African and Asian origin pearl millet populations: Genetic diversity pattern and its association with yield heterosis

K. Sudarshan Patil1,2 | Shashi Kumar Gupta1 | Balram Marathi2 | Shashibhushan Danam2 | Ramesh Thatikunta2 | Abhishek Rathore1 | Roma Rani Das1 | Kuldeep Singh Dangi2 | Om Parkash Yadav3

Abstract
Pearl millet [Pennisetum glaucum (L.) R. Br.] is a staple food crop of arid and semi-arid regions of Asia and Africa. Forty-five pearl millet populations of Asian and African origin were assessed for genetic diversity using 29 simple sequence repeat (SSR) markers. The SSR-based clustering and structure analyses showed that Asian origin–Asian bred (As-As) and African origin–African bred (Af-Af) populations were distributed across seven clusters, indicating no strong relationship among populations with their geographical origin. Most of the African origin–Asian bred (Af-As) populations had a higher average number of alleles per locus than As-As or Af-Af populations, and the majority of them clustered separately from As-As or Af-Af populations, indicating that introgression of African origin breeding materials led to the development of new gene pools adapted to the Asian region. Fourteen populations representing seven clusters were crossed according to a diallel mating design to generate 91 population hybrids (seeds of direct and reciprocal crosses were mixed) and evaluated at three locations in 2016. All the 91 hybrids when partitioned into three groups based on genetic distance (GD) between parental combinations (low, moderate, and high), revealed no correlation between GD and panmictic midparent heterosis in any of the groups, indicating that grain yield heterosis cannot be predicted based on GD. Two population hybrids (GB 8735 × ICMP 87307 and Sudan I × Ugandi) exhibited high levels of yield heterosis over standard checks and can be further utilized using different breeding schemes to develop high-yielding pearl millet cultivars.

Abbreviations: (As × Af)-As, (Asian x African) origin–Asian bred; Af-Af, African origin–African bred; Af-As, African origin–Asian bred; AMOVA, analysis of molecular variance; As-As, Asian origin–Asian bred; GD, genetic distance; HWE, Hardy–Weinberg equilibrium; ICMP, ICRISAT millet pollinator; ICMV, ICRISAT millet variety; ICRISAT, International Crops Research Institute for the Semi-Arid Tropics; MCMC, Markov chain Monte Carlo; MRC, Mandore restorer composite; MRD, modified Roger’s distance; OPV, open-pollinated variety; PCH, panmictic commercial heterosis; PCR, polymerase chain reaction; PMPH, panmictic midparent heterosis; QTL, quantitative trait locus; SRC, smut-resistant composite; SSR, simple sequence repeats.

© 2020 The Authors. Crop Science © 2020 Crop Science Society of America
1 | INTRODUCTION

Pearl millet [Pennisetum glaucum (L.) R. Br.], with a global area of ~34 million ha and production of ~31 Tg, ranks sixth among cereals after wheat (Triticum aestivum L.), rice (Oryza sativa L.), maize (Zea mays L.), barley (Hordeum vulgare L.), and sorghum [Sorghum bicolor (L.) Moench] (FAO, 2020). The majority of the pearl millet area is in Africa (~22 million ha) and Asia (~11 million ha). Recently, the crop has expanded to nontraditional areas, such as Brazil (De Assis, De Freitas, & Mason, 2018), the United States, Canada, Mexico, West Asia and North Africa, and Central Asia (Yadav & Rai, 2013) for feed and forage purposes. India is the largest producer of pearl millet in the world, with a production of 9.70 Tg from 7.5 million ha (INDIASTAT, 2018). In India, hybrids occupy ~70% (~5.0 million ha) of the area under pearl millet; the rest is under open-pollinated varieties (OPVs) or landraces (Satyavathi, 2017). Since the 1960s, Indian pearl millet breeding programs have harnessed heterosis by developing hybrid cultivars, with productivity increase of 3% per annum during 1990–2017 (Yadav, Singh, Dhillon, & Mohapatra, 2019). In contrast, African pearl millet productivity has not significantly changed during last three decades from 1988 (691 kg ha⁻¹) to 2018 (718 kg ha⁻¹), though area has increased from 15.8 to 22.1 million ha (FAO, 2020). Until recently, African countries grew only conventional landraces and locally adapted, improved OPVs. Their breeding programs have now started focusing on hybrid breeding, yet there is no viable hybrid cultivar in the seed market.

Increasing genetic gain for grain yield requires broadening the genetic base of cultivars by introducing diverse germplasm and use of modern molecular biology tools in order to achieve higher heterosis. The former strategy has been followed in pearl hybrid millet breeding programs in the past to develop a diverse range of productive parental lines for hybrids. This has included the introduction of diverse African germplasm into the Asian breeding programs (Serba, Perumal, Tesso, & Min, 2017; Yadav & Rai, 2013). Later strategy included use of linkage analysis and genome-wide association studies based on different marker systems in detecting quantitative trait loci (QTLs) for important agronomic traits (Serba et al., 2019).

