Journal of Medicinal and Industrial Plant Sciences (2025) 3 (3): 133- 150
ISSN: 2959 – 121X
DOI: https://doi.org/10.32894/MEDIP.25.3.3.10

Genetic diversity of sweet and forage corn (*Zea mays* L.) single-cross hybrids based on phenotypic cluster and principal component analysis

Neyaz Rashid Mustafa
Department of Biology, College of Education-Shaqlawa, Salahaddin University, Erbil, Iraq.
* *Corresponding author: E-mail: neyaz.mustafa@su.edu.krd*

ABSTRACT

Cluster analysis and principal component analysis are multivariate analyses used widely to assess genetic diversity. The present study was conducted in the Autumn of 2024 at the Grdarasha experimental station, College of Agriculture Engineering Sciences, Salahaddin University, Erbil, Iraq, to assess the phenotypic genetic diversity among 20 sweet and forage corn single-cross hybrids using cluster and principal component analysis. Results demonstrated that the 20 single-cross hybrids were significantly different from each other. Moreover, almost all of the traits studied showed high broad-sense heritability, which is important for selecting corn single-cross hybrids. Cluster analysis and principal component analysis revealed a high level of genetic diversity, which has implications for characterizing, conserving, and breeding sweet and forage corn single-cross hybrids, as well as for categorizing them. The hybrids under study were divided into six different groups based on the performance of phenotypic traits, indicating that the hybrids have a varied genetic background. The cluster analysis and principal component analysis were also able to separate sweet corn well from the forage corn. This indicated the differentiation of the genetic makeup of sweet corn from forage corn. Biplot analysis showed positive correlations among ear yield and several traits such as ear weight, ear length, number of kernel rows per ear, number of leaves per plant, ear height, plant height, number of leaves per ear, leaf area, stem diameter, and number of ears per plant. A correlation of the first three principal component analyses accounted for 76.26% of the variation, indicating a significant variation among the hybrids studied.

KEYWORDS: Corn; Biplot; Heritability; Single-Cross Hybrids; Genetic Diversity.

Received: 28/08/2025; **Accepted**: 22/09/2025; **Available online:** 30/09/2025

التنوع الوراثي للهجن المفردة من الذرة (*Zea mays* L.) الحلوة والعلفية استنادًا إلى تحليل المجموعة الظاهرية والمكونات الرئيسية

نياز رشيد مصطفى
قسم علوم الحياة، كلية التربية - شقلاوة، جامعة صلاح الدين، أربيل، العراق.

الملخص

التحليل العنقودي وتحليل المكونات الرئيسية هما تحليلان متعددا المتغيرات يُستخدمان على نطاق واسع لتقييم التنوع الوراثي. أُجريت هذه الدراسة في خريف عام 2024 في محطة گردەرشه التجريبية، بكلية علوم الهندسة الزراعية، جامعة صلاح الدين، أربيل، العراق، لتقييم التنوع الوراثي المظهري بين 20 هجينًا أحادي التهجين من الذرة الحلوة والعلفية باستخدام التحليل العنقودي وتحليل المكونات الرئيسية. أظهرت النتائج اختلافًا كبيرًا بين الهجن العشرين أحادية التهجين. علاوة على ذلك، أظهرت جميع الخصائص المدروسة تقريبًا معدل توريث مرتفع على النطاق الواسع ، وهو أمر مهم لاختيار هجن الذرة أحادية التهجين. كشف التحليل العنقودي وتحليل المكونات الرئيسية عن مستوى عالٍ من التنوع الوراثي، مما يؤثر على توصيف وحفظ وتربية هجن الذرة الحلوة والعلفية أحادية التهجين، بالإضافة إلى تصنيفها. قُسِّمت الهجن قيد الدراسة إلى ست مجموعات مختلفة بناءً على أداء الصفات المظهرية، مما يشير إلى أن الهجن لها خلفية وراثية متنوعة. كما تمكّن التحليل العنقودي وتحليل المكونات الرئيسية من فصل الذرة الحلوة جيدًا عن الذرة العلفية. وهذا يُشير إلى تمايز التركيب الجيني للذرة الحلوة عن الذرة العلفية. أظهر تحليل القطع الثنائية ارتباطات إيجابية بين إنتاج الكوز وعدد من الصفات، مثل وزن الكوز، وطول الكوز، وعدد صفوف الحبوب في الكوز، وعدد الأوراق في النبات، وارتفاع الكوز، وارتفاع النبات، وعدد الأوراق في الكوز، ومساحة الورقة، وقطر الساق، وعدد الكوز في النبات. وقد ساهم ارتباط تحليلات المكونات الرئيسية الثلاثة الأولى بنسبة 76.26% من التباين، مما يشير إلى تباين كبير بين الهجن المدروسة.

الكلمات المفتاحية: الذرة؛ التنوع الوراثي؛ التوريث؛ احادية التهجين؛ الاختلاف الجيني.

INTRODUCTION

Corn is distinguished by a variety of forms with highly differentiated features, both botanical and of utility character. Among the corn subspecies grown, sweet corn has become increasingly

important. Its production is constantly increasing in all countries due to its taste and nutritional value, making it a valued crop (Szymanek et al., 2006). In terms of forage, corn normally produces higher energy yields than other forage crops, particularly when it contains the leaves, stalks, and ears. It is an energy-rich feed for livestock (Brewbaker, 2003). The limiting factors of corn growers in the Kurdistan Region of Iraq are the development, improvement, maintenance, and uncontrolled quality of inbred lines. Hence, because of the absence of locally produced hybrid seeds, corn farmers must pay a high price for imported seeds, which raises the price of production (Mustafa et al. 2025).

Genetic variability is a crucial element for developing successful breeding programs and is essential for adapting to environmental changes. Genetic Diversity (GD) differs from genetic variability, which pertains to the actual phenotypic differences observed within a specific population (Šućur et al., 2023). Melchinger and Gumber (1998) reported that various assessment techniques, including phenotypic markers, pedigree information, heterosis, and molecular markers, have been used to evaluate GD in plants. Due to their affordability, ease of measurement, and speed, phenotypic markers have been widely employed to assess GD (Rahman et al., 2015). According to Zafar et al. (2022), the levels of selection, recombination, mutation, and random genetic drift all influence the amount of GD in crop germplasm. Selection and genetic drift eliminate certain alleles, while mutation and recombination introduce new variations into a population.

