Journal of Experimental Agriculture International

Volume 46, Issue 12, Page 444-464, 2024; Article no.JEAI.127627
ISSN: 2457-0591
(Past name: American Journal of Experimental Agriculture, Past ISSN: 2231-0606)

Characterization of Genetic Diversity in Fatty Acid and Nutritional Profiles Among *Brassica* Species

Achyuta Basak [a], Moumita Chakraborty [a*], Shankar K. Bhujbal [b], Sanjay. J. Jambhulkar [b], Suvendu Kumar Roy [a], Saikat Das [a], Biplab Mitra [c], Rupsanatan Mandal [a], Lakshmi Hijam [a], Manoj Kanti Debnath [d] and Md. Sabir Ahmed Mondol [e]

[a] Department of Genetics & Plant Breeding, Uttar Banga Krishi Viswavidyalaya, Pundibari, Cooch Behar, West Bengal, 736165, India.
[b] Nuclear Agriculture and Biotechnology Division, Bhabha Atomic Research Centre, Mumbai, 400085, India.
[c] Department of Agronomy, Uttar Banga Krishi Viswavidyalaya, Pundibari, Cooch Behar, West Bengal, 736165, India.
[d] Department of Agricultural Statistics, Uttar Banga Krishi Viswavidyalaya, Pundibari, Cooch Behar, West Bengal, 736165, India.
[e] Department of Agricultural Biochemistry, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, Nadia, West Bengal, 741252, India.

Authors' contributions

This work was carried out in collaboration among all authors. Author AB wrote and prepared the original draft of the manuscript, did statistical analysis and performed the laboratory experiment. Author MC conceptualized the study, did data validation, performed the methodology, and did data curation. Author SKB performed laboratory experiments. Author SJJ did true breeding genotype sharing and conceptualized the study. Author SKR conceptualized the study, wrote and prepared the original draft of the manuscript, did statistical analysis and data validation. Authors SD, BM, RM and MKD wrote, reviewed and edited the manuscript. Author BM searched for resources. Author LH searched for resources, performed the methodology and did data curation. Author MKD did statistical analysis. Author MSAM did data curation. All authors read and approved the final manuscript.

Article Information

DOI: https://doi.org/10.9734/jeai/2024/v46i123151

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

Cite as: *Basak, Achyuta, Moumita Chakraborty, Shankar K. Bhujbal, Sanjay. J. Jambhulkar, Suvendu Kumar Roy, Saikat Das, Biplab Mitra, Rupsanatan Mandal, Lakshmi Hijam, Manoj Kanti Debnath, and Md. Sabir Ahmed Mondol. 2024. "Characterization of Genetic Diversity in Fatty Acid and Nutritional Profiles Among Brassica Species". Journal of Experimental Agriculture International 46 (12):444-64. https://doi.org/10.9734/jeai/2024/v46i123151.*

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://www.sdiarticle5.com/review-history/127627

Original Research Article

Received: 11/10/2024
Accepted: 14/12/2024
Published: 20/12/2024

ABSTRACT

The present experiment was carried out to evaluate the nutritional profile of 10 advanced mutant lines of Indian mustard (*Brassica juncea*) along with 3 checks. Observation was recorded for 12 biochemical traits, and the analysis of variance (ANOVA) revealed significant variation among all characters under study. The findings demonstrate higher GCV and PCV in the case of oleic acid and arachidic acid which significantly impact these traits' inflow. Except for the total phenol content (mg GAE/g) and the total soluble protein content, all the other traits show high heritability and a significant genetic advance, indicating traits controlled by additive gene action. In the correlation study, it was found that the oleic acid, linoleic acid, linolenic acid, palmitic acid, eicosenoic acid and stearic acid are significantly negatively correlated with erucic acid. This indicates that improving these characteristics leads to decreased erucic acid content. The path diagram also indicates a high direct positive effect of total soluble protein (%) and stearic acid and a significant high direct negative effect of total phenol content and palmitic acid on total oil content. Three of the twelve PCs had eigenvalues greater than 1.0, explaining 83.36% of the variance. PC I accounted for 37.48% and PC II for 30.42% of the overall variance. Based on hierarchical cluster analysis, 13 Brassica genotypes were divided into 4 clusters. Genotypes TM-313, TM 305-1, TM 306-1, B9 show the most divergence and can be used in future breeding programs.

Keywords: Brassica juncea; genetic parameters; path analysis; cluster analysis; PCA.

1. INTRODUCTION

Identifying germplasm with a rich nutritional profile is crucial for addressing the global issue of malnutrition (Shyam et al., 2021). In recent years, there has been a significant shift in consumer perspectives; they are now increasingly focused not only on their dietary choices but also on the effects of the ingredients they consume on their health. Oil is a crucial element in our diets, serving as the most concentrated source of energy (Sharma et al., 2019). Mustard plants are part of the well-known family *Brassicaceae*, which falls under the order *Brassicales* (formerly referred to as *Capparales*). This family encompasses more than 330 genera and over 3700 species that are found across the globe (Warwick et al., 2006). The genus Brassica is increasingly recognized for its industrial significance, particularly because it includes oil-rich species like *Brassica juncea*, *B. carinata*, *B. rapa*, and *B. napus* (McVetty & Duncan, 2015). Specific nutritional values are crucial factors in choosing the oil for both edible and industrial uses. *B. juncea* stands out among mustard species for its remarkable adaptability as an oilseed crop in India, China, and Pakistan, achieving oil contents of up to 44% (Riffat Tahira et al., n.d.; Sharad Pandey et al., 2013; Stoutjesdijk et al., 2000). Brassica, especially mustard, demonstrates a variety of biological activities such as anti-cancer, anti-oxidation, anti-inflammatory, antibacterial, antiviral, anti-obesity, and anti-depression effects. It also plays a role in the prevention and treatment of diabetes and cataracts, highlighting its significance as a medicinal plant due to its biochemical compounds that are utilized in addressing diseases like cancer, diabetes, and inflammatory immune disorders (Muhammad Haseeb Anwar Tarar et al., 2021; Tian & Deng, 2020). The phenolic compounds found in mustard seeds serve as a natural source of antioxidants and anti-inflammatory agents, enhancing their nutritional value and potential health benefits (Salah et al., 2024).

Mustard is a cost-effective and nourishing food option containing bioactive elements such as glucosinolates and their derivatives, as well as

polyphenols like flavonoids and anthocyanins. It is also rich in dietary fibre, chlorophylls, β-carotene, ascorbic acid, minerals, and various volatile components (Kim, Yong-Taek et al., 2007). Among the nutritional parameters of edible oil, fatty acids such as oleic, linolenic, erucic, palmitic, and linoleic are crucial. The genotypes of Indian mustard (*B. juncea*) has been recorded as having increased levels of erucic acid and glucosinolates in the oil part (Riffat Tahira et al., n.d.). Palmitic and stearic acids are the two main saturated fatty acids, while the main unsaturated fatty acids consist of oleic, linoleic, linolenic, eicosanoic, and erucic. Oils with elevated levels of erucic acid are appropriate for industrial applications but not suitable for human consumption (Nieschlag & Wolff, 1971; Singh et al., n.d.). Consequently, the creation of cultivars that are free of commercial erucic acid and those with elevated erucic acid levels presents significant opportunities for the enhancement of brassica oilseed crops. Additional significant objectives involve enhancing the levels of oleic and linoleic acids while reducing the content of linolenic acid (Rahman et al., 2022). Linoleic acid (LA), an essential omega-6 fatty acid, plays a critical role in growth and development. Deficiencies in linoleic acid can result in compromised growth and symptoms associated with fatty acid deficiency (Connor, 1999). Additionally, linolenic acid, an omega-3 fatty acid, can be derived from LA and has been associated with a reduced risk of cardiovascular disease (Connor, 1999; Shahidi & Finley, 2001). Oleic acid, a monounsaturated fatty acid, has also been linked to a lower risk of cardiovascular disease, particularly when it replaces saturated fats in dietary patterns (Hunter, 2005; Jandacek, 2017). Furthermore, crude oil derived from plants rich in linoleic and linolenic acids is utilized in the tanning industry as well as in the production of margarine and various food products (Shahidi & Finley, 2001).

The present study involved estimating genotypic correlations, path analysis, clustering of genotypes, and conducting PCA to produce significant and valuable insights for the future enhancement program of *Brassica*. Although *Brassica* Sp, specially *Brassica juncea* had a narrow genetic base, we used mutant recombinant lines for our study to analyze the nutritional and fatty acid profiles of 13 distinct *Brassica* genotypes from two different Brassica species.

2. MATERIAL AND METHODS

Description of Experimental Site: The experiment involved the use of 12 *Brassica juncea* and 1 *Brassica rapa* genotypes (table 1) for various estimations. All the chosen genotypes were cultivated at the Instructional farm, Uttar Banga Krishi Viswavidyalaya, Pundibari during 2021-22 and Uttar Dinajpur Krishi Vigyan Kendra farm, Chopra, Uttar Dinajpur, West Bengal, during 2022-23. The research site is located at 26°19´86´´ N, 89°23´53´´ E, at an altitude of 43m above mean sea level (for Pundibari) and 26°21'18" N, 88°16'36" E, at an altitude of 76m above sea level (for Chopra). The Pundibari region experimental site belongs to the terai region with Teesta alluvial soil and the Chopra region experimental site belongs to the terai region with alluvium, mostly sandy to sandy-loam in texture, and porous.

Experimental Materials:

Table 1. List of Genotypes under study

Sl No.	Genotype Name	Pedigree	Denotation
1	TM 301-3	TM277 x TM106	G1
2	TM 303-1	TM106 x TM277	G2
3	TM 305-1	TM276 x TM277	G3
4	TM 306-1	TJD1 x Varuna	G4
5	TM 308-1	TJD-1 x NRCHB101	G5
6	TM 309-2	TJD-1 x PM26	G6
7	TM 310-3	TJD-1 x PM28	G7
8	TM 312-1	PM26 x TJD-1	G8
9	TM 313	PM27 x TJD-1	G9
10	TM 316	RH749 x TJD-1	G10
11	PM 25	SEJ 8 x Pusa Jagannath	G11
12	JD 6	PUSA BOLD x GLOSSY	G12
13	B9	Rapeseed	G13

Cultural Practices: The experiment was carried out in three replications involving 13 genotypes within a randomized complete block design for both the locations. Table 1 provides a detailed enumeration of the genotypes. Three rows of genotypes, each measuring 3 meters, were planted within each plot. The row spacing was consistently set at 30 cm, and the plant spacing was adjusted to 15 cm through careful thinning practices. All necessary interculture procedures for a thriving mustard harvest were carried out to attain a robust and competitive crop yield. The fertilizer was initially administered as a basal application of 80:40:40 kg/ha of Nitrogen, Phosphorus, and Potassium, with half of the Nitrogen later applied as a top dressing. Irrigation was implemented as required, and intercultural practices such as thinning and weeding were conducted as necessary.

