

Integrating phenotypic and SSR marker-based molecular diversity for prioritization of parental resources in elite maize (*Zea mays* L.) inbred lines

Danish Mushtaq Khanday

khandayd2@gmail.com

Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu

Praveen Singh

Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu

Ravinder Singh Sudhan

Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu

Gyanendra Kumar Rai

Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu

Manmohan Sharma

Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu

Gurdev Chand

Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu

Priyatham Kamilla

Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu

Ankita Kapoor

Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu

Akshita Mahajan

Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu

Sujeela Malik

Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu

Research Article

Keywords: Maize inbreds, phenotypic diversity, SSR markers, molecular diversity, parental selection

Posted Date: September 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10926530/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Integrating phenotypic and SSR marker-based molecular diversity for prioritization of parental resources in elite maize (*Zea mays* L.) inbred lines

Danish Mushtaq Khanday^{*1}, Praveen Singh^{*1}, Ravinder Singh Sudhan¹, Gyanendra Kumar Rai², Manmohan Sharma², Gurdev Chand³, Kamilla Priyatham¹, Ankita Kapoor¹, Akshita Mahajan¹, Sujeela Malik¹

¹ Division of Plant Breeding and Genetics, Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, Jammu 180009, India

² Institute of Biotechnology, Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, Jammu 180009, India

³ Division of Plant Physiology, Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, Jammu 180009, India

*Corresponding authors: Khandayd2@gmail.com ; prof.praveen@gmail.com

Abstract

Genetic diversity provides an important basis for the effective utilization of maize germplasm and the identification of complementary parental resources for breeding. In the present study, 54 elite maize (*Zea mays* L.) inbred lines were characterized for 14 phenological, plant-architectural, cob, kernel and grain-productivity traits during the *Kharif* seasons of 2024 and 2025, together with 20 simple sequence repeat (SSR) markers. Phenotypic variation was assessed using analysis of variance, Mahalanobis' generalized distance, Tocher clustering and principal component analysis, while molecular diversity was examined using Jaccard similarity, UPGMA clustering, principal coordinate analysis and Bayesian population-structure analysis. The inbred lines formed seven phenotypic clusters, with grain yield per plant (16.61%), test weight (14.47%), shelling percentage (9.07%), kernel rows per cob (8.27%), kernels per row (8.16%) and kernels per cob (7.97%) contributing most to phenotypic divergence. SSR analysis resolved the inbred panel into 12 molecular clusters, with the largest cluster comprising 29 lines. The markers showed variable polymorphism, with PIC values ranging from 0.02 to 0.70; *bnlg180*, *phi96100*, *umc1685*, *phi299852* and *bnlg371* were among the more informative markers. The first two principal coordinate axes explained 18.29% and 13.91% of the molecular variation, respectively, while STRUCTURE analysis indicated five inferred genetic groups ($K = 5$) with varying levels of admixture among the lines. Phenotypic and molecular classifications showed partial correspondence, indicating that the two approaches captured complementary dimensions of genetic variation within the breeding material. Integration of trait-specific agronomic performance with molecular relationships enabled the prioritization of G2, G5 and G39 for grain productivity; G13, G14, G28, G47, G48 and G49 for kernel and cob attributes; G26 for shelling efficiency; and G35 and G45 for earliness. Several prioritized lines also occupied differentiated molecular backgrounds, supporting their value as diverse parental resources. Overall, the integrated phenotypic and molecular assessment provided a practical basis for prioritizing genetically differentiated and agronomically valuable parental resources for strategic use in maize breeding.

Keywords: Maize inbreds; phenotypic diversity; SSR markers; molecular diversity; parental selection

1. Introduction

Genetic diversity within elite breeding material is a fundamental resource for sustaining genetic gain in maize (*Zea mays* L.). In hybrid breeding, however, diversity is valuable only when it can be translated into informed parental choice. The effectiveness of a breeding programme therefore depends not simply on identifying high-performing inbred lines, but on recognizing parents that combine agronomic merit with sufficient genetic differentiation and complementary attributes. This distinction is important because molecular distance alone does not consistently predict hybrid performance or combining ability; parental selection is more informative when genetic distinctiveness is considered alongside the observable breeding value of the lines (Mukri et al., 2022; Li et al., 2022).

Phenotypic characterization captures the expressed component of diversity through traits directly relevant to adaptation and breeding, including flowering behaviour, plant architecture, ear and cob characteristics, kernel number, grain weight and yield. Multivariate approaches such as Mahalanobis D^2 analysis, cluster analysis and principal component analysis allow these traits to be evaluated jointly and can distinguish genotypes according to combinations of agronomic attributes rather than individual characters. Considerable phenotypic differentiation among maize inbred panels has been reported, providing opportunities to identify lines with favourable trait combinations for breeding (Dube et al., 2023; Elmyhun et al., 2026). Nevertheless, phenotypic relationships are influenced by the evaluation environment and may therefore provide an incomplete representation of genetic relatedness. Lines from contrasting genetic backgrounds may display similar phenotypes, while genetically related materials may differ considerably in field performance.

Molecular characterization provides a complementary view by resolving genetic variation that may not be evident from agronomic expression. SSR markers remain useful for characterizing maize breeding material because of their codominant inheritance, multiallelic nature and ability to discriminate among inbred lines, despite the increasing use of high-density SNP platforms. Molecular clustering and population-structure analysis can further reveal genetic subdivisions and help identify redundancy or differentiation within breeding pools (Ayesiga et al., 2023; Mukri et al., 2022). Importantly, phenotypic and molecular datasets need not produce identical patterns of relationships. Phenotypic data reflect the expression of traits under particular environmental conditions, whereas molecular markers provide a comparatively environment-independent measure of genetic background. Their agreement can strengthen evidence for consistent

differentiation, while their discordance may expose complementary sources of variation that would be missed by either approach alone. The combined use of phenotypic and molecular information has consequently been advocated for more effective characterization and parental selection in maize (Dube et al., 2023), with recent evidence showing substantial differences between phenotypic and molecular clustering among elite inbred lines (Elmyhun et al., 2026). Despite these advances, an important practical gap remains in the characterization of regionally maintained elite maize breeding material: genetic diversity is often described, but less emphasis is placed on translating that diversity into a breeding-oriented hierarchy of parental resources. Genetic divergence by itself may identify parents that are poorly suited to breeding objectives, whereas selection based primarily on per se performance can favour genetically similar or redundant materials. Integrating multivariate agronomic divergence with molecular relationships can therefore provide a more useful basis for distinguishing lines that are not only genetically differentiated but also agronomically relevant. Such an integrated perspective has the potential to strengthen parental prioritization and guide the strategic use of elite inbred resources in hybrid breeding (Mukri et al., 2022).

In the present study, we evaluated 54 elite maize inbred lines representing diverse breeding backgrounds under rainfed subtropical conditions of Jammu and Kashmir, India. Phenotypic diversity was assessed using 14 agro-morphological, yield and yield-related traits through Mahalanobis D^2 analysis, Tocher clustering and principal component analysis, while SSR markers were used to characterize molecular polymorphism, genetic relationships and population structure. We subsequently compared phenotypic and molecular patterns to determine their correspondence and complementarity and, importantly, to identify breeding-relevant parental resources rather than diversity per se. We hypothesized that phenotypic and molecular characterization capture complementary dimensions of diversity among elite maize inbred lines and that their integration can improve the identification of genetically distinct and agronomically valuable parents. Accordingly, the study aimed to (i) characterize phenotypic diversity and divergence among the inbred lines using multivariate analyses; (ii) characterize molecular diversity and population structure using SSR markers; (iii) assess the correspondence and complementarity between phenotypic and molecular diversity; and (iv) identify parental resources with potential for strategic use in maize breeding.

2. Material and Methods

Fifty-four elite maize (*Zea mays* L.) inbred lines representing diverse breeding backgrounds were used for phenotypic and molecular characterization. The lines were maintained at the Maize Breeding Research Sub-Station, Poonch, Jammu and Kashmir, India, under the maize breeding programme of Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu (SKUAST-Jammu). Details of the inbred designations and origins are provided in Table 1. The inbred lines were evaluated during the *Kharif 2023* and *Kharif 2024* seasons at the same experimental location (33°28'–34°00' N, 73°56'–74°32' E; mean elevation 981 m above mean sea level) using a randomized complete block design (RCBD) with three replications. Standard agronomic practices recommended for maize cultivation were followed throughout crop growth. Young, healthy leaf tissue was collected from representative plants at the three- to four-leaf stage for genomic DNA extraction. Samples were immediately frozen in liquid nitrogen, transported to the Molecular Laboratory, Division of Plant Breeding and Genetics, SKUAST-Jammu, and stored at –80°C until further analysis.

2.2 Field evaluation and phenotypic characterization

Fourteen agro-morphological, phenological, yield and yield-related traits were recorded. Days to 50% tasseling, days to 50% silking and days to maturity were recorded on a plot basis. Plant height (cm), ear height (cm), stem girth (cm), cob length (cm), cob girth (cm), kernel rows per cob, kernels per row, kernels per cob, test weight (g), shelling percentage (%) and grain yield per plant (g) were recorded from five randomly selected competitive plants in each plot.

Plant height was measured from the soil surface to the insertion point of the uppermost tassel branch, while ear height was measured from the soil surface to the node bearing the uppermost ear. Stem girth was measured as the circumference of the stem at a uniform position on the stalk. Cob length was measured from the basal to the apical end, and cob girth was measured at the middle portion of the cob. Kernel rows per cob were determined by counting the longitudinal rows, whereas kernels per row were counted along a representative row. Kernels per cob were determined by counting the total kernels on the representative cob.

Grain yield per plant was determined from the shelled grain harvested from the sampled plants. For test weight, 1000 seeds were counted using an electronic grain counter and weighed

on a digital precision balance, with the value expressed as g per 1000 seeds. Shelling percentage was calculated as:

$$\text{Shelling percentage (\%)} = \text{Weight of shelled grain} / \text{Weight of ear} \times 100$$

Genotype-wise means were calculated for all 14 traits and subsequently used for Mahalanobis D² analysis, cluster analysis and principal component analysis (PCA)

2.3 DNA extraction, SSR marker analysis, PCR amplification, and gel electrophoresis

Genomic DNA was isolated from fresh leaf samples collected from 15-day-old maize seedlings using the cetyltrimethylammonium bromide (CTAB) method described by Doyle and Doyle (1990), with minor modifications. DNA integrity was assessed by electrophoresis on 1% (w/v) agarose gels, while DNA concentration and purity were determined using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). DNA samples were normalized to approximately 50 ng μL^{-1} before PCR amplification. Twenty SSR markers distributed across different regions of the maize genome were selected from the MaizeGDB database based on reported informativeness, reproducibility and genome-wide distribution. Primer sequences, repeat motifs and melting temperatures are presented in Table 2. PCR amplification was performed in a total reaction volume of 10 μL containing 5 μL Emerald 2 $\times$ PCR Master Mix (Takara Bio Inc., Tokyo, Japan), 2 μL genomic DNA (~50 ng), 1 μL each of forward and reverse primers (10 pmol μL^{-1}), and nuclease-free water. Amplification was carried out using an Applied Biosystems Veriti 96-Well Thermal Cycler (Thermo Fisher Scientific, Waltham, MA, USA). The PCR programme consisted of initial denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 45 s, primer-specific annealing at the respective temperature for 45 s and extension at 72°C for 45 s, followed by a final extension at 72°C for 7 min. Amplified products were separated on 3% (w/v) agarose gels prepared in 1 $\times$ Tris-borate-EDTA (TBE) buffer using a 100-bp DNA ladder (Thermo Fisher Scientific, Waltham, MA, USA) as the molecular size standard. Gels were stained with ethidium bromide and visualized using a gel documentation system. Only sharp, unambiguous and reproducible amplification profiles were retained for allele scoring and subsequent analyses.

2.4 SSR data scoring and genetic diversity analysis

Amplified SSR fragments were scored from the electrophoretic profiles, and only distinct and reproducible alleles were retained for analysis. The allelic data for the 54 inbred lines were organized into a genotype matrix for estimating marker polymorphism and genetic diversity parameters. The number of alleles per locus (N_a), major allele frequency (MAF), effective number of alleles (N_e), gene diversity (H_e), and polymorphism information content (PIC) were estimated using PowerMarker version 3.25 (Liu & Muse, 2005). Polymorphism information content was calculated according to Botstein et al. (1980) to assess the discriminatory ability of individual SSR loci. Shannon's information index (I) was calculated using GenAlEx version 6.5 (Peakall & Smouse, 2012). The combined estimates of allele number, allele frequency, effective number of alleles, gene diversity, Shannon's information index, and PIC were used to characterize the extent of molecular variation and evaluate the informativeness of the SSR markers among the maize inbred lines.

2.5 Genetic relationship and population structure analysis

Pairwise genetic relationships among the maize inbred lines were assessed using Jaccard's similarity coefficient, and a dendrogram was generated using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) implemented in DARwin version 6.0 (Perrier & Jacquemoud-Collet, 2006). Population structure was assessed using the Bayesian model-based clustering approach implemented in STRUCTURE version 2.3.4 (Pritchard et al., 2000). The number of genetic groups (K) was evaluated from 2 to 10, with five independent runs conducted for each K value. Each run comprised a burn-in period of 500,000 iterations followed by 500,000 Markov chain Monte Carlo (MCMC) iterations. The most likely number of genetic groups was determined using the ΔK method of Evanno *et al.* (2005), implemented using Structure Harvester (Earl & vonHoldt, 2012). Principal Coordinate Analysis (PCoA) was performed in DARwin version 6.0 using the Jaccard genetic relationship matrix to visualize the relationships among the inbred lines in reduced multidimensional space.

2.6 Statistical analysis

2.6.1 Univariate analysis

The observations recorded for the morphological traits were subjected to analysis of variance (ANOVA), and genotype-wise means were calculated from the five representative plants per plot except for phenological traits which were taken as per plot basis. The standardized

genotype means were subsequently used for multivariate analysis. Trait-wise means and ranges were calculated, and differences among genotypes were tested at the 5% probability level.

