Research Square

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

Integrative Multivariate Analysis of Pisum sativum L. Genotypes Using Agro-Morphological Traits under Contrasting Semi-Arid Highland Agro-Ecosystems

mohammed mareai

mareemohammed88@yahoo.com

Nagi zaid
Tawfiq Al-Omri

Article

Keywords: Cluster analysis, PCA, Path coefficient, Correlation, Heatmap, semi-arid agro-ecosystems

Posted Date: June 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10034537/v1

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

Additional Declarations: No competing interests reported.

Integrative Multivariate Analysis of *Pisum sativum* L. Genotypes Using Agro-Morphological Traits under Contrasting Semi-Arid Highland Agro-Ecosystems

Mohammed Mareai[1, *], Nagi zaid[2] and Tawfiq Al-Omri[3]

1. National Genetic Resource Centre, Agricultural Research and Extension Authority, Dhamar, Yemen

2. Northern Highlands Research Station, Agricultural Research and Extension Authority, Sana'a, Yemen

3. Central Highlands Research Station, Agricultural Research and Extension Authority, Dhamar, Yemen

*Corresponding author: mareemohammed88@yahoo.com

ABSTRACT

Understanding yield formation and environmental adaptation in pea (*Pisum sativum* L.) is essential for improving productivity under semi-arid highland conditions. This study evaluated eleven pea genotypes using morphological characterization and multivariate analyses across contrasting Northern and Central Highland environments of Yemen. Qualitative and quantitative descriptors revealed clear phenotypic variation among genotypes and successfully distinguished improved cultivars from local landraces. Analysis of variance showed significant differences ($P \leq 0.05$) among genotypes for all measured traits, indicating substantial genetic variability. Grain yield ranged from 0.40 to 1.20 t ha^{-1} in the Northern Highlands and from 0.50 to 1.80 t ha^{-1} in the Central Highlands, with Yahsb and Amran consistently exhibiting superior performance across environments. Principal component analysis indicated that yield variation under favorable conditions was mainly associated with seed weight (SW) and number of pods per plant (NPP), whereas phenological traits contributed more strongly under Northern Highland conditions. Correlation analysis revealed strong positive associations of grain yield with seed weight, pod length, pod width, and number of pods per plant, although the magnitude of these relationships varied between environments. Path coefficient analysis showed that seed weight exerted the greatest direct effect on grain yield in the Central Highlands (0.703), while pod length had the strongest direct effect in the Northern Highlands (0.568). Hierarchical clustering and heatmap analyses consistently grouped Yahsb, Amran, YG50075, and YG50083 into a distinct cluster across both environments, reflecting their close phenotypic similarity and superior agronomic performance. These findings demonstrate that yield determination in pea is environment-dependent and highlight Yahsb and Amran as promising genotypes for breeding and cultivation in semi-arid highland agroecosystems.

Keywords: Cluster analysis, PCA, Path coefficient, Correlation, Heatmap, semi-arid agro-ecosystems

Introduction

Pea (*Pisum sativum* L.) is an important grain legume crop worldwide due to its high nutritional value, providing a rich source of protein, essential amino acids, vitamins, and minerals, thereby contributing significantly to food security[1-3]. In addition to its nutritional importance, pea plays a key role in sustainable agriculture through biological nitrogen fixation, which improves soil fertility and reduces dependence on synthetic fertilizers[4, 5]. The productivity and adaptation of pea genotypes are strongly influenced by genetic variability, environmental conditions, and genotype × environment interactions[6]. Therefore, the assessment of genetic diversity among available germplasm is a fundamental step for developing efficient breeding strategies aimed at improving yield and stability[7- 9]. In semi-arid environments, drought stress

represents a major constraint limiting pea productivity. Under such conditions, adaptation is closely associated with phenological traits, including flowering time and maturity, as well as yield-related traits such as pod length, pod width, and number of pods per plant[10]. Phenotypic characterization based on agro-morphological traits remains a practical and widely used approach for evaluating genetic diversity, particularly in regions with limited access to molecular tools[11, 12]. Crop improvement programs rely fundamentally on the availability of genetic diversity and the effective selection of key agronomic traits associated with productivity[13]. The degree of genetic diversity and variance[14], and the degree of heritability of desired traits, determine the success of breeding programs[15]. In this context, multivariate statistical approaches have become essential tools for quantifying genetic diversity, identifying trait contributions, and classifying germplasm resources[16]. Moreover, understanding the magnitude and structure of variation[17] facilitates the identification of superior parental lines for hybridization, particularly under diverse and stress-prone environments. Among these, multivariate statistical techniques, particularly principal component analysis (PCA), have been extensively applied to reduce data dimensionality and identify the most influential traits contributing to variability among genotypes[18,19]. PCA has proven effective in revealing patterns of variation and supporting genotype selection in grain legumes[20]. Similarly, cluster analysis, especially dendrogram-based classification, provides an efficient approach for grouping genotypes based on phenotypic similarity and genetic divergence, thereby facilitating the identification of diverse parental lines for breeding programs[21,22]. Furthermore, correlation and path coefficient analyses provide valuable insights into the relationships among traits and their direct and indirect contributions to grain yield, thereby enabling the identification of effective selection criteria[23-25].

Recent studies have emphasized the importance of integrating phenotypic characterization with multivariate statistical approaches to better understand genetic variability and identify superior genotypes[26]. For instance, PCA and clustering techniques have been successfully used to reveal genetic variation patterns and enhance selection efficiency in pea and other legumes[27; 28]. he identification of diverse parental lines is essential for developing high-yielding offspring and enhancing genetic gain in breeding programs[29,30]. Furthermore, breeding efforts increasingly focus on improving both qualitative and quantitative traits[31-34], traits based on stability, and diversity[35,36].

In Yemen, agriculture plays a vital role in the national economy, with grain legumes serving as a major source of dietary protein. Approximately 54,000 hectares are cultivated annually with legumes, covering nearly 75% of national demand. Despite this importance, pea productivity remains relatively low, particularly in highland regions. This limitation is mainly attributed to the limited availability of improved varieties and the insufficient characterization of local germplasm. Local landraces are typically early-maturing and well adapted to drought-prone conditions, whereas improved cultivars, although higher yielding, often require longer growing periods and greater water availability, which restricts their adoption by farmers[37].

Despite these challenges, limited studies have comprehensively evaluated genetic variability and trait relationships of pea genotypes under contrasting highland environments in Yemen using integrated multivariate approaches. This is the first study integrating PCA, path analysis, correlation analysis**,** dendrogram and Heatmap analysis under Yemeni highland conditions. Therefore, the present study aimed to: (i) assess the magnitude of genetic variability among pea genotypes, (ii) determine the relationships among grain yield and its associated traits, and (iii) classify genotypes using multivariate techniques, including dendrogram and Heatmap analysis, in order to enhance selection efficiency and support breeding programs under water-limited conditions.

Methods

Study area

The experiment was conducted at two locations representing semi-arid highland environments in Yemen. The Northern Highlands Research Station is located in Sana'a Governorate (15.35°N, 44.21°E) at an altitude of 2,253 m above sea level, with an average annual rainfall of approximately 200 mm. The Central Highlands Research Station is located in Dhamar Governorate (14.54°N, 44.40°E) at an altitude of 2,364

m, with an average annual rainfall of approximately 300 mm. The experiments were conducted during the 2021 and 2022 growing seasons.

Plant materials

A total of eleven pea genotypes were evaluated, including two improved varieties (Yahsb and Amran) obtained from the Northern Highlands Research Station and nine local landraces collected from different regions of Yemen through the National Genetic Resource Centre. The improved varieties were originally developed through selection from ICARDA germplasm (**Table 1**).

Identifier	Governorate	Latitude	Longitude	Altitude	Local name
YM50013	Amran	15.61221	43.67863	2800	Atar
YG50015	Ibb	14.04171	44.2799	1850	Atar
YG50054	Dhamar	14.34573	44.45586	2579	Balady
YG50068	Hajjah	15.73359	43.71711	2000	Balady
YG50075	Ibb	13.59112	44.1526	2630	Bazela
YG50077	Al Mahwit	15.30551	43.3604	2100	Balady Tholathay
YG50083	Marib	15.38333	45.36667	1119	Atar
YG50085	Dhamar	14.51258	44.2819	2184	Balady
YG50086	Dhamar	14.41045	44.55943	2467	Balady

Table 1. Genotype passport descriptors for the local accessions

Treatments and design

The experiment was conducted using a randomized complete block design (RCBD) with three replications. Each plot consisted of four rows, each 3 m in length, with a spacing of 0.5 m between rows and 0.1 m between plants within rows. The distance between plots and replications was maintained at 0.5 m and 1 m, respectively. Quantitative and Qualitative traits were described for two growing seasons, 2021 and 2022. Prior to combined analysis, homogeneity of error variances across years was tested. Since the variance structure was homogeneous, data from the two growing seasons were combined and analyzed across environments using mean values over years. Standard agronomic practices recommended for pea production in each region were uniformly applied throughout the growing seasons, including irrigation, weed control, and pest management.

Analysis of Variance (ANOVA) was performed to assess the significance of differences among genotypes for all studied traits under each environment

$$Yij = \mu + Gi + Rj + \varepsilon ij$$

Where: Y_{ij}= observed value of genotype iin replication j, μ= overall mean, G_i= effect of genotype, ε_{ij}= experimental error. Mean comparisons were performed using the Least Significant Difference (LSD) test at $P \leq 0.05$[38].

Coefficient of Variation (CV%) experimental precision was evaluated using: CV (%) = ($\sqrt{MSE} / \bar{X}$) × 100, Where: MSE = mean square error, $\bar{X}$= grand mean

Correlation Analysis Phenotypic correlation coefficients were calculated to determine the relationships among traits using:

$$r_{xy} = \frac{\mathrm{Cov}(x, y)}{\sqrt{\mathrm{Var}(x)\,\mathrm{Var}(y)}}$$

Where: r_{xy}= correlation coefficient between traits xand y[23].

Path Coefficient Analysis was performed based on phenotypic correlation coefficients using standardized partial regression coefficients following Wright's method:

$$r_{iy} = p_{iy} + \sum(r_{ij} \cdot p_{jy})$$

Where: r_{iy}= correlation between trait iand yield, p_{iy}= direct effect, $r_{ij} \cdot p_{jy}$= indirect effects[39,23].

Relative contribution (%) of each trait to grain yield was estimated from the standardized path coefficients by partitioning total correlation into direct and indirect effects according to the procedure described by Dewey and Lu. This approach was used to quantify the proportional importance of each trait under contrasting environments.

Principal Component Analysis (PCA) was conducted based on the standardized correlation matrix to eliminate scale effects among variables**.** PCA is based on eigenvalue decomposition of the covariance (or correlation) matrix:

$PC = a_1 X_1 + a_2 X_2 + ... + a_n X_n$, Where: a_i= eigenvector coefficients. Eigenvalues represent explained variance .Only components with eigenvalues > 1 were retained (Kaiser criterion)[18,19.]

Cluster Analysis (Dendrogram) was performed using Euclidean distance:

$$D_{ij} = \sqrt{\sum(x_{ik} - x_{jk})^2}$$

Genotypes were grouped using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) to generate dendrograms[21,40].

Heatmap and Hierarchical Clustering Analysis

A clustered heatmap analysis was performed separately for each environment using the mean values of nine quantitative traits (DF, FL, DM, PL, PW, NPP, NSP, SW, and GY). Prior to analysis, trait values were standardized using Z-score transformation to remove scale differences among variables. Hierarchical clustering was conducted simultaneously for genotypes and traits using Euclidean distance and Ward's minimum variance method (Ward.D2). The resulting heatmaps were used to visualize trait relationships, genotype grouping patterns, and phenotypic similarities among genotypes under contrasting highland environments[41].

