한작지(Korean J. Crop Sci.), 54(1): 24~28(2009)

Interpretation of Agronomic Traits Variation of Sesame Cultivar Using Principal Component Analysis

Kang-Bo Shim*†, Chung-Dong Hwang**, Suk-Bok Pae**, Jang-Whan Park**, Jae-Cheon Byun**, and Keum-Yong Park**

**International Technical Cooperation Center, Technical Cooperation Bureau, RDA, Suwon 441-707, Korea*
***Legume and Oil Crop Research Division, Department of Functional Crop, NICS, Rural Development Administration, Miryang 627-803, Korea*

ABSTRACT This study was conducted to evaluate the growth characters and yield components of 18 collected sesame cultivars to get basic information on the variation for the sesame breeding using principal component analysis. All characters except days to flowering, days to maturity and 1,000 seed weight showed significantly different. Seed weight per 10 are showed higher coefficient of variance. Capsule bearing stem length and liter weight showed positive correlation with seed yield per 10 are. The principal components analysis grouped the estimated sesame cultivars into four main components which accounted for 83.7% of the total variation at the eigenvalue and its contribution to total variation obtained from principal component analysis. The first principal component (Z_1) was applicable to increase plant height, capsule bearing stem length and 1,000-seed weight. The second principal component (Z_2) negatively correlated with days to flowering and maturity by which it was applicable to shorten flowering and maturity date of sesame. At the scatter diagram, Yangbaek, Ansan, M1, M2, M4, M7 and M9 were classified as same group, but M10, Yanghuk, Kanghuk, M5, M6, M12 and M13 were classified as different group. This results would be helpful for sesame breeder to understand genetic relationship of some agronomic characters and select promising cross lines for the development of new sesame variety.

Keywords : sesame, principal component analysis, eigenvalue, degree of contribution

Different statistical techniques have been used in modeling crop yield, including correlation, regression and principal component analysis. Correlation coefficient is an important stastical procedure to evaluate breeding program for high yield, as well as to examine direct and indirect contribution of yield variables (Mohamed, 1999). The principal component analysis is a multivariate statistical technique for exploration and simplifying complex data sets (Everitt *et al*., 1992). Each principal component is a linear combination of the original variables, and so it is often possible to ascribe the meaning to what the components represent. Lee *et al*. (1982) and Park *et al*. (1982) classified sesame varieties into several groups on the sensitivity of day length as one group with higher sensitivity to both short and long day, another group with only higher sensitivity to short day, the other group with lower sensitivity to both short and long day. Ahn *et al*. (1984) reported that total fourteen sesame varietal groups were classified by Mahalanobias' distance (D^2), cluster, principal component analysis to identify their affinity. The purpose of this study is to provide theoretical foundations to help sesame breeders who are studying genetic correlation of the main agronomic characters and their influence in sesame productivity.

MATERIALS AND METHODS

The experiment were conducted at the Yeongnam Agricultural Research Institute, Miryang, using total eighteen promising sesame cultivars in 2005 and 2006 (Table 1). Sowing date was May 15 and a randomized complete block design with three replications was used for this experiment. Spacing between plants was 10 cm and row spacing was 30 cm in a 70 cm wide black polyethylene film mulching bed. Days to flowering was from sowing day to the time

†Corresponding author: (Phone) +82-55-350-1241
(E-mail) shimkb@rda.go.kr <Received July 8, 2008>

Table 1. Basic agronomic characters statistics of eighteen sesame cultivars

Variety	SCC†	DTF	DTM	PH (cm)	CBSL	NCP	TSW (g)	LW (g/L)	SW (kg/10a)
Yangbaek	White	44	86	129	107	82	2.86	635	97
Ansan	White	44	85	131	113	72	2.76	647	81
Kopoom	White	45	86	132	110	80	2.65	619	77
S195	White	45	86	126	100	87	2.59	626	77
Kyeonbuk2	White	48	90	138	112	75	2.74	646	71
M1	White	45	85	132	112	65	2.80	643	85
M2	White	45	86	136	117	74	2.77	640	104
M4	White	45	86	135	114	76	2.73	637	74
M7	White	45	86	133	119	95	2.74	637	104
M9	White	45	86	133	121	74	2.65	631	92
M10	White	47	86	130	109	83	2.67	624	72
Yanghuk	Black	45	85	129	104	62	2.73	502	68
Kanghuk	Black	46	86	127	103	78	2.73	493	80
M5	Black	46	88	127	112	83	2.59	550	75
M6	Black	46	87	129	110	58	2.89	523	62
M12	Black	46	88	122	104	95	2.64	548	66
M13	Black	45	88	123	105	81	2.64	562	76
Kyeongbuk4	Black	47	90	144	125	61	3.02	655	77