2.6.2 Multivariate analysis

Mahalanobis' generalized distance (D^2) between pairs of inbred lines was calculated as:

$$D_{ij}^2 = (X_i - X_j)' S^{-1} (X_i - X_j)$$

Where X_i and X_j represent the vectors of trait means for genotypes i and j , respectively, and S^{-1} represents the inverse of the pooled variance-covariance matrix (Rao, 1952). The resulting D^2 matrix was subjected to cluster analysis, and the relative contribution of individual traits to divergence was assessed from their first-ranking frequencies.

Principal component analysis (PCA) was performed using the standardized trait matrix. Eigenvalues and eigenvectors were obtained from the correlation matrix, and the percentage contribution of each principal component to total variation was calculated as:

$$\text{Variance explained by } PC_k(\%) = \lambda_k / (\sum_{i=1}^p \lambda_i) \times 100$$

where λ_k is the eigenvalue of the k^{th} principal component, and p is the total number of traits included in the analysis. Trait loadings were used to assess the contribution of individual characters to each component, and the first two principal components were used to construct the PCA biplot and visualize the relationships among genotypes and traits (Dube *et al.*, 2023).

All statistical analyses were performed using RStudio IDE (version 2026.08.1; Posit Software, PBC).

3. Results

3.1 Phenotypic variation among maize inbred lines

Analysis of variance revealed significant differences among the 54 inbred lines for all 14 traits ($p < 0.01$; Table 3). Genotype mean squares were highest for kernels per cob (18,197.99) and grain yield per plant (2,052.25). The observed ranges were 52.33–82.33 days for 50% tasseling, 56.33–87.21 days for 50% silking and 92.24–143.33 days for maturity. Plant height, ear height and stem girth ranged from 129.46–205.40 cm, 40.84–111.44 cm and 3.87–6.77 cm, respectively

(Figure 1). Kernel rows per cob, kernels per row and kernels per cob ranged from 7.66–14.34, 6.33–36.21 and 102.99–303.78, respectively. Cob girth and cob length ranged from 11.87–20.60 and 11.75–23.20 cm, respectively. Test weight, shelling percentage and grain yield per plant ranged from 117.52–232.21 g, 14.43–77.77% and 24.82–129.58 g, respectively. G39 (IB-3012) recorded the highest grain yield per plant (129.58 g), followed by G2 (IB-141; 114.09 g) and G5 (CML-367; 111.67 g). G13 (PMS-99) had the highest kernels per cob (303.78), while G5 recorded the highest test weight (232.21 g) and cob length (23.20 cm). The highest shelling percentage was recorded in G17 (CML-76; 77.77%), followed by G16 (IB-104; 73.21%) and G15 (DMR-649; 70.69%). G53 (IB-3009) was the earliest line for tasseling (52.33 days), silking (57.66 days) and maturity (92.24 days). Coefficients of variation ranged from 2.46% for plant height to 12.85% for kernels per cob (Figure 1).

3.2 Phenotypic divergence and cluster analysis

Mahalanobis D^2 analysis grouped the 54 genotypes into seven clusters using Tocher's method (Figure 2). Cluster I contained 35 genotypes, followed by Cluster II with seven and Cluster III with six. Cluster V contained three genotypes, whereas Clusters IV, VI and VII were represented by single genotypes G45, G26 and G35, respectively. The maximum inter-cluster distance was recorded between Clusters V and VII (2494.59), followed by Clusters III and V (1607.38), IV and V (1354.02), and I and V (1201.01). Distances of 1062.73, 1000.97 and 940.93 were recorded for Clusters III $\times$ VI, II $\times$ VII and V $\times$ VI, respectively. The minimum inter-cluster distances were observed between Clusters IV $\times$ VI (254.05), I $\times$ III (285.49) and I $\times$ IV (285.72) (Figure 3). Intra-cluster distances ranged from 0.00 to 176.82. Among non-singleton clusters, Cluster III showed the highest intra-cluster distance (176.82), followed by Clusters II (167.39), I (156.14) and V (139.66). Clusters IV, VI and VII showed zero intra-cluster distance because each contained a single genotype.

3.2.1 Cluster-wise variation in agronomic traits

Cluster means differed for phenological, morphological, yield and yield-related traits (Figure 4). Cluster 5 recorded the highest grain yield per plant (118.78 g), kernels per row (30.11), kernels per cob (363.02), cob girth (17.48 cm), cob length (20.64 cm) and test weight (343.43 g). It had 13.11 kernel rows per cob and 76.16% shelling, with 71.22 days to 50% tasseling, 75.44 days to 50% silking and 132.89 days to maturity. Cluster 6 recorded 91.87 g grain yield per plant and the highest shelling percentage (83.05%), with plant height of 187.56 cm, 7.67 kernels per row and 63.81 kernels per cob. Cluster 2 recorded 75.01 g grain yield per plant, 75.97% shelling, 22.81

kernels per row, 241.25 kernels per cob, 320.62 g test weight and 17.44 cm cob length. Cluster 3 recorded 12.83 kernel rows per cob, 21.11 kernels per row, 238.96 kernels per cob and 18.58 cm cob length, with grain yield of 43.70 g plant⁻¹. Cluster 4 recorded 196.90 cm plant height, 8.65 cm stem girth, 70.55% shelling and 124.33 days to maturity. Cluster 7 showed the earliest cluster mean for tasseling (68.33 days), silking (73.67 days) and maturity (122.33 days), with 49.85% shelling and 64.25 kernels per cob. Cluster 1 recorded 69.82 days to 50% tasseling, 72.74 days to 50% silking, 127.56 days to maturity and 41.60 g grain yield per plant.

3.2.2 Contribution of traits to phenotypic divergence

The relative contribution of the 14 traits to Mahalanobis D² divergence varied among characters (Figure 5). Grain yield per plant contributed the largest proportion (16.61%; 293 first-rank occurrences), followed by test weight (14.47%; 255), shelling percentage (9.07%; 160), kernel rows per cob (8.27%; 146), kernels per row (8.16%; 127) and kernels per cob (7.97%; 141). Plant height contributed 5.99% (106 first-rank occurrences), while cob girth and cob length contributed 5.51% (97) and 5.50% (82), respectively. Stem girth and ear height contributed 4.00% and 3.50%, respectively. Among phenological traits, days to 50% silking contributed 4.76%, followed by days to maturity (3.77%) and days to 50% tasseling (2.41%).

3.3 Principal component analysis of phenotypic variation

The first five principal components accounted for 73.75% of the total phenotypic variation (Figure 6). PC1 explained 23.77%, PC2 18.32%, PC3 12.37%, PC4 10.46% and PC5 8.83%. The first two components together accounted for 42.09% of the total variation. PC1 showed the highest positive loadings for cob length (0.825) and test weight (0.824), followed by kernels per row (0.680), kernel rows per cob (0.612), kernels per cob (0.576), cob girth (0.555), grain yield per plant (0.478) and shelling percentage (0.383). PC2 showed positive loadings for days to 50% tasseling (0.772), maturity (0.732), 50% silking (0.720), ear height (0.623), plant height (0.425) and stem girth (0.270), while shelling percentage (−0.268) and cob girth (−0.244) showed negative loadings. In the PC1–PC2 biplot, G5, G13, G39, G15 and G24 occupied the positive region of PC1, whereas G31, G40, G45, G46 and G35 occurred toward its negative region. G8, G32 and G12 showed relatively positive PC2 scores, while G53, G43 and G47 occurred toward the negative end (Figure 6).

3.4 SSR polymorphism and marker informativeness

The 20 SSR markers differed in their level of polymorphism among the 54 inbred lines (Table 4; Figure S1–S7). The number of polymorphic bands ranged from 1 to 5, major allele frequency (MAF) from 0.38 to 0.98, and effective number of alleles (nE) from 1.02 to 3.41. The marker *bnlg1180* showed the highest effective number of alleles (3.41), Nei's gene diversity (0.70), Shannon's information index (1.12) and PIC (0.70), with four polymorphic bands and a MAF of 0.38. *phi114* produced five polymorphic bands, with PIC and gene diversity values of 0.34. PIC values ranged from 0.02 to 0.70. The highest values were recorded for *bnlg1180* (0.70), *phi96100* (0.48), *umc1685* (0.44) and *phi299852* (0.43), followed by *bnlg371* (0.38), *phi114* (0.34), *bnlg1014* (0.32), *phi069* (0.27), *umc1354* (0.27) and *phi083* (0.25). The lowest PIC values were observed for *phi093* (0.02) and *nc013* (0.03). Nei's gene diversity ranged from 0.02 to 0.70, with the highest values for *bnlg1180* (0.70), *phi96100* (0.48), *umc1685* (0.44) and *phi299852* (0.43). Shannon's information index was highest for *phi114* (1.31) and *bnlg1180* (1.12).

3.4.1 SSR-based genetic relationships

Jaccard similarity analysis grouped the 54 inbred lines into 12 molecular clusters (Figure 7). Cluster I was the largest, comprising 29 genotypes (53.7%), followed by Cluster IX with six (11.1%). Clusters III and XII each contained four genotypes (7.4%), while Clusters V, VI and XI each contained two genotypes (3.7%). Clusters II, IV, VII, VIII and X were singleton groups (1.9% each). Cluster I comprised G1, G3, G4, G6, G8, G10, G17, G22, G25, G26, G27, G29, G30, G31, G32, G33, G34, G35, G36, G37, G38, G39, G41, G42, G43, G44, G45, G46 and G47. Cluster IX contained G18, G19, G20, G21, G23 and G24; Cluster III contained G5, G7, G9 and G28; and Cluster XII contained G51, G52, G53 and G54. Clusters V, VI and XI comprised G12–G13, G14–G40 and G49–G50, respectively. Singleton clusters contained G2, G11, G15, G16 and G48. Within-cluster Jaccard distances ranged from 0.1458 to 0.3581, with the highest values in Cluster I (0.3581), Cluster III (0.2856) and Cluster IX (0.2678). Inter-cluster distances ranged from 0.4237 to 0.7344. The maximum distance occurred between Clusters IV and XI (0.7344), followed by II $\times$ V (0.7210), IV $\times$ X (0.7097), II $\times$ IV (0.7097) and IV $\times$ XII (0.7025). Lower distances were observed between I $\times$ III (0.4237) and I $\times$ VI (0.4268). The UPGMA dendrogram reproduced the 12 molecular groups identified from the Jaccard matrix (Figure 8), including the 29-genotype Cluster I, six-genotype Cluster IX, two four-genotype clusters, three two-genotype clusters and five singleton clusters.

3.5 Molecular ordination and population structure

Principal Coordinate Analysis separated the 54 genotypes along PCoA1 and PCoA2, which accounted for 18.29% and 13.91% of the molecular variation, respectively; together, they explained 32.20% (Figure 9). G23 and G24 occurred toward the upper-left region, whereas G49 and G50 were positioned toward the upper-right. G5 showed the most positive PCoA1 position among the labelled genotypes, while G31, G40, G10, G26, G9 and G45 occurred on the negative side of PCoA1. The 12 UPGMA clusters were also represented in the PCoA distribution, with Cluster I comprising 29 genotypes and Clusters IX, III and XII comprising six, four and four genotypes, respectively. Clusters V, VI and XI contained two genotypes each, whereas Clusters II, IV, VII, VIII and X were singletons. STRUCTURE analysis identified $K = 5$ as the optimal number of genetic groups based on the ΔK criterion (Figure 10). The 54 genotypes displayed varying ancestry proportions at $K = 5$, with individual genotypes showing either predominant membership in one inferred component or mixed membership across two or more components.

3.6 Correspondence between phenotypic and molecular clustering

The phenotypic and SSR-based analyses resolved the 54 genotypes into seven and 12 clusters, respectively (Figure 11). Phenotypic Cluster I contained 35 genotypes, of which 20 were assigned to molecular Cluster I. The remaining 15 genotypes occurred in molecular Clusters III, IV, V, VI, VIII, IX, XI and XII. Phenotypic Cluster II contained seven genotypes: four in molecular Cluster I, two in Cluster IX and one in Cluster VII. The six genotypes of Phenotypic Cluster III were distributed among molecular Clusters I, III, V, VI, X and XI. G45, G26 and G35, representing Phenotypic Clusters IV, VI and VII, respectively, were all assigned to molecular Cluster I. Phenotypic Cluster V comprised G2, G39 and G5, which were assigned to molecular Clusters II, I and III, respectively. Conversely, molecular Cluster I contained genotypes representing all seven phenotypic clusters. Molecular Cluster III included genotypes from Phenotypic Clusters I, II, III and V, while Molecular Clusters V and VI contained genotypes from Phenotypic Clusters I and III. Molecular Cluster IX contained genotypes from Phenotypic Clusters I and II. Distinct cross-classifications were observed for G2, G48, G49, G50, G45, G26 and G35. G2 occurred in Phenotypic Cluster V and molecular Cluster II; G48 occurred in Phenotypic Cluster III and molecular Cluster X; G49 and G50 formed molecular Cluster XI but belonged to Phenotypic Clusters III and I, respectively; and G45, G26 and G35 belonged to separate phenotypic clusters but were all included in molecular Cluster I.

3.10 Integrated characterization of breeding resources

The combined analysis identified contrasting agronomic and molecular profiles among the 54 inbred lines (Table 5). Phenotypic Cluster 5 had the highest cluster mean for grain yield (118.78 g plant⁻¹), kernel rows per cob (13.11), kernels per row (30.11), kernels per cob (363.02), cob girth (17.48 cm), cob length (20.64 cm) and test weight (343.43 g). Its members G2, G39 and G5 occurred in molecular Clusters II, I and III, respectively. G2 combined membership in the highest-yielding phenotypic Cluster 5 with placement in singleton molecular Cluster II. G26, representing Phenotypic Cluster 6, recorded 91.87 g grain yield plant⁻¹ and 83.05% shelling and was assigned to molecular Cluster I. Phenotypic Cluster 2 recorded 75.01 g grain yield plant⁻¹, 75.97% shelling, 22.81 kernels per row, 241.25 kernels per cob and 320.62 g test weight. Phenotypic Cluster 3 recorded 12.83 kernel rows per cob, 21.11 kernels per row, 238.96 kernels per cob and 18.58 cm cob length, with its constituent genotypes distributed among different molecular clusters. Clusters 4 and 7 recorded the earliest cluster maturity values of 124.33 and 122.33 days, respectively. Singleton molecular groups included G2 (Cluster II), G11 (Cluster IV), G15 (Cluster VII), G16 (Cluster VIII) and G48 (Cluster X), while G49 and G50 formed molecular Cluster XI. G2 and G15 occurred in phenotypically differentiated groups while occupying singleton molecular clusters.

