UDC 575.630
https://doi.org/10.2298/GENSR2501037R
Original scientific article

GENETIC DIVERSITY AND POPULATION STRUCTURE OF TOMATO (*Solanum lycopersicum* L.) GENOTYPES

V.M. RASHMI[1], Parveen SHARMA[1], Manpreet KAUR[1], Akhilesh SHARMA[1],
Nimit KUMAR[1], Aparajita DWIVEDI[2], Sanjay Kumar SANADYA[2]*

[1]Department of Vegetable Science and Floriculture, CSK Himachal Pradesh Agricultural University, Palampur (Himachal Pradesh)-176062 India
[2]Department of Genetics and Plant Breeding, CSK Himachal Pradesh Agricultural University, Palampur (Himachal Pradesh)-176062 India

[1]Department of Vegetable Science and Floriculture, CSK Himachal Pradesh Agricultural University, Palampur (Himachal Pradesh)-176062 India
[2]Department of Genetics and Plant Breeding, CSK Himachal Pradesh Agricultural University, Palampur (Himachal Pradesh)-176062 India

Rashmi V.M., P. Sharma, M. Kaur, A. Sharma, N. Kumar, A. Dwivedi, S. Kumar Sanadya (2025). *Genetic diversity and population structure of tomato (Solanum lycopersicum L.) genotypes*. - Genetika, Vol 57, No.1, 37-52.

In present study, morphological and molecular markers were used to assess the genetic diversity of 35 tomato genotypes. However, only 26 tomato genotypes were used to carry out clustering analysis due to *Fusarium* wilt disease nine tomato genotypes were died. Based on clustering analysis, tomato genotypes were grouped into twelve clusters whereas cluster III with eight genotypes followed by cluster I with seven genotypes. Maximum inter-cluster distance observed between cluster V and X indicating the existence of wide range of genetic diversity. Molecular diversity and population structure of 35 tomato genotypes were analyzed using polymorphic simple sequence repeat (SSR) markers. Total 74 alleles identified with an average of 2.05 alleles per locus, while the polymorphism information content of the primers ranged from 0.07 to 0.46 with a mean

---

Corresponding author: Sanjay Kumar Sanadya, Department of Genetics and Plant Breeding, CSK Himachal Pradesh Agricultural University, Palampur (Himachal Pradesh)-176062 India, Phone No.: +91 9460037450, E-mail: sanjaypbg94@gmail.com, ORCID: 0000-0003-3328-3329, V.M.Rashmi ORCID:0000-0002-9951-8963, M. Kaur ORCID:0000-0001-6356-2092, A. Sharma ORCID: 0000-0001-6921-411X, N. Kumar ORCID: 0000-0001-9015-1686

value of 0.32. The major allele frequency and heterozygosity ranged from 0.50-0.96 and 0.07-0.52, respectively. The dendrogram clustered genotypes into two main groups A and B with 27 and 8 genotypes, respectively but group A had two sub-clusters. Similarity coefficient ranged from 0.50 to 0.88 with maximum likeness between genotype BT-20-3 (Red egg shaped) & BT-20-3 (Yellow round) and NBHIA-5 & Punjab Upma while, minimum between Swarna Rattan & KSP-1154-5. Population structure revealed three sub-populations with some admixtures. Principal coordinate analysis revealed that the genotypes were uniformly distributed across the two axes in both the plots with 35.44 % of cumulative variation. An understanding of genetic diversity and polymorphism of tomato populations will help in exploiting the tomato genetic resources improvement.

Keywords: Genetic diversity, Polymorphism, Population structure, Tomato, Clustering.

INTRODUCTION

Tomato (*Solanum lycopersicum* L.) belongs to family Solanaceae, and is the second most important vegetable crop grown worldwide. Tomato is considered as a model plant for conducting genetic research due to small genome size (950 Mb), availability of diverse genetic resources and economic importance as compared to other Solanaceae species (THE TOMATO GENOME CONSORTIUM *et al*., 2012). Tomato may be produced in a variety of agro-climatic situations with success. It is a warm-season crop that necessitates an extended growing season in order to provide a lucrative harvest. In India, tomato is cultivated in all agro-climatic zones in about 789.2 thousand hectares area and 19,759 thousand MT production with a productivity of 25 MT/ha. It is one of the most important off-season vegetable crops in Himachal Pradesh's mid and low hills, with an annual yield of 539.54 thousand MT in 13.19 thousand hectares in the state, the majority of which is shipped to plain markets (SHARMA *et al.*, 2021). Tomato was first planted as an aesthetic plant in Latin America and then transported to Europe in the sixteenth century (BERGOUGNOUX, 2014). In the 1600s, tomatoes began to be cultivated as an edible plant, spreading swiftly throughout the Old World. Disease resistance, fruit size, texture, flavor, storage ability, and other critical fruit attributes were improved by tomato breeding efforts. Although there are about 10,000 tomato cultivars available today, commercial producers are forced to use only a tiny percentage of them due to severe selecting pressures.

Underutilized existing diversity within the tomato species, could be a potential resource for enriching the genetic pool of cultivated tomato with neglected alleles that could improve production and stress adaptation (TALUCDER, 2024). Due to the intricacy of reactions to multiple biotic stresses, despite tremendous attempts were made to increase tolerance in agronomically important species, but the results produced so far with both conventional breeding have been fairly unsatisfactory. The low genetic diversity of grown tomato needs the development of markers capable of detecting tiny differences within tomato germplasm using modern molecular techniques (FOOLAD and PANTHEE, 2012). For the maintenance, conservation, and promotion of agricultural crop productivity, genetic diversity at both morphological and molecular level are critical. Plant genetic resources, such as gene banks and DNA libraries, can be used to estimate, capture, and store plant genetic diversity. Conserved genetic resources must be used for crop development to offer farmers with protection against biotic and abiotic pressures, as well as to meet future global concerns in food and nutritional security (GOVINDARAJ *et al.,* 2015). Several types of molecular markers, such as simple sequence repeats (SSR; microsatellites), amplified fragment length polymorphism (AFLP), sequence characterized amplified region (SCAR) and single nucleotide polymorphism (SNP) have been developed and widely used for genetic characterization of tomato germplasm collections and in Marker Assisted Selection (MAS) (BAUCHET *et al*., 2017). Therefore, the present study was planned to assess genetic diversity using morphological and molecular markers in tomato. The findings of this work will aid in the selection of tomato genotypes with a high genetic diversity of the genes used in crossbreeding, QTL mapping, gene tagging and other imperative genomic studies.

