

## Research Article

# Genetic diversity exploration in (*Capsicum annuum* L.) sweet pepper germplasm through chemical and morphological characterization

Nadra Khan<sup>1\*</sup>, Kashif Khaqan<sup>2</sup> and Muhammad Jamil Ahmed<sup>1</sup>

1. Department of Horticulture, Faculty of Agriculture, University of Poonch Rawalakot, 12350, AJK, Pakistan

2. Federation University, Ballarat, Victoria 3350, Australia

\*Corresponding author's email: [nadrakhan.upr@gmail.com](mailto:nadrakhan.upr@gmail.com)

### Citation

Nadra Khan, Kashif Khaqan and Muhammad Jamil Ahmed. Genetic diversity exploration in (*Capsicum annuum* L.) sweet pepper germplasm through chemical and morphological characterization. Pure and Applied Biology. Vol. 14, Issue 2, pp692-709. <http://dx.doi.org/10.19045/bspab.2025.140066>

Received: 12/01/2025

Revised: 06/03/2025

Accepted: 08/03/2025

Online First: 12/03/2025

### Abstract

The study assessed genetic variation in sweet pepper (*Capsicum annuum*), valued for its economic importance and diverse fruit types. With broad genetic diversity, it holds the potential for enhancing yield and nutrition through breeding. Ninety-five *Capsicum* genotypes were assessed for nine morpho-agronomic and six biochemical traits across different growth stages using Biodiversity International descriptors and the Association of Official Analytical Chemists. The maximum yield/plant (1809.4 g) was achieved by genotype 36595, and genotype 36606 exhibited the highest fruit weight (100.0 g). The maximum (11.0%) protein was recorded in 36582, while genotype 36633 had the highest fiber content (33.01 %). Correlation analysis revealed a significant positive correlation between fruit weight and width ( $r=0.75$ ), and yield/plant correlated positively with plant height ( $r=0.85$ ). Hierarchical cluster analysis identified six distinct clusters, with genotype 36584 as an outlier in Cluster I, genotype 36642 in Cluster II, genotype 30878 in Cluster III, genotype 36221 in Cluster IV, and genotype 36117 in Cluster VI. Principal component analysis (PCA) identified nine Principal components with eigenvalues  $>0.7$ , collectively explaining 89.97% of the variance. The first two principal components (PC1 and PC2) described 41.33% of the overall variance, with PC1 explaining 27.90% and PC2 13.43%. Genotypes 36606, 36556, 36595, and 36582, showed higher diversity and were selected as promising candidates for breeding programs. This study highlights the genetic diversity of *Capsicum annuum*, aiding superior genotype selection for crop improvement. The findings contribute to genetic resource conservation and breeding strategies, potentially enhancing yield and nutritional content, addressing food security, and promoting sustainable agriculture.

**Keywords:** *Capsicum*; Correlation; Germplasm; Genetic diversity; HCA; PCA; Proximate; Yield

### Introduction

*Capsicum* is a member of the Solanaceae family and includes various types of peppers, which are commonly referred to as chili, red pepper, and green bell pepper depending on their specific type and geographic location [1]. Evidence from archaeological studies suggests that *Capsicum* has been used as a condiment by humans for around 6,000 years [2]. This

genus is indigenous to the regions of South and Central America, with Mexico identified as a key region for its domestication and genetic diversity [3]. Among the *Capsicum* species, twenty-five are wild, out of it 5 are domesticated, with *Capsicum annuum* L. being the most widely cultivated [4].

Its plants are generally small, perennial shrubs characterized by greenish-white or

white flowers, with diverse fruit shapes and sizes [5]. They have a rapid growth cycle, maturing in 3 to 4 months, and are capable of both cross-pollination and self-pollination, facilitated by insects or wind. Optimal growth conditions include daytime temperatures between 20°C and 30°C and loamy soils that are well-drained and have a pH range of 5.5 to 6.8, enriched with lime [5].

Currently, consumers are increasingly focused on vegetables that offer high quality and nutritional value [6]. Pepper is widely recognized as a valuable vegetable for human health [7]. Its fruit is high in vitamins and antioxidants including flavonoids, carotenoids, and polyphenols, which provide substantial nutritional benefits [8]. Nutritional content may differ depending on various factors such as cultivation conditions, genotype, and fruit maturity. For instance, 100 grams of fresh hot peppers typically contain about 88 grams of water, 40 kcal of energy, and a range of vitamins and minerals [9]. In contrast, dried peppers have different nutritional profiles, including a moisture content of 13.4%, fat content of 11.9%, protein of 12.8%, carbohydrates of 56.2%, and fiber of 22.5% [10].

The sweet bell pepper available in a variety of colors such as black, green, red, purple, and orange, is cultivated worldwide for both its flavor and health benefits since ancient times [11]. The pepper's striking form and vivid hues have earned it the title of "The Christmas ornaments of the vegetable world" [12]. These peppers are consumed fresh or processed into various forms, including canned, fermented, frozen, pickled, and as seasoning as paprika powder [13]. Traditionally, phenotypic traits like fruit weight, plant height, and fruit shape have been utilized to categorize and distinguish *Capsicum* genotypes [14]. However, the complex taxonomy of *Capsicum* poses challenges for classification based solely on morphological and agronomic traits [15].

Furthermore, assessing genetic diversity is crucial for identifying beneficial traits within germplasm, which can be leveraged for breeding and improving *Capsicum* varieties [16]. Although modern high-yielding varieties are developed for optimal conditions, they may not be suitable for low-income farmers in marginal environments that face variable stress [17]. Traditional varieties, adapted to low-input agriculture, can significantly contribute to food security in these regions [18].

In addition, germplasm collections are essential for conserving genetic diversity in important crops, providing valuable resources for plant breeders, and researchers [19]. Despite the time and cost involved in genetic evaluation, effective characterization enhances the benefits derived from germplasm conservation. The domestication and breeding of *Capsicum* have focused on traits such as pericarp thickness, pungency, fruit shape, taste, color, and suitability for drying into powder, ensuring the crop's persistence and global cultivation [20].

While significant research efforts have been undertaken on the metabolite composition of pepper fruits, particularly focusing on carotenoids, capsaicinoids, vitamins, and phenolic compounds, there is a notable gap in information regarding other nutritional aspects, such as proximate composition [21]. In addition, there is an increasing demand for new pepper genotypes that produce more flavorful fruits with vibrant colors and higher concentrations of bioactive compounds beneficial to human health. In response to this demand, breeding programs must formulate strategies to enhance the content of key substances in *Capsicum* fruits to access new market opportunities. To make these strategies effective, it is essential to address the gap in the phenotypic and biochemical characterization studies of *Capsicum* peppers.

This study aims to evaluate the phenotypic variability among 95 *Capsicum* genotypes across various quantitative traits and

compare these patterns with representative genotypes from different origins. Additionally, the chemical diversity within the *Capsicum* gene pool was evaluated through proximate analysis. The ultimate objective is to identify promising genotypes for further breeding to enhance both crop performance and nutritional quality.

## Materials and Methods

### Plant material and experimental conditions

The study utilized 95 sweet pepper genotypes (Fig. 1) obtained from the Gene Bank, Bio-Resources and Conservation Institute (BCI) at the National Agricultural Research Centre (NARC) in Islamabad.