4. Discussion

4.1 Phenotypic diversity and agronomic differentiation

The seven Mahalanobis D² clusters demonstrated substantial phenotypic structuring among the 54 maize inbreds, with a large core in Cluster I (35 genotypes) and smaller groups, including three singleton clusters (G45, G26 and G35). The uneven cluster distribution agrees with previous maize studies reporting population-dependent phenotypic structuring of inbred collections (Kumar et al., 2024; Yeri et al., 2025; Sharanu et al., 2025). The broad separation observed across flowering, maturity, plant architecture, kernel, cob and yield traits further confirms that differentiation was multidimensional rather than attributable to a single character, as also reported for maize inbred panels (Yeri et al., 2025; Guo et al., 2025). The cluster means showed that agronomic value was distributed among different groups. Cluster V combined the highest grain yield with high kernel number, cob dimensions and test weight, whereas Cluster VI was distinguished by shelling percentage. Cluster III showed favourable kernel and cob attributes despite lower grain yield, while Clusters IV and VII provided earlier phenological backgrounds. Similar multi-trait differentiation has been reported in maize breeding material, where parental value is better defined by combinations of yield components and adaptation traits than by yield

alone (Bhatla et al., 2024; Bhat et al., 2024). The association of ear traits with yield potential reported by Patil et al. (2024) is also consistent with the superior kernel and cob profile of Cluster V. Thus, the present clusters represent complementary agronomic profiles rather than a simple ranking of genotypes. The D² analysis further showed that grain yield per plant (16.61%), test weight (14.47%), shelling percentage (9.07%), kernel rows per cob (8.27%), kernels per row (8.16%) and kernels per cob (7.97%) were the principal contributors to divergence. Their high first-rank frequencies reinforce their recurrent contribution across genotype comparisons. Similar importance of grain yield, kernel and ear traits in maize diversity has been reported by Korsa et al. (2024), Devi et al. (2024) and Keshamalla et al. (2024). The contribution of test weight is consistent with evidence for genetic variation in grain-weight traits (King et al., 2024), while the contribution of cob dimensions agrees with reported associations of cob length and girth with kernel yield (Sabitha et al., 2024). Phenological traits contributed less individually, indicating that the major D² separation in this panel was driven principally by productivity and yield-component variation. PCA supported this pattern. PC1 was dominated by cob length, test weight and kernel traits, whereas PC2 primarily represented flowering, maturity and plant stature. The first two components explained 42.09% of the variation, while the first five accounted for 73.75%. Similar separation of yield-related and phenological/architectural traits has been reported in maize (Mehra et al., 2025; Reddy et al., 2025; Hu et al., 2024). The agreement between D² and PCA therefore identifies a consistent phenotypic structure centred on productivity and ear architecture, with phenology providing an additional axis of variation.

4.2 Molecular diversity and genetic structure

The 20 SSR markers detected variable levels of polymorphism, with PIC values ranging from 0.02 to 0.70. *bnlg1180* was the most informative locus, combining the highest PIC, effective allele number and gene diversity, whereas *phi093* and *nc013* were weakly informative. The locus-specific variation in allele frequency, gene diversity and PIC is consistent with previous SSR-based characterization of maize inbreds (Ahmed et al., 2023; Al-Hazemawi et al., 2024; Galal et al., 2025; Islam et al., 2023). The results indicate that molecular resolution within this panel was concentrated in a subset of informative loci rather than distributed uniformly across the marker set. SSR-based UPGMA separated the 54 inbreds into 12 molecular clusters, with 29 genotypes forming the major Cluster I and the remaining lines distributed among smaller and singleton groups. Comparable molecular structuring, comprising broad genetic groups accompanied by smaller differentiated lineages, has been reported in maize (Sathua et al., 2018; Stanley et al.,

2020). Large-scale maize analyses have likewise shown that broad molecular groups can contain substantial internal variation despite shared genetic backgrounds (Tian et al., 2021). The present molecular structure should therefore be viewed as a hierarchy of marker-defined relatedness rather than as 12 completely isolated populations. PCoA provided a continuous representation of the SSR variation, with PCoA1 and PCoA2 accounting for 18.29% and 13.91%, respectively. The separation of several entries from the central distribution was consistent with the smaller molecular groups identified by UPGMA. STRUCTURE further resolved five inferred ancestry components, with several genotypes showing mixed membership. Comparable admixture and multi-layered molecular structure have been reported in elite and tropical maize germplasm (Gedil et al., 2023; Gonhi et al., 2025; Mitsuhashi, 2024). The agreement among UPGMA, PCoA and STRUCTURE was therefore broad rather than identical, reflecting their different representations of the same marker variation. Similar multi-method molecular organization has been reported in maize (Badu-Apraku et al., 2021; Chňapek et al., 2024).

4.3 Phenotypic–molecular complementarity

The most informative finding was the incomplete correspondence between phenotypic and molecular classification. The 54 inbreds formed seven phenotypic clusters but 12 SSR-based clusters, and the cross-classification showed substantial redistribution of genotypes between the two systems. Twenty of the 35 genotypes in phenotypic Cluster I occurred in molecular Cluster I, yet molecular Cluster I contained genotypes from all seven phenotypic clusters. Similarly, the three members of phenotypic Cluster V (G2, G39 and G5) occurred in three different molecular clusters. This pattern indicates that phenotypic resemblance was not equivalent to marker-defined relatedness. Similar observations have been reported in maize QPM and drought-related germplasm, where morphological and SSR classifications showed partial rather than one-to-one correspondence (Gazal et al., 2020; Singh et al., 2023). Dube et al. (2023) and Ayesiga et al. (2023) likewise demonstrated that molecular characterization can provide information additional to agronomic evaluation. The present cross-classification therefore supports the use of both datasets when characterizing elite breeding material: phenotypic data identify expressed agronomic differences, whereas SSR data provide an independent measure of marker-defined genetic background. The asymmetry of the relationship is particularly relevant. Phenotypically similar lines may occupy different molecular groups, while molecularly similar lines may span contrasting phenotypic backgrounds. This explains why either dataset alone can miss useful parental contrasts. Recent maize studies have similarly emphasized the value of integrating agronomic performance

with molecular relationships when organizing breeding material (Kumar et al., 2023; Guo et al., 2025). The present results extend this principle to a regionally maintained elite panel by showing that molecular information can refine phenotypic selection rather than merely describe diversity.

4.4 Integrated parental resources and breeding relevance

The combined results identify distinct parental functions within the panel. G2, G5 and G39 were the strongest yield-oriented candidates because they belonged to the high-performing phenotypic Cluster V but occupied three different molecular backgrounds. G13, G14, G28, G47, G48 and G49 provided contrasting kernel and cob attributes and were distributed among several molecular clusters. G26 was distinguished by high shelling percentage, while G35 and G45 represented early phenological resources. These contrasting functions are consistent with maize breeding studies showing that parental value depends on trait complementarity, combining ability and genetic background (Kamara et al., 2024; Engida et al., 2024; Mebratu et al., 2024; Manigben et al., 2024). Of particular interest were G2, G15, G48 and G49, in which favourable agronomic attributes coincided with placement in relatively small molecular groups. Such lines provide a more targeted basis for parental diversification than molecular distance alone. Similar emphasis on genetic background and heterotic-group definition has been reported for Indian maize inbreds and diverse heterotic groups (Rajasekhar et al., 2025; Xue et al., 2025). Nevertheless, the present data do not establish combining ability or heterosis. Genetic distance has not consistently predicted heterosis or SCA in maize, and direct hybrid evaluation remains necessary (Kamara et al., 2024; Engida et al., 2024). The integrated profiles also suggest specific testable crossing directions. Cluster V provides a high-productivity background that could be evaluated against Cluster VII for earliness, Cluster VI for shelling efficiency and Cluster III for kernel and cob attributes. Such complementarity-based crossing is preferable to selecting parents solely because they are statistically distant. Previous studies similarly demonstrate that heterotic grouping and parental diversity are most useful when followed by testcross and combining-ability evaluation (Xiang et al., 2024; Li et al., 2023; Bhatla et al., 2025). Differences among heterotic-grouping methods and the dependence of superior combinations on direct hybrid evaluation further caution against treating molecular divergence as a prediction of hybrid superiority (Bhat et al., 2024; Mosa et al., 2024; Kamara et al., 2025; Mosa et al., 2025).

Overall, the 54 inbreds contained a structured combination of agronomic and molecular diversity, with seven phenotypic groups and 12 SSR-defined molecular groups. Grain productivity, test weight, shelling and kernel/cob traits were the principal sources of phenotypic divergence,

while SSR analysis revealed additional differentiation within broad phenotypic backgrounds. The principal outcome is therefore not the identification of a single superior inbred, but a set of trait-specific parental resources in which agronomic merit and molecular differentiation can be considered jointly. This provides a focused framework for subsequent testcrossing, GCA/SCA estimation and multi-environment validation, which are required to establish the actual hybrid value of the identified candidates.

5. Conclusion and recommendations

The present study shows that phenotypic evaluation and SSR-based molecular characterization provide different but complementary perspectives on variation among elite maize inbred lines. Phenotypic information was useful for recognizing lines with specific agronomic attributes, while the SSR profiles helped distinguish their genetic backgrounds. Considering both types of information together therefore provides a more useful basis for deciding which inbred lines should be retained as parental resources for particular breeding objectives. The prioritized lines represent different breeding targets. G2, G5 and G39 are relevant to grain-yield improvement; G13, G14, G28, G47, G48 and G49 to kernel number and cob architecture; G26 to grain yield and shelling efficiency; and G35 and G45 to earliness-related improvement. The occurrence of favourable phenotypes in different molecular backgrounds further allows parental choice to be made with genetic differentiation in mind, rather than relying on phenotypic performance alone. Accordingly, the identified inbred lines can be taken forward for objective-specific crossing and progeny evaluation.

Author Contributions (CRediT)

Danish Mushtaq Khanday: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing.

Praveen Singh: Conceptualization, Methodology, Supervision, Validation, Visualization, Project administration, Writing – review & editing.

Ravinder Singh Sudhan: Investigation, Field evaluation, Data curation, Writing – review & editing.

Gyanendra Kumar Rai: Supervision, Investigation, Writing – review & editing.

Manmohan Sharma: Supervision, Investigation.

Gurdev Chand: Supervision, Validation.

Kamilla Priyatham: Software, Formal analysis, Investigation.

Ankita Kapoor: Investigation, Data curation, Visualization.

Akshita Mahajan: Investigation, Formal analysis.

Sujeela Malik: Investigation, Data curation.

All authors: Read and approved the final version of the manuscript.

Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgements: The authors acknowledge the Division of Plant Breeding and Genetics, Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu (SKUAST-Jammu) for institutional and academic support. The authors also acknowledge the Council of Scientific and Industrial Research (CSIR), Government of India, for its institutional support.

Data availability statement: The datasets generated and/or analyzed during the current study are not publicly available but are available from the corresponding author on reasonable request.

References

Al-Hazemawi, A. K. H., Hassawi, D. S., & Mohammed, A. A. (2024). The genetic diversity of maize inbred lines grown in Iraq using simple sequence recurrent (SSR) markers. *Iraqi Journal of Agricultural Sciences*, 55(Special), 293–300. <https://doi.org/10.36103/ijas.v55iSpecial.1908>.

Ayesiga, S. B., Rubaihayo, P., Oloka, B. M., Dramadri, I. O., Edema, R., & Sserumaga, J. P. (2023). Genetic variation among tropical maize inbred lines from NARS and CGIAR breeding programs. *Plant Molecular Biology Reporter*, 41(2), 209–217. <https://doi.org/10.1007/s11105-022-01358-2>.

Ayesiga, S. B., Rubaihayo, P., Oloka, B. M., Dramadri, I. O., Edema, R., & Sserumaga, J. P. (2023). Genetic variation among tropical maize inbred lines from NARS and CGIAR breeding programs. *Plant Molecular Biology Reporter*, 41(2), 209–217. <https://doi.org/10.1007/s11105-022-01358-2>.

Badu-Apraku, B., Garcia-Oliveira, A. L., Petroli, C. D., Hearne, S., Adewale, S. A., & Gedil, M. (2021). Genetic diversity and population structure of early and extra-early maturing maize germplasm adapted to sub-Saharan Africa. *BMC Plant Biology*, 21, 96. <https://doi.org/10.1186/s12870-021-02829-6>.

Bhat, J. S., Kumar, C. K., Mukri, G., Neelam, S., & Naidu, G. K. (2024). Heterotic grouping methods reveal differential breeding efficiencies and season-specific categorization of inbreds in tropical maize (*Zea mays* L.). *Indian Journal of Genetics and Plant Breeding*, 84(3), 402–414. <https://doi.org/10.31742/ISGPB.84.3.11>.

Bhatla, A., Angidi, S., Thomas, N., Madankar, K., & Shahi, J. P. (2025). Classification of maize inbred lines into heterotic groups based on yield and yield attributing traits. *Plant Genetic Resources: Characterization and Utilization*, 23, 27–39. <https://doi.org/10.1017/S1479262124000509>.

Bhatla, A., Angidi, S., Thomas, N., Madankar, K., & Shahi, J. P. (2025). Classification of maize inbred lines into heterotic groups based on yield and yield attributing traits. *Plant Genetic Resources: Characterization and Utilization*, 23(1), 27–39. <https://doi.org/10.1017/S1479262124000509>.

Bhatla, A., Angidi, S., Thomas, N., Madankar, K., & Shahi, J. P. (2025). Classification of maize inbred lines into heterotic groups based on yield and yield attributing traits. *Plant Genetic Resources*, 23(1), 27–39. <https://doi.org/10.1017/S1479262124000509>.

Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 32, 314–331.

Chňápek, M., Balážová, Ž., Špaleková, A., Gálová, Z., Hromadová, Z., Číšecká, L., & Vivodík, M. (2024). Genetic diversity of maize resources revealed by different molecular markers. *Genetic Resources and Crop Evolution*, 71, 4685–4703. <https://doi.org/10.1007/s10722-024-01908-5>.

