International Journal of Genetics and Genomics
2022; 10(1): 32-36
http://www.sciencepublishinggroup.com/j/ijgg
doi: 10.11648/j.ijgg.20221001.15
ISSN: 2376-7340 (Print); ISSN: 2376-7359 (Online)

Evaluation of Sunflower Genotypes Using Principal Component Analysis

Mohammed Abu

Ethiopian Institute of Agricultural Research, Holetta Research Centre, Addis Ababa, Ethiopia

Email address:
moabu1440@gmail.com

To cite this article:
Mohammed Abu. Evaluation of Sunflower Genotypes Using Principal Component Analysis. *International Journal of Genetics and Genomics*. Vol. 10, No. 1, 2022, pp. 32-36. doi: 10.11648/j.ijgg.20221001.15

Received: January 29, 2022; **Accepted**: February 21, 2022; **Published**: February 25, 2022

Abstract: Evaluation of genetic resources using morphological, physiological and biochemical data is important for effective breeding program. Principal component analysis is one of the multivariate technique used genetic resources evaluation using bi plot diagrams. The present study was conducted to evaluate sunflower genotypes for genetic diversity using multivariate analysis particularly Principal component analysis. The study was conducted during 2017/18 at central highlands of Ethiopia using 25 sunflower genotypes. The genotypes were planted using lattice design with two replication in the main season at Holetta and Adadi. The data for fifteen quantitative traits; ray floret number, leaf number, petiole length, seed yield per plant, number of seed per plant, seed yield per hectare, oil yield, oil content, head diameter, stem diameter, plant height, days to flowering, days to maturity, seed filling percentage and hundred seed weight were collected and principal component analysis was done using SAS 9.3. Eigen value greater than one was observed for the first five principal components. The first five principal components extracted showed 84.72% of total variation. The first and the second principal components contributed more than half of the total variation. The first principal component attributes 31.9% of total variation whereas, the second, the third, the fourth and the fifth principal components contributes, 22.72%, 12.25%, 10.11%, and 7.75% respectively. Different traits contribute chiefly to different principal components. Among all traits studied days to maturity and seed filling percentage contributed to the variation in three principal components out of the total principal components. The results from this study showed that there is considerable variation for the traits studied in sunflower genotypes suggesting that there is an opportunity for genetic improvement through selection directly from genotypes and or their parents.

Keywords: Genetic Diversity, Genotypes, Principal Component Analysis, Sunflower, Ethiopia

1. Introduction

Sunflower (*Helianthus annus* L.) is an oil seed crop originated from North America from where it was disseminated to other continents [1]. It is an annual crop of asteraceae family cultivated for several purposes edible oil production, animal feed, raw material for biodiesel production, ornamental purpose and medicinal value [2]. Sunflower seeds are sources of proteins, vitamins, minerals and phytochemicals [3]. It is also used in crop rotation. Sunflower has wide distribution because of its high adaptation capability. In Ethiopia it is adapted from 800m.a.s.l. to 2600m a.s.l. breeding programs in Ethiopia focus on developing improved varieties with high seed yield, oil content, early mature, medium height with good head inclination and head size. To achieve this goal development, evaluation of genotypes and estimation of genetic diversity within gene pool is important.

Many methods were used by different researchers in estimating genetic diversity and in evaluation of germplasm in sunflower [4]. Among them Principal component analysis is one of the methods used. Principal component analysis transforms the original variables into anew sets of uncorrelated variables known as principal components [5]. It is used to identify the most significant variables in the data set. The results of principal component analysis is of greater benefit to identify the parents for improving various traits or characters or component and it can also be exploited in planning and execution of future breeding program [6]. The use of multivariate tool is vital to identify genetic diversity in existing germplasm. The application of principal component will be of great value to identify genetic diversity and

evaluation of germplasm for different characters to be exploited in future breeding [6].

