Indian J. Genet. Plant Breed., (2023); 83(4): 535-545

https://doi.org/10.31742/ISGPB.83.4.10

ISSN: 0975-6906

www.isgpb.org

RESEARCH ARTICLE

Multivariate analysis among indigenous and exotic advanced inbred lines of sunflower (*Helianthus annuus* L.) for genetic improvement

H.P. Meena*, M.Sujatha, Praduman Yadav, M.Y. Dudhe, M.S. Uma[1], K. Vikas[2], R.Sasikala[3], M.K. Ghodke[4], S. Neelima[5], S.B. Sakhare[6], K. Vineeta[7], S.S. Kandhola[7], A. Misal[4], S.D. Nehru[1], Neeraj Kumar[8], Subash Chander[8], E.Umarani[9], K.Jayshree[10], S.S. Lakshmanand[11] and B.S. Yashavanth[12]

Abstract

Evaluation and validation of 358 Indian-origin and exotic inbred lines of sunflower (*Helianthus annuus* L.) was done over two years to identify trait-specific inbreds with desirable agronomical features for genetic improvement. The analysis of variance exhibited significant differences among the genotypes for most of the agronomic traits studied, which favor selection. Positive correlations were recorded among the various characters in the improvement of seed yield, particularly for plant height, head diameter, and 100 seed weight. Cluster analysis grouped 163 inbreds under Cluster V, which included genotypes with medium maturity, high oil content and good yielding ability. The clustering pattern indicated the association between geographic diversity and genetic diversity. The first principal component accounted for 75% of the total variation in the population. The most important trait that contributed more to the variation in PC1 was plant height (0.989), followed by the number of leaves per plant (0.113). The second principal component accounted for 87% of the total variation and was positively associated with days to 50% flowering (0.674) and days to maturity (0.654), whereas it was negatively associated with seed yield per plant (-0312). Several trait-specific inbred lines were identified using multivariate analysis that may serve as potential breeding material/parental lines for effective utilization in the heterosis breeding programme for developing agronomically superior heterotic sunflower hybrids.

Keywords: Agronomic traits, cluster analysis, Principal Component Analysis, seed yield, sunflower

Introduction

Sunflower (*Helianthus annuus* L.), belonging to the family Asteraceae, is one of the most important crops being cultivated worldwide as a source of premium edible oil. The wider adaptation of sunflowers to various climatic and soil conditions has expanded its cultivation as an oilseed crop worldwide (Forleo et al. 2018; Babych et al. 2021). In India, oilseed types are predominantly cultivated, where the primary goal of sunflower breeding is to develop hybrids or cultivars with high seed yield potential (> 2.5–3.0 t/ha), high oil content (> 42%) and resilience to key biotic and abiotic stresses. The major challenge for the breeders to develop high yielding genotypes with high oil content is to identify genetically diverse parents with useful genetic variability for agronomically desirable traits. In spite of significant breeding efforts put so far, the average productivity level in India is very low (825 kg/ha) and stagnant as compared to the world's productivity (2048 kg/ha) (FAOSTAT 2019). The low yields and the productivity may be attributed

ICAR-Indian Institute of Oilseeds Research, Hyderabad 560 065, India.

[1]University of Agricultural Sciences GKVK, Bengaluru 584 104, India.

[2]University of Agricultural Sciences, Raichur 641 003, India.

[3]Tamil Nadu Agricultural University, Coimbatore 413 512, India.

[4]Vasat Rao Naik MarathwadaK rishi Vidyapeeth, ORS, Latur 413 512, India

[5]Acharya N. G. Ranga Agricultural University, RARS, Nandyal, India.

[6]Dr. Panjabrao Deshmukh Krishi Vidyapeeth, Akola 484 104, India.

[7]Punjab Agricultural University, Ludhiana 141 001, India.

[8]Chaudhary Charan Singh Haryana Agriculture University, Hisar 125 004, India.

[9]Professor Jayashankar Telangana State Agricultural University, ARS, Tornala 502 114, India.

[10]Odisha University of Agriculture & Technology, Bhubaneswar 743 338, India.

---

[11]Ramkrishana Ashram KVK, Nimpith 751 003, India.

[12]ICAR-National Academy of Agricultural Research Management, Hyderabad 500 003, India.

***Corresponding Author:** H.P. Meena, ICAR-Indian Institute of Oilseeds Research, Hyderabad, India, E-Mail: hari9323@gmail.com

How to cite this article: Meena H.P., Sujatha M., Yadav P., Dudhe M.Y., Uma M.S., Vikas K., Sasikala R., Ghodke M.K., Neelima S., Sakhare S.B., Vineeta K., Kandhola S.S., Misal A., Nehru S.D., Kumar N., Chander S., Umarani E., Jayshree K., Lakshmanand S.S. and Yashavanth B.S. 2023. Multivariate analysis among indigenous and exotic advanced inbred lines of sunflower (*Helianthus annuus* L.) for genetic improvement. Indian J. Genet. Plant Breed., **83**(4): 535-544.