Devi, V., Sethi, M., Kaur, C., Raigar, O. P., Kumar, R., & Chaudhary, D. P. (2024). Genetic diversity among maize (*Zea mays*) inbred lines for methionine and yield associated traits. *The Indian Journal of Agricultural Sciences*, 94(2), 216–219. <https://doi.org/10.56093/ijas.v94i2.142323>.

Doyle, J. J., & Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, 12, 13–15.

Dube, S. P., Shayanowako, A. I. T., Kutu, F., & Sibiya, J. (2024). Performance of tropical and subtropical maize inbred lines under well-watered and drought-stressed environments. *Euphytica*, 220, 139. <https://doi.org/10.1007/s10681-024-03395-2>

Dube, S. P., Sibiya, J., & Kutu, F. (2023). Genetic diversity and population structure of maize inbred lines using phenotypic traits and single nucleotide polymorphism (SNP) markers. *Scientific Reports*, 13, 17851. <https://doi.org/10.1038/s41598-023-44961-3>.

Dube, S. P., Sibiya, J., & Kutu, F. (2023). Genetic diversity and population structure of maize inbred lines using phenotypic traits and single nucleotide polymorphism (SNP) markers. *Scientific Reports*, 13, 17851. <https://doi.org/10.1038/s41598-023-44961-3>.

Dube, S. P., Sibiya, J., & Kutu, F. (2023). Genetic diversity and population structure of maize inbred lines using phenotypic traits and single nucleotide polymorphism (SNP) markers. *Scientific Reports*, 13, 17851. <https://doi.org/10.1038/s41598-023-44961-3>.

Dube, S. P., Sibiya, J., & Kutu, F. (2023). Genetic diversity and population structure of maize inbred lines using phenotypic traits and single nucleotide polymorphism (SNP) markers. *Scientific Reports*, 13, 17851. <https://doi.org/10.1038/s41598-023-44961-3>.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>

Elmyhun, M., Abate, E., Abate, A., Menkir, A., & Teklewold, A. (2026). Genetic diversity and population structure of elite maize (*Zea mays*) inbred lines using phenotypic data and single nucleotide polymorphisms. *Plant Breeding*, 145(1), 166–176. <https://doi.org/10.1111/pbr.70026>.

Engida, B. T., Wegary, D., Keno, T., Mekonnen, T. W., et al. (2024). Combining ability and genetic distance analysis of mid-altitude sub-humid agroecology adapted maize inbred lines for high grain yield. *Heliyon*, 10(11), e32267. <https://doi.org/10.1016/j.heliyon.2024.e32267>

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>

Galal, A. A., Safhi, F. A., Al Aboud, N. M., Aljabri, M., Kucher, D. E., Kamara, M. M., El-Mogy, M. M., Ibrahim, O. M., Abd El-Moneim, D., Hassanin, A. A., El-Soda, M., Mansour, E., & Ali, A. M. (2025). Molecular diversity and genetic potential of new maize inbred lines across varying sowing conditions in arid environment. *Scientific Reports*, 15, 2809. <https://doi.org/10.1038/s41598-025-87021-8>.

Gazal, A., Nehvi, F. A., Lone, A. A., & Dar, Z. A. (2020). Studies on drought tolerance in maize inbred lines using morphological and molecular approaches. *Maydica*.

Gedil, M., Maazou, A.-R. S., Zebire, D. A., Garcia-Oliveira, A. L., Unachukwu, N., Petroli, C., Hearne, S., Everett, L. A., Kim, S.-K., & Menkir, A. (2023). Genetic structure analysis and

identifying key founder inbred lines in diverse elite sub-tropical maize inbred lines. *Scientific Reports*, 13, 11695. <https://doi.org/10.1038/s41598-023-38980-3>.

Gonhi, T., Odong, T. L., Dramadri, I. O., Ochwo-Ssemakula, M., Chiteka, Z. A., Adjei, E. A., Muungani, D., Menkir, A., Badu-Apraku, B., Adejumobi, I., Uwimana, B., Dhliwayo, T., Wegary, D., & Derera, J. (2025). Assessment of genetic diversity and heterotic alignment of CIMMYT and IITA maize inbred lines adapted to sub-Saharan Africa. *Crop Science*, 65(1), e21401. <https://doi.org/10.1002/csc2.21401>.

Gonhi, T., Odong, T. L., Dramadri, I. O., Ochwo-Ssemakula, M., Chiteka, Z. A., Adjei, E. A., Muungani, D., Menkir, A., Badu-Apraku, B., Adejumobi, I., Uwimana, B., Dhliwayo, T., Wegary, D., & Derera, J. (2025). Assessment of genetic diversity and heterotic alignment of CIMMYT and IITA maize inbred lines adapted to sub-Saharan Africa. *Crop Science*, 65(1), e21401. <https://doi.org/10.1002/csc2.21401>.

Gunundu, R., Shimelis, H., & Tesfamariam, S. A. (2025). Genetic diversity and population structure analyses of tropical maize inbred lines using single nucleotide polymorphism markers. *PLOS ONE*, 20(1), e0315463. <https://doi.org/10.1371/journal.pone.0315463>.

Guo, S., Zheng, X., Wang, S., Ai, Y., Zhao, R., & Lan, J. (2025). Genetic diversity of fresh maize germplasm revealed by morphological traits and SSR markers. *Genes*, 16(10), 1138. <https://doi.org/10.3390/genes16101138>.

Guo, S., Zheng, X., Wang, S., Ai, Y., Zhao, R., & Lan, J. (2025). Genetic diversity of fresh maize germplasm revealed by morphological traits and SSR markers. *Genes*, 16(10), 1138. <https://doi.org/10.3390/genes16101138>.

Hu, M., Tian, H., Yang, K., Ding, S., Hao, Y., Xu, R., Zhang, F., Liu, H., & Zhang, D. (2024). Comprehensive evaluation and selection of 192 maize accessions from different sources. *Plants*, 13(10), 1397. <https://doi.org/10.3390/plants13101397>.

Islam, M. A., Alam, M. S., Maniruzzaman, M., & Haque, M. S. (2023). Microsatellite marker-based genetic diversity assessment among exotic and native maize inbred lines of Bangladesh. *Saudi Journal of Biological Sciences*, 30(8), 103715. <https://doi.org/10.1016/j.sjbs.2023.103715>

Kamara, M. M., Al Aboud, N. M., Alsamadany, H., Kutby, A. M., Ben Abdelmalek, I., Abd El-Moneim, D., & Al-Mutiry, M. (2025). Combining ability and heterotic effects in newly developed early maturing and high-yielding maize hybrids under low and recommended nitrogen conditions.

Kamara, M. M., Mansour, E., Khalaf, A. E. A., Eid, M. A. M., Hassanin, A. A., Abdelghany, A. M., Kheir, A. M. S., Galal, A. A., Behiry, S. I., Silvar, C., & El-Hendawy, S. (2024). Molecular diversity and combining ability in newly developed maize inbred lines under low-nitrogen conditions. *Life*, 14(5), 641. <https://doi.org/10.3390/life14050641>

Kamara, M. M., Safhi, F. A., Al Aboud, N. M., Aljabri, M., Alharbi, S. A., Ghazzawy, H. S., et al. (2024). Genetic diversity and combining ability of developed maize lines to realize heterotic and high yielding hybrids for arid conditions. *Phyton-International Journal of Experimental Botany*, 93(12), 3465–3485. <https://doi.org/10.32604/phyton.2024.058628>

Keshamalla, N., Prasad, K. R., Neelam, S., & Mallaiah, B. (2024). Assessment of genetic variability, heritability, and genetic advance in maize (*Zea mays* L.) segregating generations. *Journal of Experimental Agriculture International*, 46(8), 1004–1011. <https://doi.org/10.9734/jeai/2024/v46i82788>.

King, K., Ferela, A., Vyn, T. J., Trifunovic, S., Eudy, D., Hurburgh, C., Lamkey, K. R., & Archontoulis, S. V. (2024). Genetic gains in short-season corn hybrids: Grain yield, yield components, and grain quality traits. *Crop Science*, 64(2), 710–725. <https://doi.org/10.1002/csc2.21199>.

Korsa, F., Dessalegn, O., Zeleke, H., & Petros, Y. (2024). Genetic variability for the yield and yield-related traits in some maize (*Zea mays* L.) inbred lines in the Central Highland of Ethiopia. *International Journal of Agronomy*, 2024, Article 9721304. <https://doi.org/10.1155/2024/9721304>.

Kumar, A., Narayan, A., & Choudhary, V. K. (2024). Morphological characterization and diversity analysis of maize inbred lines (*Zea mays* L.). *Journal of Advances in Biology & Biotechnology*, 27(8), 1033–1042. <https://doi.org/10.9734/jabb/2024/v27i81224>.

Kumar, B., et al. (2023). Genetic diversity and population structure of maize inbred lines using phenotypic traits and single nucleotide polymorphism (SNP) markers. *Scientific Reports*, 13, 16989. <https://doi.org/10.1038/s41598-023-44961-3>.

Kumar, P., Longmei, N., Jat, B. S., Choudhary, M., Yathish, K. R., Bhushan, B., Goyal, M., & Rakshit, S. (2023). Heterotic grouping of Indian baby corn lines based on combining ability.

Li, C., Guan, H., Jing, X., Li, Y., Wang, B., Li, Y., Liu, X., Zhang, D., Liu, C., Xie, X., Zhao, H., Wang, Y., Liu, J., Zhang, P., Hu, G., Li, G., Li, S., Sun, D., Wang, X., Shi, Y., Song, Y., Jiao, C., Ross-Ibarra, J., Li, Y., Wang, T., & Wang, H. (2022). Genomic insights into historical improvement of heterotic groups during modern hybrid maize breeding. *Nature Plants*, 8, 750–763. <https://doi.org/10.1038/s41477-022-01190-2>.

Li, Z., Li, C., Zhang, R., Duan, M., Tian, H., Yi, H., Xu, L., Wang, F., Shi, Z., Wang, X., Wang, J., Su, A., Wang, S., Sun, X., Zhao, Y., Wang, S., Zhang, Y., Wang, Y., Song, W., & Zhao, J. (2023). Genomic analysis of a new heterotic maize group reveals key loci for pedigree breeding. *Frontiers in Plant Science*, 14, 1213675. <https://doi.org/10.3389/fpls.2023.1213675>.

Liu, K., & Muse, S. V. (2005). PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics*, 21(9), 2128–2129. <https://doi.org/10.1093/bioinformatics/bti282>

Manigben, K. A., Beyene, Y., Chaikam, V., Tongoona, P. B., Danquah, E. Y., Ifie, B. E., Aleri, I., Chavangi, A., Prasanna, B. M., & Gowda, M. (2024). Testcross performance and combining ability of intermediate maturing drought tolerant maize inbred lines in Sub-Saharan Africa. *Frontiers in Plant Science*, 15, 1471041. <https://doi.org/10.3389/fpls.2024.1471041>

Mebratu, A., Wegary, D., Teklewold, A., & Tarekegne, A. (2024). Testcross performance and combining ability of early-medium maturing quality protein maize inbred lines in Eastern and Southern Africa. *Scientific Reports*, 14, 9151. <https://doi.org/10.1038/s41598-024-58816-y>

Mehra, K., Chandel, U., Kumar, N., Goswami, I., & Godara, S. (2025). Assessment of genetic diversity and population structure analysis in elite maize (*Zea mays* L.) inbreds adapted to North Western Himalaya. *Plant Archives*, 25(1), 367–377. <https://doi.org/10.51470/PLANTARCHIVES.2025.v25.no.1.056>.

Mitsuhashi, S. (2024). Genetic diversity among maize (*Zea mays* L.) inbred lines adapted to Japanese climates. *PLOS ONE*, 19(1), e0297549. <https://doi.org/10.1371/journal.pone.0297549>

Mosa, H. E., Abd El-Latif, M. S., Kotp, M. S., Rizk, M. S., & Mohamed, H. A. A. (2025). Classification of maize inbred lines into heterotic groups using specific combining ability effects

for grain yield. *Egyptian Journal of Plant Breeding*, 29(1), 1–15.
<https://doi.org/10.12816/ejpb.2025.424188>.

Mosa, H. E., Abd El-Moula, M. A., Motawei, A. A., El-Gazzar, I. A. I., Abd El-Latif, M. S., Rizk, M. S., & El-Mouslhy, T. T. (2024). Classifying new maize inbred lines into heterotic groups using three methods. *Egyptian Journal of Plant Breeding*, 28(1), 135–154.
<https://doi.org/10.12816/ejpb.2024.366447>.

Mukri, G., Patil, M. S., Motagi, B. N., Bhat, J. S., Singh, C., Jeevan Kumar, S. P., Gadag, R. N., Gupta, N. C., & Simal-Gandara, J. (2022). Genetic variability, combining ability and molecular diversity-based parental line selection for heterosis breeding in field corn (*Zea mays* L.). *Molecular Biology Reports*, 49(6), 4517–4524. <https://doi.org/10.1007/s11033-022-07295-3>.

Mukri, G., Patil, M. S., Motagi, B. N., Bhat, J. S., Singh, C., Jeevan Kumar, S. P., Gadag, R. N., Gupta, N. C., & Simal-Gandara, J. (2022). Genetic variability, combining ability and molecular diversity-based parental line selection for heterosis breeding in field corn (*Zea mays* L.). *Molecular Biology Reports*, 49, 4517–4524. <https://doi.org/10.1007/s11033-022-07295-3>.

Peakall, R., & Smouse, P. E. (2012). GenAEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—An update. *Bioinformatics*, 28(19), 2537–2539.
<https://doi.org/10.1093/bioinformatics/bts460>

Perrier, X., & Jacquemoud-Collet, J.-P. (2006). *DARwin software*. CIRAD, Montpellier, France.
<https://darwin.cirad.fr/>

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.

Rajasekhar, D., Naveenkumar, K. L., Kalita, S., Tatiparthi, H., Rai, M., & Sen, D. (2025). Heterotic grouping of maize (*Zea mays* L.) inbreds derived from diverse indigenous landraces of the Northeast Hill Region of India. *Plant Genetic Resources*, 23(4), 252–261.
<https://doi.org/10.1017/S1479262124000704>

Rao, C. R. (1952). *Advanced Statistical Methods in Biometrical Research*. John Wiley & Sons, New York.

