, Jamnagar: N 22.47°, E 70.06° and Hisar: N 29.15°, E 75.72°). The hybrid checks included Pioneer 86M86, Pioneer 86M88, ProAgro 9444 and HHB-67-Improved. The two Pioneer hybrids 86M86 and 86M88 are currently cultivated on large areas in India and also in some parts of Pakistan (Pioneer, Pakistan 2018). ProAgro 9444 is one of the most stable hybrids grown over large area in India since last about two decades.

Hybrids and parents were blocked separately within the replication to avoid suppressive effect of hybrids over parents. Hybrid plots were randomly assigned to the first 25 blocks of each replicate, and parental plots were randomly assigned to the last three blocks of each replicate. Each entry was planted in 2 rows of 4 m length, spaced at 75 cm between rows and 15 cm between plants.

Standard agronomic management practices were followed at all the locations for good crop growth. Basal dose of 100 kg of DAP (diammonium phosphate, containing 18% N, 46% P) was applied at the time of field preparation, and 100 kg of urea (46% N) was applied as top dressing to meet the recommended dose of 64 kg of N ha^{-1} and 46 kg of P ha^{-1}; irrigations were given soon after sowing and then

subsequently during the season as and when required. Seedlings were thinned at 15 days after sowing to maintain seedlings at a uniform spacing of 15 cm. The other cultural practices like weeding, protection against insects, pests, diseases and birds were done throughout the growing period as and when required. All the panicles in a plot were harvested for each entry. The harvested material was sun-dried for 10 to 15 days, threshed and recorded for grain yield in kilogram and converted to grain yield per hectare.

The present study investigated B × R hybrids, though A × R hybrids have to be developed for cultivation. Since A-lines of CMS system are sterile but contain exact same nuclear genotype as that of B-lines, we measured traits of B-lines parents and genotyped them in place of A-lines to obtain consistent parent phenotypic values. The phenotype derived from hybrids of B-lines and R-lines can be used to represent the crosses between A- and R-lines.

Statistical analysis

The procedure of the line × tester analysis according to Kempthorne (1957) was used for estimating general and specific combining ability effects and variances as described by Singh and Chaudhary (1977). Combined analysis of variance was carried out using PROC MIXED (SAS v9.4, SAS Institute Inc. 2017), considering environment, treatments and replication as fixed. In order to pool the data across four environments, individual environment variances were modeled to error distribution using residual maximum likelihood (REML) procedure. The phenotypic observations Z_{ijk} on accession k in replicate j of environment i were modeled as:

$$Z_{ijk} = \mu + e_i + (e/r)_{ij} + t_k + (et)_{ik} + \varepsilon_{ijk}$$

where μ is the grand mean; e_i is the fixed effect of environment i; t_k is the fixed effect of treatment k; $(e/r)_{ij}$ is the fixed effect of replication j nested with in environment i; $(et)_{ik}$ is the fixed effect of the interaction between treatment k in environment i; ε_{ijk} is the random residual effect and ~ NID(0, σ^2_ε). Least square means were estimated for all effects using combined analysis. Mean square effects of line, tester and their interactions from combined analysis of variance across environments for hybrids (Table 2) were used in

Table 2 Analysis of variance of hybrids and their parents for grain yield across 4 environments in pearl millet

Source	Degrees of freedom	Sum of squares	Mean sum of squares	*F* value	Pr > *F*
Environment (E)	3	1.502E+09	500,707,954	1421.4	< 0.0001
Replication (environment)	4	337,824,376	84,456,094	239.8	< 0.0001
Treatments	363	1.489E+09	4,100,719	11.6	< 0.0001
P versus H versus C	2	431,380,418	215,690,209	612.3	< 0.0001
Parents (P)	36	61,467,228	1,707,423	4.9	< 0.0001
Parents–line	14	20,238,512	1,445,608	4.1	< 0.0001
Parents–tester	21	27,553,239	1,312,059	3.7	< 0.0001
Parents-(line versus tester)	1	13,675,485	13,675,485	38.8	< 0.0001
Hybrids (H)	319	794,305,215	2,489,985	7.1	< 0.0001
Hybrids–line	14	107,006,074	7,643,291	5.4	< 0.0001
Hybrids–tester	21	242,446,365	11,545,065	8.1	< 0.0001
Hybrids–line × tester	284	404,965,824	1,425,936	4.1	< 0.0001
Checks (C)	6	201,408,288	33,568,048	95.3	< 0.0001
Environment × treatment	1013	1.119E+09	1,105,024	3.1	< 0.0001
Environment × (P versus H versus C)	6	55,972,003	9,328,667.2	26.5	< 0.0001
Environment × parents	107	55,796,327	521,461	1.5	0.001
Parent-environment × line	42	19,946,598	474,919	1.4	0.069
Parent-environment × tester	62	30,878,542	498,041	1.4	0.02
Environment × hybrid	882	885,572,982	1,004,051	2.9	< 0.0001
Environment × line	42	128,550,576	3,060,728	8.7	< 0.0001
Environment × tester	63	118,768,356	1,885,212	5.4	< 0.0001
Environment × (line × tester)	777	616,717,332	793,716	2.3	< 0.0001
Environment × checks	18	122,047,218	6,780,401	19.3	< 0.0001
Error	1647	580,185,396	352,268	–	–
Corrected total	3030	5.88E+09	–	–	–

the following formulas to estimate variances of general and specific combining ability (Table 3).

$$\sigma^2\mathrm{GCA} = \frac{\mathrm{MS}_{\mathrm{line}} + \mathrm{MS}_{\mathrm{tester}} - 2\mathrm{MS}_{\mathrm{line\times tester}} - \mathrm{MS}_{\mathrm{line}\times E} - \mathrm{MS}_{\mathrm{tester}\times E} + 2\mathrm{MS}_{\mathrm{line\times tester}\times E}}{rE(L+T)}$$

$$\sigma^2\mathrm{GCA} * E = \frac{\mathrm{MS}_{\mathrm{Line}\times E} + \mathrm{MS}_{\mathrm{tester}\times E} - 2\mathrm{MS}_{\mathrm{line\times tester}\times E}}{r(L+T)}$$

$$\sigma^2\mathrm{GCA}_{\mathrm{line}} = \frac{\mathrm{MS}_{\mathrm{line}} - \mathrm{MS}_{\mathrm{line}\times E} - \mathrm{MS}_{\mathrm{line\times tester}} + \mathrm{MS}_{\mathrm{line\times tester}\times E}}{rLE}$$

$$\sigma^2\mathrm{GCA}_{\mathrm{tester}} = \frac{\mathrm{MS}_{\mathrm{tester}} - \mathrm{MS}_{\mathrm{tester}\times E} - \mathrm{MS}_{\mathrm{line\times tester}} + \mathrm{MS}_{\mathrm{line\times tester}\times E}}{rTE}$$

$$\sigma^2\mathrm{GCA}_{\mathrm{line}*E} = \frac{\mathrm{MS}_{\mathrm{line}\times E} - \mathrm{MS}_{\mathrm{line\times tester}\times E}}{rT}$$

$$\sigma^2\mathrm{GCA}_{\mathrm{tester}*E} = \frac{\mathrm{MS}_{\mathrm{tester}\times E} - \mathrm{MS}_{\mathrm{line\times tester}\times E}}{rL}$$

$$\sigma^2\mathrm{SCA} = \frac{\mathrm{MS}_{\mathrm{line\times tester}} - \mathrm{MS}_{\mathrm{line\times tester}\times E}}{rE}$$

$$\sigma^2\mathrm{SCA} * E = \frac{\mathrm{MS}_{\mathrm{line\times tester}\times E} - \mathrm{MS}_{\mathrm{Error}}}{r}$$

where r number of replications, E number of environments, L number of lines and T number of testers.

