Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Unravelling Genetic Diversity and Population Structure in Cowpea [Vigna unguiculata (L.) Walp.] Germplasm Using SSR Markers

Anshu Priya
Sam Higginbottom Institute of Agriculture
Dhammaprakash Pandhari Wankhede
d.wankhede-icar@nic.in

ICAR-National Bureau Of Plant Genetic Resources
Jenny Priya Ekka
Sam Higginbottom Institute of Agriculture
Vaidurya Pratap Sahi
Sam Higginbottom Institute of Agriculture
Dixit Tagra
Sam Higginbottom Institute of Agriculture
Neha Thomas
Sam Higginbottom Institute of Agriculture
Mukesh Kumar
Bihar Agricultural University

Research Article

Keywords: Cowpea, Genetic diversity, Germplasm, Population structure, SSR markers, UPGMA, PCoA

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11459093/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

ORIGINAL ARTICLE

Unravelling Genetic Diversity and Population Structure in Cowpea [*Vigna unguiculata* (L.) Walp.] Germplasm Using SSR Markers

Anshu Priya[1], Dhammaprakash Pandhari Wankhede[2]✉, Jenny Priya Ekka[1], Vaidurya Pratap Sahi[1]✉, Dixit Tagra[1], Neha Thomas[1], Mukesh Kumar[3]

[1]Sam Higginbottom University of Agriculture, Technology and Sciences, Prayagraj, Uttar Pradesh, India – 211007

[2]National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi, India -110012

[3]Bihar Agricultural University, Sabour, Bhagalpur, India - 813210

✉**Corresponding Author:** vpsahi@shiats.edu.in and d.wankhede-icar@nic.in.

Abstract

Cowpea is an important leguminous crop known for its nutritional value and contribution to maintaining soil fertility. Therefore, it is an important crop for ensuring food security in tropical and subtropical agricultural systems. However, the narrow genetic basis of the crop limits further recombination and breeding and makes it more susceptible to evolving biotic and abiotic stresses. This study was conducted to assess the genetic diversity and population structure among 23 cowpea accessions through 27 SSR loci, out of which 26 were polymorphic. In total, the markers resulted in amplification of 89 alleles with mean number of 3.33 alleles per locus ranging from 2 to 6. The observed heterozygosity was 0.101 as against the expected heterozygosity of 0.545, while the Shannon information index value of 0.944 revealed the existence of considerable amount of genetic variation within the material. The average polymorphism information content value was 0.489 varying from 0.282 (VuSSR2) to 0.681 (CEDG143). UPGMA clustering resulted in grouping of accessions in to four clusters that were supported by principal coordinate analysis. The results revealed considerable genetic variation among the accessions which can be used in future breeding programs to select diverse parents for hybridization. This research provides baseline information about molecular markers in cowpea germplasm resources that can assist in further association mapping and marker assisted breeding for improvement of this crucial legume crop. This will also enable effective introduction of different alleles in future breeding programs.

Keywords: Cowpea, Genetic diversity, Germplasm, Population structure, SSR markers, UPGMA, PCoA

INTRODUCTION

Cowpea [*Vigna unguiculata* (L.) Walp.] is a leguminous crop cultivated across several countries, particularly in tropical and subtropical environments, where it is grown as a grain legume, vegetable, fodder and green manure in smallholder production systems. In general, it ranks among the five most widely cultivated pulses or legumes across the globe [1]. Nevertheless, cowpea productivity in India remains poor compared with other pulses, with average yields standing at around 350 kg ha^{-1} [2]. This productivity gap is all the more significant because cowpea is broadly used as an appetizing and highly nutritious food source. Seeds are reported to contain protein at 23–32% [3], carbohydrates at 50–60% [4] and fat at about 1% [5]; moreover, they are packed with several micronutrients, antioxidants, polyphenols, polyunsaturated fatty acids and fibre [2]. Because of this high-quality protein content and nutritive value, cowpea is popularly known as "poor man's meat" [6] and its regular consumption helps combat malnutrition, particularly among low-income and resource-poor communities.

Cowpea is also of great importance to farming communities beyond its nutritional role, because it contributes to soil fertility by fixing atmospheric nitrogen, thereby benefiting subsequent cereals grown in rotation [7]. This biological nitrogen fixation, together with cowpea's adaptability to moisture-limited environments, makes it a key component of sustainable cropping systems in the semi-arid tropics, where it buffers both food and income security.

Nonetheless, genetic diversity in cowpea is not very high; such low diversity can largely be explained by its predominantly self-pollinated nature, which restricts recombination and gene flow among individuals [8, 9]. A narrow genetic base renders the crop vulnerable to emerging biotic and abiotic stresses and limits the scope for genetic gain through hybridization. In view of the increasingly adverse impact of climate change and the exponential rise in the population of the world, there is an urgent need for scientific studies on the genetic diversity of cowpea [10] because genetic diversity plays a significant role in the context of breeding programs aimed at yield improvement and nutrition enhancement and increased resistance to stress factors. Higher genetic diversity offers more opportunities for discovering genes that could make plants resilient and productive, and knowledge of the genetic makeup of cowpea helps to conserve the crop and devise sound breeding strategies.

Different DNA-based molecular techniques have been used to assess the genetic diversity of cowpea, including RAPD [6], AFLP [11], RFLP [12], ISSR [13], and SSR [14, 15]. Among these techniques, SSRs (simple sequence repeats, or microsatellite markers) are widely

favoured because of their high polymorphic nature, co-dominant inheritance, multi-allelic constitution, genome-wide distribution, reproducibility, and ease of application, while requiring only small amounts of template DNA. SSR-based research on cowpea therefore ensures a more accurate assessment of genetic diversity at the molecular scale than conventional phenotypic approaches, which are influenced by environmental effects and developmental stage. SSRs have likewise proved effective in revealing diversity in the wild relatives of cowpea [14], which are of particular interest as potential sources of novel alleles for broadening the cultivated gene pool.

Equally important to the overall magnitude of diversity is population structure which is the partitioning of germplasm into genetically distinct groups and the extent of admixture among them. Characterizing the genetic relationships among accessions guides the selection of genetically diverse parents for hybridization, helps to reduce redundancy and assemble core collections for efficient evaluation, and is a prerequisite for reliable marker–trait association studies. The reliability of such estimates depends on both the degree of relatedness of the samples and the informative quality of the markers used [16]. Together, assessments of genetic diversity and population structure thus help to identify and harness the inherent genetic variation in the crop, aiding the development of superior cowpea varieties capable of meeting increased nutritional requirements while coping with climatic adversities.

Against this background, the present study was undertaken to investigate the extent of genetic diversity and population structure among 23 cowpea accessions with the aim of identifying genetically distinct genotypes that can serve as parents for future cowpea improvement programmes.