MATERIAL AND METHODS

Experimental materials

The experimental material comprised of thirty five genotypes which were collected from World Vegetable Center, Taiwan, Indian Institute of Horticultural Research, Bangalore, Indo-American Hybrids Company, Bangalore, Punjab Agricultural University, Ludhiana, University of Horticulture and Forestry, Solan and CSK Himachal Pradesh Agricultural University, Palampur (Table 1; Figure 1).

Table 1. Tomato germplasm and their sources used for diversity analysis

Sr. No.	Germplasm	Collection Source	Sr. No.	Germplasm	Collection Source
1	VT-20-2 Round red	CSKHPKV, Palampur, India	18	W-321	PAU, Ludhiana, India
2	15-2	CSKHPKV, Palampur, India	19	Punjab Ratta	PAU, Ludhiana, India
3	16-B	CSKHPKV, Palampur, India	20	NBHIA-5	PAU, Ludhiana, India
4	BT-20-3: Red Egg shaped	CSKHPKV, Palampur, India	21	Punjab Varkha Bahar-4	PAU, Ludhiana, India
5	BT-20-3: Yellow Round	CSKHPKV, Palampur, India	22	KS-227	PAU, Ludhiana, India
6	Hisar Lal	CSS HAU, Haryana, India	23	PNR-7	PAU, Ludhiana, India
7	Arka Vikas	IIHR, Bangalore, India	24	Punjab Chhuhara	PAU, Ludhiana, India
8	BWR-5	IIHR, Bangalore, India	25	Punjab Upma	PAU, Ludhiana, India
9	Naveen 2000 +	Indo-American Hybrid Seeds, Pvt. Ltd. India	26	Punjab Tropic	PAU, Ludhiana, India
10	Rakshita	Indo-American Hybrid Seeds, Pvt. Ltd. India	27	Punjab Varkha Bahar-2	PAU, Ludhiana, India
11	Swarna Rattan	PAU, Ludhiana, India	28	PKM-1	TNAU, Coimbatore, India
12	HADT- 294	PAU, Ludhiana, India	29	UHF-521	UHF, Solan, India
13	KSP-1154-5	PAU, Ludhiana, India	30	S-208	UHF, Solan, India
14	COH 1	PAU, Ludhiana, India	31	CLN 1464A	WorldVeg, Taiwan
15	Punjab Kesri	PAU, Ludhiana, India	32	EC-531802	WorldVeg, Taiwan
16	S-12	PAU, Ludhiana, India	33	EC-11975	WorldVeg, Taiwan
17	I-181	PAU, Ludhiana, India	34	EC-535580	WorldVeg, Taiwan
Bolded genotypes: Highly susceptible to wilt hence not used in agro-morphological study.			35	Hawaii-7998	WorldVeg, Taiwan

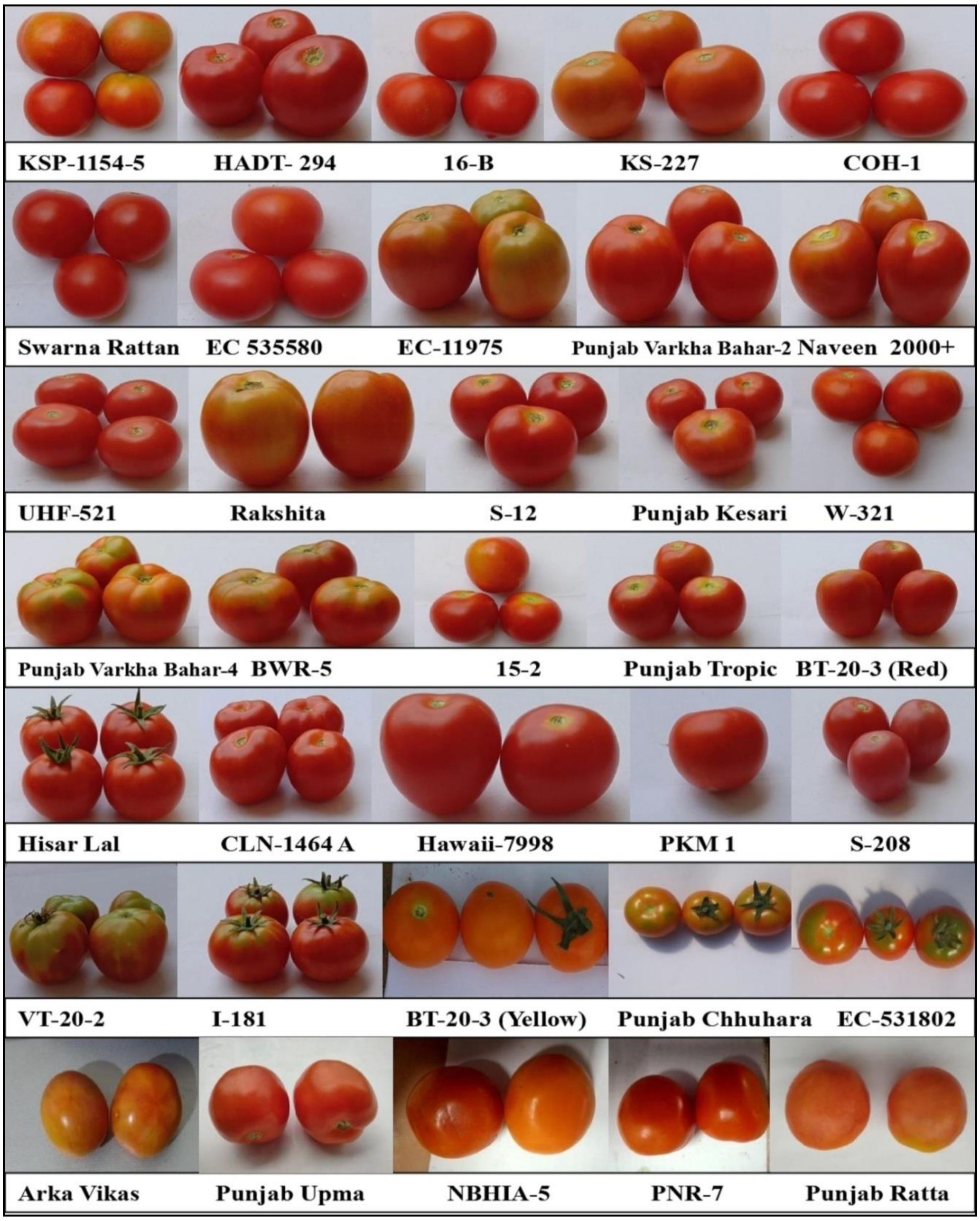

Figure 1. Phenotypic diversity for fruit shape and color in tomato germplasm

The genotypes were maintained at the Experimental Farm and Molecular Biology Laboratory of Vegetable Science and Floriculture department, CSK Himachal Pradesh Agricultural University, Palampur, India during the year 2020-21 to take up genetic diversity analysis. The experimental farm is located at an elevation of 1,290.8 meters above mean sea level, with latitudes of 32°6' N and longitudes of 76°3' E. Among the thirty-five tomato genotypes, only twenty-six were used for the morphological study, while all thirty-five were included in the molecular analysis. The remaining nine genotypes were excluded from the morphological study due to their susceptibility to bacterial Fusarium wilt under field conditions.