*Capsicum* seeds were initially cultivated in the greenhouse at the University of Poonch, Rawalakot. The seedlings were subsequently transplanted into the field using an augmented design, with 45 cm plant spacing and 75 cm row spacing. Field preparation involved ploughing the soil twice to achieve a level surface. Fertilization was performed with applications of nitrogen and phosphorus @ 100 kg/ha and 60 kg/ha, respectively. To ensure optimal crop health, essential cultural practices including hoeing, earthing-up, weeding, and pest control were conducted as needed.



Figure 1. Map of geographical representation of genotypes

### Morpho-agronomic traits (Quantitative)

A detailed set of parameters, covering various growth stages, was developed using the *Capsicum* descriptors from Biodiversity International's previous guidelines. [22] for documenting morphological traits. Proper species identification is crucially dependent on phenotypic information provided by morphological descriptors [23].

The collected data comprised quantitative parameters, primarily emphasizing morphological characteristics. The height of each pepper plant was recorded in centimeters (cm) using a meter rod. Leaf measurements were taken from ten randomly selected, fully mature plants, with width recorded at the broadest point

(cm). The length of ten mature leaves, from tip to petiole, was measured using a measuring tape (cm). A Vernier caliper was used to measure fruit width at its widest section (cm) and fruit length (cm). The weight of individual fruits was determined in grams (g) using an electronic balance. Fruit wall thickness was assessed in millimeters (mm) with a Vernier caliper. The number of locules was identified by direct counting. Fruit yield per plant (g) was calculated as the total weight of all fruits from each genotype.

### Proximate analysis (Biochemical Traits)

Mature fruits were harvested from the agricultural fields of the University of Poonch, Rawalakot, Azad Kashmir,

Pakistan. The collected samples were then transported to the Food Science and Technology Laboratory at the University of Poonch, Rawalakot, AJK, for biochemical evaluation (Fig. 2a & b). Various parameters of the fruit pulp were analyzed, including moisture content, which was determined using Method No. 925.05 as described by [24]. The ash content was

assessed following Method No. 940.26 [24], while crude fat content was analyzed using Method No. 983.23 [24]. Crude fiber was measured based on the procedure outlined in [24]. Crude protein estimation was conducted using Method No. 920.10 [24]. Carbohydrate content was calculated using the Difference Method, as specified in [24].



**Figure 2a. Genotype 36643 Sample drying for analysis, 2b. Genotype 36606 Sample drying for analysis**

### Statistical analysis

The quantitative data was analyzed using various multivariate statistical methods. Descriptive statistics were applied to determine the quantitative and chemical traits' mean, standard error, standard deviations, and variance. Pearson's correlation coefficients were calculated to assess the relationships between traits [25]. Hierarchical cluster analysis was conducted to categorize genotypes based on their similarity across multiple traits, which helped in identifying groups with comparable phenotypic and biochemical characteristics. Additionally, Principal Component Analysis (PCA) was performed using PAST software 4.17 to reduce dimensionality and highlight the principal components that account for most of the variability among the genotypes.

### Results and Discussion

#### Variability based on morpho-agronomic traits

Morphological characterization plays a key role in differentiating genotypes based on

physical traits, essential for improving critical agricultural attributes. However, it is affected by environmental factors [26]. Despite this sensitivity, it enables farmers to evaluate crop performance across different conditions. Although the genetic diversity within *Capsicum* remains underutilized, it holds promise for future selection and breeding efforts. The mean values and ranges for nine traits across all genotypes are presented in (Table 1). Plant height is a critical factor in crop management as it influences decisions regarding plant density and spacing, directly affecting yield. This trait is often selected based on assessments of mean, standard deviation, range, and variance [27]. The maximum plant height was recorded in genotype 36556 (69.3 cm). The lowest plant height (32.4 cm) was depicted in genotype 36605. Our findings are similar to [12]. Environmental factors for instance temperature, moisture levels, and soil fertility significantly influence the growth of pepper plants [5].

Leaves are vital for photosynthesis, significantly contributing to plant growth and survival. The highest leaf width of (3.6 cm) was observed in genotypes 36595, 36550, 36594, 36599, 36603, and 36649. A minimum leaf width was recorded (1.5 cm) in genotype 36625 and (1.6 cm) in genotype 36584. Maximum (5.6 cm) leaf length was obtained in genotype 36596, while minimum (2.4 cm) leaf length was scored in genotype 33832. Research has indicated that differences in leaf characters are often the result of adaptations to different environmental conditions [28].

[29] stated that certain varieties produce fruit with a greater number of locules, contributing to increased fruit width and size. The highest locule number, (4.2), was observed in genotypes 33832 and 36606. Conversely, the lowest locule number, (2.0), was recorded in genotypes 36605, 36581, 36586, 36585, 36580, 36582, 36576, 36575, 36568, 30880, 36112, 36548, 36547, 36546, 36117, 36634, 36626, and 36631.

The longest fruit was (9.9 cm) in genotype 36619. Lowest fruit length (4.0 cm) was recorded in 36621. The widest fruit (7.0 cm) was seen in genotypes 36111, 30884, and 36594. The smallest fruit widths, measuring (2.3 cm) seen in 36550, 36631, 36579, and 36586 genotypes. Fruit size is a critical trait for categorizing *Capsicum* varieties into small, medium, and large classifications. The observed variability in both fruit length and width among the *Capsicum* genotypes is consistent with earlier studies by [12], who reported substantial differences in these traits across various pepper cultivars and varieties. The variation in these traits is essential for their distinct applications in the pepper industry. Typically, fruits with longer lengths are favored for making paprika and are also utilized in cosmetics and medicinal products, whereas small-sized fruits are commonly used for eating fresh or sauce production [30].

The maximum thickness of 4.8 mm was recorded in genotypes 30883 and 36606,

while the minimum thickness of 0.9 mm was observed in genotype 36568. Fruit wall thickness is a key trait influencing both yield and marketability [31]. Fruits with thicker walls are particularly valued in the fresh market due to their extended post-harvest shelf life, enhanced visual appeal, and greater resistance to diseases and pests [32]. Thicker, and firmer fruits are generally favored for harvesting, transportation, and storage. This characteristic is also crucial in sweet pepper breeding for paprika, especially considering its importance in processing and industrial uses [33].

Significant variation in fruit weight was observed among *Capsicum* genotypes, spanning small, medium, and large categories. Fruit weight averaged  $49.3 \pm 1.7$  g, ranging from 21.0 g to 100.2 g, and a coefficient of variation (CV) of 32.6% (Table 1). The CV indicates a high degree of variability in fruit weight among the genotypes, reflecting considerable differences in fruit size. The maximum fruit weight (100.2 g) was recorded in genotype 36606. The minimum fruit weight (21.0 g) was depicted in genotype 36591. The high CV suggests that fruit weight is a trait with substantial variability, which can be advantageous in breeding programs. This variability is crucial for selecting genotypes with desirable fruit-size traits. It is a quantitatively inherited trait, influenced by multiple genetic loci, with some loci having minor effects and others major effects [34].

The highest yield was recorded in genotype 36595, which produced (1809.4 g per plant). In contrast, genotype 36605 exhibited the lowest yield of (282.4 g per plant). The high yields in certain genotypes are likely due to the substantial weight of individual fruits. Yield is a multifaceted characteristic affected by both genetic makeup and environmental conditions. Climatic factors, such as temperature and rainfall, play essential role in seed germination, growth, and fruiting. High temperatures can adversely affect fruit sets, while low temperatures can hinder

germination and growth. The present study revealed substantial variation in fruit yield across genotypes, underscoring the genetic

diversity and the potential for genetic advancements in breeding programs, in line with findings by [35].