MATERIALS AND METHODS

Plant material and DNA extraction

In the present study, 23 cowpea accessions were used. The collected accessions, along with their sources, are given in **Table 1**. The molecular experiment for genetic diversity analysis using the SSR marker was carried out in the laboratory at ICAR-NBPGR in New Delhi. For the extraction of DNA, the seeds of all the accessions were grown in separate pots under natural environmental conditions. The 3 g freshly harvested leaf sample was collected from each accession separately after 14 days for DNA extraction using the Cetyltrimethylammonium bromide method (CTAB) widely used for isolating plant DNA [17]. The quantity of the extracted DNA was measured with a spectrophotometer (ScanDrop2). Further, the purity of the DNA was evaluated using the absorbance ratio at 260 nm and 280 nm.

PCR amplification with SSR markers

The extracted DNA from 23 cowpea accessions was amplified using 27 SSR primers. The list of primers along with their annealing temperatures is listed in **Table 2**. The amplification reactions were carried out in a Bio-Rad C1000 thermal cycler. Each 20 µl reaction contained 4 µl of template DNA, 2 µl of 10X PCR buffer, 0.5 µl of dNTP mix, 0.5 µl of each forward and reverse primer, 0.3 µl of Taq DNA polymerase, and 12.2 µl of nuclease-free water. The thermal cycling program was as follows: initial denaturation at 95 °C for 5 minutes; 36 cycles of denaturation at 95 °C for 30 seconds, annealing for 30 seconds at a temperature between 51 and 60 °C depending on the primer pair, and extension at 72 °C for 60 seconds; and a final extension at 72 °C for 5 minutes.

Gel electrophoresis

The amplified PCR products were separated using gel electrophoresis method. The gel was prepared using 3% MetaPhor agrose into 1X TAE buffer with gentle mixing. The mixture was stored at 4 °C for 2–3 hours and then microwaved to avoid foaming and boiling of the gel. Ethidium bromide was added to the gel before casting. The gel was left to harden, then cooled in the fridge for about 20 minutes. Loading dye was mixed with 20 µl of the PCR product, and 10 µl of this mixture was used for loading in the gel alongside the 100 bp DNA ladder. Electrophoresis was carried out at 90 volts for 3-4 hours in 1X TAE buffer. After separation, the DNA bands were visualized under a UV transilluminator and recorded with a gel documentation system.

Band scoring and Data analysis

The bands were scored by their molecular weight using AlphaView software. The diversity obtained was then computed using GenAlEx version 6.5 [18], following the approach of Pagnotta. The parameters included the number of alleles per locus, the effective number of alleles, Shannon's information index, observed heterozygosity, expected heterozygosity, unbiased expected heterozygosity, and the fixation index. The polymorphism information content of each locus was calculated using the formula $PIC = 1 - \Sigma\, p_i^2 - \Sigma\, \Sigma\, 2\, p_i^2\, p_j^2$ ($i = 1 \ldots n$; $j = i+1 \ldots n$) where p_i and p_j are the frequencies of the ith and jth alleles at a locus and n is the number of alleles at that locus [19]. To see how the accessions were distributed across the populations, a principal coordinate analysis was carried out in GenAlEx version 6.5. In addition, a distance matrix was used to construct a dendrogram with the UPGMA algorithm in MEGA 11 [20].

RESULTS

Genetic diversity revealed by SSR markers

Genetic diversity parameters were computed for the 23 cowpea accessions using 27 SSR markers. The markers produced a total of 89 alleles across the panel. Of the 27 loci, 26 were polymorphic, giving a polymorphism percentage of 96.2 percent, while one locus, CEDGAG001, produced only a single allele. This monomorphic locus showed zero values for the number of effective alleles, Shannon's information index, both heterozygosity measures, and the polymorphism information content, and its fixation index could not be calculated, so it does not contribute to the diversity estimates described below.

The average number of alleles per locus was 3.33, which is significantly lower compared to other studies of this kind. Among the polymorphic loci, Na ranged from 2 to 6. VuSSR13, VuSSR15, VuSSR21, CEDG176, and CEDG097 carried the minimum number of alleles, two each, while VrSSR16 carried the maximum, six each. Variations in the number of accessions studied and the number and choice of markers used in different surveys are plausible reasons for such differences between studies.

The effective number of alleles expresses how many alleles are contributing at appreciable frequencies at a locus, because rare alleles add little to the effective count. Ne averaged 2.40 and ranged from 1.45 to 3.64. The comparison of Na and Ne shows that, although up to six alleles were detected, most of those alleles occurred at low frequencies. VuSSR2 had the minimum Ne of 1.45, and CEDG143 had the maximum Ne of 3.64, meaning that the two alleles at CEDG143 were the most evenly balanced in frequency of any locus in the study.

Shannon's information index (I), which combines the number of alleles at a locus with the evenness of their frequencies, ranged from 0.576 in VuSSR2 to 1.415 in CEDG143, with an average of 0.944. The loci with the highest values for this index, namely CEDG143 (1.415), VuSSR10 (1.300), VrSSR16 (1.299), and CEDG068 (1.292), were therefore the most informative about allele evenness as well as allele number. The full set of diversity parameters for all 27 loci is presented in the **Table 3**.

Heterozygosity

Heterozygosity is one of the most important measures of genetic diversity in natural populations, because it is determined by the number of alleles at a locus and by their frequencies, which together define the probability of being heterozygous at that locus. Observed heterozygosity is measured directly from the genotypes, while expected heterozygosity is the value predicted from allele frequencies under random mating, and unbiased expected heterozygosity (uHe) corrects this expectation for the finite sample of accessions.

Observed heterozygosity ranged from 0 to 0.864, with a mean of only 0.101 across all loci. The maximum Ho was obtained in VuSSR17 at 0.864, followed by VuSSR9 at 0.783 and VuSSR7 at 0.650. Ho was zero at 21 loci, including VuSSR4, VuSSR5, VuSSR6, VuSSR11, VuSSR12, VuSSR13, VuSSR15, VuSSR19, VuSSR20, VuSSR21, VuSSR2, CEDG176, CEDG143, CEDG068, CEDG097, CEDG056, CEDG180, CEDGAG001, CEDG006, VrSSR16, and VrSSR17, and it was above zero at only six loci in total. Expected heterozygosity ranged from 0.310 in VuSSR2 to 0.725 in CEDG143, with a mean of 0.545, and the corresponding uHe averaged 0.561. CEDG143 thus showed the maximum He and VuSSR2 the minimum, the same ordering seen for the effective number of alleles and Shannon's index.
The observed heterozygosity was far lower than the expected heterozygosity at nearly every locus. This large difference between Ho and He reflects a strong deficit of heterozygotes relative to the expectations of random mating, which is the expected outcome in a crop with a predominantly self-pollinating mating system.

