



Plant Archives

Journal homepage: <http://www.plantarchives.org>

DOI Url : <https://doi.org/10.51470/PLANTARCHIVES.2026.v26.supplement-1.205>

ASSESSMENT OF GENETIC DIVERGENCE AMONG GENOTYPES USING MULTIVARIATE ANALYSIS IN TOMATO (*SOLANUM LYCOPERSICUM* L.)

D. Ram Babu¹, Pavani Narigapalli³, Zakir Hussain^{1*}, Praveen Kumar Singh¹, Suman Lata¹, Alka Joshi², Navinder Saini³, Ram Vilas⁴ and Gouthami⁴

¹Division of Vegetable Science, ICAR- Indian Agriculture Research Institute, New Delhi -110012, India

²Division of Postharvest Technology, ICAR- Indian Agriculture Research Institute, New Delhi -110012, India

³Department of Genetics and Plant Breeding, ICAR- Indian Agriculture Research Institute, New Delhi -110012, India

⁴Division of Seed Science and Technology ICAR- Indian Agriculture Research Institute, New Delhi -110012, India

*Corresponding author E-mail: drzakirhussain24@gmail.com

(Date of Receiving : 04-10-2025; Date of Acceptance : 05-12-2025)

ABSTRACT

Genetic diversity is a prerequisite for effective crop improvement and sustainable yield enhancement in tomato (*Solanum lycopersicum* L.). The present study was undertaken to assess the extent of genetic divergence among diverse tomato genotypes using multivariate statistical approaches. Sixty-six tomato genotypes were evaluated for four economically important yield and quality traits, namely pericarp thickness, total soluble solids, per plant yield, and average fruit weight, under field conditions following a randomized complete block design with three replications. Multivariate analysis of variance revealed highly significant differences among genotypes for the combined expression of traits, indicating the presence of substantial genetic variability. Principal component analysis showed that the first two principal components accounted for 74.51% of the total variation, with yield-related traits contributing predominantly to genetic divergence. Mahalanobis D² analysis grouped the genotypes into two major clusters, with high inter-cluster distance suggesting wide genetic divergence among clusters. Singh's character contribution analysis revealed that average fruit weight and per plant yield together contributed more than 73% of the total genetic divergence. Hierarchical clustering further confirmed the distinct grouping pattern among genotypes. The results indicate that yield-associated traits play a major role in determining genetic diversity, and genotypes belonging to widely separated clusters may be effectively utilized as parents in hybridization programmes to exploit heterosis and broaden the genetic base for tomato improvement.

Keywords : Tomato, genetic divergence, Mahalanobis D², principal component analysis, cluster analysis, yield traits

Introduction

Tomato (*Solanum lycopersicum* L.) is one of the most important and widely cultivated vegetable crops worldwide due to its high economic value, broad adaptability and nutritional significance. It serves as a major dietary source of vitamins A, C, minerals and antioxidant compounds such as lycopene and β -carotene, which are associated with reduced risks of chronic diseases (Raiola *et al.*, 2014; Kumar *et al.*, 2022). The global demand for tomato continues to increase, driven by its diverse uses in fresh consumption and processing industries (FAO, 2023).

Despite substantial advances in tomato breeding, yield stagnation and narrowing of the genetic base remain critical concerns. Modern cultivars are often derived from a limited number of elite parents, resulting in reduced genetic variability for yield, quality, and stress tolerance traits (Bai and Lindhout, 2007; Razifard *et al.*, 2020). This genetic uniformity restricts long-term genetic gain and increases susceptibility to biotic and abiotic stresses. Therefore, the effective characterization and utilization of genetically diverse germplasm is essential for sustainable tomato improvement. Genetic diversity is the foundation of any crop breeding programme, as it determines the

potential for selection, heterosis exploitation, and development of superior recombinants (Falconer and Mackay, 1996). In self-pollinated crops like tomato, the assessment of genetic divergence among available genotypes is particularly important for identifying suitable parents for hybridization. Crosses involving genetically diverse parents have been shown to produce higher heterosis and wider segregation in subsequent generations, thereby enhancing breeding efficiency (Joshi and Dhawan, 1966). Traditional approaches to diversity assessment based on univariate analysis of individual traits are limited in their ability to capture the combined influence of multiple correlated characters. Multivariate statistical techniques overcome this limitation by simultaneously considering several traits of economic importance and providing a comprehensive assessment of genetic divergence (Rao, 1952). Among these techniques, *Mahalanobis D²* statistics is widely recognized as a robust and effective method for quantifying genetic divergence and grouping genotypes based on multiple quantitative traits (Mahalanobis, 1936; Singh and Chaudhary, 1985).