GD, SCA, MPH and BPH between any two of the 84 groups were also estimated based on the means of GD, SCA, MPH and BPH values, respectively, of all the probable combinations between the representative parents in those two groups. Heterosis for grain yield was estimated as (i) mid-parent heterosis (MPH) = $100 \times (F_1 - \mathrm{MP})/\mathrm{MP}$; (ii) better-parent heterosis (BPH) = $100 \times (F_1 - \mathrm{BP})/\mathrm{BP}$; and hybrid yield advantage over commercial check = $100 \times (F_1 - \mathrm{SC})/\mathrm{SC}$; where F_1 is hybrid yield, MP is yield mean of both parents, BP is the yield of better-yielding parent, and SC is the yield of best standard check. Pioneer 86M86, the highest yielder in the trial, was considered as standard check (SC), while ProAgro 9444 (hybrid with maximum market size in India over last 20 year period and grown across pearl millet zones and seasons) was considered as alternative check for estimation of heterosis. Pearson correlation coefficients between GD and hybrid performance, mid- and better-parent grain yield, and SCA were calculated using SAS PROC CORR procedure (SAS v9.4, SAS Institute Inc. 2017). Grain yield was estimated for all the marker groups, considering test crosses of all the representative lines of each B-line group when crossed to representative lines of all the R-groups, and vice versa. *t* test was conducted to test the significance of differences in marker groups.

Table 3 Estimates of combining ability variance for grain yield across 4 environments in pearl millet

Variance components	Estimates
$\sigma^2\mathrm{GCA}$	43,844.51
$\sigma^2\mathrm{GCA}_{\mathrm{line}}$	22,445.13
$\sigma^2\mathrm{GCA}_{\mathrm{tester}}$	75,230.28
$\sigma^2\mathrm{SCA}$	79,027.50
$\sigma^2\mathrm{GCA} * E$	45,385.24
$\sigma^2\mathrm{GCA}_{\mathrm{line}*E}$	51,523.00
$\sigma^2\mathrm{GCA}_{\mathrm{tester}*E}$	36,383.21
$\sigma^2\mathrm{SCA} * E$	257,000.00
$\sigma^2\mathrm{SCA}/\sigma^2\mathrm{GCA}$	1.80

Results

Molecular diversity

Genetic data summary calculated at each locus using R software version 3.4.1 (R Core Team 2017) using 580 hybrid parents and 8,95,791 SNP markers after a cutoff on the minor allele frequency at 5% and taxa coverage at 30% showed that the polymorphic information content (PIC) value for all 580 hybrid parents ranged from 0.091 to 0.774 with an average value of 0.238. Furthermore, the PIC value ranged from 0.000 to 0.749 (average of 0.209) and 0.000 to 0.750 (average of 0.244) for B- and R-lines, respectively. Gene diversity for all hybrid parents varied from 0.095 to 0.702, averaging 0.288. Among B- and R-lines, average gene diversity was 0.252 (ranged from 0.000 to 0.674) and 0.297 (ranged from 0.000 to 0.715), respectively. The average major allele frequency for all the hybrid parents ranged from 0.321 to 0.950 with a mean of 0.798.

Grouping of B- and R-lines and GD

The clustering pattern depicted through dendrogram (Fig. 1 and Supplementary Figure S1) clearly delineated most of B- and R-lines into separate groups. Seventy-three percent (233 lines) and seventy percent (180 lines) of B- and R-lines formed clear-cut separate groups; R-lines making further nine (G1R to G9R) subgroups and B-line making four (G13B to G16B) subgroups. In R-line groups, G1R had the highest number (46) of R-lines, while the smallest group (G5R) had 12 R-lines. In B-line groups, G15B had the highest number of 82 B-lines, while G16B had minimum (18 B) lines. Twenty percent (67) of R-lines and 23% (59) of B-lines were found in common groups (G10 to G12), mentioned as G10 B/R, G11 B/R, and 12B/R. Hence, total of 84 marker group crosses (7 B-groups × 12 R-groups) were investigated in this study.

Many of the parental lines falling in a particular group shared common pedigrees. For example 11, 12, 7, 13, 8, 6, and 6 lines having common parentage involving Mandore Restorer Composite (MRC), Medium Composite (MC 94), ICMR-312, Smut Resistant Composite (SRC), Bold Seeded Early Composite (BSEC), ICMB 03111, and Downy Mildew Resistant breeding line (DMR), respectively, were found in G1R, G6R, G7R, G11B/R, G13B, and G14B, respectively.

Hybrid parents from different breeding programs (both public and private) were found distributed across groups, the only exception was Dhule where most of the B- and R-lines were found grouped in single group G11B/R only (Fig. 2). Many lines from a particular breeding program were found in same group: 12 out of 35 R-lines from Hisar were in G1R; 6 out of 22 and 7 out of 19 R-lines from Jamnagar and Durgapura in G2R; 4 out of 18 B- and 5 out of 22 R- lines from Jamnagar in G3R; 5 out of 35 R-lines from Hisar in G4R; 9 out of 30 R-lines from Dhule in G6R; 12 out of 16 B- and 14 out of 30 R-lines from Dhule in G11B/R; and 5 out of 12 B-lines from Hisar in G15B.

Genetic distances between the 84 marker group crosses varied from 0.20 (G11B × G11R) to 0.33 (G15B × G12R and G11B × G2R). Genetic distance between all 580 parents varied from 0.016 to 0.375 with maximum number of B × R combinations (59.4%) in range of 0.28 to 0.32, followed by 23.5% in the range of 0.24 to 0.28 (Fig. 3). Among the selected representative 37 parental lines, GD varied from 0.136 to 0.349, with B-159 × R-22 having minimum GD of 0.136 and both B-96 × R-9 and B-234 × R-9 having maximum GD of 0.349.

Combined Analysis of Variance for Combining Ability

Analysis of variance for grain yield based on line × tester analysis (Table 2) showed large and highly significant variation due to locations, indicating that the materials were

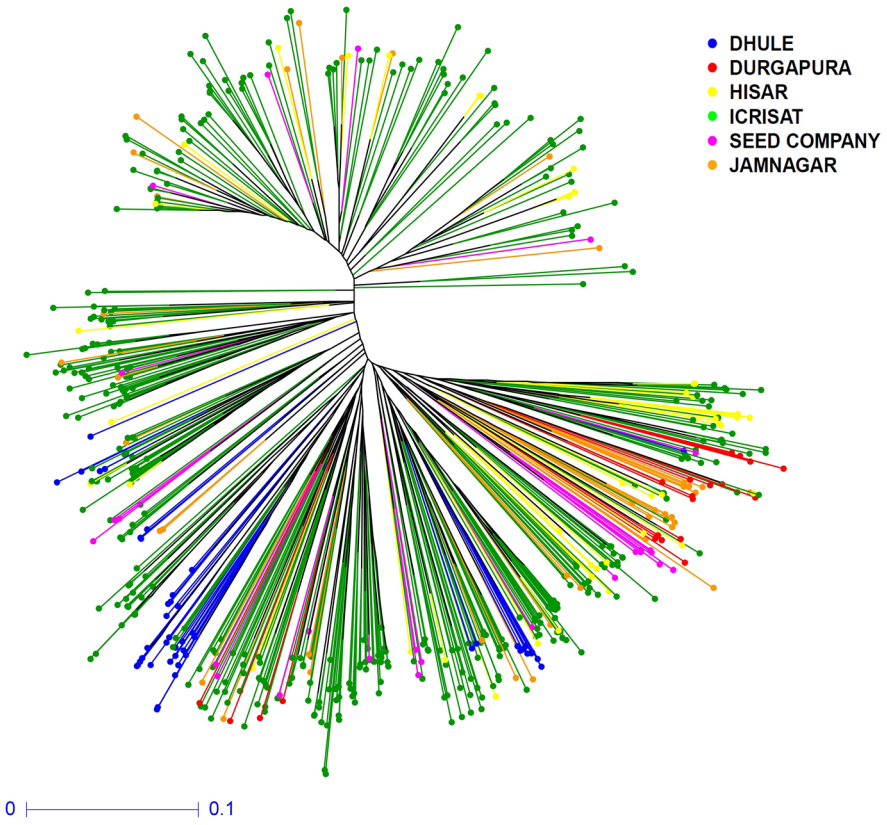

Fig. 2 Distribution of hybrid parental lines from six breeding programs in the clustering pattern of 580 hybrid parental lines of pearl millet (color figure online)