Recent characterization studies based on various agromorphological traits in pearl millet have reported significant variability among both African germplasm (Bashir, Ali, Ali, Melchinger, et al., 2014; Pucher et al., 2015; Sattler et al., 2018) and Asian germplasm (Kumari et al., 2016; Upadhyaya, Reddy, Ramachandran, Kumar, & Ahmed, 2016). Molecular marker-based studies conducted to date have also reported significant diversity among African populations and landraces (Adeoti et al., 2017; Bashir et al., 2015; Diack et al., 2017; Hu et al., 2015; Jika et al., 2017; McBenedict, Chimwamurombe, Kwembeya, & Maggs-Kolling, 2016), and also in Asian landraces (Chowdari, Davierwala, Gupta, Ranjekar, & Govila, 1998; Vom Brocke, Christinck, Weltzien, Presterl, & Geiger, 2003). However, like other crops, the utilization of available germplasm is quite limited in the current pearl millet breeding programs (Serba et al., 2017; Yadav, Bidinger, & Singh, 2009; Yadav et al., 2017). To use the genetic diversity, molecular markers have also been used to assess the utility of marker-based genetic distance (GD) to predict heterosis for grain yield in various crops and found both positive and negative relationships between marker-based GD and yield heterosis (Dias, Picoli, Rocha, & Alfenas, 2004). In pearl millet, Singh et al. (2018) and Singh and Gupta (2019) reported a significant positive correlation between molecular-based GD and better-parent heterosis for grain yield, whereas no such relationship was found in the crosses involving West African OPVs (Sattler et al., 2019).

None of the studies to date has comprehensively analyzed genetic diversity patterns among an extensive collection of both Asian and African origin populations of pearl millet using molecular markers. Therefore, this study investigated Asian and African populations of pearl millet for the extent of diversity and examined the relationship between marker-based GDs and grain-yield heterosis.

2 | MATERIALS AND METHODS

2.1 | Plant materials and DNA extraction

A series of improved pearl millet populations with genetic materials of African and Asian origins developed by the
pearl millet breeding program of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) since its inception, in collaboration with its regional partners, were used in this study. A set of 45 pearl millet populations involving 10 Asian origin–Asian bred (abbreviated as As-As), seven African origin–African bred (Af-Af), three (Asian × African) origin–Asian bred [(As × Af)-As], and 25 African origin–Asian bred (Af-As) developed at ICRISAT, Patancheru, were evaluated in this study. Pedigrees of these populations are given in Supplemental Table S1.

For DNA extraction, ~100 seeds of each population were sown in five small plastic pots (10 cm) in a glasshouse, and seedlings (20 per pot) were grown for 12–15 days. Genomic DNA was extracted from leaves of 16 randomly selected individuals per population. A set of 720 DNA samples (45 populations × 16 individuals) was isolated using NucleoSpin 96 Plant II Kit (Macherey-Nagel). Electrophoresis (0.8% agarose gel) was performed to test the quality of DNA and quantify it based on lambda DNA (MBI Fermentas). The working DNA samples were normalized uniformly at a concentration of 10 ng μl⁻¹.

### 2.2 Genotyping using SSR markers

Twenty-nine highly polymorphic simple sequence repeat (SSR) markers (Supplemental Table S2), which were mapped earlier across seven pearl millet linkage groups (Allouis, Qi, Lindup, Gale, & Devos, 2000; Budak, Pedraza, Cregan, Baenziger, & Dweikat, 2003; Qi et al., 2004; Rajaram et al., 2013; Senthilvel et al., 2008), were used to genotype the 45 populations. A step-down polymerase chain reaction (PCR) program was performed in a thermal cycler (GeneAmp, PCR System 9700; Applied Biosystems) using 384-well PCR plates. Reaction mixture of 10-μl volume consisted of 2 μl of 10 ng DNA template, 0.5 μl of 1 mM dNTPs, and 0.06 μl of 0.2 U Taq DNA polymerase, 1 μl of 10× Kappa Taq Polymerase buffer with MgCl₂ and 1 μl of primer containing 2 pM μl⁻¹ of forward and 4 pM μl⁻¹ of reverse primer, 0.2 μl of fluorescent dye (either 6-FAM [6-carboxyfluorescein], VIC [2′-chloro-7′-phenyl-1,4-dichloro-6-carboxy-fluorescein], NED, or PET). Amplification conditions in the step-down PCR program were an initial denaturation at 94 °C for 5 min, 10 cycles at 94 °C for 25 s, 64 °C (−1 °C per cycle) for 20 s, and 72 °C for 30 s, followed by 37 cycles at 56 °C for 20 s and 72 °C for 30 s, with a final extension of 72 °C for 20 min.

The DNA fragments were size fractioned on an ABI 3700 automatic DNA sequencer (Perkin-Elmer/Applied Biosystems), a fluorescence-based capillary detection system. After confirmation of amplification of 10–15 random samples per locus on agarose gel, PCR products were multiplexed with 1 μl of each dye-labeled product (FAM, VIC, NED, and PET), 7 μl of Hi-Di Formamide (Applied Biosystems), 0.1 μl of the LIZ-labeled (500[-250]) internal size standard, and 3.9 μl of deionized water, centrifuged, and denatured at 94 °C for 2 min. They were immediately cooled in ice and subjected to capillary electrophoresis on ABI 3700 automatic DNA sequencer (Perkin-Elmer/Applied Biosystems). The peaks were sized and the allele calling was done using Gene Mapper 4.0 software (Applied Biosystems, 1998). Check sample Tift 23D₂B₁ was included during the PCR of each marker and capillary electrophoresis to verify their repeatability.