Mahalanobis D^2 analysis has been extensively applied in tomato to identify genetically diverse parents and to understand the relative contribution of different traits toward overall divergence (Debnath *et al.*, 2020; Doddamani *et al.*, 2022). When combined with clustering techniques such as Tocher's method and hierarchical clustering, D^2 analysis provides valuable insights into the genetic relationships and population structure of tomato germplasm. Genotypes grouped into widely separated clusters are considered genetically divergent and are recommended for use in hybridization programmes to achieve maximum heterosis and genetic gain. Principal component analysis (PCA) is another powerful multivariate tool that reduces data dimensionality and identifies key traits responsible for variation within a population (Jolliffe, 2002). PCA has been successfully used in tomato diversity studies to determine major yield and quality traits influencing genetic differentiation and to visualize genotype dispersion in multivariate space (Pradhan *et al.*, 2021; Khandaker *et al.*, 2023). The combined application of PCA and D^2 analysis improves the reliability of genetic divergence studies and strengthens parental selection decisions. Tomato yield is a complex trait governed by several interrelated components, including average fruit weight, per plant yield, pericarp thickness, and total soluble solids. Among these, average fruit weight and per plant yield are often reported as major contributors to genetic divergence due to their high variability and direct influence on productivity (Zulfiqar *et al.*, 2020;

Kaur *et al.*, 2022). Pericarp thickness plays an important role in determining fruit firmness, shelf life, and transportability, making it a key trait for commercial cultivation (Batu, 2004).

Total soluble solids (TSS) is an important quality parameter influencing fruit taste and processing suitability. However, TSS generally contributes less to genetic divergence in cultivated tomato germplasm compared to yield-related traits, unless the population includes processing-specific lines or wild relatives (Gomez *et al.*, 2019; Rao and Kumar, 2022). Understanding the relative contribution of yield and quality traits is therefore essential for designing balanced breeding strategies. Several studies have emphasized that clustering patterns derived from multivariate analysis provide practical guidance for parental selection. Crosses between genotypes belonging to widely separated clusters have been shown to result in higher heterotic response and greater variability in segregating generations compared to crosses among closely related genotypes (Pathak and Jha, 2021; Basu *et al.*, 2022). Thus, genetic divergence analysis serves as a valuable decision-making tool in tomato breeding programmes. Although molecular markers and genomic tools have enhanced diversity analysis, phenotypic multivariate analysis remains indispensable, particularly in resource-limited breeding programmes. Phenotypic traits directly reflect agronomic performance under field conditions and capture genotype–environment interactions, which are critical for cultivar development (Malik *et al.*, 2018; Gupta *et al.*, 2023). Moreover, multivariate analysis of variance (MANOVA) allows simultaneous testing of multiple traits and improves the precision of genetic divergence estimates (Rencher, 2002; Reddy and Yadav, 2021). In many tomato-growing regions, a large number of advanced breeding lines, varieties, and hybrids are available; however, systematic evaluation of their genetic divergence for yield and quality traits is often lacking. Comprehensive assessment of genetic diversity is therefore necessary to identify elite and divergent genotypes that can be effectively utilized in future breeding programmes.

The present study was undertaken to assess genetic divergence among tomato genotypes using multivariate statistical approaches, including MANOVA, principal component analysis, *Mahalanobis D²* statistics, and cluster analysis. The investigation focused on economically important yield and quality traits, namely pericarp thickness, total soluble solids, per plant yield, and average fruit weight. The objectives were to quantify genetic diversity, identify major contributors to divergence, classify

genotypes into distinct clusters, and identify genetically diverse parents with potential utility in tomato improvement programmes.

Materials and Methods

The present study was conducted to assess genetic divergence among tomato (*Solanum lycopersicum* L.) genotypes using multivariate statistical approaches. The experiment conducted in *Kharif-2023* at IARI New Delhi. Plant material comprised sixty-six diverse tomato genotypes, including advanced breeding lines

and hybrids, selected to represent a wide range of variability for yield and quality traits. The experiment was carried out during the cropping season at the experimental farm of the institute under open field conditions, following recommended agronomic practices for tomato cultivation. The experiment was laid out in a Randomized Complete Block Design (RCBD) with three replications. Each genotype was represented by a single plot per replication. Table 1 and 2 presents the number of genotypes taken for study and traits evaluated for study in tomato.