All statistical analyses were conducted to evaluate genetic variability, trait relationships, and genotype classification across environments. Analysis of variance, mean comparison, and dendrogram construction were performed using GenStat software (18th Edition). Heatmap visualization and multivariate analyses were conducted in R software (version 4.5.2) using the FactoMineR and factoextra packages. Correlation and path coefficient analyses were performed using appropriate R packages

Data collection

Quantitative and qualitative data were recorded as shown in **Table (2).** Data were recorded from five randomly selected plants from the two central rows of each experimental plot. A set of local descriptors that were produced with the help of several scientific descriptors from the European Union descriptor (UPOV) and IBPGR[42]. Quantitative traits were initially classified into descriptor categories according to IPGRI and UPOV guidelines for germplasm characterization and dendrogram construction. Both coded quantitative descriptors and qualitative traits were used to assess morphological diversity among

genotypes. However, correlation analysis, principal component analysis, path coefficient analysis, and heatmap visualization were performed using the original quantitative measurements and plot mean values.

Quantitative and qualitative agronomic traits	
Vegetative growth traits	
Leaflet length (cm)	3=Short (2-4), 5= Medium (5-7), 7=Long (more than 7cm)
Leaflet width (cm)	2=Narrow (1.5-2.5), 3=Medium (3-4), 5=Broad (more than 4cm)
Inflorescence and fruit	
Days to 50% flowering (Number of days from sowing to 50% flowering**)**	2=Early (30-39), 3=Medium (40-49), 5=Late (more than 50)
Days to 90% maturity (Number of days from sowing to stage (90%) pods)	2=Early (70-80), 3=Medium (81-90), 5=Late (more than 90)
Diameter of flower (cm)	1=Narrow (1-2), 2=Medium (3-4), 3=Broad (more than4 cm)
The number of seed per plant	1=low (5-8), 2=Medium (9-15), 3=High (more than 15)
Length of bud the flower (mm)	3=short (3-5), 5=Medium (6-8), 7=Long (more than 8mm)
Yield-related traits	
Number of seeds/pods (Measured as a mean of 10 dry pods)	1=Low (2-3), 2=Medium (4-5), 3=High (more than 5)
100 seed weight (gm), (the average weight of two samples of 100 randomly chosen seeds**)**	1=Low (10-15), 2=Medium (16-30), 3=High (more than 30gm)
Seed yield	g/plant, seed yield after drying
Pod length (cm)	3=Short (4), 5=Intermediate (5), 7=Tall (more than 5cm)
Pod width (cm)	3=Tight (0.5-0.8), 5=Medium (1-1.2), 7=Wide (more than 1.2cm)
Qualitative phenotypic traits vegetative growth traits	
Presence of ribs on stem	1=Absent, 9=Present
The presence of stipules	1=Absent, 9=Present
Dots on the leaf	1=Absent, 9=Present
Leaf edges	1=serrated, 2=non-serrated.
Inflorescence and fruit	
Flower ground color	1=White, 2=cream, 3=Purple
The shape of the sepal apex of the flower	1=Peaked, 2=Semi-peaked, 3=Circular
Pod color	1=Yellow, 3=Green, 5=Green bluish, 7=Purple
Seeds traits	
Shape of seed	1=Cylindrical, 2=Diamond, 5=Irregular, 7=flood
Seed color	1=Green light, 2=Green dark, 3=Brown, 5=White
Cotyledon color	1=Green, 2=Yellow, 3=Orange, 5=Brown
The presence of spots on the seed	1=Absent, 9=Present
Colour of the pit of the seed	1=Same color seed, 2=Dark from seed

Table 2. Morphological and phenotypic traits were measured based on the descriptors proposed by the International Plant Genetic Resources Institute (IPGRI) and the International Union for the Protection of New Varieties of Plants (UPOV).

RESULTS

Qualitative traits

The qualitative traits of the evaluated genotypes of *Pisum sativum* revealed clear phenotypic variation that enabled differentiation among genotypes (**Table 3**). The improved cultivars Yahsb and Amran were characterized by a common set of traits, including the presence of stem ribs, stipules, and leaf dots, non-

serrated leaf margins, white flower color, circular sepal apex, diamond-shaped seeds, and green seed color. Yahsb exhibited light-green seeds, whereas Amran showed dark-green seeds. In contrast, most local genotypes (YG lines) displayed purple flowers, serrated leaf margins, irregular seed shapes, and brown seed color. These traits clearly distinguished the local germplasm from the improved cultivars and reflected their different selection histories. Among the local materials, YG50075 and YG50083 exhibited remarkable similarity to Yahsb and Amran. Both genotypes possessed white flowers, diamond-shaped seeds, light- to dark-green seed color, and similar floral characteristics, suggesting a close relationship with the improved cultivars. Their similarity was also consistent with the clustering pattern observed in the dendrogram analysis. Several qualitative traits showed little or no variation among genotypes. Stem ribs were present in most genotypes, pod color remained yellow in all accessions, cotyledon color was uniformly yellow, and seed spots were absent in all materials. The limited variation observed for these traits suggests that they have low discriminatory value within the studied germplasm. In contrast, flower color, seed shape, seed color, and leaf margin characteristics were the most informative descriptors for distinguishing improved cultivars from local landraces.

Quantitative traits in the Central Highlands

Under the more favorable Central Highland conditions, genotypic differences became more pronounced (**Table 4**). The improved cultivars Amran and Yahsb recorded the highest values for most yield-related traits. Pod length reached 6.0 cm in both cultivars compared with only 4.0–4.7 cm in most local genotypes. Similarly, pod width reached 1.5 cm in Yahsb and Amran, whereas it remained around 1.0 cm in most local lines. Seed weight was also substantially greater in the improved cultivars (21.8–22.2 g per 100 seeds) than in local genotypes, which generally ranged between 10.1 and 12.0 g. Among the local genotypes, YG50075 and YG50083 showed superior performance compared with the remaining landraces. These genotypes exhibited longer pods (5.2–6.2 cm), wider pods (1.5 cm), greater numbers of pods per plant (13.6–15.6), and higher seed weight (19.4–23.3 g), resulting in grain yields of 1.3–1.4 t ha^{-1}. In contrast, the remaining local genotypes produced only 0.5–0.7 t ha^{-1}. Flowering and maturity traits also differed among genotypes. YG50075, YG50083, Yahsb, and Amran flowered later (46–47 days) and matured later (85 days) than most local genotypes, which flowered within 33–38 days and matured within 65–72 days. These results indicate that the superior yield performance of Yahsb, Amran, YG50075, and YG50083 was associated with larger reproductive structures, greater pod development, and higher seed weight.

Quantitative traits in the Northern Highlands

Significant variation among genotypes was observed under Northern Highland conditions (**Table 5**). However, several yield-related traits declined compared with the Central Highlands, indicating a stronger environmental limitation on crop performance. The improved cultivars Yahsb and Amran maintained relatively high values for pod length (6.0 cm), pod width (1.4 cm), number of pods per plant (14–16), and seed weight (21–22 g), resulting in grain yields of 1.1 and 1.2 t ha^{-1}, respectively. This confirms their superior adaptation and stability across environments. Among local genotypes, YG50075 and YG50083 continued to outperform the remaining landraces, although their performance declined compared with the Central Highlands. Days to flowering increased markedly to 62 and 65 days, respectively, while maturity was delayed to 98 days. Despite producing relatively long pods (5.8–6.3 cm) and moderate grain yield (0.8–1.0 t ha^{-1}), these genotypes appeared more sensitive to environmental stress. The remaining local genotypes generally exhibited shorter pods (4.3–4.7 cm), narrower pod width (0.6–0.8 cm), lower numbers of pods per plant (6–11), and lower seed weight (11.2–13.5 g), resulting in grain yields ranging from 0.4

to 0.6 t ha^{-1}. These results indicate that environmental stress in the Northern Highlands reduced the expression of yield-related traits, particularly among local landraces.

Genotype performance across environments

Comparison of genotype performance across environments revealed a clear environmental effect on trait expression and grain yield. Most genotypes produced higher yields in the Central Highlands than in the Northern Highlands. The reduction was particularly evident for YG50075 and YG50083, whose grain yield decreased from 1.4 and 1.3 t ha^{-1} in the Central Highlands to 1.0 and 0.8 t ha^{-1} in the Northern Highlands, respectively. The improved cultivars Yahsb and Amran showed the highest stability across environments. Grain yield declined only slightly from 1.6–1.8 t ha^{-1} in the Central Highlands to 1.1–1.2 t ha^{-1} in the Northern Highlands. These cultivars also maintained relatively stable pod length, pod width, and seed weight across both locations. Environmental effects were particularly evident for phenological traits. For example, flowering time in YG50075 increased from 47 to 62 days and in YG50083 from 47 to 65 days when grown in the Northern Highlands. Likewise, maturity was delayed from 85 to 98 days in both genotypes. Such changes indicate a strong genotype × environment interaction affecting developmental and yield-related traits. Overall, the results demonstrate that Yahsb and Amran possess broad adaptation and stable performance across contrasting highland environments, whereas several local genotypes show greater environmental sensitivity despite possessing favorable agronomic characteristics.

Dendrogram analysis

Cluster analysis grouped the genotypes into two main clusters (**Fig. 1**A and 1B**).** The clustering pattern was highly consistent across environments, indicating stable phenotypic relationships among genotypes. Cluster I comprised the improved cultivars and their closely related farmer-maintained derivatives. This cluster was further divided into two sub-clusters. Sub-cluster I-B contained the improved cultivars Yahsb and Amran, which exhibited the highest values for most yield-related traits and shared similar qualitative characteristics. Sub-cluster I-A included YG50075 and YG50083, which were closely associated with Yahsb and Amran, reflecting their common origin and similar agronomic performance. Cluster II contained the remaining local genotypes and was also subdivided into two groups. Sub-cluster II-B consisted of YG50054, YG50013, and YG50015, whereas Sub-cluster II-A included YG50086, YG50077, YG50068, and YG50085. These genotypes generally exhibited lower values for pod size, seed weight, and grain yield compared with Cluster I, indicating substantial phenotypic divergence from the improved cultivars. The clear separation between Clusters I and II demonstrates the effectiveness of the combined qualitative and quantitative descriptors in distinguishing improved germplasm from local landraces. Furthermore, the identical clustering pattern observed across both environments suggests that the major phenotypic relationships among genotypes remained largely stable despite environmental variation.

No.	Trait	Yah-sb	Am-ran	YG5-0013	YG5-0015	YG5-0054	YG5-0068	YG5-0075	YG5-0077	YG5-0083	YG5-0085	YG5-0086
***1**	RS	9	9	9	9	9	9	9	9	1	9	9
***2**	PS	9	9	9	1	9	1	9	1	9	9	1
***3**	DL	9	9	1	1	1	1	9	1	9	1	1
***4**	LE	2	2	1	1	1	1	1	1	1	1	1
***5**	CF	1	1	3	3	3	3	1	3	1	3	3
***6**	SSF	3	3	3	3	3	1	3	1	3	1	1
***7**	PC	1	1	1	1	1	1	1	1	1	1	1
***8**	SS	2	2	7	7	7	7	2	7	2	7	7
***9**	SC	1	2	3	3	3	3	1	3	2	3	3
***10**	CC	2	2	2	2	2	2	2	2	2	2	2
***11**	SOS	1	1	1	1	1	1	1	1	1	1	1
***12**	CPS	1	1	2	2	2	2	1	2	1	2	2

Table 3. Ribs on the stem (RS), the presence of stipules(PS), dots on the leaf (DL), leaf edges (LE), Color of flower (CF), the shape of the sepal apex of the flower (SSAF), pod color (PC), seed shapes(SS), seed color (SC), cotyledon color (CC), spots on the seed (SOS) and color of pit of seed (CPS) for tested pea genotypes. Where: *1,*2,*3: 1 = Absent, 9 = Present, *4: 1= serrated, 2 = non-serrated, *5: 1 = White, 3 = Purple, *6: 1 = Peaked, 3= Circular, *7: 1 = Yellow, *8: 2=Diamond, 7=flood, *9: 1=Green light, 2=Green dark, 3=Brown, *10: 2=Yellow, *11: 1 = Absent, *12: 1 =Same color seed, 2 = Dark from see.