### 2.3 Assessing association between genetic distance and panmictic heterosis for grain yield

Fourteen parental populations were identified based on the following criteria suggested by Wang, Qiu, Larazo, Angelita, and Xie (2015): (a) representing the original population structure with seven clusters (Table 1), and (b) possessing the maximum allelic variation (representing the distribution of GD of the 45 populations) (Figure 1). Based on the number of populations in each cluster, they were designated as large (>6 populations), medium (6 populations), or small (<6 populations) clusters. Three representative populations from large clusters and one or two representative populations from small to medium clusters were identified. The representative 14 parental populations covered 84.8% of the allelic variation in the
<table>
<thead>
<tr>
<th>Cluster no.</th>
<th>No. of populations: Name of the populations</th>
<th>Total No. and name of the identified representative populations along with avg. GD in parentheses</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>As-As: WRajPop bmr, WRajPop C2, RCB-IC 948, CZ-IC 9802</td>
<td>1: Ugandi 4: AIMP 92901, ICTP 8203, HHVBC, ICMV 155</td>
</tr>
<tr>
<td></td>
<td>Af-Af: -</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Af-As: -</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>(As × Af)-As: -</td>
<td>-</td>
</tr>
<tr>
<td>II</td>
<td>Af-Af: HITIP 88 C1, CZP-86, Raj I71</td>
<td>3: SOSAT C88, GB 8735</td>
</tr>
<tr>
<td></td>
<td>Af-As: -</td>
<td>2: ICMP 87703</td>
</tr>
<tr>
<td></td>
<td>As-As: -</td>
<td>6</td>
</tr>
<tr>
<td>III</td>
<td>Af-Af: -</td>
<td>2: SCI C4 bulk, WC C3 bulk, ICMV-IS 92222</td>
</tr>
<tr>
<td></td>
<td>Af-As: -</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>As-As: -</td>
<td>-</td>
</tr>
<tr>
<td>IV</td>
<td>Af-Af: -</td>
<td>2: CZ-IC 618, MRC General bulk</td>
</tr>
<tr>
<td></td>
<td>Af-As: -</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>As-As: -</td>
<td>-</td>
</tr>
<tr>
<td>V</td>
<td>Af-Af: -</td>
<td>2: ICMV 221, Sudan II, WC C-75, JBV 3, GICKV 98771</td>
</tr>
<tr>
<td></td>
<td>Af-As: -</td>
<td>5</td>
</tr>
<tr>
<td>VI</td>
<td>Af-Af: -</td>
<td>1: CZ-IC 922</td>
</tr>
<tr>
<td></td>
<td>Af-As: -</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>As-As: -</td>
<td>-</td>
</tr>
<tr>
<td>VII</td>
<td>Af-Af: -</td>
<td>1: NWC C2</td>
</tr>
<tr>
<td></td>
<td>Af-As: -</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>As-As: -</td>
<td>-</td>
</tr>
</tbody>
</table>

*As-As, Asian origin and Asian bred; Af-Af, African origin and African bred; Af-As, African origin and Asian bred; (As × Af)-As, (Asian × African) origin and Asian bred; Mean GD of the population in relation to rest of the populations in the respective cluster.
45 populations. Modified Roger's distance (MRD) among the 14 representative populations ranged from 0.64 to 0.85, with a mean of 0.75, representing 45 populations (mean of 0.74 and range from 0.60 to 0.86).

The 14 representative populations were crossed according to full diallel mating design at ICRISAT during the 2016 summer season (February to May). At least 15–20 panicles from each parent were pollinated with bulk pollen collected from 20–25 plants of the respective crossing parent. The 14-parent diallel cross yielded 182 population hybrids (91 F1's and 91 reciprocal F1's). To generate enough quantity of seed for multilocation trials, equal quantities of seeds of each F1 and its reciprocal cross were bulked together, making the total number of crosses 91. The hybrids were evaluated along with parental populations and two popular OPV checks (ICTP 8203 and Dhanashakti) in an α-lattice design with two replications. The evaluation was conducted at three locations: ICRISAT, Patancheru, Telangana (17°30' N, 78°27' E, 545 m asl); Regional Agricultural Research Station, Palem, Telangana (16°53' N, 78°23' E, 545 m asl); and Pearl Millet Research Station, Junagadh Agricultural University, Jamnagar, Rajasthan (22°28' N, 70°04' E and 27.6 m asl), during the rainy season (June–October) of 2016. At Patancheru, each entry was planted in four rows of 4-m length (planted on 6 June), with an inter-row spacing of 75 cm and an intra-row spacing of 15 cm. At Palem (planted on 22 June) and Jamnagar (planted on 25 July), each entry was planted in four rows of 4-m length, with an inter-row spacing of 60 cm and an intra-row spacing of 12 cm. All the recommended agronomic practices were followed at all locations to obtain good crop growth. There were no major incidences of diseases and pests. All the panicles were harvested at physiological maturity (on 14 September at ICRISAT, 29 September at Palem, and 31 October 2016 at Jamnagar). The harvested material was sun dried for 10–15 d and threshed, and data on grain yield (kg ha⁻¹ adjusted to 125 g kg⁻¹ moisture) were recorded.