Various statistical techniques are used to characterize diversity within and between plant species (Ivandro Bertan et al., 2007). The level of GD can be measured using both univariate and multivariate methods. Recently, multivariate analyses have gained popularity as a means to estimate the level of genetic variation across different traits (Chavan et al., 2023). Genetic divergence between two species or individuals is measured by genetic distance, which helps evaluate the degree of genetic variation between them (Ivandro Bertan et al., 2007). To ensure all phenotypic variables contribute equally to the distance calculation, they are typically standardized before applying statistical grouping methods (Khodadadi et al., 2011). According to Muhammadi and Prasanna (2003), this standardization eliminates the effects of unit discrepancies in the measurement of each variable on variances and covariances. Fotokian et al. (2002) stated that standardization reduces differences between groups.

Cluster analysis (CA) and principal component analysis (PCA) are the two most commonly used multivariate techniques in GD research (Mohammadi and Prasanna, 2003). CA, as defined by Peeters and Martinelli (1989) and Chavan et al. (2023), is an essential method for classifying data by dividing genetic material into several homogeneous groups based on morphogenetic features. In CA, various algorithms have been employed to study genetic diversity. The most prevalent include: I) Hierarchical Clustering; II) Non-Hierarchical Clustering (e.g., K-means) (Mohammadi and Prasanna,

2003); III) Other useful algorithms (e.g., PCA, principal coordinate analysis (PCoA), and STRUCTURE); and IV) Distance/dissimilarity measures. According to the literature on plant germplasm collection structure, the most widely used clustering techniques are Ward's method (Ward, 1963) and the Unweighted Pair Group Method of Arithmetic Means (UPGMA) (Sokal and Michener, 1958).

Since PCA does not remove any samples or features, it serves as a foundation for multivariate data analysis. It is commonly used in plant sciences to reduce large datasets, enhance interpretability, and minimize information loss simultaneously (Stephen et al., 2016; Chavan et al., 2023). Additionally, PCA generates two- or three-dimensional scatter plots of individuals, allowing the genetic distances between genotypes to be represented by the geometric distances between points. Standardized values are employed to investigate how each attribute contributes to overall variability (Obeng-Antwi et al., 2011). The initial step in PCA involves calculating eigenvalues, which reflect the total variance along the PC axes. The first PC accounts for most of the variability in the original data compared to all other PCs. The second PC explains most of the remaining variability not captured by the first and is uncorrelated with it, and so forth (Jolliffe, 1986).

A biplot is a type of graph used in multivariate analysis to visualize the structure and relationships within a dataset, often using the results of principal component analysis (PCA) or singular value decomposition (SVD). It is a graphical representation that combines information about both observations and variables. It enables researchers to visualize the relationships between genotypes and traits, as well as how these relationships change across different environments. The PCA biplot (Gower and Hand, 1996) is a more modern representation that displays variables with calibrated axes and observations as points. This allows you to project the observations onto the axes and approximate the variables' initial values.

Therefore, this study aims to capture the potential phenotypic GD between a set of imported sweet and forage single-cross corn hybrids grown in the Kurdistan Region of Iraq using CA and PCA.

MATERIALS AND METHODS

Twenty single-cross hybrids of sweet and forage corn, imported from varied origins, were used in this study (Table 1). These hybrids were chosen based on their adaptability to the Kurdistan region. This experiment was conducted at the Grdarasha Experiment Station / College of Agricultural Engineering Sciences / Salahaddin University / Erbil (8 km southwest / 36.101.16" North; 44.009.25" East, and 415 meters above sea level). The climate of the region is described as semi-arid. The soil has a silty clay loam texture, and the soil pH was 7.5.

Table 1. List of The Imported Sweet and Forage Single-Cross Corn Hybrids Used in The Study.

NO.	Hybrid	Code	Types of Corn	Level of Generation
1	Arma	ARM	Sweet Corn	Single-Cross
2	Krmenia	KRM	Sweet Corn	Single-Cross
3	Snowy River	SNR	Sweet Corn	Single-Cross
4	Sugar	SUG	Sweet Corn	Single-Cross
5	Bilicious	BIL	Sweet Corn	Single-Cross
6	Burpee	BUR	Sweet Corn	Single-Cross
7	Sunny day	SUD	Sweet Corn	Single-Cross
8	Talar	TAL	Sweet Corn	Single-Cross
9	Syngenta	SGA	Forage Corn	Single-Cross
10	DKC6589	DSN	Forage Corn	Single-Cross
11	SY Batanga	SBA	Forage Corn	Single-Cross
12	Dekalb 6664	DKB	Forage Corn	Single-Cross
13	MX420	MTO	Forage Corn	Single-Cross
14	MX580	MFO	Forage Corn	Single-Cross
15	DKC5401	DFO	Forage Corn	Single-Cross
16	Agromar	AOR	Forage Corn	Single-Cross
17	MX610	MSZ	Forage Corn	Single-Cross
18	DKC6664	DSF	Forage Corn	Single-Cross
19	Reserve	RSV	Forage Corn	Single-Cross
20	NK Lucius	NLS	Forage Corn	Single-Cross

The experimental units were ploughed to a depth of 15-30 cm, fo11owed by soil rotorvation. Two seeds were manually planted per spot with four 3-meter-1ong rows at a density of 70 cm × 20 cm. The seedlings were thinned to one per point, ten days after planting. Fertilizers were applied seven days after planting, using NPK 15:15:15 at a rate of 120:120:120 kg/ha. In addition, Urea fertilizer (46% N) was applied at 15 and 35 days after planting in equal splits. Weeds were manually controlled. Over the entire plant growth cycle, dripped irrigation was used.

Hybrids were evaluated during the Autumn Season of 2024. Ten plants were selected from the middle two rows of each hybrid plot. The data were collected for ear yield (E.Y.) (Kg/ha), ear weight (E.W.) (g), ear length (E.L.) (cm), ear diameter (E.D.) (mm), number of leaves per ear (N.E.L.), number of kernel rows per ear (N.K.R.E.), number of kernels per row (N.K.R.), number of

leaves per plant (N.L.P.), stem diameter (S.D.) (mm), plant height (P.H.) (cm), ear height (E.H.) (cm), leaf area (L.A.), chlorophyll content (CHL), and number of ears per plant (N.E.P.). The hybrids were evaluated in a Randomized Complete Block Design (RCBD), with three replications. The data collected were analyzed using the PROC GLM (General Linear Model) of the Statistical Analysis System version 9.4 software (SAS Institute Inc., 2014). Subsequently, environmental, genotypic, and phenotypic variances were calculated based on the formu1a recommended by Johnson et al. (1955).

