***Original Research Article***

# Principle Component Analysis for Assessment of Genetic Diversity in Nigella (*Nigella sativa* L.) Collections


## Abstract

Seventeen land races of Nigella along with one released variety (Rajendra Shyama) as a check; collected from different parts of Bihar (**Table 1**) were evaluated in Randomized Block Design with three replications at Seed production Farm, TCA, Dholi, Bihar during *Rabi* 2015-16. Principle component analysis (PCA) showed that first three PCs had >1.00 Eigen value and accounted to 84.71 % of total variation. Rotated component matrix for various traits revealed that PC1 was strongly associated with secondary branches/plant followed by yield/plant, length of fruit, fruit per plant, primary branches/plant, height of the plant, days to 50% flowering and grains/plant. The traits that mostly contributed to PC2 were grains/plant followed by height of the plant and width of fruit whereas, days to maturity followed by width of fruit, height of the plant, days to 50% flowering and length of fruit contributed mostly to the PC3. The characters that contributed most to the PC4 were height of the plant, fruit/plant and length of fruit. Therefore, intensive selection procedures can be adopted to bring about rapid improvement of above mentioned traits. The k-mean of different clusters indicated that genotype falling in cluster III possess high values for all the traits under study indicating their potentiality as a parent in hybridization programmes for further improvement of Nigella. Highest inter-cluster distance was noted between cluster III and V indicating the genetic diversity among genotypes of these two clusters. Therefore, genotypes from these two clusters are recommended to use in hybridization programmes for further improvement.



Keywords: Nigella sativa, Principle Component Analysis, Genetic Diversity, Black Cumin


## Introduction:

Black Cumin (*Nigella sativa* L.) is an annual herbaceous plant belonging to the family Ranunculacea (Hammo, Y. H., 2008). It is popularly known as kalongi and an important seed spice crop of India. Black cumin is grown under a wide range of environments, but flourishes in cooler and dry regions (Weiss, 2002). A temperature range of 5-25$^0$C with the optimum of 12 -14$^0$C and rainfall of 400-500 mm are the most suitable to produce good crops. The crop is frost sensitive at any growth stage and this limits its distribution in Europe and the highland areas of the tropics. Black cumin can grow on all kinds of soils (Jansen, 1981) but, it prefers loamy sand soils (Datta *et al.*, 2001). It can be grown from sea level to 2500 m of altitude with a reduction in yield with increasing altitude.

The nigella producing countries other than India are Pakistan, Sri Lanka, Bangladesh, Nepal, Egypt and Iraq. In India it is cultivated commercially in Madhya Pradesh, Bihar, Punjab and Assam. It has also been noticed to occur wild in these areas. The other states

where its cultivation has been taken up on small scale are Uttar Pradesh, Rajasthan, Tamil Nadu and West Bengal.

The dried seeds of nigella are the commercial product being used in food. The seeds contain 0.5 to 1.4% essential oil which has demand in the pharmaceutical and perfume industry (Malhotra, 2004). It has been used since antiquity for culinary, seasoning and pharmacological purposes (Shah, S. H., 2008). Many medical properties have been attributed to the *Nigella sativa* L. seeds and its oil, including carminatives, diuretics, antineoplastic (antitumour),antifungal, anti-helminthic, while their oil has protective action against histamine induced bronchospasm, cough and bronchal asthma (Worthen *et al*., 1998; Khan *et al*., 2003), antidiabetics (Fararh *et al*., 2002; El-Dakhakhny *et al.*, 2002), spasmolytic and bronchodilator (Boskabady *et al*., 2004), anti-inflammatory (Hajhashemi *et al*., 2004), antibacterial (Mashhadian & Rakhshandeh, 2005), galactogogue, antioxidant (Brutis & Bucar, 2000; Kanter *et al*., 2003) and insect repellent effects (Fisher, 2002). Additionally, black seed is a valuable source of protein, carbohydrates, essential fatty acids, vitamins as well as minerals such as calcium, potassium, iron, magnesium, selenium, manganese and zinc. *Nigella sativa* L. is being considered important for both oil and bioactive compounds because their constituents have unique chemical properties and may augment the supply of edible oils (Ramadan &Morsel, 2003). Black cumin seed has higher total phospholipid content than cotton and soybean seed oil (Gunstone *et al*., 1986; Atta, 2003; Ramadan and Morsel, 2003).

