PLANT SCIENCE TODAY
Vol x(x): xx–xx
https://doi.org/10.14719/pst.9286

eISSN 2348-1900

RESEARCH ARTICLE

Genetic diversity analysis in eggplant (*Solanum melongena* L.) using Mahalanobis D^2 analysis, Principal Component Analysis (PCA) and Start Codon Targeted (SCoT)

Aniket Kumar Verma[1]*, Akhilesh Chandra Mishra[1]*, Vijay Sharma[2], Chandra Mohan Singh[2], Vikas Patel[1], Nida[1] & Shweta Yadav[1]

[1]Department of Vegetable Science, Banda University of Agriculture and Technology, Banda 210 001, India
[2]Department of Genetics and Plant Breeding, Banda University of Agriculture and Technology, Banda 210 001, India

*Correspondence author email - aniketkumarverma89@gmail.com; acm24680@gmail.com

Received: 04 May 2025; Accepted: 06 July 2025; Available online: Version 1.0: 05 August 2025

Cite this article: Aniket KV, Akhilesh CM, Vijay S, Chandra MS, Vikas P, Nida, Shweta Y. Genetic diversity analysis in eggplant (*Solanum melongena* L.) using Mahalanobis D^2 analysis, Principal component analysis (PCA) and Start Codon Targeted (SCoT). Plant Science Today (Early Access). https:/doi.org/10.14719/pst.9286

Abstract

This study evaluated morpho-molecular diversity among 34 eggplant (*Solanum melongena* L.) accessions using 10 morphological descriptors, 16 quantitative traits and 15 SCoT to parent selection and broadening the breeding base effectively. Mahalanobis D^2 analysis grouped the accessions into five clusters, with the greatest inter-cluster divergence between Clusters V and IV. PCA identified five components (eigenvalues >1) explaining 75.65 % of the total variance, with PC1 contributing 38.3 %. High-performing and genetically diverse accessions such as Kashi Uttam, Kashi Green Round, CHBR-2, BUB-18-12 and BUB-18-27 were identified through PCA biplot. SCoT markers amplified 4-10 loci/primer (mean: 6.13), all polymorphic, with SCoT-20, SCoT-13, SCoT-17 and SCoT-25 being the most informative based on PIC values. UPGMA clustering grouped accessions into three clusters (A:14, B:15, C:5), while population structure analysis identified two sub-populations (A:26, B:10). SCoT analysis revealed significant molecular variability. The lack of unique clustering across morphological and molecular data highlights the importance of integrated approaches. Crosses among divergent clusters are recommended to enhance genetic gain and cultivar development.

Keywords: D^2-statistics; genetic diversity; Mahalanobis; PCA; scot markers; *Solanum melongena*

Introduction

Eggplant (*Solanum melongena* L.; 2*n* = 2x = 24), a member of the Solanaceae family, is an economically and nutritionally significant vegetable crop cultivated worldwide. The Solanaceae family comprises approximately 2,450 species across 95 genera, making it one of the largest and most diverse plant families globally (1). Known as eggplant in India and aubergine in Europe, eggplant thrives in tropical and subtropical regions (2). It is often called the "king of vegetables" in India due to its diverse culinary applications and widespread cultivation (3). Nutritionally, eggplant is a good source of carbohydrates, proteins, fats, dietary fiber, vitamins and minerals. Its health-promoting effects are linked to bioactive compounds like chlorogenic acid and nasunin, which are known for their antioxidant and anticancer properties. In India, over 3,500 eggplant accessions are maintained by institutes such as the National Bureau of Plant Genetic Resources (NBPGR), ICAR-Indian Institute of Vegetable Research (IIVR), which play key roles in germplasm conservation and varietal improvement (4-6).

Genetic diversity is a critical aspect of crop improvement, enabling the development of superior cultivars with enhanced yield, stress tolerance and resistance to biotic and abiotic stresses. Genetic diversity in eggplant has commonly been evaluated through morphological, molecular and biochemical markers. Morphological characterization, based on traits such as plant height, leaf morphology and fruit characteristics, provides initial insights into diversity but is influenced by environmental factors, which may limit its accuracy and reproducibility (7). Quantitative traits like fruit size, shape and yield-related parameters remain essential morphological markers in eggplant breeding (8). Molecular markers, on the other hand, provide a robust and precise method for assessing genetic diversity. Techniques such as AFLP, SSR, ISSR and RAPD have been extensively used in eggplant for marker-assisted selection, DNA fingerprinting and QTL mapping (9). Among these, start codon targeted (SCoT) markers represent a relatively novel PCR-based molecular marker system and specifically targeting conserved regions near the ATG start codon (10, 11). The SCoT markers combine the simplicity of RAPD and ISSR with improved specificity and

reproducibility, making them particularly effective for genetic diversity studies (12-14). Initially applied to crops like peanut, long melon and mango, SCoT markers have demonstrated their utility in diversity assessment and genetic resource characterization in various plant species (15, 16). The evaluation of genetic resources is essential for breeding programs aimed at addressing challenges such as climate change, pest resistance and consumer preferences. Comprehensive diversity studies encompassing accessions from diverse geographic regions provide insights into the genetic base of crops and guide the selection of superior parental lines for hybridization. While previous studies have primarily focused on morphological and biochemical evaluations, integrating SCoT markers offers a novel and detailed perspective on eggplant's genetic variability. The integration of morphological traits and SCoT marker analysis will effectively reveal genetic diversity among eggplant accessions, enabling the identification of genetically distinct lines for use in future breeding programs.

Materials and Methods

Experimental materials

The experiment was conducted at the Vegetable Research Farm, BUAT, Banda, Uttar Pradesh, India with 34 eggplant accessions, comprising released varieties and advanced breeding lines (Table 1, Fig. 1). Selection was based on adaptability to the local agro-climate to ensure accurate phenotypic evaluation. The trial followed a Randomized Block Design (RBD) with three replications. Five plants per accession were randomly selected and tagged for data collection. Standard agronomic practices were followed throughout the cropping season. Border rows were excluded from observations to avoid edge effects.

Collection of data

Data were recorded as per DUS guidelines (17) on 10 botanical and 16 morphological traits across key growth stages and yield traits, including: Days to 50 % flowering (Ch1), plant height at 50 % flowering (cm) (Ch2), number of flowers per cluster (Ch3), fruit setting percentage (Ch4), days to first fruit picking (Ch5), leaf area index (Ch6), number of primary branches per plant (Ch7), fruit length (in cm) (Ch8), fruit circumference (cm) (Ch9), fruit diameter (cm) (Ch10), specific gravity of fruits (g cm^{-3}) (Ch11), total soluble solids (°Brix) (Ch12), plant height at last harvesting (cm) (Ch13), number of fruits per plant (Ch14), average fruit weight (g) (Ch15) and fruit yield per plant (kg) (Ch16). Measurements were taken from five plants per accession and averaged. Traits were analyzed for genetic variability and correlation. Observations were synchronized across accessions to reduce environmental influence and enhance data reliability.

Genomic DNA extraction

Genomic DNA was extracted from 200 mg of young, healthy and pest-free leaf tissue using a modified CTAB protocol based on previous studies (18). Leaf samples were collected at the seedling stage, placed in butter paper bags and transported on ice to the laboratory. The DNA isolation procedure involved the following steps: (i) 2 g of healthy young leaves from a single seedling were ground into fine powder using liquid nitrogen and transferred to 2 mL Eppendorf tubes. (ii) Each tube received 0.7 mL CTAB buffer with 2 % β-mercaptoethanol and was incubated at 65 °C for 45-60 min. (iii) After cooling, 0.7 mL of chloroform: isoamyl alcohol (24:1) was added, gently mixed and incubated at room temperature for 15-20 min. (IV) Samples were centrifuged at 12000 rpm for 10 min and the upper aqueous layer was transferred to a fresh tube. (V) DNA was precipitated by adding 0.7 mL chilled isopropanol, gently mixed and incubated at -20 °C for 15 min. (VI) Precipitated DNA was

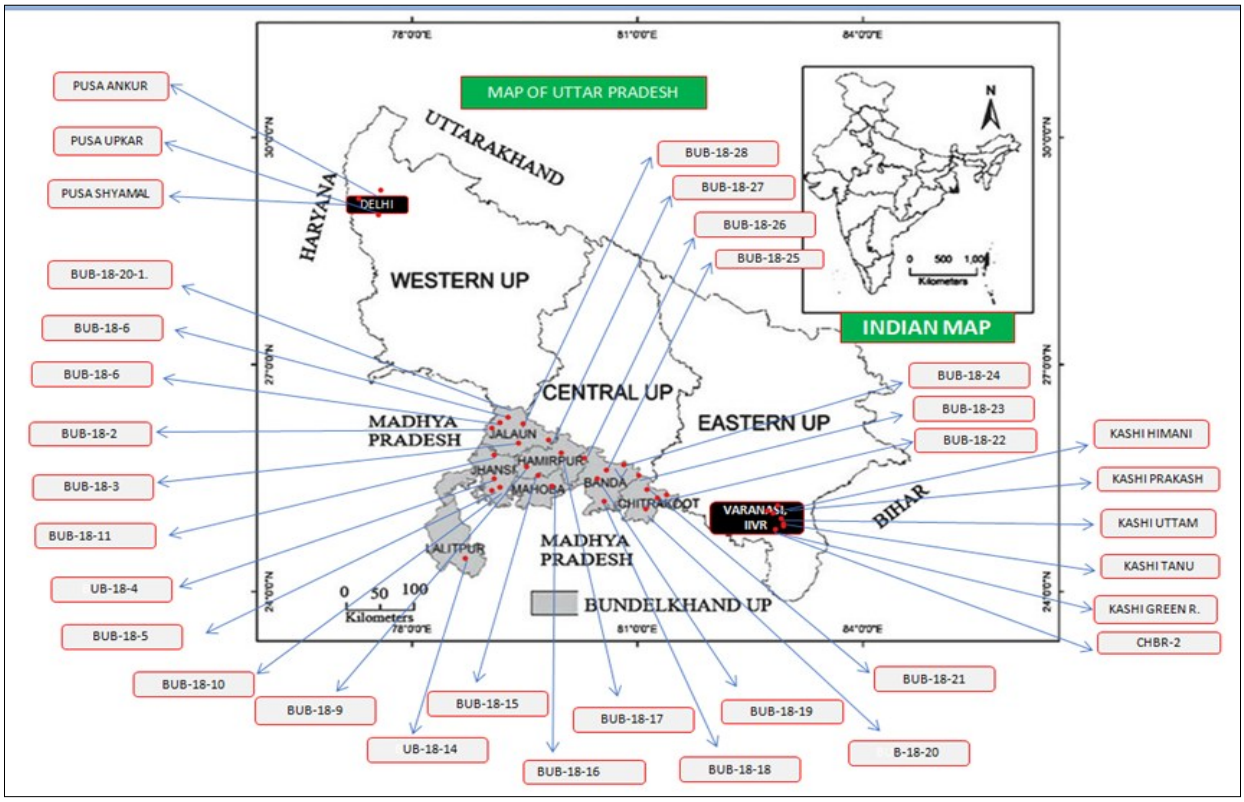

Fig. 1. Map of India indicating eggplant collection sites. Each accession is labelled with red dots on the map.