**Table 1. Descriptive summary of morphological traits**

| Traits                    | Mean±SE     | SD    | Range |        | Variance | CV % |
|---------------------------|-------------|-------|-------|--------|----------|------|
|                           |             |       | Min   | Max    |          |      |
| Plant height (cm)         | 54.7±0.9    | 8.6   | 32.4  | 69.3   | 73.2     | 15.6 |
| Leaf Width (cm)           | 2.8±0.1     | 0.6   | 1.5   | 3.6    | 0.3      | 15.6 |
| Leaf length (cm)          | 4.2±0.1     | 0.9   | 2.4   | 5.6    | 0.8      | 20.7 |
| No. locules               | 3.0±0.1     | 0.6   | 2.0   | 4.2    | 0.4      | 20.4 |
| Fruit width (cm)          | 4.4±0.1     | 1.4   | 2.3   | 7.0    | 1.9      | 21.9 |
| Fruit length (cm)         | 6.6±0.1     | 1.4   | 4.0   | 9.9    | 2.1      | 31.0 |
| Fruit wall thickness (mm) | 3.6±0.1     | 0.9   | 0.9   | 4.8    | 0.9      | 26.0 |
| Single fruit weight (g)   | 49.3±1.7    | 16.1  | 21.0  | 100.2  | 258.8    | 32.6 |
| Yield/plant (g)           | 1107.1±39.9 | 389.2 | 282.4 | 1809.4 | 151496.0 | 35.2 |

#### Variability based on biochemical traits

Water is the primary component of *Capsicum* fruits, significantly affecting their quality after harvest, as well as during storage and processing stages, which ultimately determines the fruits' usability. The highest moisture level (13.00 %) was observed in genotypes 33826 and 36115, followed by (12.87%) in genotype 36566, (12.34 %) in genotype 36594, (12.32%) in genotypes 36580 and 33829, (12.31 %) in genotype 36571, and (12.23 %) in genotype 36578. Minimum moisture content (8.01 %) was observed in genotype 30882 (Table 2). The elevated moisture content in pepper samples implies that without appropriate preservation, peppers are susceptible to microbial growth and spoilage, thus dehydration is necessary to prolong their shelf life. These findings align with research [21, 36, 37].

The highest protein level (11.0 %) was observed in genotype 36582, followed closely by (10.93 %) in genotype 36582 and (10.90 %) in genotype 36654. The lowest

protein content of (8.2 %) was recorded in genotype 36611, with (8.22 %) in genotypes 36618 and 36577, and (8.23 %) in genotype 36653. Crude protein content is calculated based on the assumption that proteins contain approximately 16% nitrogen. These findings are consistent with the research of [21].

The highest ash levels were recorded (16 %) in genotype 36578, followed by (15.82 %) in genotype 36606, (15.45 %) in genotype 30883, (15.34 %) in genotype 30882, (14.57 %) in genotype 36111, (13.24 %) in genotype 30876, and (13.21 %) in genotype 36596. The lowest ash content of (5.21 %) was observed in genotype 36112. The increase in ash content during ripening is likely due to the migration of inorganic ions to active growth regions. This observation is aligned with the research findings of [21].

Highest fiber content (33.01 %) was observed in genotype 36633, followed by (32.6 %) in genotype 36567, (32.43 %) in genotype 33826, and (32.14 %) in genotype

36610. The lowest protein (8.20 %) was observed in 36611. Highest fiber content (33.0 %) was obtained in genotype 36633, and (31.45 %) was observed in genotype 36566. whereas lowest fiber content (22.34 %) was found in genotype 36604, and (22.35 %) in genotype 36626. Fiber content can decrease during storage due to structural and biochemical changes, which may affect the texture and consumer acceptability of the pepper. These findings align with the results reported by [36].

The highest fat content (5.50 %) was observed in genotype 30884, followed by (5.34 %) in genotype 36609, and (4.31 %) in genotypes 36659 and 36110. Other notable values of fat include (4.23 %) in genotype 36562, (4.09 %) in genotype 36599, and (4.0 %) in genotype 36607. The lowest fat content (0.39 %) was found in genotype 36586. The observed fat content aligns with the findings of [21]. Elite genotypes with high-fat content exhibit consistent results across biochemical screenings, suggesting their potential for further evaluation and inclusion in genetic improvement programs.

Maximum ash content (16.0 %) was obtained in 36578, whereas minimum (5.21

%) was observed in genotype 36112. These findings align with [21, 36, 37].

The highest carbohydrate level (52.77 %) was found in genotype 36556, followed by (51.88 %) in genotype 30880, (51.58 %) in genotype 36112, and (50.30 %) in genotype 36605. The lowest carbohydrate levels were recorded at (34.11 %) in genotype 36625, (35.11 %) in genotype 33829, (35.43 %) in genotype 36607, and (35.68 %) in genotype 36633 (Table 2). The carbohydrate levels observed in this study are consistent with the research findings of [36-38]. Carbohydrates, the second most significant component in peppers after water, are essential for evaluating fruit quality traits. They are vital to the nutritional value, flavor, and functional properties of both fresh and processed foods. However, despite their high levels, carbohydrates may not always be fully digestible or nutritionally assessable, as many remain indigestible in the body [39]. Overall, there were notable differences in all morpho-agronomic and biochemical traits, reflecting substantial genetic variation among the accessions.

**Table 2. Descriptive summary of biochemical traits**

| Traits     | Mean±SE  | SD  | Range   |         | Variance | CV % |
|------------|----------|-----|---------|---------|----------|------|
|            |          |     | Minimum | Maximum |          |      |
| Moisture % | 9.6±0.1  | 1.3 | 8.01    | 13.0    | 1.7      | 13.8 |
| Protein %  | 9.1±0.1  | 0.7 | 8.20    | 11.0    | 0.4      | 7.3  |
| Fiber %    | 26.9±0.3 | 2.9 | 22.34   | 33.01   | 8.6      | 10.9 |
| Fat %      | 2.0±0.1  | 0.9 | 0.39    | 5.50    | 0.9      | 47.1 |
| Ash %      | 9.2±0.3  | 2.6 | 5.21    | 16.0    | 6.8      | 10.9 |
| CHO %      | 43.4±0.4 | 4.3 | 34.11   | 52.77   | 18.6     | 9.9  |

### Cluster analysis of morpho-agronomic and biochemical traits

The tree diagram (Fig. 3) was constructed based on 15 morpho-agronomic and biochemical traits utilizing a Euclidean distance similarity matrix. Partitioning of a cluster at (360) linkage distance, gathered into 6 distinct clusters, [40] also reported 6

clusters in 32 genotypes of chili. Cluster CI is composed of genotypes of approximately 13.68 % of the total genome; 33832, 30883, 36115, 36545, 36545, 36643, 36593, 36578, 36584, 36654, 23479, 36580, 36573, 36598. A genotype 36584 was an outlier due to high genetic distance while

minimum inter-cluster distance was observed between 36584 and 36654.