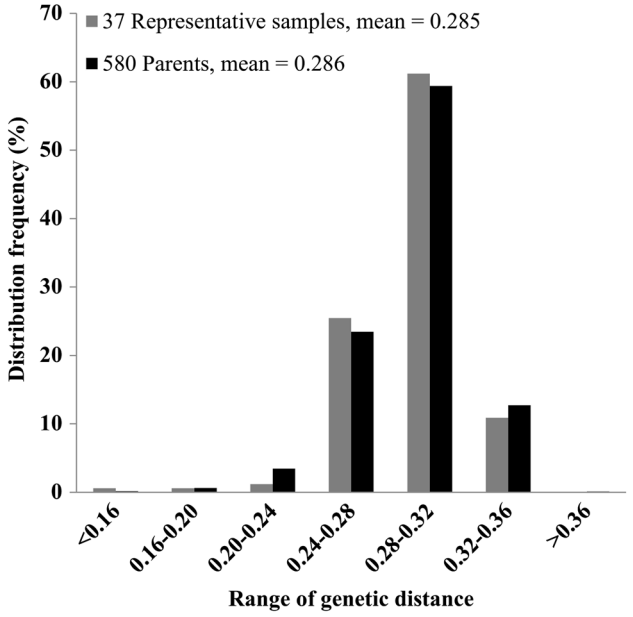

Fig. 3 Distribution of genetic distance of 580 pearl millet hybrid parents and the 37 parents (15 B- and 22 R-lines) identified as marker group representative lines

evaluated under diverse environments. Large and highly significant variation observed due to parents and due to hybrids indicated wide genetic differences among parental lines as well as among hybrids. Yield performance of both parental lines and hybrids was significantly modified by environments. The significant and relatively large percentage of the total variation attributable to the G × E suggests that hybrids responded differentially to various evaluation environments. Estimates of variances of GCA, SCA, and their interactions with environment based on line × tester analysis are provided in Table 3. SCA was found 1.8 times higher than the GCA variance.

Per se Performance of Parents, Hybrid performance, General Combining Ability (GCA), and Specific Combining Ability (SCA) Effects

The details on per se performance of parents, hybrid performance, general combining ability, and specific combining ability effects for grain yield based on pooled data of four environments are presented in Table 4. The mean values of grain yield among the B-lines varied from 932 (B110) to 2806 kg ha^{-1} (B159) with an average of 1458 kg ha^{-1}, whereas among the R-lines the range was from 1122 (R187) to 2680 kg ha^{-1} (R177) with mean of 1940 kg ha^{-1}. The GCA effects for grain yield per hectare varied from − 548.0 ($P < 0.01$) (B194) to 493.3 ($P < 0.01$) (B110) among B-lines and from − 644.6 ($P < 0.01$) (R80) to 1058.2 ($P < 0.01$) (R167) among R-lines. Out of 15 B-lines and 22 R-lines, four B- and nine R-lines exhibited positive and significant GCA effects, whereas five B- and nine R-lines showed negative and significant GCA effects (Table 4). Among hybrids, SCA effects varied from − 1704.8 ($P < 0.01$) (B110 × R177) to 2705.2 ($P < 0.01$) (B169 × R75). Fifty-eight and sixty-six hybrids possessed significant negative and positive SCA effects, respectively. Of sixty-six hybrids (mean range from 2470 to 4700 kg ha^{-1}) with specific combining ability in desirable direction, the cross combination, B-169 × R-75 (2705.2, $P < 0.01$) with H^+ gca × H^- gca parental combination, showed the highest significant SCA effect followed by B-86 × R-187 (2295.6, $P < 0.01$) (H^- gca × H^- gca) and B-37 × R-110 (2215.7, $P < 0.01$) (H^+ gca × H^+ gca) (Table 4) The mean grain yield of parents was also found positively correlated with GCA for grain yield ($r = 0.42$, $P < 0.001$).

In 84 marker-based hybrid groups, average grain yield varied from 2198 kg ha^{-1} (G12B × G8R) to 4507 kg ha^{-1} (G12B × G12R) with a mean of 3222 kg ha^{-1} (Table 5; only 30 highest yielding groups are presented). Based on performance of group crosses between B- and R-lines for grain yield, group cross G12B × G12R (4507 kg ha^{-1}) had the highest grain yield followed by G10B × G12R (4456 kg ha^{-1}), G13B × G12R (4434 kg ha^{-1}) and G11B × G12R (4255 kg ha^{-1}).

Magnitude of heterosis and its association with GD

The estimates of mid- and better-parent heterosis for grain yield are presented in Table 5. Across 84 group crosses, the range of mid-parent heterosis was from 25.5% (G12B × G5R) to 215.0% (G12B × G4R), while it was − 15.6 (G12B × G5R) to 153.6% (G14B × G11R) for better-parent heterosis. Across all the 320 hybrids, mean mid- and better-parent heterosis for grain yield was 89.0% and 60.5%, respectively. Maximum number of hybrids had mid-parent heterosis in the range of 76–100%, while better-parent heterosis was in the range of 51–75% (Fig. 4). Maximum number of parental lines had grain yield in the range of 1501–2000 kg ha^{-1}, while maximum number of hybrids was found in the range of 3001 to 5000 kg ha^{-1} (Fig. 5).

The estimates of heterosis in 320 F_1s for yield over commercial checks were estimated over the best check hybrid Pioneer 86M86 (4077 kg ha^{-1}). Hybrid ProAgro 9444, known for its wide adaptation and having highest cultivated area in India, had grain yield of 3490 kg ha^{-1}. Hybrid yield advantage over commercial check (over Pioneer 86M86) varied from − 71.5% (B-37 × R-11) to 15.3% (B-130 × R-167) at individual hybrid level, while at group level G12B × G12R had highest advantage of 10.57% followed by G10B × G12R (9.31%) and G13B × G12R (8.77%). The range of heterosis over ProAgro hybrid 9444 ranged from 34.67 (B 130 × R167) to − 66.7 (B 37 × R 11), while group cross G12B × G12R had highest heterosis of 29.1%

Table 4 GCA and SCA of pearl millet parental lines and hybrids along with grain yield (kg ha^{-1}) of representative hybrid parental lines involved in this study

	R11	R110	R115	R151	R157	R167	R171	R177	R183	R187	R2
B108	275.9	433.5*	−275.1	64.5	725.6**	951.1**	−606.7**	−288.3	116.5	−787.0**	960.7**
B110	122.2	298.3	−298.6	−1045.6**	601.5**	−198.7	467.4*	−1704.8**	1732.1**	.	395.4*
B130	357.9	129.2	287.5	41.2	236.4	616.9**	41.5	−205.8	−663.3**	−297.1	−628.4**
B159	−179.1	−836.9**	401.6*	−788.5**	−803.1**	−398.2*	180.6	358.8	582.6**	−116.9	−736.6**
B169	−490.9**	−192.9	−324.2	−119.7	−373.4*	.	472.9*	−90.3	93.9	40.4	−178.3
B191	1122.7**	63.2	658.3**	−417.0*		−144.0	−288.9	2151.8**	399.1*		−1557.6**
B194	−83.5	130.2	206.4	−9.2	186.6	489.0**	44.5	−54.2	−380.4*	332.4	220.1
B234	30.0	−593.1**	−1147.0**	833.1**	−1248.9**	65.3	−283.9	342.6	307.2	72.9	77.4
B24	40.6	−61.3	−340.0	254.9	129.6	−233.2	380.3*	104.3	−243.7	545.9**	−251.4
B3	−95.8	−309.5	139.1	222.4	67.0	90.5	−535.7**	312.9	92.4	26.8	463.5*
B37	−1586.6**	2215.7**	−544.0**	2175.2**		765.5**	342.6	618.4**	−551.5**	130.6	519.6**
B64	118.5	−158.1	434.2*	25.9	456.7*	−1585.9**	331.9	11.4	232.5	−283.0	−129.8
B65	−368.7*	−428.3*	−194.8	57.1	−152.6	−82.5	−23.2	124.9	76.8	−148.4	−88.3
B86	−127.7	223.2	687.7**	26.8	282.9	−284.8	−425.8*	−181.9	472.9*	2295.6**	782.2**
B96	318.2	664.5**	438.6*	206.6	145.6	−149.6	−115.2	151.7	−393.6*	155.7	30.6
GCA	−398.0**	116.4*	23.3	−139.7**	−154.1**	1058.2**	255.1**	265.4**	−131.2**	−300.5**	338.4**
GY (R-lines)	1362.1	2242.2	2513.7	1776.7	2288.4	2638.0	2562.7	2679.3	1682.7	1121.9	1749.9