Table 1. List of eggplant (*Solanum melongena* L.) accessions and their morphological characters used in the current study

S. No.	Accessions	Source	Fruit habit	Fruit color	Fruit shape	Stem color	Spine color	Growth habit	Spine on leaf	Spine on calyx	Spine on both	Spine on whole plant
1	Kashi Himani	IIVR, Varanasi	Solitary	Blue	Long	Green	White	Compact	No	Yes	No	No
2	Kashi Prakash	IIVR, Varanasi	Solitary	Blue	Long	Green	Absent	Spreading	No	No	No	No
3	Kashi Uttam	IIVR, Varanasi	Solitary	Blue	Round	Green	Absent	Spreading	No	No	No	No
4	Kashi Taru	IIVR, Varanasi	Solitary	Blue	Long	Green	Absent	Compact	No	No	No	No
5	Kashi Green Round	IIVR, Varanasi	Solitary	Whitish Green	Round	Green	Absent	Spreading	No	No	No	No
6	Pusa Ankur	IARI, New Delhi	Cluster	Blue	oblong	Blue	Absent	Compact	No	No	No	No
7	Pusa Upkar	IARI, New Delhi	Solitary	Blue	oblong	Blue	Blue	Spreading	No	Yes	No	No
8	CHBR-2	IIVR, Varanasi	Solitary	Blue	Round	Green	White	Spreading	No	Yes	No	No
9	BUB-18-20-1	BUAT, Banda	Solitary	Blue	Long	Green	Absent	Spreading	No	No	No	No
10	BUB-18-6	BUAT, Banda	Solitary	Blue	Round	Green	Blue	Spreading	Yes	Yes	Yes	Yes
11	BUB-18-6	BUAT, Banda	Solitary	Purple	Long	Green	Absent	Spreading	No	No	No	No
12	BUB-18-2	BUAT, Banda	Solitary	Blue	Round	Green	Absent	Compact	No	No	No	No
13	BUB-18-3	BUAT, Banda	Solitary	Green	Round	Green	Absent	Spreading	No	No	No	No
14	BUB-18-11	BUAT, Banda	Cluster	Blue	Round	Green	Absent	Compact	No	No	No	No
15	BUB-18-4	BUAT, Banda	Solitary	Green	Long	Green	Absent	Spreading	No	No	No	No
16	BUB-18-5	BUAT, Banda	Solitary	Blue	oblong	Green	White	Spreading	Yes	Yes	Yes	Yes
17	BUB-18-10	BUAT, Banda	Solitary	Blue	Round	Green	Absent	Spreading	No	No	No	No
18	BUB-18-9	BUAT, Banda	Solitary	Blue	oblong	Green	Absent	Spreading	No	No	No	No
19	Pusa Shyamal	IARI, New Delhi	Solitary	Blue	Long	Green	Absent	Compact	No	No	No	No
20	BUB-18-14	BUAT, Banda	Solitary	Blue	Long	Green	Absent	Compact	No	No	No	No
21	BUB-18-15	BUAT, Banda	Cluster	Blue	oblong	Green	Blue	Spreading	Yes	Yes	Yes	Yes
22	BUB-18-16	BUAT, Banda	Solitary	Blue	oblong	Green	Absent	Spreading	No	No	No	No
23	BUB-18-17	BUAT, Banda	Solitary	Blue	oblong	Green	Absent	Spreading	No	No	No	No
24	BUB-18-18	BUAT, Banda	Solitary	Blue	Round	Green	Absent	Spreading	No	No	No	No
25	BUB-18-19	BUAT, Banda	Solitary	Blue	oblong	Green	Absent	Spreading	No	No	No	No
26	BUB-18-20	BUAT, Banda	Solitary	Blue	Long	Green	Absent	Spreading	No	No	No	No
27	BUB-18-21	BUAT, Banda	Solitary	Blue	oblong	Green	Absent	Spreading	No	No	No	No
28	BUB-18-22	BUAT, Banda	Solitary	Blue	oblong	Green	Absent	Spreading	Yes	No	No	No
29	BUB-18-23	BUAT, Banda	Solitary	Blue	Round	Green	Absent	Compact	Yes	Yes	Yes	Yes
30	BUB-18-24	BUAT, Banda	Solitary	Blue	Long	Green	Absent	Compact	No	No	No	No
31	BUB-18-25	BUAT, Banda	Cluster	Blue	Long	Green	Absent	Spreading	No	No	No	No
32	BUB-18-26	BUAT, Banda	Solitary	Blue	Round	Green	Absent	Spreading	No	No	No	No
33	BUB-18-27	BUAT, Banda	Solitary	Blue	Long	Green	Absent	Spreading	No	No	No	No
34	BUB-18-28	BUAT, Banda	Solitary	Blue	Round	Green	White	Spreading	Yes	Yes	Yes	Yes

pelleted by centrifuging at 12000 rpm for 10 min and the supernatant was removed. (VII) The DNA pellet was rinsed with 70 % ethanol, spun at 12000 rpm for 5-10 min and the supernatant was discarded. (VIII) Residual ethanol was removed by air-drying overnight or incubating at 37 °C for 2 hrs. (IX) The dried DNA pellet was dissolved in 100 µL TE buffer (pH 8.0) overnight at 4 °C. (X) DNA quality was checked by 0.8 % agarose gel electrophoresis with EtBr staining and concentration measured spectrophotometrically; SCoT analysis was performed at Molecular Biology Laboratory, Department of Genetics and Plant Breeding, Banda University of Agriculture and Technology (BUAT), Banda, Uttar Pradesh, India.

Quality assessment of DNA and DNA quantification

DNA concentration and purity were assessed using 1 % agarose gel electrophoresis and a NanoDrop 1000 spectrophotometer. The gel was prepared by dissolving 1 g agarose in 100 mL 1X TBE buffer (45 mM Tris, 45 mM boric acid, 1 mM EDTA), heated, cooled and stained with ethidium bromide (0.5 µg mL^{-1}). After solidifying, DNA samples and standards were loaded and electrophoresis was run at 75 V for 2 hours. DNA bands were visualized using a gel documentation system and band intensity was compared to standards. For precise quantification, 1 µL of each DNA sample was analyzed with the NanoDrop, using TE buffer as a blank and results were recorded in ng Ml^{-1}.

Selection of SCoT primers and synthesis or SCoT primer optimization

Start codon targeted (SCoT) markers are relatively new molecular markers and hundreds of primer sets have been developed, though the exact number continues to expand as research progresses. The 15 specific SCoT primer (Table 2) sets used in this study were selected based on their proven ability to generate polymorphic and reproducible bands in previous studies, thereby ensuring their utility in assessing genetic diversity. Additionally, these primers were chosen for their compatibility with the target species and experimental conditions, such as their optimal annealing temperature of 52 ° C, ensuring reliable amplification of DNA across the 34 eggplant accessions. This selection maximized the likelihood of detecting meaningful genetic variation.

PCR amplification and electrophoresis

Fifteen SCoT primers (Table 2) were used to screen 34 eggplant accessions for polymorphism. PCR was performed in a 25 µL mix containing 50 ng DNA, PCR buffer, dNTPs, $MgCl_2$, primers and Taq polymerase. The cycling conditions were: 94 °C for 5 min; 35 cycles of 94 °C for 1 min, annealing at 48-60 °C for 1 min and 72 °C for 2 min; with a final extension at 72 °C for 7 min. Amplified products were separated on 2 % agarose gel with 1X TBE buffer and visualized under UV light.

A 2 % agarose gel with ethidium bromide was prepared and poured to a thickness of 0.5 cm. After solidification, the gel was placed in an electrophoresis chamber with 1X TAE buffer. DNA samples mixed with loading dye were loaded into wells. Electrophoresis was run at constant voltage until the dye migrated sufficiently. PCR products were visualized under UV light using a gel documentation system.

Statistical analysis

DUS data analysis: Binary data from 10 morphological descriptors were analyzed using NTSYS-pc software (version 2.02; Applied Biostatistics Inc., Port Jefferson, New York, USA) to construct a dendrogram. The dendrogram was generated to depict genetic relationships among the accessions. The diversity index (DI) for each descriptor was calculated using the formula:

$$H = \sum_{i=1}^{R} Pi \ In \ Pi \qquad \text{(Eqn. 1)}$$

where Pi is the proportion of the i phenotype and R represents the total number of observed phenotypes for the descriptor.

Morphological diversity analysis

Quantitative data for 16 morphological characters, excluding 10 DUS descriptors were analyzed using Mahalanobis D^2 statistics (19) and accessions were grouped into clusters following Tocher's method (20). Principal component analysis (PCA) was conducted using R software (R Foundation for Statistical Computing, Vienna, Austria), employing the EIGEN procedure to calculate correlation coefficients between the characters (21). The grouping pattern of the 34 accessions was visualized through PCA to identify key contributors to phenotypic variability.

Table 2. SCoT markers used in the study

S. No.	Primer code	Their sequences (5' 3)	Annealing temperature (°C)
1	SCoT-12	ACGACATGGCGACCAACG	52
2	SCoT-13	ACGACATGGCGACCATCG	52
3	SCoT-14	ACGACATGGCGACCACGC	52
4	SCoT-15	ACGACATGGCGACCGCGA	52
5	SCoT-16	ACCATGGCTACCACCGAC	52
6	SCoT-17	ACCATGGCTACCACCGAG	52
7	SCoT-18	ACCATGGCTACCACCGCC	52
8	SCoT-20	ACCATGGCTACCACCGCG	52
9	SCoT-21	ACGACATGGCGACCCACA	52
10	SCoT-23	CACCATGGCTACCACCAG	52
11	SCoT-25	ACCATGGCTACCACCGGG	52
12	SCoT-26	ACCÄTGGCTACCACCGTC	52
13	SCoT-27	ACCATGGCTACCACCGTG	52
14	SCoT-28	CCATGGCTACCACCGCCA	52
15	SCoT-31	CCATGGCTACCACCGCCT	52

Molecular data analysis

A binary data matrix was created for SCoT marker analysis by scoring distinct and clear DNA bands as present (1) or absent (0) from agarose gel results. Genetic relationships were analyzed using NTSYS-pc software with the UPGMA method, employing Jaccard's similarity coefficients to generate a dendrogram (22). Population structure was assessed with STRUCTURE 2.3.4 software using a Bayesian clustering approach under an admixture model with correlated allele frequencies. The optimal number of clusters (K) was determined based on ΔK values (23).