Source of support: ICAR, Govt. of India

Conflict of interest: None.

Received: June 2023 **Revised:** Sept. 2023 **Accepted:** Oct. 2023

to several reasons, including the narrow genetic base of Indian sunflower hybrid parental lines, frequent use of few selected breeding materials as parents in varietal/hybrid development programmes, genetic vulnerability and susceptibility to various diseases like *Alternaria* leaf blight and sunflower necrosis during *kharif*, downy mildew during both *kharif* and *rabi* and powdery mildew and leaf curl virus during *rabi*. Consequently, the cultivated area under sunflowers has decreased sharply (2.0–0.25 mha).

The value of breeding material depends on the number of lines and the extent of genetic diversity among the inbreds for yield and yield components. A comprehensive analysis of the genetic diversity available in inbreds currently being used across Indian breeding programmes has not been done and is mostly restricted to few germplasm accessions. Describing heterotic groups and patterns is one of the practical applications of genetic diversity analysis (Laude and Carena 2015). With the changing climatic scenario, the breeding of climate-resilient varieties is gaining importance. Hence, the characterization of available existing inbreds/parental lines of sunflowers is important to determine the genetic diversity level for increased utilization. Multivariate statistical analysis for combining desirable traits to exploit in heterosis breeding programmes has been very useful because multiple measurements of each individual under study offer a very strong theoretical basis for providing the most accurate information on real genetic distances between genotypes (Singh and Pawar 2005). Cluster analysis (CA), Principal Component Analysis (PCA), Principal Coordinate Analysis (PCoA) and Multidimensional Scaling (MDS) are the most useful and widely employed multivariate approaches in several crops, including sunflower (Mohammadi and Prasanna 2003; Riaz et al. 2020). Various breeding programmes have taken advantage of the existing diversity, resulting in the development and release of a large number of sunflower varieties and hybrids (Sujatha et al. 2022). However, the released populations and hybrids cannot be continued for long due to genetic drift or deterioration of parental lines besides susceptibility to various biotic and abiotic stresses. This demands the replacement of current populations and hybrids with new breeding materials. Keeping in view the urgent need for characterizing the existing genetic diversity and importance of the multivariate technique, the present study was undertaken to (i) determine the amount of genetic diversity in 358 inbreds of sunflower of Indian and exotic origin, (ii) whether differences in geographic origin can be used as an index of genetic diversity, (iii) to estimate correlations among the traits, and (iv) to identify the best genotypes by cluster analysis, principal component analysis and biplot display for efficient utilization and strengthening the heterosis programme in India.

Materials and methods

Plant materials and experimental design

The material for present investigation was comprised of 358 sunflower inbred lines, including newly developed fertility restorer gene pool lines (RGP), and germplasm management unit (GMU) lines from ICAR-IIOR, Hyderabad, India. The exotic collection (EC) from Russia, trait-specific lines (TSG), interspecific derivatives (ID), sunflower core germplasm lines (SCG) from USDA-ARS-MWA-PIRU, Iowa State University, USA, maintainer gene pool lines (MGP) and maintainer lines (B) of different CMS lines from different. All India Coordinated Research Project (AICRP) on Sunflower centers and four checks, *viz.,* DRSF-113, RHA-6D-1, LTRR-341 and ARM-243B. The details of the materials used in this study are presented in Table 1. The evaluation and characterization of the inbreds were conducted during the late *kharif* seasons in an Augmented Block Design (ABD) during 2017-2018 and 2018-2019 at the ICAR-Indian Institute of Oilseeds Research (ICAR-IIOR), Hyderabad, India, at the latitude of 17^0 22’31”N and a longitude of $78^0$28’27”E. All of the test materials as well as checks were divided into 9 blocks keeping two replications in each block. Each inbred was grown in two rows of 4.0 m with a 60 x 30 cm spacing between rows and plants, respectively. Thirteen plants were maintained each row. Recommended crop production and protection measures were followed to raise the healthy crop. Observations were recorded from five randomly selected plants in each inbred for eight quantitative characters, namely, days to 50% flowering (DF), days to maturity (MD), plant height (PH in cm), number of leaves per plant (NL), head diameter (HD in cm), 100 seed weight (SW in g), seed yield per plant (Y/P in g) and oil content in percent. A sample of 10 to 15 g filled seeds was drawn from each entry, and oil content was measured as percent using nuclear magnetic resonance (NMR) spectroscopy as the procedure (Praduman and Murthy 2016).