2.4 | Statistical data analysis

2.4.1 | Diversity analysis

Diversity analysis was conducted using 29 SSR markers to obtain summary statistics, which included allelic richness, as determined by the total number of alleles detected; alleles per locus; major allele frequency; gene diversity; and heterozygosity. Summary statistics for populations (individually and origin-wise) were estimated using PowerMarker version 3.25 (Liu & Muse, 2005). The breeding populations were assumed to be in Hardy–Weinberg equilibrium (HWE) since they were bred and maintained at ICRISAT through random mating. The genetic dissimilarities in each pair of populations based on the MRD matrix was computed according to Wright (1978) using SAS version 9.4 (SAS Institute, 2017). A cluster diagram was constructed based on the neighbor-joining (NJ) method using the DARwin 5.0 program (Perrier & Jacquemoud-Collet, 2006). Analysis of molecular variance (AMOVA) was performed according to Michalakis and Excoffier (1996) to divide molecular genetic variance into components attributed to variance between and within populations. The pairwise fixation index ($F_{ST}$) method was used to infer the distinctness of clusters in a neighbor-joining tree. The analysis was carried out using R program statistical software (R Core Team, 2017). The test to determine deviations from HWE and the existence of nonrandom associations of populations across all SSR loci was computed according to Guo and Thompson (1992) using PowerMarker version 3.25 (Liu & Muse, 2005).

The model-based software program STRUCTURE 2.3 was used to infer population structure by Bayesian approach using SSR marker data (Pritchard, Stephens, & Donnelly, 2000). The optimal value of $K$ (the number of clusters) was determined by performing five independent runs with $K = 2–10$. The length of burn-in of the Markov chain Monte Carlo (MCMC) interactions was set to 100,000, and data were collected across 100,000 MCMC interactions in each run. The optimal value of $K$ was computed using both the ad hoc procedure introduced by Pritchard et al. (2000) and the method developed by Evanno, Regnaut, and Goudet (2005), which improves clustering of closely related populations. The final $K$ value was determined based on (a) the rate of change in mean log probability of the data [lnP(D)] between successive $K$, (b) the stability of grouping pattern across the runs, and (c) the germplasm information on the examined material.

2.4.2 | Analysis of variance

Combined ANOVA was performed following the mixed model method of restricted maximum likelihood (REML) using the PROC MIXED procedure in SAS version 9.4 (SAS Institute, 2017), considering locations, genotypes, and replications as fixed effects and blocks as a random effect. Individual location variances were modeled to error distribution using repeated statement in SAS mixed procedure. Midparent heterosis was calculated as panmictic midparent heterosis (PMPH) according to Lamkey and Edwards (1999), and commercial heterosis was calculated as panmictic commercial heterosis (PCH). The PMPH and PCH were estimated using the following formulae:

$$\text{PMPH} = \left( \frac{\text{HP} - \text{MP}}{\text{MP}} \right) \times 100$$
where $\overline{HP}$ is hybrid performance or mean grain yield of the $F_1$ population hybrid; $\overline{MP}$ is mean grain yield of the two parental populations involved in a cross $= \frac{(P1 + P2)}{2}$; $P1$ and $P2$ are grain yields of parental Populations 1 and 2, respectively; and $\overline{CH}$ is mean grain yield of the commercial check.

ICTP 8203 was used as an OPV check to calculate PCH since it is widely cultivated in India and has desirable traits, such as early maturity, large seed size, resistance to downy mildew ($Sclerospora graminicola$), and good yielding ability in drought-affected environments. Pearson correlation coefficients ($r$) were calculated and significance tests were performed between MRD GD and PMPH using PROC CORR program in SAS version 9.4 (SAS Institute, 2017).

3 | RESULTS AND DISCUSSION

3.1 | Marker-based genetic variation in pearl millet populations

The 29 SSR loci detected 435 alleles in 720 samples (16 individuals randomly sampled from each of the 45 populations). The number of alleles per locus ranged from 6 ($Xicmp3032$) to 32 ($Xpsmp2070$), with a mean of 15 alleles per locus, indicating a high level of molecular diversity. Gene diversity ($H_e$) varied from 0.57 ($Xipes0082$) to 0.96 ($Xpsmp2079.2$), with a mean of 0.75. Observed heterozygosity ($H_o$) ranged from 0.10 ($Xpsmp2070$) to 0.67 ($Xicmp3088$), with a mean of 0.31, indicating a high level of heterozygosity within the populations (Table 2).

Among the 45 populations, the number of alleles ranged from 4.28 (SOSAT C88) to 6.07 (ICMV 97754), $H_o$ ranged from 0.57 (SOSAT C88) to 0.70 (Sudan 1), $H_e$ ranged from 0.22 (GICKV 98771) to 0.41 (RCB-IC 948), and inbreeding coefficient ranged from 0.39 (RCB-IC 948) to 0.66 (GICKV 98771) (Supplemental Table S3). Among the four categories of populations, the average number of alleles per marker ranged from 8.24 [$\{	ext{As} \times \text{Af}\}$-As] to 13.72 (Af-As). However, $H_e$, $H_o$, and inbreeding coefficient showed comparable values among the four categories of populations (Table 3).