CII composed of 19 genotypes accounted for 20 % of the whole gene pool. It contained genotypes; 36606, 36111, 36118, 36641, 36554, 36642, 36556, 36595, 36575, 36581, 36576, 36633, 36550, 36613, 36601, 36113, 36608, 36596, and 36583. A genotype 36642 is an outlier due to maximum genetic distance whereas minimum distance was seen in 36575 vs 36581.

Cluster CIII composed of 22 genotypes explained 23.1 % of the contribution of genotypes in this cluster. The genotypes; 36116, 36115, 36602, 36599, 36572, 36575, 36640, 36626, 36634, 36657, 36619, 36557, 36561, 36629, 36562, 36547, 36609, 36119, 36611, 36548, 36659, and 30878. Maximum distance was seen in genotype 30878, which was an outlier while minimum inter-cluster distance was depicted in 36634 vs 36657 and 36619 vs 36557.

The cluster CIV comprised 13 genotypes and accounted for 12.63 % of the whole set of genotypes. The genotypes 36621, 36650, 36112, 36575, 36568, 36603, 36653, 36110, 36647, 30876, 36546, and 365555. Due to maximum genetic distance, genotype 36221 was an outlier, whereas minimum genetic distance was seen in 36603, and 36653.

A cluster CV contained 22 genotypes and accounted for 23.15% of the gene pool. The genotypes; 36625, 36577, 36549, 33829, 36551, 36610, 36553, 30882, 36566, 36638, 36618, 36607, 36594, 33826, 36552, 36604, 36574, 36590, 30884, 36648, 36631, and 36579. Maximum genetic distance was seen among 30882 and 36566. Minimum inter-cluster distance was seen in 36640 and 36574.

A cluster CVI comprised of 7 genotypes and explained 7.60 % of contribution to the whole genome. The genotypes 36117, 30880, 36585, 36582, 36586, 36587, and 36605. A genotype 36117 was an outlier and minimum inter-cluster distance was seen in 36582 vs 36586.

### Correlation coefficient between morpho-agronomic and biochemical traits

The process of selecting traits to achieve genetic improvements in plant breeding can result in higher program costs and increased labor, especially when multiple traits are involved. Correlation coefficients are valuable tools in this scenario, as they help assess the impact of direct and indirect selection [41]. In this investigation, the Pearson correlation coefficient [42] was applied to evaluate the linear relationship between two traits, enabling the selection of one trait based on the other. Pearson's correlation coefficient ( $r$ ) quantifies the direction and magnitude of the linear association between two variables, expressed as the covariance of these variables standardized by their standard deviations.

The scale in (Fig. 4) ranges from -1 to 1, with circle colors representing the direction and strength of correlations between two traits. Blue circles indicate positive correlations, whereas red circles represent negative correlations between the traits. The data reveal that fruit width exhibited the strongest and most significant correlation with the number of locules ( $r = 0.51$ ). Additionally, fruit length was positively correlated with plant height, showing a significant correlation ( $r = 0.31$ ) as reported by [43]. A particularly strong and statistically significant correlation was found between fruit weight and fruit width ( $r = 0.75$ ), similar to findings by [44].

In terms of postharvest handling, selecting varieties with thicker fruit walls is vital, as these fruits demonstrate greater resistance to damage during transport and storage. Fruit wall thickness exhibited a significant positive correlation with the number of locules ( $r = 0.51$ ). Furthermore, single fruit weight showed strong and highly significant correlations with plant height ( $r = 0.50$ ), the number of locules ( $r = 0.58$ ), fruit width ( $r = 0.67$ ), and fruit wall thickness ( $r = 0.65$ ).



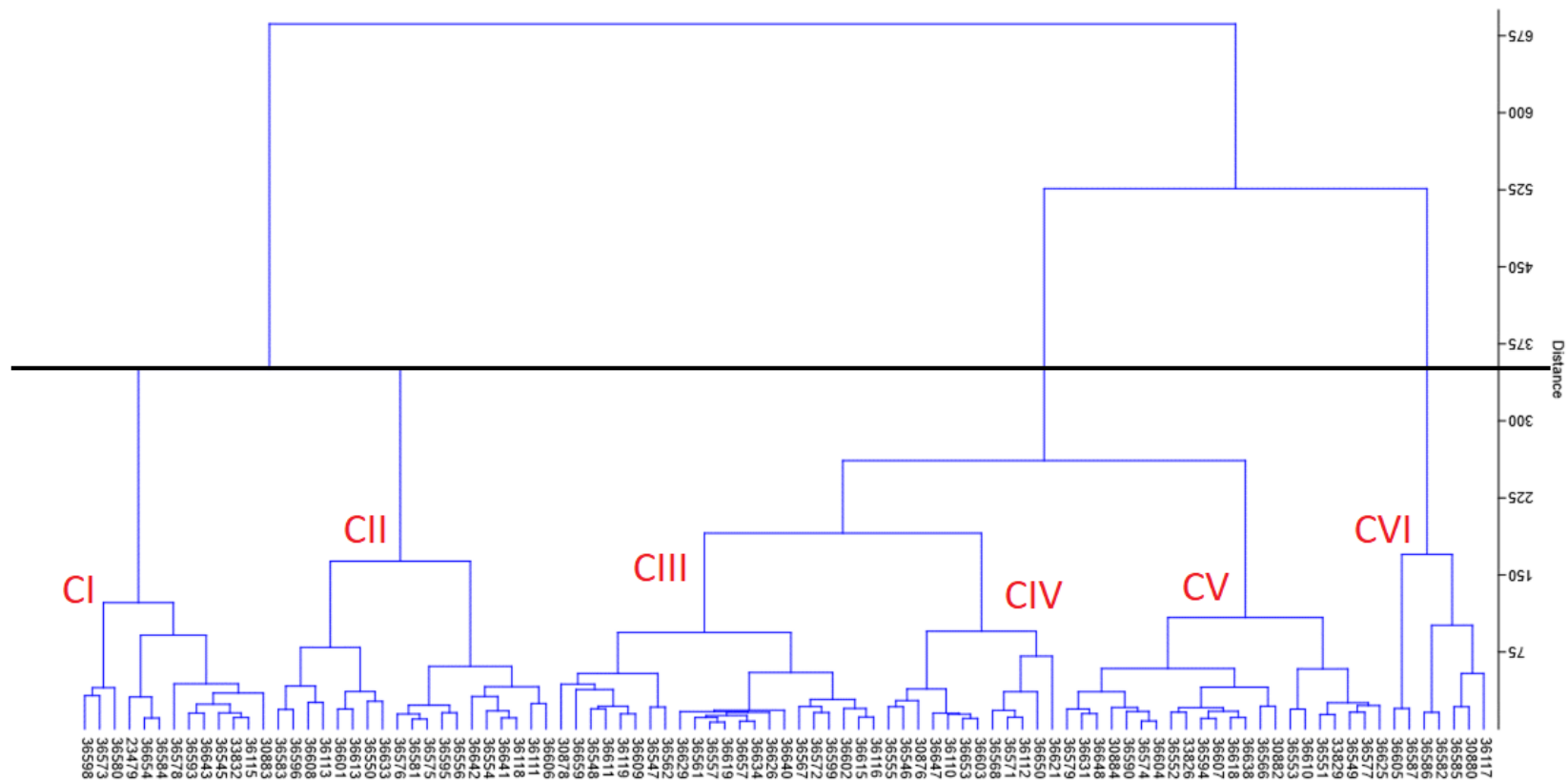
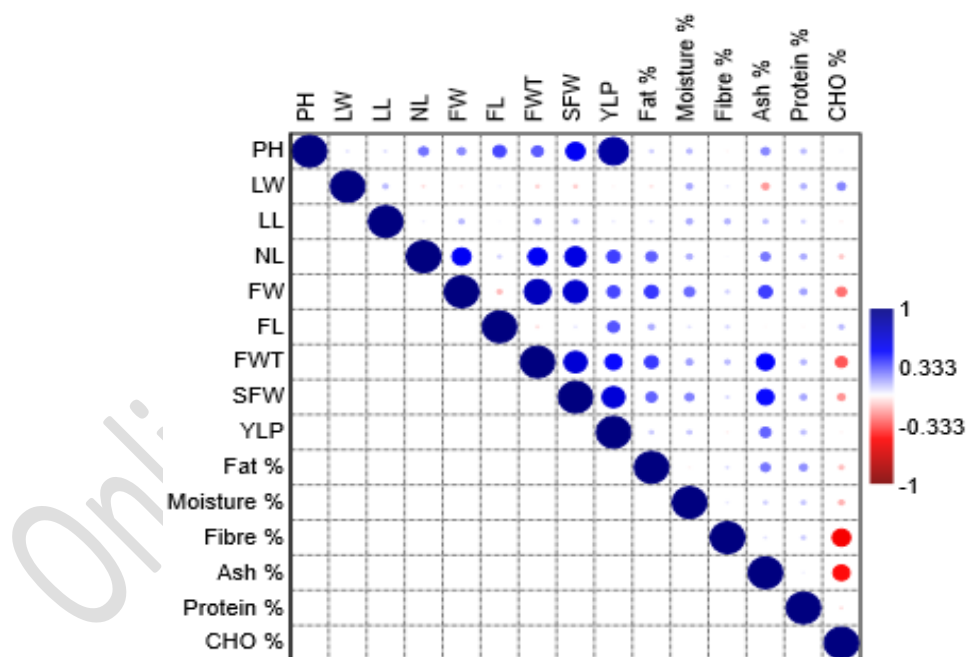