Table 1. Sunflower inbreds contributed by different AICRP centres

S. No.	Centre name	No. of inbreds contributed	Material involved	Trait
1	USDA-ARS-MWA-PIRU, Iowa State University, USA	40	Trait specific germplasm (TSG), sunflower core germplasm (SCG) lines, interspecific derivatives (ID)	High seed yield, high oil content, short height and duration
2	ICAR-Indian Institute of Oilseeds Research, Hyderabad	34	Germplasm (GMU), exotic lines (EC), restorer gene pool (RGP), maintainer gene pool lines	High seed yield, high oil content, short height and duration, resistant to powdery mildew and tolerant to Alternariaster leaf spot
3	AICRP (Sunflower) centre, University of Agricultural Sciences, GKVK, Bengaluru	69	Maintainer lines, restorer gene pool, newly developed inbreds	High seed yield, high oil content, high autogamy and early maturity
4	AICRP (Sunflower) centre, University of Agricultural Sciences, Raichur	46	Newly developed inbreds, maintainer lines, exotic collections	High seed yield, high oil content, tolerant to sunflower necrosis disease
5	AICRP (Sunflower) centre, Tamil Nadu Agriculture University, Coimbatore	35	Newly developed trait specific inbreds, maintainer lines	High seed yield, high oil content and powdery mildew resistant
6	AICRP (Sunflower) centre, Oilseeds Research Station, Latur	45	Exotic lines, Indigenous lines, maintainer lines, newly developed inbreds	High seed yield, high oil content and resistant to downy mildew
7	AICRP (Sunflower) centre, Panjabrao DeshmukhKrishi Vishavvidyalya, Akola	32	Restorer lines, maintainers	High seed yield, high oil content and tolerant to leafhopper
8	AICRP (Sunflower) centre, Regional Agriculture Research Station, Nandyal	18	Newly developed inbred lines	High seed yield and high oil content
9	AICRP (Sunflower) centre, Punjab Agriculture University, Ludhiana	39	Newly developed trait specific inbreds, maintainer and restorer lines	High seed yield and high oil content
	Total	358		

Statistical analysis

The major descriptive statistics such as mean, standard deviation and the coefficient of variation for each of the eight traits studied were calculated (Panse and Sukhatme 1964) and phenotypic correlation coefficients were calculated as per Johnson et al. (1955). Cluster analysis was utilized to group the sunflower inbreds based on their quantitative characters and to identify the patterns of morphological variations and the importance of traits in each component PCA was conducted through a correlation matrix. For determining the diversity and PCA analysis, the R version 3.1.3 package was used (Core Team 2013). The biplot was generated using the 'Facto Mine R' (factor analysis and data mining with R) package (Husson et al. 2009) with the function 'biplot'. Ward's hierarchical algorithm on squared Euclidean distances was used to cluster genotypes into similarity groups. The pseudo-F statistic and the pseudo-T 2 statistic were examined as suggested by Jobson (1992) to determine the numbers of clusters using SAS version 9.1 software (SAS Institute, Cary, NC). Only those principal components with Eigenvalues of more than one were chosen, as Jeffers (1967) proposed. In the factor analysis, the principal components method first explained the total variance and Eigenvalues were determined. The number of factors was extracted based on the screed plot. Factor loadings were obtained using Rotated varimax and were used to calculate commonalities for each trait separately (Yildirim et al. 1989). Associations between the actual traits and the respective PCs were calculated. Data were processed using the statistical programme Minitab 14 (Minitab version 14, Minitab Inc., State College, PA, USA).

Results and discussion

Variance analysis

The analysis of variance exhibited a significant variation ($p < 0.05$) among treatments, checks *vs* test entries, blocks and test entries, whereas highly significant differences were observed for plant height, days to maturity, head diameter, 100 seed weight and seed yield per plant among the test entries. A significant variation was observed in checks versus test entries for seed yield per plant. A highly significant variation across the blocks was recorded. The block differences were higher for all characters, which was due to the incorporation of several treatments under evaluation. As the number of treatments under investigation was greater, the number of blocks was also more. The obtained minimum and maximum values, mean, standard error, standard deviation, and coefficient of variation for yield and yield contributing traits are presented in Table 2. Significant variation was observed for all traits for different parameters.