The relatively high gene diversity ($H_e = 0.75$), observed heterozygosity ($H_o = 0.31$), and mean number of alleles per locus (15 alleles), obtained in the current study, indicated the presence of considerable genetic variability among the 45 pearl millet populations under investigation. The present study detected a higher number of alleles and gene diversity than found in previous studies using SSRs in

African pearl millet accessions (Adeoti et al., 2017; Diack et al., 2017). This might be due to use of more markers in our study and sampling of a larger number of individuals per population. These highly diverse populations can serve as promising sources to broaden the genetic base of

3.2 | Genetic analysis of populations

The HWE assumes a random mating system that can be assessed by computing the deviation of the $H_e$ from an unbiased estimate of $H_e$. The deviation from HWE was highly significant for all 29 SSR loci ($p < .001$). Populations achieve HWE proportions after one generation of random mating if there is no selection, mutation, or migration (Hardy, 1908; Weinberg, 1909). Given the breeding and maintenance procedures of pearl millet populations at ICRISAT (populations are random mated in isolation during their development and multiplication process), it was expected that the populations were in HWE. However, all 29 SSR loci used in the present study indicated deviation from HWE. This deviation has also been reported in previous studies in other highly cross-pollinated crops, such as maize (Qi-Lun, Ping, Ke-Cheng, & Guang-Tang, 2008; Reif et al., 2004). This could mainly be attributed to genetic causes, such as nonrandom mating, artificial subgrouping of individuals within populations, selection favoring homozygotes, and experimental errors during laboratory assays. Reif et al. (2004) mentioned that experimental errors could probably be the major cause of heterozygote deficiency within populations, apart from genuine genetic causes.

3.3 | Analysis of molecular variance

Results of AMOVA revealed that the variance within populations was significantly higher (88.94%) than that found among the populations (11.07%). The $F_{ST}$ value (0.11) indicated a moderate level of genetic differentiation among the populations, as it accounted for 11% of the total genetic variation. The intrapopulation genetic variation was found higher than many other studies conducted earlier on different pearl millet populations and landraces derived from different regions of Africa (Bashir et al., 2015; Stich et al., 2010; Dussert, Snirc, and Robert, 2015), and also on Indian landraces (Vom Brocke et al., 2003).
### Table 2: Allelic composition, gene diversity, and heterozygosity of the 29 simple sequence repeat (SSR) loci in 45 pearl millet populations

<table>
<thead>
<tr>
<th>SSR loci</th>
<th>Allele size range (bp)</th>
<th>No. of alleles</th>
<th>Major allele frequency</th>
<th>Gene diversity ($H_e$)</th>
<th>Heterozygosity ($H_o$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xctm10</td>
<td>183–219</td>
<td>19</td>
<td>0.19</td>
<td>0.89</td>
<td>0.15</td>
</tr>
<tr>
<td>Xicmp3032</td>
<td>202–217</td>
<td>6</td>
<td>0.43</td>
<td>0.66</td>
<td>0.35</td>
</tr>
<tr>
<td>Xicmp3043</td>
<td>198–228</td>
<td>10</td>
<td>0.46</td>
<td>0.66</td>
<td>0.45</td>
</tr>
<tr>
<td>Xicmp3088</td>
<td>153–192</td>
<td>13</td>
<td>0.24</td>
<td>0.85</td>
<td>0.67</td>
</tr>
<tr>
<td>Xipes0082</td>
<td>154–209</td>
<td>10</td>
<td>0.58</td>
<td>0.57</td>
<td>0.42</td>
</tr>
<tr>
<td>Xipes0152.2</td>
<td>109–137</td>
<td>14</td>
<td>0.40</td>
<td>0.77</td>
<td>0.31</td>
</tr>
<tr>
<td>Xipes0186</td>
<td>169–293</td>
<td>11</td>
<td>0.40</td>
<td>0.70</td>
<td>0.56</td>
</tr>
<tr>
<td>Xipes0200</td>
<td>174–210</td>
<td>11</td>
<td>0.35</td>
<td>0.78</td>
<td>0.33</td>
</tr>
<tr>
<td>Xipes0203</td>
<td>229–307</td>
<td>27</td>
<td>0.22</td>
<td>0.87</td>
<td>0.47</td>
</tr>
<tr>
<td>Xipes0213</td>
<td>163–181</td>
<td>7</td>
<td>0.44</td>
<td>0.62</td>
<td>0.23</td>
</tr>
<tr>
<td>Xipes0220.1</td>
<td>172–216</td>
<td>18</td>
<td>0.33</td>
<td>0.78</td>
<td>0.33</td>
</tr>
<tr>
<td>Xipes0236</td>
<td>215–263</td>
<td>15</td>
<td>0.36</td>
<td>0.74</td>
<td>0.17</td>
</tr>
<tr>
<td>Xpsmp2030</td>
<td>102–160</td>
<td>21</td>
<td>0.48</td>
<td>0.72</td>
<td>0.32</td>
</tr>
<tr>
<td>Xpsmp2068</td>
<td>112–162</td>
<td>23</td>
<td>0.15</td>
<td>0.89</td>
<td>0.24</td>
</tr>
<tr>
<td>Xpsmp2070</td>
<td>204–278</td>
<td>32</td>
<td>0.13</td>
<td>0.93</td>
<td>0.1</td>
</tr>
<tr>
<td>Xpsmp2079.2</td>
<td>117–179</td>
<td>31</td>
<td>0.07</td>
<td>0.96</td>
<td>0.2</td>
</tr>
<tr>
<td>Xpsmp2085</td>
<td>181–199</td>
<td>9</td>
<td>0.39</td>
<td>0.71</td>
<td>0.33</td>
</tr>
<tr>
<td>Xpsmp2086</td>
<td>107–149</td>
<td>18</td>
<td>0.50</td>
<td>0.71</td>
<td>0.23</td>
</tr>
<tr>
<td>Xpsmp2089</td>
<td>117–159</td>
<td>21</td>
<td>0.16</td>
<td>0.90</td>
<td>0.32</td>
</tr>
<tr>
<td>Xpsmp2090</td>
<td>189–211</td>
<td>12</td>
<td>0.32</td>
<td>0.77</td>
<td>0.28</td>
</tr>
<tr>
<td>Xpsmp2201</td>
<td>363–385</td>
<td>9</td>
<td>0.56</td>
<td>0.64</td>
<td>0.14</td>
</tr>
<tr>
<td>Xpsmp2202</td>
<td>161–185</td>
<td>13</td>
<td>0.61</td>
<td>0.59</td>
<td>0.39</td>
</tr>
<tr>
<td>Xpsmp2203</td>
<td>351–389</td>
<td>15</td>
<td>0.21</td>
<td>0.85</td>
<td>0.37</td>
</tr>
<tr>
<td>Xpsmp2207</td>
<td>315–333</td>
<td>10</td>
<td>0.59</td>
<td>0.61</td>
<td>0.37</td>
</tr>
<tr>
<td>Xpsmp2220</td>
<td>136–156</td>
<td>11</td>
<td>0.41</td>
<td>0.74</td>
<td>0.44</td>
</tr>
<tr>
<td>Xpsmp2232</td>
<td>247–275</td>
<td>11</td>
<td>0.28</td>
<td>0.78</td>
<td>0.2</td>
</tr>
<tr>
<td>Xpsmp2248</td>
<td>177–209</td>
<td>17</td>
<td>0.35</td>
<td>0.83</td>
<td>0.2</td>
</tr>
<tr>
<td>Xpsmp2249</td>
<td>154–182</td>
<td>9</td>
<td>0.61</td>
<td>0.58</td>
<td>0.19</td>
</tr>
<tr>
<td>Xpsmp2275</td>
<td>267–300</td>
<td>12</td>
<td>0.37</td>
<td>0.78</td>
<td>0.27</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>15</td>
<td>0.37</td>
<td>0.75</td>
<td>0.31</td>
</tr>
</tbody>
</table>