Biochemical Analysis: Both years freshly harvested seeds were analyzed for the diverse biochemical parameters like total oil content (%), Total soluble protein (%), Total phenol content (mg GAE/g), Palmitic acid (%), Stearic acid (%), Oleic acid (%), Linoleic acid (%), Linolenic acid (%). Arachidic acid (%), Eicosenoic acid (%), Behenic acid (%) and Erucic acid (%) content.

Oil Quantity Estimation using Soxhlet Method: Five grammes of finely crushed seed material were placed into a thimble and positioned in the extractor of the Soxhlet apparatus (fisher scientific), following the placement of cotton at the bottom. The Soxhlet extractor was employed to extract oil from the seed using n-hexane (99% extrapure) as the solvent. The extraction was performed using simple distillation to separate the solvent. The oil was let to cool in a desiccator prior to weighing. The extracted oil was securely packed in a dark brown glass container and stored for examination (W. Horwitz et al., 1970). Determine the diverse qualities and percentage yield of oil. Calculation: Weight of empty flask (g) = W_1, Weight of flask and extracted oil (g) = W_2, Weight of sample = S, **%** Oil = (W_2-W_1) × 100/S.

Estimation of Total Soluble Protein (%): The Lowry method (Lowry et al., 1951) was employed to determine the total soluble protein content. To extract the proteins from the defatted powdered material, 250 mg of powdered material from each accession was crushed thoroughly with 20 ml 0.1M Phosphate buffer (pH 7) in a pre-chilled mortar and pestle before being utilized. The blend was then precisely mixed and centrifuged for 20 minutes at 10,000 rpm, after which the protein analysis was carried out using the supernatant. 0.1 ml of protein extract was placed into a test tube using a micropipette, and the volume was adjusted to 1 ml by adding distilled water to the tube. It was then subjected to 5 ml of reagent C and allowed to sit for ten minutes. Afterward, 0.5 ml of reagent D was added and thoroughly mixed, then left to incubate at room temperature in the dark for 30 minutes. A blue color developed. The absorbance was measured with a UV–VIS spectrophotometer (Shimadzu UV-1601) at 660 nm in the presence of the blank reagent that did not include protein extract. It was necessary to develop a standard curve to measure the amount of protein present. The concentration of standard BSA (50 mg BSA was dissolved in 50 mL distilled water to make the standard stock solution, and the working standard was made by diluting the stock solution to 1:5 dilutions with distilled water) was plotted on the x-axis against absorbance on the y-axis to achieve this. The amount of protein in the sample powder was quantified in milligrams per gram of dry sample. (Here, Reagent A: 1 g Sodium Carbonate (Na_2CO_3) in 50 ml 0.1N Sodium hydroxide (NaOH); Reagent B: 0.25 g Copper sulphate ($CuSO_4$, $7H_2O$) in 50 ml 1% potassium sodium tartrate ($KNaC_4H_4O_6$); Reagent C: Combine 50 ml of reagent A and 1 ml of reagent B prior to use; Reagent D: 1N Folin-Ciocalteau reagent (FCR).

Estimation of Total Phenol Content: The Folin-Ciocalteu Reagent (FCR) method is used to estimate total phenols (Cliffe et al., 1994). To estimate the total phenol content from the samples, 15 mL of 80% ethanol was mixed with 0.5 g of powdered sample (defatted), which was then thoroughly crushed in a mortar and pestle. The homogenate was centrifuged at 10,000 rpm for 20 minutes, and the supernatant was kept. Re-extraction with 5 mL of 80 percent ethanol was carried out in a similar manner to the first extraction, and the supernatants were collected and allowed to evaporate over hot water bathed. In 50 ml of distilled water, the residue was diluted to provide a workable alternative. Following that, 0.3 ml. of the aliquot was pipetted into separate test tubes. The volume of each tube was modified to 3 ml using distilled water, and 0.5 ml of 1N Folin–Ciocalteau reagent was introduced. After allowing the mixture to react for three minutes, 2 mL of a 20% Na_2CO_3 solution was added and thoroughly mixed. Subsequently, the tubes were subjected to a boiling water bath for one minute and then allowed to cool. The

absorbance was detected immediately using a UV-VIS spectrophotometer at 650 nm. The total phenol content was determined in gallic acid equivalents using gallic acid equivalents (GAE) standard curve and expressed as mg/g of dry sample.

Oil quality/Fatty Acid profiling using Gas Chromatograph (GC): Analysis was done according to the modified methods of Seppänen-Laakso (Seppänen-Laakso et al., 2002). 100 mg of freshly harvested seeds were taken and homogenized with liquid nitrogen. Then, 1 ml of methanol and 1 ml of sodium methoxide were added, vortexed, and allowed to settle for 20-30 minutes. The mixture was transferred to a water bath at 55°C and left to cool until it reached room temperature. Afterwards, 1.5 mL of petroleum ether and 1.5 mL of Milli-Q water were added, followed by vortexing. The mixture was then left to settle for 15-30 minutes, and the upper phase was taken after phase separation for analysis using a GC. The analysis was conducted using a Shimadzu GC-17A gas chromatograph, with nitrogen as the carrier gas, and a mechanical pressure gauge controlled its flow rate. The injector and detector temperatures were set at 240°C. Ultrapure nitrogen gas was utilized as a transport medium. GC was programmed for a temperature increase of 100°C per minute, ultimately maintaining a temperature of 270°C. In gas chromatographs, the area under each peak is calculated automatically, utilizing the triangulation method, which involves measuring the peak height and width at half height. Following the computation of total peak area for each sample, the percent area under each peak was calculated to determine the percentage of the corresponding fatty acid.

Statistical Analysis: For statistical analysis, mean value of 5 tagged plants data was average for a replication. The ANOVA was done using (*Genstat for Windows, Release 11.1.0.1575. 11th Edition*, 2008). Genotypic Coefficient of Variation (GCV) and Phenotypic Coefficient of Variation (PCV) was calculated in following formula and computed in R software using 'variability' package of R-studio (Popat et al., 2020).

$$GCV=\frac{Genotypic\ standard\ deviation}{Mean}\sqrt{\sigma^2 g} \times 100$$

$$PCV=\frac{Phenotypic\ standard\ deviation}{Mean}\sqrt{\sigma^2 p} \times 100$$

$$\text{Heritability as measured by } (h^2) = \frac{\sigma^2 g}{\sigma^2 p} \times 100$$

Genetic Avance (GA) was calculated in, $G.A.= h^2 \times \sigma_p \times k$

$$\text{Genetic Advance as percentage of means (GAM)} = \frac{GA}{mean} \times 100$$

Where, σ^2_p = Phenotypic variance, σ^2_g = Genotypic variance, h^2 = Heritability in broad sense, K = Intensity of selection, σ_p = Phenotypic standard deviation.

Correlation was done using 'metan' package (Olivoto & Lúcio, 2020) in R-studio, Path coefficient analysis was done according to method by Dewey and Lu (Dewey & Lu, 1959) using GRAPES online tool. Hierarchical clustering with heat map was done using 'gplot' (Warnes et al., 2005) package in r-studio. The approach employed in this research study for multivariate principal component analysis (PCA) was developed by Hotelling (Hotelling, 1933) and is based on an earlier concept introduced by Pearson (Karl Pearson, 1904). Principal component analysis and PCA-biplot were executed using 'ggplot2' (Wickham, 2016), 'gridExtra' (Auguie, 2010), 'ggbiplot' (Vu & Friendly, 2023), and 'corrplot' (Wei & Simko, 2017). To enhance visualization, the 'factoextra' (Kassambara & Mundt, 2016) package is utilized.

3. RESULTS AND DISCUSSION

3.1 Genetic Variability for Different Biochemical Traits in *Brassica*

In this study, we utilized a comprehensive dataset that spans two years, which was pooled to perform an Analysis of Variance (ANOVA). The data was displayed in Supplementary Table 1. The genetic variations in 12 biochemical features of *Brassica sp.* were analyzed using basic statistics, including the mean with Multiple (pair-wise) comparisons using Tukey's HSD, standard error of mean (SEm±), and Least Significant Difference (LSD). The analysis of variance revealed significant differences between the genotypes for all traits, including Total oil content (%), Total phenol content (mg GAE/g), Total Soluble Protein (%), Palmitic acid (%), Stearic acid (%), Oleic acid (%), Linoleic acid (%), Linolenic acid (%), Arachidic acid (%), Eicosenoic acid (%), Behenic acid (%), and Erucic acid (%). Total phenol content ranges from 7.845 to 9.633 (mg GAE/g). Total soluble protein content ranges from 7.53 to 10.63%, with

the highest value genotype is TM-305-1. In terms of Palmitic acid, the range is from 1.747 to 2.581%, while the stearic acid is from 0.72 to 1.47. Oleic acid, linoleic acid and linolenic acid range from 7.28 to 15.18%, 11.92 to 19.63% and 6.70 to 16.46%, respectively (Samadzadeh Ghale Joughi et al., 2018). Arachidic acid, Eicosenoic acid and Behenic acid ranges from 0.49 – 1.20%, 5.17 – 10.65% and 0.547 – 1.135% respectively. Erucic acid contains 35.28 to 54.09% with having minimum amount of erucic acid presence in TM 313 (Ali et al., 2023). The erucic acid content has shown similar diversity of results as previously reported (Kumar Rai et al., 2018; Saini et al., 2016) . The oil content ranges from 25.68% to 35.43%, with the highest oil percentage found in TM-308-1. This result matches with Tahira et al., (Riffat Tahira et al., n.d.) and Mandal et al., (Mandal et al., 2002).