Figure 3. Tree diagram of pepper genotypes based on physico-chemical traits

Yield per plant was another trait with significant correlations, displaying a highly significant positive correlation with plant height ( $r = 0.85$ ) as noted by [43, 44]. Yield per plant also displayed significant and positive correlations with the number of locules ( $r = 0.33$ ) and fruit width ( $r = 0.32$ ), following findings by [43, 44]. Moreover, yield per plant was strongly correlated with fruit wall thickness ( $r = 0.44$ ) and single fruit weight ( $r = 0.63$ ). [45] stated significant and positive correlations between important economic traits, including fruit length and weight, and yield, while [46] observed a significant and positive correlation between fruit weight and yield.

In crop plants, most traits are interconnected, influencing yield in complex ways. These relationships can be either beneficial or detrimental, as highlighted by [47]. For instance, fat

content exhibited significant and positive correlations with both fruit wall thickness ( $r = 0.34$ ) and fruit width ( $r = 0.34$ ), potentially due to a higher seed count in wider fruits, which contributes to increased fat content. Ash content itself demonstrated strong correlations with fruit wall thickness ( $r = 0.48$ ), fruit width ( $r = 0.42$ ), and the number of locules ( $r = 0.33$ ). The strong correlation is likely because wider fruits with thicker fruit walls tend to have more locules, which increases ash content. As the fruit's size and structure expand, so does the ash content. Protein content was positively correlated with leaf length ( $r = 0.09$ ) and fruit wall thickness ( $r = 0.08$ ). On the other hand, carbohydrates displayed significant and negative correlations with fiber ( $r = -0.66$ ), ash content ( $r = -0.62$ ), fruit wall thickness ( $r = -0.46$ ), and fruit width ( $r = -0.38$ ).



**Figure 4.** Correlation coefficient is highly significant at (0.01), and significant at (0.05) level. Plant height (PH), Leaf Width (LW), Leaf length (LL), No. locules (NL), Fruit width (FW), Fruit length (FL), Fruit wall thickness (FWT), Single fruit weight (SFW), Yield/plant (YLP), Carbohydrates (CHO)

Due to the extensive use of *Capsicum* fruits across various sectors of the food industry, breeding programs can focus on enhancing

specific traits based on their end goals. The results suggest that selecting smaller genotypes may be beneficial for producing

fruits with lower water content, which is ideal for making paprika powdered. On the other hand, indirect selection for larger fruits might lead to higher water content, which is preferable for fresh consumption. However, it is important to consider fruit size characters, such as width and length, as there is a negative correlation between these traits and the protein, carbohydrates, and moisture which is often a more critical factor.

### Principal component analysis on the basis of morpho-chemical traits

PCA is a powerful tool for reducing the dimensionality of a dataset while preserving as much variability as possible [48]. In this study, nine principal components (PCs) were identified with eigenvalues greater than 0.7, collectively explaining 89.97% of the total variance among the traits analyzed (Table 3). Of these, the first two principal components (PC1 and PC2) are particularly noteworthy, as they account for 41.33% of the overall variance, with PC1 explaining 27.90% and PC2 explaining 13.43%.

The variability explained by each of the first nine PCs is as follows: PC1 with 27.90%, PC2 with 13.43%, PC3 with 9.48%, PC4 with 9.31%, PC5 with 7.51%, PC6 with 6.16%, PC7 with 6.03%, PC8 with 5.30%, and PC9 with 4.85%. These results highlight the complexity of the trait

structure within the *Capsicum* genotypes, indicating that while PC1 and PC2 capture the most significant patterns in the data, the other PCs are also crucial in representing the diversity among the genotypes. These components likely reflect more nuanced physico-chemical differences that contribute to the overall genetic variation observed.

The PCA findings underscore the importance of considering multiple traits in the analysis, as the additional seven components with high eigenvalues offer a more detailed understanding of the dataset's complexity. This multi-dimensional structure emphasizes the intricate genetic and physico-chemical diversity present in the *Capsicum* genotypes, which is essential for making informed decisions in breeding and selecting varieties for production. As illustrated in (Fig. 5), genotype 36606 exhibited the highest fruit weight, while genotype 36556 had the greatest plant height and fruit length. Genotype 36582 exhibited the maximum protein content, and genotype 36595 demonstrated the highest yield per plant. These genotypes, positioned further from the center, indicate higher diversity and could be valuable candidates for selection as parent plants in breeding programs.

**Table 3. Eigenvalue for physico-chemical traits**

| PC                             | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     |
|--------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| <b>Eigenvalue</b>              | 4.1   | 2.01  | 1.42  | 1.39  | 1.12  | 0.92  | 0.90  | 0.79  | 0.72  |
| <b>% variance</b>              | 27.90 | 13.43 | 9.48  | 9.31  | 7.51  | 6.16  | 6.03  | 5.30  | 4.85  |
| <b>Cumulative Variance (%)</b> | 27.90 | 41.33 | 50.81 | 60.12 | 67.63 | 73.79 | 79.82 | 85.12 | 89.97 |



protein content (0.2626), fiber content (0.18088), plant height (0.16838), yield per plant (0.10509), number of locules (0.090929), and single fruit weight (0.061251). The most significant negative loadings were seen for fat content (-0.47724), followed by ash content (-0.24532), fruit length (-0.11801), carbohydrate content (-0.10025), number of locules (-0.06889), and fruit wall thickness (-0.02668). This component is therefore referred to as the "moisture factor" (Fig. 9). In the fifth principal component (PC5), the maximum positive loadings were associated with protein content (0.71375), followed by fat content (0.41589), leaf width (0.23993), fiber content (0.23601), fruit length (0.2101), number of locules (0.13092), plant height (0.061587), yield per plant (0.025101), and fruit width (0.005667). The most significant negative loadings were observed for ash content (-0.36996), followed by moisture content (-0.25488), leaf length (-0.22693), carbohydrate content (-0.07577), single fruit weight (-0.07474), and fruit wall thickness (-0.02928). As a result, this component is referred to as the "protein factor" (Fig. 10).