These results indicated the presence of sufficient variation at the population level in the materials involved in this study, which can be used to select parents to generate new synthetic populations with good adaptation and yield, as suggested by Hartings et al. (2008).

### 3.4 Genetic distance, cluster analysis, and population structure analysis

The highest GD (0.86) was found between Sudan II and ICMP 97774 and the lowest GD (0.56) was found between ICMP 96201 and CZ-IC 922. Cluster analysis, based on the genetic dissimilarity matrix, grouped 45 pearl millet populations into seven major clusters (Figure 2). All the clusters showed significant $F_{ST}$ values, indicating their distinctness (Supplemental Table S4). The number of populations in each cluster varied from five to nine. The As-As populations were distributed across four clusters; Cluster I (4 populations), Cluster II (3 populations), Cluster IV (2 populations), and Cluster VI (1 population). Similarly, Af-Af populations were found distributed across four clusters: Cluster I (1 population), Cluster II (2 populations), Cluster III (3 populations), and Cluster VII
FIGURE 2 Unweighted neighbor-joining tree of 45 pearl millet populations (10 As-As [Asian origin and Asian bred], 7 Af-Af [African origin and African bred], 3 (As × Af)-As [(Asian × African) origin and Asian bred], and 25 Af-As [African origin and Asian bred]) based on modified Roger’s distance matrix using simple sequence repeat (SSR) markers. The filled arrows denote As-As populations, the unfilled arrows denote Af-Af populations, the dotted arrows denote (As × Af)-As populations, and the rest are Af-As populations. CI–CVII denote Clusters I–VII.

The population-structure analysis revealed that all 45 populations could be delineated into three main groups at K = 3 (Figure 3) on the basis of the rate of change in lnP(D).

Group 1 consisted of 24 populations (53.3%), Group 2 had nine populations (20%), and Group 3 had the remaining 12 populations (26.7%). Group 1 was dominated by 14 Af-As populations (58.3%), four As-As (16.7%) populations, and three populations each of (As × Af)-As (12.5%) and Af-Af (12.5%) populations. Among the nine populations of Group 2, eight (88.9%) were Af-As populations, and one (11.1%) was an Af-Af population. Group 3 had an intermix of six As-As (50%), three Af-Af (25%), and three Af-As (25%) populations. Among the 25 Af-As populations, 22 (88%) were found to be together in the first two groups; out of 10 As-As populations, six (60%) were found to be in Group 3, and four (40%) were found to be in Group 1, whereas seven Af-Af populations were found to be distributed across these groups, with three populations each in Groups 1 and 3.