$\sigma_e^2 = MSe$ ……..1

$\sigma_g^2 = \frac{MSg - MSe}{r}$ ……..2

$\sigma_p^2 = \sigma_g^2 + \sigma_e^2$ ……..3

Were, $MSe =$ error mean square, $MSg =$ genotypic mean square, and $r =$ replication.

Then, h^2_b was calculated based on the formula proposed also by Johnson et al. (1955):

$h_b^2 = \frac{\sigma_g^2}{\sigma_p^2} \times 100$ …….4

h^2_b estimates were categorized based on the scale proposed by McWhirter (1979).

> 20%= Low\ 20-50% = Moderate\ < 50% = High

Since the units for phenotypic data varied among traits, the data were first weighted according to the formula described by Milligan and Cooper (1988), using NTSYS-pc (Rohlf, 2002).

$s = \frac{x_{ij} - \bar{x}}{\sigma}$ ……..……..5

Where: $s =$ standardized value, $x_{ij} =$ observation from i^{th} genotype in j^{th} block, $\bar{x} =$ mean value of measured trait, and $\sigma =$ standared deviation of measured trait.

The standardized data were utilized to construct the resemblance matrix of genetic distance among the hybrids based on the average Euclidean distance (Sokal and Michener, 1958), as follows:

$d_{ij} = \sqrt{\frac{1}{n} \sum k (x_{ki} + x_{kj})^2}$ ……6

Where, d_{ij} = genetic distance between two hybrids, x_{ki} = phenotypic observation of k^{th} trait in i^{th} hybrid, x_{kj} = phenotypic observation of k^{th} trait in j^{th} hybrid, and n = sample size.

There are many computer programs available to do CA and PCA. These include more specialized programs like SAS, SPSS, NTSYS-pc, Genetix, ADE-4, GenAlEx, and PCAgen (Weising et al., 2005), and recently, Grapes (Gopinath et al., 2020) and Agri Analyze (trio Radhika et al., 2023)

could be added, which not only compute multivariate statistics but also graphically present the analysis's findings. In the present study, the CA was then constructed using the UPGMA following the Sequential Agglomerative Hierarchical and Nested (SAHN) method by software Numerical Taxonomy System (NTSYS-pc: Version 2.1) (Rohlf, 2002), to observe GD based on phenotypic distance coefficients. A biplot analysis was run to visualize the correlation of traits with single-cross hybrids. The data matrix with columns representing traits, and rows representing the genotypes, was first standardized and then subjected to PCA to obtain the information on the traits most effective in discriminating genotypes on the first three PCs using the R package through the General R-shiny based Analysis Platform Empowered by Statistics in Agriculture part-1 (grapes Agril) Version 1.1.0 (Gopinath et al., 2020).

RESULTS AND DISCUSSION

Performance of hybrids

The Analysis of Variance (ANOVA) is a statistical formula used to compare variances across different groups' means (or averages). The results of ANOVA conducted for each trait are shown in Table 2. The block effect was insignificant at $p \leq 0.05$ for all measured traits. The impacts of hybrids were significant ($p < 0.05$) for all traits measured, except for E.D. The existence of reasonable genetic variability among these traits could be utilized in sweet and forage corn breeding programs separately to produce new varieties possessing the desired combinations of these traits. This finding was consistent with previous reports (Woldemariam, 2004; Mustafa, 2021).

Broad-sense heritability (h^2_B)

h^2_B is crucial for assessing the proportion of total phenotypic variance attributable to genetic factors. Bhardwaj et al. (2020) suggested that selection for traits with very high h^2_B (70% or more), should be easy because the phenotype and genotype synchronize closely due to the minimal impact of the environment on the phenotype, while traits with low h^2_B are extremely difficult or practically unsuitable for selection because of the tendency of the environments to mask genotypic impact. The results showed high h^2_B for most traits studied (Table 2). Indicating that a good match between phenotypic and genotypic values for most traits measured, and thereby a low environmental impact on the expression of these traits, and selection for these traits would be simple. High h^2_B for the traits is controlled by a multigene, which could be useful for selection (Ranjitha et al. 2018). However, environmental factors influenced E.D. (3.88%), which indicates a low h^2_B. Additionally, there were no significant differences for this particular trait (ED) among the hybrids, as exhibited by ANOVA. N.R.E. (47.36%) shows moderate h^2_B. Similarly, high h^2_B in sweet corn was reported by Mustafa (2021). Mustafa *et al*. (2021) noted moderate to high estimates for all traits measured at two locations

for forage corn hybrids and lines.

Table 2. Mean Squares in ANOVA and Broad-Sense Heritability (h^2_B) for 14 Phenotypic Traits Measured on 20 Sweet And Forage Corn Single-Cross Hybrids

S.O.V	d.f.	Mean Square						
		E.Y.	E.W.	E.L.	E.D.	N.E.L.	N.K.R.E	N.K.R.
Replications	2	25300.32	974.13	6.19	3179.94	4.46	1.50	9.26
Hybrids	19	535954.54**	10085.64**	15.67**	3352.79^{ns}	18.83**	4.81*	112.61**
Error	38	24028.74	963.41	3.37	2990.47	0.98	1.30	18.07
C.V.%		19.23	19.29	9.93	18.63	10.45	7.65	13.51
h^2_B		87.65	75.93	54.88	3.88	85.85	47.36	63.55

Cont… Table2

S.O.V	d.f.	Mean Square						
		N.L.P.	S.D.	P.H.	E.H.	L.A.	CHL	N.E.P.
Replications	2	0.41	0.74	438.54	14.50	3000.56	121.04	0.008
Hybrids	19	23.09**	44.88**	5513.32**	1913.51**	39097.97**	60.09**	0.63**
Error	38	0.92	2.85	155.98	31.79	2200.32	16.82	0.09
C.V.%		8.39	10.70	7.97	10.33	11.12	7.99	20.99
h^2_B		88.92	83.09	91.96	95.17	84.82	46.16	66.66

S.O.V: source of variations, d.f.: degree of freedom, C.V.: coefficient of variation.