In the sixth principal component (PC6), the highest positive loadings were contributed by leaf length (0.80499), followed by fat content (0.19646), carbohydrate content (0.098745), plant height (0.097476), protein content (0.090773), ash content (0.077181), yield per plant (0.066462), leaf width (0.040522), and fruit wall thickness (0.026129). The most significant negative loadings were associated with moisture content (-0.39848), number of locules (-0.15711), fiber content (-0.10662), fruit width (-0.04995), single fruit weight (-0.00393), and fruit length (-0.00121). Therefore, this component is identified as the "leaf length factor" (Fig. 11).

In the seventh principal component (PC7), the highest positive loadings were

associated fruit length (0.5215), followed by fat content (0.39962), moisture content (0.34763), number of locules (0.21706), leaf length (0.21113), fruit width (0.11254), leaf width (0.069056), fruit wall thickness (0.022281), and carbohydrate content (0.021614). The most significant negative loadings were contributed by fiber content (-0.36222), followed by ash content (-0.23883), yield per plant (-0.16283), plant height (-0.15324), single fruit weight (-0.06534), and fiber content (-0.01598). Therefore, this component is identified as the "fruit length factor" (Fig. 12).

In the eighth principal component (PC8), the highest positive loadings were recorded for fiber content (0.38681), followed by leaf width (0.29819), number of locules (0.18286), fruit width (0.13487), fruit wall thickness (0.12492), carbohydrate content (0.1185), single fruit weight (0.11546), and yield per plant (0.06475). The most significant negative loadings were detected for protein content (-0.42353), followed by moisture content (-0.40758), ash content (-0.27311), fat content (-0.11307), fruit length (-0.15492), leaf length (-0.04241), and plant height (-0.00938). This component is therefore referred to as the "fiber factor" (Fig. 13).

In the ninth principal component (PC9), the highest positive loadings were contributed by leaf width (0.56833), followed by ash content (0.34664), fat content (0.27187), plant height (0.078495), fruit wall thickness (0.047894), and yield per plant (0.023869). The most significant negative loadings were observed for a number of locules (-0.28767), leaf length (-0.20039), protein content (-0.16559), fiber content (-0.16324), carbohydrate content (-0.12578), single fruit weight (-0.0901), ash content (-0.0298), fruit width (-0.00499), and fruit length (-0.00066). This component is therefore identified as the "leaf width factor" (Fig. 14).

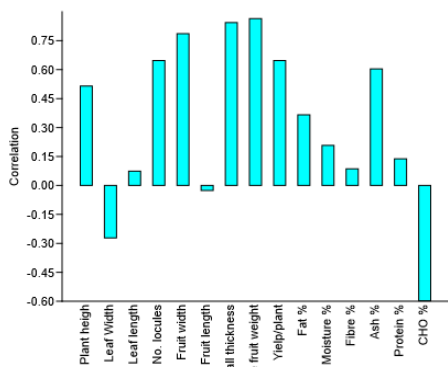


Figure 6. Factor loading for PC1 for physico-chemical traits

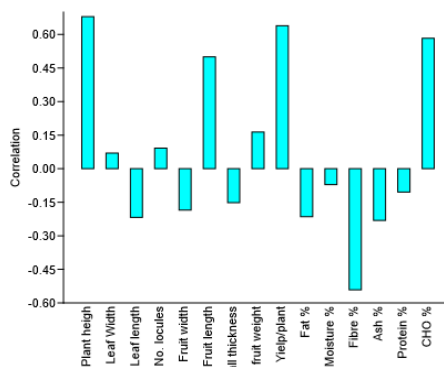


Figure 7. Factor loading for PC2 for physico-chemical traits

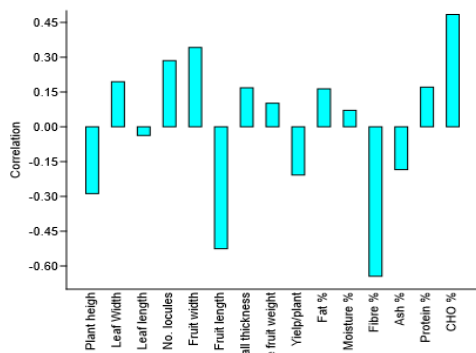


Figure 8. Factor loading for PC3 for physico-chemical traits

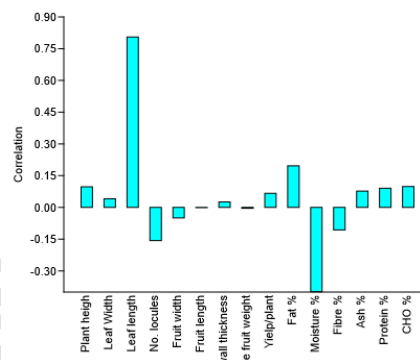


Figure 9. Factor loading for PC4 for physico-chemical traits

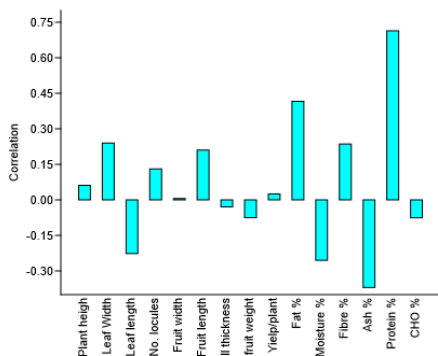


Figure 10. Factor loading for PC5 for physico-chemical traits

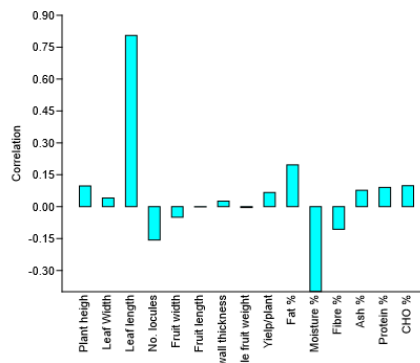
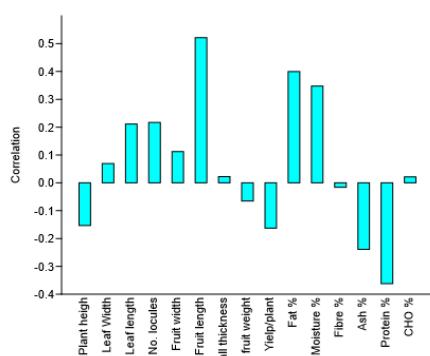
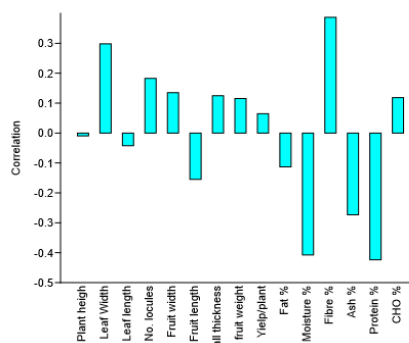