Fixation index

The fixation index (F), also known as the inbreeding coefficient, compares observed with expected heterozygosity and expresses the degree to which a locus departs from Hardy Weinberg proportions. Values of 1 indicate complete homozygosity, values near 0 indicate proportions close to random mating, and negative values indicate an excess of heterozygotes. In this study F ranged from -0.304 to 1.00, with an average of 0.822. Twenty primers showed the maximum value of 1.00, indicating complete fixation at those loci. Only three loci, VuSSR7 (-0.086), VuSSR9 (-0.333) and VuSSR17 (-0.304), recorded negative values, showing heterozygote excess at these positions, which is consistent with them also recording the three highest observed heterozygosity. The remaining polymorphic loci had positive values below 1, namely VuSSR3 (0.861), VuSSR10 (0.619), and VuSSR8 (0.607). The high average F of 0.861 agrees with the self-pollinating nature of cowpea, which drives homozygosity within accessions.

Polymorphism information content

The polymorphism information content of the SSR loci was calculated as the measure of allelic diversity at a particular locus, since it reflects both the number of alleles and the distribution of their frequencies. PIC ranged from 0.282 in VuSSR2 to 0.681 in CEDG143, with a mean of 0.498, while the monomorphic CEDGAG001 contributed a value of zero. The most informative markers were CEDG143 (0.681), CEDG068 (0.649), VuSSR (0.628) and CEDG006 (0.623), and these are therefore the most suitable of the assayed primer pairs for future diversity and fingerprinting work in this material. A mean PIC above 0.5 confirms that

the SSR set used here was sufficiently polymorphic to resolve genetic differences among the 23 accessions.

Visualising the relationship among genotypes using cluster analysis

To visualise the genetic relationships among the accessions, a distance matrix was converted into a dendrogram with the UPGMA (unweighted pair group method with arithmetic mean) algorithm. Since UPGMA clusters based on the distance between accessions or clusters with the lowest pairwise distance, accessions that are close to one another on the tree are the ones that share the most genetic similarities. Four major clusters have been generated based on the SSR marker data for the accessions, which are summarised in Table 4.

Cluster I had seven genotypes, Cluster III had four while Cluster II and Cluster IV had six genotypes each. This clustering is clearly illustrated in the dendrogram below, where branch lengths correspond to genetic distances between accessions or clusters. The first cluster from the top consists of Cluster IV, which includes genotypes EC240808, EC240826, EC240841, EC240891, EC240971 and Kashi Kanchan. Cluster III, included EC240648, EC240755, EC240765 and EC240829, forms the middle group. Cluster II, with EC240703, EC240714, EC240700, EC240665, EC724539 and EC240693B and Cluster I, with EC219922, EC240636, EC240663, EC149457, EC149471, EC170072, and EC170579-B, occupy the lower portion of the tree. In the variety Kashi Kanchan of SHUATS in Cluster IV with five germplasms of NBPGR were included, instead of being kept separate from them, suggesting that there is some genetic similarity between the released variety and the germplasms. The dendrogram structure is fully consistent with the cluster formation in Table 4.

Principal coordinate analysis

For further validation of the clustering procedure, a principal coordinate analysis was performed. PCoA transforms the distance matrix between accessions into a few axes such that accessions that cluster together on the plot are genetically similar, and the percentage of the total genetic variability explained by each axis is given next to the axis. The cumulative percentage variation explained by the first three axes using the SSR dataset is 57.31 percent. Taking individual axes into account, the first, second, and third axes explain 28.42, 16.54 and 12.35 percent variation, respectively see Table 5.

All the accessions that have been grouped separately in the UPGMA dendrogram have grouped separately in all three axes in the PCoA plot. In summary, the SSR assay detected a moderate but useful level of diversity, with 89 alleles over 26 polymorphic loci, a mean PIC of

0.498, and a clear division of the 23 accessions into four genetic clusters, while the very low observed heterozygosity and high fixation index of 0.822 reflect the strongly self-pollinating mating system of the crop.

DISCUSSION

Gene diversity and allelic pattern

The number of alleles per locus ranged from one to six, a relatively modest range compared with earlier reports. Sarr et al. [21] recorded two to 15 alleles per locus, averaging 6.2, in cultivated and wild cowpea from Senegal. Similarly, one to 16 alleles were reported among 22 Senegalese varieties and inbred lines [22]. In Sudanese germplasm, Ali et al. detected 129 alleles across 16 loci, averaging 8.1 per locus [23]. Such ranges are generally obtained from large, geographically diverse collections, whereas smaller panels such as the present one are likely to reveal fewer alleles. Nevertheless, the range observed here is comparable with or wider than those reported for Ghanaian cowpea [24] and recently screened Indian genotypes [25]. Differences among studies are therefore more plausibly attributable to sample size, marker number and selection, and scoring criteria than to diversity alone [26, 27, 28]**.** In particular, allele counts should not be interpreted as evidence of low diversity, because cowpea is predominantly self-pollinating and has a narrower genetic base than many other crops [29, 30].

Although some loci amplified as many as six alleles, most showed relatively low-frequency alleles. This pattern is reflected in the difference between the mean observed and effective numbers of alleles. The effective number represents the number of equally frequent alleles required to produce the observed heterozygosity and therefore accounts for allele-frequency distribution rather than simply allele counts. VuSSR2 had the lowest effective allele number, indicating an uneven distribution, whereas CEDG143 had the highest, suggesting a more even contribution of alleles to diversity. VrSSR16 were the most allele-rich loci, with six alleles, followed by VuSSR7, VuSSR10 and CEDG143, with five each. VuSSR13, VuSSR15, VuSSR21, CEDG176, and CEDG097 carried only two alleles. Effective allele numbers were generally well below observed counts, as expected in a predominantly homozygous crop [31]. Polymorphism information content measures marker discriminatory power by incorporating both allele number and frequency distribution [42, 31]. Values above 0.50 are considered highly informative [32], whereas those between 0.25 and 0.50 are informative [33]. In the present study, PIC ranged from 0.282 at VuSSR2 to 0.681 at CEDG143, with a mean of 0.498; 18 of the 26 polymorphic loci had values of at least 0.50. This mean exceeded those reported for

Ghanaian cowpea [24], Senegalese varieties [22], Chinese germplasm [34], Korean accessions [34], and recently evaluated Indian genotypes [25]. It was lower only than the value reported for wild cowpea, where PIC reached 0.870 [14]. These findings indicate substantial genetic variability in the Indian material and highlight loci suitable for genotype identification and parental selection [31].