No.	Trait	Yah-sb	Am-ran	YG5-0013	YG5-0015	YG5-0054	YG5-0068	YG5-0075	YG5-0077	YG5-0083	YG5-0085	YG5-0086
***1**	DFL	2	2	1	1	1	1	2	1	2	1	1
***2**	LB	5	5	3	3	3	3	5	3	5	3	3
***3**	LL	5	5	3	3	3	3	5	3	5	3	3
***4**	LW	3	3	2	2	2	2	3	2	3	2	2
***5**	DF	3	3	2	2	2	2	3	2	3	2	2
***6**	NSP	2	2	2	2	2	2	2	2	2	3	2
***7**	PW	7	7	5	5	5	5	7	5	7	5	5
***8**	NPP	2	3	1	2	1	2	3	1	2	1	1
***9**	PL	7	7	3	3	3	5	7	3	5	3	3
***10**	DM	3	3	2	2	2	2	5	2	5	2	2
***11**	SW	2	2	1	1	1	1	2	1	2	1	1

Table 4. Diameter of flower (DFL), Length of bud the flower (LB), leaflet length (LL), leaflet width (LW), days to flowering (DF), pod width (PW), the number of seed per pod (NSP), the number of seed per plant (NPP), pod length (PL), days to maturity (DM) and seed weight (SW**)** for tested pea genotypes at the central highlands**.** Where: *1: 1=Narrow, 2=Medium, *2: 3=short, 5=Medium, *3: 3=Short, 5=Medium, *4: 2=Narrow, 3=Medium, *5: 2=Early, 3=Medium, 5=Late, *6: 2=Medium, 3=Hight, *7: 3=Narrow, 5=Medium, 7=Broad, *8: 1=law, 2=Medium, 3=Hight, *9: 3=Short, 5=Medium, 7=Tall, *10: 2=Early, 3=Medium, 5=Late, *11: 1=Law, 2=Medium

No.	Trait	Yah-sb	Am-ran	YG5-0013	YG5-0015	YG5-0054	YG5-0068	YG5-0075	YG5-0077	YG5-0083	YG5-0085	YG5-0086
***1**	DFL	2	2	1	1	1	1	2	1	2	1	1
***2**	LB	5	5	3	3	3	3	5	3	5	3	3
***3**	LL	5	5	3	3	3	3	5	3	5	3	3
***4**	LW	3	3	2	2	2	2	3	2	3	2	2
***5**	DF	3	3	2	2	2	2	5	2	5	2	2
***6**	NSP	2	2	2	2	2	2	3	2	3	2	2
***7**	PW	7	7	3	3	3	3	5	3	5	3	3
***8**	NPP	2	3	1	2	2	1	2	1	2	1	1
***9**	PL	7	7	3	5	5	5	7	5	7	5	5
***10**	DM	3	3	2	2	2	2	5	2	5	2	2
***11**	SW	2	2	1	1	1	1	1	1	1	1	1

Table 5. Diameter of flower (DFL), Length of bud the flower (LB), leaflet length (LL), leaflet width (LW), days to flowering (DF), pod width (PW), the number of seed per pod (NSP), the number of seed per plant (NPP), pod length (PL), days to maturity (DM) and seed weight (SW**)** for tested pea genotypes at the northern highlands**.** Where: *1: 1=Narrow, 2=Medium, *2: 3=short, 5=Medium, *3: 3=Short, 5=Medium, *4: 2=Narrow, 3=Medium, *5: 2=Early, 3=Medium, 5=Late, *6: 2=Medium, 3=Hight, *7: 3=Narrow, 5=Medium, 7=Broad, *8: 1=law, 2=Medium, 3=Hight, *9: 3=Short , 5=Medium, 7=Tall, *10: 2=Early, 3=Medium, 5=Late, *11: 1=Law, 2=Medium.

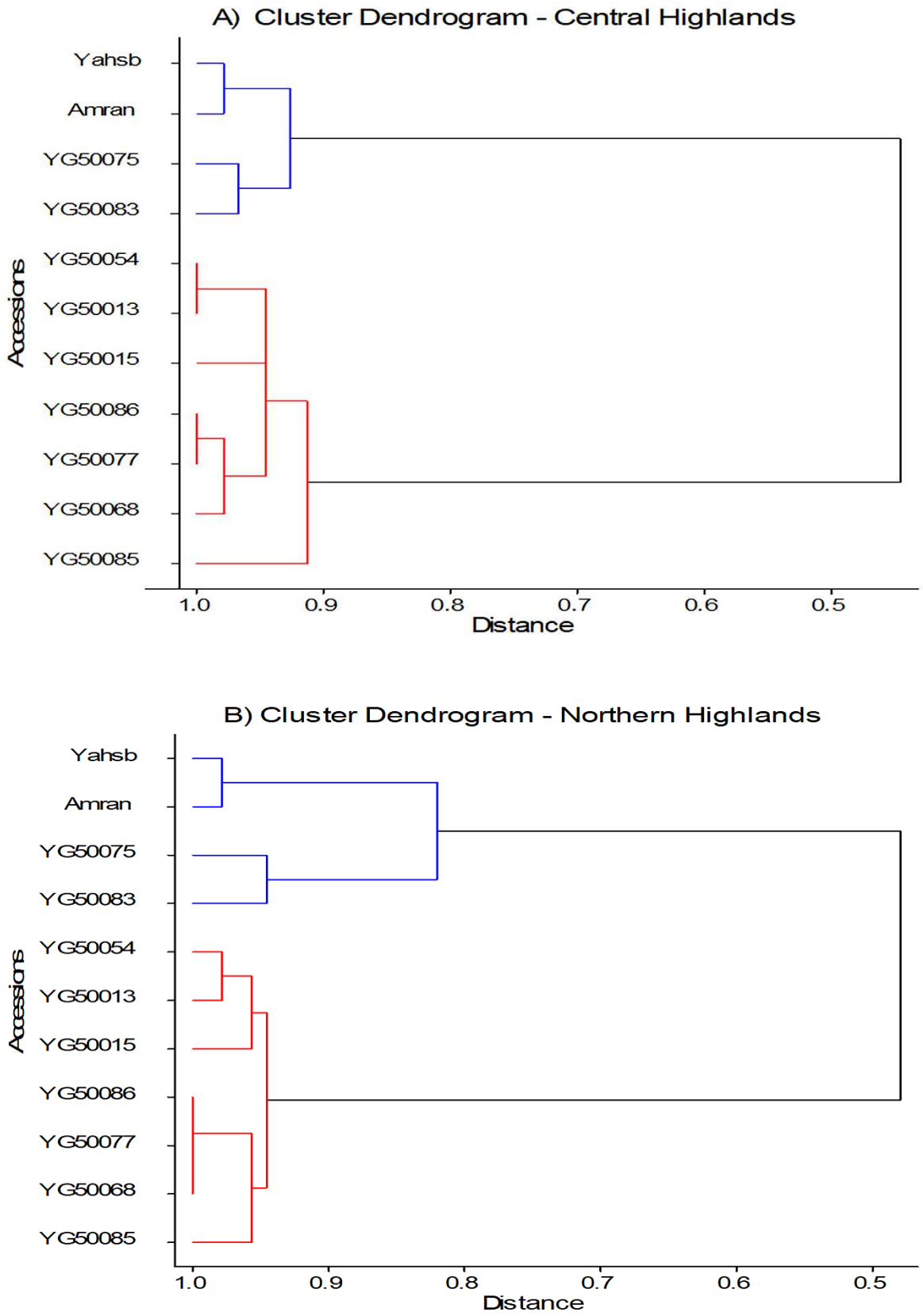


Figure 1. Dendrogram showing hierarchical clustering of *Pisum sativum* genotypes based on Euclidean distance and UPGMA method under (a) Central Highlands. (b) Northern Highlands and Genotypes are grouped into two major clusters: Cluster I includes improved genotypes (Yahsb and Amran) along with YG50075 and YG50083, characterized by superior agronomic performance, while Cluster II contains local

genotypes with lower yield potential. Differences in clustering patterns between environments indicate the influence of genotype × environment interaction.

The Mean performance

Central Highlands Research Station

Analysis of variance revealed significant differences among genotypes for all studied traits (**Table 6**). Indicating substantial variability within the evaluated germplasm. Days to flowering (DF) ranged from 33 to 47 days, while days to maturity (DM) varied from 65 to 90 days. The earliest flowering and maturing genotypes were YG50054 and YG50085, whereas Yahsb, Amran, YG50075, and YG50083 exhibited the longest growth duration. Significant variation was also observed for flower length (FL), pod length (PL), and pod width (PW). Flower length ranged from 10 to 20 mm, with YG50075 recording the highest value, followed by YG50083 and Yahsb. Pod length varied from 4.0 to 6.2 cm, with YG50075 producing the longest pods, while Yahsb and Amran also recorded high values. Pod width ranged from 1.0 to 1.5 cm, with the highest values observed in YG50075 and YG50083. Yield component traits showed considerable variation among genotypes. The number of pods per plant (NPP) ranged from 6.8 to 15.6, with Amran and YG50075 producing the highest values, followed by Yahsb and YG50083. The number of seeds per pod (NSP) varied from 3.5 to 5.6, with YG50085 recording the highest value. Significant differences were also observed for 100-seed weight (SW), which ranged from 10.23 to 23.30 g. YG50083 produced the heaviest seeds, followed by Amran and Yahsb. Grain yield (GY) ranged from 0.50 to 1.80 t ha^{-1}. Amran recorded the highest grain yield, followed by Yahsb, YG50075, and YG50083. These four genotypes consistently ranked among the highest-performing entries for several agronomic and yield-related traits. The coefficient of variation ranged from 2.9% to 9.5%, indicating good experimental precision.

Northern Highlands Research Station

Analysis of variance revealed significant differences among genotypes for all studied traits under Northern Highlands conditions (**Table 7).** Days to flowering (DF) ranged from 30 to 65 days, whereas days to maturity (DM) varied from 70 to 98 days. YG50086 was the earliest genotype for both flowering and maturity, while YG50075 and YG50083 exhibited the latest flowering and maturity dates. Flower length (FL) ranged from 10 to 20 mm, with YG50075 and YG50083 recording the highest values. Pod length (PL) varied from 4.13 to 6.30 cm, and the longest pods were observed in YG50075, followed by Yahsb and Amran. Pod width (PW) ranged from 0.60 to 1.40 cm, with Yahsb and Amran exhibiting the highest values. Substantial variation was observed for yield components. The number of pods per plant (NPP) ranged from 5.7 to 16.0, with Amran producing the highest value, followed by Yahsb. The number of seeds per pod (NSP) ranged from 4.0 to 6.17, with YG50083 and YG50075 recording the highest values. Seed weight (SW) varied from 11.3 to 22.0 g, with Amran and Yahsb producing significantly heavier seeds than the remaining genotypes. Grain yield (GY) ranged from 0.40 to 1.20 t ha^{-1}. Amran recorded the highest grain yield, followed by Yahsb, YG50075, and YG50083. Similar to the Central Highlands environment, Amran, Yahsb, YG50075, and YG50083 ranked among the highest-performing genotypes for several agronomic and yield-related traits. The coefficient of variation ranged from 3.74% to 15.8%, indicating acceptable experimental precision.