Morphological data and its analysis

The experiment was laid out in a randomized block design with three replications. Seedlings were transplanted at a spacing of 70 cm × 30 cm. All the recommended cultural practices were adopted for raising the crop successfully. The observations were recorded on five randomly selected plants per replication for each genotype on sixteen characters: days to 50% flowering, days to first harvest, internodal length (cm), number of nodes, plant height (cm), plant survival (%), number of fruits per plant, number of fruits per cluster, average fruit weight (g), number of locules per fruit, fruit shape index, pericarp thickness (mm), total soluble solid (TSS) (% Brix), ascorbic acid (mg/100g), titrable acidity (%), yield per plant (kg). Multivariate analysis was done utilizing Mahalanobis D^2 statistic and genotypes were grouped into different clusters following Tocher's method (SINGH *et al*., 2022).

Genotyping

The entire genomic DNA was extracted from each genotype using the CTAB technique and quantification was done using Nanodrop spectrophotometer (A_{260}/A_{280}) and Gel-electrophoresis. Amplification of genomic DNA was carried out in thermal cycler with denaturation followed by annealing with SSR primers and extension processes and amplified products were quantified using agarose gel electrophoresis visualized using the Gel-Doc system. For all genotypes, exclusive DNA bands were evaluated as present (1) or absent (0). Binary data were used to generate a Jaccard's similarity coefficient through UPGMA (Unweighted Pair-Group Method with Arithmetic Averages) method. PIC value calculates the informativeness of a particular DNA marker (SERROTE *et al*., 2020). Using the software STRUCTURE version 2.3.4 (PRITCHARD *et al.,* 2000), model-based cluster analysis was performed to determine the genetic structure and number of clusters in the data set. The number of hypothesized populations (K) was varied between 2 and 10 and the analysis was carried out twice and the true k was determined according to the method described by (KAUR *et al.,* 2023). Principal Coordinates Analysis (PCoA) in GenALEx 6.5 was used to visualise the genetic relationship patterns in the matrix. Maximum likelihood, POPGENE and 1000 bootstraps were used to assign individual genotypes into groups, variety of genetic variation parameters and test branch robustness, respectively. Structure analysis was done to estimate population structure (Q matrix) using STRUCTURE (PRITCHARD *et al.,* 2000; FALUSH *et al.,* 2003) and express as membership probability.

RESULTS

Clustering analysis

Clustering of genotypes under study is presented in Figure 2. Based on D^2 values, the 26 genotypes were grouped into twelve highly divergent clusters (Table 2). Some of genotypes were so divergent in all the characters; hence each single genotype formed a separate cluster.

The two clusters were having maximum number of genotypes. Cluster III was biggest with 8 genotypes *viz*., CLN-1464A, UHF-521, S-12, Punjab Kesri, Punjab Varkha Bahar-4, Hawaii-7998, Punjab Varkha Bahar-2, EC-535580 followed by cluster I with 7 genotypes viz., COH-1, PKM-1, BWR-5, 16-B, 15-2, VT-20-2 Round red, KSP-1154-5 and the remaining 10 clusters were solitary with one genotype in each cluster (Table 2).

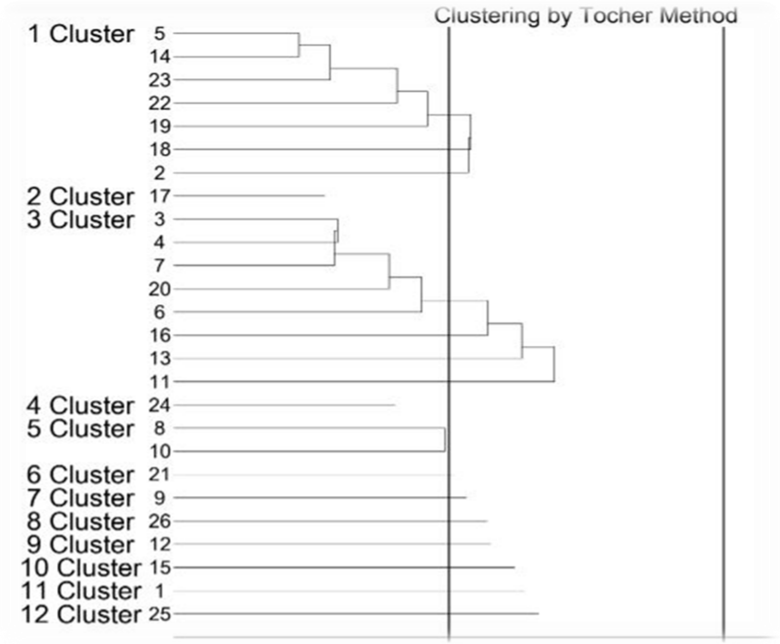

Figure 2. Dendrogram showing clustering pattern for divergence of tomato genotypes. 1=HADT-294, 2=KSP-1154-5, 3=CLN-1464A, 4=UHF-521, 5=COH-1, 6=Punjab Kesri, 7=S-12, 8=I-181, 9=W-321, 10=EC-11975, 11=Punjab Varkha Bahar-4, 12=KS-227, 13=EC-535580, 14=PKM-1, 15=S-208, 16=Punjab Varkha Bahar-2, 17=Arka Vikas, 18=VT-20-2 Round Red, 19=15-2, 20=Hawaii-7998, 21=Naveen 2000+, 22=16-B, 23=BWR-5, 24=BT-20-3 (Red egg shaped), 25=BT-20-3(Yellow round), 26=Rakshita.

Intra-cluster distances represent genetic divergence within clusters, while inter-cluster distances indicate cluster diversity (Table 3). The cluster III had the maximum D^2 value (8.1) followed by Cluster I (7.56). The inter cluster distance was minimum between cluster II and VI (7.13) indicating close relationship and similarity for most of the characters of the genotypes included in these clusters. The maximum inter cluster distance was observed between clusters V and X (18.81) followed by between clusters X and XII (17.52) indicating wider genetic diversity among the genotypes included in these groups.