†SCC: Seed coat color, DTF: Days to flowering, DTM: Days to maturity, PH: Plant height, CBSL: Capsule bearing stem length, NCP: Number of capsule per plant, TSW: 1,000-seed weight, LW: Liter weight, SW: Seed weight per 10a.

that one of buds in plant was bursting, and days to maturity was from sowing day to the time that one of capsules in plant was dehiscent.

Variance of some agronomic characters were analyzed. Correlation coefficient and principal component analysis using SAS program (proc princomp) was conducted to identify relationship of some agronomic characters of sesame.

RESULTS AND DISCUSSION

Simple correlation analysis of agronomic characters

Table 2 showed mean values, standard deviation and F values for all estimated variables of sesame. All variables except days to flowering, days to maturity and 1,000 seed weight showed significantly different. Seed weight per 10

Table 2. Statistic data for the estimated agronomic characters of sesame cultivars

	DTF†	DTM	PH (cm)	CBSL	NCP	TSW (g)	LW (g/L)	SW (kg/10a)
Mean	45.1	86.5	130.9	110.9	76.7	2.73	601.0	79.8
Maximum	47.5	89.5	144.0	125.0	95.0	3.02	493.0	104.0
Minimum	43.5	85.0	122.0	100.0	58.0	2.59	655.0	62.0
S. D.	1.04	1.53	5.33	6.66	10.58	0.11	54.74	12.18
C. V.	2.38	1.60	4.08	5.99	13.8	4.03	9.11	15.25
F value	0.53	0.69	2.65*	3.31*	5.26*	0.05	27.22*	6.06*

†DTF: Days to flowering, DTM: Days to maturity, PH: Plant height, CBSL: Capsule bearing stem length, NCP: Number of capsule per plant, TSW: 1,000-seed weight, LW: Liter weight, SW: Seed weight per 10a.
*Significant at the 5% level

Table 3. A matrix of simple correlation coefficients for the estimated variables of sesame cultivar

Traits	X1	X2	X3	X4	X5	X6	X7
Days to flowering (X1)	1.000						
Days to maturity (X2)	0.735**	1.000					
Plant height (X3)	0.307*	0.276	1.000				
Capsule bearing stem length (X4)	0.123	0.229	0.807**	1.000			
Number of capsule per plant (X5)	-0.093	0.001	-0.459*	-0.274	1.000		
1,000-seed weight (X6)	0.118	0.202	0.618*	0.467*	-0.634*	1.000	
Liter weight (X7)	-0.083	0.030	0.627*	0.578*	0.078	0.205	1.000
Seed weight per 10a (Y)	-0.048	-0.329*	0.267	0.433*	0.252	0.108	0.478*

*, **Significant at the 5%, 1% level respectively. ns: Not significant

Table 4. Eigenvalue and its contribution to total variation obtained from principal component analysis.

Principal component	Z_1 †	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7
Eigenvalue	3.135	2.197	1.364	0.473	0.360	0.253	0.142
Proportion (%)	39.19	27.46	17.04	5.92	4.50	3.17	1.78
Cumulative contribution (%)	39.19	66.65	83.69	89.61	94.12	97.28	99.06

† Z_1~Z_7: the 1st~7th principal component eigenvectors

are showed higher coefficient of variance.

Table 3 showed simple correlation coefficients of estimated variables with each other. All characters had positive correlation with seed weight per 10 are, except days to flowering, plant height, number of capsule per plant and 1,000 -seed weight. In detail, capsule bearing stem length (0.48) and liter weight (0.43) positively correlated with seed weight, but days to flowering (-0.05) and days to maturity (-0.33) negatively correlated with seed weight per 10 are.

Similar results were also observed by Yingzhong. Z. *et al*. (2002), Uzo (1985) and Singh *et al*. (1995).

Table 4 and Figure 1 demonstrated that an increase in the number of principal components was associated with a decrease in eigenvalues. This trend reached its maximum at the three factors. Therefore, it is reasonable to assume that the principal components analysis had grouped the estimated sesame variables into three main components which accounted for 83.7% of the total variation.