Genetic diversity is pre-requisite for any crop improvement programme, as it helps in the development of superior recombinant (Manonmani & Fazlullah Khan., 2003).Genetic distance estimates for population grouping can be estimated by different methods as it is crucial to understand the usable variability existing in the population panel (Nachimuthu *et al.*, 2014). One of the approaches is to apply multivariate analysis. Principal Component Analysis (PCA) is a powerful tool in modern data analysis because this is a well-known multivariate statistical technique which is used to identify the minimum number of components, which can explain maximum variability out of the total variability (Anderson, 1972and Morrison, 1978) and also to rank genotypes on the basis of PC scores. Principal components are generally estimated either from correlation matrix or covariance matrix. Considering the importance of PCA this study is conducted on diverse Nigella genotypes for identification of traits responsible for yield differences.

**Material and Methods:**

Seventeen land races of Nigella along with one released variety (Rajendra Shyama) as a check, collected from different parts of Bihar (**Table 1**) were evaluated in Randomized Block Design with three replications at Seed production Farm, TCA, Dholi, Bihar during *Rabi* 2015-16.The experimental site is located at 25.59 N latitude and 85.75 E longitudes and has altitude of 51.20 m above mean sea level. Soil of TCA, Dholi, Bihar is mainly young alluvium and calcareous. Standard agronomic practices were adopted with row to row and plant to plant spacing of 25×5 cm., recommended dose of fertilizer was applied during the time of crop period. Data was recorded for ten different traits *viz.*, height of the plant, primary branches per plant, secondary branches per plant, days to 50 per cent flowering, length of fruit, width

of fruit, days to maturity, fruit per plant, grains per plant and yield per plant. Five competitive plants from each plot were randomly chosen for recording of data except for days to 50 per cent flowering and days to maturity whose data were recorded on plot basis.

**Table 1. Name, Source and collection year of the Nigella genotypes**

| S. No. | Name | Source | Collection Year |
|---|---|---|---|
| **1)** | RN-20 | East Champaran, Bihar | 2000 |
| **2)** | RN-22 | Muzaffarpur, Bihar | 2001 |
| **3)** | RN-25 | Bhagalpur, Bihar | 2001 |
| **4)** | RN-65 | Nalanda, Bihar | 2002 |
| **5)** | RN-66 | Banka, Bihar | 2002 |
| **6)** | RN-68 | Jammui, Bihar | 2002 |
| **7)** | RN-69 | Munger, Bihar | 2003 |
| **8)** | RN-70 | Nawada, Bihar | 2003 |
| **9)** | RN-71 | Gopalganj, Bihar | 2004 |
| **10)** | RN-72 | Siwan, Bihar | 2004 |
| **11)** | RN-73 | Chhapra, Bihar | 2004 |
| **12)** | RN-74 | Arrah, Bihar | 2005 |
| **13)** | RN-75 | Khagaria, Bihar | 2005 |
| **14)** | RN-76 | Gopalganj, Bihar | 2005 |
| **15)** | RN-77 | Vaishali, Bihar | 2006 |
| **16)** | RN-78 | Darbhanga, Bihar | 2006 |
| **17)** | RN-79 | Sitamarhi, Bihar | 2007 |
| **18)** | Rajendra Shyama | Released variety (Check) | |

The data were subjected to pooled analysis for genetic divergence by using statistical package WINDOSTAT version 9.2 (INDOSTAT service, Hyderabad). Intra and inter-cluster Euclidean distances generated were used to describe the relationship among the genotypes. Cluster analysis was carried out to construct the dendrogram, depicting the relationship among genotypes based on the genetic distances.