Genotype	DF	FL	DM	PL	PW	NPP	NSP	SW	GY
50013	37 ^{a}	10 ^{c}	72 ^{c}	4.0 ^{d}	1.0 ^{c}	7.4 ^{d}	4.6 ^{cd}	10.23 ^{h}	0.50 ^{e}
50015	37 ^{a}	13 ^{b}	70 ^{d}	4.0 ^{d}	1.0 ^{c}	10.8 ^{c}	5.2 ^{b}	11.2 ^{fg}	0.70 ^{d}
50054	33 ^{c}	11 ^{c}	65 ^{f}	4.0 ^{d}	1.0 ^{c}	7.0 ^{d}	5.4 ^{ab}	10.50 ^{g}	0.50 e
50068	37 ^{a}	12 ^{bc}	70 ^{d}	4.7 ^{c}	1.0 ^{c}	8.6 ^{d}	4.4 ^{d}	11.50 ^{f}	0.60 ^{de}

50075	47 [b]	20 [a]	85 [a]	6.2 [a]	1.5 [b]	15.6 [a]	5.0 [bc]	19.40 [d]	1.40 [c]
50077	38 [a]	10 [c]	72 [c]	4.0 [d]	1.0 [c]	7.0 [d]	4.5 [cd]	12.00 [e]	0.60 [de]
50083	47 [b]	19 [a]	85 [a]	5.2 [b]	1.5 [b]	13.6 [b]	4.2 [de]	23.30 [a]	1.30 [c]
50085	33 [c]	11 [c]	65 [f]	4.3 [d]	1.0 [c]	9.0 [c]	5.6 a	10.60 [g]	0.60 [de]
50086	37 [a]	10 [c]	70 [d]	4.3 [d]	1.0 [c]	6.8 [d]	3.5 [f]	11.2 [fg]	0.50 [e]
Yahsb	47 [b]	15 [b]	90 [a]	6.0 [a]	1.4 [b]	14.0 [ab]	4.2 [de]	21.80 [c]	1.60 [b]
Amran	46 [b]	10 [c]	90 [a]	6.0 [a]	1.4 [b]	15.6 [a]	4.8 [c]	22.20 [b]	1.80 [a]
LSD (0.05)	2.1	1.8	4.5	0.6	0.2	1.9	0.28	0.44	0.14
LSD (0.01)	2.9	2.5	6.1	0.8	0.3	2.6	0.38	0.60	0.19
CV %	4.8	5.2	3.9	6.5	5.1	7.4	4.6	2.9	9.5

Table 6. Mean performance and mean separation of 11 genotypes in the central Highlands. DF = Days to flowering; FL = Flower length; DM = Days to maturity; PL = Pod length; PW = Pod width; NPP = Number of pods per plant; NSP = Number of seeds per pod; SW = 100-seed weight; GY = Grain yield. Means followed by the same letter within each column are not significantly different according to LSD test at $P \leq 0.05$ and $P \leq 0.01$.

The comparison between the two environments showed clear differences in genotype performance (**Table 8**). Overall, grain yield was higher in the Central Highlands than in the Northern Highlands, indicating more favorable growing conditions. Genotypes Amran and Yahsb consistently exhibited high yield across both environments, demonstrating good stability. In contrast, genotypes such as YG50075 and YG50083 showed reduced performance in the Northern Highlands, indicating sensitivity to environmental conditions. Some genotypes (e.g., YG50013, YG50054, and YG50086) maintained similar performance across environments, although at relatively low yield levels, suggesting stability but limited productivity.

Genotype	DF	FL	DM	PL	PW	NPP	NSP	SW	GY
50013	38 [c]	12 [c]	77 [c]	4.13 [c]	0.80 [c]	7 [cd]	5.40 [ab]	12 [c]	0.50 [de]
50015	38 [c]	12 [c]	78 [c]	4.70 [c]	0.60 [d]	11 [bc]	5.20 [ab]	11.4 [c]	0.60 [d]
50054	38 [c]	13 [c]	75 [c]	4.50 [c]	0.60 [d]	11 [bc]	4.53 [bc]	13.5 [b]	0.50 [de]
50068	38 [c]	11 [c]	75 [c]	4.57 [c]	0.70 [cd]	7 [cd]	4.00 [c]	11.5 [c]	0.40 [e]
50075	62 [a]	20 [a]	98 [a]	6.30 [a]	1.00 [b]	12 [b]	6.00 [a]	14 [b]	1.00 [b]
50077	38 [c]	11 [c]	75 [c]	4.50 [c]	0.80 [c]	8 [cd]	4.47 [bc]	12 [c]	0.50 [de]
50083	65 [a]	20 [a]	98 [a]	5.80 [ab]	1.00 [b]	11 [bc]	6.17 [a]	13 [b]	0.80 [c]
50085	38 [c]	12 [c]	77 [c]	4.50 [c]	0.60 [d]	5.7 [d]	4.57 [bc]	11.4 [c]	0.50 [de]
50086	30 [d]	10 [c]	70 [d]	4.50 [c]	0.70 [cd]	7 [cd]	4.67 [bc]	11.3 [c]	0.50 [de]
Yahsb	46 [b]	15 [b]	90 [b]	6.00 [a]	1.40 [a]	14 [ab]	4.33 [bc]	21 [a]	1.10 [ab]
Amran	47 [b]	10 [c]	90 [b]	6.00 [a]	1.40 [a]	16 [a]	4.67 [bc]	22 [a]	1.20 [a]
LSD (0.05)	4.11	2.46	5.20	0.77	0.23	2.98	1.22	1.12	0.10
CV (%)	5.57	8.53	3.74	10.1	15.2	14.3	15.8	6.9	10.6

Table 7. Mean performance and mean separation of 11 genotypes in the Northern Highlands. DF = Days to flowering; FL = Flower length; DM = Days to maturity; PL = Pod length; PW = Pod width; NPP = Number of pods per plant; NSP = Number of seeds per pod; SW = 100-seed weight; GY = Grain yield. Means followed by the same letter within each column are not significantly different according to the LSD test at $P \leq 0.05$.

Genotype	Central (t/ha)	Northern(t/ha)	Response
50013	0.50 e	0.50 de	Stable high but low yielding
50015	0.70 d	0.60 d	Stable but low yielding
50054	0.50 e	0.50 de	Stable high but low yielding
50068	0.60 de	0.40 e	Environment sensitive
50075	1.40 c	1.00 b	Environment sensitive
50077	0.60 de	0.50 de	Stable but low yielding

50083	1.30 c	0.80 c	Environment sensitive
50085	0.60 de	0.50 de	Stable but low yielding
50086	0.50 e	0.50 de	Stable high but low yielding
Yahsb	1.60 b	1.10 ab	Environment sensitive
Amran	1.80 a	1.20 a	Environment sensitive

Table 8. Direct comparisons between the two environments**.**

Principal component analysis (PCA)

Central Highlands Research Station

Principal component analysis (PCA) revealed a clear difference between the two studied environments in explaining the variance between traits. Principal component analysis revealed substantial phenotypic variability among the evaluated pea genotypes under the Central Highlands environment (**Fig. 2A.**). The first two principal components (PC1 and PC2) explained 92.1% of the total variation, with PC1 accounting for 80.5% and PC2 accounting for 11.6%. The PCA biplot demonstrated that seed weight (SW), grain yield (GY), number of pods per plant (NPP), and pod length (PL) contributed most strongly to genotype discrimination along PC1. These traits were positively associated and clustered in the same vector direction, indicating their strong contribution to yield improvement under this environment. Days to flowering (DF), days to maturity (DM), and pod width (PW) also showed positive associations with grain yield, although their contribution was lower than SW and NPP. In contrast, number of seeds per pod (NSP) was negatively associated with most yield-related traits, as indicated by its opposite vector orientation. Genotypes Amran, Yahsb, YG50075, and YG50083 were positioned near the positive side of PC1 and closely associated with high SW, GY, and NPP, suggesting superior adaptation and productivity under Central Highlands conditions. Conversely, YG50054, YG50085, and YG50086 were located away from yield-associated vectors, indicating comparatively lower agronomic performance. The PCA results indicate that SW, NPP, and PL are the major contributors to grain yield variation in the Central Highlands environment and may therefore serve as effective selection criteria in pea breeding programs targeting highland agroecologist.

Northern Highlands Research Station

In contrast, Under Northern Highlands conditions, the first two principal components explained 91.6% of the total phenotypic variation, with PC1 and PC2 accounting for 68.2% and 23.4%, respectively. Unlike the Central Highlands environment, phenological traits including DF, DM, and FL contributed strongly to genotype discrimination. These traits were closely associated with PC1, indicating that developmental timing played a major role in genotype adaptation under Northern Highlands conditions. Grain yield showed strong positive associations with pod-related traits such as PL, PW, NPP, and SW. Genotypes Amran and Yahsb were positioned near the positive region of the yield vectors, indicating superior agronomic performance and wider adaptability. YG50075 exhibited a distinct position in the biplot, suggesting differential environmental responsiveness relative to other genotypes (**Fig. 2B**.). The PCA results indicate that phenological traits were more important determinants of genotype performance in the Northern Highlands environment, whereas pod and seed traits remained critical for yield improvement.

This suggests a greater influence of environmental factors on growth traits such as flowering days (DF) and ripening days (DM). A change in vector orientations was also observed between the two locations, reflecting a significant genotype-environment interaction (G×E). Some varieties, such as Amran and Yahsb, showed relative stability across the two environments, maintaining positions close to the yield vector in both analyses, indicating their suitability for cultivation in different environments. In contrast,

other varieties, such as YG50075 and YG50083, exhibited a clear environmental response, shifting their position between the two environments, indicating their sensitivity to environmental conditions. These results highlight the importance of assessing varieties in multiple environments to identify both stable and

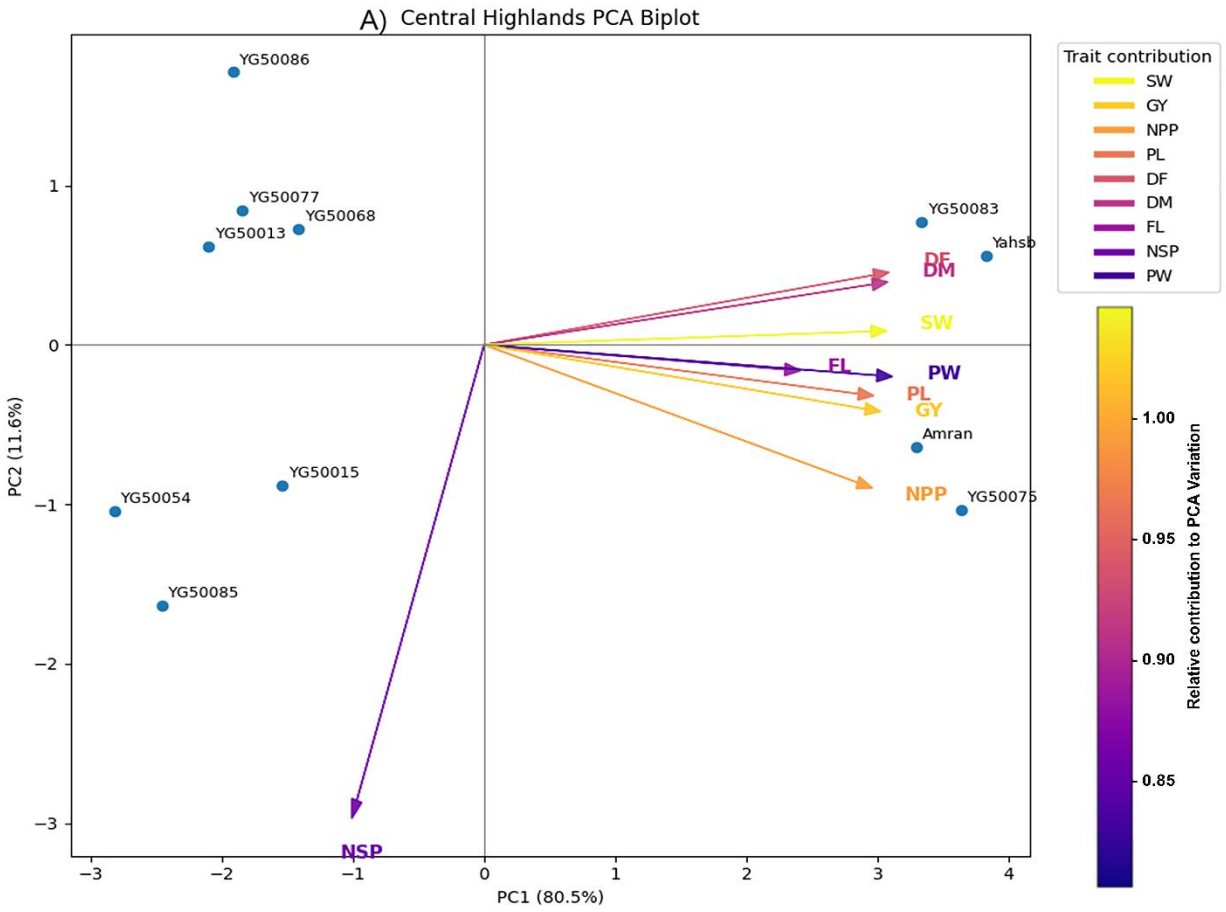
The Correlation analysis

Central Highlands Research Station

Pearson's correlation analysis revealed strong positive relationships between grain yield and most agronomic traits. Grain yield exhibited highly positive correlations with: SW (r = 0.95), NPP (r = 0.95), PW (r = 0.96) and PL (r = 0.94). Similarly, DF and DM showed an almost perfect positive association (r = 0.98), reflecting the close physiological relationship between flowering and maturity duration (**Fig. 3A).** Pod weight (PW) also exhibited strong positive correlations with SW (r = 0.98) and GY (r = 0.96), suggesting that increased pod development contributed substantially to final grain yield. In contrast, NSP displayed weak to negative correlations with most yield-related traits, including SW (r = −0.34) and GY (r = −0.19), indicating limited contribution of this trait to productivity under Central Highlands conditions. The correlation analysis confirms that improving seed weight, pod development, and pod number would likely enhance grain yield under Central Highlands conditions. The weak association between NSP and GY suggests that selection based solely on seed number per pod may not effectively improve productivity.