Shannon's information index, which reflects allele number and frequency evenness, had a mean of 0.944 and was highest at CEDG143. This result indicates moderate diversity and agrees with the allele-frequency patterns described above.

Heterozygosity combines allele number and frequency to estimate the probability that an individual is heterozygous at a locus. Mean expected heterozygosity was 0.545, whereas mean observed heterozygosity was only 0.101. The average gene diversity expressed by expected heterozygosity which exceeded the mean values reported for Senegalese cowpea [21], Senegalese varieties [22] and Korean germplasm [35] and was close to the average gene diversity of 0.60 reported for Sudanese cowpea [23]. Observed heterozygosity, however, was substantially lower than expected at nearly all loci, indicating a pronounced heterozygote deficit. Similar patterns have been documented in Togolese accessions [35], Korean germplasm [35] and 370 accessions from the IITA Genetic Resources Centre evaluated by Fatokun et al. [36] and Seo et al. [37]. This deficit is consistent with a broader pattern in cowpea germplasm and can be attributed primarily to predominant self-pollination, along with selection pressure and repeated use of farmer-preferred materials, which may further reduce genetic variation [38, 39]**.**

The fixation index measures deviation from Hardy-Weinberg equilibrium through comparison of heterozygosity in the observed versus expected genotype distributions. The mean fixation index of 0.822 and complete fixation at 18 loci confirm the extensive homozygosity expected due to self-fertilization [31]. Similarly, the mean inbreeding coefficient of 0.830 was reported for the Togolese cowpea population [35]. VuSSR7, VuSSR9 and VuSSR17 loci had negative fixation indices, suggesting heterozygote excess. This finding is contrary to the general trend and may be a result of increased gene flow between subpopulations, statistical error, or sampling bias, especially in small samples [31]**.** Given that these two loci also had the highest observed heterozygosity, negative fixation index may be genuine heterozygotes rather than a departure from the inbreeding nature of the species.

Genetic relationships among cowpea accessions

Cluster analysis using UPGMA and Principal Coordinate Analysis was carried out to investigate the relationships between the accessions. PCoA is a statistical method that helps in

detecting genetic structure among populations and sub-populations by plotting a distance matrix on a series of orthogonal axes [18] . First three coordinates accounted for 57.31 percent of the total variation in the data set. Grouping based on these coordinates was congruent with the grouping pattern observed through UPGMA dendrogram. UPGMA grouped the 23 accessions into four clusters where Cluster I was the largest with seven accessions followed by Clusters II and IV, which comprised of six accessions while Cluster III consisted of four accessions each. Accessions from both the sources were interspersed in each other. Kashi Kanchan grouped in Cluster IV along with five NBPGR accessions whereas NBPGR accessions existed in all the four clusters. Thus, genetic relationships did not closely correspond to source or collection history. Instead, the released variety shared measurable similarity with conserved germplasm rather than forming a distinct group.

The four clusters formed in the present set of 23 accessions confirm considerable genetic diversity. Similar kinds of sub-divisions have been found in other large sets of cowpeas like three major clusters in Sudanese cowpea [23], five branches in Ghanaian germplasm [24], three groups in cultivated African cowpea with 65 percent similarity [40], and four major populations in global and Korean germplasm [37]. These results provide the possibility of selecting divergent parents for making segregating populations with diverse alleles.

The results obtained above have practical significance in the improvement of cowpea germplasm. The accessions are maintained at ICAR-NBPGR and their SSR profiles will aid in reducing redundancy, developing a core collection for focused evaluation, and forming crossing blocks.

The high level of marker polymorphism observed (96.2 per cent) contrasts with the low microsatellite polymorphism of 4 to 12 per cent reported in several earlier cowpea studies [29, 30], which underlines the informativeness of the primer set used here for Indian germplasm. Earlier global surveys identified Indian cowpea as one of the more variable groups, with a gene diversity of 0.35 [28], yet molecular characterisation of Indian accessions remains limited in comparison with African collections, which dominate the cowpea diversity literature [41]. The present study adds SSR based baseline data for Indian germplasm conserved in the national genebank and demonstrates that even a modest number of accessions can harbour enough diversity for meaningful breeding applications. Future work should move towards functional characterisation and association mapping to link the polymorphic loci identified here with agronomically important traits, which would accelerate the use of this germplasm in cowpea improvement [1, 10].

CONCLUSION

This study characterized the genetic diversity and population structure of 23 Indian cowpea accessions using 27 SSR markers, 26 of which were polymorphic. The markers amplified 89 alleles, averaging 3.33 alleles per locus, with a mean effective number of alleles of 2.40 and a Shannon information index of 0.944, indicating a moderate but useful level of genetic variability. The expected heterozygosity was 0.545 while the observed heterozygosity was only 0.101. High fixation index and complete fixation of the 18 loci were yet another confirmation of the self-pollinating nature of cowpea. Polymorphism information content had a mean value of 0.498 and 18 out of 26 polymorphic loci had an informative value of more than 0.50 showing that the panel of SSR used here was informative enough to discriminate varieties, fingerprint and assess genetic diversity. CEDG143 was the most informative locus in terms of high polymorphism information content and effective allele number. Hierarchical cluster analysis using UPGMA and principal coordinate analysis showed similar results grouping accessions into four clusters explaining 57.31% of the variation among the first three coordinates. Cluster affiliation was not linked to either accession origin or collection history. As an example, Kashi Kanchan, a released variety, clustered with NBPGR germplasms in Cluster IV, reflecting the presence of significant genetic similarities between developed and conserved accessions and not the division along origin lines. This information has provided baseline information about SSRs in cowpea germplasm, which have received lesser attention compared to their African counterparts. The generated clusters can be used as a means for selection of divergent parents in cowpea breeding programs. In addition, the informative loci identified in this study can support core collection development and future marker-assisted selection. Extending this work through association mapping could help link these markers to agronomically important traits and accelerate their application in cowpea improvement.

Acknowledgement

The authors express their gratitude to ICAR-National Bureau of Plant Genetic Resources, New Delhi and Sam Higginbottom University of Agriculture, Technology and Sciences, Prayagraj, Uttar Pradesh for providing the necessary facilities. The authors also express their sincere gratitude to the anonymous reviewers for their constructive comments and valuable suggestions, which significantly improved the manuscript.

Contributors

The research was designed by JPE in collaboration with AP, VPS, DPW, and NT. The experiment was conducted by AP with assistance from JPE. Molecular work was carried out under the laboratory of DPW and conducted by AP. The manuscript was drafted by DT, MK

and NT and edited by JPE and VPS. All authors reviewed and approved the final manuscript for publication.

