



Deciphering Trait Interrelationships in Sunflower (*Helianthus annuus* L.): A Multivariate Approach to Optimise Seed Yield and Oil Quality

Yukta Sharma, Vineeta Kaila

► To cite this version:

Yukta Sharma, Vineeta Kaila. Deciphering Trait Interrelationships in Sunflower (*Helianthus annuus* L.): A Multivariate Approach to Optimise Seed Yield and Oil Quality. 2025, pp.17 - 29. <10.9734/jsrr/2025/v31i83346>. <hal-05212304>

HAL Id: hal-05212304

<https://hal.science/hal-05212304v1>

Submitted on 18 Aug 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License



Deciphering Trait Interrelationships in Sunflower (*Helianthus annuus* L.): A Multivariate Approach to Optimise Seed Yield and Oil Quality

Yukta Sharma ^{a*} and Vineeta Kaila ^a

^a Oilseed Section, Department of Plant and Genetics, Punjab Agricultural University, Ludhiana, Punjab, 141004, India.

Authors' contributions

This work was carried out in collaboration between both authors. Authors YS and VK planned and performed the work, accomplished literature searches, analyzed data, and wrote the manuscript. Both authors read and approved the final manuscript.

Article Information

DOI: <https://doi.org/10.9734/jsrr/2025/v31i83346>

Open Peer Review History:

This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: <https://pr.sdiarticle5.com/review-history/140661>

Original Research Article

Received: 20/05/2025

Published: 28/07/2025

ABSTRACT

To investigate the complex relationships between seed yield and oil content, with their related agronomic traits in sunflower, this correlation analysis study was done at Punjab Agricultural University (30.9010° N and 75.85.73° E) using 113 genotypes, which were grown in spring 2022 in a randomised block design with three replications. Eighteen quantitative traits were analysed, revealing key correlations, like seed yield per plant (economic yield) positively associated with structural traits like plant height, head diameter, stem girth, number of leaves per plant, days to maturity, biological yield per plant, number of seeds per head, hundred seed weight and seed volume weight, while oil content was positively correlated with seed volume weight and inversely correlated with hundred seed weight, indicating a trade-off between yield and oil content. Additionally, oleic acid content demonstrated strong positive correlations with palmitic and linoleic

*Corresponding author: E-mail: yukta9872@gmail.com;

Cite as: Sharma, Yukta, and Vineeta Kaila. 2025. "Deciphering Trait Interrelationships in Sunflower (*Helianthus Annuus* L.): A Multivariate Approach to Optimise Seed Yield and Oil Quality". *Journal of Scientific Research and Reports* 31 (8):17-29. <https://doi.org/10.9734/jsrr/2025/v31i83346>.

acid content. These insights provide direction for breeding programs aiming to optimise sunflower yield and quality.

Furthermore, this research applied principal component analysis (PCA) to elucidate the genetic and phenotypic interrelations influencing oil content and fatty acid composition, the results of which identified five principal components, with the first three accounting for 93% of the total variation, out of which the first two components captured about 77% of the variation. The scatter plot and biplot analysis revealed that oleic acid and linoleic acids dominate the first component with a strong inverse correlation, while the second component is significantly influenced by the positive correlations of palmitic and stearic acids, thereby explaining 77% of variations in oil parameters, and offering actionable insights for enhancing sunflower breeding programs targeting oil quality optimisation.

Keywords: *Sunflower; correlation analysis; principal component analysis; yield; oil parameters; scatter-plot; biplot.*

1. INTRODUCTION

Sunflower (*Helianthus annuus* L.) is an oilseed crop belonging to the family Asteraceae, with a significant contribution to edible oil production and biofuel industries. Therefore, enhancing seed yield and oil content is a critical goal for breeders, particularly in the face of rising, here, correlation analysis serves as a pivotal tool for breeders, facilitating the comprehension of relationships between diverse traits inherent in sunflower and guiding breeders to gain invaluable insights into the associations among traits such as yield, disease resistance, flowering patterns, and seed quality through meticulous correlation scrutiny as also favoured by several previous authors like Manivannan et al., (2005), Hladni et al., (2016), Nasim et al., (2016) and Baraiya et al., (2018). This understanding guides the selection of optimal parental lines by emphasising combinations that exhibit favourable correlations, thereby enhancing the probability of yielding progeny with desired traits. Moreover, correlation analysis expedites the breeding process by focusing on traits strongly correlated with improved sunflower performance, leading to the development of superior varieties.

The oil accounts for 80% value of this crop and is considered to be of superior quality because it contains about 90% fatty acid, 9% phytosterols and 1% vitamin E (alpha-tocopherol), and oleic and linoleic (typically 90% unsaturated fatty acids) are the primary fatty acids present in sunflower oil along with 10% saturated fatty acids like palmitic and stearic acid (Škorić et al., 2015). This is primarily used in the form of salad dressings, cooking oil or in the preparation of margarine and non-dairy creamers, bakery applications, and spray coating oils for cereal,

crackers and dried fruit, thereby it has huge market demand (Hu et al., 2010 and Hassan et al., 2013). However, the complexity of sunflower traits, influenced by genetic and environmental factors, poses challenges in breeding programs aimed at improving the oil content and oil quality of this crop. Therefore, principal component analysis (PCA), a multivariate technique used previously by authors like Hilli and Immadi (2025) in sunflower, to simplify the complexity by reducing the dimensionality of the data and highlights the most influential traits, has been used in this experiment to understand which fatty acids significantly contribute to overall oil content and how different acids interrelate.

2. MATERIALS AND METHODS

2.1 Plant Material Used

In this experiment, a diverse set of 113 sunflower genotypes was used, comprising seven inbred lines, thirteen tester lines, ninety-one hybrids crossed in Line $\times$ Tester fashion, and two commercial checks. These materials were selected to represent a wide range of genetic variability for key agronomic and oil quality traits, enabling robust analysis of trait correlation and principal component. The details of the genotype used are in Table 1.

2.2 Experimental Site

The field evaluation was conducted during Spring 2022 at the sunflower breeding area of the Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana. The site is located at 30.9010° N latitude and 75.85.73° E longitude. The experiment followed a randomised complete block (RCB) design with three replications to ensure the reliability of data and reduce environmental error.