Table 2. Agro-morphological characteristics of 358 sunflower inbreds

S. No.	Variable	Range		Mean ± SE	Standard deviation (SD)	CV (%)
		Minimum	Maximum			
1	Days to 50% flowering	45.0	76.0	59.5 ± 0.314	6.24	10.48
2	Days to maturity	75.0	105.0	89.4 ± 0.305	6.06	6.78
3	Plant height (cm)	51.8	192.2	112.1 ± 1.124	22.33	14.92
4	Head diameter (cm)	5.0	20.4	10.1 ± 0.122	2.43	20.11
5	Number of leaves per plant	13.6	44.6	24.9 ± 0.241	4.80	15.24
6	Seed yield per plant (g)	2.4	25.4	12.8 ± 0.229	4.56	29.72
7	100 seed weight (g)	1.6	9.8	4.4 ± 0.076	1.51	17.38
8	Oil content (%)	14.5	38.8	29.5 ± 0.225	4.47	15.17

CV=Coefficient of variation

A major objective in sunflower breeding programmes is to improve the crop by reducing plant height. The minimum and maximum values for days to 50% flowering were 45 days and 76 days, respectively, with a general mean value of 59.5 days. RGP-58-P_4-S_1, TSG-47, IC-625774, EC-85818, HA-248B, TSG-114, RGP-58-P_4-S_2, R-272, SCG-36, CPI-3, GP_6-127 had <50 days mean values for this trait. These lines can be utilised for the development of early hybrids, especially for spring and *rabi* seasons. Seed yield ranged from 2.4 to 25.4 g/plant and high mean values for this trait were observed in EC-198073, EC-198075, ID-97, GPB-29, GP_6-1102, TSG-240 and SCG-22. Likewise, the minimum and maximum values for 100 seed weight were 1.6 and 9.8 g, with a general mean value of 4.4 g. The genotypes including SCG-66, IC-601951, ID-32, EC-198073, EC-276189, EC-210309, TSG-76, SCG-31, PM-134, EC-191897 and TSG-347 had high mean values for this trait. A wide range was observed for plant height, which varied from 51.8 to 192.2 cm with a general mean plant height of 112.1 cm. In the studies of plant height ranged from 124.3 to 222.9 cm. Due to the low mean value of plant height, the effects of tolerance to lodging, the genotypes, including TSG-369, RGP-60-P_4, RGP-95-P_1-S_1, TSG-283, TSG-284, RGP-32-P_1-S_1, RGP-28, TSG-271, GPB-26-1, COSF-12B and NDLB-8 will be preferable. Oil content is one of the essential traits for developing high oil-yielding hybrids in sunflowers. The range for this trait varied from 14.5 to 38.8% with a mean of 29.5%. COSF-6B, COSF-7B, COSF-3B, GP_6-11, R-83, RC-72, TSG-331, R3Brown (Br), NDI-7, R-393, CMS-57B and HA-302B are the inbreds for improving this trait. The range in variation in respect of the traits studied in the present study has been worked out earlier by several researchers in different populations (Razi and Assad 1999). The values of oil content and seed yield depend on the genotype and the environmental conditions in which the genotypes are grown (Marinkovic et al. 2003). The highest coefficient of variation was revealed for head diameter (23.94%), while the lowest was exhibited for days to maturity (6.78%) followed by days to 50% flowering (10.48%). However, the highest coefficient of variation for days to 50% flowering, plant height, head diameter and 100 seed weight were also observed earlier (Nooryazdan et al. 2010) but in a different set of material.

Association between seed yield and yield components

Understanding the relationships between achene yield and the factors that influence yield is essential for designing an efficient breeding programme (Velkov 1980). Yield is a complex character, a function of several component characteristics, and is greatly influenced by the environment. Hence, it is important to understand the relationships between different traits and achene yield to increase the usefulness of selection criteria to be followed while developing new varieties. Correlation analysis was used to determine the relationship between characters (Table 3). The character interrelationships indicated that among the eight traits studied, achene yield/plant had a non-significant positive correlation with plant height (0.157), head diameter (0.188), number of leaves per plant (0.052) and 100 seed weight (0.055) at p = 0.01. Recently, Riaz et al. (2020) reported that plant height and 100 seed weight exhibited a significant positive correlation with achene yield, while Radić et al. (2013) found a positive and significant association between seed yield and 100 seed weight. Sincik and Goksoy (2014) reported a significant positive correlation between seed yield and plant height, head diameter and 100 seed weight. A significant correlation between seed yield, oil yield and oil content was reported by Balalić et al. (2012), whereas Sefaoglu (2023) recorded no significant correlation between oil content with head size in sunflowers. However, the preset findings do not support the conclusion of Shojaei et al. (2022), who reported that the correlation between 100 seed weight and seed yield was positive and significant. This disagreement among results is presumably associated with differences in the genetic material and environmental conditions where these studies were conducted.