Table 1: Tomato 66 genotypes evaluated in the present study

K-127	K-55	K-54	K-61	K-70	K-110
K-63	K-73	K-125	K-69	K-64	K-84
K-83	K-192	H-81	K-74	K-92	K-82
K-130	K-47	K-52	K-91	K-217	K-122
K-49	K-132	H-147	K-37	K-38	K-77
K-46	K-131	H-26	K-113	H-412	K-134
K-44	H-48	K-33	K-35	K-120	K-72
K-133	K-80	K-78	K-34	K-87	H-162
K-39	K-32	K-220	K-79	K-65	K-19
K-21	K-60	K-20	K-98	K-111	K-50
K-391	K-412	K-42	K-43	K-51	K-516

Table 2: Traits evaluated and methods of observation in tomato

Trait	Method of observation	Unit
Pericarp thickness	Fully mature, marketable fruits were cut transversely at the equatorial region and pericarp thickness was measured using a digital vernier caliper. Measurements were taken from multiple fruits per genotype per replication and averaged.	cm
Total soluble solids (TSS)	Fresh fruit juice was extracted and TSS was measured using a digital refractometer calibrated with distilled water prior to use. Readings were recorded at room temperature and averaged.	°Brix
Per plant yield	Per plant yield was calculated as the cumulative weight of all marketable fruits harvested from an individual plant during the cropping period. Mean values were computed for each genotype per replication.	kg plant ⁻¹
Average fruit weight	Average fruit weight was derived by dividing the total fruit yield per plant by the corresponding number of fruits harvested. Replication-wise mean values were used for analysis.	g

Statistical Analysis

Genotype-wise mean data pooled over replications were subjected to multivariate statistical analysis. Multivariate analysis of variance (MANOVA) was performed to test the significance of genotypic differences for the combined set of traits. Principal component analysis (PCA) was carried out to identify major sources of variation and to visualize genetic relationships among genotypes. Mahalanobis D² statistics were computed to estimate genetic distances among genotypes based on pooled variance covariance matrices. Genotypes were grouped into clusters using Tocher's method, intra and inter-cluster distances were calculated to assess the extent of genetic divergence.

The relative contribution of individual traits to total genetic divergence was estimated using Singh's

method. In addition, hierarchical clustering was performed using average linkage to generate a dendrogram depicting genetic relationships among the genotypes. All statistical analyses were performed using the R statistical software with appropriate multivariate analysis packages.

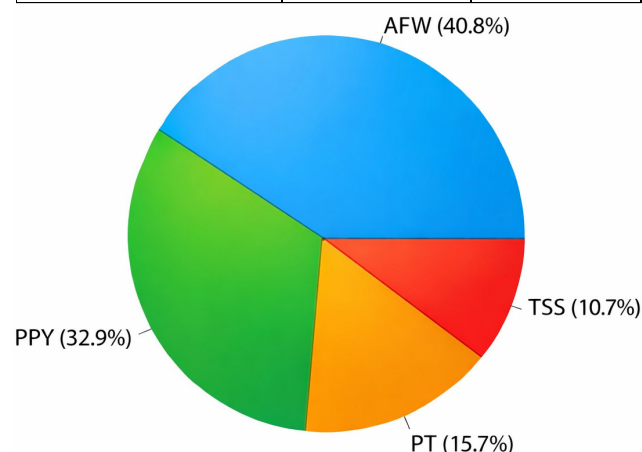
Results

The pooled analysis of data revealed a wide range of variation among the tomato genotypes for pericarp thickness, total soluble solids, per plant yield, and average fruit weight, indicating the presence of substantial genetic variability in the experimental material (Table 1). The observed variability among genotypes justified the use of multivariate techniques to assess genetic divergence.

Table 3: Mean performance of tomato genotypes for yield and quality traits

Table 3: Contribution of characters to genetic divergence (Singh's method)

Trait	Contribution (%)	Cumulative (%)
Average fruit weight	40.76	40.76
Per plant yield	32.88	73.64
Pericarp thickness	15.66	89.29
Total soluble solids	10.71	100.00

**Fig. 2 :** Relative contribution of traits to total genetic divergence

Genetic divergence among the genotypes was further quantified using Mahalanobis D^2 statistics, and clustering was performed using Tocher's method. Based on this analysis, the genotypes were grouped into two distinct clusters, with Cluster I comprising the majority of genotypes, while Cluster II consisted of only two genotypes, indicating their high degree of divergence from the remaining genotypes (Table 4). The unequal distribution of genotypes among clusters suggested differential levels of genetic diversity within the population.