**Results and Discussion:**

The analysis of variance indicated that mean squares for genotype were highly significant for all the traits under study. For establishing genetic relationship among the traits and their genetic discrimination Principal Component analysis was performed. Association between traits emphasised by this method may correspond to genetic linkage between loci controlling the traits or a pleotropic effect (Iezzoni and Pritts, 1991)

The genetic variation present in breeding population was divided into four principal components (PCs) which explained 91.69% of total variation (**Table 3**). According to Brejda *et al.*, 2000, the PC with eigen values >1 and which explained at least 5% of the variation in the data will be considered as principal components. The first three PCs had >1.00 eigen value and accounted to 84.71 % of total variation. It indicates that the identified characters within these components exhibited immense influence on the phenotype of the genotypes.

The first principal component ($PC_1$) explained 55.47% of the total variation. The second principal component ($PC_2$) explained 17.14% variation individually and 72.62% cumulative variation. The third principal component ($PC_3$) explained 12.08% variation individually and 84.71% cumulatively.

Rotated component matrix for various traits revealed that$PC_1$ was strongly associated with secondary branches/plant followed by yield/plant, length of fruit, fruit per plant, primary branches/plant, height of the plant, days to 50%flowering and grains/plant. The traits that mostly contributed to $PC_2$ were grains/plant followed by height of the plant and width of fruit whereas, days to maturity followed by width of fruit, height of the plant, days to 50% flowering and length of fruit contributed mostly to the $PC_3$. The characters that contributed most to the $PC_4$ were height of the plant, fruit/plant and length of fruit. Therefore, intensive selection procedures can be adopted to bring about rapid improvement of above mentioned traits.

**Cluster analysis:**

Cluster analysis helps in selection of suitable genotype(s) or parent to use in hybridization programme for the manipulation of desirable traits. Choice of proper parent(s) plays a vital role for a successful plant breeding programme. Parents having more genetic distance believed to create higher variations by generating higher recombination frequency, which increase the genetic gain in selection. The grouping of the mutant lines was done by K-mean clustering pattern. The distribution of 18 lines along with check into five clusters and their cluster means are presented in **Table 5** and **6** respectively. Cluster III and IV comprised of only one line forming the smallest cluster followed by cluster V which was comprised of only two lines. Cluster I and II comprised of nine and five nigella lines respectively. The check Rajendra Shyama was clustered in Cluster V. The k-mean of different clusters indicated that genotype falling in cluster III possess high values for all the traits under study. The genotypes in cluster I have minimum values for height of the plant, primary branches/plant, secondary branches/plant, days to 50% flowering, length of fruit, days to maturity and yield/plant while, genotypes in cluster II have minimum values for width of fruit, fruits/plant and grains/plant. It indicates that representative lines can be chosen from particular diverse groups based on their cluster mean and can be involved in improvement programmes for Nigella improvement.

The character contribution of various clusters towards the genetic diversity by Tochers clustering method indicated that Grains per plant and Yield per plant were the major contributors towards total divergence (**Table 7).**

Intra and inter cluster distance was carried out on the basis of ten morphological traits in between eighteen genotypes (Table 8.). In general, inter-cluster were greater than the intra-cluster distances indicating the considerable amount of genetic diversity among the genotypes studied. The average intra-cluster distance between genotypes was maximum (1.48) for the cluster II followed by Cluster I (0.57). Cluster III, IV and V did not show intra-cluster distance indicating their genetic relatedness. The highest inter-cluster distance was noted

between cluster III and V (22.32) followed by cluster III and IV (19.54). The least inter-cluster distance was observed for cluster I and II (1.69) indicating their genetic relatedness.

3-D Plot diagram was constructed on the first three principle components (**Fig. 1**). Researchers use 3-D plot in principle component analysis to visually assess which components explain most of the variability in the data. In 3-D diagram, Rajendra shyama and RN-77 were plotted at distant end whereas, RN-78 and RN-20 were plotted at other end of 3-D plot indicating their effectiveness in breeding programme for improvement of Nigella.