Table 1. Genotypes used in multivariate study

Type	Lines/ Entries
Inbreed lines	CMS 67A, CMS 73A, CMS 82A, CMS 84A, CMS 86A, CMS 88A and CMS 103A
Tester lines	HOHAL-17-1, HOAL-2-P ₃ , HOHAL-11-4, HOAL-30-2, HOAL-21-2, HOHAL-75-4, HOHAL-85-1, HOHAL-28-1, HOHAL-21-1, HOHAL-70-2, HOAL-41-2, HOHAL-85-3 and HOAL-17-2
Hybrids	91 hybrids derived from Line × Tester crosses
Commercial checks	PSH 1962 and PSH 2080

Table 2. Traits recorded in this study

Trait category	Lines/ Entries
Phenological trait	days to flowering initiation, days to 50% flowering, days to maturity
Morphological traits	plant height, plant diameter, stem girth, number of leaves per plant, biological yield per plant, seed yield per plant (economic yield), number of seeds per head, hundred seed weight, seed volume weight and harvest index
Yield-related traits	
Biochemical traits	Oil content, stearic acid (C18:0), palmitic acid (C16:0), linoleic acid (C18:2) and oleic acid (C18:1)

2.3 Data Recorded

The eighteen traits assessed in this study were broadly categorised into phenological, morphological, yield-related and biochemical characteristics, and are presented in Table 2. Observations were made on five randomly selected plants from each genotype in each replication and averaged for phenological and morphological characters, while observations for yield-related characters were made as a population and biochemical traits were analysed using Nuclear Magnetic Resonance Spectrometry (NMR) for oil content and Gas Liquid Chromatography (GLC) for fatty acid composition.

2.4 Statistical Analysis

The correlation coefficient analysis was conducted to evaluate the degree and direction of linear relationship among eighteen agronomic and biochemical characters. This analysis was carried out using R software and guided by the standard statistical procedure as applied in plant breeding Pearson (1901), Al-Jibouri et al., (1958) and Singh and Chaudhary (1977), and its significance was tested using a t-test (Panse and Sukhatme, 1985).

The principal component analysis (PCA) was used for oil quality parameters for germplasm characterisation and categorising them according to their traits, by calculating the eigenvectors and eigenvalues of the covariance matrix. It was performed using R software, and results were

visualised using biplot via 'FactoMineR' (Factor analysis and data mining with R) package with the command 'biplot' (Husson et al., 2014 and R Core Team, 2013). This analysis enabled dimension reduction while retaining maximum variation and facilitated the identification of trait combinations contributing to genetic variation.

3. RESULTS AND DISCUSSION

3.1 Correlation Analysis of Agronomic and Biochemical Traits

The correlation analysis of thirteen agronomical traits and oil content is depicted in Table 3. Understanding trait correlation is vital in crop improvement as positively correlated traits may be simultaneously enhanced through selection, while negative associations require balanced breeding strategies.

The days to flowering initiation (DFI) exhibited a highly significant and positive correlation with days to fifty per cent flowering (DTF) (0.709***), suggesting synchronous phenological development, aligning with the previous findings of Rao (1987) and Tariq et al., (1992). However, DFI showed significant negative association with plant height (-0.237***), head diameter (-0.184***), number of leaves per plant (-0.196***), and hundred seed weight (-0.15**), suggesting that earlier flowering genotypes tend to have smaller plant and seed structures, which is in opposition to the findings of Reddy and Kumar (1996).

Table 3. Correlation coefficient for various agronomic traits and oil content

Traits of variation	DFI	DTF	HGT	HDIA	STG	NLPP	DTM	BYLD	SYP	NSH	HSW	VWT	HI	OC
DFI	1	0.709***	-0.237***	-0.184***	-0.025	-0.196***	0.041	0.014	0.001	0.052	-0.15**	-0.021	-0.012	-0.009
DTF	0.709***	1	-0.144**	-0.144**	-0.154**	-0.161**	0.227***	-0.039	-0.018	0.054	-0.164**	0.134*	-0.076	-0.022
HGT	-0.237***	-0.144**	1	0.788***	0.476***	0.51***	0.39***	0.473***	0.505***	0.448***	0.248***	0.273***	-0.209***	0.094
HDIA	-0.184***	-0.144**	0.788***	1	0.543***	0.467***	0.416***	0.57***	0.607***	0.507***	0.366***	0.162**	-0.225***	-0.078
STG	-0.025	-0.154**	0.476***	0.543***	1	0.185***	0.165**	0.526***	0.435***	0.385***	0.145**	0.031	-0.215***	-0.026
NLPP	-0.196***	-0.161**	0.51***	0.467***	0.185***	1	0.2***	0.326***	0.297***	0.233***	0.306***	0.145**	-0.175**	-0.042
DTM	0.041	0.227***	0.39***	0.416***	0.165**	0.2***	1	0.313***	0.239***	0.183***	0.257***	0.29***	-0.29***	-0.077
BYLD	0.014	-0.039	0.473***	0.57***	0.526***	0.326***	0.313***	1	0.661***	0.584***	0.237***	0.158**	-0.49***	-0.057
SYP	0.001	-0.018	0.505***	0.607***	0.435***	0.297***	0.239***	0.661***	1	0.922***	0.301***	0.156**	0.094	0.012
NSH	0.052	0.054	0.448***	0.507***	0.385***	0.233***	0.183***	0.584***	0.922***	1	-0.026	0.163**	0.137*	0.042
HSW	-0.15**	-0.164**	0.248***	0.366***	0.145**	0.306***	0.257***	0.237***	0.301***	-0.026	1	0.022	-0.074	-0.149**
VWT	-0.021	0.134*	0.273***	0.162**	0.031	0.145**	0.29***	0.158**	0.156**	0.163**	0.022	1	-0.082	0.112*
HI	-0.012	-0.076	-0.209***	-0.225***	-0.215***	-0.175**	-0.29***	-0.49***	0.094	0.137*	-0.074	-0.082	1	0.074
OC	-0.009	-0.022	0.094	-0.078	-0.026	-0.042	-0.077	-0.057	0.012	0.042	-0.149**	0.112*	0.074	1

*, ** and *** means significant at 1%, 5% and 10% level of significance respectively.

where, DFI = days to flowering initiation (in days), DTF = days to fifty percent flowering (in days), HGT = plant height (in cm), HDIA = head diameter (in cm), STG = stem girth (in cm), NLPP = number of leaves per plant, DTM = days to maturity (in days), BYLD = biological yield per plant (in grams), SYP = seed yield per plant or economic yield per plant (in grams), NSH = number of seeds per head, HSW = hundred seed weight (in grams), VWT = seed volume weight (in grams /100ml), HI = harvest index (in %), OC = oil content (in %)

Table 4. Correlation coefficient among oil content and various fatty acid parameters

Traits of Variation	Oil Content (%)	Palmitic Acid (%)	Stearic Acid (%)	Oleic Acid (%)	Linoleic Acid (%)
Oil content (%)	1	0.03	0.01	0.06	0.05
Palmitic acid (%)	0.03	1	0.12*	0.4***	0.34***
Stearic acid (%)	0.01	0.12*	1	0.14**	0.03
Oleic acid (%)	0.06	0.4***	0.14**	1	0.99***
Linoleic acid (%)	0.05	0.34***	0.03	0.99***	1

*, ** and *** means significant at 1%, 5% and 10% level of significance respectively.