Principal component analysis (PCA) is a reliable tool for successfully selection of parents in breeding program of any crop [7]. Some germplasms of sunflower which differs in different characters are available in Ethiopia but limited research work has been done in evaluation of those available germplasms using multivariate tools. Taking into account the importance of principal component analysis the present study was conducted to evaluate genetic diversity among sunflower genotypes and grouping them on their similarity based on different traits using principal component analysis.

2. Materials and Methods

The study was conducted at Holeta Agriculture Research center (HARC) and Adadi testing site of HARC during 2017/18. Description of the study area is presented in Table 1.

Table 1. *Description of the study area.*

Descriptor	Locations	
	Holletta	**Adadi**
Latitude	9°00' N	08°31' N
Longitude	38°30' E	38°13' E
Altitude (m a. s. l.)	2400	2383
Mean annual rain fall (mm)	1144 mm	1105
Mean annual temperature (°C)	Min=6°C Max=22°C	16.9
Soil type	Nitosols	Light brown
Soil drainage	Well drained	Well drained
Soil pH	6.0	7.62
Organic C (%)	1.18	1.16
N (%)	0.16	0.15

Table 2. *Experimental materials used for the study.*

Entry no.	Genotype	Origin	Source	Status
1	Adadi-3-SPS-2/4	India	HARC	PVT
2	Adadi-3-SPS-5/4	India	HARC	PVT
3	Adadi-3-SPS-9/4	India	HARC	PVT
4	NK-FERTI-SPS-1/4	France	HARC	PVT
5	NK-FERTI-SPS-4/4	France	HARC	PVT
6	NK-KONDI-SPS-2/4	France	HARC	PVT
7	NK-KONDI-SPS-7/4	France	HARC	PVT
8	NK-NEOMA-SPS-2/4	France	HARC	PVT
9	Brazil Long seed PL2-SPS-3/4	Brazil	HARC	PVT
10	Brazil Long seed PL4-SPS-4/4	Brazil	HARC	PVT
11	NK-FERTI-SPS-8/1	France	HARC	PVT
12	Brazil Long Seed PL6-SPS-7/4	Brazil	HARC	PVT
13	Brazil Long Seed PL9-SPS-4/4	Brazil	HARC	PVT
14	H-45-SPS-5/4	India	HARC	PVT
15	NK-FERTI-SPS-7/4	France	HARC	PVT
16	NK-KONDI-SPS-7/4	France	HARC	PVT
17	NK-NEOMA-SPS-7/4	France	HARC	PVT
18	VSFH-180-SPS-5/4	India	HARC	PVT
19	VSFH-1044-SPS-1/4	India	HARC	PVT
20	VSFH-1044-SPS-2/4	India	HARC	PVT
21	VSFH-1044-SPS-3/4	India	HARC	PVT
22	VSFH-1044-SPS-9/4	India	HARC	PVT
23	VSFH-1044-SPS-10/4	India	HARC	PVT
24	VSFH-2006-SPS-2/4	India	HARC	PVT
25	Oissa	Released	HARC	Breeder seed

2.1. Experimental Materials Used for Study

The experimental materials were obtained from the National oilseed Coordinating Center, Holletta Agricultural Research Center. Twenty four sunflower genotypes introduced and adapted to central highland of Ethiopia and one standard check were included in the study (Table 2). The genotypes were introduced from, France, Hindi and Brazil and were being maintained and used for different breeding purposes at Holletta Agricultural Research Center. The check variety included in the study (Oissa) is breeder seed variety released by Ethiopian institute of Agricultural Research and recommended for high and mid altitude agro ecologies of Ethiopia.

2.2. Experimental Design and Trial Management

The experiment was replicated twice using simple lattice Design (5x5). Each plot had a length of 4m and 4 rows with row to row spacing of 75cm and plant-to-plant 25cm respectively. The seed sown at two locations (Holetta and Adadi) and all seeds emerged. Two seeds were dibbled in each hill for better emergency and thinning was carried after emergency to retain one plant per hill. Sowing date was done on June, 29/2017. Planting was performed manually and hand weeding was carried out twice. All other agronomic practices recommended for the area were followed during the crop growth period.