Results

Diversity at the morphological level

Morphological diversity and characterization

The evaluation of morphological diversity among the 34 eggplant accessions (Supplementary Fig. 1) revealed significant variation in key traits. Consistent and stable scores were observed across all 10 morphological descriptors, reflecting the reliability of these traits in characterizing the accessions. Notable variability was documented in important attributes such as fruit shape, stem color, leaf size and flower characteristics (Supplementary Fig. 2). The Shannon diversity index, calculated for the 10 traits, ranged from 0.22 (stem color) to 1.10 (fruit shape), with a mean of 0.50 and a range of 0.88 (Table 3), highlighting moderate overall diversity. Among the descriptors, fruit shape exhibited the highest diversity index (1.10), underscoring its significance in differentiating accessions. In contrast, stem color showed limited variability, with the lowest diversity index (0.22), indicating a relatively uniform trait across the studied accessions. These findings emphasize the importance of morphological characterization in identifying and categorizing genetic variation within eggplant accessions. The considerable variability observed in traits such as fruit shape provides a valuable resource for selecting diverse accessions for breeding programs aimed at trait improvement and heterosis exploitation.

Analysis of variance

The ANOVA (Table 4) for 16 morphological traits of eggplant accessions revealed highly significant differences among the accessions, indicating substantial variability. The coefficient variation ranged from 3.05 % to 6.97 %, with the highest variation observed in average fruit weight (6.97 %), followed by fruit setting percentage (6.73 %), plant height at 50 % flowering (6.59 %), number of primary branches per plant (6.58 %), number of flowers per cluster (6.39 %), fruit yield per plant (6.09 %) and fruit circumference (6.32 %) (Supplementary table 1).

Analysis of genetic divergence based on morphological markers using Mahalanobis D2 statistics

The Clustering of 34 eggplant accessions based on D^2 values using Tocher's method resulted in five distinct clusters across 16 traits. Cluster composition is detailed in Table 5 and illustrated in Fig. 2. The results indicated that a maximum number of 13 accessions appeared in cluster III followed by 9 accessions in cluster IV, 7 accessions in cluster II, 4 accessions in cluster V and 1 accession in cluster I. This indicated that there is a presence of diversity among the 34 accessions studied. Average intra-cluster distance values ranged from 2.572 (cluster IV) to 3.904 (cluster I) indicating minimum and maximum heterogeneity among the accessions within the cluster. Maximum intra-cluster genetic distance was observed in cluster V (3.904) followed by cluster II (2.886), cluster I (2.875), cluster III (2.681) and cluster IV (2.572) (Table 5). The maximum inter-cluster distance was observed between clusters V and IV (6.872) followed by clusters IV and I (5.292), clusters V and II (4.862) clusters IV and II (4.676) while minimum inter-cluster distance was noticed between clusters II and I (2.856) (Table 6).

Table 3. Diversity indices of 10 morphological descriptors in eggplant accessions (pooled)

Trait	Class or scale	Frequency	Relative frequency (%)	Diversity index
Fruit habit	Solitary	30	88.24	0.36
	Cluster	4	11.76	
Fruit color	Blue	32	94.12	0.26
	Whitish Green	1	2.94	
	Purple	1	2.94	
	Green	0	0.00	
Fruit shape	Long	12	35.29	1.10
	Round	12	35.29	
	oblong	10	29.41	
Stem color	Green	32	94.12	0.22
	Blue	2	5.88	
Spine color	White	4	11.76	0.65
	Absent	27	79.41	
	Blue	3	8.82	
Growth habit	Compact	9	26.47	0.58
	Spreading	25	73.53	
Spine on leaf	Yes	6	17.65	0.45
	No	28	82.35	
Spine on calyx	Yes	8	23.53	0.55
	No	26	76.47	
Spine on both	Yes	5	14.71	0.42
	No	29	85.29	
Spine on whole plant	Yes	5	14.71	0.42
	No	29	85.29	
Mean				0.50
Minimum				0.22
Maximum				1.10
Range				0.88

Table 4. ANOVA, means, range and coefficient of variation for various morphological traits in 34 eggplant accessions (pooled)

Characters	Mean squares			Means	Range		Coefficient of variation (%)
Degree of freedom (DF)	Replication (2)	Treatment (33)	Error (66)		Min.	Max.	
Days to 50 % flowering	14.66	87.33	8.09	58.25	50.00	69.67	4.88
Plant height at 50 % flowering (cm)	1.55	63.53	6.51	38.73	28.08	45.73	6.59
Number of flowers per cluster	0.01	45.62	0.05	3.62	1.20	24.09	6.39
Fruit setting percentage	2.02	1265.07	13.00	53.58	18.94	92.41	6.73
Days to first fruit picking	12.11	110.60	9.22	73.14	62.49	90.00	4.15
Leaf area index	0.012	1.982	0.010	1.85	0.53	4.22	5.35
Number of primary branches per plant	3.42	8.94	0.31	8.41	5.98	11.86	6.58
Fruit length (cm)	0.04	41.53	0.30	11.48	6.49	22.32	4.73
Fruit circumference (cm)	1.48	74.03	1.14	16.92	10.73	28.53	6.32
Fruit diameter (cm)	0.11	7.97	0.10	5.51	3.51	9.63	5.60
Specific gravity of fruits (g cm^{-3})	0.006	0.077	0.001	0.90	0.50	1.30	3.36
Total soluble solids (°Brix)	0.002	1.984	0.019	4.55	3.18	5.60	3.05
Plant height at last picking (cm)	19.51	217.28	11.76	75.93	58.98	94.87	4.52
Number of fruits per plant	6.26	1322.01	1.63	28.78	11.02	138.45	4.44
Average fruit weight (g)	16.87	10961.84	66.03	116.65	15.67	251.09	6.97
Days to 50 % flowering	0.093	1.347	0.026	2.66	1.46	3.57	6.09

Table 5. Clustering pattern of 34 accessions of eggplant based on Mahalanobis D^2 statistics (pooled)

S. No.	Clusters	Number of accessions	Name of clustered accessions
1	I	1	BUB-18-15
2	II	7	Pusa Ankur, Pusa Shyamal, BUB-18-14, BUB-18-21, Kashi Prakash, BUB-18-19, BUB-18-25
3	III	13	BUB-18-20-1, Kashi Himani, BUB-18-11, Kashi Taru, BUB-18-6, BUB-18-3, BUB-18-10, BUB-18-23, BUB-18-24, BUB-18-26, BUB-18-9, BUB-18-20, BUB-18-2
4	IV	9	BUB-18-22, BUB-18-17, BUB-18-4, Pusa Upkar, Kashi Green Round, Kashi Uttam, CHBR-2, BUB-18-27, BUB-18-12
5	V	4	BUB-18-5, BUB-18-16, BUB-18-18, BUB-18-28

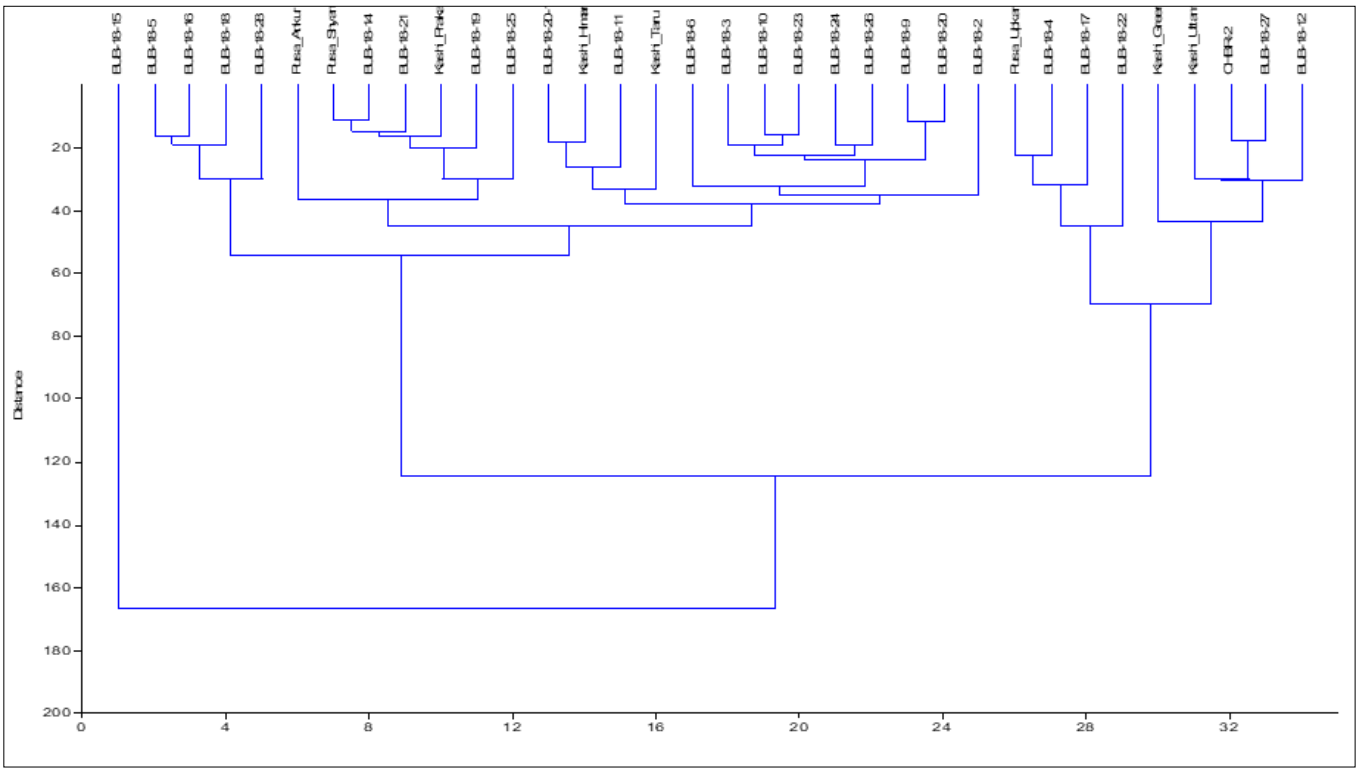

Fig. 2. Dendrogram of 34 eggplant accessions on 16 morphological traits using Mahalanobis D^2 clustering pattern.