Both MRD matrix-based cluster analysis and population structure analysis showed comparable results for grouping populations. Both analyses revealed a separate clustering of a majority of Af-As populations but did not reveal a clear-cut distinction between As-As and Af-Af populations. Pearl millet populations within Clusters IV, V, VI, and VII (mostly Af-As populations) were found together in Groups 1 and 2 according to the structure analysis. Similarly, populations in Cluster I (As-As and Af-As) belonged to Group 3. Overall, the results showed poor differentiation of populations (As-As and Af-Af) from different geographical (Asian and African) regions which are supported by results of previous studies reporting the absence of any specific structuration among pearl millet accessions from different regions of West Africa (Adeoti et al., 2017) and among landraces of different regions of Sudan and West Africa (Bashir et al., 2015). Based on the
FIGURE 3  Population structure analysis showing the grouping of 45 pearl millet populations. The color bands represent different groups: red denotes Group 1, blue denotes Group 2, and green denotes Group 3. The populations are denoted by circles: lined circle = As-As (Asian origin and Asian bred); full circle = Af-Af (African origin and African bred); thick dotted circle = (As × Af)-As [(Asian × African) origin and Asian bred]; and the uncircled ones = Af-As (African origin and Asian bred). K, number of clusters

clustering pattern, it can be inferred that the molecular GD between the 45 African and Asian populations did not show their clear grouping by geographical origin. This may be attributed to the evolutionary process and the manner in which dispersal of pearl millet from the African continent occurred across India during the domestication process (Fuller & Boivin, 2009; Manning, Pelling, Higham, Schwenniger, & Fuller, 2011), and also because of the spread of germplasm from West Africa to Asian countries, leading to vast gene exchange among the gene pools, resulting in reduced nucleotide diversity (Hu et al., 2015).

Some of the breeding populations related by parentage, or bred for the same region, were found in same clusters, indicating some genetic relatedness; for example, in Cluster II, three As-As populations bred using germplasm from Rajasthan (a northwestern state of India that grows 50% of pearl millet in India) grouped together, and these were found to be grouped separately from three Af-Af populations, which were bred using West African germplasm. Cluster III had a majority of African origin populations from Uganda and Nigeria, except for EC C6, which was an (As × Af)-As population. Cluster IV was composed of four Af-As populations and two As-As populations (CZ-IC 618 and ICMV 94135 and ICMV 94132) derived from the Large Grain Populations (LaGraP), whereas Sudan I and ICMP 00552 were bred using materials originating in Senegal. The two As-As populations were developed using Mandor Restorer Composite (MRC), a composite developed with 12 breeding lines adapted to arid environment such as drought-prone environments of western Rajasthan (Yadav et al., 2012). Also, Cluster V contained five populations belonging to the Af-As populations. Two of these populations (GICKV 98771 and GICKV 96752) were derived from smut-resistant composite (SRC) of African origin. Cluster VI was composed of six populations, which included three Asian adaptation- and origin-linked populations. Two of these populations (ICMS 7704 and ICMP 87307) were developed in Asia through Asian × African crosses, and the third As-As population (CZ-IC 922) was developed at Jodhpur (Rajasthan) using EC C6 (African origin). The other three populations in Cluster VI were derived from African origin genetic material (e.g., SRC [ICMV 93752 and ICMP 96201]) and a bold seeded early composite (ICMV 88908). Also, Cluster VII consisted of seven populations of Af-As, three of which had SRC in their pedigree (ICMP 98107, ICMP 97774, and ICMP 97754), and four had different genetic backgrounds developed using materials with African parentage: Uganda (ICMP 87237), Zimbabwe (ICMP 96601), Nigeria (NWC C2), and half-sib progenies of African origin material HHVBC (ICMP 99901). These instances of genetically related populations found in common clusters indicated that markers used in this study were able to differentiate populations sharing common gene pools.

The overall grouping pattern of 45 populations showed that Af-As populations clustered separately from As-As and Af-Af populations. This might have happened because a diverse range of African germplasm sources was used during the early phase of the Indian pearl millet improvement program to develop several gene pools and composites (e.g., early composite, medium composite, late composite, smut-resistant composite, high-tillering composite, bold-seeded composite, dwarf composite, and high head-volume composite). Selection among these gene pools yielded promising Af-As OPVs, such as WC-C75, ICMP 155, ICTP 8203, ICMP 221, JBV 3, and CZP 9802, which are adapted to Indian conditions and were found to be grouped together. This use of breeding materials of African origin in Asian (Indian) pearl millet improvement programs led to the fixation of alleles favoring adaptation to Asian regions. Serba et al. (2019) also reported that West African genetic accessions have been used extensively in Indian breeding programs. This seems to have led to the differentiation of Af-As populations from the As-As and Af-Af populations, as observed in the present study.
The genetic divergence of Af-As populations from As-As and Af-Af was further evidenced by the larger number of alleles per locus (13.7) observed in Af-As populations compared with As-As (11.6) and Af-Af (11.4) populations. Such allelic differences between interregional populations (Af-As) and intraregional populations (As-As and Af-Af) suggested that breeding efforts involving interregional germplasm to develop improved cultivars and germplasm with a higher number of adaptive alleles should continue.

3.5 Grain yield of population hybrids, panmictic midparent heterosis, and its association with genetic distance

Analysis of variance for grain yield revealed highly significant variance attributable to locations (environments), indicating that the materials were evaluated under diverse environments (Table 4). Large and significant genotypic variation observed in population hybrids and parental populations indicated the existence of adequate genetic variation in parental populations and their hybrids. Significant “hybrids vs. parents” variance indicated the presence of significant heterosis for grain yield in hybrids. Location × hybrid and location × parent interactions were highly significant, suggesting that different genotypes (hybrids or parental populations) had differential response at different test locations.