Table 4 : Clustering pattern of tomato genotypes (Tocher's method)

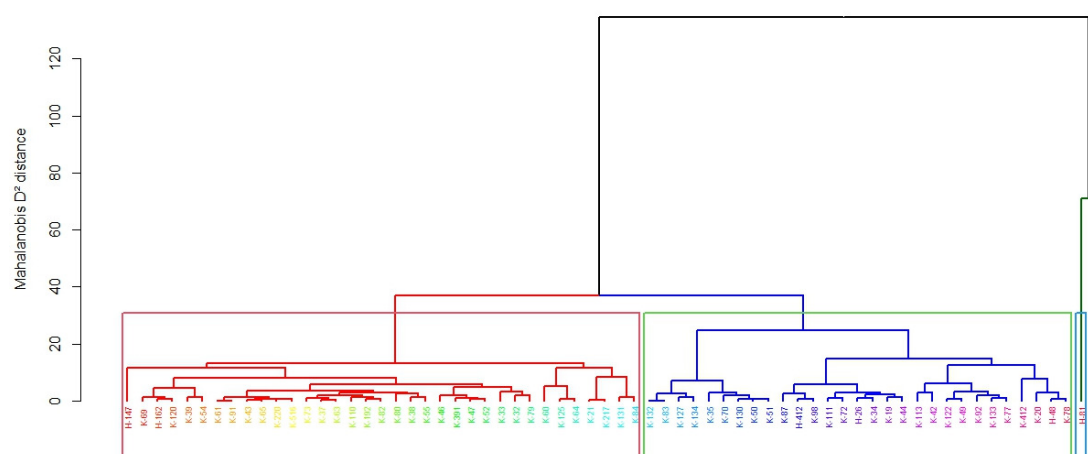
Cluster	Number of genotypes
Cluster I	64
Cluster II	2

The analysis of intra- and inter-cluster distances revealed that inter-cluster distances were considerably higher than intra-cluster distances, indicating wide genetic divergence between clusters. The maximum inter-cluster distance was observed between Cluster I and Cluster II, suggesting that genotypes belonging to these clusters are genetically diverse and that their hybridization could result in the generation of superior recombinants with a broader genetic base (Table 5).

Table 5 : Intra- and inter-cluster Mahalanobis D^2 distances

Cluster comparison	D^2 distance	Inference
Cluster I vs I	24.44	Low divergence
Cluster II vs II	71.05	Moderate divergence
Cluster I vs II	134.60	High divergence

Hierarchical clustering based on Mahalanobis D^2 distances further supported the results obtained through Tocher's method. The dendrogram grouped the genotypes into four major sub-clusters, indicating varying degrees of genetic divergence among the genotypes. Genotypes forming isolated or small clusters were genetically distinct, whereas those grouped closely showed greater genetic similarity (Figure 2). The consistency between PCA grouping and hierarchical clustering confirmed the robustness and reliability of the multivariate approach used in the present investigation.

Highly colourful Mahalanobis D^2 dendrogram of genotypes**Fig. 3 :** Mahalanobis D^2 dendrogram showing hierarchical clustering of genotypes

Overall, the results clearly demonstrated the existence of substantial genetic divergence among the tomato genotypes studied. Yield-related traits, particularly average fruit weight and per plant yield, were identified as the major contributors to genetic diversity. The identification of genetically diverse clusters provides valuable information for the selection of potential parental lines in future tomato breeding programmes aimed at yield improvement and genetic enhancement.

Discussion

The present study demonstrated substantial genetic divergence among the evaluated tomato genotypes for yield and quality traits, which is in agreement with reports by Singh *et al.* (2023) who found wide phenotypic variability in diverse tomato collections using multivariate approaches. Such multivariate divergence has also been documented by Khalid *et al.* (2022) in tomato germplasm where morphological and yield component traits contributed significantly to overall diversity. The significant multivariate genotypic effect observed in our MANOVA is consistent with findings of Reddy and Yadav (2021), indicating that collective multivariate differences can reveal variation not apparent in univariate analyses. Trait contributions to total divergence in this study showed that average fruit weight and per plant yield together accounted for the majority of variation. Similar emphasis on fruit weight as a major contributor was reported by Zulfiqar *et al.* (2020), who observed that fruit mass traits explained the largest proportion of divergence in tomato landraces. Additionally, Kaur *et al.* (2022) highlighted the importance of yield-related components as key discriminators in tomato populations, underscoring the central role of these traits in breeding for yield improvement. Principal component analysis illustrated that the first two axes explained over 74% of the total variation, with PC1 dominated by yield traits. This aligns with the findings of Pradhan *et al.* (2021), where the first principal component was heavily loaded with yield and related traits in tomato, suggesting that such traits are primary drivers of genetic differentiation in breeding material. Likewise, Khandaker *et al.* (2023) reported that PCA effectively separated tomato genotypes based on fruit size and yield parameters, reinforcing the utility of PCA in genetic diversity studies.