	R203	R21	R22	R243	R3	R69	R70	R75	R80	R83	R9	GCA	*GY (B-lines)*
B108	40.9	194.1	1596.2**	−29.8	37.7	400.2*	73.1	−41.1	−601.8**		−25.8	−77.8	1155.2
B110	494.9**	147.4	−193.6	1110.7**	−933.3**	−154.7	47.4	−511.5**	−337.1	−297.2	313.3	493.3**	931.8
B130	−226.5	−461.6*	413.5*	88.7	−155.6	361.2	−443.6*	329.3	−381.1*	−240.5	364.2	−87.8*	1323.9
B159	705.0**	262.3	187.2	−165.4	66.5	−407.3*	919.4**	421.1*	−59.6	599.3**	−13.3	416.6**	2806.1
B169	132.6	841.5**	1126.0**	−45.1	851.6**	−629.1**	−737.7**	2705.2**	.	402.0*	121.2	196.5**	1786.4
B191	−293.8	−326.9	1054.1**		−1438.6**	1039.3**	40.8	−722.3**	285.6	1879.8**	16.0	59.6	1403.2
B194	186.6	−413.9*	33.8	90.4	154.5	51.8	377.3*	−559.7**	−128.0	−624.9**	−190.6	−548.0**	955.2
B234	−319.0	1090.0**	−599.5**	116.3	724.7**	348.2	215.6	484.7**	−125.6	−289.5	1153.8**	−15.7	1626.6
B24	−45.0	−212.6	253.6	−523.8**	−120.6	−312.9	225.4	51.2	171.4	254.5	−377.2*	−212.4**	1215.4
B3	399.8*	−616.3**	392.1*	−186.6	54.9	−156.6	−28.9	−212.8	−162.9	354.5	154.8	157.2**	1369.4
B37	−177.3	−640.4**	−626.2**	−41.5	−261.6	−125.4	35.9	777.1**	−637.5**	−572.7**	−802.3**	−130.4**	1502.2
B64	−201.6	−279.2	−551.6**	−177.5	929.5**	−347.5	−29.5	498.7**	724.3**	227.1	−290.1	12.3	1674.2
B65	−155.1	975.4**	−312.8	−11.7	721.2**	60.7	−325.2	−268.6	1066.8**	−19.6	110.8	−1.2	1834.9
B86		191.5	−204.8	−368.6*	112.8	.	−352.3	−99.8	−402.3*	22.3	−431.5*	−214.1**	940.4
B96	8.0	−784.2**	−320.5	−109.5	−412.3*	−89.4	314.9	−430.7*	54.3	−346.4	1209.9**	−4.5	1345.0
GCA	−62.4	−511.5**	395.0**	25.6	−578.7**	168.4**	185.2**	−154.1**	−644.6**	121.5*	24.4		
GY (R-lines)	2013.8	1464.4	2073.6	1853.6	1591.0	1979.6	1713.7	1659.9	1670.5	2392.3	1670.8		

*, **Significant at 0.05, 0.01 levels of probability, respectively

Table 5 Genetic distance and hybrid yield across four environments, specific combining ability (SCA) and yield heterosis (MPH, BPH, and hybrid yield advantage over commercial checks) for 30 highest yielding marker group crosses in pearl millet

Sl. no.	Cross combination of marker groups	Genetic distance	Hybrid yield (kg ha^{-1})	SCA	MPH (%)	BPH (%)	Yield advantage over ProAgro 9444 (%)	Yield advantage over Pioneer 86M86 (%)
1	G12B × G12R	0.25	4507.4	− 198.7	152.5	70.9	29.2	10.6
2	G10B × G12R	0.31	4456.1	109.3	96.0	64.1	27.7	9.3
3	G13B × G12R	0.31	4434.3	310.7	116.7	68.1	27.1	8.8
4	G11B × G12R	0.31	4255.3	65.3	99.6	61.3	21.9	4.4
5	G12B × G4R	0.28	4234.0	395.4	215.8	141.9	21.3	3.9
6	G15B × G12R	0.33	4112.5	211.8	104.8	55.9	17.8	0.9
7	G12B × G1R	0.29	4031.7	156.2	142.1	68.2	15.5	− 1.1
8	G12B × G3R	0.27	3857.2	802.8	171.0	101.6	10.5	− 5.4
9	G16B × G4R	0.29	3779.8	166.9	132.7	116.0	8.3	− 7.3
10	G12B × G6R	0.28	3761.7	528.6	175.2	109.3	7.8	− 7.7
11	G12B × G2R	0.30	3732.6	217.8	204.7	146.8	6.9	− 8.4
12	G13B × G5R	0.28	3729.5	893.3	76.3	39.2	6.9	− 8.5
13	G11B × G7R	0.27	3725.1	658.9	122.6	116.6	6.7	− 8.6
14	G11B × G5R	0.31	3711.0	342.6	72.4	38.5	6.3	− 9.0
15	G10B × G5R	0.27	3687.6	76.5	56.8	34.2	5.7	− 9.5
16	G14B × G12R	0.31	3628.8	18.2	94.9	37.6	4.0	− 11.0
17	G14B × G4R	0.28	3576.5	344.3	153.6	104.4	2.5	− 12.3
18	G10B × G6R	0.22	3576.4	166.0	84.8	56.1	2.5	− 12.3
19	G11B × G4R	0.32	3553.1	77.4	110.5	103.1	1.8	− 12.8
20	G10B × G1R	0.29	3504.1	− 21.1	59.8	35.0	0.4	− 14.0
21	G12B × G9R	0.28	3498.6	36.6	138.9	76.1	0.2	− 14.2
22	G13B × G1R	0.29	3482.6	480.2	76.0	45.6	− 0.2	− 14.6
23	G16B × G12R	0.32	3462.4	− 747.7	68.1	31.3	− 0.8	− 15.1
24	G15B × G4R	0.29	3430.8	207.0	120.2	93.1	− 1.7	− 15.8
25	G10B × G3R	0.28	3419.6	100.5	74.3	47.2	− 2.0	− 16.1
26	G16B × G1R	0.28	3414.1	− 14.9	74.4	42.5	− 2.2	− 16.3
27	G16B × G6R	0.28	3400.3	17.8	104.5	87.6	− 2.6	− 16.6
28	G16B × G5R	0.27	3377.6	162.2	61.2	26.1	− 3.2	− 17.2
29	G15B × G5R	0.27	3373.8	− 19.7	65.9	25.9	− 3.3	− 17.2
30	G12B × G10R	0.26	3361.6	− 75.6	136.8	80.5	− 3.7	− 17.5
Mean of all 84 group crosses		0.29	3221.9	35.0	93.8	63.6	− 7.7	− 21.0

followed by G10B × G12R (27.68%) and G12B × G12R (27.0%).

There was no significant correlation ($r = 0.054$) between grain yield of hybrids and GD between parental lines (Table 6). Significant positive correlation was found between grain yield of hybrids and mean grain yield of parental lines ($r = 0.42$, $P < 0.001$) and with grain yield of better-parent ($r = 0.44$, $P < 0.001$).