The box plot analysis (Fig. 1) showed a lot of variation in twelve traits among 13 *Brassica* species. The highest median values were found in Erucic acid (about 50) and Total Oil Content (about 30), with Erucic acid having a wide range of values. Most of the traits had median values below 20, while Total phenol content, Total soluble protein, Oleic acid, Linoleic acid, and Linolenic acid showed moderate ranges between 5 and 20. Palmitic acid, Stearic acid, Arachidic acid, and Eicosenoic acid had consistently low values close to zero, indicating a low presence or expression of these traits in the species studied. Some traits, especially Erucic acid, Oleic acid, and Linoleic acid had outliers outside their usual ranges, showing extreme values in certain species. The interquartile ranges differed greatly between traits, with Erucic acid and Total Oil Content having the largest spreads. In contrast, Palmitic acid, Stearic acid, and Eicosenoic acid had more tightly grouped distributions. This pattern suggests significant diversity in character expression across the studied *Brassica* species, with some traits showing high variability while others remained relatively consistent across the species examined.

For Total phenol content (mg GAE/g) (7.13 and 7.16) the GCV and PCV were both low (less than 10%). Moderate PCV and GCV (10-25%) were observed for Total oil content (%) (11.88 and 12.07), Total soluble protein content (%) (10.57 and 12.28), Palmitic acid content (%) (11.68 and 12.70), Linoleic acid content (%) (12.06 and 12.37), Linolenic acid content (%) (19.16 and 19.59), Stearic acid content (%) (21.98 and 22.27), Eicosenoic acid content (%) (22.16 and 22.68), Behenic acid content (%) (17.88 and 19.87) and Erucic acid content (%) (10.83 and 11.16) (Table 2). Elevated PCV and GCV (>25%) were noted for oleic acid content (%) (25.28 and 26.01) and arachidic acid content (%) (26.96 and 27.53). This outcome is consistent with references (Nagib Ali et al., 2007; S, 2013). Though the phenotypic coefficient of variation (PCV) was greater than the genotypic coefficient of variation (GCV) for every trait that was observed (Table 2) for all the characters but the PCV and GCV values for the different characteristics showed little variation, indicating that genetic factors have a stronger impact on the expression of these traits than environmental factors. The genetic advance, expressed as a percentage of the mean, was divided into three categories: high genetic advance (above 20%), moderate genetic advance (10-20%), and low genetic advance (less than 10%). All the biochemical traits under study having higher GA, except total phenol content (mg GAE/g) (14.61) and total soluble protein content (18.75) had a moderate GA in percentage of mean (Table 2). All traits exhibit high heritability. Except for the total phenol content (mg GAE/g) and the total soluble protein content (18.75), all the other traits show high heritability along with a significant high genetic advance. Consequently, these traits can be directly used for the future improvement of *Brassica juncea* genotypes. Similar findings were recorded by (Dupont et al., 1990; S, 2013).

3.2 Character Association Studies between Traits

The correlation of the characters and the magnitude of their relationship with other characters (Fig. 2) revealed that Erucic acid content (%) was significantly positively associated with behenic acid content (%) (0.7) and significantly negatively correlated with Linolenic acid (-0.61), Palmitic acid (-0.69), Linoleic acid (-0.61), Eicosenoic acid (-0.92), oleic acid (-0.49) and stearic acid (-0.34). Total oil content is significantly negatively correlated with Palmitic acid (-0.35) and total phenol content (-0.62). A Fatty acid, Arachidic acid is significantly positively correlated with Behenic acid (0.39), Eicosenoic acid (0.43), Oleic acid (0.76) and Stearic acid (0.94), whereas it was significantly negatively correlated with Linoleic acid (-0.63) and Linolenic acid (-0.46). Linolenic acid is significantly positively correlated with Palmitic acid (0.53), as previously suggested similar result (Riffat Tahira et al., n.d.), also with Linoleic acid (0.65) and Eicosenoic acid (0.36),

whereas significantly negatively correlated with total phenol content (-0.43), behenic acid (-0.79) and erucic acid (-0.61). Oleic acid exhibits a significant positive correlation with stearic acid (0.85), Arachidic acid (0.76) and Eicosenoic acid (0.73) whereas a significant negative correlation with erucic acid (-0.49) (J S Chauhan et al., 2007). Linoleic acid exhibits a significant positive correlation with both Linolenic acid (0.65) and Palmitic acid (0.76). In contrast, it shows a significant negative correlation with Arachidic acid (-0.63), Stearic acid (-0.44), Erucic acid (-0.61), Behenic acid (-0.66), and Total phenol content (mg GAE/g) (-0.41).

When exploring the correlations among various fatty acids in mustard, it is essential to consider the biosynthetic pathways that contribute to these relationships. Fatty acids are produced through a process called fatty acid synthesis (FAS), which mainly happens in the cytoplasm (Harwood, 2020). This process begins with a compound called acetyl-CoA, which is changed into malonyl-CoA by an enzyme known as acetyl-CoA carboxylase. The fatty acid synthase complex then builds on this starting point through a series of reactions, leading to the creation of long-chain fatty acids (Angeles & Hudkins, 2016; Magnuson et al., 1993).

The strong correlation between erucic acid and behenic acid can be understood because both are long-chain fatty acids made through similar processes of elongation. On the other hand, the significant negative relationships that erucic acid has with linolenic, linoleic, palmitic, oleic, and stearic acids may indicate that they compete with each other during the process of making these fatty acids. Hence, the positive relationship between arachidic acid and behenic, eicosenoic, oleic, and stearic acids supports the idea that these fatty acids are interconnected in metabolism. However, the negative relationship with linoleic acid shows that there is competition in making these fatty acids (Brown et al., 2009; Harwood, 2020). Understanding the relationships involved in the production of fatty acids can provide important insights into how these processes interact and this knowledge can be further used for modification of the level of any fatty acid content in mustard genotype. This can affect the fatty acid composition in mustard, influencing its nutritional value and its applications in various industries.

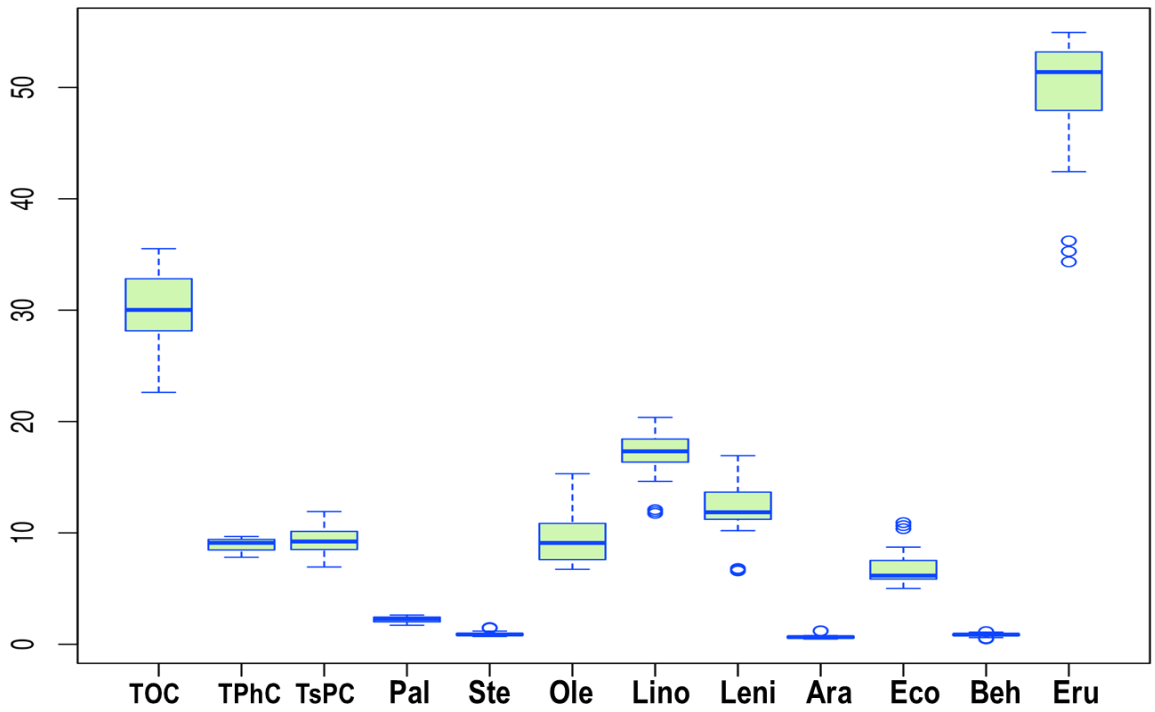


Fig. 1. Box plot showing different means with standard error of twelve characters of 13 *Brassica sp*. TOC=Total Oil Content (%), TsPhC=Total Phenol content (mg GAE/g), TsPC=Total Soluble Protein (%), Pal= Palmitic acid (%), Ste=Stearic acid (%), Ole=Oleic acid (%), Lino=Linoleic acid (%), Leni=Linolenic acid (%), Ara=Arachidic acid (%), Eco=Eicosenoic acid (%), Beh=Behenic acid (%) and Eru=Erucic acid (%)

Table 2. Genetic variability parameters for twelve biochemical traits in 13 genotypes of Indian Mustard

Characters	Genotypic Variance	Phenotypic Variance	Genotypic Coefficient of Variance (GCV)	Phenotypic Coefficient of Variance (PCV)	Heritability (%)	Genetic Advance (GA)	Genetic Advance (GAM-% of Mean)
TOC	12.76	13.19	11.88	12.07	96.75	7.2	24.06
TPhC	0.41	0.41	7.13	7.16	99.05	1.31	14.61
TsPC	0.97	1.31	10.57	12.28	74.14	1.75	18.75
Pal	0.07	0.08	11.68	12.7	84.63	0.49	22.13
Ste	0.04	0.04	21.98	22.27	97.43	0.42	44.70
Ole	6.11	6.47	25.28	26.01	94.46	4.95	50.62
Lino	4.22	4.44	12.06	12.37	95.02	4.12	24.22
Leni	5.40	5.64	19.16	19.59	95.68	4.68	38.61
Ara	0.03	0.03	26.96	27.53	95.89	0.37	54.39
Eco	2.23	2.33	22.16	22.68	95.53	3.0	44.62
Beh	0.02	0.02	17.88	19.87	80.99	0.28	33.14
Eru	28.87	30.64	10.83	11.16	94.24	10.74	21.68

Where, TOC=Total Oil Content (%), TPhC=Total Phenol content (mg GAE/g), TsPC=Total Soluble Protein (%), Pal= Palmitic acid (%), Ste=Stearic acid (%), Ole=Oleic acid (%), Lino=Linoleic acid (%), Leni=Linolenic acid (%), Ara=Arachidic acid (%), Eco=Eicosenoic acid (%), Beh=Behenic acid (%) and Eru=Erucic acid (%)