Genetic Dissimilarities (dij)

dij refers to populations of organisms that have significant genetic differences from one another, often due to geographical separation or varying environmental pressures. *dij* among hybrids obtained from phenotypic characterization using average Euclidean distance are shown in Table 3. The highest phenotypic dissimilarity was obtained between SUD and MSZ, ARM and MSZ, BIL and MSZ, ARM and DKB, SUD and DKB, BUR and DKB, BUR and MSZ, BIL and DKB, SUD and MFO, SUG and MSZ, KRM and MSZ, SUD and DSF, and DFO and MSZ (*dij* = 2.46, 2.34, 2.32, 2.32, 2.30, 2.27, 2.26, 2.18, 2.12, 2.07, 2.02, 2.00, and 2.00, respectively). Due to the dissimilarity in gene pools of the populations, natural selection may act on traits that enable a population to adapt to changing environments. The greater the GD within a population (Bruford et al., 2017), the more adaptable it is likely to be. Several investigations on the use of *dij* for sweet and forage corn are available in the literature (Mustafa, 2021; Ismael, 2023; Abu Sin, 2019). In contrast, SUG and BIL

were found to be the most identical hybrids based on their phenotypic performances, with a dissimilarity value of 0.33. This suggests that these single-cross hybrids have been developed from similar source populations and have exhibited similar performance for most of the phenotypic traits measured in the field evaluation (Table 3) . We could conclude that h^2_B and *dij* are inversely related; a higher h^2_B (meaning a greater proportion of phenotypic variation is due to genetic factors) typically corresponds to lower *dij* within a population. This is because high heritability implies that individuals within the population are more genetically similar for the trait in question.

Cluster Analysis (CA)

The CA based on phenotypic distance coefficients exhibited six distinct clusters, designated as 1 to 6 (Figure 1). Group 1, consisting of seven sweet corn hybrids ARM, KRM, SNR, SUG, BIL, BUR, and SUD, was found to be separated from the other hybrids. Group 2 consists of one sweet corn hybrid, TAL. However, foliage corn hybrids SGA, DSF, and MFO were found to be placed in Group 3. Hybrid MSZ formed Group 4, while hybrids SBA, NLS, DSN, RSV, AOR, DFO, and MTO were left in Group 5. Hybrid DKB was found in group six. The clear separation of sweet corn hybrids (Groups 1 and 2) and foliage corn hybrids (Groups 3-6) through cluster analysis highlights the distinct breeding goals and genetic makeup of each type. In general, there is an indication that significant differences in phenotypic performance existed among the hybrids studied. This indicates that phenotypic diversity among the hybrids studied was considerably high for effective hybrid development. According to Heryanto et al. (2022), differences in diversity and clustering were likely caused by differences in the number and type of traits, as well as the quantity and background of hybrids used in each study. Similarly, Mustafa (2021) obtained five main groups from 27 tropical sweet corn inbred lines studied based on phenotypic traits. Several researchers have attempted to classify corn based on specific sets of phenotypic traits and proposed different recommendations, with overall yield and ear traits being highly emphasized as selection criteria (Ismael, 2023; Abu Sin, 2019).

Table 3. Genetic Dissimilarities Among 20 Sweet and Forage Corn Single-Cross Hybrids Were Determined Using Average Euclidean Distance Measured on 14 Phenotypic Traits. **Genetic Dissimilarity**

Traits	ARM	KRM	SNR	SUG	BIL	BUR	SUD	TAL	SGA	DSN	SBA	DKB	MTO	MFO	DFO	AOR	MSZ	DSF	RSV
ARM																			
KRM	0.96																		
SNR	0.91	0.60																	
SUG	0.92	0.72	0.58																
BIL	0.86	0.83	0.75	0.33															
BUR	0.81	1.04	0.99	0.69	0.61														
SUD	1.06	0.88	1.12	0.86	0.74	0.67													
TAL	1.98	1.42	1.33	1.61	1.83	1.98	1.96												
SGA	1.66	1.36	1.15	1.39	1.62	1.67	1.84	1.08											
DSN	1.52	1.12	1.19	1.33	1.52	1.49	1.55	1.39	0.75										
SBA	1.50	1.06	1.25	1.17	1.28	1.28	1.27	1.66	1.25	0.70									
DKB	2.32	1.85	1.80	1.97	2.18	2.27	2.30	1.56	1.31	1.38	1.59								
MTO	1.64	1.37	1.22	1.27	1.42	1.38	1.55	1.58	0.95	0.76	0.77	1.49							
MFO	1.95	1.76	1.54	1.75	1.97	1.89	2.12	1.67	0.84	0.94	1.46	1.60	1.06						
DFO	1.39	1.13	1.42	1.48	1.56	1.47	1.46	1.82	1.40	0.84	0.87	1.80	1.32	1.58					
AOR	1.32	0.91	1.10	1.17	1.33	1.34	1.36	1.56	0.99	0.62	0.89	1.63	1.14	1.21	0.75				
MSZ	2.34	2.02	1.74	2.07	2.32	2.26	2.46	1.55	1.14	1.38	1.82	1.59	1.31	1.00	2.00	1.77			
DSF	1.87	1.42	1.28	1.58	1.81	1.92	2.00	1.07	0.46	0.84	1.36	1.30	1.12	0.94	1.53	1.07	1.14		
RSV	1.38	1.10	1.17	1.40	1.56	1.55	1.61	1.30	0.78	0.55	1.00	1.45	1.11	1.20	0.74	0.69	1.56	0.92	
NLS	1.40	1.06	1.30	1.25	1.30	1.25	1.25	1.92	1.52	0.99	0.61	1.80	1.10	1.71	0.74	1.01	1.99	1.68	1.13