Heterotic groups

B-lines of G12B when crossed to 22 representative lines of all 12 R-line groups gave highest mean grain yield of 3448 kg ha^{-1}, followed by lines of G10B (3330 kg ha^{-1}), G16B (3237 kg ha^{-1}) (Table 7), while B-line groups G13B, G15B, and G14B produced hybrids with lowest grain yields. Similarly, R-lines of G12R when crossed to all the 15 representative lines of 7 B-line groups showed highest mean grain yield of 4122 kg ha^{-1}, followed by lines of G4R (3559 kg ha^{-1}), G1R (3382 kg ha^{-1}) and so on, while groups G10R, G11R and G8 produced hybrids with lowest grain yields. Based on the t test and mean performance for grain yield of the marker groups, G12B when crossed with all other R-line groups had highest yield performance of hybrids, mid-parent and better-parent heterosis followed by G10B. Whereas among R-line marker groups, G12R had high GCA when crossed with most of the B-line groups (G12B, G10B, and G11B), followed by G4R (Table 4). B-line (B-110) in G12B and R-167 (G12R) had highest GCA among all the breeding

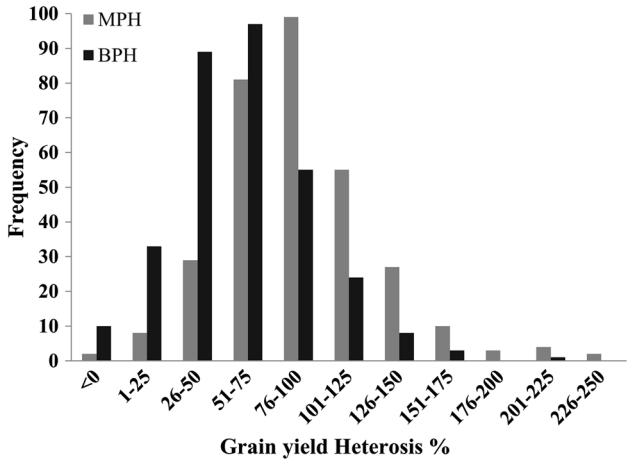

Fig. 4 Distribution of mid-parent and better-parent heterosis (in %) for grain yield of all the 320 pearl millet hybrids involved in the study

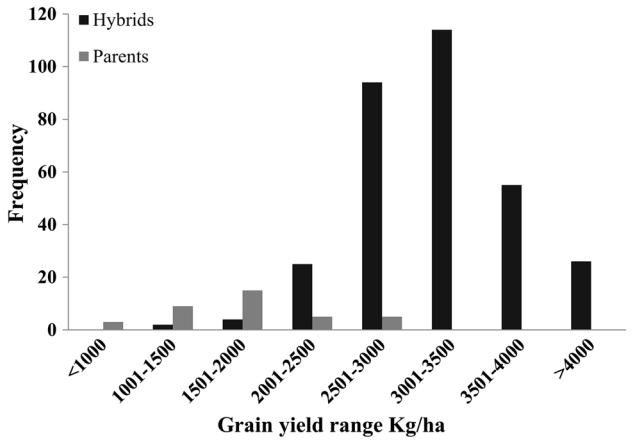

Fig. 5 Distribution of grain yield (kg ha^{-1}) of representative 37 parents and their 320 hybrids in pearl millet

Table 6 Correlations between genetic distance (GD), hybrid grain yield (HGY), SCA, and mid- and better-parent yield, among all the 320 hybrids and among 84 marker-based groups in pearl millet

Sl. no.	Trait	All hybrids ($n=320$)		Marker-based groups ($n=84$)	
		GD	HGY	GD	HGY
1	HGY	0.054	–	0.135	–
2	SCA	0.004	0.610**	−0.037	0.439**
3	MPGY	−0.039	0.420***	–	–
4	BPGY	−0.028	0.440***	–	–

HGY hybrid grain yield, *MPGY* mid-parent grain yield, *BPGY* better-parent grain yield

** and ***Significant at 0.01 and < 0.001 levels of probability

lines, while B-159 from G10B and R-2 (G4R) had highly significant GCA values.

Discussion

Connections between historical strategy in the development of hybrid parents in pearl millet and observed diversity pattern

Majority of hybrid parents derived from six breeding programs delineated into clear-cut separate B- and R-line groups. Some previous studies have also reported similar clustering pattern of B- and R-lines into two separate mega-groups based on genotyping or phenotypic evaluations (Nepolean et al. 2012; Gupta et al. 2015; Singh et al. 2018; Ramya et al. 2018). The existence of B- and R-lines as separate groups has been found responsible to behave as two separate broad heterotic pools, as B × R hybrids reported significantly higher levels of heterosis than B × B or R × R hybrids (Singh et al. 2018). This can be explained by historical strategy of trait-based breeding followed by pearl millet breeding programs in the utilization of germplasm and breeding material in the development of B- and R-lines. The programs utilized germplasm having highly productive traits, like earliness, bold grain size, high tillering potential, along with photo-insensitiveness toward seed/maintainer parent (B-line) development, while global or South Asian germplasm having better adaptation to South Asian ecology along with taller height and good pollen production traits was utilized more toward pollinator/restorer parent (R-line) development. ICRISAT used most of the *Iniadi* germplasm, which refers to the germplasm from "*Togo*" region (region comprising of West African countries of Ghana, Togo, Benin, and Burkina Faso) which was highly productive, early, bold seeded, photo-insensitive, dark gray seed color in the B-line development program (Andrews and Kumar 1996). Also, gene diversity and PIC indicated R-lines to be genetically more diverse than the B-lines, which is due to involvement of mostly non-*iniadi* and more diverse germplasm in the development of R-lines, as reported earlier in pearl millet (Nepolean et al. 2012; Gupta et al. 2015; Singh et al. 2018; Ramya et al. 2018).

Though pearl millet breeding programs have avoided using same germplasm in both the B- and R-line development to maintain them as separate heterotic groups, but some groups (G10, G11, and G12) in the present study had mix of both B- and R-lines indicating the existence of genetic closeness between some of the B- and R-lines. This genetic relatedness between B- and R-lines was observed earlier also in a study based on ICRISAT hybrid parents, which was primarily due to the involvement of some trait-specific donor B-line (when specific trait was not available

Table 7 Mean grain yield of all the hybrids when representative parents of different B-line marker groups were crossed to all the representative R-lines and vice versa, along with MPH and BPH in pearl millet

Marker-based groups	Mean hybrid yield (kg ha^{-1})*	MPH (%)	BPH (%)
B-line group	Crossed to all 22 representative *R*-lines		
G12B	3447.9^{a}	138.5	78.4
G10B	3329.9^{ab}	71.1	45.2
G16B	3236.6^{b}	90.3	69.1
G11B	3234.1^{bc}	84.0	67.4
G13B	3192.9^{bc}	83.0	61.7
G15B	3129.9^{bc}	90.6	62.5
G14B	3000.8^{c}	102.4	62.0
R-line group	Crossed to all 15 representative *B*-lines		
G12R	4122.4^{a}	104.7	55.6
G4R	3559.3^{b}	127.9	97.8
G1R	3381.8^{bc}	77.9	39.7
G6R	3344.7^{bcd}	107.4	78.1
G5R	3338.6^{bc}	61.4	24.1
G3R	3226.0^{cde}	93.5	63.1
G2R	3058.5^{ef}	110.6	87.8
G7R	3055.7^{def}	93.8	71.5
G9R	3032.8^{def}	80.0	51.2
G10R	3006.0^{ef}	79.5	50.9
G11R	2822.2^{fg}	116.2	89.8
G8R	2657.4^{g}	75.3	57.1

*T-test, mean values followed by same letter(s) do not differ significantly at $p = 0.05$

in restorer parent background) in the R-line breeding or vice versa (Gupta et al. 2015).

This study involving hybrid parents from six breeding programs showed that parental lines from any of these programs were distributed across different marker-based groups indicating the existence of significant genetic diversity in both public and private sector pearl millet breeding programs. Also, genetic relatedness was observed between parental lines due to the presence of common parents in their pedigrees. Such genetic similarities found between lines bred from a particular program can happen due to involvement of same set of few parents in breeding or due to directional selection for same set of adaptive traits in that program.

Yield and heterosis

Significant G × E effects were detected for hybrid yields due to the differential response of hybrids to various environments modulated by various changes in climatic and edaphic factors across locations. It is an important component in pearl millet hybrid breeding to evaluate G × E effect for developing products adapted to different environments as this crop is grown in highly unpredictable marginal environments. It is difficult to develop a "universal high-yielding or high-heterotic" hybrid for different environments in crops like pearl millet which are grown in the arid and semiarid tropics, because of divergent and fluctuating environments in these ecologies. The only pearl millet hybrid, which occupied largest acreage across different agroecologies/zones and cultivated across seasons over a longer period in India, has been ProAgro 9444. ProAgro 9444 had grain yield of 3490 kg ha^{-1} in our study, while Pioneer 86M86 was the highest yielder (4077 kg ha^{-1}).