Table 3. Path coefficient table or 12 biochemical traits of 13 *Brassica sp.*

Character	TsPhC	TsPC	Plm	Str	Olc	Lnlc	Lnln	Arc	Ecs	Bhn	Erc	Cor_ TOC
TsPhC	**-0.06**	-0.09	0.14	0.10	-0.83	0.67	0.99	-0.25	0.29	-0.14	-1.46	**-0.63****
TsPC	0.01	**0.40**	-0.04	-0.82	0.69	-0.28	-0.29	0.49	0.05	-0.01	0.07	**0.28**
Palmitic	0.01	0.03	**-0.66**	-0.42	-0.12	-1.26	-1.27	0.48	-0.56	0.19	3.19	**-0.38***
Stearic	0.00	-0.11	0.09	**3.17**	-3.43	0.77	0.62	-1.59	-0.83	-0.07	1.55	**0.18**
Oleic	-0.01	-0.07	-0.02	2.76	**-3.94**	0.41	0.69	-1.28	-0.92	-0.01	2.21	**-0.18**
Linoleic	0.02	0.07	-0.51	-1.51	0.99	**-1.62**	-1.46	1.10	-0.34	0.24	2.80	**-0.21**
Linolenic	0.02	0.05	-0.37	-0.88	1.21	-1.06	**-2.24**	0.76	-0.52	0.29	2.89	**0.15**
Arachidic	-0.01	-0.12	0.19	3.03	-3.04	1.07	1.02	**-1.66**	-0.53	-0.16	0.36	**0.17**
Ecosenoic	0.01	-0.02	-0.29	2.08	-2.87	-0.44	-0.92	-0.70	**-1.26**	0.19	4.25	**0.04**
Behenic	-0.02	0.01	0.37	0.64	-0.10	1.13	1.89	-0.75	0.70	**-0.34**	-3.45	**0.07**
Erucic	-0.02	-0.01	0.45	-1.07	1.89	0.99	1.40	0.13	1.16	-0.26	**-4.60**	**0.09**

Residual value = 0.0874

Whereas, Cor_TOC=Correaltion with Total Oil Content (%) (as dependable variable), TsPhC=Total Phenol content (mg GAE/g), TsPC=Total Soluble Protein (%), Plm= Palmitic acid (%), Str=Stearic acid (%), Olc=Oleic acid (%), Lnlc=Linoleic acid (%), Lnln=Linolenic acid (%), Arc=Arachidic acid (%), Ecs=Eicosenoic acid (%), Bhn=Behenic acid (%) and Erc=Erucic acid (%)

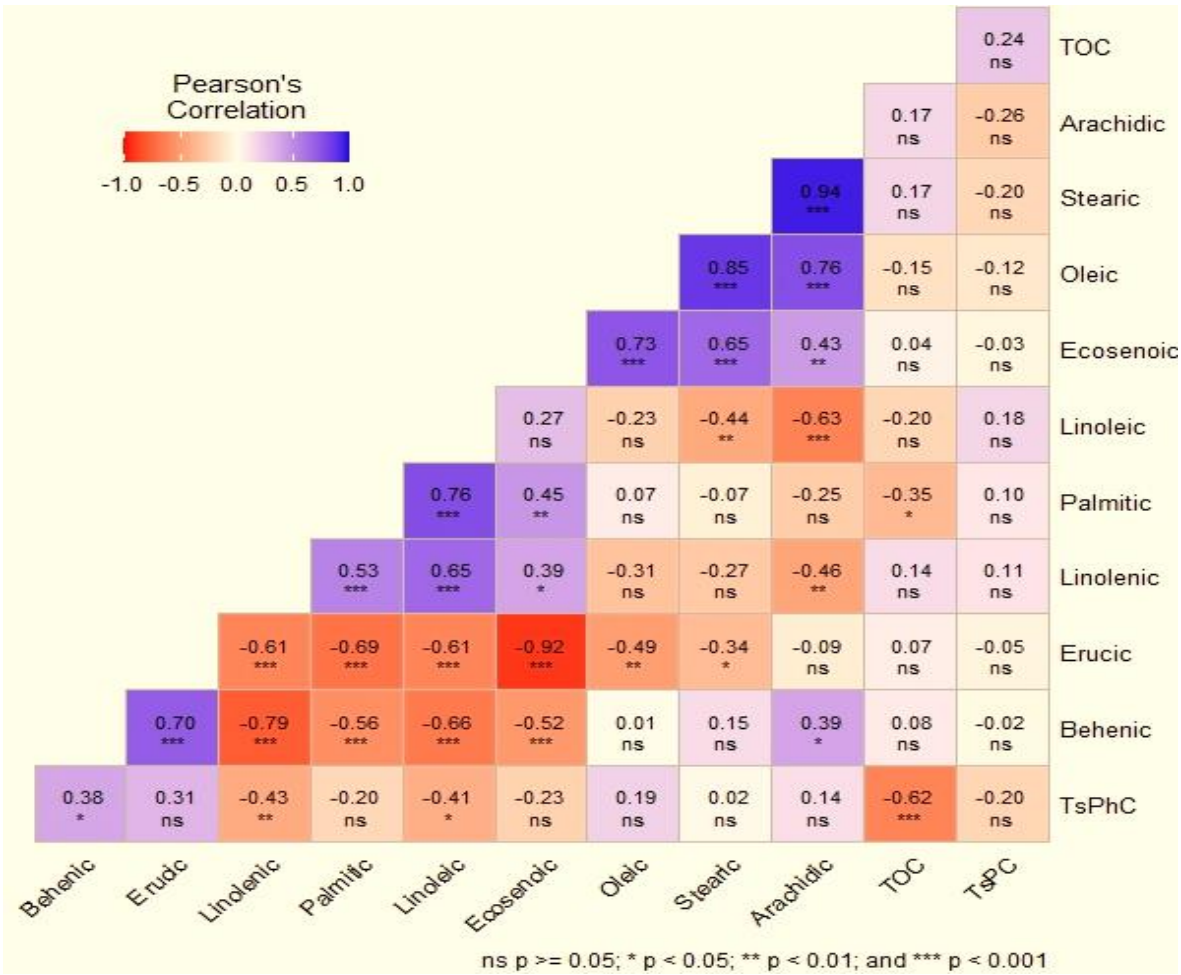


Fig. 2. Pearson's correlation among biochemical traits/parameters of 13 Mustard genotypes **: significant at 1% level, *: significant at 5% level, *: highly significant (p < 0.001); TOC=Total Oil Content (%), TsPhC=Total Phenol content (mg GAE/g), TsPC=Total Soluble Protein (%), Palmitic= Palmitic acid (%), Stearic=Stearic acid (%), Oleic=Oleic acid (%), Linoleic=Linoleic acid (%), Linolenic=Linolenic acid (%), Arachidic=Arachidic acid (%), Ecosenoic=Eicosenoic acid (%), Behenic=Behenic acid (%) and Erucic=Erucic acid (%)**

Path analysis is the method of statistical research which focuses on testing the association and relationship between a group of observed variables. The relations formed in this process can be of two types: direct or indirect. A direct relationship means one variable is associated with another. On the contrary, an indirect relation involves when one variable relates to another through a third, which in turn is directly associated with the outcome variable (TOC) (Valenzuela & Bachmann, 2017).

This path analysis (Table 3) illustrates the intricate relationships among various fatty acids, biochemical components, and Total Oil Content (TOC) in an oilseed crop study. The analysis indicates that stearic acid has the strongest positive direct effect (3.17) on total Oil Content, followed by total soluble protein content which also exerts a significant positive direct effect (0.4). Total phenol content (mg GAE/g) (-0.06) and all other different fatty acids under study except stearic acid demonstrate a negative direct effect. The soluble protein content with total phenolic content and different fatty acids exhibits contrasting behaviors due to competitive pathways involving nitrogen and carbon. Nitrogen is available to the plant for the synthesis of protein, while carbon is mainly used for energy-intensive such as oil synthesis. Since both nitrogen and carbon share the same substrate, acetyl CoA, this competition leads to an inverse relationship between protein content and the synthesis of phenolics and fatty acids (Mawlong et al., 2024).

The residual value of 0.0874 indicates that the path model fits well, as it shows a small amount of the observed variable is not explained by the model. This suggests that the model accounts for a high proportion of the variation in Total Oil Content. These relationships are consistent with established pathways of fatty acid biosynthesis while potentially uncovering new insights into the intricate interactions between different fatty acid components and total oil accumulation (Bates, 2016). From a breeding perspective, these results suggest that selecting for higher stearic acid content could be an effective strategy for enhancing total oil content. However, the negative relationship with palmitic acid, oleic acid and linoleic acid implies that breeders should carefully consider these competing effects. The presence of both positive and negative relationships among various fatty acids reflects the complex nature of lipid metabolism, indicating that multiple pathways contribute to the final determination of oil content. This path analysis proves to be quite valuable as it aids in distinguishing direct impacts from indirect impacts through other associated characteristics by deconstructing the genotypic correlation coefficient (Alves & Cargnelutti Filho, 2017).

3.3 Hierarchical Cluster Analysis of 13 *Brassica* Genotypes

An analysis was conducted on multiple biochemical characteristics using hierarchical cluster analysis. Cluster analysis shows that 13 Indian mustard genotypes have been divided into 4 major clusters (Fig. 3), i.e., Cluster I (G4, G3, G2, G7, G5, G6, G1), Cluster II (G10, G12, G11, G8), Cluster III (G9) and Cluster IV (G13). Shyam et al., 2021 reported similar results, having studied 188 genotypes and identified 18 different clusters based on biochemical parameters.