Northern Highlands Research Station

The correlation analysis in the Northern Highlands revealed strong and highly positive associations among most agronomic traits. Grain yield exhibited strong positive correlations with: PL (r = 0.93), PW (r = 0.93), NPP (r = 0.92) and SW (r = 0.87). Days to flowering and days to maturity were also strongly correlated (r = 0.96), indicating synchronized developmental progression among genotypes. **(Fig. 3B**). Pod width showed high positive associations with SW (r = 0.92) and GY (r = 0.93), confirming the importance of pod development for yield enhancement. In contrast, NSP showed relatively weak associations with grain yield (r = 0.25), indicating limited contribution to yield determination under Northern Highlands conditions. The correlation results suggest that pod architecture and seed development traits are the primary determinants of grain yield under Northern Highlands conditions.ecologically specialized varieties.


A) Central Highlands PCA Biplot
YG50086
YG50077
YG50068
YG50013
YG50083
Yahsb
DF
DM
SW
FL
PW
PL
GY
Amran
YG50015
YG50054
NPP
YG50075
YG50085
NSP
PC2 (11.6%)
PC1 (80.5%)
Trait contribution
SW
GY
NPP
PL
DF
DM
FL
NSP
PW
Relative contribution to PCA Variation
1.00
0.95
0.90
0.85

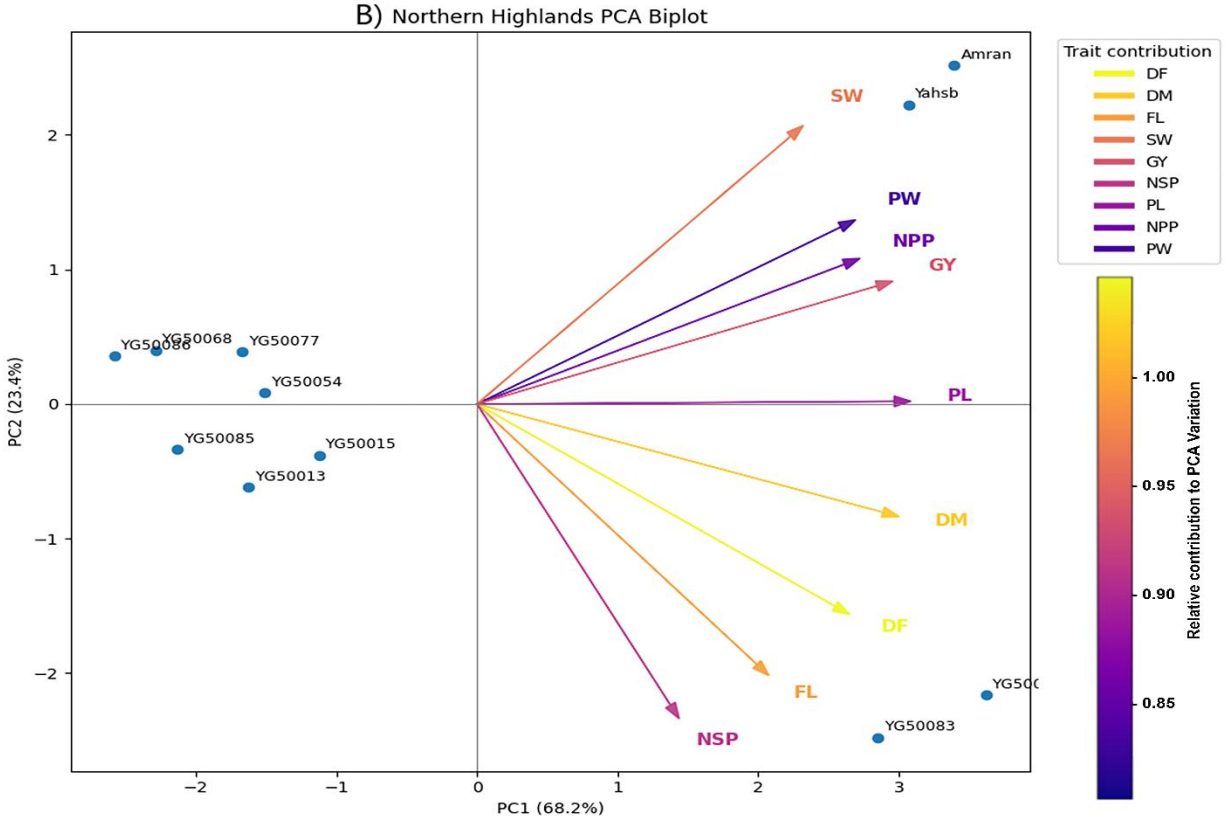

B) Northern Highlands PCA Biplot
Amran
SW
Yahsb
PW
NPP
GY
YG50086
YG50068
YG50077
YG50054
PL
YG50085
YG50015
YG50013
DM
DF
FL
YG50083
NSP
PC2 (23.4%)
PC1 (68.2%)
Trait contribution
DF
DM
FL
SW
GY
NSP
PL
NPP
PW
Relative contribution to PCA Variation
1.00
0.95
0.90
0.85

Figure 2. Principal component analysis (PCA) biplots illustrating the relationships among pea genotypes and agronomic traits under (a) Central Highlands and (b) Northern Highlands environments. The first two principal components accounted for most of the observed variation and revealed differences in trait contribution and genotype adaptation patterns between environments. Genotype positions relative to trait vectors indicate their phenotypic associations and yield-related performance. DF = Days to flowering; FL = Flower length; DM = Days to maturity; PL = Pod length; PW = Pod width; NPP = Number of pods per plant; NSP = Number of seeds per pod; SW = 100-seed weight; GY = Grain yield.

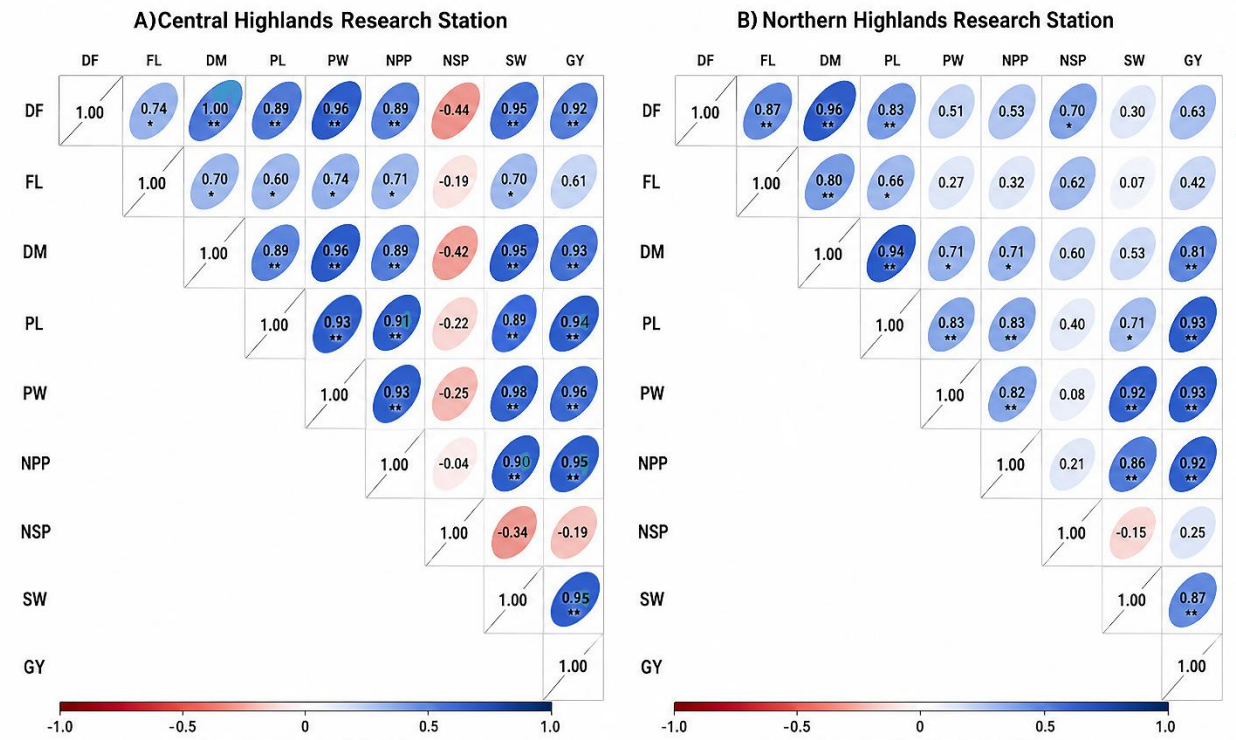


Figure 3. Pearson correlation heatmaps showing the magnitude and direction of associations among agronomic traits of pea genotypes under (a) Central Highlands and (b) Northern Highlands environments. Positive and negative correlations are represented by contrasting color intensities, highlighting key trait relationships associated with grain yield formation in each environment, where * Significant at $P \leq 0.05$ and ** Significant at $P \leq 0.01$. DF = Days to flowering; FL = Flower length; DM = Days to maturity; PL = Pod length; PW = Pod width; NPP = Number of pods per plant; NSP = Number of seeds per pod; SW = 100-seed weight; GY = Grain yield.

The path coefficient analysis

Central Highlands Research Station

Path coefficient analysis revealed substantial variation in the contribution of agronomic traits to grain yield (GY) between the Central Highlands and Northern Highlands Path coefficient analysis partitioned the correlation coefficients into direct and indirect effects on grain yield. **(Table 9).** The path coefficient analysis revealed that seed weight (SW) exerted the highest positive direct effect on grain yield (0.703), indicating that seed filling efficiency was the principal determinant of productivity under Central Highlands conditions. Pod length (PL) and days to maturity (DM) also showed substantial positive direct effects of 0.353 and 0.365, respectively, suggesting that prolonged reproductive duration and improved pod development contributed significantly to grain yield improvement. Pod width (PW) exhibited a negative direct effect (-0.346) despite maintaining the highest positive correlation with grain yield (r = 0.963). This indicates that the positive association of PW with grain yield resulted mainly from indirect effects through seed weight and maturity-related traits rather than from its own direct contribution.

Number of pods per plant (NPP) showed a moderate positive direct effect (0.248) and strong indirect contributions through SW and PL, confirming the importance of reproductive efficiency in determining yield performance. Conversely, days to flowering (DF) and flowering length (FL) exhibited negative direct effects on grain yield, suggesting that delayed flowering and prolonged flowering duration reduced productivity directly. However, both traits maintained positive total correlations with grain yield because of their indirect associations with maturity and seed-filling traits. The number of seeds per pod (NSP) displayed weak direct and indirect effects and showed a negative overall association with grain yield (-0.188), indicating limited importance in yield determination under the Central Highlands environment. Overall, the analysis demonstrated that grain yield under Central Highlands conditions was primarily influenced by seed filling, maturity duration, and pod development traits.