Reddy, P. R. K., Kumar, M. V. N., Selvarangam, V., Sunil, N., & Dhandapani, A. (2025). Identification of genetically diverse maize (*Zea mays* L.) inbred lines through cluster and principal

- component analysis. *International Journal of Research in Agronomy*, 8(4), 542–547. <https://doi.org/10.33545/2618060X.2025.v8.i4g.2828>.
- Sabitha, N., Mohan Reddy, D., Lokanatha Reddy, D., Sudhakar, P., & Ravindra Reddy, B. (2024). Association analysis over seasons among morphological, physiological and yield components with kernel yield in maize (*Zea mays* L.). *Journal of Advances in Biology & Biotechnology*, 27(5), 151–156. <https://doi.org/10.9734/jabb/2024/v27i5774>.
- Sathua, S. K., Shahi, J. P., Mahato, A., Gayatonde, V., & Kumar, P. (2018). Molecular diversity analysis of maize (*Zea mays* L.) inbreds using SSR markers. *Electronic Journal of Plant Breeding*, 9(3), 1122–1129. <https://doi.org/10.5958/0975-928X.2018.00140.0>.
- Sharanu, S. G., Patil, A., Kuchanur, P. H., Kisan, B., Yeri, S. B., Vinayan, M. T., & Zaidi, P. H. (2025). Genetic diversity of tropical maize (*Zea mays* L.) inbred lines using phenotypic clustering. *Journal of Advances in Biology & Biotechnology*, 28(10), 1466–1477. <https://doi.org/10.9734/jabb/2025/v28i103161>.
- Sharanu, S. G., Patil, A., Kuchanur, P. H., Kisan, B., Yeri, S. B., Vinayan, M. T., & Zaidi, P. H. (2025). Genetic diversity of tropical maize (*Zea mays* L.) inbred lines using phenotypic clustering. *Journal of Advances in Biology & Biotechnology*, 28(10), 1466–1477. <https://doi.org/10.9734/jabb/2025/v28i103161>.
- Sharanu, S. G., Patil, A., Kuchanur, P. H., Nair, S., Kisan, B., Yeri, S. B., Vinayan, M. T., & Zaidi, P. H. (2025). Assessing molecular diversity of tropical maize inbred lines using single nucleotide polymorphic (SNP) markers. *Journal of Experimental Agriculture International*, 47(10), 490–500. <https://doi.org/10.9734/jeai/2025/v47i103831>.
- Singh, D., Kushwaha, N., Swapnil, Parveen, R., Mohanty, T. A., Suman, S. K., Kumar, R., Kumar, A., & Singh, M. K. (2023). Assessment of Quality Protein Maize (QPM) inbreds for genetic diversity using morphological characters and simple sequence repeats markers. *Electronic Journal of Plant Breeding*, 14(3), 1276–1284. <https://doi.org/10.37992/2023.1403.145>.
- Sserumaga, J. P., Makumbi, D., Assanga, S. O., Mageto, E. K., Njeri, S. G., Jumbo, B. M., & Bruce, A. Y. (2020). Identification and diversity of tropical maize inbred lines with resistance to common rust (*Puccinia sorghi* Schwein). *Crop Science*, 60(6), 2971–2989. <https://doi.org/10.1002/csc2.20345>.

Stanley, A., Menkir, A., Agre, P., Ifie, B., Tongoona, P., Unachukwu, N., Meseka, S., Mengesha, W., & Gedil, M. (2020). Genetic diversity and population structure of maize inbred lines with varying levels of resistance to *Striga hermonthica* using agronomic trait-based and SNP markers. *Plants*, 9(9), 1223. <https://doi.org/10.3390/plants9091223>.

Tian, H., Yang, Y., Yi, H., Xu, L., He, H., Fan, Y., Wang, L., Ge, J., Liu, Y., Wang, F., & Zhao, J. (2021). New resources for genetic studies in maize (*Zea mays* L.): A genome-wide Maize6H-60K single nucleotide polymorphism array and its application. *The Plant Journal*, 105(4), 1113–1122. <https://doi.org/10.1111/tpj.15089>.

Xiang, Y., Xia, C., Li, L., Wei, R., Rong, T., Liu, H., & Lan, H. (2024). Genomic prediction of yield-related traits and genome-based establishment of heterotic pattern in maize hybrid breeding of Southwest China. *Frontiers in Plant Science*, 15, 1441555. <https://doi.org/10.3389/fpls.2024.1441555>.

Xue, Y., Zhao, Y., Zhang, Y., Wang, R., Li, X., Liu, Z., Wang, W., Zhu, S., Fan, Y., Xu, L., Zhao, W., Zhao, J., & Wang, F. (2025). Insights into the genomic divergence of maize heterotic groups in China. *Journal of Integrative Plant Biology*, 67(6), 1467–1486. <https://doi.org/10.1111/jipb.13884>

Zhang, H., Li, D., Lamkey, K. R., Stine, W. H., Yu, J., & Stine, H. H. (2024). New maize inbred development method I: Inter-heterotic group composite and tester choice. *Crop Science*, 64(4), 2170–2187. <https://doi.org/10.1002/csc2.21260>.

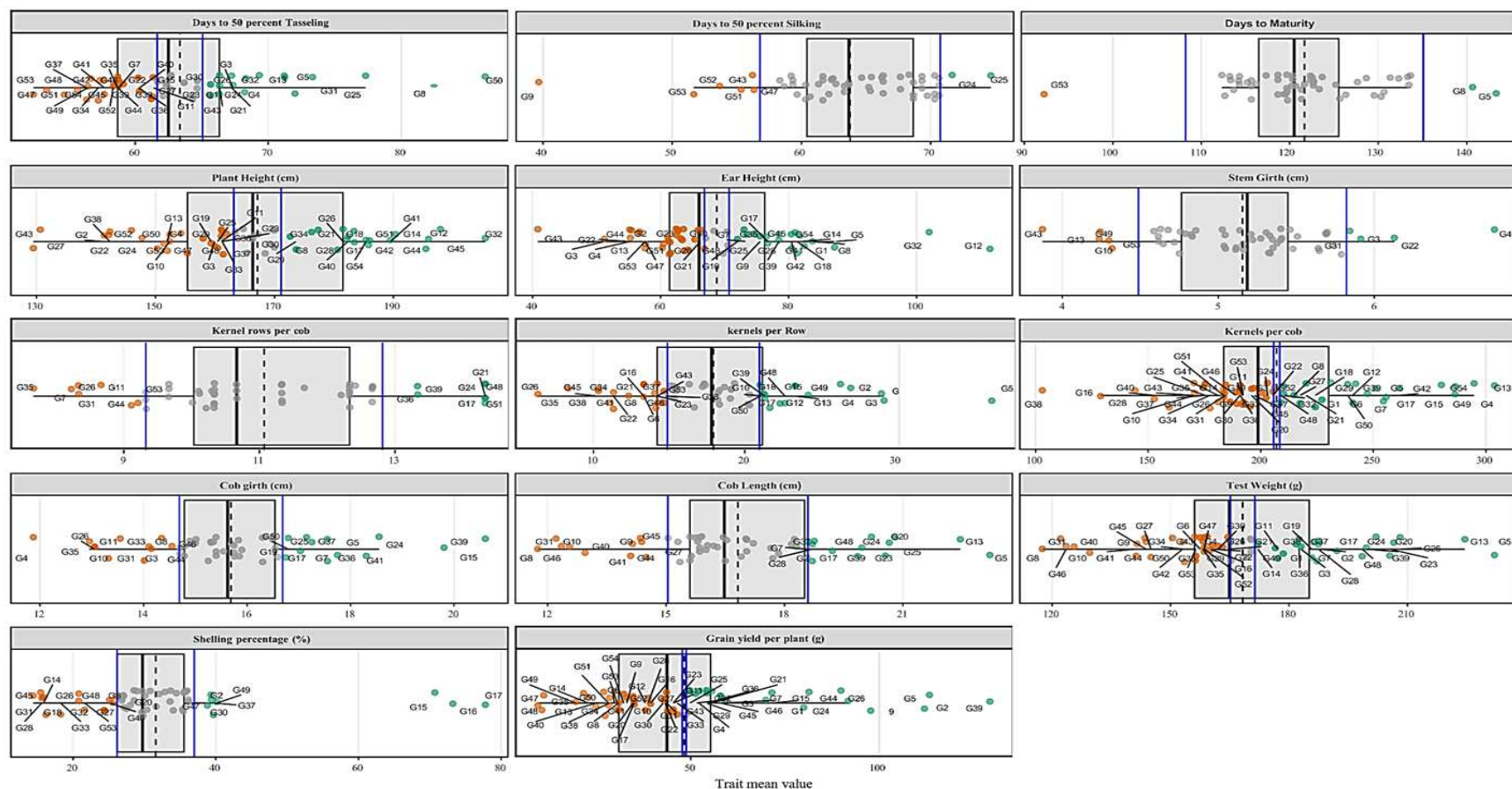


Figure.1 Trait-wise distribution of maize inbred lines for phenological, morphological, yield, and yield-contributing traits: Each panel represents the mean performance of individual genotypes for a specific trait. The solid black vertical line indicates the population mean, while the grey box represents the inter-quartile range and the whiskers denote the extent of variability. Blue vertical lines indicate statistical thresholds (mean $\pm$ critical difference) used to identify significant deviations from the mean

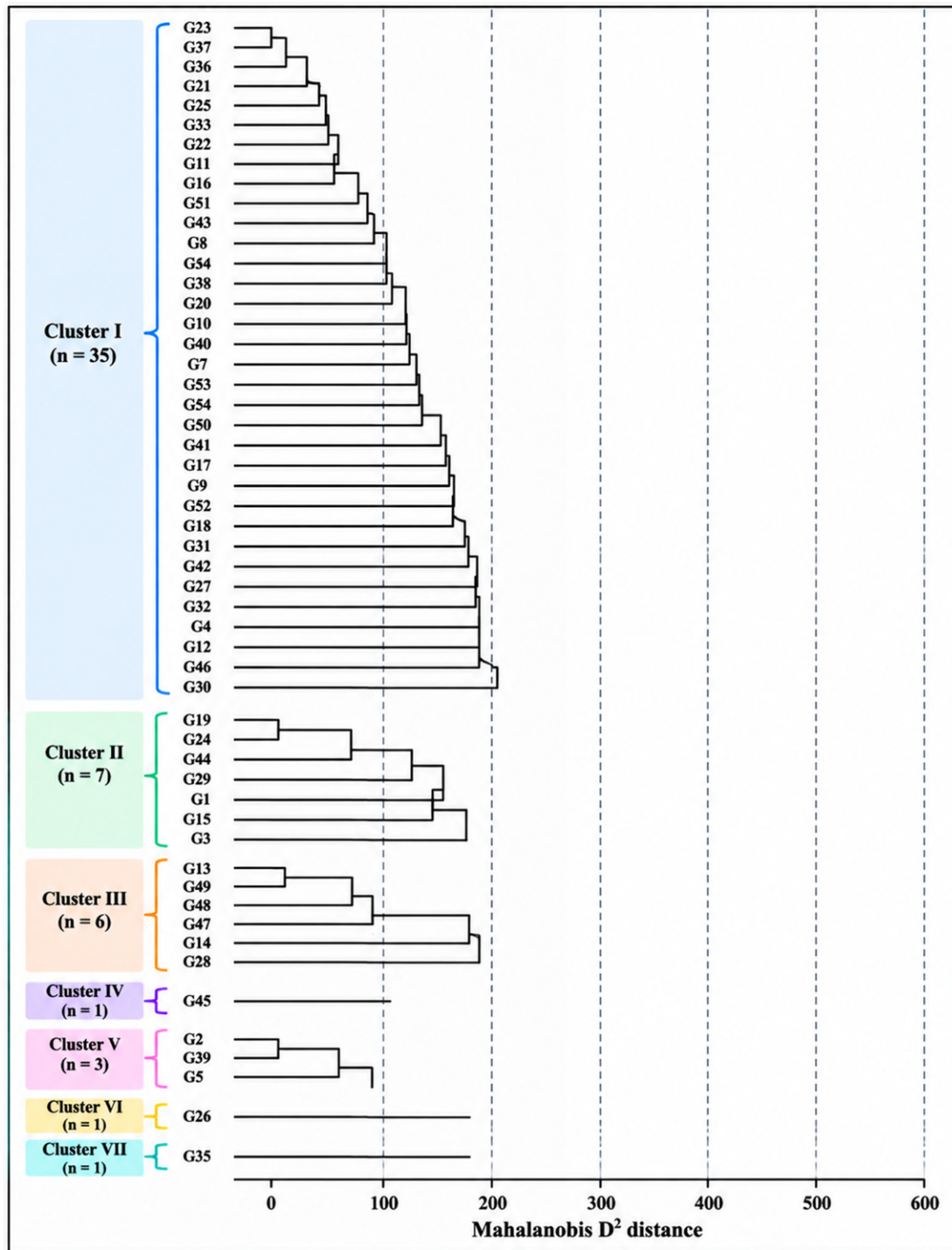


Figure 2. Mahalanobis D²-based Tocher clustering of maize genotypes: Tocher clustering of 54 maize genotypes based on Mahalanobis D² distance, resolving the genotypes into seven phenotypically differentiated clusters. The horizontal axis represents Mahalanobis D² distance, with increasing values indicating progressively greater multivariate phenotypic divergence among genotypes or clusters. The dendrogram depicts the hierarchical arrangement and relative separation of genotypes within the identified cluster

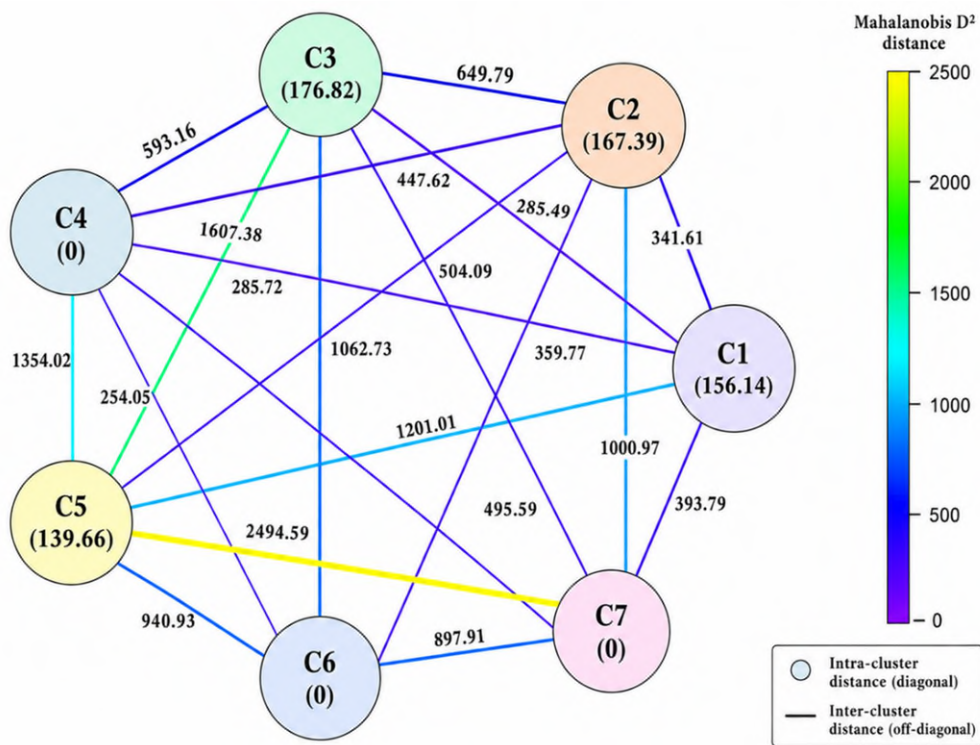


Figure 3. Intra- and inter-cluster Mahalanobis D^2 distances among maize genotypes: Network representation of intra- and inter-cluster Mahalanobis D^2 distances among the seven phenotypic clusters identified in the 54 maize genotypes. Values within the cluster nodes represent intra-cluster distances, whereas values connecting clusters represent inter-cluster distances. Line colour intensity corresponds to the magnitude of Mahalanobis D^2 distance, with higher values indicating greater multivariate phenotypic divergence.