Table 2. Cluster classification of 26 genotypes of tomato

Cluster	Number of genotypes	Genotypes
1	7	COH-1, PKM-1, BWR-5, 16-B, 15-2, VT-20-2 Round red, KSP-1154-5
2	1	Arka Vikas
3	8	CLN-1464A, UHF-521, S-12, Punjab Kesri, Punjab Varkha Bahar-4, Hawaii-7998, Punjab Varkha Bahar-2, EC-535580
4	1	BT-20-3 (Red egg shaped)
5	2	I-181, EC-11975
6	1	Naveen 2000 +
7	1	W-321
8	1	Rakshita
9	1	KS-227
10	1	S-208
11	1	HADT-294
12	1	BT-20-3 (Yellow round)

Table 3. Average intra (bold) and inter-cluster D^2 values for twelve clusters in 26 genotypes of tomato

CLUSTERS	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
I	**7.56**	8.27	10.77	9.35	11.53	9.8	11.34	9.99	10.02	14.12	9.7	9.7
II		**0.00**	10.09	11.05	9.41	7.13	14.48	8.73	11.31	15.84	9.9	11.69
III			**8.1**	11.76	13.22	11.2	13.05	13.08	9.86	10.45	11.15	14.64
IV				**0.00**	15.7	9.78	11.67	11.38	9.99	12.52	13.46	9.55
V					**7.02**	11.27	13.94	11.54	13.27	18.81	10.37	11.83
VI						**0.00**	11.51	7.55	9.65	15.68	10.01	9.9
VII							**0.00**	12.51	7.67	14.05	9.17	10.31
VIII								**0.00**	12.53	15.66	8.87	10.54
IX									**0.00**	12.8	8.71	11.57
X										**0.00**	14.8	17.52
XI											**0.00**	12.65
XII												**0.00**

Relative contribution of different characters towards divergence

The character's proportionate contribution to overall divergence always indicates the variability associated with that character. Titrable acidity (33.23 %) contributed the most to yield divergence among the yield contributing characters, followed by TSS (18.77%). %). Number of locules per fruit (11.08 %), ascorbic acid (9.85 %), average fruit weight (5.85 %), days to first harvest (4.31 %), fruit shape index (3.69 %), yield per plant (2.77 %), pericarp thickness (2.46%), plant survival percent (2.15%), number of fruits per plant (2.15%), number of nodes (1.85) and plant height (1.85) were the remaining characters which contributed towards divergence in descending order (Figure 3). The remaining characters viz., days to 50% flowering, internodal length and number of fruits per cluster did not contribute to the total divergence. Greater contribution represents the potential for using those traits in crop development. These traits are less changeable from genotype to genotype and their transmission to subsequent generations is nearly zero. As a result, these characteristics should not be prioritized when selecting genotypes for breeding. Highest intra-cluster distance was observed for clusters III, I, V and on the other hand, highest inter-cluster distance was observed between clusters V and X followed by clusters X and XII. So, it is desirable to select genotypes from clusters showing high inter-cluster distance also with high fruit yield as parents in recombination breeding programmes as these are much diverse from each other.

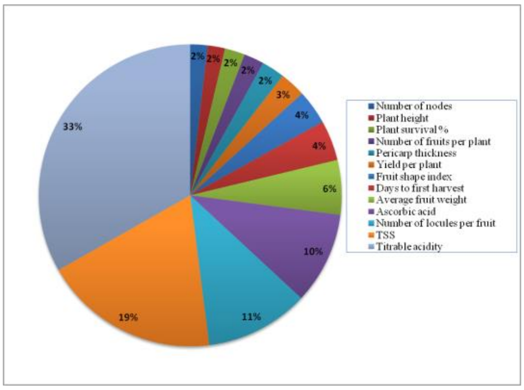

Figure 3. Relative contribution of characters studied toward genetic divergence

Molecular Characterization

SSR and marker informativeness

The gel representation of 35 germplasm with primer T-70 is revealed in Figure 4. The total molecular variability parameters such as PIC, heterozygosity, major allele frequency, number of alleles and allele size across all 35 germplasm are presented in Table 4. Out of 60 SSR primers, 36 primers exhibited high polymorphism. A total of 74 amplicons were created, with sizes ranging from 100 to 320 bp. The total number of alleles from 36 primers observed was 74 with a mean of 2.05 alleles per locus and three alleles were identified in primers like TOM 31A-32A and T-70. Major allele frequency varied from 0.50 (TOM 47-48) to 0.96 (STI24) with an average value of 0.67. The polymorphic information content (PIC), ranged from 0.07 (STI24) to 0.46 (T-70), with an average value of 0.32 per primer. Similarly, heterozygosity varied from 0.07 (STI24) to 0.52 (T-70) with an average value of 0.41.

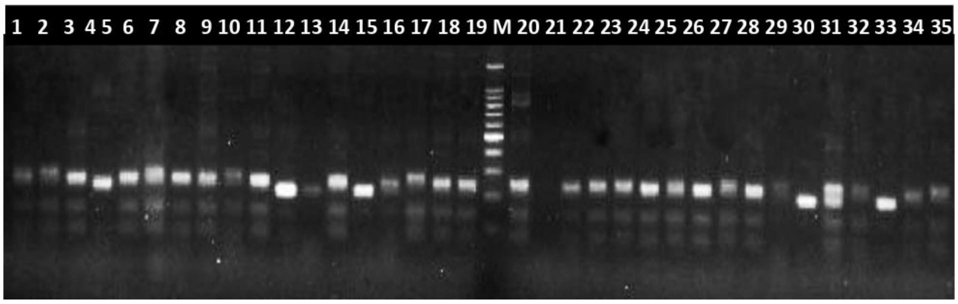

Figure 4. DNA profile of 35 germplasm of tomato showing polymorphism with SSR primer T-70 (M is 100 bp ladder)

Table 4. List of SSR primer observed product size and their PIC during diversity study