Trait	DF	FL	DM	PL	PW	NPP	NSP	SW	r(GY)
DF	**-0.146**	-0.127	0.364	0.316	-0.332	0.220	-0.047	0.671	0.919
FL	-0.108	**-0.172**	0.257	0.240	-0.258	0.175	-0.020	0.494	0.608
DM	-0.146	-0.121	**0.365**	0.315	-0.333	0.220	-0.045	0.671	0.929
PL	-0.131	-0.117	0.326	**0.353**	-0.324	0.227	-0.023	0.624	0.936
PW	-0.140	-0.128	0.351	0.330	**-0.346**	0.232	-0.027	0.691	0.963
NPP	-0.130	-0.121	0.325	0.323	-0.324	**0.248**	-0.004	0.634	0.950
NSP	0.065	0.033	-0.155	-0.078	0.088	-0.010	**0.105**	-0.236	-0.188
SW	-0.139	-0.121	0.349	0.313	-0.340	0.224	-0.035	**0.703**	0.954

Table 9. Direct effects (bold diagonal values) and indirect effects of agronomic traits on grain yield under Central Highlands conditions. DF = Days to flowering; FL = Flower length; DM = Days to maturity; PL = Pod length; PW = Pod width; NPP = Number of pods per plant; NSP = Number of seeds per pod; SW = 100-seed weight; GY = Grain yield.

Northern Highlands Research Station

Under Northern Highlands conditions, pod length (PL) exerted the highest positive direct effect on grain yield (0.568), indicating that pod development was the primary determinant of productivity under the relatively stressful northern environment. Days to maturity (DM) also exhibited a strong positive direct effect (0.409), suggesting that genotypes with prolonged reproductive duration achieved greater assimilate accumulation and improved adaptation. **(Table 10).** Pod width (PW), seed weight (SW), and number of pods per plant (NPP) showed relatively smaller direct effects but substantial indirect contributions through pod length and maturity duration, demonstrating the integrated role of developmental and reproductive traits in determining grain yield. Days to flowering (DF) exhibited a strong negative direct effect (-0.506), indicating that delayed flowering reduced grain yield directly. Nevertheless, DF maintained a positive total correlation with grain yield due to strong indirect effects through PL and DM. Flowering length (FL) showed only a negligible negative direct effect and relatively weak indirect contributions, suggesting a limited role in yield determination under Northern Highlands conditions. Compared with the Central Highlands environment, the Northern Highlands conditions showed greater dependence on developmental adaptation and pod architecture traits rather than seed-filling traits alone. Therefore, pod length and maturity duration can be considered the most reliable selection criteria for improving pea productivity under Northern Highlands conditions.

Trait	DF	FL	DM	PL	PW	NPP	NSP	SW	r(GY)
DF	**-0.506**	-0.029	0.392	0.470	0.101	0.065	0.095	0.041	0.628
FL	-0.442	**-0.034**	0.328	0.377	0.053	0.039	0.084	0.010	0.416
DM	-0.485	-0.027	**0.409**	0.532	0.139	0.087	0.081	0.074	0.810
PL	-0.419	-0.022	0.384	**0.568**	0.163	0.102	0.053	0.099	0.928
PW	-0.260	-0.009	0.290	0.473	**0.196**	0.100	0.011	0.127	0.928
NPP	-0.268	-0.011	0.291	0.473	0.160	**0.123**	0.029	0.120	0.918
NSP	-0.356	-0.021	0.247	0.225	0.016	0.026	**0.134**	-0.020	0.250

SW	-0.150	-0.002	0.217	0.405	0.179	0.106	-0.020	**0.139**	0.874

Table 10. Direct effects (bold diagonal values) and indirect effects of agronomic traits on grain yield under Northern Highlands conditions. DF = Days to flowering; FL = Flower length; DM = Days to maturity; PL = Pod length; PW = Pod width; NPP = Number of pods per plant; NSP = Number of seeds per pod; SW = 100-seed weight; GY = Grain yield.

Path coefficient analysis revealed clear differences in the relative importance of yield-related traits between the two environments. Under Central Highlands conditions, grain yield was primarily influenced by seed weight (SW), which exhibited the highest positive direct effect, followed by days to maturity (DM) and pod length (PL). These results indicate that seed filling efficiency and reproductive development were the major determinants of productivity under favorable environmental conditions. In contrast, under Northern Highlands conditions, pod length (PL) showed the strongest positive direct effect on grain yield, followed by days to maturity (DM), indicating that pod development and phenological adaptation played a greater role under the relatively stressful environment. Although pod width (PW), number of pods per plant (NPP), and seed weight (SW) exhibited lower direct effects, they contributed substantially through indirect pathways. Overall, the analysis demonstrated that the hierarchy of yield-determining traits changed across environments. Grain yield under favorable conditions was mainly associated with seed filling and reproductive efficiency, whereas under Northern Highlands conditions it depended more strongly on pod development and developmental adaptation. Nevertheless, pod length, days to maturity, and seed weight consistently showed important contributions across both environments and may therefore serve as reliable selection criteria for improving pea productivity under highland conditions.

Heatmap analysis

Central Highlands Research Station

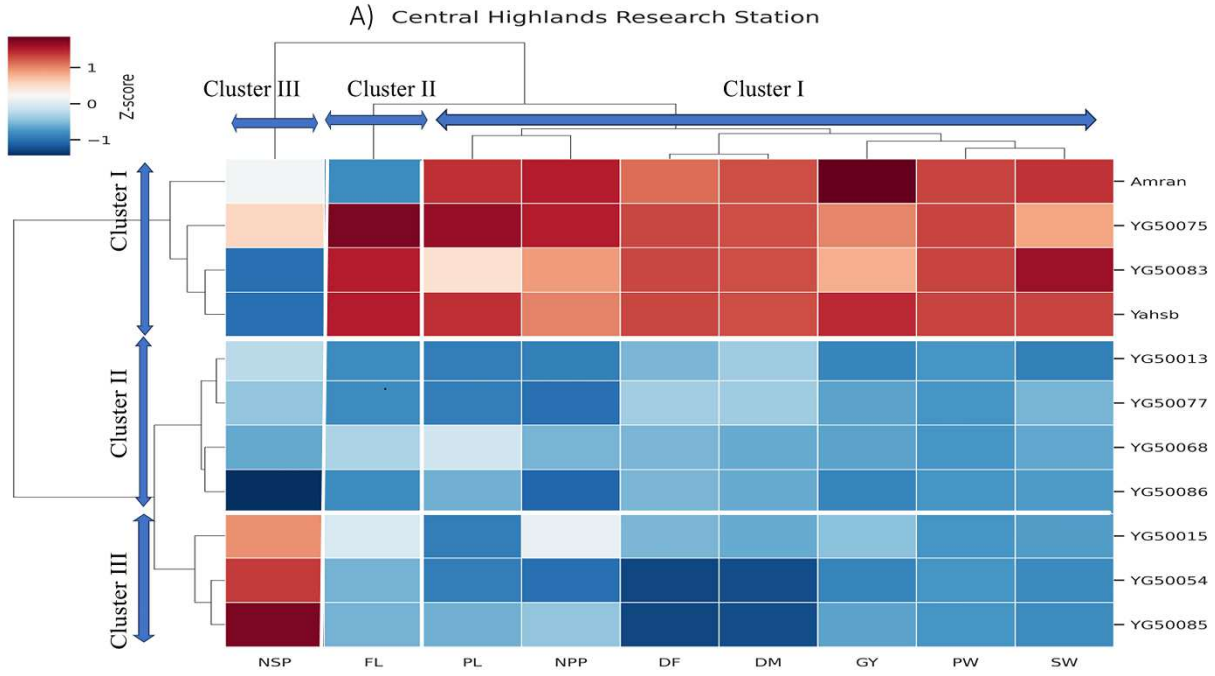
The heatmap combined with hierarchical clustering revealed a clear separation of the evaluated pea genotypes into two major groups (Fig. 4A). The first group consisted of the improved genotypes Yahsb and Amran together with YG50075 and YG50083, which exhibited consistently high standardized values for grain yield (GY), seed weight (SW), pod length (PL), pod width (PW), and number of pods per plant (NPP). The clustering pattern indicates a high degree of phenotypic similarity among these genotypes, supporting the hypothesis that YG50075 and YG50083 share a close relationship with the improved cultivars. In contrast, the remaining local genotypes formed a separate cluster characterized by lower values for most yield-related traits. Trait clustering further grouped GY, SW, NPP, PL, and PW within the same cluster, reflecting their strong positive association and their collective contribution to productivity under Central Highlands conditions.

Northern Highlands Research Station

Under Northern Highlands conditions, the heatmap also separated the genotypes into two major clusters (Fig. 4B). Yahsb and Amran remained associated with relatively high values for grain yield and several agronomic traits, whereas YG50075 and YG50083 formed a closely related subgroup characterized by delayed flowering and maturity as well as larger pod dimensions. Several local genotypes, including YG50013, YG50054, YG50077, YG50085, and YG50086, clustered together because of their relatively low and uniform trait values. Compared with the Central Highlands, a stronger contribution of phenological traits such as days to flowering (DF) and days to maturity (DM) was evident in determining genotype differentiation. The trait dendrogram revealed close associations among DF, DM, FL, and pod-related traits, suggesting that developmental timing played an important role in genotype adaptation under the Northern Highlands environment.

Overall, the heatmap analysis confirmed the existence of distinct phenotypic groups among the evaluated genotypes and highlighted the close relationship between the improved cultivars and genotypes YG50075 and YG50083. Furthermore, the clustering of grain yield with pod length, pod width, number of pods per

plant, and seed weight was consistent with the correlation, PCA, and path coefficient analyses, indicating that these traits represent the principal determinants of productivity across environments. The differences observed between environments further demonstrate the influence of environmental conditions on trait expression and genotype grouping patterns.



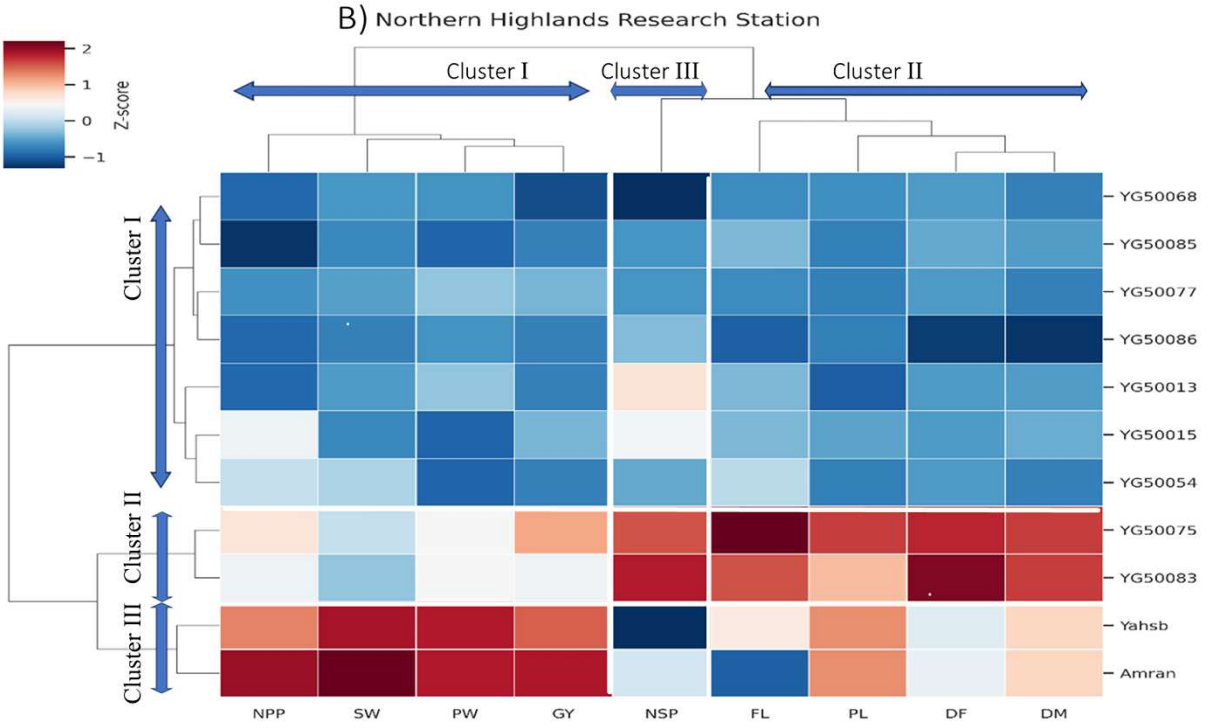


Figure 4. Clustered heatmaps of standardized quantitative traits for eleven pea genotypes evaluated under (a) Central Highlands and (b) Northern Highlands environments. Trait values were standardized using Z-scores prior to analysis. Hierarchical clustering based on Euclidean distance and Ward's method was used to simultaneously classify genotypes and traits, providing an integrated visualization of phenotypic variation, trait associations, and genotype grouping patterns across environments. DF = Days to flowering; FL = Flower length; DM = Days to maturity; PL = Pod length; PW = Pod width; NPP = Number of pods per plant; NSP = Number of seeds per pod; SW = 100-seed weight; GY = Grain yield.