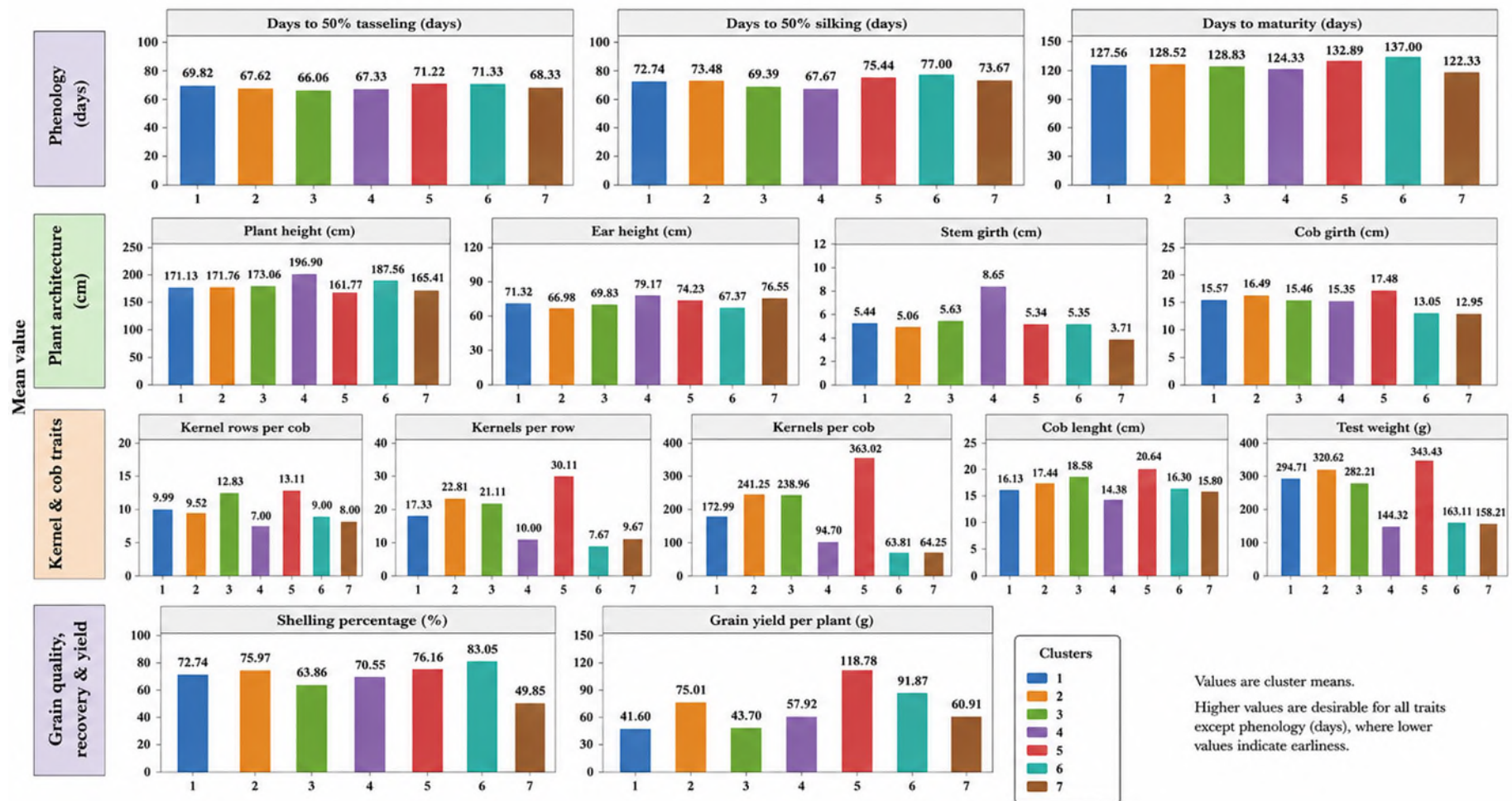


Figure 4. Cluster-wise mean performance of maize genotypes for phenological, agro-morphological, kernel, cob and yield-related traits: Cluster-wise mean values for the measured traits across seven phenotypic clusters. Each panel represents one trait, with bars showing the corresponding cluster mean. Phenological traits are expressed in days, while other traits are presented in their respective measurement units

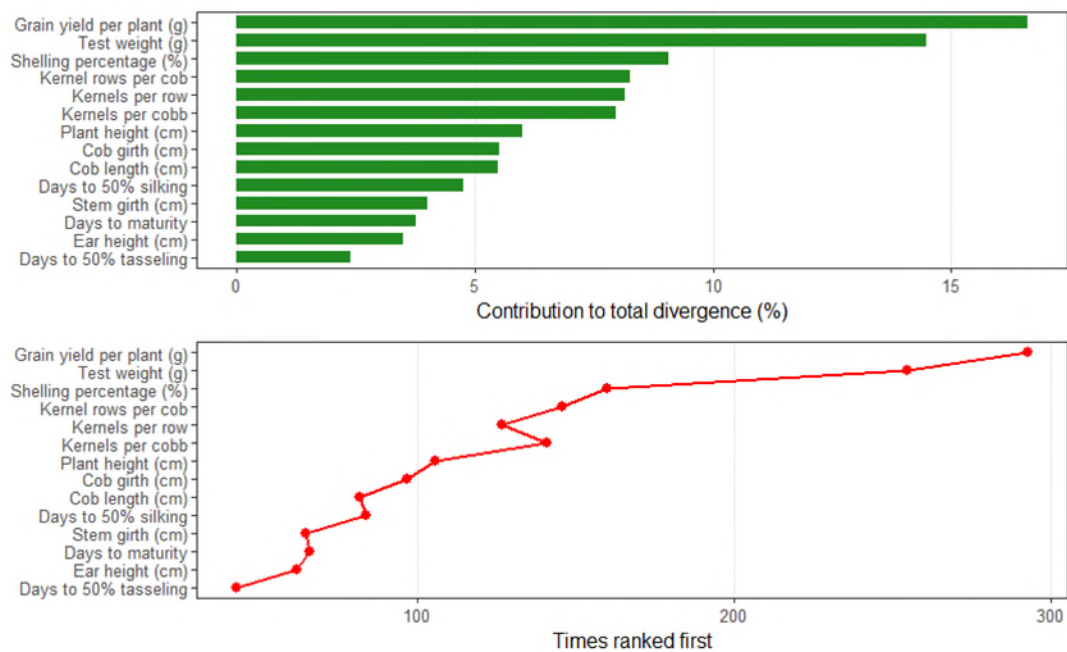


Figure 5. Relative contribution of agronomic traits to Mahalanobis D^2 phenotypic divergence among maize genotypes: Relative contribution (%) of individual traits to total Mahalanobis D^2 divergence (upper panel) and their frequency of occurrence as the first-ranked contributor (lower panel) among the 54 maize genotypes. Bars represent trait-wise percentage contributions, whereas points and connecting lines indicate the number of times each trait ranked first in pairwise divergence.

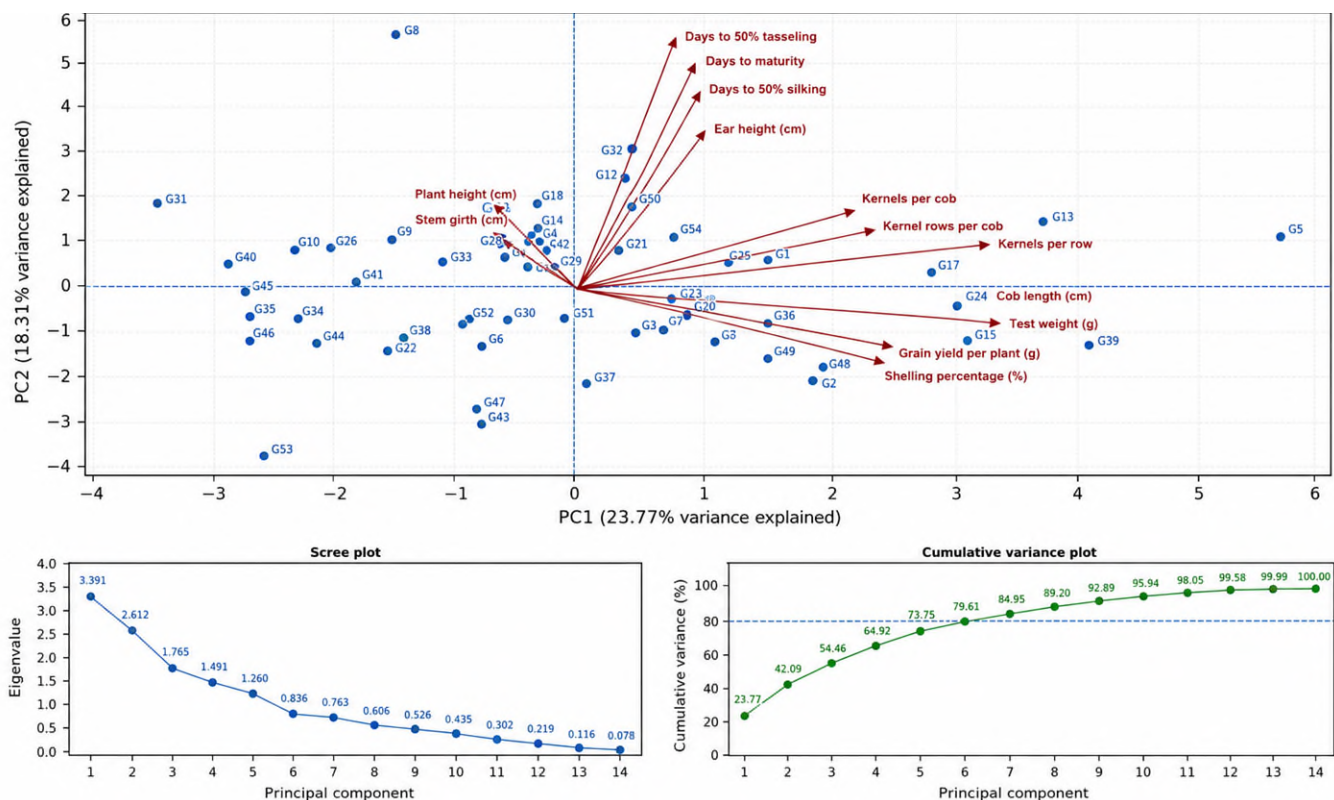


Figure 6. Principal component analysis of phenotypic trait-space architecture in maize: PCA visualization of 54 maize genotypes evaluated for 14 agronomic traits. The upper panel displays genotype scores on PC1 and PC2, with arrows representing trait loading vectors; arrow direction indicates the association of each trait with the principal components, while arrow length reflects its contribution to the displayed trait space. The lower-left panel shows the eigenvalue profile across the 14 principal components (scree plot), and the lower-right panel shows the cumulative proportion of phenotypic variance explained by successive components. Dashed reference lines indicate the zero axes in the biplot and the 80% cumulative-variance level in the lower-right panel.