**Conclusion:**

Based on above discussion, PCA analysis revealed the possibility for improvement of Nigella through various agro-morphological traits. In 3-D diagram (Fig.1) Rajendra shyama and RN-77 were found most divergent mutant lines with RN-78 and RN-20 which can be utilized effectively in breeding programme for improvement of Nigella. The genotype (RN-20) from cluster III have higher mean value for all the traits indicating their potentiality as a parent in hybridization programmes for further improvement of Nigella. The highest inter-cluster distance was noted between cluster III and V (22.32) indicating the genetic diversity among genotypes of these two clusters. Therefore, genotypes from these two clusters were recommended for use in hybridization programmes for further improvement.

**References:**


Anderson, T. W. 1972. An Introduction to Multivariate Analysis. Wiley Eastem Pvt. Ltd. New Delhi.

Atta, M.B. 2003. Some characteristics of Nigella (*Nigella sativa* L.) seed cultivated in Egypt and its lipid profile. *Food Chem*., **83**: 63-68.

Boskabady, M.H., B. Shermohammadi, P. Jandaghi and S.Kiani. 2004. Possible mechanism (s) for the relaxant effectof aqueous and macerated extracts from *Nigella sativa* ontracheal chains of guinea pig. *BMC Pharmacology*, **4**: 1-6.

Brejda, J. J., Moorman, T. B., Karlen, D. L. and Dao, T. H., 2000.Identification of regional soil quality factors and indicators. I. Central and Southern High- Plains. *Soil Sci. Soc. Am. J.***64**: 2115-2124.

Brutis, M. and F. Bucar. 2000. Antioxidant activity of *Nigella sativa* essential oil. *Phytother. Res.*, **14**: 323-328.

Datta, S., Sharanghi A. B., Pariari A., and Chatterjee R. R. 2001. Prospect of minor spices cultivation in West Bengal. Paper presented in VIII State Science Congress, 28 Feb-2nd–March, 2001, Kalyani Publisher, West Bengal, India.

El-Dakhakhny, M., N. Mady, N. Lembert and H.P. Ammon. 2002.The hypoglycemic effect of *Nigella sativa* oil is mediated by extrapancreatic actions. *PlantaMedica*., **68**: 465-466.

Fararh, K. M., Y. Atoji, Y. Shimizu and T. Takewaki. 2002.Isulinotropic properties of Nigella sativa oil instreptozotocin plus nicotinamide diabetic hamster. *Res.Veteria. Sci*., **73**: 279-282.

Fisher, C., 2002. Spices of life. *Chemistry in Britain*. **38**: 40-42.

Gunstone, F. D., J. L. Harwood and F. B. Padley. 1986. The Lipid Handbook. Chapman and Hall Ltd. The University Press, Cambridge.

Hajhashemi, V. A.,Ghannadi and H. Jafarabadi. 2004. Blackcumin seed essential oil, as a potent analgesic and anti-inflammatory drug. *Phytother. Res*., **18**: 195-199.

Hammo, Y.H. 2008. Effect of high levels of nitrogen and phosphorous fertilizer, pinching, and seed rate on growth and yield components of *Nigella sativa* L. 1-vegetative growth and seed yield. *Mesopotamia J of Agric*.**36**: 1-8.

Jansen, M.C. 1981. Spices, condiments and medicinal plants in Ethiopia, their taxonomy and agricultural significances. Center for Agricultural Publishing and Documentation, Wageningen. 294p.

Kanter, M., I. Meral, S. Dede, H. Gunduz, M. Cemek, H. Ozbekand I. Uygan. 2003. Effect of *Nigella sativa* L. and *Urtica dioica* L., on lipid peroxidation, antioxidant enzyme systems and some liver enzymes in CC14-treated rats. *J.Vet. Med*., **50**: 264-268.

Khan, N., S. Sharma and S. Sultana. 2003. *Nigella sativa* (black cumin) ameliorates potassium bromate-induced early events of carcinogenesis: diminution of oxidative stress. *Human Experi. Toxicol*.,**22**: 193-203.