Table 6. Average intra (diagonal) and inter-cluster (above diagonal) distances of five clusters (pooled)

Cluster	I	II	III	IV	V
I	2.875				
II	2.856	2.886			
III	3.478	2.995	2.681		
IV	5.292	4.676	3.989	2.572	
V	4.310	4.862	3.924	6.872	3.904

The cluster mean for 16 quantitative characters of 34 accessions was presented in Table 7. Cluster I observed that cluster mean value for minimum days to 50 % flowering (51.98), days to first fruit picking (66.59) and maximum fruit setting percentage (64.62 cm). Cluster II showed maximum cluster mean for fruit length (16.43 cm), specific gravity of fruit (0.99 g cm^{-3}) and fruit yield per plant (3.30 kg). Cluster IV promising for leaf area index (2.53), number of primary branches per plant (9.17), fruit circumference (26.20 cm), fruit diameter (8.31 cm), plant height at 50 % flowering (40.69 cm) and average fruit weight (233.58 g). Cluster V showed maximum cluster mean values for the number of flowers per cluster (8.07), total soluble solids (5.26), plant height at last picking (88.26 cm) and number of fruits per plant (49.09).

Principal Component Analysis (PCA)

The estimated Eigenvector values, percentage of variance, cumulative percentage (Table 8) and principal component of 34 eggplant accessions for 16 traits are presented in Table 9 and The PC biplot (Fig. 3.) depicts the accessions distribution and factors, as well as the difference in traits between the principal components, demonstrating the way these traits contributed to eggplant accessions variation. The first five principal components (PCs) accounted for 75.65 % of the total observed variation. Among the five principal components, the first principal component (PC) contributed 27.48 % of the total variation, this was due to the positive factor loading effect of all studied traits mainly for plant height at 50 % flowering (0.64), days to 50 % flowering (0.53), fruit yield per plant (0.03) and average fruit weight (0.02).

Table 7. Cluster mean for 16 characters in accessions of eggplant (pooled)

Clusters	DFPF	PHFPF	NFPC	FSP	DFFP	LAI	NPBPP	FL	FC	FD	SGF	TSS	PHLP	NFPP	AFV	FYPP
I	51.98	33.70	1.79	64.62	66.59	1.63	7.48	12.85	15.60	5.30	0.88	4.39	72.90	31.51	98.90	2.97
II	58.24	39.65	3.40	54.60	72.62	2.38	8.73	16.43	14.68	4.71	0.99	4.18	78.19	29.87	118.34	3.30
III	61.63	40.68	3.56	52.96	77.36	1.79	8.65	9.23	16.58	5.46	0.92	4.39	72.96	23.48	102.23	2.24
IV	63.17	40.69	2.16	51.83	78.96	2.53	9.17	10.77	26.20	8.31	0.81	4.91	70.60	13.75	233.58	3.15
V	55.35	38.61	8.07	39.68	68.77	0.86	8.04	7.82	13.31	4.22	0.84	5.26	88.26	49.09	51.04	1.71

DFPF: Days to 50 % flowering; PHFPF: plant height at 50 % flowering (cm); NFPC: number of flowers per cluster; FSP: fruit setting percentage; DFFP: days to first fruit picking; LAI: leaf area index; NPBPP: number of primary branches per plant; FL: fruit length (cm); FC: fruit circumference (cm); FD: fruit diameter (cm); SGF: specific gravity of fruits (g cm^{-3}); TSS: total soluble solids (°Brix); PHLP: plant height at last picking (cm); NFPP: number of fruits per plant; AFV: average fruit weight (g); FYPP: fruit yield per plant (kg).

Table 8. Eigenvalues and important variation percentage for different traits of eggplant accessions

S. No.	Principle component	Eigenvalue	Importance variability (%)	Accumulated variability %
1	PC1	4.40	27.48	27.48
2	PC2	2.79	17.43	44.91
3	PC3	2.11	13.19	58.09
4	PC4	1.70	10.64	68.73
5	PC5	1.11	6.92	75.65
6	PC6	0.94	5.85	81.50
7	PC7	0.79	4.95	86.45
8	PC8	0.66	4.15	90.60
9	PC9	0.44	2.73	93.33
10	PC10	0.41	2.58	95.91
11	PC11	0.27	1.68	97.59
12	PC12	0.18	1.11	98.70
13	PC13	0.11	0.70	99.40
14	PC14	0.05	0.30	99.70
15	PC15	0.03	0.17	99.86
16	PC16	0.02	0.14	100.00

DFPF: Days to 50 % flowering; PHFPF: plant height at 50 % flowering (cm); NFPC: number of flowers per cluster; FSP: fruit setting percentage; DFFP: days to first fruit picking; LAI: leaf area index; NPBPP: number of primary branches per plant; FL: fruit length (cm); FC: fruit circumference (cm); FD: fruit diameter (cm); SGF: specific gravity of fruits (g cm^{-3}); TSS: total soluble solids (°Brix); PHLP: plant height at last picking (cm); NFPP: number of fruits per plant; AFV: average fruit weight (g); FYPP: fruit yield per plant (kg).

Table 9. Principal component analysis for different traits of eggplant accessions

S. No.	Principle component characters	Factors loading				
		PC1	PC2	PC3	PC4	PC5
1	DFPF	0.53	0.25	-0.56	-0.06	0.60
2	PHFPF	0.64	0.60	0.24	-0.45	0.60
3	NFPC	-0.03	-0.06	0.52	0.35	-0.16
4	FSP	0.37	0.37	0.43	0.51	0.36
5	DFFP	-0.16	0.19	-0.02	0.34	-0.16
6	LAI	-0.02	-0.19	-0.21	0.05	0.04
7	NPBPP	-0.02	0.16	-0.09	0.38	0.10
8	FL	-0.33	0.46	-0.16	0.20	-0.20
9	FC	-0.01	-0.26	0.00	0.11	0.07
10	FD	-0.05	-0.22	0.16	0.19	-0.01
11	SGF	0.11	-0.12	-0.22	0.24	0.15
12	TSS	-0.04	-0.03	0.11	0.01	0.03
13	PHLP	-0.09	-0.04	-0.04	-0.05	0.03
14	NFPP	-0.13	0.01	-0.01	-0.02	0.15
15	AFV	0.02	0.01	-0.08	-0.01	-0.02
16	FYPP	0.03	0.00	-0.01	0.01	-0.02

DFPF: Days to 50 % flowering; PHFPF: plant height at 50 % flowering (cm); NFPC: number of flowers per cluster; FSP: fruit setting percentage; DFFP: days to first fruit picking; LAI: leaf area index; NPBPP: number of primary branches per plant; FL: fruit length (cm); FC: fruit circumference (cm); FD: fruit diameter (cm); SGF: specific gravity of fruits (g cm^{-3}); TSS: total soluble solids (°Brix); PHLP: plant height at last picking (cm); NFPP: number of fruits per plant; AFV: average fruit weight (g); FYPP: fruit yield per plant (kg).

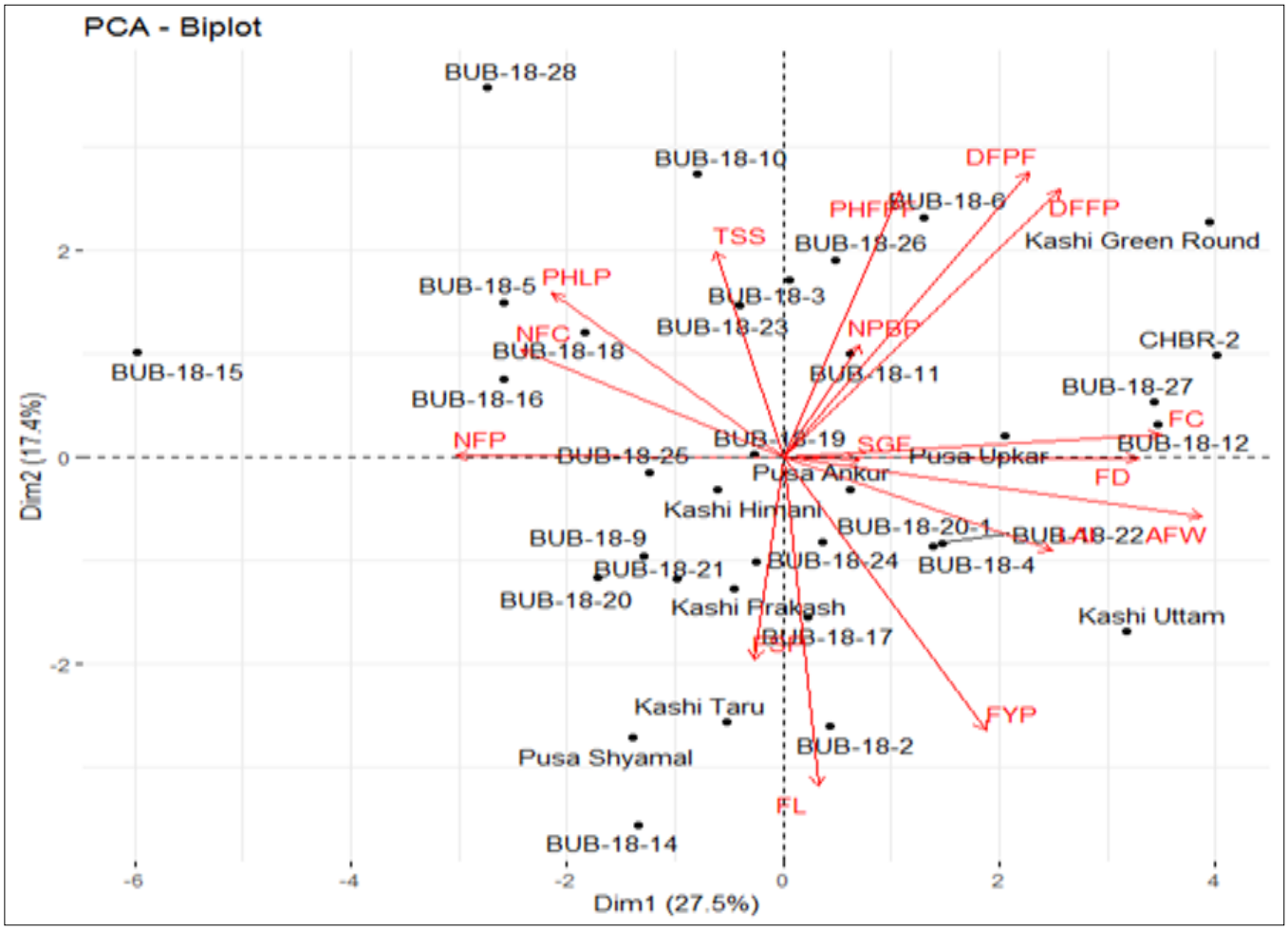

Fig. 3. Biplot of 34 eggplant accessions.

DFPF: Days to 50 % flowering; PHFPF: plant height at 50 % flowering (cm); NFC: number of flowers per cluster; FSP: fruit setting percentage; DFFP: days to first fruit picking; LAI: leaf area index; NPBP: number of primary branches per plant; FL: fruit length (cm); FC: fruit circumference (cm); FD: fruit diameter (cm); SGF: specific gravity of fruits (g cm^{-3}); TSS: total soluble solids (°Brix); PHLP: plant height at last picking (cm); NFP: number of fruits per plant; AFV: average fruit weight (g); FYP: fruit yield per plant (kg).