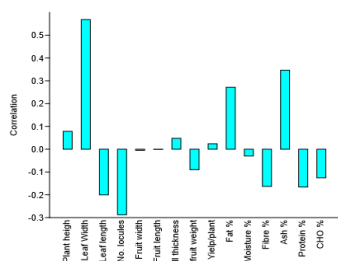
Figure 11. Factor loading for PC6 for physico-chemical traits



**Figure 12. Factor loading for PC7 for physico-chemical traits**



**Figure 13. Factor loading for PC8 for physico-chemical traits**

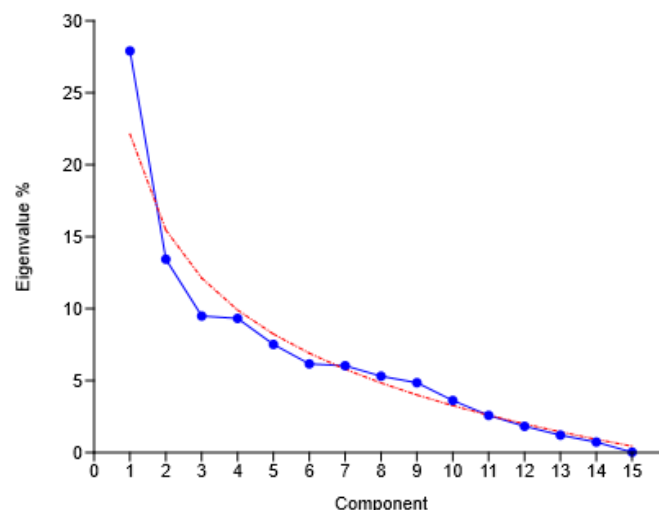


**Figure 14. Factor loading for PC9 for physico-chemical traits**

### Scree plot based on physico-chemical traits

A Scree plot (Fig. 15) illustrates the distribution of variance associated with each principal component (PC) based on their eigenvalues. The plot reveals that PC1 accounted for the highest proportion of variance at 27.09%, with an eigenvalue of 4.1. Subsequent principal components showed a gradual decrease in variance. The

plot features a semi-curved line that levels off after the ninth PC, indicating minimal variance in the remaining components. This analysis suggests that PC1 captures the most significant variation compared to the other principal components. Consequently, focusing on traits associated with PC1 could be advantageous for character selection.



**Figure 15. Scree plot diagram of several factors**

### Conclusion

The study revealed significant variation in agro-morphological and biochemical traits among 95 *Capsicum* genotypes. Genotypes 36606, 36556, 36595, and 36582 were identified as promising candidates for breeding, showing high fruit yield, optimal fruit weight, and length, desirable plant height, and notable biochemical profiles. These results highlight the connection between phenotypic traits and biochemical composition, suggesting that selective breeding can enhance both agronomic performance and nutritional value in future *Capsicum* cultivars. This study contributes to a deeper understanding of the genetic variability within *Capsicum annuum* and its potential applications in breeding programs to develop superior cultivars that meet both agricultural and consumer demands.

### Authors' contributions

Conceived and designed the experiments: N Khan & MJ Ahmed, Performed the experiments: N Khan, Analyzed the data: N Khan, Contributed materials/ analysis/ tools: K Khaqan & MJ Ahmed, Wrote the paper: N Khan & K Khaqan.

### Acknowledgments

The author extends acknowledgment to the Department of Horticulture, University of Poonch Rawalakot, Pakistan, the Laboratory of Food Science and

Technology, University of Poonch Rawalakot, Pakistan.

### References

1. Younes AH & Mustafa YF (2023). Sweet Bell Pepper: A Focus on Its Nutritional Qualities and Illness Alleviated Properties. *Ind J Clin Biochem*.
2. Hill TA, Ashrafi H, Wo SRC, Yao J, Stoffel K, Truco MA, Kozik R, Michelmore RW & Deynze AV (2013). Characterization of (*Capsicum annuum*) genetic diversity and population structure based on parallel polymorphism discovery with a 30K Unigene Pepper Gene Chip. *Plos One* 8(2): 1-16.
3. Sclavo Castillo, D (2024). Culture, Gender, and Flavour in the Conservation of Chile Pepper in Mexico, 1970s-present (Doctoral dissertation).
4. Baruah J, Begum T & Lal M (2024). Estimation of environment stability for fruit yield and capsaicin content by using two models in *Capsicum chinense* Jacq. (Ghost Pepper) with multi-year evaluation. *Peer J* 12: e17511.
5. Khan N, Ahmed MJ, Shah SZA, Shehzad T, Ahmed M, Bashir S & Hamid A (2020). Morphological and agronomic characterization of (*Capsicum annuum* L.) germplasm in Pakistan. *Pure and Appl Biol* 9(2): 1603-1612.
6. Khan N, Ahmed MJ & Shah SZA (2019). Comparative analysis of mineral content and proximate composition from chilli