Similarly, days to fifty per cent flowering (DFI) showed a strong positive correlation with days to maturity (0.227***), again stabilizing the synchronization of phenological development of this crop, and also shows positive association with the seed volume weight (0.134*), which shows that delayed flowering might enhance the yield potential, in accordance with the findings of Chikkadevaiah et al., (2002) and Kaya et al., (2007). But still, this character shows a negative association with key structural traits like plant height (-0.144**), head diameter (-0.144***), stem girth (-0.154**) number of leaves per plant (-0.161**), hundred seed weight (-0.164**), reinforcing that early flowering genotypes have smaller plant structure and seed size, similar to the findings of Habib et al., (2007).

Plant height (HGT) was further a key determinant of yield-related traits, showing a positive correlation with the head diameter (0.788***), stem girth (0.476***), number of leaves per plant (0.51***), days to maturity (0.39***), biological yield per plant (0.473***), seed yield per plant (0.505***), number of seeds per head (0.448***), hundred seed weight (0.248***) and seed volume weight (0.243***). These results align with the reports of Mogali and Virupakshappa (1994), Khan et al., (2007), Ilahi et al., (2009), Patil (2011), Rao (2013), Kumari et al., (2012), Iqbal et al., (2013) and Sincik and Goksoy (2014) who also identified that positive relationships of plant height with these traits indicate that taller plants with robust structural features contribute directly to higher yield, however, it did show a strong negative correlation with harvest index (-0.012***), due to increasing in biological yield with increase in plant height. However, this character had no significant correlation with oil content (0.094) as opposed to the study of Arshad et al., (2019).

Similarly, head diameter (HDIA) had a positive influence on plant height (0.788**), stem girth (0.543***), number of leaves per plant (0.467**), days to maturity (0.416***), biological yield per plant (0.57***), seed yield per plant (0.607***), number of seeds per head (0.507***), hundred seed weight (0.366***) and seed volume weight (0.162***), but negative influence on days to flowering initiation (-0.184***), days to fifty per cent flowering (-0.144**) and harvest index (-0.225***), reinforcing its role as a critical yield parameter, similar to the reports of Doddamani et al., (1997), Lal et al., (1997) Amorim et al., (2008), Machikowa and Saetang (2008),

Darvishzadeh et al., (2011) and Sujatha and Nadaf (2013).

Further, stem girth (STG) was another vital structural trait showing a significant positive association with yield parameters like biological yield per plant (0.526***), seed yield per plant (0.435***), number of seeds per head (0.385***) and hundred seed weight (0.145**), in accordance to the findings of Behradfar et al., (2009) and Jockovic et al., (2015), reflecting influence on overall plant vigor and productivity, but it has negative correlation with harvest index (0.215***).

Number of leaves on the main stem (NLPP) correlated highly significantly and positively with traits such as plant height (0.51***), head diameter (0.467***), stem girth (0.185***), days to maturity (0.2***), biological yield per plant (0.326***), seed yield per plant (0.297***), number of seeds per head (0.233***), hundred seed weight (0.306***) and seed volume weight (0.145**), and negatively correlated with harvest index (-0.175**), reinforcing the conclusions of Teklewold et al., (2000) and Vidhyavathi et al., (2005), and highlighting it's role in photosynthetic efficiency and assimilate supply, thereby it's role in influencing phenological, morphological and yield-related characters.

On the other hand, days to maturity (DTM) exhibited strong positive correlations with biological yield per plant (0.313***), seed yield per plant (0.661***), number of seeds per head (0.584***), hundred seed weight (0.237***) and seed volume weight (0.29***), suggesting that genotypes with extended growth periods can achieve better yields, as supported by Arshad et al., (2007), who linked extended maturity to enhanced seed yield.

Biological Yield per plant (BYLD), a significant yield determinant, showed robust association with seed yield per plant (0.661***), number of seeds per head (0.584***), hundred seed weight (0.237***) and seed volume weight (0.158**), reinforcing the direct impact of biomass accumulation on reproductive output. While, Seed yield per plant (economic yield) (SYP) was positively correlated with plant height (0.505***), head diameter (0.607***), stem girth (0.435***), biological yield per plant (0.661***), number of seeds per head (0.922***), hundred seed weight (0.301***) and seed volume weight (0.156**), indicating that these traits contribute significantly to economic yield and so improvement in

these traits will lead direct improvement in the complex character like seed yield. A similar trend was also followed by the trait number of seeds per head (NSH). These associations align with studies by Yasin and Singh (2010), and Dan et al., (2012), emphasising the contribution of morphological and yield components to productivity.

However, traits like hundred seed weight (HSW), although positively correlated with the most structural and yield parameters, had a significant negative association with oil content (-0.149^{**}), as also reported by Gjorgjieva et al., (2015), suggesting a potential trade-off between seed mass and oil concentration. While seed volume weight (VWT), an indicator of seed density, was positively associated with all the yield parameters and oil content (0.112^{*}), indicating that denser seeds may contribute to better oil content.

Oil content (OC) showed generally weak correlation with the phenological and structural traits, but exhibited a moderate positive correlation with seed volume weight (0.112^{*}) and significant negatively associated with head seed weight (-0.149^{**}), highlighting the inherent trade-offs between seed size and oil yield, as presented in a similar study by Arshad et al., (2010) and Singh et al., (2018), but opposed to the findings of Rani et al., (2016).

