

# **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 **Kkalongi** 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<sup>0</sup>C with the optimum of 12 -14<sup>0</sup>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

Comment [w1]: should mention objective of study in abstract

Formatted: Font: Italic

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

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

60 Genetic diversity is pre-requisite for any crop improvement programme, as it helps in  
61 the development of superior recombinant (Manonmani & Fazlullah Khan., 2003). Genetic  
62 distance estimates for population grouping can be estimated by different methods as it is  
63 crucial to understand the usable variability existing in the population panel (Nachimuthu  
64 *et al.*, 2014). One of the approaches is to apply multivariate analysis. Principal Component Analysis  
65 (PCA) is a powerful tool in modern data analysis because this is a well-known multivariate  
66 statistical technique which is used to identify the minimum number of components, which  
67 can explain maximum variability out of the total variability (Anderson, 1972 and Morrison,  
68 1978) and also to rank genotypes on the basis of PC scores. Principal components are  
69 generally estimated either from correlation matrix or covariance matrix. Considering the  
70 importance of PCA this study is conducted on diverse *Nigella* genotypes for identification of  
71 traits responsible for yield differences.

Comment [w2]: where is objective of this study in this part

72  
73

#### 74 **Material and Methods:**

75 Seventeen land races of *Nigella* along with one released variety (Rajendra Shyama) as a check,  
76 collected from different parts of Bihar (**Table 1**) were evaluated in Randomized Block Design  
77 with three replications at Seed production Farm, TCA, Dholi, Bihar during *Rabi* 2015-16. The  
78 experimental site is located at 25.59 N latitude and 85.75 E longitudes and has altitude of  
79 51.20 m above mean sea level. Soil of TCA, Dholi, Bihar is mainly young alluvium and  
80 calcareous. Standard agronomic practices were adopted with row to row and plant to plant  
81 spacing of 25×5 cm., recommended dose of fertilizer was applied during the time of crop  
82 period. Data was recorded for ten different traits *viz.*, height of the plant, primary branches  
83 per plant, secondary branches per plant, days to 50 per cent flowering, length of fruit, width

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

87 **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)	

88

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

#### 94 **Results and Discussion:**

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

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

Comment [w3]: discussion part is completely lacking

106 The first principal component (PC<sub>1</sub>) explained 55.47% of the total variation. The second  
107 principal component (PC<sub>2</sub>) explained 17.14% variation individually and 72.62% cumulative  
108 variation. The third principal component (PC<sub>3</sub>) explained 12.08% variation individually and  
109 84.71% cumulatively.

110 Rotated component matrix for various traits revealed that PC<sub>1</sub> was strongly associated  
111 with secondary branches/plant followed by yield/plant, length of fruit, fruit per plant, primary  
112 branches/plant, height of the plant, days to 50% flowering and grains/plant. The traits that  
113 mostly contributed to PC<sub>2</sub> were grains/plant followed by height of the plant and width of fruit  
114 whereas, days to maturity followed by width of fruit, height of the plant, days to 50%  
115 flowering and length of fruit contributed mostly to the PC<sub>3</sub>. The characters that contributed  
116 most to the PC<sub>4</sub> were height of the plant, fruit/plant and length of fruit. Therefore, intensive  
117 selection procedures can be adopted to bring about rapid improvement of above mentioned  
118 traits.

#### 119 **Cluster analysis:**

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

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

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

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

148 3-D Plot diagram was constructed on the first three principle components (**Fig. 1**).  
149 Researchers use 3-D plot in principle component analysis to visually assess which  
150 components explain most of the variability in the data. In 3-D diagram, Rajendra shyama and  
151 RN-77 were plotted at distant end whereas, RN-78 and RN-20 were plotted at other end of 3-  
152 D plot indicating their effectiveness in breeding programme for improvement of Nigella.

153

#### 154 **Conclusion:**

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

#### 164 **References:**

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

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

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

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

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

177 Datta, S., Sharanghi A. B., Pariari A., and Chatterjee R. R. 2001. Prospect of minor spices  
178 cultivation in West Bengal. Paper presented in VIII State Science Congress, 28 Feb-  
179 2nd-March, 2001, Kalyani Publisher, West Bengal, India.

180 El-Dakhkhny, M., N. Mady, N. Lembert and H.P. Ammon. 2002. The hypoglycemic effect  
181 of *Nigella sativa* oil is mediated by extrapancreatic actions. *Planta Medica.*, **68**: 465-  
182 466.