On average, our hybrids produced 60.5% higher yield than their better parents, with a wide range of variation in yield and better-parent heterosis (− 22.7 to 211.5%), indicating good opportunity to select heterotic hybrids for future. The highest yielding hybrid had 71% higher yield than its better-parent and 30% higher yield than the most stable hybrid 9444, while had about 15.3% higher yield than the best check hybrid (86M86). Thirty-eight hybrids among the 320 experimental hybrids out yielded the most adaptable check hybrid ProAgro 9444 by more than 10% heterosis. The parents involved in these 38 heterotic hybrids were mainly from the groups G12R (28.9%), G10B (26.3%), G12B (21.0%), G13B (18.4%), and G4R (10.5%). The parents in the G11B, G15B, G16B, G2R, G7R, and G9R contributed only 5.2, 7.8, 7.8, 5.2, 5.2, and 5.2% each, respectively, to these 38 hybrids.

Positive correlation found between grain yield of hybrids and mean grain yield of parental lines ($r = 0.42$, $P < 0.001$) or with grain yield of better-parent ($r = 0.44$, $P < 0.001$) indicated that continued selection for higher yielding parents as normally practiced in all pearl millet breeding programs

can enhance opportunity for development of high-yielding hybrids. Also, the mean grain yield of parents was found positively correlated with GCA for grain yield ($r = 0.42$, $P < 0.001$) indicating that selection of breeding lines for high grain yield will lead to some extent, indirect selection for high GCA. In this context, it is also important to note that high grain yield potential of parental lines has a direct bearing on hybrid seed production economy.

Identification of heterotic groups in hybrid parents of pearl millet

Eight out of 84 marker group crosses had heterosis of > 10% for grain yield than ProAgro 9444, while 6 marker group crosses had grain yield higher than the best performing hybrid 86M86. These results indicated that it is possible to identify hybrid combinations with yields higher than the best commercial hybrid available in the market based on the heterotic group information.

Earlier, the relative performance of hybrid parents of known origin and pedigree was commonly used, to combine parents from different genetic backgrounds to develop heterotic hybrids. To define heterotic groups in any crop, Menz et al. (2004) suggested that initially germplasm need to be classified into groups based on the estimates of genetic diversity. Molecular markers have been recently used in pearl millet hybrid parents to assess the genetic relatedness; however, there is scarce information on assessing heterotic groups in hybrid parents available across breeding programs and specifically in relation to best performing hybrids available at farmer's field. Pucher et al. (2016) investigated heterotic patterns among geographically close versus distant pearl millet populations of Western and Central African (WCA) origin based on phenotypic evaluation and could not find clear heterotic groups, and the probable reason identified was the presence of high levels of genetic admixture in WCA pearl millet populations (due to a combination of protogyny facilitated outcrossing and extremely robust wind and insect borne pollen in pearl millet), whereas Ramya et al. (2018) recently identified heterotic pools in pearl millet in investigation of 384 hybrid parents, which involved parental lines from one breeding program only.

The present study involving inbred parental lines from six breeding programs attempted to reveal information on heterotic gene pools in the existing hybrid parents. Based on a lead from previous study that the existing B- and R-gene pools in pearl millet behave as two major heterotic groups (Singh et al. 2018), this study investigated hybrids which represented 84 B- × R-marker-based group of crosses. The basic criterion to identify heterotic groups was based on the performance of hybrids among these groups of crosses. Based on the marker group crosses involving hybrid parents available across ICRISAT and pearl millet breeding programs (both public and private sector), parental lines of G12B when crossed to parents of all R-line groups had highest hybrid yield performance, MPH and BPH than other B-line groups and was designated as HGB1. Parents of G12R were found best combiners with parental lines of most of the B-lines (G12B, G10B, G13B, and G11B), so was designated as HGR1. The highest heterosis was observed between these HGB1 (B-lines) × HGR1 (R-lines) marker groups. Similarly, the parents (B-lines) of group G10B and of G4R (R-lines) also had high hybrid yield performance, BPH and MPH and were significantly different in mean hybrid yields than HGB-1 and HGR-1 groups, hence were designated as HGB-2 and HGR-2, respectively. Among R-line marker groups, lines of G12R had high GCA with lines of the most of the B-line groups (G12B, G10B, and G11B), followed by G4. G12B and G10B B-line groups combined with G12R and G4 R-line groups are our preferred choice of hybrids from available hybrid parents to achieve higher-yielding hybrids than the currently available best commercial checks. On average, the hybrids derived using these heterotic groups produced 10.6% (HGB-1 × HGR-1) and 9.3% (HGB-2 × HGR-1) of grain yield over best commercial check Pioneer 86M86. The parents with low or no possibility of producing high-performing hybrids, such as those in G14B and G15B of B-lines and G8R, G11R, G10R, G9R, G7R of R-line groups have limited value for developing high-performing hybrids within currently available germplasm and they have to be further improved to combine with germplasm from other possible heterotic pools or to be diverged with more breeding efforts. It should be noted that the core set of parents from this study generally fits South Asian environment based on hybrid parents available in ICRISAT-India and pearl millet breeding programs in India. Further heterotic groups could be changed or enhanced when new parents with new traits/germplasm are integrated and adapted to the targeted cropping region.

Though the identification of heterotic groups in our study was primarily based on the hybrid performance between marker group crosses, we also need to understand the patterns of SCA and GCA in this breeding material. The public and private sector pearl millet breeding programs involved here and many others in India have been advancing progenies based on performance per se of lines to develop high-yielding seed parental lines to economize the production cost in hybrid seed production plots. Hence, there is not much information available about the status of SCA and GCA variances in the existing B- and R-gene pools of pearl millet. Our study indicated high SCA:GCA ratio (about 2 times, as shown in Table 3) in this representative set of breeding materials, indicating predominance of SCA variance over GCA variance in pearl millet hybrid parents, which was different than many of the established maize hybrid breeding programs of the USA and Europe where low SCA:GCA ratio

was observed (Duvick et al. 2004; Melchinger et al. 2003; Parrisseaux and Bernardo, 2004; Reif et al. 2005; Schrag et al. 2006). Heterotic groups were identified quite early in maize, mainly flint and dent, which helped maize breeders to focus on the development of breeding lines having high GCA within the heterotic groups and simultaneously for high SCA between the heterotic groups. Such clear-cut breeding strategy was missing in the other crops, like in pearl millet, where information about heterotic groups was missing. Though some of the US and European maize breeding programs have now shown reduction in these differences between SCA and GCA variances over a period of time (Melchinger et al. 2003; Schrag et al. 2006) due to continuous selection, some of the other breeding programs, like of the University of Hohenheim, Germany, counterbalanced this reduction by continuous introgression of new germplasm (Fischer 2008). Also, with predominance of GCA variance over SCA variance, early testing became more effective in these programs and superior hybrids could be identified and selected based on their predictions from GCA effects. Considering this scenario in other crops, we suggest that there is need to investigate SCA and GCA variances in the existing B- and R-line heterotic pools of pearl millet to better understand the contribution of GCA and SCA variances toward heterosis and to plan controlled introgression of new germplasm in hybrid breeding programs to balance diversity (long-term selection gain) versus short-term selection gain.