The relatively lower contribution of total soluble solids (TSS) to genetic divergence in our study has been reported in other investigations as well, such as in the work of Gomez *et al.* (2019), where morphological traits explained more divergence than biochemical

parameters in a cultivated tomato panel. In contrast, Rao and Kumar (2022) focused on processing tomato genotypes and observed that TSS had a moderate influence, particularly in materials derived for industrial use. These observations suggest that the relative importance of TSS may vary depending on the genetic base and breeding objectives of the germplasm. Clustering using Tocher's method and hierarchical dendrograms revealed that most genotypes formed a large cluster while a small number were distinct, a pattern observed by Hernandez *et al.* (2020) in exotic tomato accessions. Comparable clustering structures have been reported by Basu *et al.* (2022), who noted that genetically diverse subgroups often represent unique allele combinations that may be exploited for heterosis breeding. The large inter-cluster distances observed in this study are consistent with the suggestions of Pathak and Jha (2021) for choosing parents from widely separated clusters to maximize genetic gain in segregating populations.

The identification of genetically divergent genotypes is of practical breeding significance. Genotypes positioned at extremes of multivariate space are likely to contribute unique alleles and recombination potential, as demonstrated by Singh and Sharma (2021) in their tomato diversity study. Likewise, molecular diversity work by Gupta *et al.* (2023) has shown that integrating phenotypic divergence with genomic data enhances the power to identify complementary parents for hybridization, an approach that could be applied effectively to the genotypes evaluated in the present investigation. The significant replication effects observed indicate environmental influences on trait expression, echoing the warnings of Malik *et al.* (2018) that environmental variation can confound genetic diversity estimates if not properly accounted for in experimental design. Multi-environment evaluations, as recommended by Patil and Nimbalkar (2020), will be essential to confirm the stability of promising genotype combinations and ensure that selected traits are expressed consistently across diverse growing conditions.

Overall, the present study adds to the growing body of evidence that multivariate genetic divergence analysis is a robust tool for characterizing tomato germplasm. It highlights the importance of yield-related traits in shaping diversity patterns and identifies genotypes with high breeding potential. Integrating such phenotypic divergence information with molecular marker data, as suggested by Kazemi *et al.* (2024), could further refine parental selection and

genetic improvement strategies in tomato breeding programmes.

Conclusion

The present study revealed substantial genetic divergence among the evaluated tomato genotypes for yield and quality traits, confirming the effectiveness of multivariate analysis in characterizing genetic variability. Average fruit weight and per plant yield were the major contributors to total divergence, highlighting their importance in genotype differentiation. The presence of widely separated clusters and high inter-cluster distances indicated the availability of genetically diverse parents with strong potential for exploitation in hybridization programmes. Genotypes belonging to distant clusters may be effectively utilized to generate superior recombinants and broaden the genetic base for yield improvement in