2.3. Data Collected and Statistical Data Analysis

Data were collected on plot basis and plant bases for ray floret number, leaf number, petiole length, yield/plant, number of seed per plant, yield per hectare, oil content, oil yield, stem diameter, head diameter, days to flowering, days to maturity, seed filling percentage and hundred seed weight as follows.

2.3.1. Phenological and Agronomic Data Collected

(i). Number of Leaves per Plant

Leaves of 10 randomly selected plants were counted after attaining of maximum number of leaves.

(ii). Height of Plant (cm)

At full development of crop 10 plants were sampled randomly and heights of plants were recorded from ground level to the point of attachment of disk with the stem.

(iii). Head Diameter (cm)

Lengths of disks of 10 randomly taken plants were measured from one edge of the disc to the other and their averages were taken.

(iv). Petiole Length (cm)

The lengths of petioles were recorded for 10 randomly taken plants per plot and average was computed.

(v). Number of Leaves on Main Stem

Numbers of leaves for 10 plants per plot were counted at flowering stage and averages were computed.

(vi). Ray Floret Number

Ray florets of 10 randomly taken plants were counted at flowering stage and averages were computed.

(vii). Days to 50% Flowering

Numbers of days from the date of sowing to the day on which 50 per cent of plants flowers were recorded.

(viii). Days to Maturity

Number of days from the date of sowing to the day on which back of capitulum in 50 per cent of plants in a line turned to lemon yellow color was recorded.

(ix). Stem Diameter (cm)

Stem diameter was measured for 10 plants randomly taken from a plot in centimeters at one third height of the plant from ground level with the help of vernier calipers.

2.3.2. Seed and Other Biochemical Data Collected

(i). Seed Filling Percentage (%)

Total number of seeds per head was calculated from number of filled seeds and unfilled seeds of the head and the seed filling percentage was computed using the following formula.

$$\text{Seed filling percent} = \frac{\text{number of filledd seeds}}{\text{total number of filled and unfilled seeds}} \times 100$$

(ii). Hundred Seed Weight (g)

One hundred filled seeds were sampled randomly, weighed and the weight was recorded in grams.

(iii). Seed Yield Per Plant (g)

The total weight of seeds per plant was recorded in grams for 10 plants and averages were computed.

(iv). Oil Content (%)

The seed oil content was determined through non-destructive method by utilizing nuclear magnetic resonance (NMR) technique in the laboratory of Oil Seeds Research at HARC. Oil percentage (%) was determined by using nuclear magnetic resonance spectrometry (NMRS). Sample of 10g of seeds were dried in an oven for 3hr at 78°C and cooled for 3hours. Then oil contents of seeds were measured using nuclear magnetic resonance machine.

(v). Seed Yield Per Hectare (kg)

Seed Yield recorded per plot was converted to hectare for all plots.

(vi). Oil Yield (kg/ha)

Oil yield in kg/ha was calculated by using the following formula.

$$\text{Oil yield } \left(\frac{\text{kg}}{\text{ha}}\right) = \frac{\text{seed yield } \left(\frac{\text{kg}}{\text{ha}}\right) \text{ x oil content (\%)}}{100}$$

Data were subjected to statistical analysis using excel and SAS 9.3. Soft wares.

3. Results and Discussion

In the present study five principal components which account for most of the variability have been extracted, since five components had Eigen value greater than one. They account for 84.72% of the total variability in the original data (Table 3). Interpretation of the principal components is depends on finding which variables are most strongly correlated with each component, i.e., which of these values are large in magnitude, the farthest from zero in either direction influence the clustering more than those with lower value closer to zero [8].

Loadings of 0.30 or higher can be considered important [9]. The first principal component which accounted for 31.9% of total variability among genotypes were chiefly originated from traits such as yield per plant, head diameter, stem diameter, Ray floret number, yield per plant, number of seed per plant, yield per hectare, head diameter, stem diameter, plant height, days to flowering, days to maturity and seed filling percentage.