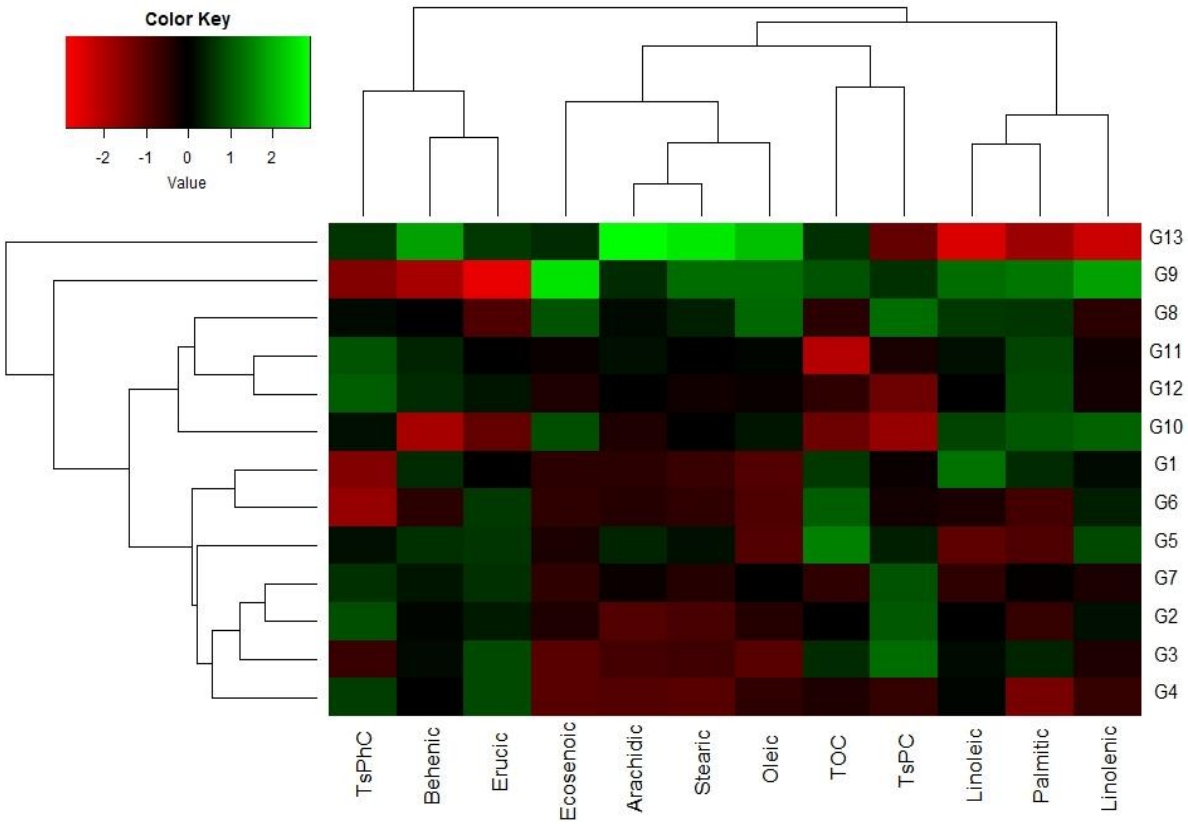


Fig. 3. The evaluated genotypes are divided into different clusters based on various biochemical parameters using heatmap and hierarchical clustering. Low performance for the corresponding characters is indicated by red colors, while high performance is indicated by blue colors. Whereas, TOC=Total Oil Content (%), TsPhC=Total Phenol content (mg GAE/g), TsPC=Total Soluble Protein (%), Palmitic= Palmitic acid (%), Stearic=Stearic acid (%), Oleic=Oleic acid (%), Linoleic=Linoleic acid (%), Linolenic=Linolenic acid (%), Arachidic=Arachidic acid (%), Ecosenoic=Eicosenoic acid (%), Behenic=Behenic acid (%) and Erucic=Erucic acid (%) (Genotypes name along with denotation are mentioned in Table 1)

3.4 Principal Component Analysis

PCA is a valuable method in genetic diversity research for reducing the complexity of data. It eliminates connections between components and identifies the variables that contribute the most to genetic variation, which can be chosen for further genotype characterization.

PC1 has an eigenvalue of 4.490 and a variance of 37.484%, having positive loading values (Supplementary Table 2) for Total soluble protein content (%), Palmitic acid, Linoleic acid, Linolenic acid and Eicosenoic acid while the rest of the factors are loaded as negative. In PC2, the eigenvalue and cumulative variance are 3.651 and 67.907% respectively. Total soluble protein Content, Linoleic acid content, Linolenic acid content, Behenic acid and Erucic acid show positive loading values, others were loaded as negative. PC3 has an eigenvalue of 1.854. Total phenol content (mg GAE/g), Palmitic acid, Oleic acid, and Linoleic acid content are loaded as negative, while the rest of the characters are loaded as positive. The individual PCA plots (Fig. 4) indicate that the *Brassica* population in our study exhibits considerable diversity for the examined variables. This divergence can be advantageous for selecting varied ancestors for forthcoming hybridization initiatives (Kishore et al., 2024).

The analysis of contribution patterns in the PCA corrplot reveals distinct clustering among variables, particularly focusing on the first twelve dimensions (Fig. 4). The primary sources of variation are concentrated in the first four dimensions, with significant contributions emerging particularly from Dimensions 1 through 4.

In Dimension 1 (Fig. 5), key contributors include several fatty acids—specifically Linoleic, Linolenic, and Behenic acids—along with moderate contributions from Palmitic acid and Total phenol content (TsPhC). This suggests a potential metabolic link between phenolic compounds and fatty acid synthesis,

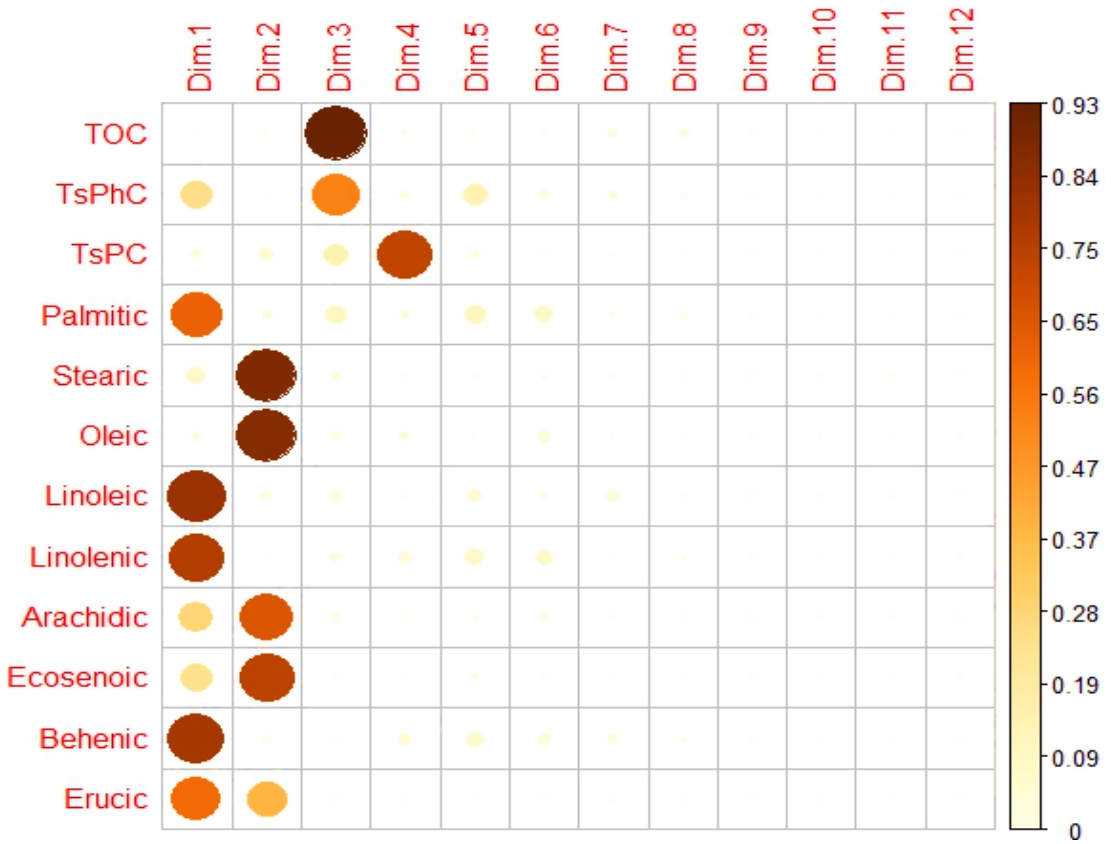


Fig. 4. Corrplot showing the contribution of each character on the first twelve components.; Whereas, TOC=Total Oil Content (%), TsPhC=Total Phenol content (mg GAE/g), TsPC=Total Soluble Protein (%), Palmitic= Palmitic acid (%), Stearic=Stearic acid (%), Oleic=Oleic acid (%), Linoleic=Linoleic acid (%), Linolenic=Linolenic acid (%), Arachidic=Arachidic acid (%), Ecosenoic=Eicosenoic acid (%), Behenic=Behenic acid (%) and Erucic=Erucic acid (%)