Table 5 showed that the first principal component (Z_1) positively correlated with plant height, capsule bearing stem length and 1,000 seed weight. Meanwhile, the second principal component (Z_2) negatively correlated with days to flowering and days to maturity. The third component (Z_3) showed highly positive corelation with number of capsule per plant. Variables which significantly correlated with the first three eigenvectors were those with the greatest variability. Data in table 4 showed three main principal components account for 83.7% of the total variation. Therefore, plant height, capsule bearing stem length, 1,000-seed weight and number of capsule per plant were shown to be important variables affecting sesame grain yield increasement.

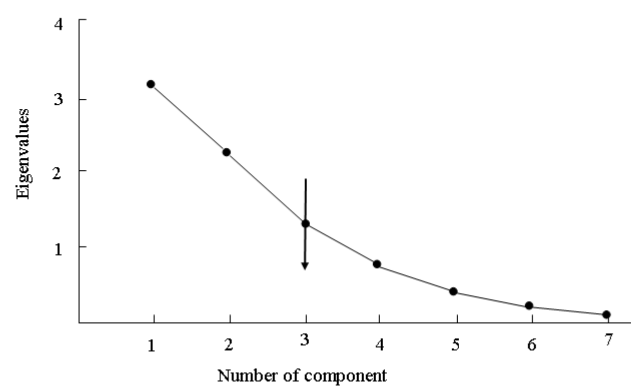

Fig. 1. Screeplot showing eigenvalues in response to the number of components for the estimated variables of sesame.

Table 6 showed that plant height, number of capsule per plant and seed weight per plant corresponded to same response,

Table 5. Correlation coefficients between agronomic characters and principal components.

Traits	Principal component eigenvectors						
	Z_1 †	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7
Days to flowering	0.153	-0.548*	0.319	-0.079	-0.178	0.631*	0.039
Days to maturity	0.192	-0.475*	0.410	0.362	0.179	-0.496*	-0.389*
Plant height	0.539*	-0.003	0.024	-0.228	-0.100	0.247	-0.103
Capsule bearing stem length	0.492*	0.122	0.120	0.037	-0.573*	-0.393*	0.438*
Number of capsule per plant	0.275	0.201	0.658**	0.242	0.149	0.124	0.516*
1,000-seed weight	0.415*	-0.091	-0.391*	0.499*	0.495*	0.158	0.379*
Liter weight	0.358	0.313	0.325	-0.503*	0.546*	-0.075	-0.061
Seed weight per 10a	0.175	0.558*	0.154	0.499*	-0.178	0.309	-0.481*

*, **Significant at the 5%, 1% level respectively

Table 6. Classification of characters by the degree of contribution to the first six principal components.

Principal component	Class	Corresponding traits
Z_1	+	Plant height, capsule bearing stem length, 1,000 seed weight, liter weight
	−	Number of capsule per plant
Z_2	+	Seed weight per 10a, liter weight, number of capsule per plant
	−	Days to flowering, days to maturity, plant height, 1,000-seed weight
Z_3	+	Number of capsule per plant, days to maturity, liter weight
	−	1,000-seed weight
Z_4	+	Seed weight per 10a, 1,000-seed weight, days to maturity
	−	Liter weight, plant height, days to flowering
Z_5	+	Liter weight, 1,000-seed weight
	−	Capsule bearing stem length, seed weight per 10a, days to flowering
Z_6	+	Days to flowering, seed weight per 10a
	−	Days to maturity, Capsule bearing stem length, liter weight

and days to maturity, days from flowering to maturity and days to flowering corresponded to another response at the classification of characters by the degree of contribution to the first six principal components.

Table 7 showed Eigenvectors of eighteen sesame cultivars on the first four principal components. In pervious table, 1st principal component (Z_1) accounted for about 39.19% of the variation in grain yield; Z_2 for 27.46%, and Z_3 for 17.04%. Therefore, it will be helpful if we understand main reaction direction of eighteen sesame cultivars on the first three principal components eigenvectors. Kyeongbuk2 and Kyeonbuk4 showed higher positive scores at the first component, but lower negative scores at the second component. On the other hand, Suwon195 and M12 showed different direction. Both of them were shown higher negative scores at the first component, but opposite direction; one is positive, another is negative.

Figure 2 showed overall direction of those eighteen sesame cultivars on the first two principal components. At the scatter diagram of the first two principal components eigenvectors by the principal component analysis, four distinct groups were identified.