S. N.	Primers	Forward primer (5'-3')	Reverse primer (5'-3')	No. of alleles	Major allele frequency	PIC	Heterozygosity
1.	TOM 11-28	ATTGTAATGGTGATGCTCTTCC	CAGTTACTACCAAAAATAGTCAAACAC	2	0.56	0.37	0.50
2.	TOM 31A-32A	AATGTCCTTCGTATCCTTTCGT	CTCGGTTTTAATTTTTGTGTCT	3	0.86	0.24	0.26
3.	TOM 41-42	GAAATCTGTTGAAGCCCTCTC	GACTGTGATAGTAAGAATGAG	2	0.52	0.37	0.51
4.	TOM 43-44	GCAGGAGATAATAACAGAATAAT	GGTAGAAGCCCGAATATCATT	2	0.52	0.37	0.51
5.	TOM 47-48	CAAGTTGATTGCATTACCTATTG	TACAACAACATTTCTTCTTCCTT	2	0.50	0.38	0.51
6.	TOM 49-50	AGAAAACTTTTTGAATGTTGC	ATTACAATTTAGAGAGTCAAGG	2	0.85	0.22	0.25
7.	TOM 57-58	TCTAAGTGGATGACCATTAT	GCAGTGATAGCAAATGAAAAC	2	0.52	0.37	0.51
8.	T-62	GTGACCACATGAGATATCCAGA	CAGTTGTCCATATTGTGTGGG	2	0.52	0.37	0.51
9.	T-70	AACATGCGGAGAAAAATT	GGAACACGTCCCAAAAATGT	3	0.65	0.46	0.52
10.	T-107	GCACAAATAATTTTTCAAGACCAA	AAAAACGGACATAGCTTTGTACT	2	0.53	0.37	0.51
11.	SSR 111	TTCTTCCCTTCCATCAGTTCT	TTTGCTGCTATACTGCTGACA	2	0.63	0.36	0.48
12.	SSR 43	CTCCAAATTGGGCAATAACA	TTAGGAAGTTGCATTAGGCCA	2	0.90	0.16	0.18
13.	STI032	TGGGAAGAATCCTGAAATGG	TGCTCTACCAATTAACGGGCA	2	0.62	0.36	0.48
14.	STI0003	ACCAATCCACCATGTCAATGC	CTCATGGATGGTGTCATTGG	2	0.55	0.37	0.50
15.	STI0023	GCGAATGACAGGACAAGAGG	TGCCACTGCTACCATAACCA	2	0.58	0.37	0.50
16.	STM1045	GTACATATATATACAATTATCTAACCG	TTCTCTATGTTAGGCTAGAGTG	2	0.74	0.31	0.39
17.	STI24	CGCCATTCTCTCAGATCACTC	GCTGCAGCAGTTGTTGTTGT	2	0.96	0.07	0.07
18.	SSR565	GAGGATGATGAGAACTCGCC	TCAGAGGCTTCTGGGTCAGT	2	0.53	0.37	0.51
19.	SSR46	CCGAGGCGAATCTTGAATAC	GCACCATCTCTTGTGCCTCT	2	0.67	0.35	0.45
20.	SSR 4	TTCTTCGGAGACGAAGGGTA	CCTTCAATCCTCCAGATCCA	2	0.56	0.37	0.50
21.	AI773078	GATGGACACCCTTCAATTTATGGT	TCCAAGTATCAGGCACACCAGC	2	0.60	0.36	0.49
22.	SSR20	GAGGACGACAACAACAACGA	GACATGCCACTTAGATCCACAA	2	0.70	0.33	0.43
23.	SSR74	ACTCACCATGGCTGCTTCTT	TTTCTTGAAGGGTCTTTCCC	2	0.52	0.37	0.51
24.	X90937	TGCCCATGACGTTCCATC	GACAGACAGAGAGACAGATTTAGG	2	0.61	0.36	0.49
25.	SSR 9	CCCTTTGCAAGTTCTTCTTCA	TTCATGAGCCAACATAGGAGG	2	0.85	0.23	0.27
26.	AI895937	CATAATCACAAGCTTCTTTCGCCA	CATATCCGCTCGTTTCGTTATGTAAT	2	0.77	0.29	0.36
27.	AI780156	TCCAATTTCAGTAAGGACCCCTC	CCGAAAACCTTTGCTACAGAGTAGA	2	0.70	0.33	0.43
28.	AW037347	GCCACGTAGTCATGATATACATAG	GCCTCGGACAATGAATTG	2	0.71	0.32	0.41
29.	SSR47	TCCTCAAGAAATGAAGCTCTGA	CCTTGGAGATAACAACCACAA	2	0.83	0.24	0.29
30.	TOM236-237	GTTTTTTCAACATCAAAGAGCT	GGATAGGTTTCGTTAGTGAACT	2	0.81	0.26	0.33
31.	AI491173	GCACGAGCACATATAGAAGAGAATCA	CCATTTCATCATATCTCTCAGCTTGC	2	0.62	0.36	0.48
32.	AI897173	CCTCTCTTCACCTCTTTACAATTTCC	CACTGGTCATTAAGTCTACAGCC	2	0.83	0.24	0.28
33.	TMS6	CTCTCTCAATGTTTGTCTTTC	GCAAGGTAGGTAGCTAGGGA	2	0.92	0.13	0.15
34.	U2108	CATTTTATCATTTATTTGTGTCTTG	ACAAAAAAAAGGTGACGATACA	2	0.76	0.30	0.37
35.	LEMDDNa	ATTCAAGGAACTTTTAGCTCC	TGCATTAAGGTTCATAAATGA	2	0.57	0.37	0.50
36.	LE20592	CTGTTTACTTCAAGAAGGCTG	ACTTTAACTTTATTATTGCCACG	2	0.64	0.35	0.47
Overall				**74**	**24.21**	**11.45**	**14.91**
Average				**2.05**	**0.67**	**0.32**	**0.41**

Genetic diversity

Thirty-five tomato genotypes were divided into two main clusters (A and B). Cluster A was separated into two sub-clusters comprising of total of 27 germplasm, while Cluster B had contained eight genotypes namely, Swarna Rattan, CLN-1464A, Punjab Kesri, Punjab Tropic, I-181, NBHIA-5, Punjab Upma and Punjab Chhuhara which were basically introduced from PAU, (Punjab, India) and one line from AVDRC (Taiwan) (Figure 5). Based on UPGMA analysis, Jaccard's similarity coefficient varied from 0.50 to 0.88 with maximum similarity between genotype BT-20-3 (Red Egg shaped) & BT-20-3 (Yellow Round), and NBHIA-5 & Punjab Upma, whereas minimum between Swarna Rattan & KSP-1154-5. Based on Neighbor joining analysis, genotypes were grouped into three clusters as depicted using the color codes in Figure 6. Cluster I (Red), Cluster II (Blue) and Cluster III (Green) comprised of twenty, five, ten genotypes, respectively. Principal Coordinate Analysis (PCoA) showed that first three coordinates accounted for 35.44 % cumulative variation among 35 genotypes (Figure 6) with the first and second coordinates explaining 17.66% and 10.45% of the total variation, respectively.