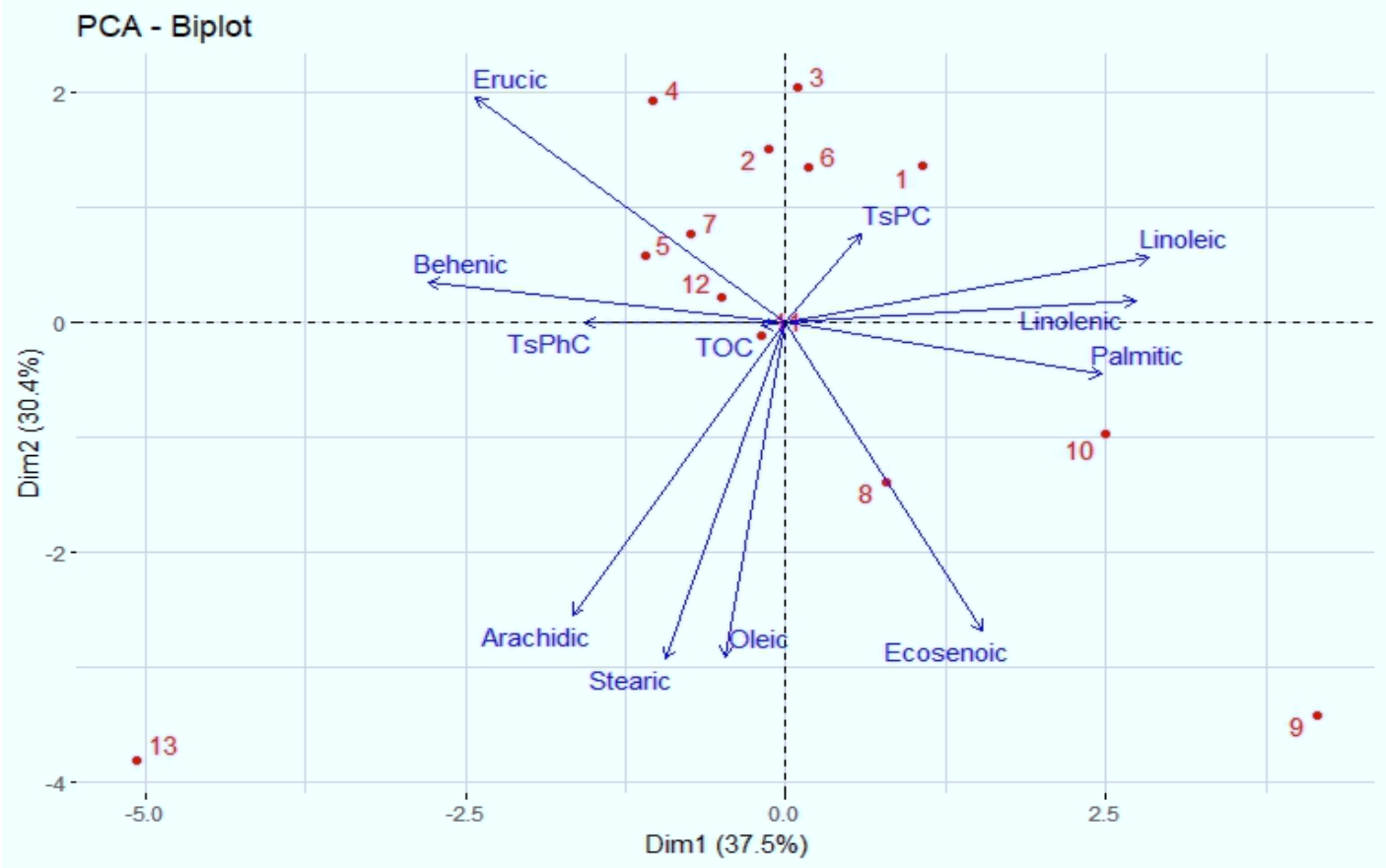


Fig. 5. Biplot of 13 Indian Mustard genotypes for 12 biochemical traits; TOC=Total Oil Content (%), TsPhC=Total Phenol content (mg GAE/g), TsPC=Total Soluble Protein (%), Palmitic= Palmitic acid (%), Stearic=Stearic acid (%), Oleic=Oleic acid (%), Linoleic=Linoleic acid (%), Linolenic=Linolenic acid (%), Arachidic=Arachidic acid (%), Ecosenoic=Eicosenoic acid (%), Behenic=Behenic acid (%) and Erucic=Erucic acid (%); 1=G1, 2=G2, 3=G3, 4=G4, 5=G5, 6=G6, 7=G7, 8=G8, 9=G9, 10=G10, 11=G11, 12=G12, 12=G13

likely influenced by energy distribution or regulatory mechanisms. Dimension 2 highlights Stearic and Oleic acids as major contributors, supported by Arachidic and Ecosenoic acids. The relationship between Stearic and Oleic acids is significant due to the former acting as a precursor for the latter through delta-9 desaturase (Gratraud et al., 2009; Nakamura & Nara, 2004; Si et al., 2023). In Dimension 3, Total Oil Content (TOC) shows significant contributions, alongside moderate input from TsPhC. Dimension 4 is primarily characterized by a strong contribution from Total soluble Protein (TsPC), indicating that protein content varies independently of oil traits. This separation allows for potential optimization of both oil and protein content, although trade-offs may occur at extreme values. Dimensions 5 through 12 exhibit considerably weaker contributions across all variables.

A PCA biplot includes vectors that extend from the origin to illustrate each trait, enhancing the visualization of the interrelationships among traits. The length of the vector for a trait indicates the magnitude of its effects on other traits (Yan & Tinker, 2005). The most effective way to observe the relationship between genotypes and traits is through the polygon view of the PCA biplot, as long as the biplot adequately represents the total variation. The cosine of the angle formed by two vectors can be used as an estimate for the correlation coefficient between two characteristics (Yan & Rajcan, 2002). According to this concept, two characters exhibit positive correlation when the angle between their vectors is less than 90 degrees, while a negative correlation occurs when the angle exceeds 90 degrees (Weikai Yan & Manjit S. Kang, 2002).

The Total phenol (mg GAE/g) showed a positive relationship with oil content and Behenic acid content, but an inverse relationship with Palmitic acid and Linolenic acid. Additionally, it was noted that Erucic acid is also positively correlated with total oil content, although to a lesser extent. However, genotypes with high erucic acid content are not preferred. Furthermore, the correlation between Linoleic acid, Linolenic acid, and Palmitic acid was found to be positive. Similarly, Arachidic acid was positively correlated with Stearic acid, Oleic acid, and Eicosenoic acid. The arrow vector of Eicosenoic acid forms an angle of 180^0 with Erucic acid, indicating the opposite in genotype ranking.

In the biplot diagram of PCA, PC1 and PC2 demonstrate this study's spread and variety of variables and genotypes. G9, G3, G4, and G13 each have one vector and also exhibit extreme values for any traits under study. Whether these genotypes are superior or not, they can serve as potential donor parents for upcoming breeding programme.

4. CONCLUSION

The current study focuses on characterizing *Brassica* genotypes for breeding programs and provides comprehensive guidelines based on multivariate analysis of biochemical traits. In the near future, if our objective is to develop Indian mustard genotypes with superior oil quality, we should aim to increase oleic acid content to over 50 percent, while keeping linoleic and linolenic acid levels below 40 percent and 5 percent, respectively. There was a notable inverse correlation between oleic acid and erucic acid, as indicated by the correlation analysis. Based on the observed patterns of genetic variability among the tested genotypes for various traits, it can be concluded that genotypes with low erucic acid content may be developed through hybridization with low erucic acid containing genotypes as parents. Path analysis revealed that stearic acid, linolenic acid, arachidic acid, ecosenoic acid, behenic acid, erucic acid and total soluble protein content have positive direct effect on total oil content. Genotypes TM-313, TM 305-1, TM 306-1, and B9 show the greatest divergence and may be utilized in future breeding programs. The variability observed among a diverse range of mustard genotypes suggests significant potential for enhancements aimed at both nutritional and industrial applications.

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.

ACKNOWLEDGEMENTS

The authors are thankful to the Dean, faculty of Agriculture & Directorate of Research for providing all the facilities for the present investigation. The authors would also like to thank BARC-UBKV MOU programme for providing the advanced mustard mutants (Trombay Mustard Mutants) for carrying out the present research whole. The authors are also thankful to the head, UDKVK, for providing

experimental field facilities and the Directorate of Research for providing funds to carry out this research. The authors are thankful to Dr. **Archana N. Rai** for critically reading the manuscript and for her suggestions.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

Ali, F., Ali, F., Bibi, A., S. Dessoky, E., Almowallad, S., AlShaqhaa, M. A., AL-Balawi, S. M., Darwish, D. B. E., Allohibi, A., Omara, M. Y., & Althobaiti, F. (2023). Morphological, Biochemical, and Molecular Characterization of Exotic *Brassica* Germplasm. *ACS Omega*, *8*(47), 44773–44783. Available: https://doi.org/10.1021/acsomega.3c05688

Alves, B. M., & Cargnelutti Filho, A. (2017). Genotypic correlation and path analysis in early and super-early maize genotypes. *Genetics and Molecular Research*, *16*(2). Available: https://doi.org/10.4238/gmr16029671

Angeles, T. S., & Hudkins, R. L. (2016). Recent advances in targeting the fatty acid biosynthetic pathway using fatty acid synthase inhibitors. *Expert Opinion on Drug Discovery*, *11*(12), 1187–1199. Available: https://doi.org/10.1080/17460441.2016.1245286

Auguie, B. (2010). *gridExtra: Miscellaneous Functions for "Grid" Graphics* (p. 2.3) [Dataset]. Available: https://doi.org/10.32614/CRAN.package.gridExtra

Bates, P. D. (2016). Understanding the control of acyl flux through the lipid metabolic network of plant oil biosynthesis. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1861*(9), 1214–1225. Available: https://doi.org/10.1016/j.bbalip.2016.03.021

Brown, A. P., Slabas, A. R., & Rafferty, J. B. (2009). Fatty Acid Biosynthesis in Plants—Metabolic Pathways, Structure and Organization. In H. Wada & N. Murata (Eds.), *Lipids in Photosynthesis* (Vol. 30, pp. 11–34). Springer Netherlands. Available: https://doi.org/10.1007/978-90-481-2863-1_2

Cliffe, S., Fawer, M. S., Maier, G., Takata, K., & Ritter, G. (1994). Enzyme assays for the phenolic content of natural juices. *Journal of Agricultural and Food Chemistry*, *42*(8), 1824–1828. Available: https://doi.org/10.1021/jf00044a048

Dewey, D. R., & Lu, K. H. (1959). A Correlation and Path-Coefficient Analysis of Components of Crested Wheatgrass Seed Production[1]. *Agronomy Journal*, *51*(9), 515–518. Available: https://doi.org/10.2134/agronj1959.00021962005100090002x

Dupont, J., White, P. J., Carpenter, M. P., Schaefer, E. J., Meydani, S. N., Elson, C. E., Woods, M., & Gorbach, S. L. (1990). Food uses and health effects of corn oil. *Journal of the American College of Nutrition*, *9*(5), 438–470. Available: https://doi.org/10.1080/07315724.1990.10720403

Genstat for Windows, Release 11.1.0.1575. 11th Edition. (2008). VSN International Ltd., Oxford. Genstat.co.uk

Harwood, J. L. (2020). Fatty acid biosynthesis. In D. J. Murphy (Ed.), *Plant Lipids* (1st ed., pp. 27–66). Blackwell. Available: https://doi.org/10.1201/9780367813123-2

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, *24*(6), 417–441. Available: https://doi.org/10.1037/h0071325

J S CHAUHAN, V P S BHADAURIA, MAHARAJ SINGH, K H SINGH, & ARVIND KUMAR. (2007). QualiSy characteristics and their interrelationships in Indian rapeseed-mustard (Brassiea sp) varieties. *The Indian Journal of Agricultural Sciences*, *77*(9), 616–620.