Funding

The present manuscript has no funding source to declare.

Conflict of interest

The authors declare no known financial conflicts of interest or personal relationships that could have influenced the work presented in this paper.

Availability of data and material

This article contains all the data generated and analysed during the course of this study.

Ethical approval

The plant genetic resources used in this study were obtained from the ICAR-National Bureau of Plant Genetic Resources (ICAR-NBPGR), New Delhi, and Sam Higginbottom University of Agriculture, Technology and Sciences, Prayagraj. No wild plant material was collected by the author for this study. The use of cowpea accession was conducted in accordance with applicable institutional and national guidelines and regulations governing the use of plant genetic resources.

Permission to collect plant material

The cowpea accessions used in this study were received for research use, and the research based on these accessions was undertaken with the permission of the National Genebank, ICAR-NBPGR and SHUATS

Compliance with ethical standard

This study was conducted in accordance with institutional and ethical guidelines. The research did not involve human or animal participants. All experimental procedures were performed in accordance with standard laboratory and biosafety protocols.

Consent to participate

Not applicable

Consent to publication

Not applicable

REFERENCES

1. Boukar O, Fatokun CA, Huynh BL, Roberts PA, Close TJ. Genomic Tools in Cowpea Breeding Programs: Status and Perspectives. Frontiers in Plant Science. 2016 June 3;7. https://doi.org/10.3389/fpls.2016.00757

2. Indian Institute of Pulses Research [Internet]. Icar-iipr.org.in. https://www.icar-iipr.org.in/cowpea/

3. Cruz FJR, de Almeida HJ, Santos DMM dos. Growth, nutritional status and nitrogen metabolism in Vigna unguiculata (L.) Walp is affected by aluminum. Australian Journal of Crop Science. 2014 July 1;8(7):1132–9.

4. Elharadallou IIKSB. Functional properties of cowpea (Vigna ungiculata l.walp), and Lupin (lupinus termis) flour and protein isolates. Journal of Nutrition & Food Sciences. 2013;03(06). https://doi.org/10.4172/2155-9600.1000234
5. Kirse-Ozolina A, Karklina D. Integrated evaluation of Cowpea (Vigna unguiculata (L.) Walp.) And maple pea (pisum sativum var. arvense L.) spreads. Agronomy Research. 2015 July;13:956–68.

6. Diouf D, Hilu KW. Microsatellites and RAPD markers to study genetic relationships among cowpea breeding lines and local varieties in Senegal. Genetic Resources and Crop Evolution. 2005 Dec;52(8):1057–67. https://doi.org/10.1007/s10722-004-6107-z
7. Mortimore, M., J.SinghB. B.HarrisF.BladeS. F. (1997). “Cowpea in traditional cropping systems,” in *Advances in Cowpea Research. Co-publication of International Institute of Tropical Agriculture (IITA) and Japan International Research Center for Agricultural Sciences (JIRCAS).* Eds. SinghB. B.Mohan RajD. R.DashiellK. E.JackaiL. E. N. (Ibadan, Nigeria: IITA), pp. 99–Pp:113.
8. Boukar O, Abberton M, Oyatomi O, Togola A, Tripathi L, Fatokun C. Introgression breeding in cowpea [vigna unguiculata (L.) Walp.]. Frontiers in Plant Science. 2020 Sept 16;11. https://doi.org/10.3389/fpls.2020.567425

9. Edema R, Adjei EA, Ozimati AA, Tusiime SM, Badji A, Ibanda A, et al. Genetic Diversity of Cowpea Parental Lines Assembled for Breeding in Uganda. Plant Molecular Biology Reporter. 2023 July 18;41(4):713–25. Available from: https://doi.org/10.1007/s11105-023-01394-6

10. Omomowo OI, Babalola OO. Constraints and prospects of improving cowpea productivity to ensure food, nutritional security and environmental sustainability. Frontiers in Plant Science. 2021 Oct 22;12. https://doi.org/10.3389/fpls.2021.751731

11. Fang J, Chao CCT, Roberts PA, Ehlers JD. Genetic diversity of cowpea [vigna unguiculata (L.) Walp.] In four west African and USA breeding programs as determined by AFLP analysis. Genetic Resources and Crop Evolution [Internet]. 2007 Sept 1 [cited 2026 Sept 2];54(6):1197–209. https://doi.org/10.1007/s10722-006-9101-9
12. Fatokun CA, Danesh D, Young ND, Stewart EL. Molecular taxonomic relationships in the genus Vigna based on RFLP analysis. Theoretical and Applied Genetics. 1993 Mar;86(1):97–104. https://doi.org/10.1007/bf00223813

13. Ghalmi N, Malice M, Jacquemin JM, Ounane SM, Mekliche L, Baudoin JP. Morphological and molecular diversity within Algerian cowpea (Vigna unguiculata (L.) Walp.) Landraces. Genetic Resources and Crop Evolution. 2009 Sept 11;57(3):371–86. https://doi.org/10.1007/s10722-009-9476-5

14. Ogunkanmi, L.A, Ogundipe, O.T., Ng, N.Q., Fatokun, C.A. Genetic diversity in wild relatives of cowpea (Vigna unguiculata) as revealed by simple sequence repeats (SSR) markers. Journal of Food, Agriculture & Environment. 2008;6(3&4):132–7.

15. Wang ML, Barkley NA, Gillaspie GA, Pederson GA. Phylogenetic relationships and genetic diversity of the USDA Vigna germplasm collection revealed by gene-derived markers and sequencing. Genetics Research [Internet]. 2008 Dec 1 [cited 2026 Sept 2];90(6):467–80. https://doi.org/10.1017/S0016672308009889
16. Songok S, Ferguson M, Muigai A, Silim S. Genetic diversity in PIGEONPEA [cajanus cajan (L.) MILLSP.] Landraces as revealed by simple sequence repeat markers. Afr J Biotech. 2010 May;9.