- 183 Fararh, K. M., Y. Atoji, Y. Shimizu and T. Takewaki. 2002. Insulinotropic properties of  
184 Nigella sativa oil instreptozotocin plus nicotinamide diabetic hamster. *Res.Veteria.*  
185 *Sci.*, **73**: 279-282.
- 186 Fisher, C., 2002. Spices of life. *Chemistry in Britain*. **38**: 40-42.
- 187 Gunstone, F. D., J. L. Harwood and F. B. Padley. 1986. The Lipid Handbook. Chapman and  
188 Hall Ltd. The University Press, Cambridge.
- 189 Hajhashemi, V. A., Ghannadi and H. Jafarabadi. 2004. Blackcumin seed essential oil, as a  
190 potent analgesic and anti-inflammatory drug. *Phytother. Res.*, **18**: 195-199.
- 191 Hammo, Y.H. 2008. Effect of high levels of nitrogen and phosphorous fertilizer, pinching,  
192 and seed rate on growth and yield components of *Nigella sativa* L. 1-vegetative  
193 growth and seed yield. *Mesopotamia J of Agric.* **36**: 1-8.
- 194 Jansen, M.C. 1981. Spices, condiments and medicinal plants in Ethiopia, their taxonomy and  
195 agricultural significances. Center for Agricultural Publishing and Documentation,  
196 Wageningen. 294p.
- 197 Kanter, M., I. Meral, S. Dede, H. Gunduz, M. Cemek, H. Ozbek and I. Uygan. 2003. Effect of  
198 *Nigella sativa* L. and *Urtica dioica* L., on lipid peroxidation, antioxidant enzyme  
199 systems and some liver enzymes in CC14-treated rats. *J.Vet. Med.*, **50**: 264-268.
- 200 Khan, N., S. Sharma and S. Sultana. 2003. *Nigella sativa* (black cumin) ameliorates  
201 potassium bromate-induced early events of carcinogenesis: diminution of oxidative  
202 stress. *Human Experi. Toxicol.*, **22**: 193-203.
- 203 Mahalanobis. 1936. On the generalized distance in statistics. *Proc. Nat. Inst. Sciences India*,  
204 **2**: 49-55.
- 205 Malhotra, S.K. 2004. Minor seed spices 2 – Parsley, caraway, black, caraway and nigella. In  
206 Fifty Years of Spices Research in India. P.N. Ravindran (ed.) IISR, Calicut, India (in  
207 press).
- 208 Manonmani, S. and Khan, A. F., 2003. Analysis of genetic diversity for selection of parents in  
209 rice. *Oryza*. **40**(3/4): 54-56.
- 210 Mashhadian, N.V. and H. Rakhshandeh. 2005. Antibacterial and antifungal effects of  
211 *Nigella sativa* extracts against *S.aureus*, *P. aeruginosa* and *C. albicans*. *Pak. J.*  
212 *Medi. Sci.*, **21**(1): 47-52.
- 213 Morrison, D. E., Multivariate Statistical Methods 2nd ed. 4th Print. 1978. McGraw Hill  
214 Kogakusta Ltd.
- 215 Nachimuthu, V. V., Robin, S., Sudhakar, D., Raveendran, M., Rajeswari, S. and Manonmani,  
216 S., 2014. Evaluation of Rice Genetic Diversity and Variability in a Population Panel

217 by Principal Component Analysis. *Indian Journal of Science and Technology*.  
218 7(10):1555-1562.

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

222 Shah, S. H. 2008. Effects of nitrogen fertilization on nitrate reductase activity, protein, and oil  
223 yields of *Nigella sativa* L. as affected by foliar GA3 application. *Turk J Bot.***32**: 165-  
224 170.

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

227 Worthen, D.R., O.A. Ghosheh and P.A. Crooks. 1998. The In vitro anti tumor activity of some  
228 crude and purified components of black seed (*Nigella sativa* L.). *Anticancer Res.*,  
229 **18**: 1527-1532.