Mahalanobis. 1936. On the generalized distance in statistics. *Proc. Nat. Inst. Sciences India*, **2**:49–55.

Malhotra, S.K. 2004. Minor seed spices 2 – Parsley, caraway, black, caraway and nigella. In FiftyYears of Spices Research in India. P.N. Ravindran (ed.) IISR, Calicut, India (in press).

Manonmani, S. and Khan, A. F.,2003. Analysis of genetic diversity for selection of parents in rice. *Oryza*. **40**(3/4): 54-56.

Mashhadian, N.V. and H. Rakhshandeh. 2005. Antibacterial andantifungal effects of *Nigellasativa* extracts against *S.aureus*, *P. aeroginosa* and *C. albicans*. *Pak. J. Medi. Sci*.,**21**(1): 47-52.

Morrison, D. E., Multivariate Statistical Methods 2nd ed. 4th Print. 1978. McGraw Hill Kogakusta Ltd.

Nachimuthu, V. V., Robin, S., Sudhakar, D., Raveendran, M., Rajeswari, S. and Manonmani, S., 2014. Evaluation of Rice Genetic Diversity and Variability in a Population Panel

by Principal Component Analysis. *Indian Journal of Science and Technology*. **7**(10):1555-1562.

Ramadan, M.F. and J.T. Morsel. 2003. Analysis of glycolipids from black cumin (*Nigella sativa* L.), coriander(*Coriandrum sativum* L.) and Niger (*Guizotia abyssinicacass*.) oil seeds. *Food Chemi*.,**80**: 197-204.

Shah, S. H. 2008. Effects of nitrogen fertilization on nitrate reductase activity, protein, and oil yields of *Nigella sativa* L. as affected by foliar GA3 application. *Turk J Bot*.**32**: 165-170.

Weiss EA. 2002. Spice Crops. CABI Publishing. CABI International, Wallingford, Oxon, UK.

Worthen, D.R., O.A. Ghosheh and P.A. Crooks. 1998. The Invitro anti tumor activity of some crude and purified components of black seed (*Nigella sativa* L.). *Anticancer Res.*, **18**: 1527-1532.

**Table 2. ANOVA for various characters of Nigella**

| Source of Variation | Replications(MSS) | Genotypes(MSS) | Error(MSS) |
|---|---|---|---|
| Degree of freedom | 2 | 17 | 34 |
| Height of the plant | 13.55 | 207.76** | 35.37 |
| Primary branches per plant | 0.36 | 5.02** | 3.04 |
| Secondary branches per plant | 0.96 | 204.87** | 16.66 |
| Days to 50 per cent flowering | 1.40 | 39.48** | 10.34 |
| length of fruit | 0.003 | 0.04* | 0.016 |
| Width of fruit | 0.011 | 0.080** | 0.028 |
| Days to maturity | 0.46 | 4.89** | 1.49 |
| Fruit per plant | 15.79 | 283.29** | 72.79 |
| Grains per plant | 19.79 | 1042.92** | 46.09 |
| Yield per plant | 0.28 | 12.03** | 1.48 |

**Table 3. Eigen values and variability explained by each principal components (PCs)**

| | PC1 | PC2 | PC3 | PC4 |
|---|---|---|---|---|
| **Eigene Value (Root)** | 5.547 | 1.714 | 1.208 | 0.698 |
| **% Var. Exp.** | 55.47 | 17.14 | 12.08 | 6.98 |
| **Cum. Var. Exp.** | 55.47 | 72.62 | 84.71 | 91.69 |