Across locations, grain yields of population hybrids ranged from 1,652 (EC C6 × GB 8735) to 2,992 kg ha⁻¹.
(GB 8735 × ICMP 87307), and grain yields of parental populations ranged from 1,322 (ICMS 7704) to 2,667 kg ha⁻¹ (Sudan I). Across locations, PMPH varied from −21.7 (EC C6 × ICMP 98107) to 62.1% (ICMS 7704 × ICMP 155), with a mean of 6.4%, and 12 population hybrids had significant positive PMPH (Supplemental Table S5).

Hybrid yield advantage over OPV check ICTP 8203 varied from −3.9 (EC C6 × GB 8735) to 74.0% (GB 8735 × ICMP 87307), with a mean of 30.6% across locations. Out of 91 hybrids, 60 showed >25% heterosis over OPV check ICTP 8203 (Supplemental Table S5). More than half of the population hybrids investigated in this study had significant positive and higher heterosis compared with the standard check ICTP 8203. This clearly demonstrated that the performance of such hybrids could be significantly higher than that of the commercial check, making them a highly suitable base material for further crop improvement. Earlier, Yadav et al. (2012) reported significant heterosis for grain yield while evaluating crosses between Indian and African composites. The highest-yielding population hybrid (GB 8735 × ICMP 87307) had the highest PCH of 74.0%, followed by hybrids Sudan I × Ugandi and HHVBC × ICMP 87307 with ~50% yield heterosis. All these high heterotic cross combinations can be used to improve combining ability and to develop high-yielding, improved populations adapted to regions where OPVs are cultivated. They can also be improved for combining ability and used as base parental populations to derive superior inbreds for hybrid breeding, following reciprocal recurrent selection.

The correlation coefficient between GD of parental populations and PMPH of these 91 population hybrids was investigated. When all the hybrids were partitioned into three groups with high, moderate, and low GD within each group, the correlation coefficient between GD and PMPH was nonsignificant: r was −.24 (p < .26) in 26 hybrids with high GD (0.78–0.86); r was −.13 (p < .39) in 44 hybrids with moderate GD (0.69–0.77); and r was −.09 (p < .68) in 21 hybrids with low GD (0.60–0.68). This indicated that prediction of grain yield heterosis based on GD calculated using molecular markers was not possible in pearl millet populations. Earlier, Dias et al. (2004) reviewed the association of GD with hybrid performance in several crops in 54 different studies and reported that 28 studies showed a positive correlation, whereas the remaining 26 showed a negative correlation or inconclusive results. Also, studies in pearl millet by Chowdari et al. (1998) using RAPD (random amplified polymorphic DNA) markers and by Gupta et al. (2018) using SSR markers did not find a significant correlation between marker-based GDs and midparent heterosis for yield. In the case of West African pearl millet populations, Sattler et al. (2019) also did not find a significant correlation between SSR-based MRD and PMPH. On the contrary, one of the studies in pearl millet (Singh & Gupta, 2019) and a few other studies in maize (Reif et al., 2003a, 2003b) found a significant positive association between GD and yield heterosis. Our study used a set of neutral markers, though non-neutral markers could have helped in finding a relationship between GD and PMPH more accurately if the markers were linked to yield-related QTLs, as earlier suggested by several workers (Bashir et al., 2015; Jordan et al., 2003; Zhang, Gao, Saghai-Maroof, Yang, & Li, 1995). Hence, it is necessary that genetic diversity should be investigated using a higher number of markers that are evenly distributed across the genome and the markers linked to the yield-related traits may help to predict grain yield heterosis. The results suggested the need to generate detailed information on heterotic pools among diverse populations originating and bred in Africa and Asia to enable breeders to use such materials to develop a new series of improved cultivars.

4 | CONCLUSIONS

The present study on SSR marker-based genetic diversity showed the presence of significant genetic variation among African and Asian pearl millet populations. Populations of Af-As group formed clusters separate from those formed by As-As or Af-Af populations. This showed that targeted breeding efforts made in the past had led to the development of new gene pools adapted to the Asian region because of the introgression of materials of Asian and African origin. This introgression of African gene pools into Asian pearl millet breeding pools has contributed significantly to the enhancement of pearl millet productivity in India, which should be continued. No correlation between marker-based GD and heterosis for grain yield indicated that heterosis could not be predicted on the basis of GD. Therefore, in addition to GD between the parents, combining ability patterns also need to be assessed for better prediction of heterosis. This study also underlined the need to further investigate the existence of heterotic groups among pearl millet populations to identify heterotic populations to develop high-yielding OPVs or to derive superior inbreds for hybrid development. The study also identified some parental combinations that can further be used for the development of superior populations and inbreds in pearl millet cultivar development programs.

ACKNOWLEDGMENTS

This research was financially supported by the CGIAR Research Program on Dryland Cereals Scholarship (CRP-DC) and the S.M. Sehgal Foundation Fund (USA) for carrying out a doctoral study of K. Sudarshan Patil at ICRISAT. We gratefully acknowledge Dr. C. T. Hash (formerly Principal Scientist, ICRISAT, Niger) for helping with
pedigrees of populations involved in the study. Thanks are also due to Dr. Marilyn L. Warburton (USDA-ARS Corn Host Plant Resistance Research Unit, PO Box 9555, Mississippi State) for her valuable inputs and suggestions in planning the molecular diversity work.

**CONFLICT OF INTEREST**
The authors declare that there is no conflict of interest.

**ORCID**
K. Sudarshan Patil https://orcid.org/0000-0002-8218-2674

**REFERENCES**


SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of the article.