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256 **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

257

258

259

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

	PC1	PC2	PC3	PC4
<b>Eigene Value (Root)</b>	5.547	1.714	1.208	0.698
<b>% Var. Exp.</b>	55.47	17.14	12.08	6.98
<b>Cum. Var. Exp.</b>	55.47	72.62	84.71	91.69

261

262 **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		

274

275

276

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

278	<b>Group K</b>	<b>No. of genotypes</b>	<b>Genotypes within clusters</b>
279		<b>Within clusters</b>	
280			
281	I	9	RN-25, RN-65, RN-66, RN-69, RN-70, RN-73, RN-74
282			RN-76 & RN-79
283	II	5	RN-22, RN-68, RN-71, RN-72, RN-75
284	III	1	RN-20
285	IV	1	RN-78
286	V	2	RN-77 & Rajendra Shyama
287			

288 **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

289

290 **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%

291

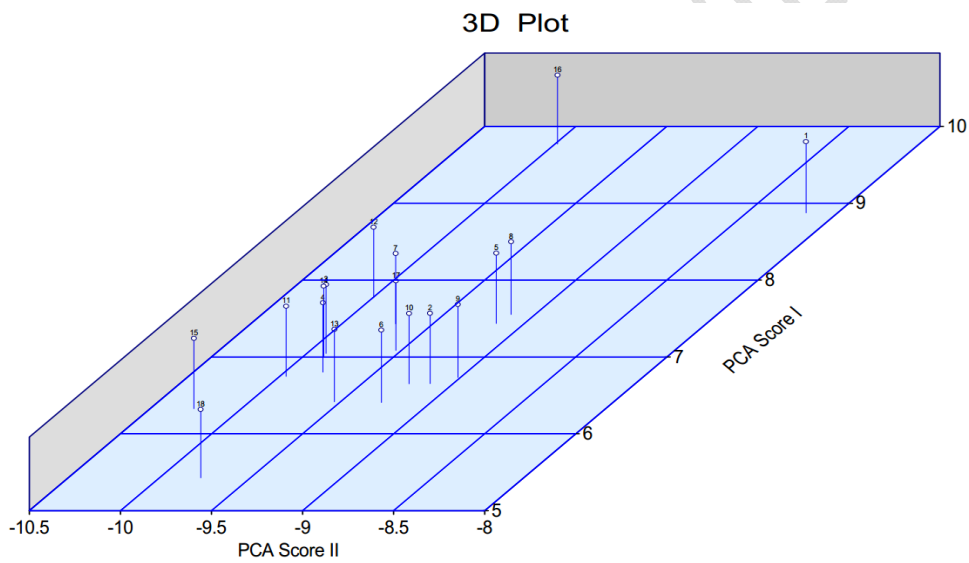
292 **Table 8. Estimates of Intra (diagonal) and Inter-cluster distances in 18 genotypes of**

293 **Nigella**

294 Cluster	I	II	III	IV	V
295 I	<b>0.57</b>	1.69	2.64	10.94	11.29
296 II		<b>1.48</b>	4.53	8.06	11.32
297 III			<b>0.00</b>	19.54	22.32
298 IV				<b>0.00</b>	5.63
299 V					<b>0.00</b>

300

301



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

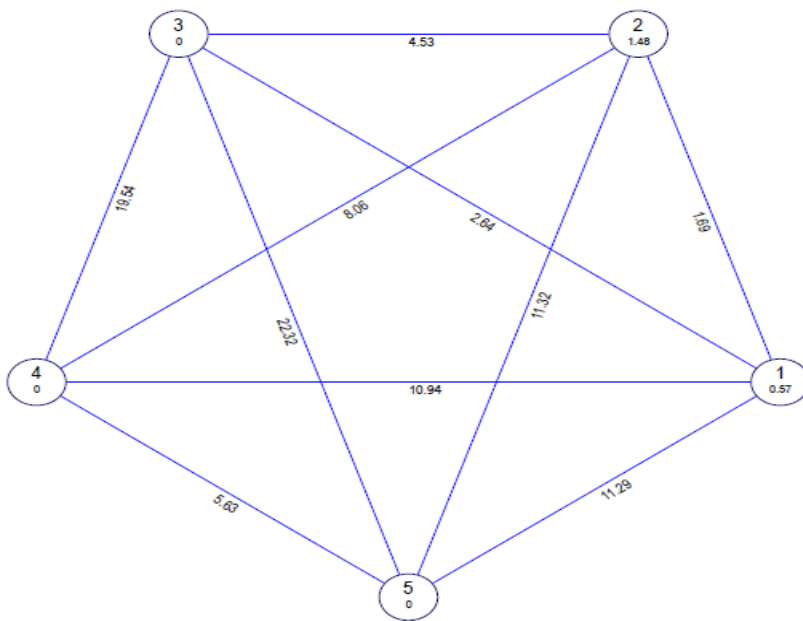


Fig.2: Mahalanobis Euclidean Distance (Not to the Scale) by Tocher Method

304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319

UNDER PEER REVIEW