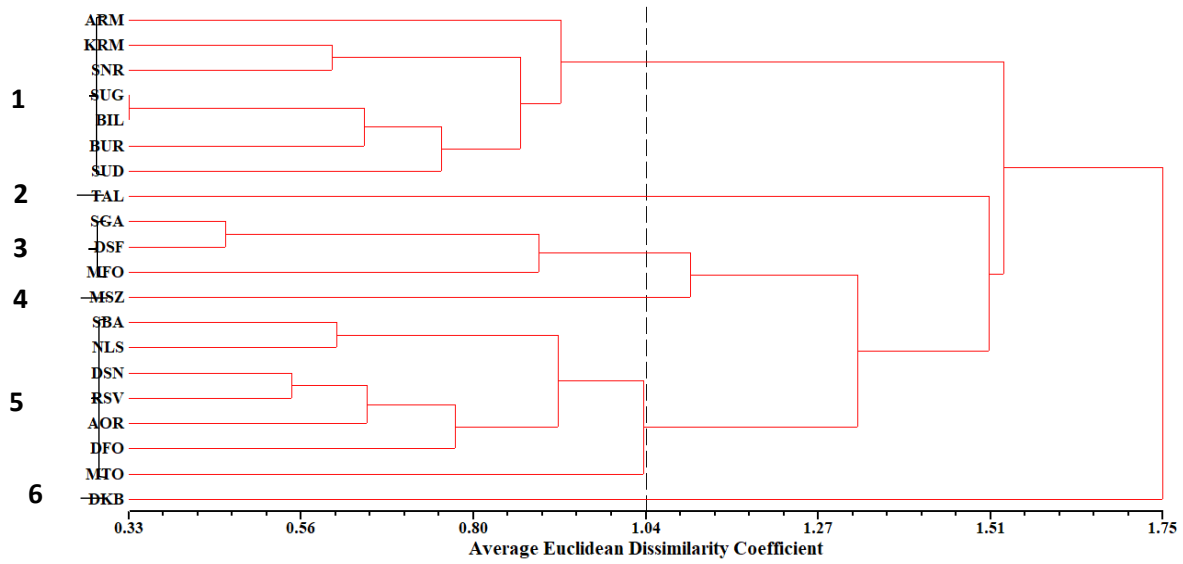


Figure 1. A dendrogram using UPGMA tree showing the grouping of 20 sweet and forage corn single-cross hybrids was conducted based on normalized average Euclidean genetic distance coefficients using 14 phenotypic traits.

Principal Component Analysis (PCA)

PCA is most commonly used to condense the information contained in a large number of original variables into a smaller set of new composite dimensions, with a minimum loss of information. PCA was executed using the standardized data obtained from the phenotypic traits of the 20 hybrids to classify the main traits that differentiated the hybrids. The eigenvalues of the PCs among the calculated standardized data on the hybrids are presented in Table 4 and Figure 2. The eigenvalues obtained indicate that they could provide a good description of the data. The first PC accounted for 49.32% of the variation out of the total variation (76.26%) among the hybrids. The PC1 was able to differentiate the hybrids by positive associations with E.Y., E.W., E.L., E.D., N.E.L., N.K.R.E., N.L.P., S.D., P.H., E.H., L.A., CHL, and N.E.P. (Table 5). Hence, PC1 was found to be associated with the yield traits of the hybrids. Mustafa et al. (2024) performed biplot analysis for eight corn hybrids in two seasons, Autumn 2020 and Spring 2021. In the first season, the total variation explained was 79.92% (66.33% and 13.59% for PC1 and PC2, respectively). However, the second season explained 60.90% of the total variation (38.29% and 22.61% for PC1 and PC2, respectively).

Biplot

Biplot analysis (Figure 3) illustrated genotype-by-trait relationships, enabling the visualization of potential parent genotypes with favorable trait combinations. A positive correlation exists between two parameters if the angle between their vectors is less than 90 degrees, and vice versa. A biplot depicts the relationship between many traits in this way (also, a measure of 90° angle

between two parameters will be treated as no correlation). And right angles, which are equal to 90°, or small vectors, indicate no correlation. The principal component Biplot expressed that variables are imposed as vectors on the graph (Latif et al*.,* 2015). Acute angles (less than 90^{o} cosine angle) were observed between E.Y. and several traits such as E.W., E.L., N.K.R.E., N.L.P., E.H., P.H., N.E.L., L.A., S.D., and N.E.P., indicating positive correlations (Figure 4) for forage corn hybrids DSF, SGA, MFO, DKB, MTO, RSV, DSN. Conversely, obtuse angles (greater than 90° cosine angle) were noted between N.K.R. and E.H., P.H., L.A., E.D., S.D., and N.E.P. for forage corn hybrids MFO, DKB, MTO, RSV, DSN, suggesting an absence of correlation among these traits.

Table 4. Principal Component Analysis Of 14 Phenotypic Traits Associated With 20 Sweet And Forage Corn Single-Cross Hybrids.

Principal Components	**Eigenvalue**	**Percentage of Variance**	**Cumulative Percentage of Variance**
PC1	6.91	49.32	49.32
PC2	2.30	16.46	65.78
PC3	1.47	10.50	76.28
PC4	1.18	8.40	84.68
PC5	0.66	4.69	89.37
PC6	0.52	3.71	93.08
PC7	0.32	2.30	95.38
PC8	0.29	2.08	97.46
PC9	0.16	1.18	98.57
PC10	0.09	0.62	99.19
PC11	0.05	0.33	99.52
PC12	0.03	0.24	99.77
PC13	0.02	0.16	99.93
PC14	0.01	0.08	100

Table 5. Principal Components (PCs) For Phenotypic Traits Measured On 20 Sweet and Forage Corn Single-Cross Hybrids.

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
E.Y.	0.32	0.21	0.14	-0.12	0.25	-0.12	0.45	-0.07	0.35	-0.28	0.31	-0.44	0.15	-0.12
E.W.	0.37	0.03	0.01	-0.09	-0.02	-0.19	0.22	-0.03	0.10	0.53	-0.03	0.16	0.14	0.66
E.L.	0.22	0.29	-0.47	0.03	-0.19	0.38	-0.25	-0.002	0.43	0.33	-0.04	-0.10	0.15	-0.28
E.D.	0.18	-0.05	0.35	0.46	-0.67	0.13	0.14	0.31	-0.06	-0.07	0.10	-0.12	0.15	0.006
N.E.L.	0.11	0.42	0.44	-0.11	0.36	0.32	-0.03	0.51	-0.05	0.09	-0.26	0.18	-0.07	-0.07
N.K.R.E.	0.20	0.31	0.46	0.002	-0.17	-0.004	-0.36	-0.67	0.09	-0.10	-0.03	0.11	-0.22	0.03
N.K.R.	-0.008	0.53	-0.40	0.04	-0.23	0.03	0.42	-0.07	-0.26	-0.28	-0.28	0.06	-0.27	0.15
N.L.P.	0.35	-0.002	-0.20	0.003	0.11	0.29	-0.17	0.09	-0.30	-0.16	0.69	0.23	-0.20	0.14
S.D.	0.31	-0.24	-0.1	-0.19	-0.16	-0.25	-0.20	0.35	0.43	-0.34	-0.22	0.08	-0.44	0.10
P.H.	0.36	-0.09	-0.07	0.03	0.14	0.03	-0.30	-0.008	-0.44	0.01	-0.30	-0.67	-0.01	0.132
E.H.	0.37	-0.07	-0.10	-0.07	0.03	-0.09	-0.06	-0.07	-0.16	-0.36	-0.25	0.41	0.65	-0.16
L.A.	0.35	-0.13	0.03	-0.10	-0.07	-0.25	0.27	-0.05	-0.27	0.40	-0.007	0.13	-0.31	-0.60
CHL	0.05	0.23	-0.12	0.74	0.32	-0.47	-0.17	0.09	0.08	0.03	0.04	0.09	-0.05	-0.05
N.E.P.	0.19	-0.42	0	0.39	0.28	0.50	0.31	-0.27	0.18	-0.06	-0.26	0.11	-0.18	0.04