Karl Pearson. (1904). III. Mathematical contributions to the theory of evolution.—XII. On a generalised Theory of alternative Inheritance, with special reference to Mendel's laws. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, *203*(359–371), 53–86. Available: https://doi.org/10.1098/rsta.1904.0015

Kassambara, A., & Mundt, F. (2016). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses* (p. 1.0.7) [Dataset]. Available: https://doi.org/10.32614/CRAN.package.factoextra

Kim, Yong-Taek, Kim, Boh-Kyung, & Park, Kun-Young. (2007). Antimutagenic and Anticancer Effects of Leaf Mustard and Leaf Mustard Kimchi. *The Korean Society of Food Science and Nutrition*, *12*(2), 84–88.

Kishore, P., Roy, S., & Das, S. (2024). Multivariate analysis of grain yield attributes and spot blotch incidence in wheat (Triticum aestivum L.) under sub-Himalayan agroclimatic condition. *Electronic Journal of Plant Breeding*, *15*(1). Available: https://doi.org/10.37992/2024.1501.013

Kumar Rai, G., Bagati, S., Kumar Rai, P., & Kumar Rai, S. (2018). Fatty Acid Profiling in Rapeseed Mustard (Brassica species). *International Journal of Current Microbiology and Applied Sciences*, *7*(05), 148–157. Available: https://doi.org/10.20546/ijcmas.2018.705.019

Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951). Protein measurement with the Folin phenol reagent. *The Journal of Biological Chemistry*, *193*(1), 265–275.

Magnuson, K., Jackowski, S., Rock, C. O., & Cronan, J. E. (1993). Regulation of fatty acid biosynthesis in Escherichia coli. *Microbiological Reviews*, *57*(3), 522–542. Available: https://doi.org/10.1128/mr.57.3.522-542.1993

Mandal, K. G., Saha, K. P., Ghosh, P. K., Hati, K. M., & Bandyopadhyay, K. K. (2002). Bioenergy and economic analysis of soybean-based crop production systems in central India. *Biomass and Bioenergy*, *23*(5), 337–345. Available: https://doi.org/10.1016/S0961-9534(02)00058-2

Mawlong, I., Singh, V. V., Ram, B., Garg, P., Rani, R., Kumar, M. S. S., Gurung, B., & Rai, P. K. (2024). Exploring biochemical traits for heat stress tolerance in Indian mustard germplasm. *Vegetos*. Available: https://doi.org/10.1007/s42535-024-00927-y

McVetty, P. B. E., & Duncan, R. W. (2015). Canola, Rapeseed, and Mustard: For Biofuels and Bioproducts. In V. M. V. Cruz & D. A. Dierig (Eds.), *Industrial Crops* (Vol. 9, pp. 133–156). Springer New York. Available: https://doi.org/10.1007/978-1-4939-1447-0_7

Muhammad Haseeb Anwar Tarar, Hoda Zahoor, Nida Talat, Rashida Humma, Muhammad Humayun Ajaz, Syed Muhammad Zaigham Zia, & Sara Siddique. (2021). Biochemical, Biotechnical Significance of Mustard and its Role in Agricultural Based Industries. *Scholars Bulletin*, *7*(7), 185–189.

Nagib Ali, Farhatullah, Naqib Ullah Khan, Malik Ashiq Rabbani, Ijaz Hussain, Sardar Ali, Sher Aslam, & Muhammad Qasim Kakar. (2007). Genetic diversity of Brassica Rapa L. Indigenous landraces based on cluster and principal component analyses. *Pakistan Journal of Botany*, *49*(5), 1891–1901.

Nieschlag, H., & Wolff, I. (1971). Industrial uses of high erucic oils. *J. Am. Oil Chem. Soc*, *18*, 723–727.

Olivoto, T., & Lúcio, A. D. (2020). metan: An R package for multi-environment trial analysis. *Methods in Ecology and Evolution*, *11*(6), 783–789. Available: https://doi.org/10.1111/2041-210X.13384

Popat, R., Patel, R., & Parmar, D. (2020). *variability: Genetic Variability Analysis for Plant Breeding Research* (p. 0.1.0) [Dataset]. Available: https://doi.org/10.32614/CRAN.package.variability

Rahman, J., Sultana, F., Fatima, K., Hasan, M. M., Gain, N., Hossain, M. S., Chowdhury, A. K., & Rahman, A. (2022). GENETIC DIVERSITY OF FIELD MUSTARD (BRASSICA RAPA L.) AND THEIR SATURATED AND UNSATURATED FATTY ACIDS ASSOCIATION. *SABRAO Journal of Breeding and Genetics*, *54*(2), 249–266. Available: https://doi.org/10.54910/sabrao2022.54.2.4

Riffat Tahira, Ihsanullah, Ata-ur-Rehman, & Mahjabeen Saleem. (n.d.). Studies on variability for quality traits, association and path analysis in raya (Brassica juncea) Germplasm. *Int. J. Agri. Biol.*, *17*, 381–386.

S, K. (2013). Genetic analysis of oil content and quality parameters in Indian mustard (Brassica juncea (L) Czern and Coss). *Scholarly J Agril Sci*, *3*, 299–304.

Saini, N., Singh, N., Kumar, A., Vihan, N., Yadav, S., Vasudev, S., & Yadava, D. K. (2016). Development and validation of functional CAPS markers for the FAE genes in Brassica juncea and their use in marker-assisted selection. *Breeding Science*, *66*(5), 831–837. Available: https://doi.org/10.1270/jsbbs.16132

Salah, H. A., Elsayed, A. M., Bassuiny, R. I., Abdel-Aty, A. M., & Mohamed, S. A. (2024). Improvement of phenolic profile and biological activities of wild mustard sprouts. *Scientific Reports*, *14*(1), 10528. Available: https://doi.org/10.1038/s41598-024-60452-5

Samadzadeh Ghale Joughi, E., Majidi Hervan, E., Shirani Rad, A. H., & Noormohamadi, Gh. (2018). *Fatty acid composition of oilseed rapeseed genotypes as affected by vermicompost application and different thermal regimes* [PDF]. 640.5Kb. Available: https://doi.org/10.15159/AR.18.002

Seppänen-Laakso, T., Laakso, I., & Hiltunen, R. (2002). Analysis of fatty acids by gas chromatography, and its relevance to research on health and nutrition. *Analytica Chimica Acta*, *465*(1–2), 39–62. Available: https://doi.org/10.1016/S0003-2670(02)00397-5

Sharad Pandey, Manoj Kabdal, & M.K. Tripathi. (2013). Study of inheritance of erucic acid in Indian mustard (Brassica juncea L Czern & Coss). *Octa J. Biosci*, *1*, 77–84.

Sharma, A., Aacharya, M., Punetha, H., Sharma, S., Kumari, N., Singh, K. H., & Rai, P. K. (2019). Biochemical Characterization and their Correlations in Brassica juncea Genotypes. *International Journal of Current Microbiology and Applied Sciences*, *8*(01), 2408–2417. Available: https://doi.org/10.20546/ijcmas.2019.801.254

Shyam, C., Tripathi, M. K., Tripathi, N., Tiwari, S., & Sikarwar, R. S. (2021). Genetic Variations in fatty Acids and Oil Compositions among 188 Indian Mustard Brassica juncea (Linn.) Czern & Coss genotypes. *Current Journal of Applied Science and Technology*, 9–28. Available: https://doi.org/10.9734/cjast/2021/v40i4631629

Singh, K. H., Shakya, R., & Mahawar, R. (n.d.). Genetic diversity and patterns of variation among Indian mustard (Brassica juncea (L.) Czernj.& Cosson) varieties. *SABRAO J. Breed. Genet*, *46*(2), 329–339.

Stoutjesdijk, P. A., Hurlestone, C., Singh, S. P., & Green, A. G. (2000). High-oleic acid Australian Brassica napus and B. juncea varieties produced by co-suppression of endogenous Delta12-desaturases. *Biochemical Society Transactions*, *28*(6), 938–940.

Tian, Y., & Deng, F. (2020). Phytochemistry and biological activity of mustard (*Brassica juncea*): A review. *CyTA - Journal of Food*, *18*(1), 704–718. Available: https://doi.org/10.1080/19476337.2020.1833988

Valenzuela, S., & Bachmann, I. (2017). Path Analysis. In J. Matthes, C. S. Davis, & R. F. Potter (Eds.), *The International Encyclopedia of Communication Research Methods* (1st ed., pp. 1–9). Wiley. Available: https://doi.org/10.1002/9781118901731.iecrm0181

Vu, V. Q., & Friendly, M. (2023). *ggbiplot: A Grammar of Graphics Implementation of Biplots* (p. 0.6.2) [Dataset]. Available: https://doi.org/10.32614/CRAN.package.ggbiplot

W. Horwitz, P. Chichilo, & H. Reynolds. (1970). *Official methods of analysis of the Association of Official Analytical Chemists*. Washington, DC, USA: Association of Official Analytical Chemists.

Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., Schwartz, M., Venables, B., & Galili, T. (2005). *gplots: Various R Programming Tools for Plotting Data* (p. 3.2.0) [Dataset]. Available:https://doi.org/10.32614/CRAN.package.gplots

Warwick, S. I., Francis, A., & Al-Shehbaz, I. A. (2006). Brassicaceae: Species checklist and database on CD-Rom. *Plant Systematics and Evolution*, *259*(2–4), 249–258. Available: https://doi.org/10.1007/s00606-006-0422-0

Wei, T., & Simko, V. (2017). *R package "corrplot": Visualization of a Correlation Matrix*. Available: https://github.com/taiyun/corrplot

Weikai Yan & Manjit S. Kang. (2002). *GGE Biplot Analysis* (1st ed.). CRC Press.

Wickham, H. (with Sievert, C.). (2016). *ggplot2: Elegant graphics for data analysis* (Second edition). Springer international publishing.