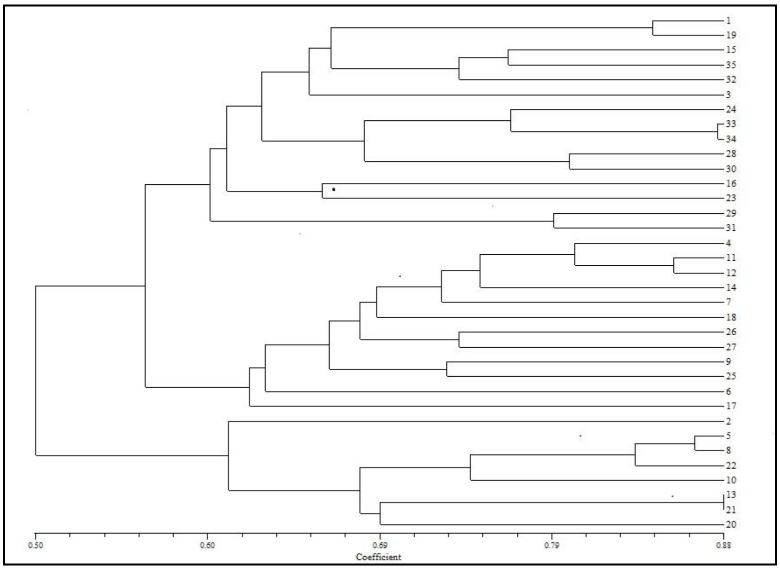

Figure 5. Dendrogram depicting genetic relationships among the tomato germplasm constructed by NTSYS–PC (version 2.02) using UPGMA method

Structure analysis

The STRUCTURE analysis divided the population into three groups. The differentiations at K = 3 were nearly equivalent to pedigree knowledge with a few outliers. In group 1 (Red) consists of 13 genotypes, group 2 (Green) comprises 10 genotypes while 12 genotypes in group 3 (Blue) (Figure 7a). The germplasm generated by the NTSYS software were confirmed using STRUCTURE analysis at K = 3. As a result of this, it was established that the germplasm that were separated according to cluster analysis were almost identical to those that were divided according to structure analysis, with a few minor differences. The Evanno test gave a maximum peak value for Delta K at K = 3 in the plots of L (K) versus Delta confirming three sub-groups among tomato genotypes. The individual membership coefficient at K = 3 had maximum mean probability of likelihood value of L (K) = -2359.4, leading to assignment of tomato genotypes to three (K = 3) sub-populations (Figure 7b).

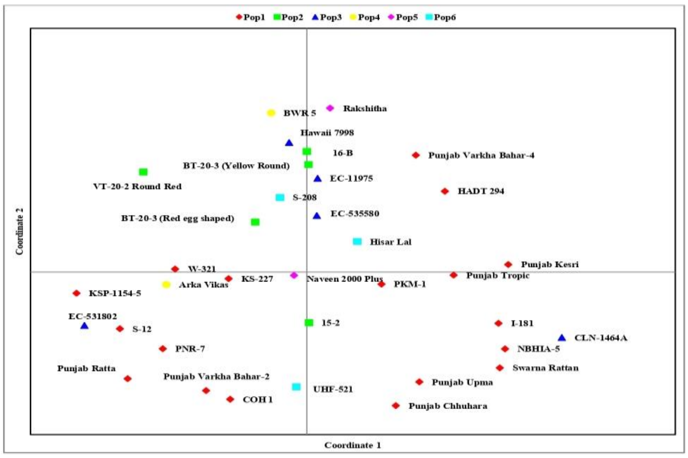

Figure 6. PCoA scatter diagram analysis showing the distributions of 35 tomato germplasm

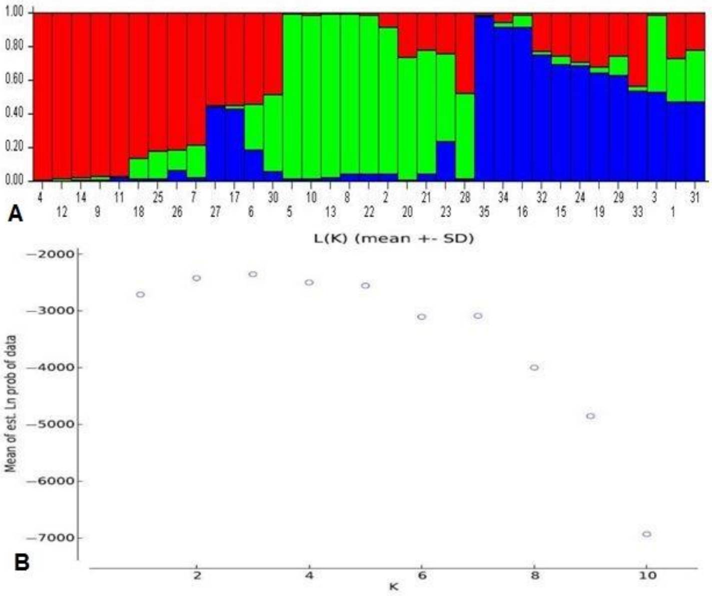

Figure 7. Genetic structure of 35 tomato germplasm, considering K= 3. Colors (red blue and green) represent the three groups, defined by the K value. Tomato germplasm showing more than one color may have an intermixed genetic makeup, resulted from crossing. a) The vertical axis indicates the membership value b) Estimation of the number of populations for K ranging from 1 to 10 by calculating delta K values.

Structure analysis

The STRUCTURE analysis divided the population into three groups. The differentiations at K = 3 were nearly equivalent to pedigree knowledge with a few outliers. In group 1 (Red) consists of 13 genotypes, group 2 (Green) comprises 10 genotypes while 12 genotypes in group 3 (Blue) (Figure 7a). The germplasm generated by the NTSYS software were confirmed using STRUCTURE analysis at K = 3. As a result of this, it was established that the germplasm that were separated according to cluster analysis were almost identical to those that were divided according to structure analysis, with a few minor differences. The Evanno test gave a maximum peak value for Delta K at K = 3 in the plots of L (K) versus Delta confirming three sub-groups among tomato genotypes. The individual membership coefficient at K = 3 had maximum mean probability of likelihood value of L (K) = -2359.4, leading to assignment of tomato genotypes to three (K = 3) sub-populations (Figure 7b).

DISCUSSION

Sufficient information about genetic variation is necessary to be utilized by breeding, conservation and management of genetic resources programs. On the other hand, the amount of variation can quantify the information needed for good germplasm conservation in crops and is considered a prerequisite for attempting any successful breeding operations. In our study, the morphological data represented a wide amount of variation among the genotypes. The cluster analysis revealed that 26 genotypes were sorted into twelve clusters. Among them cluster III had maximum genotypes followed by cluster I. This was in agreement with results of (CHERNET *et al.,* 2014), (MEENA and BAHADUR, 2017). The lowest average genetic divergence values (inter-cluster value), indicating a close association between the genotypes included in these clusters and less heterosis responses will be utilized by hybridizing the genotypes found in these clusters. Similar results were reported in tomato by (CHERNET *et al.*, 2014) and he observed the largest inter-cluster distance between Cluster IV and V (1805.00) and similarly in other study of (MEENA and BAHADUR, 2017) the highest inter-cluster values were found between clusters III and VI (47922.37). The relative contribution of the character towards total divergence always hints on the variability associated with that character. More is the relative contribution; maximum is the possibility of exploitation of that character in crop improvement. Our study revealed that, titrable acidity (33.23 %) contributed the most to yield divergence among the yield contributing characters, followed by TSS (18.77%). The contribution to divergence of features like days to 50% flowering, internodal length and number of fruits per cluster was found to be 0%. Previous study, (SINGH *et al.,* 2008) confirmed importance of the plant height, number of fruits per plant and fruit yield per plant to divergence.