Likewise the Second principal component (PC2) which accounted for 22.72% of the total variability among genotypes were attributed to discriminatory traits such as yield per plant, yield per hectare, oil yield per hectare, seed filling percentage and hundred seed weight. Conversely, leaf number, oil content, plant height, number of seed per plant and days to maturity have the highest negative loadings in PC2. The third principal component which contributed 12.25% of total variability among genotypes was due to discriminatory traits namely,

number of seed per plant, head diameter and stem diameter. Whereas seed yield per hectare, oil yield per hectare and oil content are chief contributors with negative loading in PC3. The fourth principal component which accounted for 10.11% of total variability among genotypes were due to leaf number, days to maturity, seed filling percentage and hundred seed weight while, ray floret number and head diameter contributed negatively to PC4. Finally leaf number contributed chiefly to the variation of PC5 (1.08%).

Of all quantitative traits studied days to maturity (days) and seed filling percentage (%) contributed to the variation in three principal components out of the five principal components. The present study showed that the sunflower genotypes manifested variation for the traits studied. This suggests opportunity for genetic improvement through selection directly from genotypes and or selection of diverse parents for hybridization program.

Many scientists applied principal component analysis in his cultivar development program and stated the potential use of principal component in the selection of superior genotypes in sunflower. [10] also applied principal component analysis for estimation of genetic diversity in sunflower. Numerous yield related traits and reported the contribution of the first two components in assessment of total variation of sunflower hybrids [7]. The importance of principal component analysis also suggested by other sunflower researchers in breeding and germplasm improvement [11-13].

As Naila Gandahi et al. [14] Showed in the principal component analysis the first three components explained 91.60% of the total variations, which is consider a very huge genetic variance for the studied traits among the sunflower genotype and reported that the first, second and third components accounted 46.50%, 32.90% and 12.20% of the variation found in the Eigen vector analysis respectively. [14] reported the first principal component, seed yield plant-1 (0.48), plant height (0.45) and head diameter (0.44) were the most important contributing characters whereas days to heading (0.51), days to maturity (0.50) and seed index (0.49) were the important traits that chiefly contributes to the second principal component.

Table 3. *Eigen vectors Eigen values, proportion and cumulative variance for the first five PCA of 25 sunflower genotypes based on 15 quantitative traits.*

Traits	Eigenvectors				
	PC1	**PC2**	**PC3**	**PC4**	**PC5**
Ray floret number (no)	0.43	-0.21	0.03	-0.62	0.23
Leaf number (no)	-0.065	-0.3	0.25	0.48	0.74
Petiole length (cm)	0.29	-0.27	0.36	0.2	-0.26
yield (g)/plant	0.69	0.55	0.13	-0.16	0.245
number of seed/plant	0.6	-0.25	0.56	-0.27	0.18
yield (kg)/hectare	0.62	0.62	-0.4	0.016	0.05
Oil content	0.39	-0.571	-0.58	-0.05	0.15
Oil yield (kg)/hectare	0.68	0.24	-0.67	-0.00056	0.122
Head diameter (cm)	0.69	0.27	0.35	-0.349	0.075
stem diameter (cm)	0.68	0.099	0.365	0.135	-0.53
plant height (cm)	0.67	-0.38	-0.061	0.296	0.00123
Days to flowering (days)	0.685	-0.52	0.12	0.29	-0.134
Days to maturity (days)	0.662	-0.374	0.087	0.44	-0.085
Seed filling percentage (%)	0.324	0.83	-0.07	0.32	-0.029
Hundred seed weight (g)	-0.21	0.766	0.291	0.314	0.193
Eigen value	4.47	3.18	1.71	1.41	1.08
Proportion	31.9	22.72	12.25	10.11	7.75
Cumulative	31.9	54.62	66.86	76.97	84.72

4. Conclusion

Five principal components have been extracted, since five components had Eigen value greater than or equal to one and they accounted for 84.72% of total variation. Head diameter, days to flowering and yield plant and oil yield showed the maximum contribution in PC1. Whereas, hundred seed weight and yield per hectare in PC2, number of seed per plant, hundred seed weight and leaf number in PC3, leaf number and days to maturity in PC4 PC5 showed the maximum contribution. Maximum contribution for number of leaves per plant, yield per plant and ray floret number were observed in PC5.