A significant negative association of days to 50% flowering was observed with the head diameter (-0.210), the number of leaves/plant (-0.170), seed yield/plant (-0.205), 100 seed weight (-0.279) and oil content (-0.042). A strong negative association between these traits indicates that

Table 3. Phenotypic correlation coefficients for eight traits in 358 sunflower accessions

Character	Days to 50% flowering	Days to maturity	Plant height (cm)	Head diameter (cm)	Number of leaves/plants	Seed yield/plant (g/plant)	100 seed weight (g)	Oil content (%)
Days to 50% flowering	1.00	0.99**	0.093**	-0.210**	0.170**	-0.205*	-0.279*	-0.042*
Days to maturity		1.00	0.089**	-0.212**	0.172*	-0.206*	-0.280*	-0.044*
Plant height (cm)			1.000	0.571**	0.518**	0.157**	0.089*	0.111**
Head diameter (cm)				1.000	0.397**	0.188**	0.204**	0.006
Number of leaves/plants					1.000	0.0526**	-0.026	-0.059*
Seed yield/plant (g)						1.000	0.132**	0.099*
100 seed weight (g)							1.000	0.055*
Oil content (%)								1.000

*, **Significant at 0.05, 0.01probability level

selection for early types will affect the traits mentioned above. The characters, *viz.,* plant height, head diameter, 100 seed weight, number of leaves/plant and head diameter, are important selection indices for oil and seed yield improvement. In fact, in a breeding programme, the choice of correlated component traits should be taken into consideration for selecting genetically diverse parents.

Clustering pattern of the sunflower inbreds

Cluster analysis illustrates the pattern of resemblance

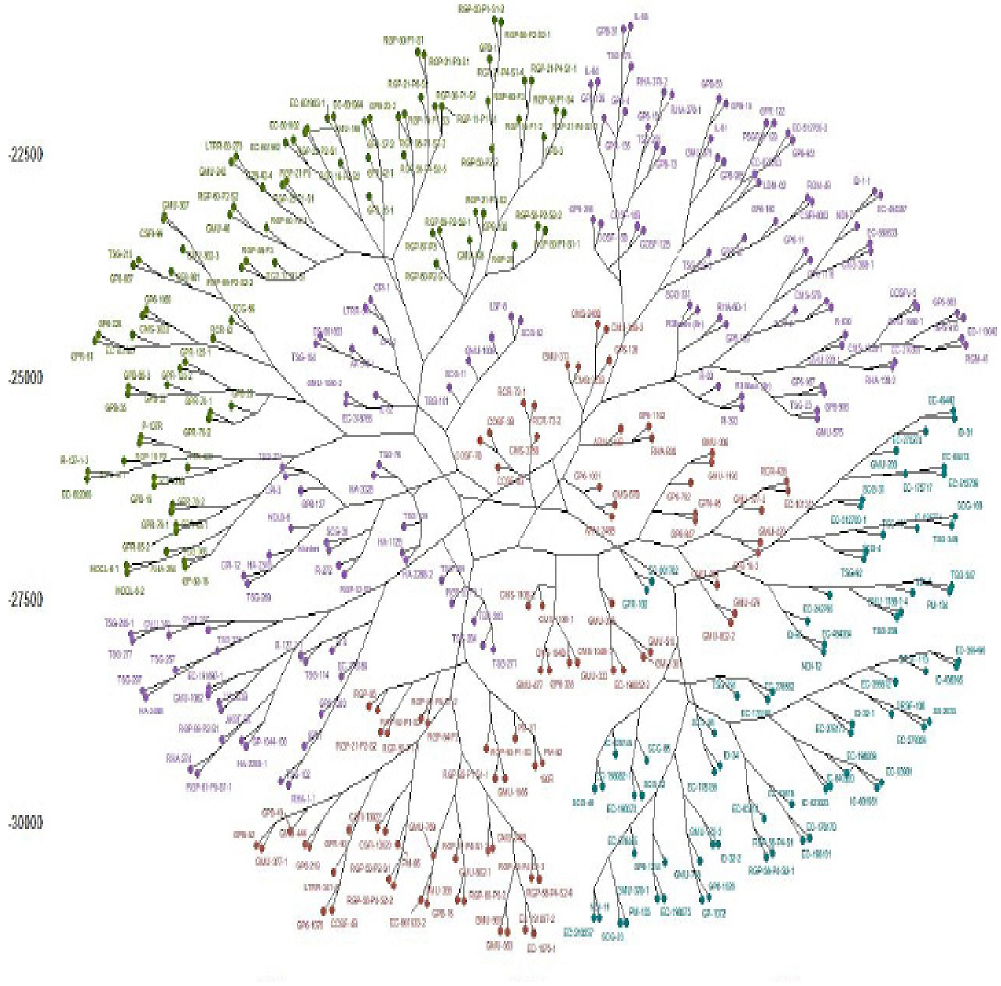

Fig. 1. Tative traits

Table 4. Mean values of eight characters in different clusters of sunflower

Cluster no.	Number of genotypes	Characters							
		DF	DM	PH	HD	NL	SY	SW	OC
Cluster I	33	66.58	96.35	82.84	7.74	21.51	7.45	3.03	27.90
Cluster II	100	61.69	91.31	156.55	12.86	29.37	14.65	4.93	28.93
Cluster III	15	52.00	82.00	71.67	8.53	19.26	13.73	5.20	29.67
Cluster IV	09	60.74	90.64	128.36	11.24	26.97	13.36	4.66	30.11
Cluster V	163	52.27	82.27	112.97	10.62	23.94	12.16	5.48	29.62
Cluster VI	01	60.30	90.30	111.04	9.97	25.96	17.45	4.08	30.71
Cluster VII	36	64.31	94.08	100.53	8.96	24.54	9.38	3.16	27.03
Cluster VIII	01	53.91	83.86	94.09	9.57	21.16	13.96	5.43	29.75