The second principal component shared about 17.43 % of the total variation and the major positive contributing was fruit length (0.46). The third PC accounted for 13.19 % of the total variations for number of flowers per cluster (0.52) and total soluble solids (0.11). The fourth principal component contributed 10.64 % of total variation for significant trait to fruit setting percentage (0.51), number of primary branches per plant (0.38), days to first fruit picking (0.34), specific gravity of fruits (0.24), fruit diameter (0.19) Fruit circumference (0.11) and Leaf Area Index (0.04). The fifth principal component shared 6.92 % of the total variation for positively significant Number of fruits per plant of fruits per plant (0.15) and plant height at last picking (0.03). The top principal component scores (PC scores) for all the traits were estimated in five principal components and are presented in Table 10.

Principal component scores of eggplant accessions

The principal component scores (PC1, PC2, PC3, PC4 and PC5) exhibited both positive and negative values (Table 11), providing a basis for developing precise selection indices based on the variability explained by each component. High scores in a specific principal component indicate superior performance of accessions for the associated traits. Accessions such as Kashi Uttam, Kashi Green Round, CHBR-2, BUB-18-12 and BUB-18-27 (Table 12) showed the highest scores in PC1, reflecting significant variability in traits like days to 50 % flowering, plant

Table 10. Interpretation of rotated component matrix for the traits having maximum values in each PC

PC 1	PC 2	PC 3	PC 4	PC 5
Plant height at 50 % flowering	Fruit length	Number of flowers per cluster	Fruit setting percentage	Number of fruits per plant
Days to 50 % flowering		Total soluble solids	Number of primary branches per plant	Plant height at last picking
Fruit yield per plant			Days to first fruit picking	
Average fruit weight			Specific gravity of fruits	
			Fruit diameter	
			Fruit circumference	
			Leaf area index	

Table 11. PC scores of eggplant accessions

S. No.	Accessions	PC 1	PC 2	PC 3	PC 4	PC 5
1	Kashi Himani	9.010	11.666	-4.605	11.089	3.866
2	Kashi Prakash	9.157	10.729	-4.384	12.255	6.836
3	Kashi Uttam	12.734	10.317	-1.196	10.023	5.437
4	Kashi Taru	9.082	9.447	-6.062	12.844	5.523
5	Kashi Green Round	13.480	14.228	-3.403	11.303	4.854
6	Pusa Ankur	10.219	11.674	-3.335	11.844	6.599
7	Pusa Upkar	11.625	12.178	-3.091	11.892	5.410
8	CHBR-2	13.552	12.948	-2.931	11.735	4.839
9	BUB-18-20-1	9.956	11.176	-4.400	11.415	5.658
10	BUB-18-6	10.876	14.271	-4.320	12.518	3.591
11	BUB-18-12	13.024	12.291	-1.443	10.051	6.285
12	BUB-18-2	10.021	9.410	-0.535	8.146	5.239
13	BUB-18-3	9.658	13.674	-4.155	10.899	5.612
14	BUB-18-11	10.209	12.961	-3.897	10.227	6.032
15	BUB-18-4	10.971	11.128	-5.121	11.190	6.133
16	BUB-18-5	7.050	13.458	-3.947	8.696	5.145
17	BUB-18-10	8.813	14.683	-4.019	10.779	6.921
18	BUB-18-9	8.331	11.040	-3.581	10.588	4.008
19	Pusa Shyamal	8.232	9.301	-4.152	11.267	5.645
20	BUB-18-14	8.277	8.468	-3.453	12.644	4.025
21	BUB-18-15	3.719	12.976	1.791	13.955	5.208
22	BUB-18-16	7.048	12.718	-4.677	8.556	5.648
23	BUB-18-17	9.821	10.452	-3.306	9.445	7.059
24	BUB-18-18	7.801	13.170	-3.973	8.914	5.438
25	BUB-18-19	9.335	12.001	-2.580	10.339	5.879
26	BUB-18-20	7.914	10.829	-3.182	9.563	6.102
27	BUB-18-21	8.634	10.812	-4.138	10.774	6.730
28	BUB-18-22	11.054	11.159	-3.975	10.414	4.377
29	BUB-18-23	9.204	13.430	-2.292	10.121	4.486
30	BUB-18-24	9.355	10.980	-4.122	11.737	4.263
31	BUB-18-25	8.378	11.838	-3.679	11.141	6.324
32	BUB-18-26	10.080	13.863	-3.692	10.480	5.435
33	BUB-18-27	12.980	12.506	-1.855	11.110	7.262
34	BUB-18-28	6.897	15.506	-4.489	10.189	6.111

Table 12. Selected eggplant accessions in each principal component

S. No.	PC 1	PC 2	PC 3	PC 4	PC 5
1	Kashi Uttam	BUB-18-28	BUB-18-15	BUB-18-14	BUB-18-17
2	Kashi Green Round	BUB-18-10		BUB-18-15	BUB-18-27
3	CHBR-2	Kashi Green Round		BUB-18-6	
4	BUB-18-12	BUB-18-6		Kashi Taru	
5	BUB-18-27			Kashi Prakash	

height at 50 % flowering, average fruit weight and fruit yield per plant, which are key yield-related attributes. The highest PC score of BUB-18-28 followed by BUB-18-10, Kashi Green Round and BUB-18-6 in PC2 was mainly related to fruit length. The highest PC score was obtained by BUB-18-15 in PC3 for characters namely number of flowers per cluster and total soluble solids. In PC4, high PC scores were recorded for characters viz., fruit setting percentage, days to first fruit picking, leaf area index, number of primary branches per plant, fruit circumference and specific gravity of fruits by the accessions BUB-18-14, BUB-18-15 BUB-18-6 Kashi Taru and Kashi Prakash. The highest PC score was obtained by BUB-18-17 and BUB-18-27 in PC5 for character's plant height at last picking and number of fruits per plant.

Analysis of genetic diversity based on molecular markers

The SCoT primers were used for molecular marker analysis among the 34 eggplant accessions. Among 15 SCoT primers, all primers showed polymorphism ranging from 11.11 % to 100 %. These 15 primers produced 92 bands out of 34 eggplant accessions, of which 49 were polymorphic (Table 13). The amplicons (bands) in SCoT 25 were 4 and SCoT 31 was 10 and the average was 6.13. The amplicons (bands) in SCoT-25 were 4 and SCoT-31 was 10, averaging 6.13. The polymorphic amplicons of the primer ranged from 1 (SCoT-16, SCoT-21, SCoT-26 and SCoT-28) to 8 (SCoT-13 in Fig. 4.) and averaged 3.26. The percentage of polymorphic bands varied from 11.11 % (SCoT-28) to 100 % (SCoT-13, SCoT-17, SCoT-20 and SCoT-25), averaging 55.02 %. Some other primers amplifying a higher number of polymorphic amplicons and thus resulting in higher levels of polymorphic loci were SCoT-27 (80.00 %) and SCoT-12 (60.00 %). The PIC value for the SCoT marker ranged from 0.11 for SCoT 26 and SCoT 28 (lowest value) to 0.38 for SCoT-20 (highest value). SCoT-20, SCoT-13, SCoT-17 and SCoT-25 recorded the highest PIC value 0.38, 0.36, 0.34 and 0.31 followed by SCoT-27 (0.30) and SCoT-12 (0.25), respectively. SCoT marker bands are like RAPD profiles and the size of the amplified product ranges from 200 to 1400 bp (Supplementary plates 1).

Cluster analysis based on molecular markers

Cluster analysis of 34 eggplant accessions was conducted using Jaccard's similarity coefficient and UPGMA in NTSYSpc-2.02i software. The similarity coefficient ranged from 0.0 % to 0.1 %. A UPGMA-based dendrogram grouped all accessions into three major clusters. Cluster A comprises 14 accessions, Cluster B comprises 15 accessions and Cluster C comprises 5 accessions. Cluster A is further divided into two sub-clusters, A1 and A2, as shown in Fig. 5 and Table 14. The sub-cluster A1 comprised of 10 accessions viz., Kashi Taru, Pusa Ankur, Pusa Upkar, BUB-18-11, BUB-18-3, BUB-18-6, BUB-18-20-1, BUB-18-12, CHBR-2 and BUB-18-2. Sub-cluster A2 comprised four accessions viz., BUB-18-4, BUB-18-21, BUB-18-14 and BUB-18-16. Cluster B was divided into three sub-clusters: B1, B2 and B3. Sub-cluster B1 consisted of 11 accessions: Kashi Uttam, BUB-18-05, BUB-18-18, Pusa Shyamal, BUB-18-24, BUB-18-25, BUB-18-17, BUB-18-

Table 13. Number of scorable and polymorphic SCoT bands obtained in the PCR amplified DNA of eggplant accessions generated

S.No.	Primer code	Their sequence (5' 3)	Molecular weight range (bp)	Number of amplified loci	Polymorphic molecular weight range (bp)	Number of polymorphic loci	Polymorphic loci (%)	PIC
1	SCoT-12	ACGACATGGCGACCAACG	300-1300	10	400-900	6	60	0.25
2	SCoT-13	ACGACATGGCGACCATCG	300-1400	8	400-1400	8	100	0.36
3	SCoT-14	ACGACATGGCGACCACGC	300-1200	6	400-1200	3	50	0.24
4	SCoT-15	ACGACATGGCGACCGCGA	300-1200	6	300-1200	3	50	0.20
5	SCoT-16	ACCATGGCTACCACCGAC	700-1400	5	1400	1	20	0.16
6	SCoT-17	ACCATGGCTACCACCGAG	700-1300	5	900-1300	5	100	0.34
7	SCoT-18	ACCATGGCTACCACCGCC	300-1200	5	300-1200	2	40	0.24
8	SCoT-20	ACCATGGCTACCACCGCG	200-1400	6	200-1400	6	100	0.38
9	SCoT-21	ACGACATGGCGACCCACA	200-1300	6	200	1	16.67	0.16
10	SCoT-23	CACCATGGCTACCACCAG	600-1400	6	600-1100	2	33.34	0.21
11	SCoT-25	ACCATGGCTACCACCGGG	400-1300	4	400-1300	4	100	0.31
12	SCoT-26	ACCÄTGGCTACCACCGTC	600-1400	7	600	1	14.29	0.11
13	SCoT-27	ACCATGGCTACCACCGTG	700-1400	5	900-1200	4	80	0.30
14	SCoT-28	CCATGGCTACCACCGCCA	300-1300	9	400	1	11.11	0.11
15	SCoT-31	CCATGGCTACCACCGCCT	400-1300	4	400-1300	2	50	0.21
Average				6.13		3.26	55.02	0.24
Total				92		49		
Minimum				4		1	11.11	11
Maximum				10		8	100	0.38