Marker-based GD and hybrid yield/SCA for grain yield had no correlation, as was also reported in other studies in different crops (Riday et al. 2003; Geleta et al. 2004; Dias et al.2004; Teklewold and Becker 2006; Zeid et al. 2003; Ghaly and Al-Sowayan 2014; Gupta et al. 2018). Earlier, Charcosset et al. (1991) demonstrated that the association between marker heterozygosity and QTLs affecting the target trait is a function of the linkage disequilibrium between them. This necessary condition is generally fulfilled in the case of with-in group hybrids because of linkage disequilibrium generated by drift; while no such correlation can be expected between-group hybrids since linkage disequilibrium differs randomly from one heterotic group to the other (Charcosset and Essioux 1994). Hence, distances based on markers cannot predict performance for between-group hybrids, the same result was found in our study for between-heterotic group hybrids, and was also reported earlier (Melchinger et al. 1992; Boppenmeier et al. 1992). Use of markers unlinked to the trait in the estimation of genetic distance could also be other possible reasons for low magnitude of correlations or poor correlations between genetic distances. To overcome this problem, Bernardo (1992) suggested identifying of specific marker loci with tight linkage to those chromosomal segments that determine the expression of the traits of interest. Schrag et al. (2007) determined the haplotype block structure of experimental germplasm from a hybrid maize breeding program to develop models for hybrid performance prediction based on haplotype blocks. They compared the prediction based on haplotype blocks with other approaches and found that prediction based on variable haplotype block length resulted in an improved prediction of hybrid performance compared with the use of single markers. Therefore, estimation of genetic distance between inbred lines based on 'informative markers' has been suggested by some researchers to possibly improve the correlation between genetic distance and hybrid performance. Our results showed that parental groups can be formed first by molecular markers, which may not predict the best hybrid combination, but have practical value in assigning the existing and new hybrid pearl millet germplasm/parental lines into heterotic groups and thus increasing the opportunities to develop desirable hybrids from the best heterotic groups, which is consistent with a previous study in rice (Xie et al. 2013) and maize (Lanza et al. 1997).

Acknowledgements The research reported here was financially supported by CGIAR Research Program on Dryland Cereals (CRP-DC), ICRISAT-Pearl Millet Hybrid Parent Research Consortium (PMH-PRC), and S.M. Sehgal Foundation Endowment Fund, ICRISAT.

Author contribution statement SKG, VS, RKV, RKS, AR, and RG conceived and designed the study; SKG, DY, LDS, KDM, HDP, Su. KG, and RK contributed to experimental materials; SKG, MB, DY, LDS, KDM, HDP, and RK conducted the experiments; SKG, KSP, AR, VC, RRD, AK, RKV, OPY, KNR, and Su. KG analyzed and/or interpreted the data; SKG, KSP, and VC drafted the manuscript; OPY, KNR, SKG, AR, RKV, RKS, and RG revised the manuscript critically for important intellectual content; All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical Standards This article does not contain any studies with human participants or animals performed by any of the authors.

References

Akinwale RO, Badu-Apraku B, Fakorede MAB, Vroh-Bi I (2014) Heterotic grouping of tropical early-maturing maize inbred lines based on combining ability in Striga-infested and Striga-free environments and the use of SSR markers for genotyping. F Crop Res 156:48–62. https://doi.org/10.1016/j.fcr.2013.10.015

Andrews DJ, Kumar KA (1996) Use of the west African pearl millet *iniadi* in cultivar development. Plant Genet Resour Newslett 105:15–22

Bernardo R (1992) Relationship between single-cross performance and molecular marker heterozygosity. Theor Appl Genet 83:628–634. https://doi.org/10.1007/BF00226908

Boppenmaier J, Melchinger AE, Brunklaus-Jung E, Geiger HH, Herrmann RG (1992) Genetic diversity for RFLPs in European maize inbreds. I. Relation to performance of flintYdent crosses for forage traits. Crop Sci 32:895–902

Charcosset A, Essioux (1994) The effect of population structure on the relationship between heterosis and heterozygosity at marker loci. Theor Appl Genet 89:336–343

Charcosset A, Lefort-Buson M, Gallais A (1991) Relationship between heterosis and heterozygosity at marker loci: a theoretical computation. Theor Appl Genet 81:571–575

dos Dias LAS, de Picoli EAT, Rocha RB, Alfenas AC (2004) A priori choice of hybrid parents in plants. Genet Mol Res 3:356–368

Duvick D, Smith J, Cooper M (2004) Long-term selection in a commercial hybrid maize breeding program. In: Janick J (ed) Plant breeding reviews. Wiley, Hoboken, pp 109–152

Fischer S, Mohring J, Schon CC, Piepho HP, Klein D, Schipprack W, Utz HF, Melchinger AE, Reif JC (2008) Trends in genetic variance components during 30 years of hybrid maize breeding at the University of Hohenheim. Plant Breed 127:446–451

Fischer S, Maurer HP, Würschum T, Möhring J, Piepho HP, Schön CC, Thiemt EM, Dhillon BS, Weissmann EA, Melchinger AE, Reif JC (2010) Development of heterotic groups in triticale. Crop Sci 50:584–590. https://doi.org/10.2135/cropsci2009.04.0225

Geleta LF, Labuschagne MT, Viljoen CD (2004) Relationship between heterosis and genetic distance based on morphological traits and AFLP markers in pepper. Plant Breed 123:467–473. https://doi.org/10.1111/j.1439-0523.2004.01017.x

Ghaly SMA, Al-Sowayan SS (2014) A high B1 field homogeneity generation using free element elliptical four-coil system. Am J Appl Sci 11:534–540. https://doi.org/10.3844/ajassp.2014.534.540

Gupta SK, Nepolean TV, Sankar SM, Rathore A, Das RR, Rai KN, Hash CT (2015) Patterns of molecular diversity in current and previously developed hybrid parents of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. Am J Plant Sci 06:1697–1712. https://doi.org/10.4236/ajps.2015.611169

Gupta SK, Nepolean TV, Chinna GS, Rai KN, Hash CT, Das RR, Rathore A (2018) Phenotypic and molecular diversity-based prediction of heterosis in pearl millet (*Pennisetum glaucum* L. (R.) Br.). Crop J 6:271–281. https://doi.org/10.1016/j.cj.2017.09.008

ICRISAT (2012) Hybrid parents research consortium. The jewels of ICRISAT, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, pp 48–51

Kapila RK, Yadav RS, Plaha P, Rai KN, Yadav OP, Hash CT, Howarth CJ (2008) Genetic diversity among pearl millet maintainers using microsatellite markers. Plant Breed 127:33–37. https://doi.org/10.1111/j.1439-0523.2007.01433.x

Kempthorne O (1957) An introduction to genetic statistics. Wiley, New York

Lanza LLB, de Souza Jr CL, Ottoboni LMM, Vieira MLC, de Souza AP (1997) Genetic distance of inbred lines and prediction of maize single-cross performance using RAPD markers. Theor Appl Genet 94:1023–1030. https://doi.org/10.1007/s001220050510

Mace ES, Buhariwalla HK, Crouch JH (2003) A high-throughput DNA extraction protocol for tropical molecular breeding programs. Plant Mol Biol Rep 21:459–460. https://doi.org/10.1007/bf02772596

Melchinger AE (1999) Genetic diversity and heterosis. In: Coors JG, Pandey S (eds) The genetics and exploitation of heterosis in crops. ASA, CSSA and SSSA, Madison, pp 99–118

Melchinger AE, Gumber RK (1998) Overview of heterosis and heterotic groups in agronomic crops. In: Lamkey KR, Staub JE (eds) Concepts and breeding of heterosis in crop plants. CSSA, Madison, pp 29–44

Melchinger AE, Boppenmeier J, Dhillon BS, Pollmer WG, Herrmann RG (1992) Genetic diversity for RFLPs in European maize inbreds: II. Relation to performance of hybrids within versus between heterotic groups for forage traits. Theor Appl Genet 84:672–681

Melchinger AE, Geiger HH, Utz HF, Schnell FW (2003) Effect of recombination in the parent populations on the means and combining ability variances in hybrid populations of maize. Theor Appl Genet 106:332–340

Menkir A, Melake-Berhan A, The C, Ingelbrecht I, Adepoju A (2004) Grouping of tropical mid-altitude maize inbred lines on the basis of yield data and molecular markers. Theor Appl Genet 108:1582–1590. https://doi.org/10.1007/s00122-004-1585-0

Menz MA, Klein RR, Unruh NC, Rooney WL, Klein PE, Mullet JE (2004) Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. Crop Sci 44:1236–1244. https://doi.org/10.2135/cropsci2004.1236

Mula RP, Rai KN, Kulkarni VN, Singh AK (2007) Public-private partnership and impact of ICRISAT's pearl millet hybrid parents research. J SAT Agric Res 5:1–5

Nepolean T, Gupta SK, Dwivedi SL, Bhattacharjee R, Rai KN, Hash CT (2012) Genetic diversity in maintainer and restorer lines of pearl millet. Crop Sci 52:2555–2563. https://doi.org/10.2135/cropsci2011.11.0597