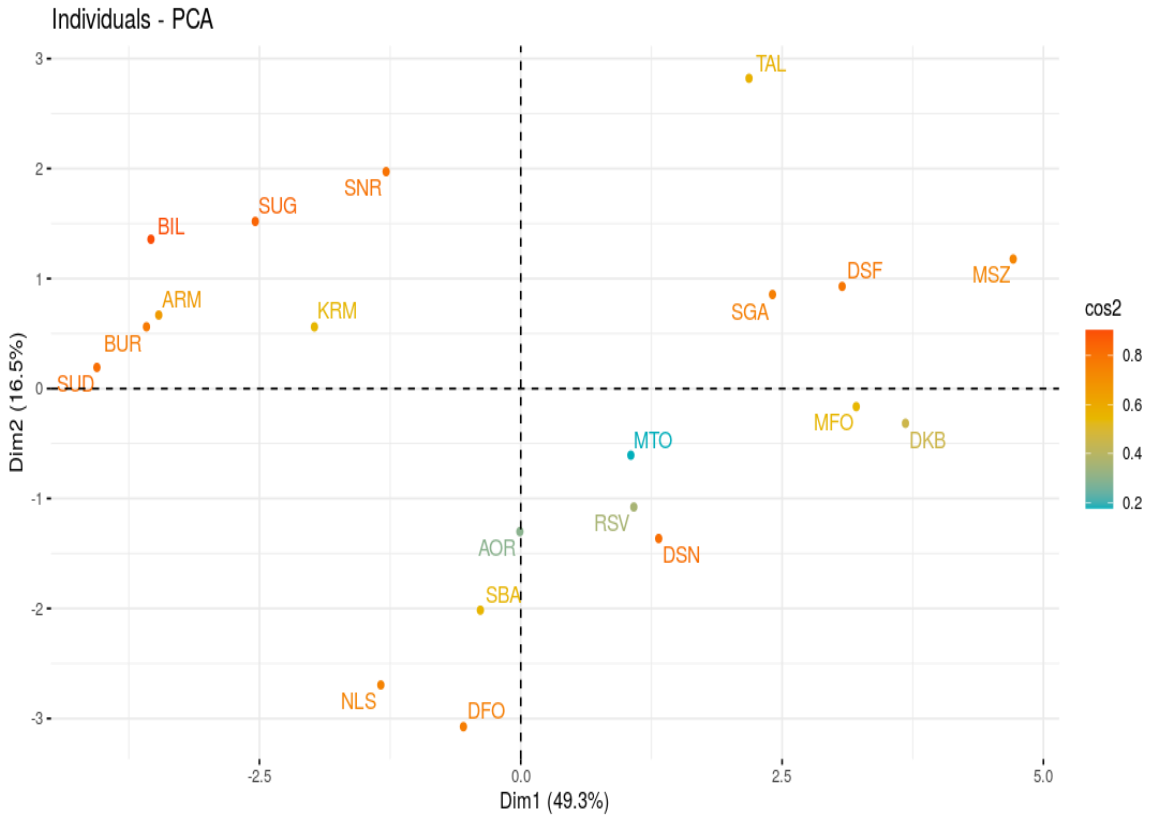


Figure 2. Principal component analysis (PCA) of 14 traits among the 20 sweet and forage corn single-cross hybrids.

The biplot analysis revealed the many ways in which distinct qualities contributed to total variation. In addition, it highlighted how traits and genotypes influence plant yield and illustrated the diversity of the examined genotypes. Due to this diversity and variability, the relevant trait in the germplasm can be enhanced, resulting in enhanced yield performance.

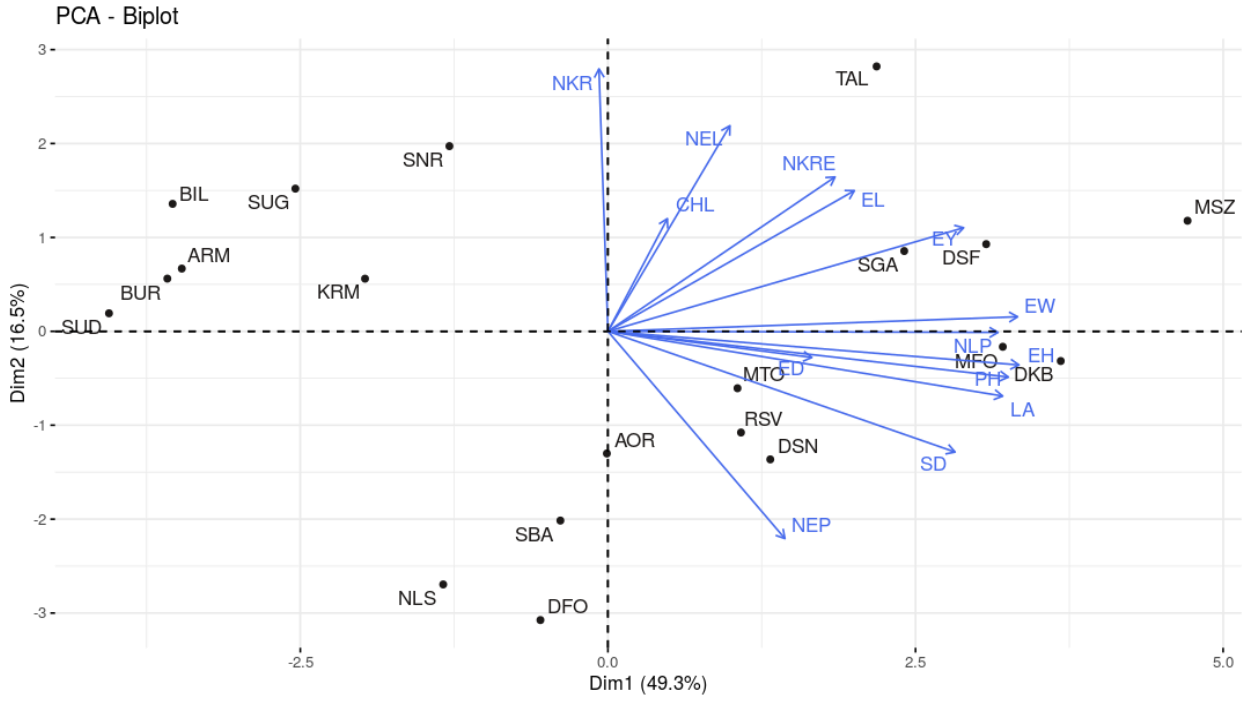
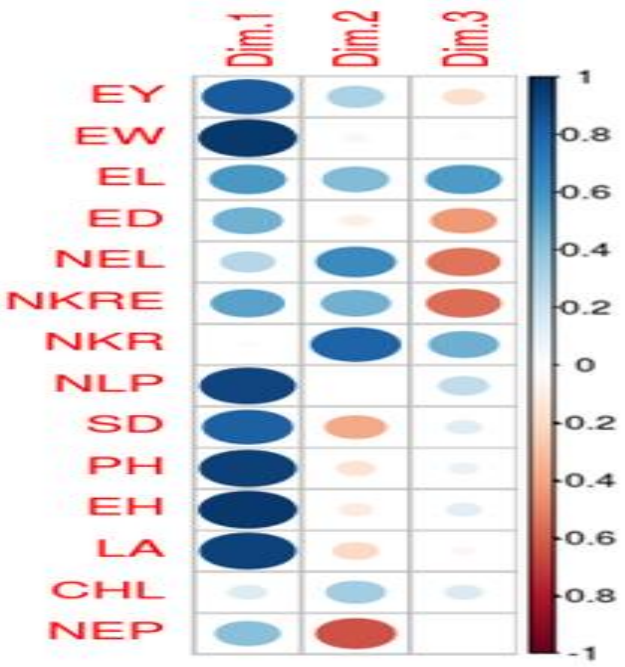