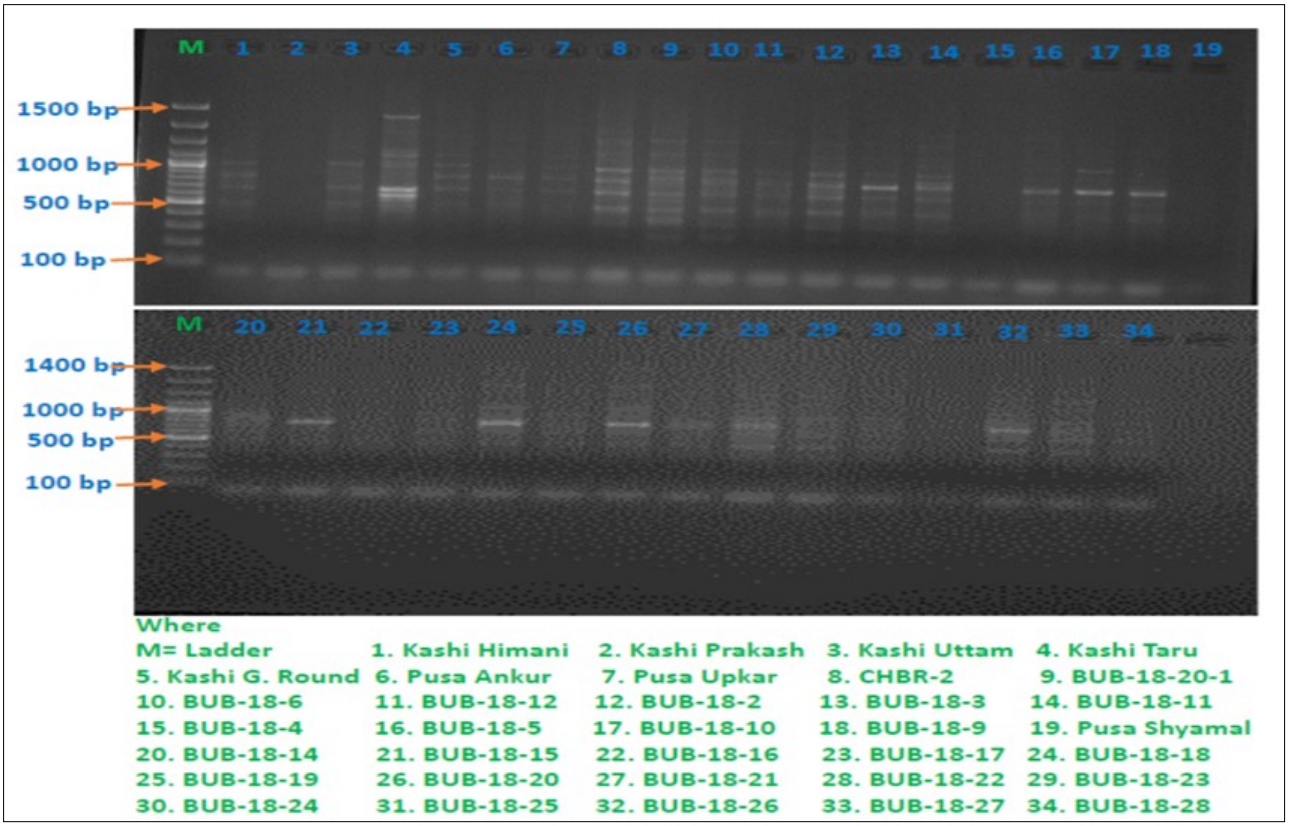

Fig. 4. SCoT amplification profile of primer SCoT-13 in 34 eggplant accessions.

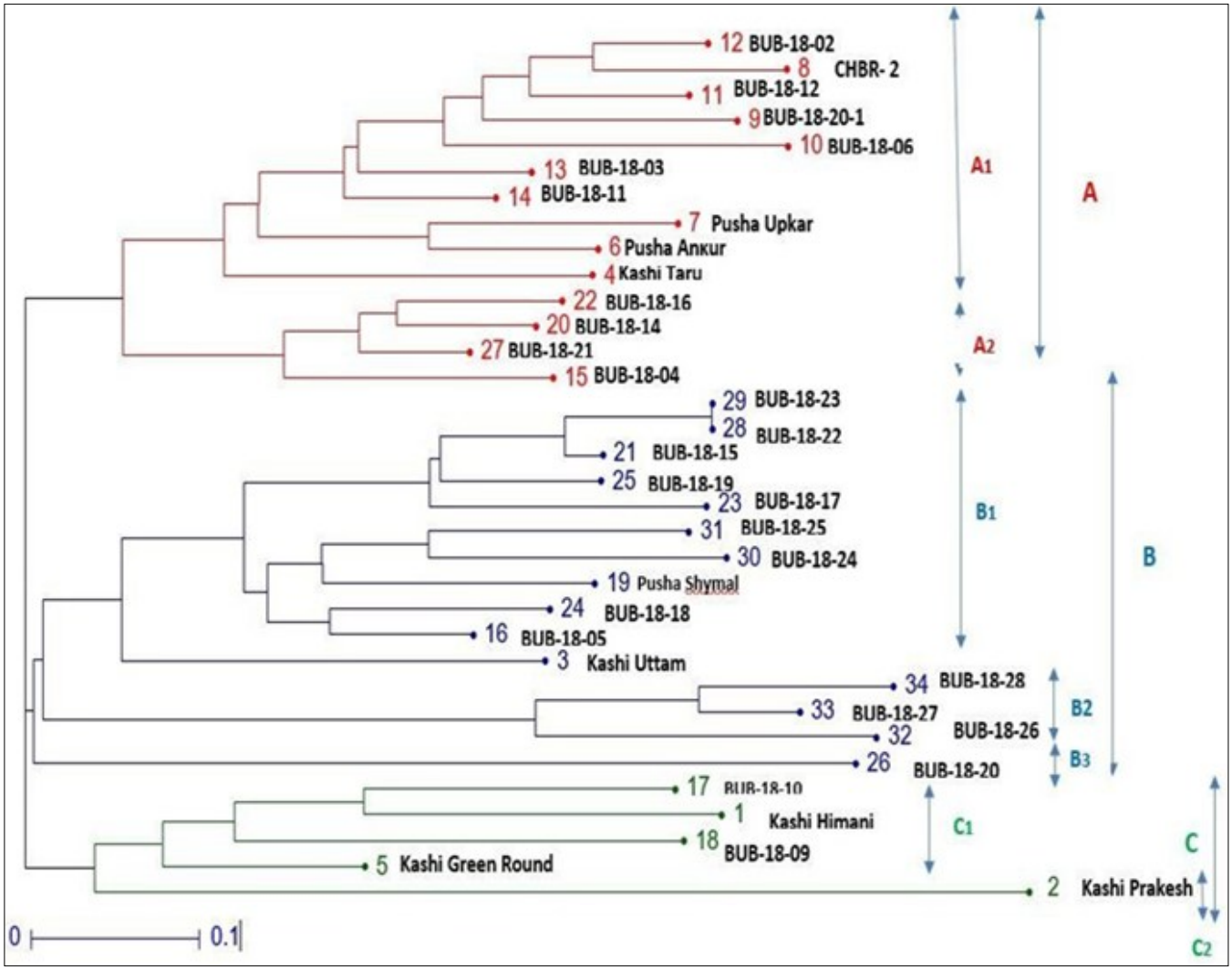

Fig. 5. Dendrogram generated by the Jaccard similarity coefficient, using UPGMA clustering via NTSYS$_{PC}$ software, showed the relationship between 34 different eggplant accessions with the SCoT marker.

Table 14. Clustering of eggplant accessions based on SCoT data (DARWin, 2.02)

Cluster number	Sub-group of clusters	Number of accessions	Accessions
A	A_1	10	Kashi Taru, Pusa Ankur, Pusa Upkar, BUB-18-11, BUB-18-3, BUB-18-6, BUB-18-20-1, BUB-18-12, CHBR-2 and BUB-18-2
	A_2	4	BUB-18-4, BUB-18-21, BUB-18-14 and BUB-18-16
B	B_1	11	Kashi Uttam, BUB-18-05, BUB-18-18, Pusa Shyamal, BUB-18-24, BUB-18-25, BUB-18-17, BUB-18-19, BUB-18-15, BUB-18-22 and BUB-18-23
	B_2	3	BUB-18-26, BUB-18-27 and BUB-18-28
	B_3	1	BUB-18-20
C	C_1	4	Kashi Himani, Kashi Green round, BUB-18-09 and BUB-18-10,
	C_2	1	Kashi Prakash

19, BUB-18-15, BUB-18-22 and BUB-18-23. The sub-cluster B2 comprised 3 accessions viz., BUB-18-26, BUB-18-27 and BUB-18 -28, whereas the accession BUB-18-20 was also noticed in sub-cluster B3. Two sub-clusters, C1 and C2, have been formed from Cluster C. Cluster B1 comprises four accessions viz., Kashi Himani, Kashi Green round, BUB-18-09 and BUB-18-10 and B2 comprises Kashi Prakash accessions only. The diversity of the population can be seen in the three different groups of the population.

Population structure analysis

Bayesian analysis structured the 34 eggplant accessions into two distinct subpopulations, as determined by the sharp peak of Delta K at K = 2. These subpopulations were categorized into pure and admixture groups (Fig. 6.). Accessions with a probability score > 0.70 were classified as pure, while those with < 0.70 were considered admixtures. Subgroup A (red) comprised 26 accessions (76.47 %) and subgroup B (green) included 10 accessions (29.41 %). Subgroup A had 18 accessions score (^{3} 70 %) pure and 6 accessions score (^{3} 70 %) admixture, while subgroup B had 9 accessions score (^{3} 70 %) pure and 1 accessions score (^{3} 70 %) admixtures.

Discussion

Thirty-four eggplant accessions were evaluated using 10 morphological descriptors covering plant, stem, leaf, flower and fruit traits. The observed traits were distinct and consistent, indicating phenotypic stability. These characteristics are valuable for classification and breeding efforts. The genetically stable accessions identified may be suitable for cultivation in various agroclimatic zones (24).

ANOVA revealed significant differences among the 34 eggplant accessions for all traits, indicating substantial phenotypic variability. This diversity is valuable for selecting genetically distinct parents to enhance yield through heterosis. Genetic divergence was measured using Mahalanobis D^2 statistics, showing higher divergence between clusters and closer similarity within clusters. Accessions from distinct clusters are ideal for breeding. Similar findings were reported in previous studies for supporting the effectiveness of this method in diversity assessment (25, 26).