17. Healey A, Furtado A, Cooper T, Henry RJ. Protocol: A simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. Plant Methods. 2014;10(1):21. https://doi.org/10.1186/1746-4811-10-21

18. PEAKALL R, SMOUSE PE. Genalex 6: Genetic analysis in excel. Population genetic software for teaching and research. Molecular Ecology Notes. 2005 Mar;6(1):288–95. https://doi.org/10.1111/j.1471-8286.2005.01155.x
19. Taller J. PICcalc: An Online Program to Calculate Polymorphic Information Content for Molecular Genetic Studies Sandor Nagy• Peter Poczai• Istvan Cernak• Ahmad Mousapour Gorji• Geza Heged} us•. Biochem Genet. 2012;50:670-2. https://doi.org/10.1007/s10528-012-9509-1
20. Tamura K, Stecher G, Kumar S. MEGA11: Molecular evolutionary genetics analysis version 11. Molecular Biology And Evolution. 2021;38(7):3022–7. https://doi.org/10.1093/molbev/msab120
21. Sarr A, Bodian A, Gbedevi KM, Ndir KN, Ajewole OO, Gueye B, et al. Genetic diversity and population structure analyses of wild relatives and cultivated cowpea (Vigna unguiculata (L.) Walp.) From Senegal using simple sequence repeat markers. Plant Molecular Biology Reporter. 2020 July 27;39(1):112–24. https://doi.org/10.1007/s11105-020-01232-z
22. Badiane FA, Gowda BS, Cissé N, Diouf D, Sadio O, Timko MP. Genetic relationship of cowpea (Vigna unguiculata) varieties from Senegal based on SSR markers. Genetics and Molecular Research. 2012;11(1):292–304. https://doi.org/10.4238/2012.february.8.4
23. Ali ZB, YAO KN, Odeny DA, Kyalo M, Skilton R, Eltahir IM. Assessing the genetic diversity of cowpea [Vigna unguiculata (L.) Walp.] Accessions from Sudan using simple sequence repeat (SSR) markers. African Journal of Plant Science. 2015 July 31;9(7):293–304. https://doi.org/10.5897/ajps2015.1313
24. Asare AT, Gowda BS, Galyuon IKA, Aboagye LL, Takrama JF, Timko MP. Assessment of the genetic diversity in cowpea (*Vigna unguiculata L.* Walp.) Germplasm from Ghana using simple sequence repeat markers. Plant Genetic Resources. 2010 Apr 13;8(2):142–50. https://doi.org/10.1017/s1479262110000092
25. Gavade RT, Sawardekar SV, Kunkerkar RL, Mane AV, Dodake SB, Palshetkar MG, et al. Determination of genetic diversity of cowpea [vigna unguiculata L. (walp.)]

Genotypes using SSR marker. Journal of Advances in Biology & Biotechnology. 2025 Nov 7;28(11):679–93. https://doi.org/10.9734/jabb/2025/v28i113267

26. Lacape JM, Nguyen TB, Courtois B, Belot JL, Giband M, Gourlot JP, et al. QTL analysis of cotton fiber quality using multiple *gossypium hirsutum* × *gossypium barbadense* backcross generations. Crop Science. 2005 Jan;45(1):123–40. https://doi.org/10.2135/cropsci2005.0123a

27. Varshney RK, Thudi M, Roorkiwal M, He W, Upadhyaya HD, Yang W, et al. Resequencing of 429 chickpea accessions from 45 countries provides insights into genome diversity, domestication and agronomic traits. Nature Genetics. 2019 May 1;51(5):857–64. https://doi.org/10.1038/s41588-019-0401-3

28. Xiong H, Shi A, Mou B, Qin J, Motes D, Lu W, et al. Genetic diversity and population structure of cowpea (Vigna unguiculata L. Walp). Parida SK, editor. PLOS ONE. 2016 Aug 10;11(8):e0160941. https://doi.org/10.1371/journal.pone.0160941

29. Symphorien A, Jeninah K, Kassim S, Paul G, Richard E, Achille EA, et al. Microsatellites markers associated with resistance to flower bud thrips in a cowpea F2 population derived from genotypes TVU-123 and WC36. African Journal of Biotechnology. 2018 June 20;17(25):767–778. https://doi.org/10.5897/ajb2018.16480

30. Li CD, Fatokun CA, Ubi B, Singh BB, Scoles GJ. Determining genetic similarities and relationships among cowpea breeding lines and cultivars by microsatellite markers. Crop Science. 2001 Jan;41(1):189–197. https://doi.org/10.2135/cropsci2001.411189x

31. Chesnokov YuV, Artemyeva AM. Evaluation of the measure of polymorphism. Sel'skokhozyaistvennaya Biologiya. 2015 Oct;50(5):571–578. https://doi.org/10.15389/agrobiology.2015.5.571eng

32. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics. 1980 May;32(3):314. https://pmc.ncbi.nlm.nih.gov/articles/PMC1686077/

33. Palumbo F, Barcaccia G. Critical aspects on the use of microsatellite markers for assessing genetic identity of crop plant varieties and authenticity of their food

derivatives. Rediscovery of Landraces as a Resource for the Future. 2018 Sept 12. https://doi.org/10.5772/intechopen.70756

34. Abdoul MIS, Christian F, Bunmi O, Ousmane B. Current status of molecular tools development for cowpea [Vigna unguiculata (L.) Walp.] Improvement. African Journal of Biotechnology. 2023 Aug 31;22(8):26–137. https://doi.org/10.5897/ajb2022.17541

35. Dagnon YD, Palanga KK, Bammite D, Bodian A, Akabassi GC, Foncéka D, et al. Genetic diversity and population structure of cowpea [vigna unguiculata (L.) Walp.] Accessions from Togo using SSR markers. Kumar A, editor. PLOS ONE. 2022 Oct 5;17(10):e0252362. https://doi.org/10.1371/journal.pone.0252362

36. Fatokun C, Girma G, Abberton M, Gedil M, Unachukwu N, Oyatomi O, et al. Genetic diversity and population structure of a mini-core subset from the world cowpea (Vigna unguiculata (L.) Walp.) Germplasm collection. Scientific Reports. 2018 Oct 30;8(1). https://doi.org/10.1038/s41598-018-34555-9

37. Seo E, Kim K, Jun TH, Choi J, Kim SH, Muñoz-Amatriaín M, et al. Population structure and genetic diversity in Korean cowpea germplasm based on SNP markers. Plants. 2020 Sept 1;9(9):1190. https://doi.org/10.3390/plants9091190

38. Chipeta MM, Kafwambira J, Yohane E. Cowpea genetic diversity, population structure and genome-wide association studies in Malawi: Insights for breeding programs. Frontiers in Plant Science. 2025 Jan 20;15. https://doi.org/10.3389/fpls.2024.1461631

39. Diallo S, Badiane FA, Kabkia BNPA, Diédhiou I, Diouf M, Diouf D. Genetic diversity and population structure of cowpea mutant collection using SSR and ISSR molecular markers. Scientific Reports. 2024 Dec 30;14(1). https://doi.org/10.1038/s41598-024-83087-y

40. Ogunkanmi, LA, Ogundipe, OT, Fatokun, CA. Molecular characterization of cultivated cowpea (Vigna unguiculata L. Walp) using simple sequence repeats markers. African Journal of Biotechnology. 2014 Aug 20;13(34):3464–72. https://doi.org/10.5897/ajb2013.13166

41. Carvalho M, Lino-Neto T, Rosa E, Carnide V. Cowpea: A legume crop for a challenging environment. Journal of the Science of Food and Agriculture. 2017 Mar 17;97(13):4273–84. https://doi.org/10.1002/jsfa.8250

42. El-Sobky MA, Eissa RA, Abdel-Lateif KS, Fahmi AI, El-Zanaty AM, Hassan MM, et al. Genetic diversity assessment of Trichoderma spp. isolated from various Egyptian locations using its gene sequencing marker, rep-PCR, and their cellulolytic activity. Egyptian Journal of Biological Pest Control. 2024 Apr 25;34(1). https://doi.org/10.1186/s41938-024-00784-6

Table 1: List of cowpea accessions used in the study and their respective sources.