Yan, W., & Rajcan, I. (2002). Biplot Analysis of Test Sites and Trait Relations of Soybean in Ontario. *Crop Science*, *42*(1), 11–20. Available: https://doi.org/10.2135/cropsci2002.1100

Yan, W., & Tinker, N. A. (2005). An Integrated Biplot Analysis System for Displaying, Interpreting, and Exploring Genotype × Environment Interaction. *Crop Science*, *45*(3), 1004–1016. Available: https://doi.org/10.2135/cropsci2004.0076

SUPPLEMENTARY DATA

Supplementary Table 1. Analysis of variance for 12 biochemical traits of *Brassica* Sp

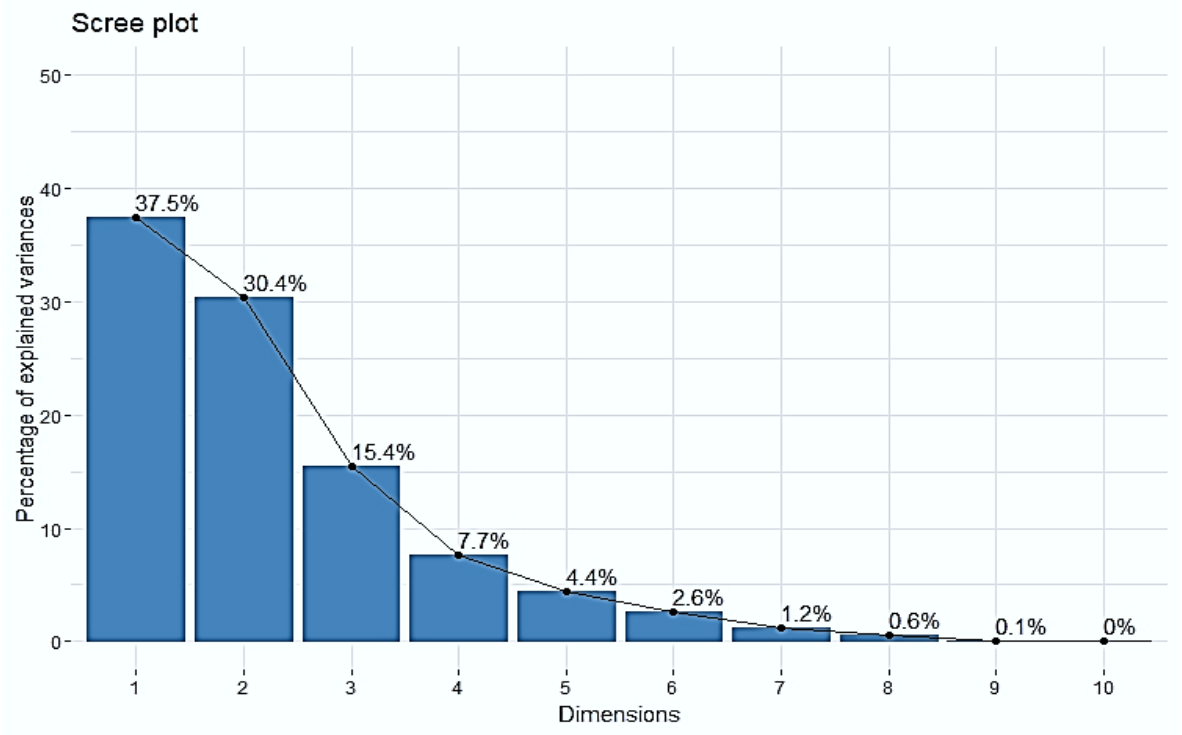
Genotype	TOC**	TsPhC**	TsPC**	Pal**	Ste**	Ole**	Lino**	Leni**	Ara**	Eco**	Beh**	Eru**
TM-301-3	32.46^{de}	7.98^{a}	9.15^{abcd}	2.35^{def}	0.79^{abc}	7.32^{a}	19.68^{g}	12.43^{cde}	0.58^{abc}	5.95^{abc}	0.93^{bc}	49.63^{c}
TM-303-1	30.02^{bc}	9.50^{def}	10.3^{cd}	2.05^{abcd}	0.75^{ab}	8.65^{abc}	17.01^{cd}	12.63^{def}	0.50^{a}	6.16^{bc}	0.86^{b}	51.38^{cd}
TM-305-1	31.92^{cd}	8.51	10.63^{d}	2.34^{cdef}	0.78^{ab}	7.28^{a}	17.34^{def}	11.28^{bcd}	0.53^{ab}	5.17^{a}	0.87^{b}	54.07^{d}
TM-306-1	28.82^{ab}	9.39^{cde}	8.7^{abc}	1.85^{a}	0.72^{a}	8.40^{ab}	17.2^{cde}	10.7^{b}	0.49^{a}	5.23^{ab}	0.84^{b}	54.09^{d}
TM-308-1	35.43^{f}	9.06^{b}	9.7^{bcd}	1.97^{ab}	0.97^{ef}	7.45^{a}	14.79^{b}	14.11^{fg}	0.75^{e}	6.25^{c}	0.94^{bc}	52.77^{cd}
TM-309-2	33.96^{ef}	7.85^{a}	9.05^{abcd}	2.01^{abc}	0.82^{abc}	7.47^{a}	16.4^{cd}	12.97^{ef}	0.59^{abc}	5.91^{abc}	0.77^{b}	53.06^{cd}
TM-310-3	28.03^{a}	9.28^{c}	10.30^{cd}	2.19^{bcde}	0.84^{bcd}	9.80^{bc}	15.89^{bc}	11.36^{bcd}	0.64^{bcd}	5.94^{abc}	0.88^{b}	52.45^{cd}
TM-312-1	28.25^{ab}	9.03^{b}	10.62^{d}	2.39^{ef}	1.00^{f}	12.83^{d}	18.42^{efg}	10.99^{bc}	0.70^{de}	8.14^{d}	0.84^{b}	44.69^{b}
TM-313	33.53^{def}	8.00^{a}	9.89^{cd}	2.58^{f}	1.18	12.88^{d}	19.66^{g}	16.46	0.76^{e}	10.65	0.55^{a}	35.28^{a}
TM-316	25.68	9.08^{b}	7.53^{a}	2.49^{ef}	0.92^{def}	10.37^{c}	18.63^{fg}	14.83^{g}	0.60^{abcd}	8.05^{d}	0.55^{a}	43.55^{b}
PM-25	2.75	9.56^{ef}	8.99^{abcd}	2.43^{ef}	0.93^{def}	10.01^{bc}	17.4^{def}	11.66^{bcde}	0.70^{de}	6.48^{c}	0.92^{bc}	49.45^{c}
JD-6	28.12^{ab}	9.63^{f}	7.98^{ab}	2.45^{ef}	0.89^{cde}	9.47^{bc}	16.98^{cd}	11.55^{bcde}	0.67^{cde}	6.16^{bc}	0.93^{bc}	50.91^{cd}
B9	32.1^{de}	9.33	8.11^{ab}	1.75^{a}	1.47	15.18	11.92^{a}	6.7^{a}	1.20	7.5^{d}	1.14^{c}	53.14^{cd}
Grand Mean	30.08	8.94	9.31	2.22	0.93	9.78	17.02	12.13	0.67	6.74	0.85	49.57
SEm (±)	0.38	0.03	0.33	0.06	0.02	0.35	0.27	0.28	0.02	0.18	0.04	0.76
LSD	1.10	0.11	0.98	0.19	0.05	1.0	0.79	0.83	0.06	0.54	0.12	2.24
CV%	2.2	0.7	6.2	5.0	3.5	6.1	2.8	4.1	5.6	4.8	8.6	2.7

*Means with common letter are do not significantly different at P ≤ 0.05, based on Tukey HSD least significant difference test. Symbol ** & * indicate the significant at 5% and 1% respectively; TOC=Total Oil Content (%), TPhC=Total Soluble Phenol content (mg GAE/g), TsPC=Total Soluble Protein (%), Pal= Palmitic acid (%), Ste=Stearic acid (%), Ole=Oleic acid (%), Lino=Linoleic acid (%), Leni=Linolenic acid (%), Ara=Arachidic acid (%), Eco=Eicosenoic acid (%), Beh=Behenic acid (%) and Eru=Erucic acid (%)*

Supplementary Table 2. Proportion of variance, cumulative proportion, eigen values and factor loadings of *Brassica* genotypes with respect to different PC's

Parameter	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12
TOC	-0.029	-0.006	0.709	0.103	0.057	-0.155	-0.408	0.529	-0.003	0.101	-0.005	0.022
TPhC	-0.023	-0.002	-0.536	-0.139	0.519	-0.300	-0.384	0.274	0.203	0.092	0.06	0.022
TsPC	0.089	0.127	0.282	-0.889	0.214	-0.011	0.185	-0.09	0.115	0.054	0.019	0.01
Pal	0.371	-0.077	-0.244	-0.167	-0.461	-0.520	0.236	0.466	-0.096	-0.036	0.019	0.013
Ste	-0.140	-0.489	0.113	0.015	-0.06	-0.134	0.096	-0.054	0.628	-0.402	-0.35	0.121
Ole	-0.071	-0.487	-0.122	-0.196	0.025	0.360	-0.040	0.211	-0.448	0.154	-0.369	0.407
Lino	0.426	0.092	-0.133	-0.092	-0.334	0.208	-0.532	-0.168	0.364	0.193	0.079	0.379
Leni	0.410	0.030	0.125	0.207	0.406	-0.477	0.034	-0.368	-0.207	0.031	-0.261	0.360
Ara	-0.250	-0.425	0.099	0.035	-0.124	-0.267	0.140	-0.241	0.097	0.671	0.345	0.042
Eco	0.230	-0.448	0.042	-0.050	0.168	0.048	-0.145	-0.064	-0.191	-0.46	0.661	0.02
Beh	-0.418	0.056	0.019	-0.230	-0.375	-0.350	-0.434	-0.355	-0.340	-0.248	-0.114	-0.021
Eru	-0.362	0.326	0.035	0.614	-0.051	-0.023	0.264	0.132	0.042	-0.173	0.304	0.735
Eigen values	4.490	3.651	1.854	0.920	0.527	0.317	0.149	0.070	0.011	0.002	0.001	0.000
Total variance (%)	37.484	30.423	15.449	7.667	4.391	2.643	1.245	0.580	0.095	0.017	0.005	0.000
Cumulative variance %	37.484	67.907	83.357	91.024	95.415	98.058	99.303	99.882	99.977	99.995	100.000	100.000

Where, TOC=Total Oil Content (%), TsPhC=Total Phenol content (mg GAE/g), TsPC=Total Soluble Protein (%), Pal= Palmitic acid (%), Ste=Stearic acid (%), Ole=Oleic acid (%), Lino=Linoleic acid (%), Leni=Linolenic acid (%), Ara=Arachidic acid (%), Eco=Eicosenoic acid (%), Beh=Behenic acid (%) and Eru=Erucic acid (%)



Supplementary Fig. 1. Scree plot with percentage of variance for each PC's

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/127627