Parrisseaux B, Bernardo R (2004) In Silico mapping of quantitative trait loci in maize. Theor Appl Genet 109:508–514

Perrier X, Flori A, Bonnot F (2003) Data analysis methods. In: Hamon P, Seguin M, Perrier X, Glaszmann JC (eds) Genetic diversity of cultivated tropical plants. Science Publishers, Enfield, pp 43–76

Pioneer, Pakistan (2018) http://www.pioneercom/home/site/pakistan/products/millet/. Accessed 10 June 2018

Pucher A, Sy O, Sanogo MD, Angarawai II, Zangre R, Ouedraogo M, Boureima S, Hash CT, Haussmann BIG (2016) Combining ability patterns among West African pearl millet landraces and prospects for pearl millet hybrid breeding. Field Crop Res 195:9–20. https://doi.org/10.1016/j.fcr.2016.04.035

R Development Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org/. Accessed 20 Jan 2018

Ramya AR, Ahamed ML, Satyavathi CT, Rathore A, Katiyar P, Raj AGB, Kumar S, Gupta R, Mahendrakar MD, Yadav RS, Srivastava RK (2018) Towards Defining Heterotic Gene Pools in Pearl Millet [*Pennisetum glaucum* (L.) R. Br.]. Front Plant Sci 8:1–11. https://doi.org/10.3389/fpls.2017.01934

Rao VN, Rao KPC, Gupta SK, Kizito M, Kumara CD, Nagaraj N, Singh RN, Singh SS, Singh SP (2018) Impact of ICRISAT Pearl Millet Hybrid Parents Research Consortium (PMHPRC) on the Livelihoods of Farmers in India. Nagubadi Research Services, Hyderabad and Monitoring, Evaluation, Impacts and Learning (MEIL) Research Program on Innovation Systems for Drylands (ISD), Research Report vol 75, p 100, ICRISAT, Patancheru, Hyderabad

Reif JC, Melchinger AE, Xia XC, Warburton ML, Hoisington DA, Vasal SK, Beck D, Bohn M, Frisch M (2003) Use of SSRs for establishing heterotic groups in subtropical maize. Theor Appl Genet 107:947–957. https://doi.org/10.1007/s00122-003-1333-x

Reif JC, Hallauer AR, Melchinger AE (2005) Heterosis and heterotic patterns in maize. Maydica 50:215–223

Reif JC, Zhao Y, Würschum T, Gowda M, Hahn V (2013) Genomic prediction of sunflower hybrid performance. Plant Breed 132:107–114. https://doi.org/10.1111/pbr.12007

Riday H, Brummer EC, Campbell TA, Luth D, Cazcarro PM (2003) Comparisons of genetic and morphological distance with heterosis between *Medicago sativa* subsp. *sativa* and subsp. *falcata*. Euphytica. https://doi.org/10.1023/A:1023050126901

Rogers JS (1972) Measures of genetic similarity and genetic distance. In: Wheeler MR (ed) Studies in Genetics, vol 7213. University of Texas Publication, Texas, pp 145–173

SAS Institute Inc. (2017) SAS OnlineDoc® 9.3. SAS Institute Inc, Cary

Satyavathi CT (2017) Review of pearl millet research. Project coordinator Review. In: 52nd annual group meeting, PAU Ludhiana April 28–30, 2017. AICRP on Pearl millet Jodhpur, Rajasthan, India

Schrag TA, Melchinger AE, Sorensen AP, Frisch M (2006) Prediction of single-cross hybrid performance for grain yield and grain dry matter content in maize using AFLP markers associated with QTL. Theor Appl Genet 113:1037–1047

Schrag TA, Maurer HP, Melchinger AE, Piepho HP, Peleman J, Frisch M (2007) Prediction of single-cross hybrid performance in maize using haplotype blocks associated with QTL for grain yield. Theor Appl Genet 114:1345–1355

Singh RK, Chaudhary BD (1977) Biometrical methods in quantitative genetic analysis, 3rd edn. Kalyani Publishers, Ludhiana

Singh AM, Rana MK, Singh S, Kumar S, Kumar D, Arya L (2013) Assessment of genetic diversity among pearl millet [*Pennisetum* glaucum (L.) R.Br.] cultivars using SSR markers. Range Manag Agrofor 34(1):77–81

Singh S, Gupta SK, Thudi M, Das RR, Vemula A, Garg V, Varshney RK, Rathore A, Pahuja SK, Yadav DV (2018) Genetic diversity patterns and heterosis prediction based on SSRs and SNPs in hybrid parents of pearl millet. Crop Sci 58(6):2379–2390

Stich B, Haussmann BIG, Pasam R, Bhosale S, Hash CT, Melchinger AE, Parzies HK (2010) Patterns of molecular phenotypic diversity in pearl millet [*Pennisetum glaucum* (L.) R.Br.] from West Central Africa their relation to geographical environmental parameters. BMC Plant Biol 10:1–10. https://doi.org/10.1186/1471-2229-10-216

Suwarno WB, Pixley KV, Palacios-Rojas N, Kaeppler SM, Babu R (2014) Formation of heterotic groups and understanding genetic effects in a provitamin a biofortified maize breeding program. Crop Sci 54:14–24. https://doi.org/10.2135/cropsci2013.02.0096

Teklewold A, Becker HC (2006) Comparison of phenotypic and molecular distances to predict heterosis and F_1 performance in Ethiopian mustard (*Brassica carinata* A. Braun). Theor Appl Genet 112:725–759. https://doi.org/10.1007/s00122-005-0180-3

Ullah A, Ahmad A, Khaliq T, Akhtar J (2017) Recognizing production options for pearl millet in Pakistan under changing climate scenarios. J Integr Agric 16:762–773. https://doi.org/10.1016/S2095-3119(16)61450-8

Varshney RK, Shi C, Thudi M et al (2017) Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. Nat Biotechnol 35:969–976. https://doi.org/10.1038/nbt.3943

Wang K, Qiu FL, Larazo W, Dela Paz MA, Xie FM (2015) Heterotic groups of tropical indica rice germplasm. Theor Appl Genet 128:421–430. https://doi.org/10.1007/s00122-014-2441-5

Xie FM, He ZZ, Esguerra MQ, Qiu FL, Ramanathan V (2013) Determination of heterotic groups for tropical Indica hybrid rice germplasm. Theor Appl Genet 127:407–417. https://doi.org/10.1007/s00122-013-2227-1

Yadav OP, Rai KN (2013) Genetic improvement of pearl millet in India. Agric Res 2:275–292. https://doi.org/10.1007/s40003-013-0089-z

Zeid M, Schon CC, Link W (2003) Genetic diversity in recent elite faba bean lines using AFLP markers. Theor Appl Genet 107:1304–1314. https://doi.org/10.1007/s00122-003-1350-9

Affiliations

S. K. Gupta[1] · K. Sudarshan Patil[1] · Abhishek Rathore[1] · Dev Vart Yadav[2] · L. D. Sharma[3] · K. D. Mungra[4] · H. T. Patil[5] · Suresh K. Gupta[6] · Ramesh Kumar[2] · Vaibhav Chaudhary[1] · Roma R. Das[1] · Anil Kumar[1] · Vikas Singh[1,7] · Rakesh K. Srivastava[1] · Rajeev Gupta[1] · M. Boratkar[1] · Rajeev K. Varshney[1] · K. N. Rai[1] · O. P. Yadav[8]

1 International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, Telangana, India

2 Chaudhary Charan Singh Haryana Agricultural University (CCSHAU), Hisar, Haryana, India

3 Sri Karan Narendra Agriculture University (SKNAU), Durgapura, Rajasthan, India

4 Junagadh Agricultural University (JAU), Jamnagar, Gujarat, India

5 Mahatma Phule Krishi Vidyapeeth (MPKV), Dhule, Maharashtra, India

6 Hytech Seed India Pvt Ltd, Hyderabad, Telangana, India

7 International Rice Research Institute (IRRI), South Asia Hub, ICRISAT, Hyderabad, Telangana, India

8 ICAR-Central Arid Zone Research Institute (CAZRI), Jodhpur, Rajasthan, India