**Table 4. Rotated component matrix for various traits**

| Trait | PC1 | PC2 | PC3 | PC4 |
|---|---|---|---|---|
| Height of the Plant (cm) | 0.27047 | 0.40103 | 0.29174 | 0.40761 |
| Primary Branches/Plant | 0.35236 | 0.02250 | | |
| Secondary Branches/Plant | 0.41041 | | | |
| Days to 50%Flowering | 0.25405 | | 0.26469 | 0.02384 |
| Length of Fruit (cm) | 0.38104 | 0.07956 | 0.18150 | 0.32709 |
| Width of Fruit (cm) | | 0.31455 | 0.33182 | |
| Days to Maturity | | | 0.67470 | |
| Fruit/Plant | 0.37405 | | 0.04416 | 0.33129 |
| Grains/Plant | 0.17236 | 0.51885 | 0.39816 | |
| Yield/ Plant (in G) | 0.39107 | 0.06985 | | |

**Table 5. Distribution of Nigella genotypes in various clusters**

| Group K | No. of genotypes Within clusters | Genotypes within clusters |
|---|---|---|
| I | 9 | RN-25, RN-65, RN-66, RN-69, RN-70, RN-73, RN-74 RN-76 & RN-79 |
| II | 5 | RN-22, RN-68, RN-71, RN-72, RN-75 |
| III | 1 | RN-20 |
| IV | 1 | RN-78 |
| V | 2 | RN-77 & Rajendra Shyama |

**Table 6. Mean characteristics (K-Mean) on various traits for each cluster in Nigella genotypes**

| Cluster | Height of the plant | Primary branches/ plant | Secondary branches/ plant | Days to 50% flowering | Length of fruit | Width of fruit | Days to maturity | Fruits/ plant | Grains/ plant | Yield/plant |
|---|---|---|---|---|---|---|---|---|---|---|
| **I** | 37.852 | 5.148 | 16.056 | 42.481 | 1.131 | 0.917 | 77.444 | 38.500 | 47.444 | 5.471 |
| **II** | 40.200 | 5.567 | 16.100 | 44.633 | 1.190 | 0.875 | 83.267 | 37.900 | 39.167 | 5.544 |
| **III** | 76.333 | 10.333 | 31.667 | 69.167 | 2.072 | 1.655 | 138.000 | 68.833 | 108.50 | 12.677 |
| **IV** | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| **V** | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 7.Contribution percentage of traits towards genetic divergence**

| S. No. | Source | **Times Ranked 1st** | **Contribution (%)** |
|---|---|---|---|
| 1. | Height of the plant | 10 | 6.54% |
| 2. | Primary branches per plant | 0 | 0.00% |
| 3. | Secondary branches per plant | 2 | 1.31% |
| 4. | Days to 50 per cent flowering | 6 | 3.92% |
| 5. | length of fruit | 2 | 1.31% |
| 6. | Width of fruit | 4 | 2.61% |
| 7. | Days to maturity | 4 | 2.61% |
| 8. | Fruit per plant | 1 | 0.65% |
| 9. | Grains per plant | 57 | 37.25% |
| 10. | Yield per plant | 67 | 43.79% |

**Table 8. Estimates of Intra (diagonal) and Inter-cluster distances in 18 genotypes of Nigella**

| Cluster | I | II | III | IV | V |
|---|---|---|---|---|---|
| I | **0.57** | 1.69 | 2.64 | 10.94 | 11.29 |
| II | | **1.48** | 4.53 | 8.06 | 11.32 |
| III | | | **0.00** | 19.54 | 22.32 |
| IV | | | | **0.00** | 5.63 |
| V | | | | | **0.00** |