In the group I, Yangbaek, Ansan, M1, M2, M4, M7 and M9 were included. Group II contained S195 and Kyeonbuk2. In the group III, M10, Yanghuk, Kanghuk, M5, M6, M12 and M13 which are black seed coating cultivars were included. Those results could be applied to the interpretation of the affinity of 18 sesame cultivars for the cross materials selection or determination of breeding target in the development of sesame variety.

Table 7. Eigenvectors of eighteen sesame cultivars on the first four principal components

Entry	Cultivar	Eigenvectors			
		1st component	2nd component	3rd component	4th component
1	Yangbaek	0.013	1.281	-0.361	1.595
2	Ansan	0.155	1.043	-0.829	-0.927
3	Kopoom	-0.247	0.373	0.194	-1.103
4	S195	-1.212	0.402	0.608	-1.067
5	Kyeonbuk2	1.028	-1.699	1.483	-0.707
6	M1	0.459	0.671	-1.051	-0.785
7	M2	0.914	1.230	0.032	0.621
8	M4	0.362	0.303	-0.130	-1.282
9	M7	0.443	1.531	1.245	1.363
10	M9	0.523	0.942	0.272	-0.544
11	M10	-0.261	-0.456	0.765	-1.339
12	Yanghuk	-0.811	-0.497	-2.032	-0.351
13	Kanghuk	-0.983	-0.531	-0.628	1.364
14	M5	-0.668	-0.629	0.897	0.308
15	M6	0.093	-1.439	-1.878	0.558
16	M12	-1.437	-0.886	1.133	0.774
17	M13	-1.083	-0.336	0.324	0.773
18	Kyeongbuk4	2.713	-1.301	-0.042	0.750

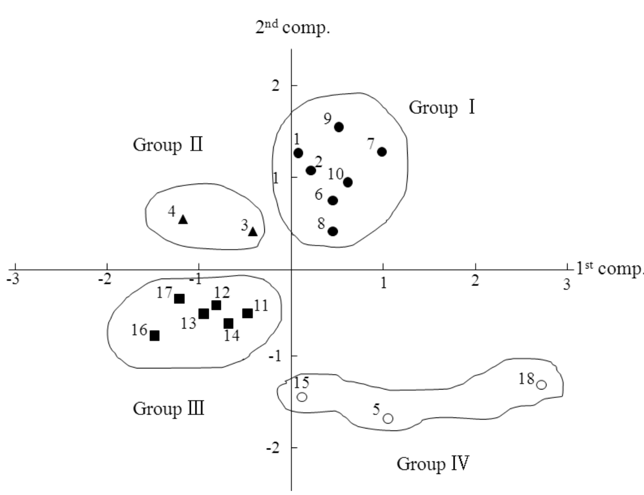

Fig. 2. Scatter diagram of the first two principal components eigenvectors of of 18 sesame cultivars by the principal component analysis. The number are shown in Table 7.

LITERATURE CITED

Ahn S. N. and Y. A. Chae. 1984. Varietal Classification on the basis of multivariate analysis in sesame. Korean J. Breed. 16(3) : 340-348.

Everitt, B. S., G Dunn. 1992. Applied multivariate data analysis. Oxford University Press. New York, NY.

Mohamed, N. A., 1999. Some statistical procedures for evaluation of the relative contribution for yield in sugar beet using some statistical techniques. Bulletin of the Faculty of Agriculture, University of Cairo 44(1) : 253-266.

Lee, J. I., S.T. Lee, G.C. Um, and C. H. Park. 1982. Effect of planting date on agronomic characteristics and varietal differences in sesame varieties. Korean. J. Crop Sci. 27(3) : 268-275.

Park, C. H. and J. I. Lee. 1982. Response of flowering in sesame varieties. Educational-industrial cooperation. 180(29) : 1-48.

SAS Institute, 2006. The SAS Enterprise Guide. SAS Institute, Inc., Cary, NC.

Sigh, P. K., R. K. Dixit, and R. K. Yadav. 1997. Estimates of genetic parameters, character association and path analysis in sesame. Crop Res., 13(10) : 115-119.

Uzo, J. O. 1985. Yield, yield components and nutritional attributes of cultivated sesame, S. indicum and its endemic wild relatives in Nigeria. FAO Pl. Prod. and Prot. Paper:Sesame and Safflower: status and potentials, pp. 166-176.

Yingzhong, Z. and W. Yishou. 2002. Genotype correlations and path coefficient analysis in sesame. Sesame and safflower newsletter No. 17 : 10-12.