The correlation among oil content and fatty acid composition traits is presented in Table 4. Notably, Oil content exhibited negligible correlations with individual fatty acids, including palmitic acid (0.03), stearic acid (0.01), oleic acid (0.06) and linoleic acid (0.05), suggesting that total oil concentration is relatively independent of its constituent fatty acid proportions.

Palmitic acid showed a significant positive correlation with stearic acid (0.12^{*}), oleic acid (0.4^{***}), and linoleic acid (0.34^{***}), indicating a biosynthetic association among saturated and unsaturated fatty acids. Similarly, stearic acid was also significantly positively correlated with palmitic acid (0.12^{*}), oleic acid (0.14^{**}) and weakly with linoleic acid (0.03). These relationships reflect similar trends reported by Rather et al., (1998) and Tilak et al., (2016).

The most profound and biologically significant correlation was observed between oleic acid and linoleic acid (0.99^{***}), indicating a tight relationship between these two key unsaturated

fatty acids. This finding aligns with studies by Hladni et al., (2011) and Kholghi et al., (2011), who emphasised that the balance between oleic acid and linoleic acid is a critical factor in determining sunflower oil quality, as high oleic acid content typically corresponds to low linoleic acid levels and vice versa. These findings provide valuable insights into breeding strategies aimed at optimising both oil content and fatty acid profiles in sunflower genotypes.

3.2 Principal Component Analysis of Oil Content and Other Fatty Acids

The principal component analysis (PCA) was performed on five key oil quality traits – oil content, palmitic acid, stearic acid, oleic acid and linoleic acid, to identify variation among 113 sunflower genotypes. As shown in Table 5, the first two principal components (PCs) had eigen value greater than 1 and collectively explained 76.69% of the total variability, while first three PCs capture about 93.23% of the variability, but has an eigen value of less than one, therefore, only first two of these components were retained for meaningful interpretation. The high cumulative variance indicated in the first two PCs indicates that PCA effectively reduced dimensionality while preserving essential trait variability, in agreement with the findings of Pérez-Vich et al., (1998) and Manin et al., (2021).

The first principal component (Dim.1) captures the highest variance (45.802%) and was heavily influenced by a strong positive loading from oleic acid (0.975), and a strong negative association from linoleic acid (-0.979), along with a moderate positive contribution from stearic acid (0.534). This indicates that Dim.1 primarily represents the trade-off axis between oleic and linoleic acids, which are biochemically antagonistic and critical determinants of sunflower oil quality. These findings are consistent with the work of the Acuña and Natera (2007) and Ullah (2018), who also identified that oleic acid and linoleic acid are key determinants of first principal component, however, it contradicts the results of Ahmadian et al., (2019), who found that palmitic and stearic acid were primary contributors to the first principal component, while oleic and linoleic acid dominated the second component. Moreover, a strong negative correlation of linoleic acid with saturated acids like palmitic, stearic and arachidic acid was also explained by Hosni et al., (2022) through PCA.

The second principal component (Dim.2) explained 30.89% of the variation, and was characterised by the high positive association for palmitic acid (0.615) and stearic acid (0.734), and a strong negative correlation for oil content (-0.782). When compared with Dim.1, the Dim.2 represents an entirely different set of combinations of variables of being positively or negatively correlated with the trait, therefore,

capturing another significant dimension of the data variation, and suggesting that oil content is inversely related to saturated fatty acids, thereby representing an additional layer of biochemical differentiation among genotypes. These findings align with the work of Chernova et al., (2021), who identified the role of specific saturated fatty acids in shaping genetic variation in oil content.

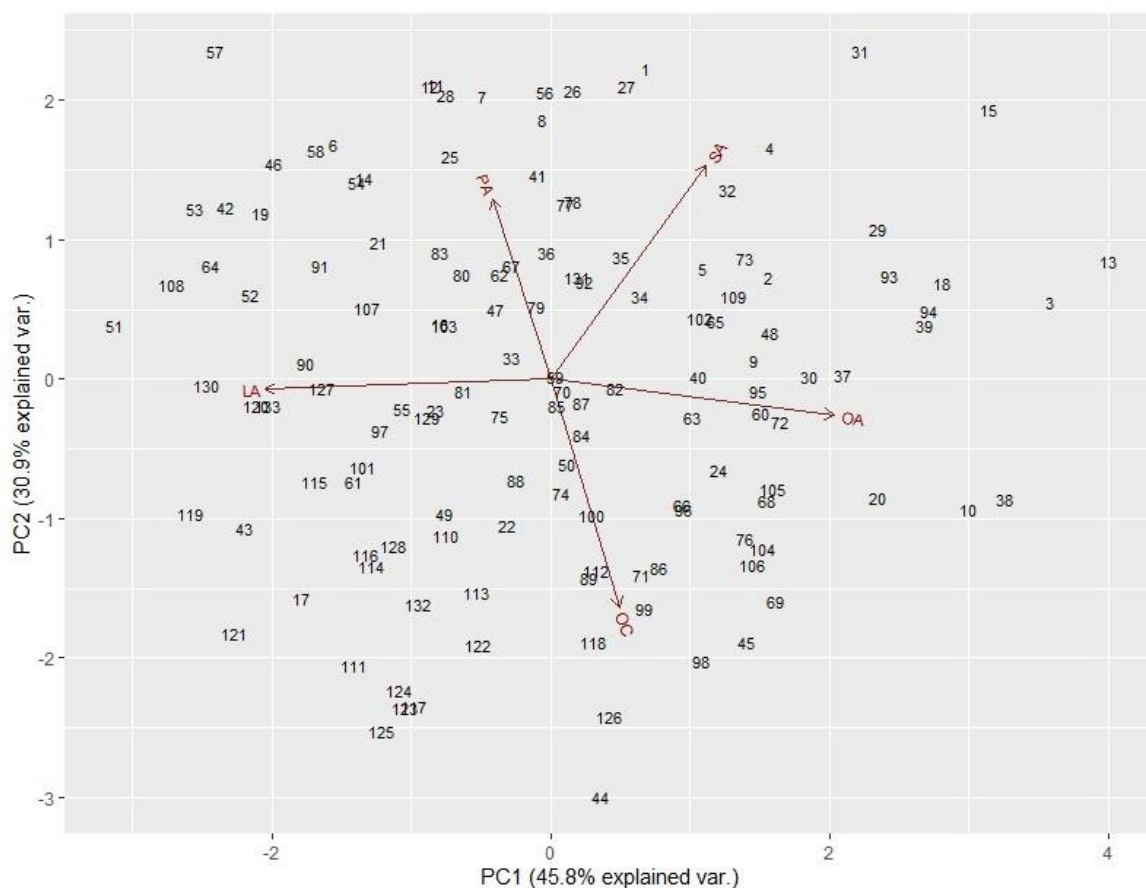


Fig. 1. Scatter-Plot between first principal component (Dim.1) and second principal component (Dim.2)