Discussion

The present study demonstrated significant genetic variability among pea genotypes across contrasting highland environments, as evidenced by the highly significant ANOVA results. This finding is consistent with Annicchiarico et al.[7], who reported that multi-environment trials in legumes often reveal substantial genotypic variation due to differential adaptation. Similarly, Mohammadi et al.[9] emphasized that high phenotypic variability is a prerequisite for effective selection in breeding programs. However, unlike these general reports, the current study clearly shows that variability is not only present but also environment-dependent, which strengthens its relevance for semi-arid systems. The higher grain yield observed in the Central Highlands compared to the Northern Highlands indicates that favorable environmental conditions enhance the expression of yield potential. This aligns with the findings of Berger et al[43], who indicate that environmental conditions and phenological adaptation strongly influence reproductive performance in grain legumes. However, in contrast to their findings, in which yield variation was attributed mainly to climatic factors, the present study shows that genotype × environment interaction (G×E) plays a critical role, as evidenced by the differential ranking of genotypes across environments.

The qualitative and coded quantitative descriptors provided a clear morphological differentiation among the evaluated genotypes. The improved cultivars (Yahsb and Amran) were consistently characterized by white flowers, green seeds, and diamond-shaped seeds, whereas most local genotypes exhibited purple flowers, brown seeds, and irregular seed shapes. Interestingly, YG50075 and YG50083 shared most descriptor states with the improved cultivars and were grouped together in the cluster analysis. This pattern suggests a close genetic relationship and supports field observations indicating that these genotypes may represent farmer-maintained derivatives of the released cultivars. Such morphological differentiation is particularly valuable for germplasm characterization and parental selection in breeding programs,as reported by Ton et al[20].

The consistent superiority of Yahsb and Amran across both environments suggests their wide adaptability and stability. This observation agrees with Olivoto et al.[44], and Sabaghnia et al.[45] who reported that stable genotypes maintain high performance across diverse environments due to balanced trait expression. Similarly, Yan et al.[46] highlighted that stable genotypes are typically located near the ideal genotype position in multivariate analyses. In the present study, this was clearly reflected in the PCA results, where Yahsb and Amran remained closely associated with yield vectors across both environments, confirming their robustness. A particularly interesting finding was the behavior of YG50075 and YG50083. Although these genotypes clustered closely with Yahsb and Amran and shared most qualitative descriptors, their performance declined more markedly under Northern Highland conditions. This suggests that farmer-mediated selection maintained much of the original phenotypic identity of these materials, while gradual adaptation under local cultivation may have altered their environmental responsiveness. Such divergence between phenotypic similarity and environmental stability highlights the importance of multi-environment evaluation when characterizing farmer-maintained germplasm. This partially agrees with Hanci et al.[47], and Pratap et al.[48] who reported that farmer-maintained varieties may share phenotypic similarity with improved cultivars due to continuous farmer selection and local adaptation while exhibiting divergence in performance due to local adaptation. However, unlike their findings, the current study

demonstrates that such adaptation may lead to reduced stability, particularly under stress environments, suggesting possible genetic drift or selection pressure under farmer conditions. The PCA analysis revealed a clear shift in trait importance between environments. In the Central Highlands, yield was primarily associated with seed weight (SW) and number of pods per plant (NPP), indicating that yield formation is driven by reproductive efficiency. This is in agreement with Georgiev et al.[49], who found that yield components dominate under favorable conditions. In contrast, in the Northern Highlands, phenological traits such as days to flowering (DF) and days to maturity (DM) contributed more to variation. This agrees with Berger et al.[43], who demonstrated that phenological adaptation plays an important role in crop performance across contrasting environments. Correlation analysis further confirmed these patterns. In the Central Highlands, strong positive correlations between grain yield and traits such as PL, PW, NPP, and SW suggest that biomass accumulation and assimilate partitioning are key determinants of yield. These results are consistent with Sharma et al.[50], who reported similar associations in pea under optimal conditions. However, in the Northern Highlands, yield was strongly associated with both phenological and yield traits, indicating a combined effect of adaptation and productivity. This partially agrees with Singh et al.[51] and Singh & Chaudhary[24] who reported the importance of both adaptive and yield-related traits in determining productivity under variable environments. However, the present study provides stronger evidence that both trait groups jointly contribute to grain yield under stress conditions. Path coefficient analysis provided deeper insight into the causal relationships among traits. In the Central Highlands, seed weight (SW) showed the highest direct effect on yield, confirming its primary role under favorable conditions. This is consistent with findings by Sharma et al.[50] and Ton et al.[52], who reported that seed weight and other yield-component traits exert substantial direct effects on grain yield in pea. Conversely, in the Northern Highlands, pod length (PL) and days to maturity (DM) had stronger direct effects, highlighting the importance of structural and developmental traits under stress. This partially agrees with Chisa et al[53], who reported substantial direct contributions of pod-related and yield-component traits to seed yield. The estimation of relative trait contribution further clarified the environmental shift in yield determination mechanisms. Under the Central Highlands conditions, seed weight contributed most strongly to grain yield, whereas under the Northern Highlands, pod-related and phenological traits contributed more substantially. Similar observations were reported in legumes by Dewey and Lu[23], who emphasized that partitioning correlation coefficients into direct and indirect effects improves the biological interpretation of complex yield systems. These findings support the concept that environmental stress alters the hierarchy of yield-determining traits and consequently modifies selection priorities in breeding programs.

The heatmap coupled with hierarchical clustering provided an integrated visualization of genotype performance and trait relationships across environments. The clustering patterns were largely consistent with those obtained from the dendrogram and PCA analyses, providing additional support for the observed phenotypic relationships among genotypes. In addition, the close association of grain yield with pod length, pod width, number of pods per plant, and seed weight was consistent with the correlation, PCA, and path coefficient analyses. These findings demonstrate that heatmap-based clustering is an effective complementary tool for visualizing complex multivariate relationships and genotype–trait associations under contrasting environmental conditions (Wilkinson and Friendly)[54]. Cluster analysis supported the multivariate results by grouping genotypes according to their phenotypic similarity. The consistent clustering of Yahsb, Amran, YG50075, and YG50083 confirms their phenotypic similarity, consistent with the grouping patterns reported by Hluyako et al[55]. However, the clear separation of local genotypes indicates substantial genetic divergence, which is valuable for breeding programs aiming to exploit heterosis. This agrees with Behpouri et at.[56] who demonstrated the usefulness of multivariate approaches for identifying phenotypic divergence and yield-related diversity among genotypes. A notable finding is the consistently low contribution of number of seeds per pod (NSP) to grain yield. Despite being a yield component, NSP showed weak correlations and minimal direct effects. This observation is consistent with the weak direct and indirect contributions of seed number traits reported by Srivastava et al.[25]. However, the present study clearly demonstrates that this compensation is stable across environments, reinforcing its biological significance.

Overall, the integration of multiple statistical approaches (ANOVA, PCA, correlation, path coefficient analysis, heatmap visualization, and cluster analysis) provides a robust framework for understanding yield determination. Similar multi-environment analytical frameworks have been recommended by Kusi, et al.[57] for evaluating genotype performance and environmental responses and Azam et al.[58] , who emphasized the integration of multiple phenotypic traits for identifying superior genotypes, but the present study advances this by explicitly linking trait importance to environmental conditions. Importantly, the results confirm that yield in pea is highly environment-dependent. Under favorable conditions, yield is primarily driven by seed weight and reproductive traits, whereas under stress conditions, phenological adaptation and plant structure become more critical. Therefore, breeding strategies should focus on combining high yield potential with environmental stability, while also targeting specific adaptation to semi-arid conditions.

Conclusion

This study demonstrated substantial genetic variability among pea (*Pisum sativum* L.) genotypes under contrasting highland environments, highlighting the strong influence of environmental conditions on yield formation and trait expression. Grain yield was primarily determined by different trait combinations depending on the environment, with yield components such as seed weight and number of pods per plant playing a dominant role under favorable conditions, while phenological traits, particularly pod length and days to maturity, were more influential under stress conditions. The integration of correlation, principal component analysis, path coefficient analysis, and heatmap-based hierarchical clustering provided consistent evidence that seed weight, pod length, pod width, number of pods per plant, and maturity-related traits are the principal determinants of grain yield. Heatmap analysis further confirmed the close phenotypic relationship among Yahsb, Amran, YG50075, and YG50083 and highlighted the environmental influence on genotype grouping patterns.

Cluster analysis further revealed strong phenotypic similarity between the improved cultivars (Yahsb and Amran) and the farmer-maintained genotypes YG50075 and YG50083, supporting their close morphological relationship. However, the reduced environmental stability observed in YG50075 and YG50083 emphasizes the importance of multi-environment evaluation when characterizing locally maintained germplasm. Moreover, the integration of qualitative and quantitative descriptors proved effective for distinguishing improved cultivars from local genotypes and provided valuable information for germplasm characterization and parent selection in future breeding programs.

Overall, the findings underscore that yield improvement in pea requires a dual strategy that combines selection for high-yielding components under favorable conditions with adaptive traits under stress environments. The applied analytical framework offers a robust approach for identifying stable and high-performing genotypes, thereby contributing to the development of resilient pea varieties adapted to variable highland agro-ecologies.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files

Reference

1. Mudryj, A. N., Yu, N. & Aukema, H. M. Nutritional and health benefits of pulses. *Appl. Physiol. Nutr. Metab.* **39**, 1197–1204 (2014). https://doi.org/10.1139/apnm-2013-0557
2. Dahl, W. J., Foster, L. M. & Tyler, R. T. Review of the health benefits of peas (*Pisum sativum* L.). *Br. J. Nutr*. **108**, S3–S10 (2012). https://doi.org/10.1017/S0007114512000852
3. FAOSTAT. Crops production database. http://www.fao.org/faostat/en/#data/QC (2021).