Sr. No.	Genotype name	Source
1	EC149457	NBPGR, New Delhi
2	EC149471	NBPGR, New Delhi
3	EC170072	NBPGR, New Delhi
4	EC170579-B	NBPGR, New Delhi
5	EC219922	NBPGR, New Delhi
6	EC240636	NBPGR, New Delhi
7	EC240648	NBPGR, New Delhi
8	EC240663	NBPGR, New Delhi
9	EC724539	NBPGR, New Delhi
10	EC240665	NBPGR, New Delhi
11	EC240693-1	NBPGR, New Delhi
12	EC240700	NBPGR, New Delhi
13	EC240703	NBPGR, New Delhi
14	EC240714	NBPGR, New Delhi
15	EC240755	NBPGR, New Delhi
16	EC240765	NBPGR, New Delhi
17	EC240808	NBPGR, New Delhi
19	EC240829	NBPGR, New Delhi
20	EC240841	NBPGR, New Delhi
21	EC240891	NBPGR, New Delhi
22	EC240917	NBPGR, New Delhi
23	Kashi Kanchan	SHUATS, Prayagraj

Table 2: Details of SSR markers used for molecular characterization of the cowpea genotypes, including primer sequences and annealing temperatures.

Marker	Forward primer	Reverse primer	Annealing temp. (°C)
VuSSR3	GATTTAGGATGCATCAGTCCAAC	CTGGATACTGCAACGGGAG	60
VuSSR4	ACTATCACACCCAAAAGCACACT	CAAGAGCAGTGAAGAAGAAGAGC	60
VuSSR5	GGAAAAGAAACTCTAGCTTGGCA	ATCCCCATTGCATCTTCTTATTT	61
VuSSR6	ATTTGCACTTTTCGACTCCAGTA	GAAGAAGAGAAGTTATCGTGCCA	60
VuSSR7	TAGTGCTGCTGGTAATTCTGGAT	CTCTTCTTCAAAGGGGTTGACTT	60
VuSSR8	ACACTTCCTAGGTGATTGTAGCG	ATCACCCAACCTTAACTCTCACA	60
VuSSR9	AACTCAGAGACAAAACTCACCCA	CCACTCGTTTTAGCTTCTTCAAA	60
VuSSR10	TCATTCTTGTTCTGTGCTTCTCA	ATTAATGGGACAAGAAATGCAGA	60
VuSSR11	CTCCTCCTACTACTCTTCCCGAC	AAAGAGAGGAAGGTGGTGGTTT	60
VuSSR12	CTGTCGGAGAAATTGTAGTTTGG	CAACAATCACAACACCAAGAAGA	60
VuSSR13	AAGTTGCAGAAGCAGAAGTGTTC	TCTCGTTTTGTCTCGTGAAGAAG	61
VuSSR15	GCTTTAAACAGGGGAACCATTAT	CAAAACAAAGCATCAAGTAACCA	59
VuSSR17	AGCACCATTTCTGATAAGTTCCA	ACATCGGACACAACCTCACTTAT	60
VuSSR19	GTTCTGATGCACGCTCATCTT	AACACACTCAATACAGCAACGGT	61
VuSSR20	AGCAACAAGAAGAGAAGGTCTCC	GATCCGAGTTTTCGGAATCA	60
VuSSR21	TGGGATATCTCTCTGGATGAAAA	TCTCTGTTCGCTCAGAAGTAGGT	60
VuSSR2	AAACATTTTTCTTAAACCGCCAC	GGCGTGGATATTAGGGATTACTT	60
CEDG176	GGTAACACGGGTTCAGATGCC	CAAGGTGGAGGACAAGATCGG	57
CEDG143	GATGAACTCGTCTCGCTCATCG	CTGGACGCGTCTACTCAGAC	60
CEDG068	TCTCCATAGGAACCCCTGAAAG	TGGGATCAGTGAATTCGCCAG	57.7
CEDG097	GTAAGCCGCATCCATAATTCCA	TGCGAAAGAGCCGTTAGTAGAA	60
CEDG056	TTCCATCTATAGGGGAAGGGAG	GCTATGATGGAAGAGGGCATGG	57.7

CEDG180	GGTATGGAGCAAAACAATC	GTGCGTGAAGTTGTCTTATC	60
VrSSR16	TCTCCATCCCCATCTTCATC	GGAGAGATCTGCGACCTTTG	60
VrSSR17	AACTTCGTCCTGCGCTTAAA	AGCATGACCACACCAATCAA	58
CEDGAG001	CTCATCAGGGACATCCTCCC	GATCGTGATCGATCCAACGGTC	56.1
CEDG006	AATTGCTCTCGAACCAGCTC	GGTGTACAAGTGTGTGCAAG	51.3

Table 3: Genetic diversity and polymorphism parameters of the SSR loci used for characterization of 23 cowpea genotypes.