Figure 3. Biplot analysis for 20 sweet and forage corn single-cross hybrids and the traits measured.

Trait Association Analysis

A correlation plot was measured, providing knowledge on how much each variable contributes to the first three PCs and showing the extent of variable contribution for each corresponding PC (Figure 4). Observing the plot revealed that PC1 accounted for 49.3% of the variance in the dataset and was largely influenced by nearly all input variables, with the exceptions of E.L., N.E.L., N.K.R., and Chl. In relation to PC2, which explains 16.5% of the variance, an overproportional part of the variance is correlated with the N.E.L., N.K.R., and Chl. E.L. was the



primary factor influencing PC3, contributing to 10.50% of the variance. Jolliffe and Cadima (2016) noted that the first three PCs retained after reducing dimensions explained 76.28% of the variability in the datasets.

Figure 4. Correlation plot of variables VS the three principal components

CONCLUSION

In conclusion, the evaluation of 20 hybrids revealed considerable variability in their performance, with high heritability (h^2_B) for the majority of traits. Multivariate statistical analysis (CA and PCA) of phenotypic data clustered the 20 sweet and forage corn hybrids into six distinct heterotic groups. Positive correlations, as determined by biplot analysis, were observed between E.Y. and several traits, including E.W., E.L., N.K.R.E., N.L.P., E.H., P.H., N.E.L., L.A., S.D., and N.E.P. This knowledge will aid in identifying genotypes that can enhance the genetic foundation in programs aimed at improving sweet and forage corn. A correlation plot (PCs 1–3) explained 76.26% of the variation across the datasets, indicating significant variation among the traits.

Additional research should focus on collecting, characterizing, and utilizing imported sweet

and forage corn single-cross hybrids to select the most adaptive hybrid to be grown in the Kurdistan Region of Iraq. Since phenotypic data could be affected by environmental factors, to verify these findings, analysis of molecular variations could be performed on the single-cross hybrids as an additional tool to assist in the selection process for superior hybrids. At the end, CA and PCA are appropriate tools to distinguish genotypes; therefore, I suggest using them continuously for further research.

REFERENCES

Abu Sin, M. B. (2019). *Genetics and combining ability of corn (Zea mays L.) genotypes for forage utilization* (Ph.D. thesis). Department of Crop Science, Faculty of Agriculture, Universiti Putra Malaysia. http://psasir.upm.edu.my/id/eprint/84425/

Bhardwaj, A., Sharma, A., & Lata, H. (2020). Genetic variability for pod yield and related traits in garden pea (Pisum sativum L.). *Electronic Journal of Plant Breeding, 11*(4), 1233–1238. https://doi.org/10.37992/2020.1104.199

Brewbaker, J. L. (2003). *Corn production in the tropics: The Hawaii experience*. University of Hawaii at Manoa, College of Tropical Agriculture and Human Resources.

Bruford, M. W., Davies, N., Dulloo, M. E., Faith, D. P., & Walters, M. (2017). Monitoring changes in genetic diversity. In *The GEO handbook on biodiversity observation networks* (pp. 107–128). Springer International Publishing. https://doi.org/10.1007/978-3-319-27288-7_5

Chavan, S., Bhadru, D., Swarnalatha, V., & Mallaiah, B. (2023). Evaluation of variations for phenotypic traits by multivariate techniques in sweet corn (Zea mays L. saccharata). *Journal of Crop and Weed, 19*(1), 164–172. https://doi.org/10.22271/09746315.2023.v19.i1.1675

Fotokian, M., Shahnejat Bushehri, A., & Taleie, A. (2002). Cluster analysis based on PCA in rice genotypes. Paper presented at the 6th International Conference of Statistics, University of Tarbiat Modares, Iran, 26–28.

Gopinath, P. P., Parsad, R., Joseph, B., & Adarsh, V. S. (2020). *GRAPES: General Rshiny Based Analysis Platform Empowered by Statistics* (Version 1.0.0) [Software]. https://www.kaugrapes.com/home

Heryanto, F. S. S., Wirnas, D., & Ritonga, A. W. (2022). Diversity of twenty-three sweet corn (Zea mays L. saccharata) varieties in Indonesia. *Biodiversitas, 23*(11), 6075–6081. https://doi.org/10.13057/biodiv/d231164

Ismael, N. B. (2023). Characterization and grouping of imported single cross Zea mays L. hybrids based on morphological variations. *Journal of Medicinal and Industrial Plants, 1*(1), 41–51.

Ivandro Bertan, de Carvalho, F. I. F., & de Oliveira, A. C. (2007). Parental selection strategies in plant breeding programs. *Journal of Crop Science and Biotechnology, 10*(4), 211–222.