The genetic diversity and population structure in tomato was investigated for improvement of various traits using crop breeding practices. A limited number of SSR molecular markers are used in Indian genotypes. It has been investigated that SSR markers show high polymorphism in tomato. In our study, we have determined the genetic diversity using Sixty SSR markers in 35 genotypes of tomato comprising a wider geographical distribution of genotypes. Among 60 SSRs primers, 36 primers showed high polymorphism and a total of 74 alleles were identified with an amplicon size ranging from 100-320 bp. The number of alleles observed was 2-3 with a mean of 2.05 alleles per locus. Similarly, (BENOR *et al.,* 2008) observed an amplicon size ranging from 100-400 bp and a total of 64 alleles with an average 4 alleles per locus, respectively.

The polymorphic information content (PIC), a measure related to marker discrimination, ranged from 0.066 (STI24) to 0.461 (T 70), with a mean of 0.32 per primer. Heterozygosity varied from 0.070 (STI24) to 0.523 (T 70) with a mean of 0.41. The results were in agreement with earlier studies on tomato suggesting the role of SSR markers for identification of genotypes, DNA fingerprinting and maintenance of genotypes in the gen banks. Based on UPGMA analysis with Jaccard's similarity coefficient varied from 0.50 to 0.88. Similarly, (IBRAHIM and ERDINC, 2020) reported Jaccard's similarity coefficient ranging from 0.360 to 0.960 and 0.261 to 0.941, in tomato respectively.

Minimum Jaccard's similarity coefficient was observed in Swarna Rattan and KSP-1154-5 showing maximum diversity among genotypes. The clustering formed by the UPGMA dendrogram was moderately validated by projecting individual genotypes into a two-dimensional multivariate space in PCoA diagram. As per UPGMA method the tomato genotypes were divided into two main clusters A (A1-15 and A2-12) and B (8). Similar results were reported by SINGH *et al.*, (2015) which grouped 24 germplasm into two main distinct clusters. Various clustering methods were employed to assess genetic relationship of different genotypes or germplasm. Based on Neighbor Joining tree thirty-five genotypes were grouped into three clusters as represented by using color codes. Cluster I with 20 genotypes formed a larger cluster followed by cluster III (10 genotypes) and cluster II (5 genotypes).

PCoA is a multivariate strategy for grouping data based on similarity coefficients or variance-covariance values that provides more information about main groups, whereas cluster analysis provides higher resolution among closely related populations. PCoA explores correlations between many quantitative variables by constructing a small number of linear combinations (principal components) that retain as much information as feasible from the original data. Principal Coordinate Analysis showed that first three coordinates accounted for 35.44 % cumulative variation among 35 genotypes (Figure 6) with the first and second coordinates explaining 17.66% and 10.45% of the total variation respectively. The population structure analysis grouped the genotypes into 3 groups including genotypes having admixtures. As a result, pedigree information was combined with cluster membership to determine the division of red, green and blue groupings. Similarly, 75 Italian landraces were categorized into four groups (CORRADO *et al.,* 2014) and, 276 USDA-GRIN collections were grouped into three populations (ALATAWI *et al.,* 2024). The increased variance should be recorded for germplasm preservation and agricultural enhancement breeding strategies.

Issues of agricultural importance, such as biodiversity loss, as well as the fact that modern farming practices require low inputs, all represent new problems for researchers in the current universal agenda (BAI *et al.,* 2018). In light of recent climate change, it is clear that characterizing and conserving genetic diversity of genotypes is critical in order to save the gene pool's "richness". Future research should concentrate on phenotypic variation in order to discover potential plant material for tomato breeding projects. It is necessary to develop a broad collection of accession and crosses between populations should be made rather than within populations. This could offer a firm platform for preventing genetic erosion, as well as a significant boost to breeding programmes for crop development.

CONCLUSION

This study could be used to estimate genetic variation within a group of elite genotypes to employ in tomato improvement in India. Cluster III had maximum number of genotypes and showed highest intra cluster distance. The maximum inter-cluster distance between cluster V and cluster X (18.81), indicating the existence of wide genetic variability. A total of 35 tomato genotypes were assessed using 36 polymorphic SSR markers. The experiment depicted total number of 74 amplicons, with an overall average of 2.05 alleles per locus. T-70 primer was recorded to have good marker informativeness. Based on UPGMA cluster analysis, maximum similarity (less diverse) was observed between genotypes BT-20-3: Red egg shaped & BT-20-3: Yellow round and NBHIA-5 & Punjab Upma whereas minimum similarity (more diverse) between Swarna Rattan & KSP-1154-5. The population structure depicted three main populations including admixture genotypes. It may be further utilized in future projects related to QTLs identification, genome wide association studies, DNA fingerprinting and preservation of tomato germplasm across India and other countries.

ACKNOWLEDGEMENT

Authors are thankful to the Department of Vegetable Science and Floriculture, CSK Himachal Pradesh Agricultural University, Palampur, India for providing necessary research facilities.

Received, March 15th, 2024
Accepted January 25th, 2025

REFERENCES

ALATAWI, I., H., XIONG, H., ALKABKABI, K., CHIWINA, Q., LUO, K.S., LING, Y., QU, R., DU, A., SHI (2024): Genetic diversity and population structure of tomato (*Solanum lycopersicum*) from the USDA-GRIN germplasm collection. Agronomy, *15*(1): 22.

BAI, Y., C., KISSOUDIS, Z., YAN, R.G.F., VISSER, G., VAN DER LINDEN (2018): Plant behavior under combined stress: tomato responses to combined salinity and pathogen stress. The Plant J., *93*: 781–793.

BAUCHET, G., S., GRENIER, N., SAMSON, J., BONNET, L., GRIVET, M., CAUSSE (2017): Use of modern tomato breeding germplasm for deciphering the genetic control of agronomical traits by genome-wide association study. TAG, *130*: 875–889.

BENOR, S., M., ZHANG, Z., WANG, H., ZHANG (2008): Assessment of genetic variation in tomato (*Solanum lycopersicum* L.) inbred lines using SSR molecular markers. Journal of Genetics and Genomics and Genomics, *35*: 373–379.

BERGOUGNOUX, V. (2014): The history of tomato: from domestication to biopharming. Biotechnology Advances, *32*: 170–189.