References

[1] Maia Filho, FF, RF Pereira, IF Alves, SN Cavalcante, EF Mesquita and TC Suassuna. 2013. Growth and phytomass of sunflower variety 'Embrapa 122/V-2000' fertilized with cattle manure in two soils. Agropecu. Scientific Semiarid. 9: 67-75.

[2] Nobre, R. G., H. R. Gheyi, F. A. L. Soares, and J. F. Cardoso. 2011. Sunflower production under saline stress and nitrogen fertilization. Rev. Bras. Soil Science. 35: 929-937.

[3] Dutta, Anand, and Sudhir K. Dutta., 2003. "Vitamin E and its role in the prevention of atherosclerosis and Carcinogenesis: a review." Journal of the American College of Nutrition 22 (4): 258-268.

[4] Dong, G., Liu, G., Li, K. 2007. Studying genetic diversity in the coregermplasm of confectionary sunflower (*Helianthus annuus* L.) in China based on AFLP and morphological analysis. Russian Journal of Genetics 43 (6): 627-635.

[5] Mohammadi SA, Prasanna BM. 2003. Analysis of genetic diversity in crop plants: Salient statistical tools and consideration. Crop Science 43, 1235-1248.

[6] Venujayakanth, B., Dudhat A S., Swaminathan B., and Anurag ML. 2017. Assessing Crop Genetic Diversity Using Principle Component Analysis: A Review. Trends in Biosciences 10 (2): 523-528.

[7] Nazir, A., Farooq, J., Mahmood, A., Shahid, M., & Riaz, M. 2013. Estimation of genetic diversity for CLCuV, earliness and fiber quality traits using various statistical procedures in different crosses of *Gossypiumhirsutum L*., 43 (4).

[8] Chahal, G. and Gosal, S. S. (2002) Principles and Procedures of Plant Breeding: Biotechnological and Conventional Approaches. Narosa Publishing House, New Delhi, Vol. 21, 64-89.

[9] Kline, P., 2014. An easy guide to factor analysis. London: Routledge.

[10] Kholghi, M., Bernousi, I., Darvishzadeh, R. and Pirzad, A. 2011. Correlation and path-coefficient analysis of seed yield and yield related trait in Iranian confectionery sunflower populations. African Journal of Biotechnology, 10 (61): 13058-13063.

[11] Masvodza DR, Gasura E, Zifodya N, Sibanda P & Chisikaurayi B (2015). Genetic diversity analysis of local and foreign sunflower germplasm (*Helianthus annuus*) for the national breeding program: Zimbabwe. *J Cereals Oilseeds* 6 (1): 1-7.

[12] Ghaffari M (2004). Use of principal component analysis method for selection of superior three ways crosses hybrids in sunflower. *Seed and Plant* 19 (4): 513-527.

[13] Ghaffoor A & Arshad M (2008). Multivariate analysis for quantitative traits to determine genetic diversity of black gram (*Vigna mungo* L. Hepper) germplasm. *Pak J Bot* 40 (6): 2307-2313.

[14] Naila Gandahi, Aftab Ahmed Mahar, Abdul Wahid Baloch, Sarfraz Ahmed Ansari, Tauqeer Ahmad Yasir, Munaiza Baloch, Liaquat Ali Bhutto, Asad Mari and Tanweer Fatah Abro (2017). Assessment of genetic diversity for quantitative traits in sunflower germplasm. Pure and Applied Biology. Vol. 6, Issue 1, pp 261-266.