between genotypes based on evolutionary relationships or phenotypic performance. It is utilized to group similar lines/germplasm into one group and differentiate them from others. The cluster mean values for yield component characters of 358 sunflower inbreds, including checks, are presented in Table 4. A large range among the eight clusters was observed for different traits, indicating the existence of a great extent of diversity in the material. The inbred dendrogram tree based on Euclidean distance exhibited four major clusters (Fig. 1). The major cluster (V) consisted of 163 lines followed by cluster II, which consisted of 100 lines, and cluster 6 consisted of 36 lines, contributing 45.53, 27.93 and 10.05% of total genotypic strength, respectively (Supplementary Table S1). Clusters VI (TSG-277) and VIII (RGP-60-P_2-S_2) are solitary clusters containing one genotype each, indicating the distinctiveness of those genotypes compared to other inbred lines included in the study. On the other hand, the 40 genotypes from USDA, USA were spread over seven clusters, except cluster VIII, indicating greater genetic similarity between the genotypes originating from the USA. These findings reveal that geographical diversity is not always a direct cause of genetic diversity. Genotypes within a cluster exhibit a narrow range of genetic variability, whereas different clusters manifest wider variability. Sahari et al. (2013) also observed similar clustering among sunflower populations. From the present analysis of data, it can be concluded that the days to 50% flowering was highest in cluster I (66.6 days) and the lowest in cluster III (52.0 days), indicating the presence of late-maturing inbreds. Cluster II recorded the highest mean values for plant height (156.5 cm) and the lowest in cluster III (71.7 cm). The character head diameter was the highest in cluster II (12.9 cm) and the lowest in cluster I (7.7 cm). The genotypes with the highest mean value for seed yield were recorded in cluster VI (17.4) and the lowest in cluster I (7.5). Genotypes in cluster VI followed by cluster IV, recorded the highest mean value for oil content (30.7 and 30.1), while the least was recorded by genotypes in cluster VII (27.0). Furthermore, cluster analysis revealed that inbreds in each cluster possessed some specific characteristics. Particularly noteworthy are the extreme mean values of two traits, i.e., days to flowering and days to maturity, in clusters III and I, despite the fact that both clusters included entries from the USDA, USA and different parts of India. This demonstrated that geographic distribution cannot be taken to be a single criterion of genetic diversity. Further, it is possible to collect

Table 5. Factor analysis of traits under study via principal components analysis in 358 sunflower genotypes

Trait	Factors loadings					
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Days to 50% flowering	-0.03	-0.67	-0.22	-0.05	-0.03	0.07
Days to maturity	-0.03	-0.65	-0.22	-0.04	-0.03	0.07
Plant height (cm)	-0.99	0.03	0.06	-0.07	-0.09	-0.03
Head diameter (cm)	-0.06	0.07	0.01	0.11	0.05	0.16
Number of leaves/plants	-0.11	-0.07	-0.02	0.71	0.69	0.02
Seed yield per plant (g)	-0.05	0.31	-0.95	0.05	-0.05	-0.03
100 seed weight (g)	-0.01	0.09	0.01	-0.03	-0.01	0.98
Oil content (%)	-0.02	0.04	-0.06	-0.69	0.72	-0.01
Eigenvalues	490.4	83.30	38.37	7.47	13.27	6.35
Proportional variance (%)	22.11	9.11	6.18	4.17	3.63	2.51
Cumulative variation (%)	75.15	87.92	93.80	96.48	98.52	99.49

wide variability even from relatively nearer regions.

The clustering clearly demonstrated that none of the clusters contained inbred lines with all suitable characters which could be selected and exploited directly in the breeding programme. Thus, in order to develop a desirable genotype, F_1 hybrids or segregants, hybridization is needed between the genotypes of various clusters. Cluster means revealed differences in the value of a particular character between the groups, indicating the importance of a particular cluster for choosing a desirable parent for a hybridization programme. The clusters contained sunflower inbred lines of heterogenous origin coming from different parts of the country. Even the exotic lines from USDA, USA were grouped with indigenous materials. This suggested a lack of parallelism between genetic and geographical diversity. Murthy and Arunachalam (1966) demonstrated that genetic drift and selection in different environmental conditions might result in greater diversity among inbreds than geographical distance. Hence, the selection of parental material for hybridization simply based on geographic diversity may not be successful. Frequent interchange of valuable breeding materials from one location/area to another and the subsequent selection can also be responsible for the distribution of gene complexes over distant geographical locations. Therefore, the selection of parental material for hybridization should be based on genetic diversity rather than geographic distance. From the above discussion, it may be concluded that intercrossing the inbreds from various clusters may generate substantial variability and it is likely to introduce high yield and transgressive segregants for population improvement programmes.