Johnson, H. W., Robinson, H. F., & Comstock, R. E. (1955). Estimates of genetic and environmental variability in soybean. *Agronomy Journal, 47*(7), 314–318. https://doi.org/10.2134/agronj1955.00021962004700070009x

Jolliffe, I. T. (1986). *Principal component analysis*. Springer-Verlag.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374*, 20150202. https://doi.org/10.1098/rsta.2015.0202

Khodadadi, M., Fotokian, M. H., & Miransari, M. (2011). Genetic diversity of wheat (Triticum aestivum L.) genotypes based on cluster and principal component analyses for breeding strategies. *Australian Journal of Crop Science, 5*(1), 17–24.

Latif, A., Bilal, M., Hussain, S. B., & Ahmad, F. (2015). Estimation of genetic divergence, association, direct and indirect effects of yield with other attributes in cotton (Gossypium hirsutum L.) using biplot correlation and path coefficient analysis. *Tropical Plant Research, 2*(2), 120–126.

McWhirter, K. S. (1979). Breeding of cross-pollinated crops. In *Plant breeding* (pp. 79–121). Australian Vice-Chancellors' Committee.

Melchinger, A. E., & Gumber, R. K. (1998). Overview of heterosis and heterotic groups in agronomic crops. In *Concepts and breeding of heterosis in crop plants* (Vol. 25, pp. 29–44). Crop Science Society of America. https://doi.org/10.2135/cssaspecpub25.c3

Milligan, G. W., & Cooper, M. C. (1988). A study of standardization of variables in cluster analysis. *Journal of Classification, 5*, 181–204. https://doi.org/10.1007/BF01897163

Mohammadi, S. A., & Prasanna, B. M. (2003). Analysis of genetic diversity in crop plants—Salient statistical tools. *Crop Science, 43*(4), 1235–1248. https://doi.org/10.2135/cropsci2003.1235

Mustafa, B. S., Ismael, N. B., Mustafa, N. R., Kakarash, S. A., & Abdulazeez, S. D. (2024). Chlorophyll content and leaf area correlated with corn (Zea mays L.) yield components in F1 hybrids. *The Indian Journal of Agricultural Sciences, 94*(4), 352–357. https://doi.org/10.56093/ijas.v94i4.140666

Mustafa, N. R. (2021). *Diversity, performance, and selection of tropical sweet corn inbred lines, and their combining abilities in hybrid combinations* (Ph.D. thesis). Universiti Putra Malaysia. https://core.ac.uk/download/529928169.pdf

Mustafa, N. R., Kakarash, S. A., Ismael, N. B., & Abdulazeez, S. D. (2025). Formation of potential heterotic groups of maize inbred lines using variation at simple sequence repeat loci. *Iraqi Journal of Agricultural Sciences, 56*(2), 657–667. https://doi.org/10.36103/tnhhjw11

Mustafa, N. R., Saleh, G. B., & Kashiani, P. (2021). Genetic potential of tropical sweet corn hybrids and combining ability among parental inbred lines. *Australian Journal of Crop Science,*

15(10), 1279–1288. https://doi.org/10.21475/ajcs.21.15.10.p3189

Obeng-Antwi, K., Craufurd, P. Q., Menkir, A., Ellis, R. H., & Sallah, P. Y. K. (2011). Intra-landrace variability of two landraces in Ghana. *International Journal of Science and Advanced Technology, 1*(9), 23–40.

Peeters, J. P., & Martinelli, J. A. (1989). Hierarchical cluster analysis as a tool to manage variation in germplasm collections. *Theoretical and Applied Genetics, 78*, 42–48. https://doi.org/10.1007/BF00299751

Rahman, S., Mia, M. M., Quddus, T., Hassan, L., & Haque, M. A. (2015). Assessing genetic diversity of maize (Zea mays L.) genotypes for agronomic traits. *Research in Agriculture, Livestock and Fisheries, 2*(1), 53–61. https://doi.org/10.3329/ralf.v2i1.23029

Ranjitha, M. C., Vaddoria, M. A., & Jethava, A. S. (2018). Genetic variability, heritability and genetic advance in garlic (Allium sativum L.) germplasm. *International Journal of Pure and Applied Bioscience, 6*(4), 401–407. http://dx.doi.org/10.18782/2320-7051.6784

Rohlf, F. (2002). *NTSYS-pc: Numerical taxonomy system* (Version 2.1). Exeter Publishing Ltd.

SAS Institute Inc. (2014). *SAS/STAT® user's guide* (Version 9.4). SAS Institute Inc.

Sokal, R. R., & Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *The University of Kansas Science Bulletin, 38*, 1409–1438.

Stephen, A., Antonia, Y. T., Patrick, T., Kingsley, B. A., & Richard, A. A. (2016). Genetic diversity in lowland, midaltitude and highland African maize landraces by morphological trait evaluation. *African Journal of Plant Science, 10*(11), 246–257. https://doi.org/10.5897/AJPS2016.1448

Šućur, R., Mladenov, V., Banjac, B., Trkulja, D., Mikić, S., Šumaruna, M., & Börner, A. (2023). Phenotypic marker study of worldwide wheat germplasm. *Italian Journal of Agronomy, 19*(1), 1–7. https://doi.org/10.4081/ija.2023.2194

Szymanek, M., Dobrzański Jr., B., Niedzióka, I., & Rybczyński, R. (2006). *Sweet corn harvest and technology: Physical properties and quality*. Bohdan Dobrzański Institute of Agrophysics, Polish Academy of Sciences.

Trio, R., Geetaben, & Bhumikaben. (2023). *Agri Analyze* [Website]. https://www.agrianalyze.com

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Weising, K., Nybom, H., Pfenninger, M., Wolff, K., & Kahl, G. (2005). *DNA fingerprinting in plants: Principles, methods, and applications*. CRC Press.

Woldemariam, M. N. (2004). *Genetic studies and selection for ear length of sweet corn (Zea mays L.)* (Ph.D. thesis). Universiti Putra Malaysia. http://psasir.upm.edu.my/id/eprint/339/

Zafar, M. M., Zhang, Y., Farooq, M. A., Ali, A., Firdous, H., Haseeb, M., Fiaz, S., Shakeel, A.,

Razzaq, A., & Ren, M. (2022). Biochemical and associated agronomic traits in *Gossypium hirsutum* L. under high temperature stress. *Agronomy, 12*(6), 1–19. https://doi.org/10.3390/agronomy12061310