where genotypes 1 to 7 are female inbred lines, in order stated in Table 1; genotypes 8 to 20 are male tester lines, in order stated in Table 1, genotypes 21 to 111 are the hybrids crossed in Line $\times$ Tester fashion, starting from 67A $\times$ HOHAL-17-1, 67A $\times$ HOHAL-2-P₃to..... 103A $\times$ HOHAL-17-2; and genotype 112 = PSH 1962, genotype 113 = PSH 2080. PC1 = first principal component, PC 2 = second principal component, OC = oil content, PA = palmitic acid, SA = stearic acid, OA = oleic acid and LA = linoleic acid.

Table 5. EIGEN values and variance of 5 principal components

Principal components	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
Oil content	0.236	-0.782	0.490	0.305	-0.003
Palmitic acid	-0.199	0.615	0.753	-0.119	0.011
Stearic acid	0.534	0.734	-0.111	0.405	0.007
Oleic acid	0.975	-0.122	0.013	-0.163	0.091
Linoleic acid	-0.979	-0.034	-0.083	0.156	0.091
EIGEN value	2.290	1.545	0.827	0.322	0.017
variance (%)	45.802	30.890	16.541	6.431	0.335
cumulative variance (%)	45.802	76.693	93.234	99.665	100.000

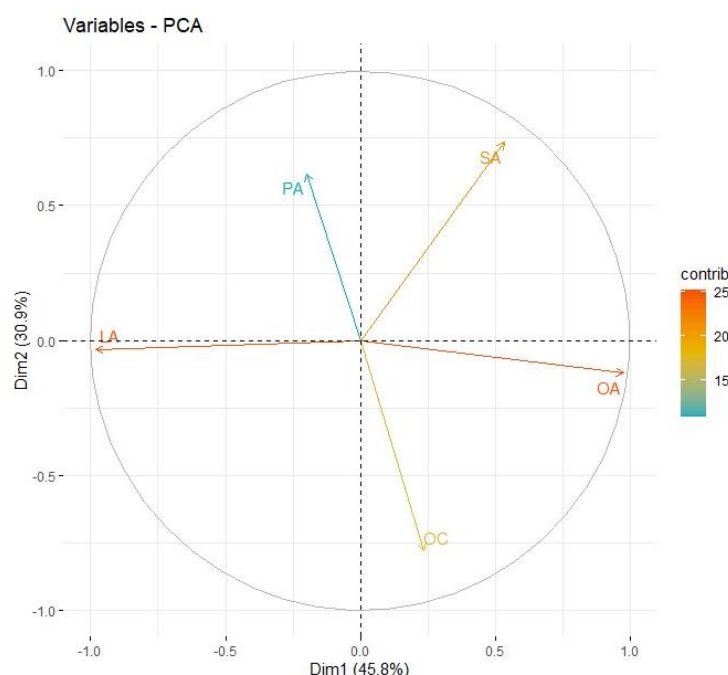


Fig. 2. Biplot between first principal component (dim.1) and second principal component (dim.2)

where Dim1 = first principal component, Dim2 = second principal component, OC = oil content, PA = palmitic acid, SA = stearic acid, OA = oleic acid and LA = linoleic acid

Together, the first two PCs capture about 77% of the variance, while the third, fourth and fifth PCs seem to have very little variability, confirming their suitability for dimension reduction and trait interpretation in sunflower genotypes. Therefore, the result of this PCA highlighted the importance of balancing unsaturated fatty acids for oil quality improvement.

To further interpret the trait variation among sunflower genotypes, the scatter plot analysis and bi-plot analysis were done for the first two principal components (Dim.1 and Dim.2), which together accounted for 76.69% of the variance. In the scatter plot (Fig. 1), each point represents a genotype positioned according to its score on Dim.1 and Dim.2. Genotypes located closer to the origin have a low score for these components., thereby they exhibit average closer to the original variables, while those which are positioned farther from the origin, demonstrate the extreme values of trait expression. Genotypes on the right side (positive Dim.1) area are associated with higher oleic content and lower linoleic acid levels, whereas those in the left side of the plot (negative Dim.1) display elevated linoleic acid levels but reduced oleic acid content. Similarly, genotypes on the upper end of the plot (positive Dim.2), tend to

accumulate higher levels of palmitic and stearic acids, while those in the lower half (negative Dim.2) are characterized by lower oil content and varied levels of other fatty acids. Similarly, Zimmerman and Fick (1973) also explained that a decrease in oleic acid content occurs with increasing linoleic and palmitic acid content, through this method.

The biplot analysis (Fig.2) provides a visual contribution of the traits to the first two principal components, through direction (arrow) and its strength (color). The vector for oleic acid (OA) is oriented strongly toward the right side, indicating a high positive impact on the Dim.1, while linoleic acid (LA) points leftward, reflecting a strong negative association with this component. Thereby, the antagonistic orientation reinforces the well-established biochemically inverse relationship of OA and LA in sunflower oil biosynthesis. On the other hand, Oil content displays a moderate negative vector along Dim.2, and suggests an inverse association with the saturated fatty acids (PA and SA). Both palmitic acid (PA) and stearic acid (SA) have a positive correlation with the second principal component (Dim.2), and their proximity and parallel orientation imply a shared pattern of accumulation.

Overall, the spatial distribution of the genotypes and trait vectors in the biplot indicates the multivariate structure of oil composition traits. It helps to identify the genotypes with a favorable high oleic acid and low linoleic acid profile, along with desirable saturated acid combinations, thus offering critical insights for strategic selection in sunflower breeding programs aimed at quality improvement.

4. CONCLUSION

This study elucidates the interrelationship among phenological, morphological, yield, and quality traits (oil content and fatty acid composition) in sunflower by employing multivariate analysis in 113 genotypes. The correlation analysis identified the importance of structural traits like plant height, head diameter and biological yield to be reliable indicators of advanced seed yield while addressing the trade-off with the oil quality. Moreover, the correlation analysis of fatty acid composition further underscored the importance of the critical balance between oleic acid and linoleic acid.