Locus	Na	Ne	I	Ho	He	uHe	F	PIC
VuSSR3	3.000	1.559	0.612	0.050	0.359	0.368	0.861	0.310
VuSSR4	3.000	2.532	0.999	0.000	0.605	0.621	1.000	0.527
VuSSR5	3.000	2.689	1.036	0.000	0.628	0.658	1.000	0.551
VuSSR6	3.000	2.462	0.974	0.000	0.594	0.613	1.000	0.511
VuSSR7	5.000	2.492	1.126	0.650	0.599	0.614	-0.086	0.537
VuSSR8	3.000	1.496	0.619	0.130	0.332	0.339	0.607	0.304
VuSSR9	4.000	2.421	1.057	0.783	0.587	0.600	-0.333	0.525
VuSSR10	5.000	2.909	1.300	0.250	0.656	0.673	0.619	0.616
VuSSR11	3.000	2.359	0.972	0.000	0.576	0.592	1.000	0.513
VuSSR12	4.000	2.778	1.168	0.000	0.640	0.656	1.000	0.581
VuSSR13	2.000	1.710	0.606	0.000	0.415	0.428	1.000	0.329
VuSSR15	2.000	1.899	0.666	0.000	0.473	0.492	1.000	0.361
VuSSR17	4.000	2.960	1.180	0.864	0.662	0.678	-0.304	0.597
VuSSR19	4.000	3.108	1.253	0.000	0.678	0.699	1.000	0.628
VuSSR20	3.000	2.528	0.991	0.000	0.604	0.625	1.000	0.521
VuSSR21	2.000	1.761	0.624	0.000	0.432	0.444	1.000	0.339
VuSSR2	3.000	1.449	0.576	0.000	0.310	0.317	1.000	0.282
CEDG176	2.000	1.841	0.649	0.000	0.457	0.471	1.000	0.352
CDEG143	5.000	3.636	1.415	0.000	0.725	0.744	1.000	0.681
CDEG068	4.000	3.306	1.292	0.000	0.698	0.717	1.000	0.649
CEDG097	2.000	2.000	0.693	0.000	0.500	0.519	1.000	0.375
CEDG056	3.000	2.455	0.965	0.000	0.593	0.610	1.000	0.505
CEDG180	4.000	3.073	1.220	0.000	0.675	0.702	1.000	0.613
VrSSR16	6.000	2.882	1.299	0.000	0.653	0.669	1.000	0.597
VrSSR17	3.000	2.359	0.972	0.000	0.576	0.592	1.000	0.513
CEDGAG001	1.000	1.000	0.000	0.000	0.000	0.000	#N/A	0.000
CEDG006	4.000	3.175	1.224	0.000	0.685	0.703	1.000	0.623

Abbreviations: Na: number of alleles; Ne: effective number of alleles; I: Shannon's information index; Ho: observed heterozygosity; He: expected heterozygosity; uHe: unbiased expected heterozygosity; F: fixation index; PIC: polymorphism information content.

Table 4: Distribution of cowpea genotypes into distinct clusters based on SSR marker-derived genetic diversity.

Clusters	Genotypes
Cluster I	EC170579-B, EC170072, EC149471, EC149457, EC240663, EC240636, EC219922
Cluster II	EC240703, EC240714, EC240700, EC240665, EC724539, EC240693B
Cluster III	EC240648, EC240755, EC240765, EC240829
Cluster IV	EC240808, EC240826, EC240841, EC240891, EC240917, Kashi Kanchan

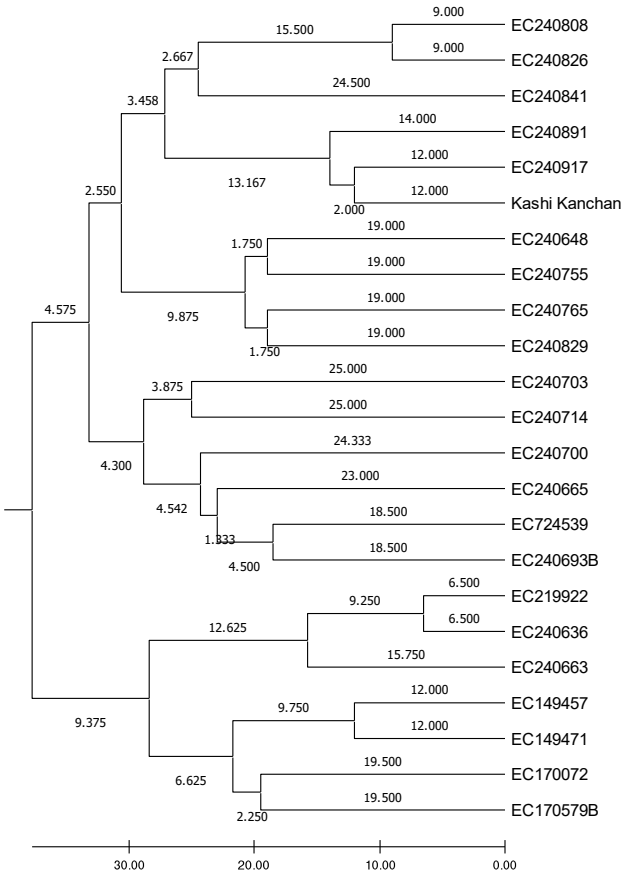


Figure 1: UPGMA-based dendrogram showing genetic relationships among 23 cowpea genotypes inferred from SSR marker data. The genotypes were grouped into distinct clusters based on their genetic similarity.

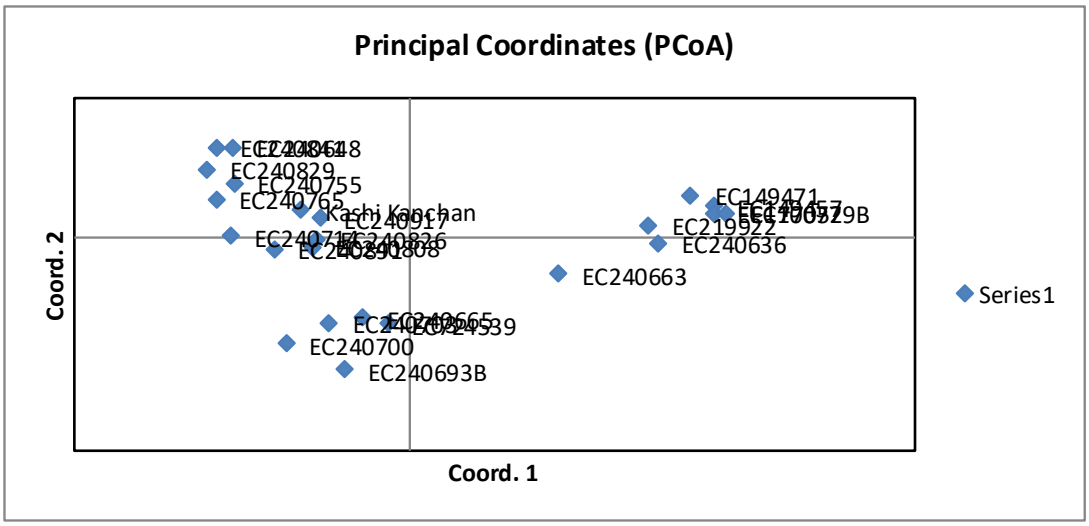


Figure 2: Principal coordinate analysis (PCoA) showing the genetic relationships and distribution of 23 cowpea genotypes based on SSR marker data. The first two coordinate axes represent the major components of molecular variation among the genotypes.

Table 5: Percentage variation explained by the first three principal coordinate axes and their cumulative contribution to the total molecular variation among cowpea genotypes.

Axis	**1**	**2**	**3**
%	28.42	16.54	12.35
Cum %	28.42	44.96	57.31

Supplementary Files

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

- Scoring.xlsx