CHERNET, S., D., BELEW F., ABAY (2014): Genetic diversity studies for quantitative traits of tomato (*Solanum lycopersicon* L.) genotypes in Western Tigray, Northern Ethiopia. J. Plant Breeding and Crop Sci., *6*: 105–113.

CORRADO, G., M., CARAMANTE, P., PIFFANELLI, R., RAO (2014): Genetic diversity in Italian tomato landraces: Implications for the development of a core collection. Scientia Horticulturae, *168*: 138–144.

EVANNO, G., S., REGNAUT, J., GOUDET (2005): Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology, *14*: 2611–2620.

FALUSH, D., M., STEPHENS J.K., PRITCHARD (2003): Inference of population structure using multilocus genotype data:

Linked loci and correlated allele frequencies. Genetics, *164*: 1567–1587.

FOOLAD, M.R., D.R., PANTHEE (2012): Marker-Assisted Selection in Tomato Breeding. Critical Reviews in Plant Sciences, *31*: 93–123.

GOVINDARAJ, M., M., VETRIVENTHAN, M., SRINIVASAN (2015): Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its analytical perspectives. Genetics Research International, *2015*: 1–14.

IBRAHIM, A.S., C., ERDINC (2020): Determination of genetic relations among tomato accessions in Sulaymaniyah region through ISSRs markers. Yuzuncu Yıl University Journal of Agricultural Science, *30*: 810–820.

KAUR, M., P., SHARMA, A., SHARMA, H., LATA, N., KUMAR (2023): SSR analysis to assess genetic diversity and population structure in parthenocarpy cucumber (*Cucumis sativus* L.). Journal of Horticultural Sciences, *18*(1): 46–52.

MEENA, O.P., V., BAHADUR (2017): Principal component and cluster analysis of indigenous tomato genotypes based on morphological indicators. Research Journal of Biotechnology, *12*: 50-58.

PRITCHARD, J.K., M., STEPHENS, P., DONNELLY (2000): Inference of population structure using multilocus genotype data. Genetics, *155*: 945–959.

SERROTE, C.M.L., L.R.S., REINIGER, K.B., SILVA, S.M., RABAIOLLI, S., DOS, C.M. STEFANEL (2020): Determining the polymorphism information content of a molecular marker. Gene, *726*: 144175.

SHARMA, H., S., SOOD, V. KUMAR (2021): Economic analysis of tomato cultivation in Himachal Pradesh. Indian Journal of Economics and Development, *17*(4): 960-965.

SINGH, A.K., J.P., SHARMA, S., KUMAR, S., CHOPRA (2008): Genetic divergence in tomato (*Lycopersicum esculentum* Mill.). SKUAST Journal of Research, *7*: 1–8.

SINGH, B., A., GOSWAMI, VAISHALI (2015): Morphological and molecular characterization of tomato (*Lycopersicum esculentum* Mill) genotypes. Vegetos, *28*: 67–75.

SINGH, H. P., O.P., RAIGAR, R.K. CHAHOTA (2022): Estimation of genetic diversity and its exploitation in plant breeding. The Botanical Review, *88*(3): 413–435.

TALUCDER, M.S.A., U.B., RUBA, M.A.S., ROBI (2024): Potentiality of Neglected and Underutilized Species (NUS) as a future resilient food: A systematic review. Journal of Agriculture and Food Research, *16*: 101116.

THE TOMATO GENOME CONSORTIUM (2012): The tomato genome sequence provides insights into fleshy fruit evolution. Nature, *485*(7400): 635–641.

GENETIČKI DIVERZITET I STRUKTURA POPULACIJE GENOTIPOVA PARADAJZA (*Solanum lycopersicum* L.)

V.M. RASHMI[1], Parveen SHARMA[1], Manpreet KAUR[1], Akhilesh SHARMA[1],
Nimit KUMAR[1], Aparajita DWIVEDI[2], Sanjay Kumar SANADYA[2*]

[1]Departman za povrtarstvo i cvećarstvo, CSK Himačal Pradeš Poljoprivredni univerzitet,
Palampur (Himačal Pradeš)-176062 Indija
[2]Departman za genetiku i oplemenjivanje biljaka, CSK Himačal Pradeš Poljoprivredni univerzitet, Palampur (Himačal Pradeš)-176062 Indija

Izvod

U ovoj studiji, morfološki i molekularni markeri su korišćeni za procenu genetske raznovrsnosti 35 genotipova paradajza. Međutim, samo 26 genotipova paradajza je korišćeno za sprovođenje klister analize. Zbog bolesti fuzarijumskog uvenuća, devet genotipova paradajza je uginulo. Na osnovu analize, genotipovi paradajza su grupisani u dvanaest klastera, gde je klaster III sa osam genotipova sledio klaster I sa sedam genotipova. Maksimalna međuklasterna udaljenost primećena je između klastera V i X, što ukazuje na postojanje širokog spektra genetske raznovrsnosti. Molekularna raznolikost i populaciona struktura 35 genotipova paradajza analizirane su korišćenjem polimorfnih markera jednostavnog sekvencijalnog ponavljanja (SSR). Ukupno je identifikovano 74 alela sa prosekom od 2,05 alela po lokusu, dok se sadržaj informacija o polimorfizmu prajmera kretao od 0,07 do 0,46 sa srednjom vrednošću od 0,32. Frekvencija glavnih alela i heterozigotnost kretale su se od 0,50-0,96 i 0,07-0,52, respektivno. Dendrogram je grupisao genotipove u dve glavne grupe A i B sa 27 i 8 genotipova, respektivno, ali grupa A je imala dva podklastera. Koeficijent sličnosti kretao se od 0,50 do 0,88, sa maksimalnom sličnošću između genotipa BT-20-3 (crveni oblik jajeta) i BT-20-3 (žuti okrugli) i NBHIA-5 i Punjab Upma, dok je minimalna između Swarna Rattan i KSP-1154-5. Struktura populacije otkrila je tri podpopulacije sa nekim primesama. Analiza glavnih koordinata pokazala je da su genotipovi ravnomerno raspoređeni duž dve ose na oba dijagrama sa 35,44% kumulativne varijacije. Razumevanje genetske raznolikosti i polimorfizma populacija paradajza pomoći će u poboljšanju iskorišćavanja genetskih resursa paradajza.

Primljeno 15.III.2024.
Odobreno 25. I. 2025.

How to cite this article: Rashmi V.M., P. Sharma, M. Kaur, A. Sharma, N. Kumar, A. Dwivedi, S. Kumar Sanadya (2025). *Genetic diversity and population structure of tomato (Solanum lycopersicum L.) genotypes.* - Genetika, Vol 57, No.1, 37-52.