The principal component analysis (PCA) identified the first two components to be highly informative, capturing the eigenvalues of more than one and covering the total variance of about 76.7%. It clearly highlighted the inverse relationship between oleic acid and linoleic acid. This integration of PCA and correlation analysis provides a robust framework for characterising genetic diversity and guiding targeted sunflower breeding for yield and quality traits.

DISCLAIMER (ARTIFICIAL INTELLIGENCE)

Author(s) hereby declare that no generative AI technologies such as large language models (ChatGPT, Copilot, etc.) And text-to-image generators have been used during the writing or editing of this manuscript, except for Grammarly: a free AI writing assistance tool was downloaded and installed in the PC for grammar checks.

ACKNOWLEDGEMENT

We gratefully acknowledge the great support of Dr Pankaj Sharma (Plant Pathologist (Sunflower), Oilseed section, Dept. of PBG, PAU) in the successful execution of this project and Dr Sanjula Sharma (Biochemist (Oilseed), Oilseed section, Dept. of PBG, PAU) for providing the laboratory facilities to perform biochemical analysis.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

- Acuña, A., Malavé, R., & Natera, J. R. M. (2007). Lipid composition of sunflower seeds (*Helianthus annuus* L.) using multivariate analysis. *UDO Agricultural Review*, 6(1), 27–32.
<http://hdl.handle.net/1807/45363>
- Ahmadian, S., Enferadi, T., & Alemzadeh, A. (2019). Assessment of genetic diversity of cultivated sunflower in terms of oil content, fatty acid compositions and seed traits. *Helia*, 42, 229–246.
<https://doi.org/10.1515/helia-2019-0009>
- Al-Jibouri, H. A., Miller, P. A., & Robinson, H. F. (1958). Genotypic and environmental variances and covariances in an upland cotton cross of interspecific origin. *Journal of Agronomy*, 50(10), 633–636.
<https://doi.org/10.2134/agronj1958.00021962005000100020x>
- Amorim, E. P., Ramos, N. P., Ungaro, M. R. G., & Kiihl, A. (2008). Correlation and path analysis in sunflower. *Bragantia, Campions*, 67(2), 307–316.
<https://doi.org/10.1590/S0006-87052008000200006>
- Arshad, M., Jan, S., Awan, S., Azam, S., Khalid, S., & Khan, M. A. (2019). Investigation of genetics divergence in newly developed local sunflower (*Helianthus annuus* L.) hybrids. *Pakistan Journal of Agriculture Research*, 32(1), 33–40.
<http://dx.doi.org/10.17582/journal.pjar/2019/32.1.33.40>
- Arshad, M., Kashif, M., & Ayub, Khan, M. (2007). Genetic divergence and path coefficient analysis for seed yield traits in sunflower (*Helianthus annuus* L.) hybrids. *Pakistan Journal of Botany*, 39(6), 2009–2015.
[http://www.pakbs.org/pjbot/PDFs/39\(6\)/PJ B39\(6\)2009.pdf](http://www.pakbs.org/pjbot/PDFs/39(6)/PJ B39(6)2009.pdf)
- Arshad, M., Khan, A., & Jadoon, S. A. (2010). Factor analysis in sunflower (*Helianthus annuus* L.) to investigate desirable hybrids. *Pakistan Journal of Botany*, 42(6), 4393–4402.
https://www.researchgate.net/publication/265334995_Factor_analysis_in_sunflower_Helianthus_annuus_L_to_investigate_desirable_hybrids

- Baraiya, V. K., Jagtap, P. K., Sangani, J. L., & Malviya, A. V. (2018). Correlation and path analysis in sunflower (*Helianthus annuus* L.). *Journal of Pharmacognosy and Phytochemistry*, 7(5), 2730–2732.
https://www.researchgate.net/profile/Vipul-Baraiya/publication/328904209_Correlation_and_path_analysis_in_sunflower_Helianthus_annuus_L/links/5beabf43a6fdcc3a8dd21b58/Correlation-and-path-analysis-in-sunflower-Helianthus-annuus-L.pdf
- Behradfar, A., Gorttapeh, A. H., Zardashty, M. R., & Talat, F. (2009). Evaluation correlated traits for seed and oil yield in sunflower (*Helianthus annuus* L.) through path analysis in under condition relay cropping. *MEDWELL Research Journal of Biological Sciences*, 4(1), 82–85.
<https://www.researchgate.net/publication/259383353>
- Chernova, A. I., Gubaev, R. F., Singh, A., Sherbina, K., Goryunova, S. V., Martynova, E. U., Goryunov, D. V., Boldyrev, S. V., et al. (2021). Genotyping and lipid profiling of 601 cultivated sunflower lines reveals novel genetic determinants of oil fatty acid content. *BMC Genomics*, 22, 505.
<https://doi.org/10.1186/s12864-021-07768-y>
- Chikkadevaiah, H. C., Sujatha, H. L., & Nandini, R. C. (2002). Correlation and path analysis in sunflower. *Helia*, 25(37), 109–118.
<https://doi.org/10.2298/hel0237109c>
- Dan, G., Manivannan, N., & Vindhiyavarman, P. (2012). Genetic divergence of inbred lines in sunflower (*Helianthus annuus* L.). *Electronic Journal of Plant Breeding*, 3(1), 652–656.
<https://www.ejplantbreeding.org/index.php/EJPB/article/view/523>
- Darvishzadeh, R., Maleki, H. H., & Sarrafi, A. (2011). Path analysis of the relationships between yield and some related traits in diallel population of sunflower (*Helianthus annuus* L.) under well-watered and water-stressed conditions. *Australian Journal of Crop Science*, 5(6), 674–680.
<https://search.informit.org/doi/abs/10.3316/informit.281912150560452>
- Doddamani, I. K., Patil, S. A., & Ravikumar, R. L. (1997). Relationship of autogamy and self-fertility with seed yield and yield components in sunflower (*Helianthus annuus* L.). *Helia*, 20(26), 95–102.
<https://agris.fao.org/agris-search/search.do?recordID=YU1998000021>
- Gjorgjieva, B., Karov, I., Mitrev, S., Ruzdik, N. M., Kostadinovska, E., & Kovacevik, B. (2015). Correlation and path analysis in sunflower (*Helianthus annuus* L.). *Helia*, 38(63), 201–210.
<https://doi.org/10.1515/helia-2015-0008>
- Habib, H., Sadaqat, Mehdi, S., Ashfaq, Anjum, M., & Ahmad, R. (2007). Genetic association and path analysis for oil yield in sunflower (*Helianthus annuus* L.). *International Journal of Agriculture and Biology*, 9(2), 359–361.
<http://www.fspublishers.org>
- Hassan, S. M. F., Iqbal, M. S., Rabbani, G., Naeem-ud-Din, N., Shabbir, G., Riaz, M., & Noorka, I. R. (2013). Correlation and path analysis for yield and yield components in sunflower (*Helianthus annuus* L.). *African Journal of Biotechnology*, 12(16), 1968–1971.
<https://doi.org/10.5897/AJB12.817>
- Hilli, H. J., & Immadi, S. (2025). Principal component analysis and diversity studies in sunflower lines (*Helianthus annuus*). *Indian Journal of Agricultural Research*, 59(4), 560.
<https://doi.org/10.18805/IJARE.A-5996>
- Hladni, N., Jovic, S., Mijic, A., Miklic, V., & Miladinovic, D. (2016). Correlation and path analysis of yield and yield components of confectionary sunflower. *Genetika*, 48(3), 827–835.
<https://doi.org/10.2298/GENSR1603827H>
- Hladni, N., Jovic, S., Miklic, V., Saftic-Pankovic, D., & Kraljevic-Balalic, M. (2011). Interdependence of yield and yield components of confectionary sunflower hybrid. *Genetika*, 43(3), 583–594.
<https://scispace.com/pdf/interdependence-of-yield-and-yield-components-of-13fkh1pbqt.pdf>
- Hosni, T., Abbes, Z., Abaza, L., Medimagh, S., Salah, B., & Kharrat, M. (2022). Biochemical characterization of seed oil of Tunisian sunflower (*Helianthus annuus* L.) accessions with special reference to its fatty acid composition and oil content. *Journal of Food Quality*.
<https://doi.org/10.1155/2022/2875072>
- Hu, J., Seiler, G., & Kole, C. (2010). *Genetics, genomics and breeding of sunflower*. Routledge, USA.
<https://doi.org/10.1201/b10192>