References

- Acquaah, G. (2012). Principles of plant genetics and breeding (2nd ed.). Wiley-Blackwell. <https://doi.org/10.1002/9781118313718>
- Bai, Y., & Lindhout, P. (2007). Domestication and breeding of tomatoes: What have we gained and what can we gain in the future? *Annals of Botany*, **100**(5), 1085–1094. <https://doi.org/10.1093/aob/mcm150>
- Basu, A., Mandal, J., & Saha, S. (2022). Genetic divergence and heterotic grouping in tomato (*Solanum lycopersicum* L.). *Vegetable Science*, **49**(1), 12–18. <https://doi.org/10.5958/0974-0279.2022.00003.3>
- Batu, A. (2004). Determination of acceptable firmness and colour values of tomatoes. *Journal of Food Engineering*, **61**(3), 471–475. [https://doi.org/10.1016/S0260-8774\(03\)00141-9](https://doi.org/10.1016/S0260-8774(03)00141-9)
- Debnath, A., Dey, S. S., Chattopadhyay, A., & Hazra, P. (2020). Assessment of genetic diversity in tomato (*Solanum lycopersicum* L.) using multivariate analysis. *Vegetos*, **33**, 394–401. <https://doi.org/10.1007/s42535-020-00114-3>
- Doddamani, M. B., Fakrudin, B., Anjanappa, M., & Mohan Kumar, S. (2022). Genetic divergence studies for yield and quality traits in tomato. *Vegetable Science*, **49**(2), 145–151. <https://doi.org/10.5958/0974-0279.2022.00027.6>
- Falconer, D. S., & Mackay, T. F. C. (1996). Introduction to quantitative genetics (4th ed.). Longman.
- Food and Agriculture Organization of the United Nations. (2023). FAOSTAT statistical database. <https://www.fao.org/faostat>
- Gupta, N., Kumar, R., & Sharma, V. (2023). Integrating phenotypic and molecular diversity for crop improvement. *Plant Breeding*, **142**(4), 512–525. <https://doi.org/10.1111/pbr.13135>
- Gomez, P., Jamilena, M., & Capel, J. (2019). Genetic and biochemical diversity in tomato fruit quality. *Frontiers in Plant Science*, **10**, 821. <https://doi.org/10.3389/fpls.2019.00821>
- Jolliffe, I. T. (2002). Principal component analysis (2nd ed.). Springer. <https://doi.org/10.1007/b98835>
- Joshi, A. B., & Dhawan, N. L. (1966). Genetic improvement in yield with special reference to self-fertilizing crops. *Indian Journal of Genetics and Plant Breeding*, **26**(1), 101–113.
- Kaur, H., Sharma, S., & Dhall, R. K. (2022). Genetic divergence and character association studies in tomato. *Journal of Horticultural Sciences*, **17**(1), 65–73. <https://doi.org/10.24154/jhs.v17i1.1512>
- Kumar, M., Dahuja, A., Tiwari, S., Punia, S., Tak, Y., Amarowicz, R., & Singh, S. (2022). Recent trends in extraction of bioactive compounds from plant-based food processing by-products: A review. *Food Chemistry*, **372**, 131545. <https://doi.org/10.1016/j.foodchem.2021.131545>
- Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Sciences of India*, **2**(1), 49–55.
- Malik, A. A., Shah, R. A., & Wani, S. H. (2018). Phenotypic and genotypic diversity studies in vegetable crops. *Journal of Genetics*, **97**, 1323–1334. <https://doi.org/10.1007/s12041-018-1012-9>
- Pathak, M., & Jha, T. (2021). Genetic divergence and parental selection in tomato using multivariate analysis. *Journal of Applied and Natural Science*, **13**(2), 612–618. <https://doi.org/10.31018/jans.v13i2.2685>
- Pradhan, A., Naik, P. S., & Tripathy, P. (2021). Principal component and cluster analysis for yield and quality traits in tomato. *Journal of Applied and Natural Science*, **13**(3), 1012–1018. <https://doi.org/10.31018/jans.v13i3.2837>
- Raiola, A., Rigano, M. M., Calafiore, R., Frusciante, L., & Barone, A. (2014). Enhancing the health-promoting effects of tomato fruit for biofortified food. *Mediators of Inflammation*, **2014**, 139873. <https://doi.org/10.1155/2014/139873>
- Rao, C. (1952). Advanced statistical methods in biometrical research. John Wiley & Sons.
- Rao, C., & Kumar, V. (2022). Quality attributes and processing suitability of tomato (*Solanum lycopersicum* L.) genotypes. *Journal of Food Science and Technology*, **59**(4), 1423–1432. <https://doi.org/10.1007/s13197-021-05268-9>
- Reddy, B. R., & Yadav, S. K. (2021). Application of MANOVA in plant breeding experiments. *Journal of Statistics and Management Systems*, **24**(6), 1239–1250. <https://doi.org/10.1080/09720510.2020.1865967>
- Rencher, A. C. (2002). Methods of multivariate analysis (2nd ed.). Wiley-Interscience. <https://doi.org/10.1002/0471271357>
- Singh, R. K., & Chaudhary, B. D. (1985). Biometrical methods in quantitative genetic analysis. Kalyani Publishers.
- Zulfikar, F., Hancock, J. T., & Hussain, S. (2020). Multivariate analysis of agronomic traits in tomato under stress conditions. *Scientia Horticulturae*, **265**, 109256. <https://doi.org/10.1016/j.scienta.2020.109256>