In this study, inter-cluster distances were notably higher than intra-cluster distances, indicating greater genetic variation between clusters than within them. The highest intra-cluster distance was in Cluster V (3.904), suggesting more internal diversity, while the maximum inter-cluster distance was between Clusters V and IV (6.872), showing they are genetically most distinct. These clusters offer strong potential for hybridization to generate diverse and superior progenies (27). Lower inter-cluster distances, such as between Clusters IV and I (5.292), suggest closer genetic relationships. Overall, the findings confirm that accessions within clusters are more genetically similar than those across clusters(28, 29).

Cluster mean analysis is a useful approach for identifying promising accessions in breeding, reducing the need to evaluate less suitable lines. In this study, Cluster I showed the days to 50 % flowering (51.98 days) and days to first fruit picking (66.59 days), along with the highest fruit-setting percentage (64.62 %). Cluster II had the highest fruit length (16.43 cm), specific gravity of fruit (0.99 g/cm^3) and fruit yield per plant (3.30 kg), indicating strong yield potential. Cluster IV stood out for traits like leaf area index (2.53), number

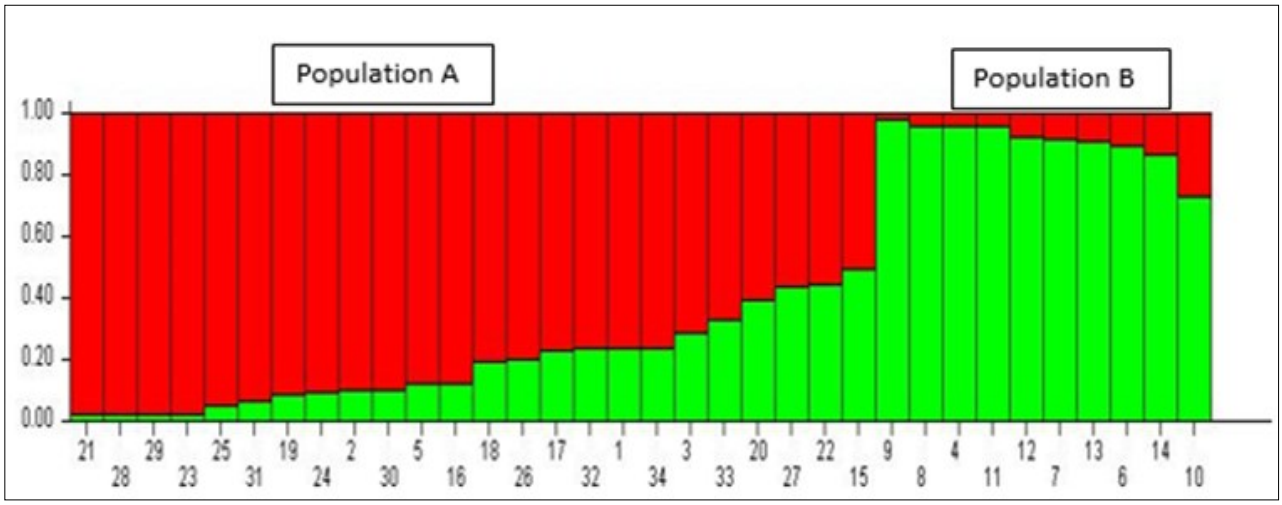

Fig. 6. Analysis of population structure using STRUCTURE 3.2.1 shows a distribution of 34 eggplants in two subpopulations.

of primary branches (9.17), fruit circumference (26.20 cm), fruit diameter (8.31 cm), plant height at 50 % flowering (40.69 cm) and average fruit weight (233.58 g), making it ideal for improving plant structure and fruit quality. Cluster V excelled in number of flowers per cluster (8.07), total soluble solids (5.26 ° Brix), plant height at last picking (88.26 cm) and number of fruits per plant (49.09), making it valuable for enhancing yield and quality through hybridization. Similar findings were reported earlier studies (30-32)

PCA was employed to identify key traits contributing to phenotypic variability for effective trait selection in breeding. Of the 16 components analyzed, five with eigenvalues above 1.00 accounted for 75.65 % of total variation. PC1 (27.48 %) was mainly linked to plant height at 50 % flowering, days to 50 % flowering, fruit yield per plant and average fruit weight. PC2 (17.43 %) was strongly related to fruit length, while PC3 (13.19 %) correlated with number of flowers per cluster and total soluble of solids. PC4 (10.64 %) was linked to traits like fruit setting percentage, days to first fruit picking, number of branches per plant, specific gravity of fruit and fruit diameter. PC5 (6.92 %) was associated with fruit count and plant height at last picking. Similar results were reported in previous studies (33-36).

Accessions like Kashi Uttam, Kashi Green Round, CHBR-2, BUB-18-12 and BUB-18-27 showed high positive scores in PC1, indicating their association with traits such as days to 50 % flowering, plant height at 50 % flowering, average fruit weight and fruit yield per plant. Similarly, BUB-18-14, BUB-18-15, BUB-18-6, Kashi Taru and Kashi Prakash performed well in PC4 for traits like fruit setting percentage, days to first fruit picking, leaf area index, number of primary branches per plant, fruit circumference and specific gravity of fruits. Other accessions also scored high in PC2, PC3 and PC5, highlighting their value for traits like fruit length, number of flowers per cluster, total soluble solids and plant height at last picking. These accessions, due to their favorable trait associations, are ideal for breeding high-yielding cultivars. Comparable observations were reported in former studies (37, 38).

Genetic diversity and population structure are key to crop improvement, as they help breeders identify diverse, superior parents for hybridization (39, 40). While morphological traits have been widely used to assess diversity in eggplants, their sensitivity to environmental factors reduces reliability. Combining morphological and molecular tools offers a more accurate understanding of genetic variation (41). Among molecular markers, SCoT markers are effective due to their dominant nature, high polymorphism and reproducibility in diversity analysis (10, 42)

In this study, 15 SCoT markers were used to assess genetic diversity among 34 eggplant accessions, producing 92 alleles with an average of 3.2 alleles per locus. Four markers (SCoT-13, 17, 20, 25) showed 100 % polymorphism, while SCoT -27 and 12 showed 75-80 %, indicating high variability (43, 44). About half the primers exhibited 50-100 % polymorphic loci, confirming their effectiveness in genetic diversity analysis. These results reveal a broad genetic base, useful for selecting parents in breeding programs targeting yield, stress resistance and improved plant traits.

UPGMA cluster analysis grouped the 34 eggplant accessions into three main clusters and seven sub-clusters. Accessions like BUB-18-20 and Kashi Prakash showed distinct genetic divergence, making them useful for breeding. Similar results were seen in Turkish eggplants, where molecular and morphological data showed minimal correlation (45). Population structure analysis identified two subpopulations, with no link to collection sites or morphological traits, indicating the need for more polymorphic markers. This underscores the value of integrating both approaches to better assess genetic diversity (46).

Combining morphological and molecular analyses provided deeper insights into genetic diversity among eggplant accessions. Distinct genotypes like Kashi Prakash and BUB-18-20 offer strong potential for hybridization to enhance yield and plant traits. Expanding molecular marker use in future research will help identify key genetic variations. This integrated strategy is essential for selecting diverse parents and strengthening the genetic base of eggplant varieties (47, 48).

Conclusion

This study assessed genetic diversity in 34 eggplant accessions using both morphological traits and SCoT molecular markers. Mahalanobis D^2 analysis grouped accessions into three clusters, while UPGMA-based molecular analysis (via NTSYSpc 2.20) formed three main clusters (A, B, C), further divided into sub-clusters. The clustering patterns differed between methods, likely due to morphological traits reflecting gene expression at specific stages, while molecular data captures overall genomic variation. No accession was consistently distinct in both analyses, highlighting the need for integrated approaches. This combined method offers a more accurate understanding of diversity, supporting its use in eggplant breeding efforts.

Acknowledgements

The authors express their gratitude to the Department of Vegetable Science and the Molecular Biology Laboratory, Department of Genetics and Plant Breeding, Banda University of Agriculture and Technology, Banda, Uttar Pradesh, India, for providing laboratory and field facilities essential for this study.

Authors' contributions

AKV and ACM carried out conceptualization. AKV and ACM carried out methodology. Software by CMS and AKV. Validation by AKV. Formal analysis carried out by AKV and VP. Investigation done by AKV. Resources provided by ACM. Data curation was carried out by AKV, N and SY. Writing original draft preparation by AKV. Writing-review and editing done by CMS and VS. Visualization by VS. Supervision by ACM. All authors have reviewed and approved the final version of the manuscript for publication.

Compliance with ethical standards

Conflict of Interest: The authors confirm that there are no conflicts of interest.