Factor analysis

The main goal of factor analysis is to explain as much variance as possible in a data set by using the smallest number of factors and the smallest number of items or variables within each factor. Factor analysis, which separates traits correlated with each other into independent clusters, can be used successively to determine characteristics that can be utilized as a selection criterion (Kaya et al. 2002). Factor analysis was performed based on principal component analysis and provisional factors were rotated by the Varimax method (Table 5). In factor 1, all traits being recorded possessed negative factor loading. Plant height showed the highest factor loading value (-0.99), followed by the number of leaves per plant (-0.11). The inbreds in the PC-I were more likely to be integrated with early maturity and less number of leaves per plant. A positive factor-loading value was observed for all the traits in none of the factors.

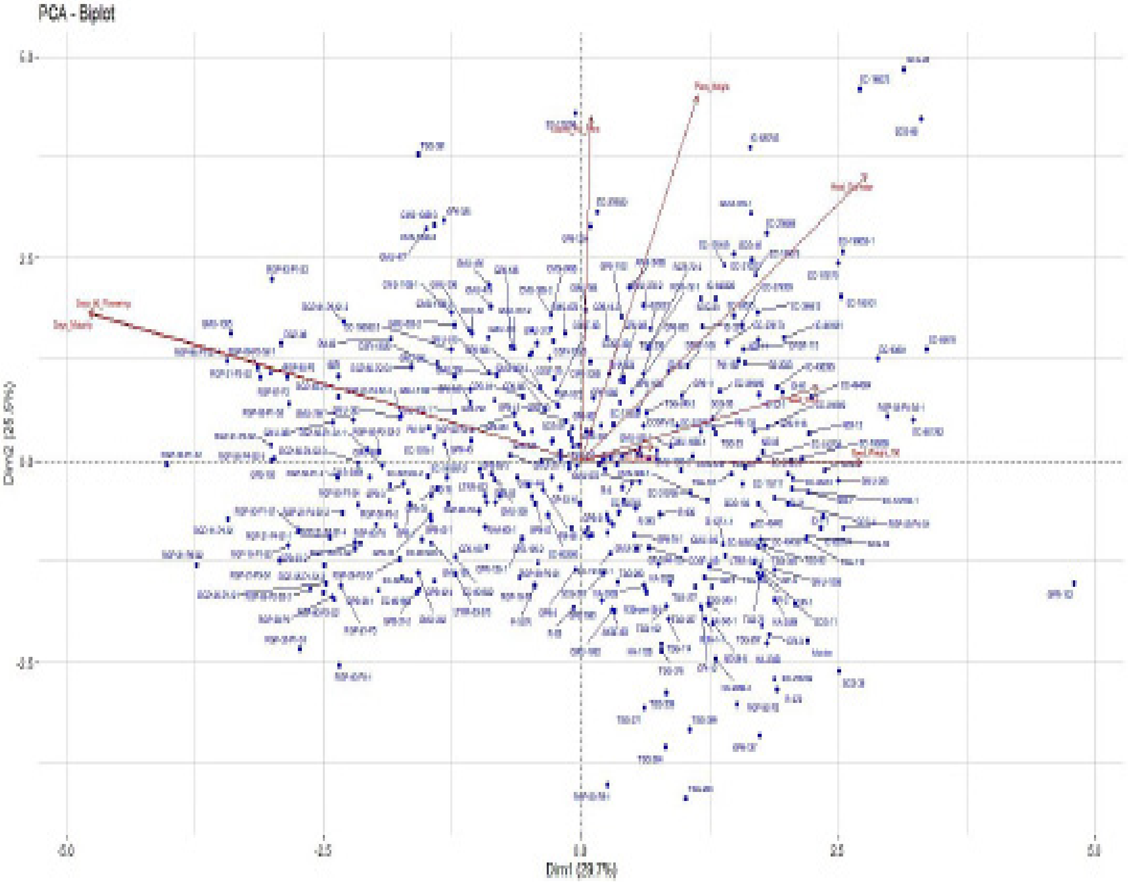

Fig. 2. Components

Table 6. Eigen value of principal components

	Principal component axis							
	1	2	3	4	5	6	7	8
Eigenvalues	490.38	83.30	38.37	7.47	13.27	6.35	3.18	0.10
Proportional variance (%)	22.11	9.11	6.18	4.17	3.63	0.01	0.01	0.01
Cumulative variation (%)	0.75	0.87	0.93	0.96	0.98	0.99	0.99	1.00