- Husson, F., Josse, J., Lê, S., & Mazet, J. (2014). *FactoMineR: Multivariate exploratory data analysis and data mining with R*. R package version 1, 102–123. https://www.researchgate.net/publication/284673888_FactoMineR_Multivariate_Exploratory_Data_Analysis_and_Data_Mining_with_R
- Ilahi, F., Tahir, M. H. N., & Sadaqat, H. A. (2009). Correlation and path co-efficient analysis for achene yield and yield components in sunflower. *Pakistan Journal of Agricultural Science*, 46(1), 20–24. <https://www.cabidigitallibrary.org/doi/full/10.5555/20093272941>
- Iqbal, M., Ijaz, U., Smiullah, I. M., Mahmood, K., Najeebullah, M., Abdullah, N. S., & Sadaqat, H. A. (2013). Genetic divergence and path coefficient analysis for yield related attributes in sunflower (*Helianthus annuus* L.) under less water conditions at productive phase. *Plant Knowledge Journal*, 2(1), 20–23. <https://www.cabidigitallibrary.org/doi/full/10.5555/20133405946>
- Jockovic, M., Jovic, S., Jeromela, A. M., Ciric, M., Canak, P., Miklic, V., & Cvejic, S. (2015). Biomorphological association and path analysis in sunflower (*Helianthus annuus* L.). *Helia*, 38(63), 189–199. <https://doi.org/10.1515/helia-2015-0004>
- Kaya, Y., Evc, G., Durak, S., Pekcan, V., & Gocer, T. (2007). Determining the relationships between yield and yield attributes in sunflower. *Turkish Journal of Agriculture*, 31(4), 237–244. <https://journals.tubitak.gov.tr/cgi/viewcontent.cgi?article=2009andcontext=agriculture>
- Khan, H., Muhammad, S., Shah, R., & Iqbal, N. (2007). Genetic analysis of yield and some yield components in sunflower. *Sarhad Journal of Agriculture*, 23(4), 985–990. <https://journals.tubitak.gov.tr/agriculture/vol31/iss4/3/>
- Kholghi, M., Bernoussi, I., Darvishzadeh, R., & Pirzad, A. (2011). Correlation and path coefficient analysis of seed yield and yield related trait in Iranian confectionery sunflower populations. *African Journal of Biotechnology*, 10(61), 13058–13063. <https://doi.org/10.5897/AJB11.1452>
- Kumari, S., Sheoran, R. K., Singh, A., & Avtar, R. (2012). Genetic variability, heritability and character association in sunflower. *Research on Crops*, 13(2), 566–572. <http://www.croprosearch.org>
- Lal, G. S., Bhadoriya, V. S., & Singh, A. K. (1997). Genetic association and path analysis in elite lines of sunflower. *Crop Research*, 13(3), 6. <https://www.cabidigitallibrary.org/doi/full/10.5555/19981604778>
- Machikowa, T., & Saetang, C. (2008). Correlation and path coefficient analysis on seed yield in sunflower. *Suranaree Journal of Science and Technology*, 13(2), 243–248. <https://www.researchgate.net/publication/242390885>
- Manin, L., Pizzo, J., Rydlewski, A., Santos, P., Galuch, M., Zappiello, C., Santos, O., & Visentainer, J. (2021). Evaluation of the adulteration of edible and cosmetic sunflower oils by GC-FID and ESI-MS. *Research on Crops*, 32(11), 2086–2094. <https://doi.org/10.21577/0103-5053.20210101>
- Manivannan, N., Muralidharan, V., & Subbalakshmi, B. (2005). Correlation analysis in sunflower. *Legume Research*, 28(1), 71–73. <https://arccarticles.s3.amazonaws.com/webArticle/articles/lr281019.pdf>
- Mogali, S., & Virupakshappa, K. (1994). Inter-character association and path coefficient analysis in sunflower (*Helianthus annuus* L.). *Indian Journal of Genetics and Plant Breeding*, 54(4), 366–370. <https://www.isgpb.org/journal/index.php/IJGPB/article/view/2715>
- Nasim, W., Belhouchette, H., Tariq, M., Fahad, S., Hammad, H. M., Mubeen, M., Munis, M. F. H., Chaudhary, H. J., et al. (2016). Correlation studies on nitrogen for sunflower crop across the agroclimatic variability. *Environmental Science and Pollution Research*, 23, 3658–3670. <https://link.springer.com/article/10.1007/s11356-015-5613-1>
- Panase, V. G., & Sukhatme, P. V. (1985). *Statistical methods for agricultural research* (8). ICAR, New Delhi. <https://search.worldcat.org/title/statistical-methods-for-agricultural-workers/oclc/630116948?referer=di&ht=edition>
- Patil, L. C. (2011). Research note: Correlation and path analysis in sunflower populations. *Electronic Journal of Plant Breeding*, 2(3), 442–447. <http://sites.google.com/site/ejplantbreeding>
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin*