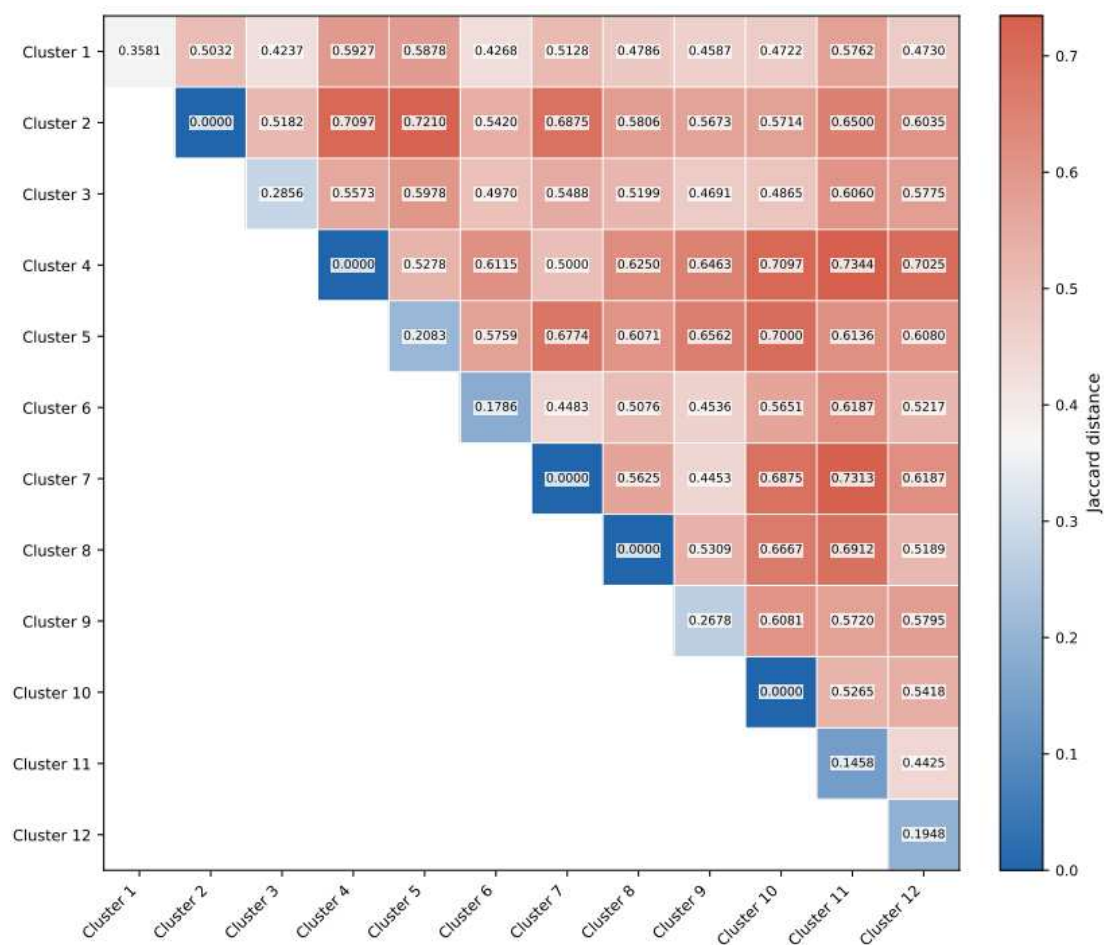


Figure 7. Intra- and inter-cluster genetic distance among maize genotypes based on Jaccard's coefficient: Upper-triangular matrix showing intra- and inter-cluster Jaccard distances among the 12 SSR-based molecular clusters. Values on the diagonal represent within-cluster distances, whereas off-diagonal values represent distances between clusters. Colour intensity represents the magnitude of genetic distance, with lower values indicating closer molecular relationships and higher values indicating greater molecular divergence.

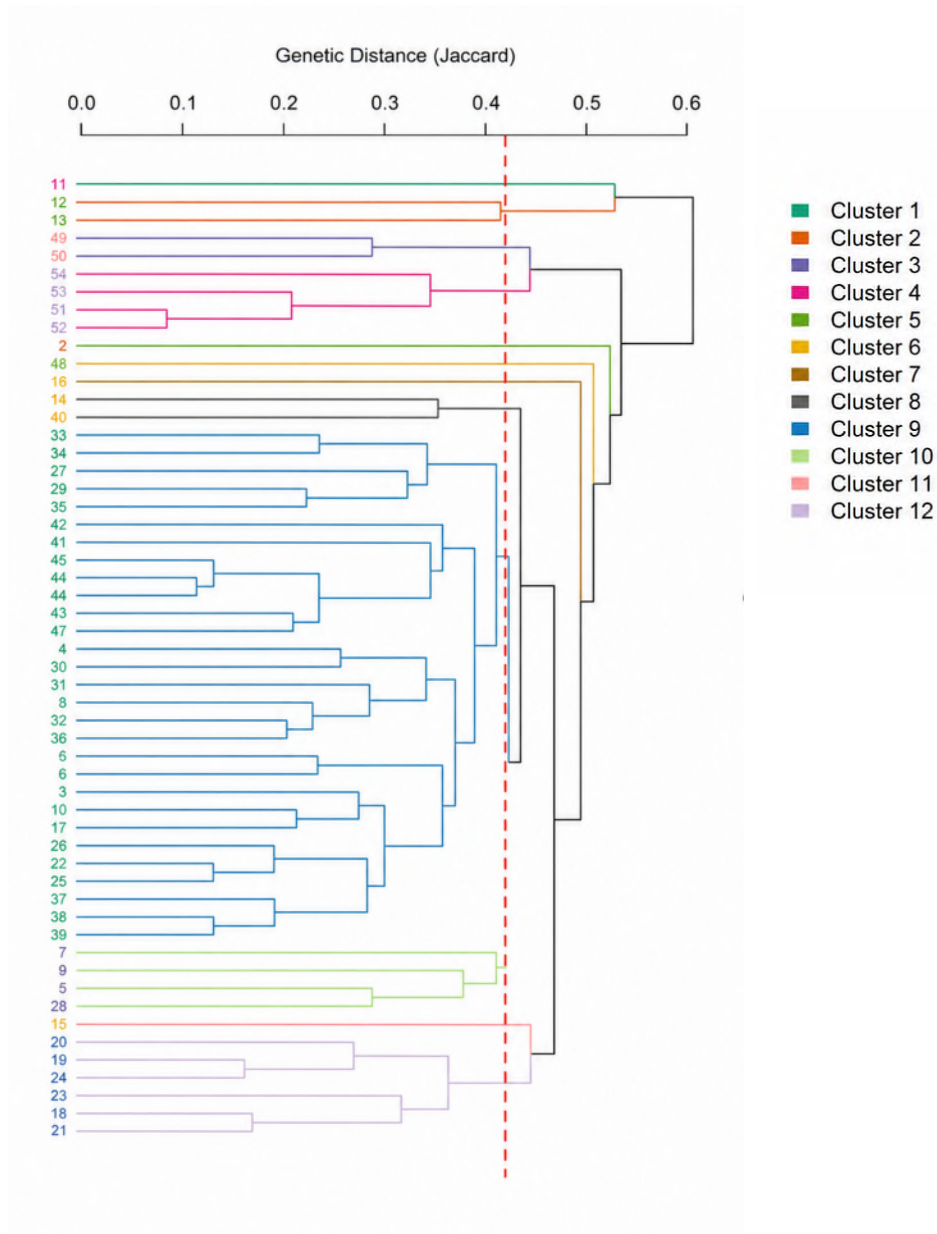


Figure 8. UPGMA clustering of maize genotypes based on SSR-derived Jaccard genetic distance: UPGMA dendrogram showing the genetic relationships among 54 maize genotypes based on Jaccard genetic distance calculated from SSR marker profiles. The horizontal axis represents genetic distance, with larger values indicating greater genetic dissimilarity. The red dashed line indicates the selected clustering threshold (~ 0.42), while branch colours distinguish the 12 molecular clusters identified at this threshold. Genotype numbers at the branch tips correspond to the entries used in the SSR analysis.

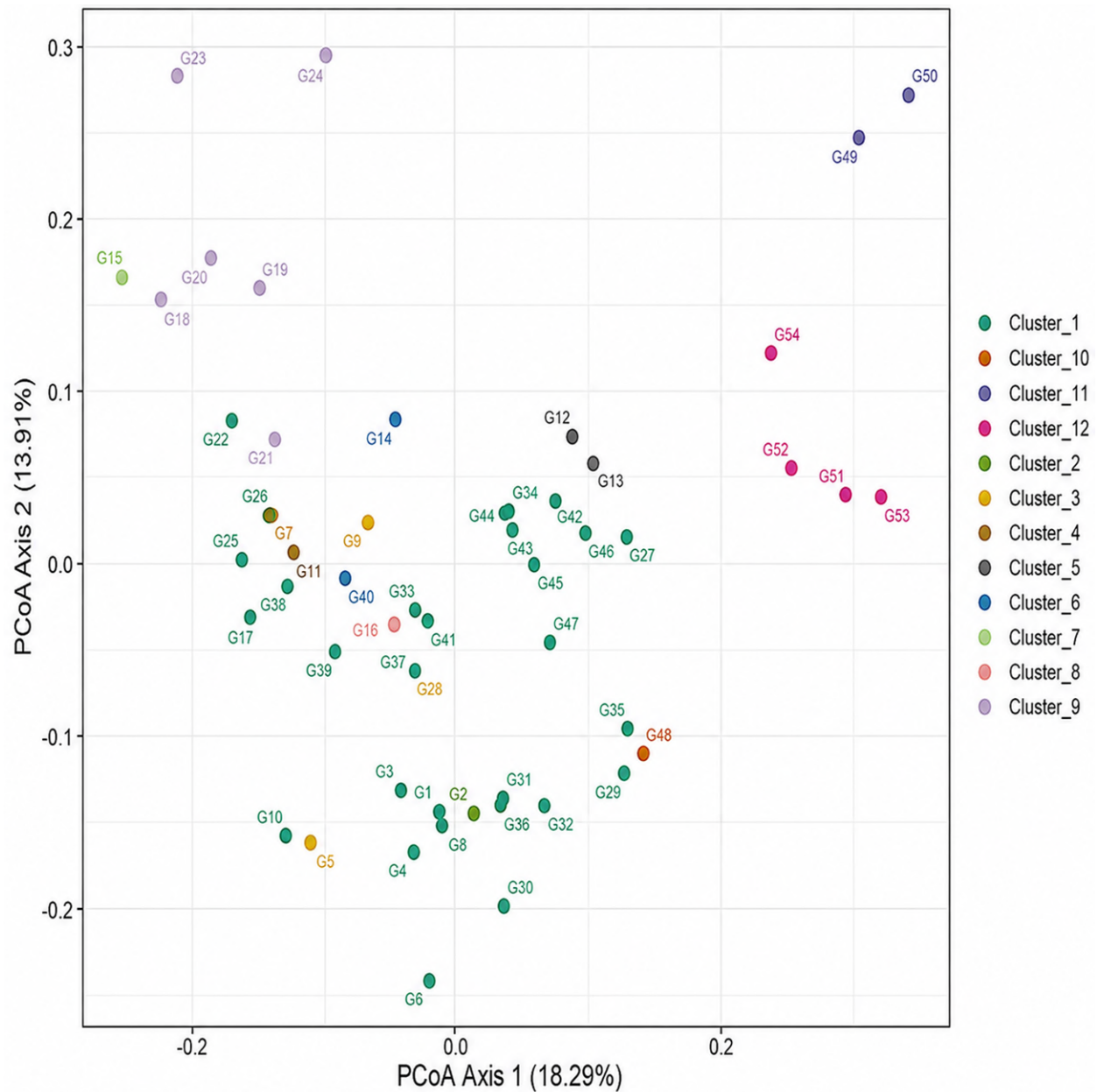


Figure 9. Principal coordinate analysis of SSR-based molecular relationships among maize genotypes: Principal Coordinate Analysis (PCoA) of 54 maize genotypes based on SSR marker profiles. Each point represents an individual genotype, with colours indicating its assignment to the corresponding molecular cluster. The horizontal and vertical axes represent PCoA1 and PCoA2, explaining total molecular variation, respectively. Genotype positions reflect their relative molecular relationships, with greater separation in the ordination space indicating greater differentiation in SSR profiles.

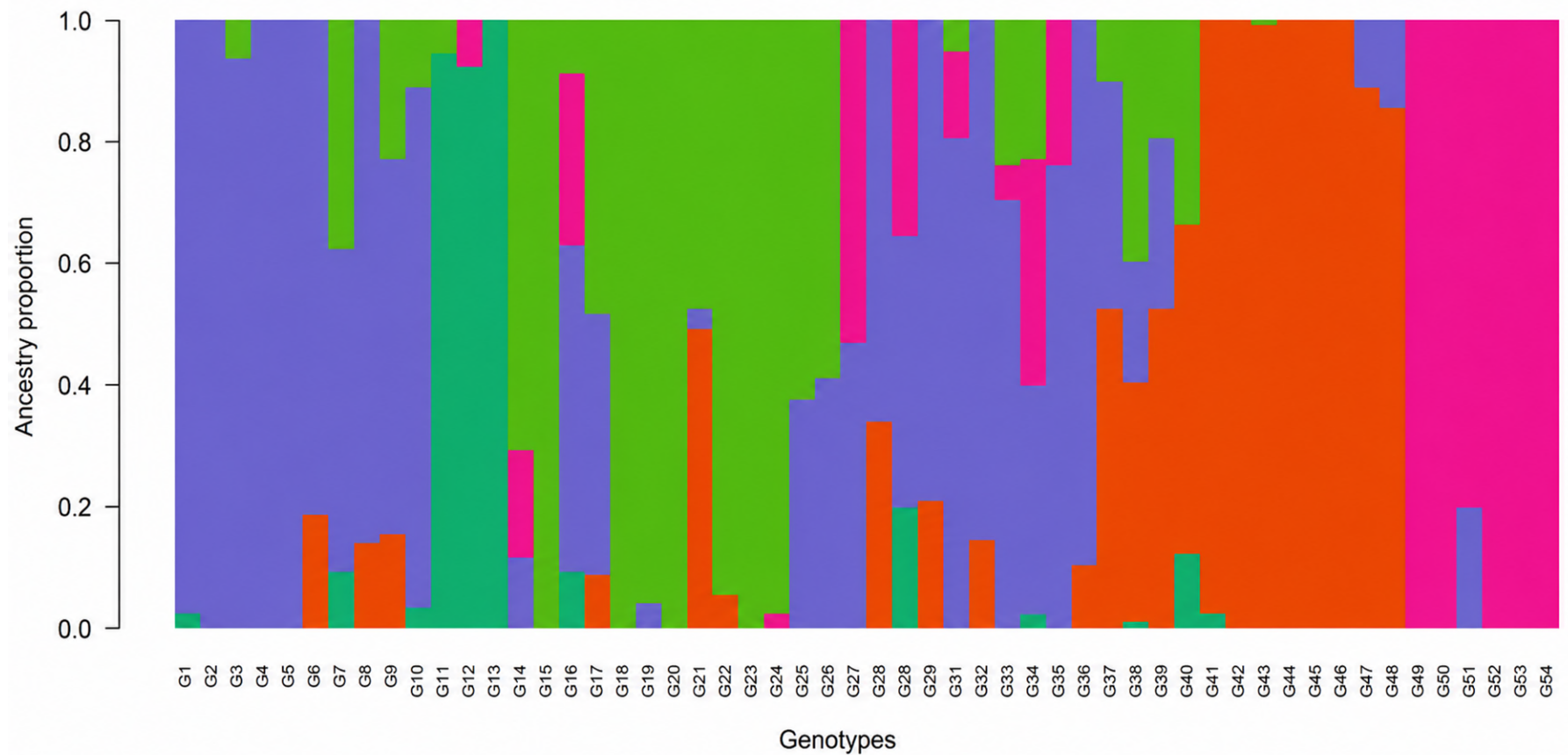


Figure 10. Bayesian population structure of maize genotypes based on SSR markers: Stacked bar plot showing the estimated ancestry proportions of the 54 maize genotypes at $K = 5$. Each vertical bar represents one genotype, and the colored segments within each bar indicate the relative proportion of membership in the five inferred genetic components. The x-axis denotes genotype identity, while the y-axis represents the estimated ancestry proportion ranging from 0 to 1.