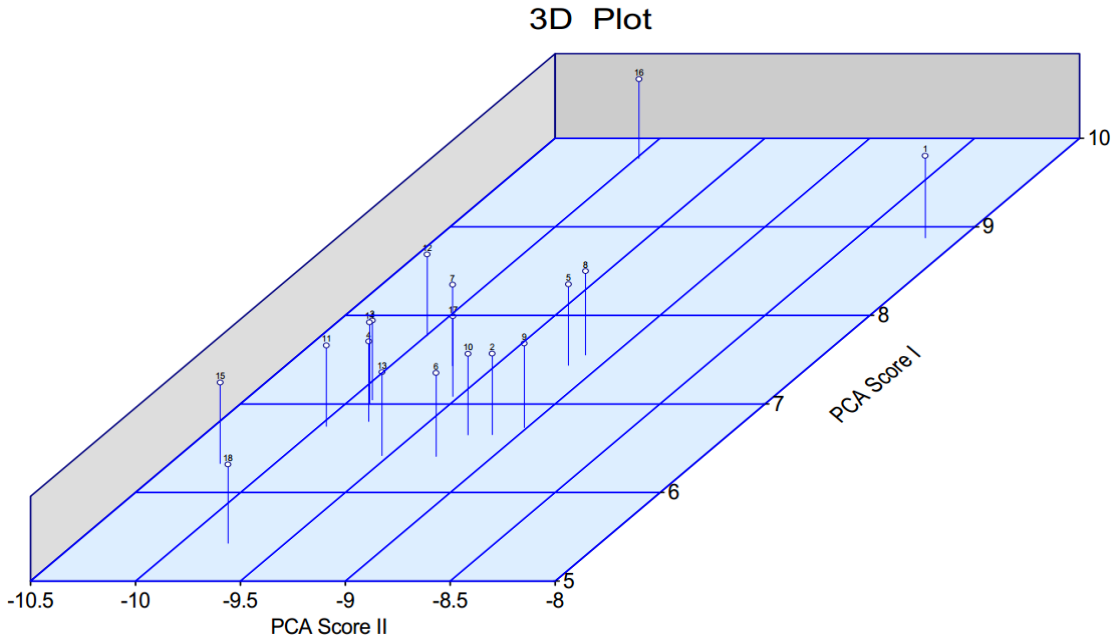


Fig. 1: 3-D distribution of the diverse genotype based on principal components.

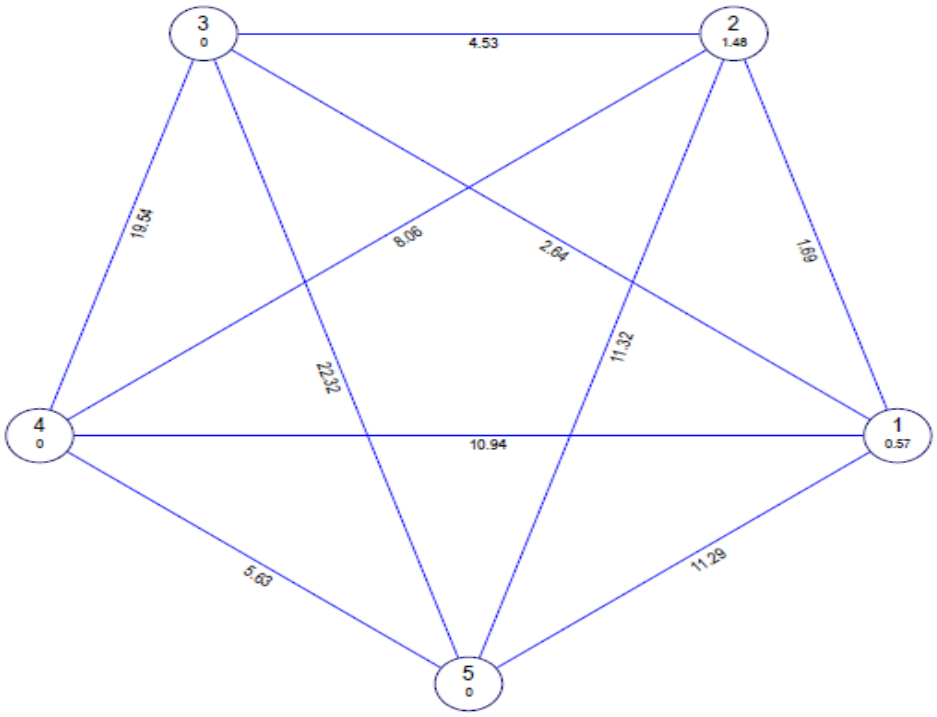


Fig.2: Mahalonobis Eucladean Distance (Not to the Scale) by Tocher Method