Ethical isssues: None

References

1. Bhanushree N, Saha P, Tomar BS, Lyngdoh YA, Gopala KS, Gurung B, et al. Genetic analysis and identification of molecular marker linked to the gene for fruit skin colour in eggplant (*Solanum melongena* L). Veg Sci. 2018;45:149-53. https://doi.org/10.61180/4em4mk71
2. Chithra K, Devaraju M, Srinivasa V, Varalakshmi B, Asha AB. Genetic investigation in segregating generation of brinjal (*Solanum melongena* L). Natl Acad Sci Lett. 2022;45:5-8. https://doi.org/10.1007/s40009-021-01070-x
3. Verma AK, Mishra AC, Tripathi PK. Evaluation of brinjal genotypes (*Solanum melongena* L) for growth, yield and quality characters in the Bundelkhand region of UP, India. Int J Plant Soil Sci. 2023;35:1690-9. https://doi.org/10.9734/ijpss/2023/v35i183444
4. Salem N, Msaada K, Hammami M, Limam F, Vasapollo G, Marzouk B. Variation in anthocyanin and essential oil composition and their antioxidant potentialities during flower development of borage (*Borago officinalis* L). Plant Biosyst. 2014;148:444-59. https://doi.org/10.1080/11263504.2013.778349
5. Taher D, Solberg SØ, Prohens J, Chou Y, Rakha M, Wu T. World Vegetable Centre eggplant collection: Origin, composition, seed dissemination and utilization in breeding. Front Plant Sci. 2017;8:1484. https://doi.org/10.3389/fpls.2017.01484
6. Docimo T, Francese G, Ruggiero A, Batelli G, De Palma M, Bassolino L. Phenylpropanoids accumulation in eggplant fruit: Characterization of biosynthetic genes and regulation by a MYB transcription factor. Front Plant Sci. 2016;6:1233. https://doi.org/10.3389/fpls.2015.01233
7. Ullah S, Ijaz U, Shah TI, Najeebullah M, Niaz S. Association and genetic assessment in brinjal. Eur J Biotechnol Biosci. 2014;2:41-5.
8. Hassan I, Jatoi SA, Arif M, Siddiqui SU. Genetic variability in eggplant for agro-morphological traits. Sci Tech Dev. 2015;34 (1):34-40.
9. Kalia RK, Rai MK, Kalia S, Singh R, Dhawan A. Microsatellite markers: an overview of the recent progress in plants. Euphytica. 2011;177:309-34. https://doi.org/10.1007/s10681-010-0286-9
10. Collard BCY, Mackill DJ. Start Codon Targeted (SCoT) polymorphism: A simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant Mol Biol Rep. 2008;27:86-93. https://doi.org/10.1007/s11105-008-0060-5
11. Xiong FQ, Tang RH, Chen ZL, Pan LH, Zhuang WJ. SCoT: A novel gene-targeted marker technique based on the translation start codon. Mol Plant Breed. 2009;7:635-8.
12. Sultana S, Islam MN, Hoque MdE. DNA fingerprinting and molecular diversity analysis for the improvement of brinjal (*Solanum melongena* L) cultivars. J Adv Biotechnol Exp Ther. 2018;1:1-6. https://doi.org/10.5455/jabet.d1
13. Laxman L, Nandi, Saha P, Behera TK, Lyngdoh YA, Munshi AD, et al. Genetic characterisation and population structure analysis of indigenous and exotic eggplant (*Solanum* spp) accessions using microsatellite markers. J Hortic Sci Biotechnol. 2020;96:73-86. https://doi.org/10.1080/14620316.2020.1763211
14. Chauhan H, Sirohi U, Aastha V, Sharma R, Mohit, Chaudhary V, et al. Identification of diverse genotypes in brinjal (*Solanum melongena* L) based on RAPD markers analysis. Ecol Environ Conserv. 2022;28:248-54. http://doi.org/10.53550/EEC.2022.v28i03s.037
15. Luo C, He XH, Chen H, Ou SJ, Gao MP, Brown JS, et al. Genetic diversity of mango cultivars was estimated using SCoT and ISSR markers. Biochem Syst Ecol. 2011;39:676-84.
16. Xiong FQ, Zhong RC, Han ZQ, Jiang J, He LQ, Zhuang WJ, Tang R. Start codon targeted polymorphism for evaluation of functional genetic variation and relationships in cultivated peanut (*Arachis hypogaea* L) genotypes. Mol Biol Rep. 2011;38:3487-94. https://doi.org/10.1007/s11033-010-0459-6
17. Protection of Plant Varieties and Farmers' Rights Authority. Test guidelines applied for all varieties, hybrids and parental lines of brinjal/eggplant (*Solanum melongena* L). Plant Var J. 2009;3 (11):185-96. https://plantauthority.gov.in/sites/default/files/fbrinjal.pdf
18. Murray MG, Thompson WF. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 1980;8(19):4321-5. https://doi.org/10.1093/nar/8.19.4321
19. Mahalanobis PC. On the generalized distance in statistics. Proc Natl Acad Sci India. 1936;2:49-55. https://www.scirp.org/reference/ReferencesPapers?ReferenceID=1649365
20. Rao CR. Advanced statistical methods in biometrical research. New York: John Wiley and Sons Inc; 1952. https://www.scirp.org/reference/referencespapers?referenceid=268780
21. R Software. A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2023. https://www.R-project.org/
22. Rohlf FJ. NTSYS-PC: Numerical taxonomy and multivariate analysis system, version 2.1. Exeter Publications, New York; 2000. https://wwwresearchgatenet/publication/246982444
23. Pritchard JK, Stephens P, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155:945-59. https://doi.org/10.1093/genetics/155.2.945
24. Singh B, Chaubey T, Pandey S, Singh RK, Upadhyay DK, Jha A, et al. Categorization of diverse and stable extant cultivars of brinjal by using pheno-morphometric DUS characters. Indian J Plant Genet Resour. 2023;36(1):1-18. https://doi.org/10.5958/0976-1926.2023.00036.1.01
25. Begum F, Islam AA, Rasul MG, Mian MK, Hossain MM. Morphological diversity of eggplant (*Solanum melongena*) in Bangladesh. Emir J Food Agric. 2013;25:45-51. https://doi.org/10.9755/ejfa.v25i1.4937
26. Sharma A, Sharma S, Kumar N, Rana RS, Sharma P, Kumar P. Morpho-molecular genetic diversity and population structure analysis in garden pea (*Pisum sativum* L) genotypes using simple sequence repeat markers. PLOS ONE. 2022;17(9):e0273499. https://doi.org/10.1371/journal.pone.0273499
27. Kumar S, Rattan P, Sharma JP, Gupta RK. D2 analysis for fruit yield and quality components in tomato (*Lycopersicon esculentum* Mill). Ind J Plant Gen Res. 2010;23:318-20. https://ispgr.in/index.php/ijpgr/article/download/1678/1514
28. Mohanty KK, Mishra H, Barik S. Morphological profiling and assessment of genetic divergence of brinjal (*Solanum melongena* L). J Pharm Phytochem. 2021;10:602-7. http://www.phytojournal.com
29. Dash SP, Singh J, Sharma D, Thakur P. Genetic divergence study in brinjal (*Solanum melongena* L). J Ent Zool Stud. 2020;8:1277-81. https://www.entomoljournal.com
30. Sanga L, Pandey AK, Warade SDL, Hazarika BN, Singh S. Assessment of wild brinjal (*Solanum gilo*) genotypes of North-Eastern region. Int J Curr Microbiol Appl Sci. 2017;6:1451-8. https://doi.org/10.20546/ijcmas.2017.610.171
31. Nand N, Adarsh A, Kumar A, Akhtar S, Kumar R, Ray PK. Morphological characterization of different genotypes of brinjal (*Solanum melongena* L). Int J Curr Microbiol Appl Sci. 2018;7:2218-26. https://doi.org/10.20546/ijcmas.2018.701.267

32. Quamruzzaman AKM, Rashid MA, Ahmad S, Moniruzzaman M. Genetic divergence analysis in eggplant (*Solanum melongena* L). Bangladesh J Agric Res. 2009;34(4):705-12. https://doi.org/10.3329/bjar.v34i4.5845

33. Barik S, Ponnam N, Acharya GC, Singh TH, Dash M, Sahu GS, et al. Genetic variability, character association and diversity studies in brinjal (*Solanum melongena* L). Electron J Plant Breed. 2021;12 (4):1102-10. https://www.ejplantbreeding.org/index.php/EJPB/article/view/3822

34. Chaudhary AK, Yadav GC, Maurya RK, Anjana CS, Prajapati J. Estimates of genetic variability, yield and quality traits of brinjal (*Solanum melongena* L). Int J Environ Clim Chang. 2023;13(9):583-9. https://doi.org/10.9734/ijecc/2023/v13i92273

35. Singh B, Chaubey T, Singh RK, Upadhyay K, Jha A, Pandey S. Genetic variation in phenatic traits of extant varieties of brinjal (*Solanum melongena* L). J Appl Hort. 2024;26(1):107-11. https://doi.org/10.37855/jah.2024.v26i01.20

36. Kousalya R, Praneetha S, Irene, Vethamoni P, Ravichandran V, Iyanar K, et al. Unveiling the genetic variability, character association and principal component analysis for yield and yield contributing traits in brinjal (*Solanum melongena*) genotypes. Indian J Agric Sci. 2024;94(10):1039-44. https://doi.org/10.56093/ijas.v94i10.152248

37. Baraki F, Berhe M. Evaluating the performance of sesame (*Sesamum indicum* L) genotypes in different growing seasons in northern Ethiopia. Int J Agron. 2019;7:1-7. https://doi.org/10.1155/2019/7804621

38. Mukhthambica K, Bisen R, Ramya KT. Principal component analysis for yield and yield-related traits in sesame (*Sesamum indicum* L). Biol Forum Int J. 2023;15(3):227-32.

39. Dar AA, Mahajan R, Lay P, Sharma S. Genetic diversity and population structure of *Cucumis sativus* L by using SSR markers. Biotechnology. 2017;7:307. https://doi.org/10.1007/s41207-024-00713-x

40. Oliveira MM, Sousa LB, Reis MC, Silva Junior ES, Cardoso DBO, Hamawaki OT. Evaluation of genetic diversity among soybean (*Glycine max*) genotypes using univariate and multivariate analysis. Genet Mol Res. 2017;16:1-10. https://pubmed.ncbi.nlm.nih.gov/28613377/

41. Raja WH, Yousuf N, Qureshi I, Sharma OC, Singh DB, Kumawat KL. Morpho-molecular characterization and genetic diversity analysis across wild apple (*Malus baccata*) accessions using simple sequence repeat markers. South Afr J Bot. 2022;145:378-85. https://doi.org/10.1016/j.sajb.2021.08.020

42. Bhandari HR, Bhanu AN, Srivastava K. Assessment of genetic diversity in crop plants - an overview. Plants Agric Res. 2017;7:279-86. https://doi.org/10.15406/apar.2017.07.00255

43. Mohamed A, García-Martínez S, Loumerem M, Carbonell P, Ruiz JJ, Boubaker M. Assessment of genetic diversity among local pea (*Pisum sativum* L) accessions cultivated in the arid regions of Southern Tunisia using agro-morphological and SSR molecular markers. Genet Resour Crop Evol. 2019;66:1189-203. https://doi.org/10.1007/s10722-019-00784-8

44. Sharma R, Dar AA, Mahajan R, Sharma S. Molecular and biochemical characterization of Indian germplasm of *Pisum sativum* L. Proc Natl Acad Sci India Sect B Biol Sci. 2020;90:103-11. https://doi.org/10.1007/s40011-018-01069-3

45. Cericola F, Portis E, Toppino L, Barchi L, Acciarri N, Ciriaci T, et al. The population structure and diversity of eggplant from Asia and the Mediterranean Basin. PLOS ONE. 2013;8:e73702. https://doi.org/10.1371/journal.pone.0073702

46. Kaur A, Kaur N, Bassi G, Kaur N, Dhatt AS. Morphological and molecular markers-based assessment of genetic diversity in eggplant. Indian J Hort. 2020;77(1):116-24. https://doi.org/10.5958/0974-0112.2020.00012.2

47. Knapp S, Vorontsova MS, Prohens J. Wild relatives of the eggplant (*Solanum melongena* L: Solanaceae): New understanding of species names in a complex group. PLOS ONE. 2013;8:12. https://doi.org/10.1371/journal.pone.0057039

48. Nandi LL, Saha P, Behera TK, Lyngdoh YA, Munshi AD, Saha ND, et al. Genetic characterisation and population structure analysis of indigenous and exotic eggplant (*Solanum* spp) accessions using microsatellite markers. J Hortic Sci Biotechnol. 2021;96:73-86. https://doi.org/10.1080/14620316.2020.1763211

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.