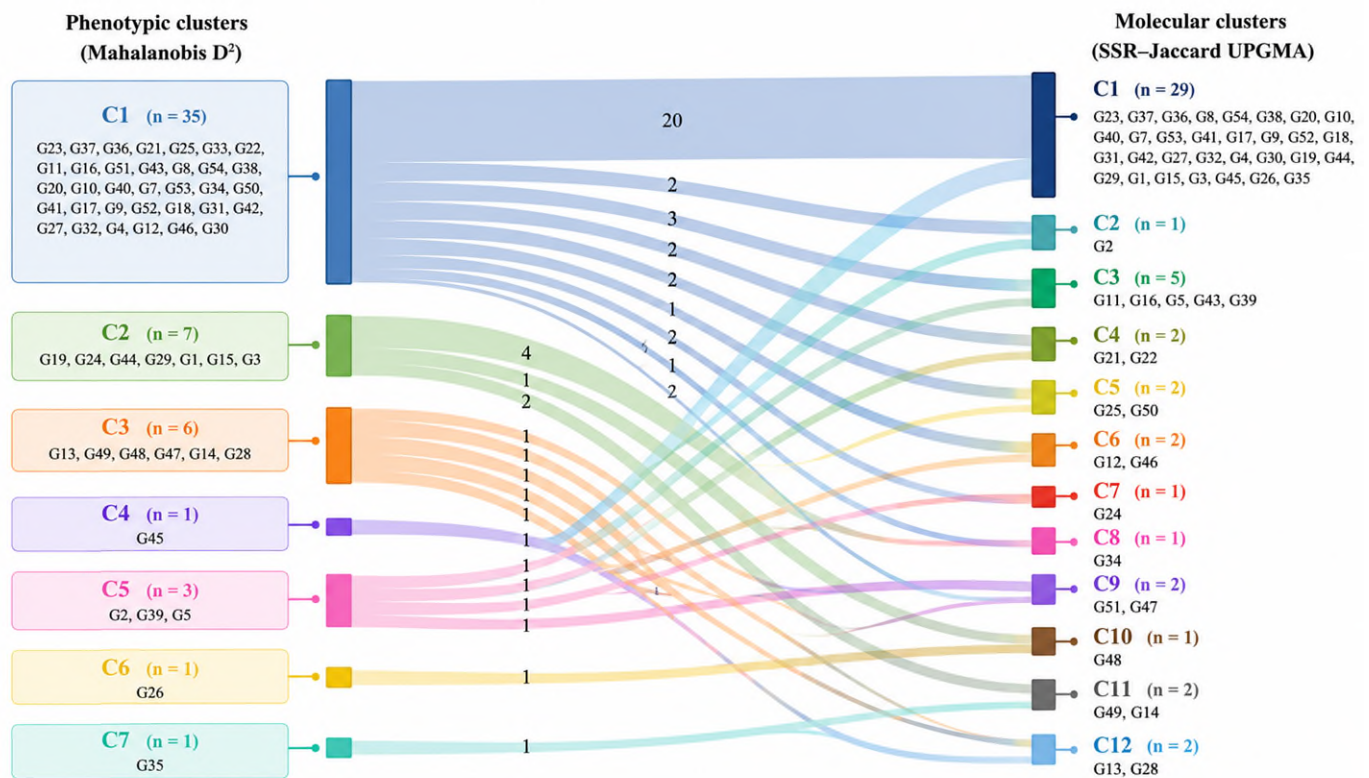


Figure 11. Correspondence between phenotypic and molecular clustering of 54 maize genotypes: Correspondence between Mahalanobis D²-based phenotypic clusters and SSR–Jaccard UPGMA molecular clusters of 54 maize genotypes. The left-hand boxes represent the seven phenotypic clusters, with cluster size indicated by the number of genotypes (n). The right-hand boxes represent the 12 molecular clusters obtained from SSR-based UPGMA analysis, also labelled by cluster size (n). Connecting bands trace the genotypes shared between each phenotypic and molecular cluster; band width is proportional to the number of shared genotypes. Thus, broad bands indicate greater overlap between the corresponding classifications, whereas narrow or multiple bands indicate that a phenotypic group is distributed across several molecular clusters. The figure provides a direct visual comparison of the two independent classifications and identifies concordant as well as differently resolved genotype groupings.

Table 1. Origin and designation of the 54 maize inbred lines used in the study

S.no	Genotype Code	Inbred Name	Origin	Source
1	G1	CML- 141	CIMMYT	MBRSS
2	G2	IB-141	Local	MBRSS
3	G3	CML- 158	CIMMYT	MBRSS
4	G4	CML- 150	CIMMYT	MBRSS
5	G5	CML-367	CIMMYT	MBRSS
6	G6	CML-77	CIMMYT	MBRSS
7	G7	CML-177	CIMMYT	MBRSS
8	G8	CML-258	CIMMYT	MBRSS
9	G9	IB-6128	Local	MBRSS
10	G10	PMS-14	MBRSS	MBRSS
11	G11	IB-3009	Local	MBRSS
12	G12	PMS-75	MBRSS	MBRSS
13	G13	PMS-99	MBRSS	MBRSS
14	G14	M-4-1-3-1	Local	MBRSS
15	G15	DMR-649	IIMR	MBRSS
16	G16	IB-104	local	MBRSS
17	G17	CML-76	CIMMYT	MBRSS
18	G18	IB-6066	Local	MBRSS
19	G19	CML-142	CIMMYT	MBRSS
20	G20	CML-149	CIMMYT	MBRSS
21	G21	CML-52	CIMMYT	MBRSS
22	G22	CML-324	CIMMYT	MBRSS
23	G23	IB-6118	Local	MBRSS
24	G24	PMS-116	Local	MBRSS
25	G25	CML-269	CIMMYT	MBRSS
26	G26	CML-150	CIMMYT	MBRSS
27	G27	IB-6047	Local	MBRSS
28	G28	CML-447	CIMMYT	MBRSS
29	G29	CML-2157	CIMMYT	MBRSS
30	G30	CML-150	CIMMYT	MBRSS
31	G31	CML-258	CIMMYT	MBRSS
32	G32	CML-447	CIMMYT	MBRSS
33	G33	IB-6074	Local	MBRSS
34	G34	NP 38-1-1-1	Local	MBRSS
35	G35	IB-3012	Local	MBRSS
36	G36	NP-44-1-1-1	Local	MBRSS
37	G37	IB-8018	Local	MBRSS
38	G38	IB 50164	Local	MBRSS
39	G39	IB-3012	Local	MBRSS
40	G40	IB-4007	Local	MBRSS
41	G41	IB-8009	Local	MBRSS
42	G42	IB-8033	Local	MBRSS
43	G43	VL Makka	VPKAS	MBRSS
44	G44	M8	Local	MBRSS
45	G45	IB-17	Local	MBRSS
46	G46	R2-3 (RM)	Local	MBRSS
47	G47	IB-2040	Local	MBRSS
48	G48	IB-7841	Local	MBRSS
49	G49	PMS-100	MBRSS	MBRSS
50	G50	PMS-116	MBRSS	MBRSS
51	G51	IB-6120	Local	MBRSS
52	G52	IB-6118	Local	MBRSS
53	G53	IB-3009	Local	MBRSS
54	G54	Black Maize	Local	MBRSS

Table 2. SSR markers used for molecular characterization of the 54 maize inbred lines

SSR marker	Avg. T _m (°C)	Chromosome	Bin / map position	Annealing temp. (°C)
<i>phi083</i>	68.4	2	2.04	52
<i>phi96100</i>	64.3	2	2.00	54
<i>phi093</i>	77.1	4	4.08	60
<i>phi299852</i>	65.0	6	6.08	58
<i>phi112</i>	69.3	7	7.01	56
<i>phi114</i>	70.4	7	7.02	60
<i>umc1304</i>	66.3	8	8.02	54
<i>umc1161</i>	66.9	8	8.06	56
<i>phi065</i>	67.0	9	9.03	54
<i>phi084</i>	72.5	10	10.04	54
<i>bnlg1890</i>	67.0	4	4.08	55
<i>nc130</i>	64.4	5	5.00	54
<i>phi069</i>	71.7	7	7.05	60
<i>nc013</i>	71.0	5	5.00	54
<i>umc1354</i>	64.3	1	1.00	55
<i>umc1685</i>	68.9	1	1.01	56
<i>umc2226</i>	60.8	1	1.01	54
<i>bnlg371</i>	71.7	2	2.05	55
<i>bnlg1014</i>	69.6	1	1.01	52
<i>phi090</i>	70.9	2	2.08	60

Note: Avg. T_m = average melting temperature of the primer pair; SSR = simple sequence repeat; bin/map position refers to the maize genetic map position. Annealing temperatures represent the temperatures used for PCR amplification of the respective markers.

Table.3 Analysis of variance revealing variation among 54 maize genotypes for phenological, morphological and yield-related traits

Source of Variation	Degrees of Freedom	Days to 50 percent Tasseling	Days to 50 percent Silking	Days to Maturity	Plant Height (cm)	Ear Height (cm)	Stem Girth (cm)	Kernel Rows per Cob	Kernels per Row	Kernels per Cob	Cob Girth (cm)	Cob Length (cm)	Test Weight (g)	Shelling Percentage (%)	Grain Yield per Plant (g)
Replication	2	1.167	21.74	55.93	3.94	0.59	0.39	2.29	7.46	0.33	1.83	4.90	6.67	45.62	0.092
Genotypes	53	135.40**	109.86**	197.486**	862.06**	440.576**	0.842**	9.140**	93.103**	18197.99**	8.131***	18.911**	140.443**	457.793**	2052.254**
Error	106	1.11	10.53	8.67	6.02	1.41	0.16	1.15	3.52	0.69	0.37	1.20	3.65	11.09	0.10

* significant at $p < .05$; ** significant at $p < .01$.

Table 4. SSR marker polymorphism and genetic diversity statistics across the maize genotypes

Marker	NPB	MAF	nE	h	I	PI
<i>bngl180</i>	4	0.38	3.41	0.70	1.12	0.70
<i>phi96100</i>	3	0.68	1.95	0.48	0.65	0.48
<i>umc1685</i>	2	0.74	1.79	0.44	0.44	0.44
<i>phi299852</i>	3	0.70	1.76	0.43	0.75	0.43
<i>bnlg371</i>	2	0.74	1.62	0.38	0.57	0.38
<i>phi114</i>	5	0.68	1.53	0.34	1.31	0.34
<i>bnlg1014</i>	3	0.81	1.47	0.32	0.58	0.32
<i>phi069</i>	2	0.83	1.38	0.27	0.45	0.27
<i>umc1354</i>	1	0.85	1.37	0.27	0.13	0.27
<i>phi083</i>	3	0.75	1.34	0.25	0.69	0.25
<i>umc2226</i>	3	0.85	1.33	0.24	0.62	0.24
<i>phi065</i>	3	0.83	1.22	0.18	0.76	0.18
<i>phi090</i>	1	0.90	1.21	0.17	0.08	0.17
<i>nc130</i>	3	0.90	1.19	0.16	0.49	0.16
<i>umc1161</i>	3	0.90	1.18	0.15	0.55	0.15
<i>phi084</i>	3	0.92	1.16	0.13	0.30	0.13
<i>phi112</i>	3	0.94	1.10	0.09	0.43	0.09
<i>umc1304</i>	3	0.94	1.09	0.09	0.49	0.09
<i>nc013</i>	2	0.98	1.03	0.03	0.17	0.03
<i>phi093</i>	3	0.94	1.02	0.024	0.66	0.02

NPB: Number of polymorphic bands; **MAF:** Major allele frequency; **NE:** Effective number of alleles; Nei's gene diversity; **I:** Shannon's information index; **PI:** Polymorphic information content

Table 5. Elite maize inbreds identified as breeding-relevant parental sources through integrated morphological, phenological and molecular diversity

Breeding target	Priority inbred(s)	Key phenotypic advantage	Molecular distinctness	Parent-prioritization value
Grain yield & yield components	G2, G5, G39	High grain yield and favourable kernel/cob traits	G2–CII; G5–CIII; G39–CI	Priority parents for coordinated yield improvement
Kernel number & cob architecture	G13, G14, G28, G47, G48, G49	Superior kernel rows, kernels/row, kernels/cob and cob length	Distributed across CI, CIII, CV, CVI, CX and CXI	Complementary parents for kernel-number and cob improvement
Grain recovery & shelling	G26	High grain yield (91.87 g) and highest shelling (83.05%)	CI	Priority parent for grain recovery and shelling
Yield + kernel number + shelling	G15	High yield, kernels/row, kernels/cob and shelling	CVII	Distinct parent combining yield-component advantages
Earliness	G35	Earliest tasseling, silking and maturity	CI	Priority parent for earliness
Early maturity + stem strength	G45	Earlier maturity with high stem girth	CI	Parent for combining earliness and stem girth
Yield + kernel/cob traits + molecular distinctness	G2	Superior grain yield, kernel/cob traits and test weight	CII (singleton)	Strongest integrated candidate for parent prioritization
Kernel/cob traits + molecular distinctness	G48, G49	Favourable kernel and cob traits	G48–CX; G49–CXI	Parents combining favourable phenotype with differentiated backgrounds
Molecular diversity within phenotypic core	G11, G16	Intermediate agronomic profiles	G11–CIV; G16–CVIII; both singleton clusters	Molecular-diversity sources for broadening the major phenotypic group

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SUPLIMENTARYFIGURES.docx](#)