4. Peoples, M. B. et al. The contributions of nitrogen-fixing crop legumes to the productivity of agricultural systems. *Symbiosis* **48**, 1–17 (2009). https://doi.org/10.1007/BF03179980
5. Foyer, C. H. et al. Neglecting legumes has compromised human health and sustainable food production. *Nat. Plants* **2**, 16112 (2016). https://doi.org/10.1038/nplants.2016.112
6. Acquaah, G. Principles of Plant Genetics and Breeding (2nd ed.) (Wiley-Blackwell, 2012). https://doi.org/10.1002/9781118313718
7. Annicchiarico, P. Genotype × environment interactions: Challenges and opportunities for plant breeding and cultivar recommendations (FAO, 2002). https://www.researchgate.net/publication/322951880
8. Yan, W. & Kang, M. S. GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists and Agronomists (CRC Press, 2003). https://doi.org/10.1201/9781420040371
9. Mohammadi, S. A. & Prasanna, B. M. Analysis of genetic diversity in crop plants—salient statistical tools and considerations. *Crop Sci*. **43**, 1235–1248 (2003). https://doi.org/10.2135/cropsci2003.1235
10. Bashir, I., Ishtiaq, S., Fiaz, S. & Sajjad, M. Association of yield attributing traits in pea (*Pisum sativum* L.) germplasm. *Banat J. Biotechnol*. **8**, 43–49 (2017). https://doi.org/10.7904/2068-4738-VIII(15)-43
11. Nyirenda, Y. E. et al. Phenotypic divergence analysis in pigeonpea (*Cajanus cajan* (L.) Millsp.) germplasm accessions. *Agronomy* **10**, 1682 (2020). https://doi.org/10.3390/agronomy10111682
12. Smith, J. S. C. & Smith, O. S. Fingerprinting crop varieties. *Adv. Agron*. **47**, 85–140 (1992). http://dx.doi.org/10.1016/S0065-2113(08)60489-7
13. Siddique, M. N. A., Sarker, U. & Mian, M. A. K. Genetic diversity in restorer line of rice. In Proc. Int. Conf. Plant Breeding and Seed for Food Security, 137–142 (2009).
14. Rai, P. K., Sarker, U., Roy, P. C. & Islam, A. Character association in F4 generation of rice (*Oryza sativa* L.). *Bangladesh J. Plant Breed. Genet*. **26**, 39–44 (2013).
15. Karim, D. et al. Phenotypic and genotypic correlation coefficient of quantitative characters and character association of aromatic rice. *J. Biosci. Agric. Res.* **1**, 34–46 (2014).
16. Uyeda, J. C., Caetano, D. S. & Pennell, M. W. Comparative analysis of principal components can be misleading. *Syst. Biol.* **64**, 677–689 (2015).
17. Malviya, N., Kumar, K. & Upadhyay, D. K. Genetic divergence in Indian mustard (Brassica juncea L.). Bangladesh J. Bot. **50**, 37–44 (2021).
18. Jolliffe, I. T. Principal Component Analysis (2nd ed.) (*Springer*, 2002). https://doi.org/10.1007/b98835
19. Abdi, H. & Williams, L. J. Principal component analysis. Wiley Interdiscip. *Rev. Comput. Stat*. **2**, 433–459 (2010). https://doi.org/10.1002/wics.101
20. Ton, A. et al. Characterization of some local pea (*Pisum sativum* L.) genotypes for agro-morphological traits and mineral concentrations. *Turk. J. Agric. For*. **46**, (2022). https://doi.org/10.55730/1300-011X.2975
21. Sneath, P. H. & Sokal, R. R. Numerical Taxonomy: The Principles and Practice of Numerical Classification (W. H. Freeman, 1973). https://www.scirp.org/reference/referencespapers?referenceid=1229485
22. Govindaraj, M., Vetriventhan, M., & Srinivasan, M. Importance of Genetic Diversity Assessment in Crop Plants and Its Recent Advances: An Overview of Its Analytical Perspectives. *Genetics Research International*. **14**, (2015). http://dx.doi.org/10.1155/2015/431487
23. Dewey, D. R. & Lu, K. H. A correlation and path-coefficient analysis of components of crested wheatgrass seed production. *Agron. J.* **51**, 515–518 (1959). https://doi.org/10.2134/agronj1959.00021962005100090002x
24. Singh, R. K. & Chaudhary, B. D. Biometrical Methods in Quantitative Genetic Analysis (Kalyani Publishers, 1979). https://www.scirp.org/reference/referencespapers?referenceid=4046345
25. Srivastava, A., Sharma, A., Singh, T. & Kumar, R. Correlation coefficient and path coefficient in field pea (*Pisum sativum* L.). *Int. J. Curr. Microbiol. Appl. Sci.* **7**, 549–553 (2018). https://doi.org/10.20546/ijcmas.2018.702.069

26. Ceccarelli, S. Efficiency of plant breeding. *Agron. Sustain. Dev.* **35**, 423–438 (2015). https://doi.org/10.2135/cropsci2014.02.0158
27. Brhane, H. & Hammenhag, C. Genetic diversity and population structure analysis of a diverse panel of pea (*Pisum sativum*). *Front. Genet.* **15**, 1396888 (2024). https://doi.org/10.3389/fgene.2024.1396888
28. Shrestha, D. S. et al. Agro-morphological characterization and diversity assessment of pea (*Pisum sativum* L.) germplasm conserved in the genebank of Nepal. *Int. J. Res. Publ. Rev.* **4**, 422–435 (2023). https://doi.org/10.55248/gengpi.2023.4106
29. Kumar, R. & Kumar, M. Estimation of genetic divergence in garden pea (*Pisum sativum* var. Hortense L.) genotype to facilitate the selection of potential parents for hybridization program. *Legume Res*. **39**, 709–712 (2016).
30. Kumar, S. et al. Evaluation of seed yield and genetic divergence in germplasm of *Pisum sativum* L. var. arvense. *J. Pharmacogn. Phytochem*. **6**, 1016–1021 (2017).
31. Sarker, U. & Oba, S. Nutrients, minerals, pigments, phytochemical and radical scavenging activity in Amaranthus blitum. *Sci. Rep*. **10**, 3868 (2020). https://doi.org/10.1038/s41598-020-59848-w
32. Sarker, U. & Ercisli, S. Salt eustress induction in red amaranth (*Amaranthus gangeticus*) augments nutritional, phenolic acids and antiradical potential of leaves. *Antioxidants* **11**, 2434 (2022). https://doi.org/10.3390/antiox11122434
33. Azam, M. D., Sarker, U. & Uddin, M. S. Screening maize (*Zea mays* L.) genotypes for phosphorus deficiency at the seedling stage. *Turk. J. Agric. For*. **46**, 3 (2022). https://doi.org/10.55730/1300-011X.3044
34. Hasan-Ud-Daula, M. & Sarker, U. Variability, heritability, character association, and path coefficient analysis in advanced breeding lines of rice (*Oryza sativa* L.). *Genetika* **52**, 711–726 (2020). https://doi.org/10.2298/GENSR2002711H
35. Sarker, U., Oba, S., Alsanie, W. F. & Gaber, A. Characterization of phytochemicals, nutrients, and antiradical potential in slim amaranth. *Antioxidants* **11**, 1089 (2022). https://doi.org/10.3390/antiox11061089
36. Hossain, M. A., Sarker, U. Azam, M.G. Kobir, R.M.S. Integrating BLUP, AMMI, and GGE models to explore GE interactions for adaptability and stability of winter lentils (*Lens culinaris* Medik). *Plants* **12**, 2079 (2023). https://doi.org/10.3390/plants12112079
37. Mareai, M., Maqbool, M., Al-Omari, T. Participatory evaluation and agro-morphological characterization of pea genotypes in dry environments. *Agric. Res. J.* **62** 399-407 (2025). https://doi.org/10.5958/2395-146X.2025.00047.9
38. Gomez, K. A. & Gomez, A. A. Statistical Procedures for Agricultural Research (2nd ed.) (Wiley, 1984). https://www.scirp.org/reference/referencespapers?referenceid=2253909
39. Wright, S. Correlation and causation. J. Agric. Res. **20**, 557–585 (1921). https://www.scirp.org/reference/referencespapers?referenceid=1706123
40. Everitt, B. S., Landau, S., Leese, M. & Stahl, D. Cluster Analysis (5th ed.) (Wiley, 2011). http://dx.doi.org/10.1002/9780470977811
41. Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics, 32(18), 2847–2849. https://doi.org/10.1093/bioinformatics/btw313
42. UPOV. Protocol for Distinctness, Uniformity and Stability Tests: Pea (*Pisum sativum* L.) (CPVO-TP/007/1, European Union, 2003). Microsoft Word - TP PISUM SATIVUM 007.docx
43. Berger, J. D. et al. Chickpea evolution has selected for contrasting phenological mechanisms among different habitats. *Euphytica* **180**, 1–15 (2011). https://hdl.handle.net/20.500.11766/13465
44. Olivoto, T. et al. Mean Performance and Stability in Multi-Environment Trials I: Combining Features of AMMI and BLUP Techniques. *Agron. J.* **111**, 2949–2960 (2019). https://doi.org/10.2134/agronj2019.03.0220
45. Sabaghnia, N., Karimizadeh, R., & Mohammadi, M. (2012) Genotype by environment interaction and stability analysis for grain yield of lentil genotypes Žemdirbystėm *Agriculture* **99**, 305–312 (2012). https://www.researchgate.net/publication/286380675_
46. Yan, W. & Tinker, N. A. Biplot analysis of multi-environment trial data. *Can. J. Plant Sci*. **86**, 623–645 (2006). https://doi.org/10.4141/P05-169

47. Hanci, F., & Cebeci, E. Determination of morphological variability of different pisum genotypes using principal component analysis. *Legume Res.* **42**, 162–167 (2018). https://doi.org/10.18805/LR-438
48. Pratap, V. et al. Multivariate analysis of quantitative traits in field pea (*Pisum sativum* var. arvense). *Legume Res.* **47**, 917–921 (2024). https://doi.org/10.18805/LR-4604
49. Georgieva, N., Nikolova, I. & Kosev, V. Association study of yield and its components in pea (*Pisum sativum* L.). *Int. J. Pharmacogn*. **2**, 536–542 (2015). https://doi.org/10.13040/IJPSR.0975-8232.IJP.2(11).536-42
50. Sharma, D., Kumari, M., Sandeep, S., & Thakur, N. (2023). Correlation and path coefficient analysis for yield and component traits in pea (*pisum sativum* l.). *Bangladesh J. Bot*. **52**, 907–914 (2023). https://doi.org/10.3329/bjb.v52i4.70570
51. Singh, S. K. et al. Estimation of correlation coefficient among yield and attributing traits of field pea (*Pisum sativum* L.). *Legume Res.* **41**, 20–26 (2018). https://doi.org/10.18805/LR-3449
52. Ton A, Karaköy T, Anlarsal AE, Turkeri M. Genetic variability, heritability and path analysis in field pea (*Pisum sativum* L.). *Fresenius Environ. Bull.* **27**, 2275–2279 (2018). https://www.researchgate.net/publication/330358333
53. Chisa, A. et al. Correlation and path coefficient analysis for seed yield and agronomic traits of *pigeon pea*. *J. Agric. Sci*. **13**, 33 (2021). https://doi.org/10.5539/jas.v13n11p33
54. Wilkinson, L., & Friendly, M. The History of the Cluster Heat Map. The American Statistician, 63(**2**), 179–184 (2009***).*** https://doi.org/10.1198/tas.2009.0033
55. Hluyako, L. L. et al. Characterisation of Pigeon Pea (*Cajanus cajan*) Landraces Grown in Two Climatic Zones in KwaZulu-Natal Province, South Africa. *S. Afr. J. Plant Soil* **34**, 191–199 (2017). https://doi.org/10.1080/02571862.2016.1266044
56. Behpouri, A. et al. Use of multivariate analysis and machine learning methods to characterize traits contributing to wheat yield diversity, *Span. J. Agric. Res.* **21** (2023). https://www.researchgate.net/publication/366894485
57. Kusi, F. et al. Genotype × environment interaction effect on grain yield of cowpea. *PLoS One* **20**, e0314464 (2025). https://doi.org/10.1371/journal.pone.0314464
58. Azam, M.G., Sarker, U., Hossain, M.A. et al. Phenotypic diversity in qualitative and quantitative traits for selection of high yield potential field pea genotypes. *Sci Rep* **14**, 18561 (2024). https://doi.org/10.1038/s41598-024-69448-7

Acknowledgements

The authors sincerely thank Agricultural Research and Extension Authority (AREA) for providing research facilities and financial support for this study. They also thank National Genetic Resource Centre (NGRC), for their kind assistance in the proper execution of these trials at their fields and Northen Highlands Research Station and Central Highlands Research Station for recording observations on field pea reaction meticulously.

Author contributions

M.M. and N.Z. initiated the research work and conceived the study; M.M. and N.Z. performed the experiments; M.M. and N.Z. performed statistical analysis; M.M. and N.Z, and T.A. drafted, edited, interpreted data, and prepared the manuscript. M.M. and N.Z, and T.A. contributed to revising the manuscript and English editing the revised manuscript**.**

Funding

The Authors declare that there was no fund receive for conducting this research

Declarations

Competing interests

The authors declare no competing interests.

Consent for publication

All authors have given consent for publication.

Additional information

Correspondence and requests for materials should be addressed to M.M.

Supplementary Files

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

- Supplementarymaterial.docx
- correlationtablesbystation.xlsx
- heatmapclusteringresults.xlsx