- pepper (*Capsicum annuum* L.) germplasm. *Pure and Appl Biol* 8(2): 1338-1347.
7. Nunez IM, Verni MC, Argañaraz Martinez FE, Babot JD, Terán V, Danilovich M E & Arena ME (2024). Novel Lactic Acid Bacteria Strains from Regional Peppers with Health-Promoting Potential. *Fermentation*, 10(4): 209.
  8. Nadeem MF, Muhammad A, Khan A, Saed MR & Riaz A (2011). Antioxidant potential of bell pepper (*Capsicum annuum* L.). *Pak J of Food Sci* 21(14): 45-51.
  9. Hygrotech (2000). Production of *Capsicum*. Pyramid, Pretoria, South Africa. 26 pp.
  10. FAO 2009. Food composition table for vegetables and fruits, [www.fao.org/infoods/index\\_en.htm](http://www.fao.org/infoods/index_en.htm).
  11. Dadasaheb MGA (2024). Effect of training techniques on yield and quality of capsicum under protected cultivation (*Capsicum annuum* var. *grossum* sendt.) (doctoral dissertation, mahatma phule krishi vidyapeeth).
  12. Rathore R & Bahadur V (2024). Assessment of Capsicum (*Capsicum annuum*) Varietal Performance in Natural Ventilation Polyhouse Settings. *Inter J of Environ and Climate Change* 14(2): 128-134.
  13. Moulin MM, Santos TO, Ramos HCC and Sudre CP (2022). Transferable polymorphic microsatellite markers from *Capsicum annuum* to *Capsicum baccatum*. *Aus J of Crop Sci* 16: 227-232.
  14. Weerakoon SR & Somaratne S (2010). Agro-morphological characterization and relationships among Mustard germplasm (*Brassica juncea* L.) Czern and Coss in Sri Lanka. A classification tree approach. *J of Agric Sci* 5(2): 89-97.
  15. Costa FRD, Pereira TNS, Victoria AP, Dcampos KP, Rodrigues D, Silva DH & Pereir MG (2006). Genetic diversity among *Capsicum* accessions using RAPAD markers. *Crop Breeding and Appl Biotechnol* 6: 18-23
  16. Zohoungbogbo HP, Vihou F, Achigan-Dako EG & Barchenger DW (2024). Current knowledge and breeding strategies for management of aphid-transmitted viruses of pepper (*Capsicum* spp.) in Africa. *Frontiers in Plant Sci* 15: 1449889.
  17. Evenson RE & Gollin D (2000). Assessing the impact of the Green Revolution, 1960. *Sci* 300: 758–762.
  18. Ceccarelli S & Grando S (2002). Plant breeding with farmers requires treating the assumptions of conventional plant breeding: lesson from the ICCRDA Barely Program, in Farmers, Scientists and Plant Breeding Integrating Knowledge and Practice, Eds., CABI International, New York, NY, USA.
  19. Zhao D, Sapkota M, Lin M, Beil C, Sheehan M, Greene S, & Irish BM (2024). Genetic diversity, population structure, and taxonomic confirmation in annual medic (*Medicago* spp.) collections from Crimea, Ukraine. *Front Plant Sci* 15: 1339298.
  20. Devi J, Sagar V, Kaswan V, Ranjan JK, Kumar R, Mishra GP & Verma RK (2021). Advances in breeding strategies of bell pepper (*Capsicum annuum* L. var. *grossum* Sendt.). *Advances in Plant Breeding Strategies: Vegetable Crops: Volume 9: Fruits and Young Shoots* 3-58.
  21. Emmanuel-Ikpeme C, Henry P & Okiri OA (2014). Comparative evaluation of the nutritional, phytochemical and microbiological quality of three pepper varieties. *J of Food Sci* 2: 74-80.
  22. IPGRI (1995). Descriptors for *Capsicum* (*Capsicum* spp.). Roma: International Plant Genetic Resources Institute (IPGRI), Roma. 51p.
  23. Batool S, Kousar R, Younas Z, Afzal M, Khan MA & Latif S (2024). Morphological Marker-Based Evaluation of Genetic Diversity in Seventeen Foreign *Capsicum annuum* Accessions.
  24. AOAC 1990 Official Methods of Analysis. Association of Official Analytical Chemists, Gaithersburg, Maryland, USA.
  25. Singh RK & Chaudhry BD (1985). Biometric method in quantitative genetic analysis. Kalyani Publishers, New Dehli.
  26. Al-Mamun M, Rafii M, Oladosu Y, Misran AB, Berahim Z, Ahmad Z & Khan, MH (2022). Genetic diversity among kenaf mutants as revealed by qualitative and quantitative traits. *J of Nat Fibers* 19(11): 4170-4187.
  27. Aliyu L, Ahmed KK & Magaji MD (2000). Correlation and multiple regression analysis of morphological characters and components of yield in pepper (*Capsicum annuum* L.). *Crop Res* 9: 318-323.

28. Siso S, Camarero JJ and Gil-Pelegri E (2001). Relationship between hydraulic resistance and leaf morphology in broadleaf *Quercus* species: a new interpretation of leaf lobation. *Trees Str and Fun* 15: 341-345.
29. Lippman Z & Tanksley SD (2001). Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species *pimpinellifolium* L. *esculentum*, var. Giant Heirloom. *Genetics* 158: 413–422.
30. Leite WS, Pavan BE, Alcântara Neto F & Matos Filho CHA (2016). Multivariate Exploratory Approach and Influence of Six Agronomic Traits on Soybean Genotypes Selection. *Nativa* 4(4): 206-210.
31. De Mori G & Cipriani G (2023). Marker-assisted selection in breeding for fruit trait improvement: A review. *Inter J of Mol Sci* 24(10): 8984.
32. Abud HF, Araujo RF, Pinto CMF & Araujo EF (2018). Caracterização morfológica dos frutos de pimentas malagueta e biquinho. *Rev Bras Agropec Sustent* 8(1): 478.
33. Ribeiro CSC, Carvalho SIC, Henz GP & Reifschneider FJB (2008). Brasília: Embrapa Hortaliças, pp: 55-69.
34. Zhao J, Sauvage C, Bitton F & Causse M (2022). Multiple haplotype-based analyses provide genetic and evolutionary insights into tomato fruit weight and composition. *Hort Res* 9: uhab009.
35. Elias HT, Vidigal MCG, Gonela A & Vogt GA (2007). Variabilidade genética em germoplasma tradicional de feijão-preto em Santa Catarina. *Pesquisa Agropecuaria Brasileira* 42: 1443-1449.
36. Sharma J, Sharma P, Chaudhary BP (2017). Estimation of Proximate Composition of Selected Species of Capsicum (*Capsicum annuum* and *Capsicum chinense*) Grown in India. *Inter J of Pure and Appl Biosci* 5 (3): 369-372.
37. Esyas K, Shimelis A, Ashebir F, Negussie R, Tilahun B & Gulelat D (2011). Proximate composition, mineral content and antinutritional factors of some capsicum (*capsicum annum*) varieties grown in ethiopia. *Bull of Chem Soc of Ethiopia* 25(3): 451-454.
38. Bahare S, Alan JH, Maria C, Miquel M, Karina R, Guiomar M, Karl RM, Mehdi S, William NS, Muhammad N, Zubaida Y & Javad S (2018). Potential Phytopharmacy and Food Applications of *Capsicum* spp. A Comprehensive Review. *NPC* 13: 1543-1556.
39. Gloria A, Oyelola B, Adenika T and Anthony J (2010). Comparative analysis of the chemical composition of three spices. *J of Korean Soc of Food Sci and Nut* 37(8): 1090-1095.
40. Manju PR & Sreelathakumary I (2004). Genetic divergence in hot chilli (*Capsicum chinense* Jacq.). *Capsicum and Eggplant Newsl.* pp. 69-72.
41. Santana JGS, Santos PS, Freitas LS & Soares FS (2021). Phenotypic characterisation of the germplasm bank of mangaba (*Hancornia speciosa* Gomes), a unique Brazilian native fruit, with emphasis on its high vitamin C content. *Newzeland J of Crop and Hort Sci* 49(4): 361-373.
42. Pearson K (1920). Notes on the history of correlation. *Biometrika* 13(1): 25-45.
43. Mubarak BS (2002). Evaluation of chilli (*Capsicum annuum* L.) germplasm for productivity, its component traits and resistance to some biotic stresses, M.Sc. (Agri.) Thesis, University of Agricultural Sciences, Dharwad. pp. 29.
44. Rego RE, Rego MM, Cruz CD, Finger LF & Casali VWD (2011). Phenotypic diversity, correlation and importance of variables for fruit quality and yield traits in Brazilian peppers (*Capsicum baccatum*). *Genetic Resources and Crop Evolution* 58: 909-918.
45. Sreelathakumary I & Rajamony L (2002). Variability, heritability, and correlation studies in chilli (*Capsicum* spp.) under shade. *Indian J of Hort Sci* 59: 77-83.
46. Munshi AD, Behra TK & Gyanendra S (2002). Correlation and path coefficient analysis in chilli. *Indian J of Hort* 57(2): 157-159.
47. Karim KMR, Rafii MY, Misran A, Ismail MF, Harun AR, Ridzuan R, Chowdhury MFN, Hosen M, Yusuff O & Haque MA, (2022). Genetic Diversity Analysis among *Capsicum annuum* Mutants. *Agro* 12: 2436.
48. Jolliffie IT (1986). Principal Component Analysis. Springer, New York.