- Philosophical Magazine and Journal of Science*, 2(11), 559–572.
<https://pca.narod.ru/pearson1901.pdf>
- Pérez-Vich, B., Velasco, L., & Fernández-Martínez, J. (1998). Determination of seed oil content and fatty acid composition in sunflower through the analysis of intact seeds, husked seeds, meal and oil by near-infrared reflectance spectroscopy. *Journal of the American Oil Chemists' Society*, 75, 547–555.
<https://doi.org/10.1007/S11746-998-0064-1>
- R Core Team. (2013). *R: A language and environment for statistical computing* [Software].
<http://www.r-project.org/>
- Rani, R., Sheoran, R. K., & Chander, S. (2016). Association analysis for yield and component traits in sunflower (*Helianthus annuus* L.). *Journal of Oilseeds Research*, 33(3), 201–204.
<https://doi.org/10.56739/jor.v33i3.137998>
- Rao, N. G. L. (1987). Studies on correlation and path coefficient analysis in sunflower. *Mysore Journal of Agriculture Science*, 21, 94–95.
<https://www.cabidigitallibrary.org/doi/full/10.5555/19881673043>
- Rao, V. T. (2013). Association analysis in sunflower (*Helianthus annuus* L.). *International Journal of Applied Biology and Pharmaceutical Technology*, 4(1), 1–4.
https://www.researchgate.net/publication/349442091_Association_analysis_in_sunflower
- Rather, A. G., Sandha, G. S., Shafiq, A., Wani, & Seikh, F. A. (1998). Correlation and path analysis in sunflower (*Helianthus annuus* L.). *Journal of Oilseed Research*, 15(1), 67–70.
<https://isor.in/wp-content/uploads/2022/10/vol-15-1-1998.pdf>
- Reddy, M. D., & Kumar, K. A. (1996). Performance of sunflower at different times of sowing during post rainy season in North Telangana Zone of Andhra Pradesh. *Journal of Oilseed Research*, 13(2), 313–316.
<https://www.researchgate.net/publication/348482497>
- Sincik, M., & Goksoy, A. T. (2014). Investigation of correlation between traits and path analysis of confectionary sunflower genotypes. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 42(1).
<https://doi.org/10.15835/nbha4219429>
- Singh, R. K., & Chaudhary, B. D. (1977). *Biometrical methods in quantitative genetic analysis*. Kalyani Publishers, New Delhi.
<https://www.cabidigitallibrary.org/doi/full/10.5555/19801689021>
- Singh, V. K., Sheoran, R. K., & Chander, S. (2018). Correlation analysis for seed yield and its component traits in sunflower. *Journal of Pharmacognosy and Phytochemistry*, 7(3), 2299–2301.
<https://www.phytojournal.com/archives/2018/vol7issue3/PartAE/7-3-228-855.pdf>
- Škorić, D., Sakač, Z., & Demurin, Y. (2015). Genetic variability of oil quality components in sunflower as a function of developing hybrids with novel oil. *Journal of ASM, Life Sciences, Chisinau. Moldova*, 2(326), 13–146.
https://ibn.idsi.md/sites/default/files/j_nr_file/BuletinCongres%20Nr%202,%202015%20final2.pdf#page=138
- Sujatha, K., & Nadaf, H. L. (2013). Correlation for yield and yield related trait in mutant and segregating genotypes in sunflower (*Helianthus annuus* L.). *Molecular Plant Breeding*, 4(32), 265–266.
<https://doi.org/10.5376/mpb.2013.04.0032>
- Tariq, M., Drees, G., & Tahir, A. (1992). Genetic variability and correlation studies in sunflower. *Sarhad Journal of Agriculture*, 8(6), 659–663.
<https://agris.fao.org/agris-search/search.do?recordID=PK9600698>
- Teklewold, A. H., Jayaramaiah, & Jagadeesh, B. N. (2000). Correlation and path analysis of physio morphological characters of sunflower (*Helianthus annuus* L.) as related to breeding methods. *Helia*, 23(32), 105–114.
<https://www.cabidigitallibrary.org/doi/full/10.5555/20013034487>
- Tilak, I. S., Kisan, B., Patil, A., Shankergoud, I., Kadirvel, Ghante, V., & Kulkarni, V. (2016). Genotypic and phenotypic correlations and path coefficient analysis for oleic acid content in sunflower. *Karnataka Journal of Agricultural Science*, 29(4), 476–478.
<https://www.researchgate.net/publication/325846189>
- Ullah, Z. (2018). Grouping and selection of 32 single cross sunflower hybrids using principal component analysis. *Eurasian Journal of Agricultural Research*, 2, 4–12.
<https://dergipark.org.tr/en/pub/ejar/issue/39001/449230>
- Vidhyavathi, R., Mahalakshmi, P., Manivannan, N., & Muralidharan, V. (2005). Correlation

- and path analysis in sunflower (*Helianthus annuus* L.). *Agricultural Science Digest*, 25(1), 6–10.
<https://www.indianjournals.com/ijor.aspx?target=ijor:asd&volume=25&issue=1&article=002>
- Yasin, A. B., & Singh, S. (2010). Correlation and path coefficient analyses in sunflower. *Journal of Plant Breeding and Crop Science*, 2(5), 129–133.
<http://www.academicjournals.org/jpbcs>
- Zimmerman, D., & Fick, G. (1973). Fatty acid composition of sunflower (*Helianthus annuus* L.) oil as influenced by seed position. *Journal of the American Oil Chemists' Society*, 50, 273–275.
<https://doi.org/10.1007/BF02641352>

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of the publisher and/or the editor(s). This publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© Copyright (2025): Author(s). The licensee is the journal publisher. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here:
<https://pr.sdiarticle5.com/review-history/140661>