In factor 2, traits like seed yield per plant (0.31), followed by 100 seed weight (0.09), head diameter (0.06) and oil content (0.04) possessed positive factor loadings. The inbreds in the PC-II were more likely to be associated with high seed yield/plant, high test weight, large head diameter and high oil content. The highest factor-loading value was positive for plant height (0.063), followed by head diameter (0.013) and 100 seed weight (0.005) in Factor 3. In Factor 4, the highest positive factor-loading value was possessed by the number of leaves per plant (0.704), followed by head diameter (0.111) and seed `seed yield per plant and oil content were observed in Factor 6 and the highest positive value was observed for 100 seed weight (0.982), followed by head diameter (0.156) and days to 50% flowering (0.071). The present study revealed that most of the yield contributing traits with higher direct effects and oil content were scattered throughout all factors, and thus the best hybrids could be among various population groups due to the diverse nature of the inbred lines. The present findings support the results obtained earlier by Elizondo-Barron (1991).

The first three factors explained 93.0% of the total variation (Table 6). The first factor, determining 22.11% of the variation, has an important role in justifying the alteration of the number of leaves per plant and seed yield per plant. So, this factor was named as the "Seed Yield" factor. If the selection had been complemented based on the first factor, this selection would have been the most effective in terms of seed yield per plant. The third factor explained 6.18% of the alteration, while the fourth factor justified 4.17% of the total variance. The present findings are comparable with those of Tousi Mojarad and Bihamta (2007) and Mollasadeghi et al. (2011), who worked on wheat.

Principal component analysis (PCA)

The most significant variables in the dataset were identified using principal component analysis. The importance of a character towards the PC could be seen from the corresponding Eigenvalues. The PCA revealed that the first

Table 7. Promising sunflower inbreds for different economic characters

S. No.	Character	Best check performance	No. of accessions identified	Inbreds
1	Days to 50% flowering (<50 days)	RHA-6D-1	13	RGP-58-P_4-S_1, RGP-58-P_4-S_2, HA-248B, TSG-274, TSG-283, CPI-3, EC-85818, IC-625774, EC-276189, Morden, SCG-36, R-272, GP6-127
2	Seed yield/plant (>20 g/plant)	ARM-243B	28	TSG-369, CPI-1, CPI-2, EC-198075, EC-49442, EC-68413, EC-175717, EC-276189, EC-210309, GP_6-1116, GMU-200, SCG-52, R-272, ID-97, SCG-22, CMS-38B, PM-155, EC-198073, GMU-1006, RHA-834, CMS-104B, GPR-125, COSF-7B, COSF-15B, TSG-240, GP6-1001, RGM-41, RGP-21-P_4-$S_{1\text{-}3}$
3	Oil content (>35%)	RHA-6D-1	19	COSF-6B, COSF-7B, COSF-3B, CMS-335B, RGP-93-P_1-S_3, NDI-7, GP_6-1116, GP_6-109, ID-1-1, TSG-339, R3Brown (Br), GMU-520-1, CMS-57B, R-8, R-83, R-630, R-393, GP6-127, RCR-72
4	100 seed weight (>7.0 g)	LTRR-341	18	SCG-66, IC-601951, EC-93901, EC-276189, EC-210309, EC-198089, TSG-347, EC-494357, EC-198073, EC-276582, EC-276173, Morden, TSG-76, SCG-31, RHA-138-2, HA-302B, EC-191897, PM-134
5	Plant height (<90 cm)	RHA-6D-1	36	TSG-369, RGP-32-P_1-S_1, TSG-283, TSG-284, RGP-52-P_3, RGP-61-P_5-S_1, TSG-271, TSG-274, TSG-297, NDLB-8, EC-276189, SCG-36, R-272, TSG-339, HA-234B, GP_6-127, EC-601962, CMS-246B, COSF-12B, COSF-14B, COSF-15B, GP_6-1063, P-107R, EC-601964, RGP-89-P_2-S_1, GPB-26-1, GPB-42-4, RGP-19-P_1-S_3, RGP-11-P_1-S_1, RGP-28, RGP-21-P_6-S_2, RGP-21-P_4-S_2, RGP-21-P_3-S_1, RGP-60-P_4, RGP-60-P_2-S_2, RGP-95-P_1-S_1
6	Head diameter (>15 cm)	ARM-243B	12	SCG-66, TSG-101, EC-93901, EC-85818, IC-601951, IC-623023, IC-625745, IC-640320, EC-178159, EC-178168, EC-276465, EC-198073
7	Number of leaves per plant (>30)	ARM-243B	24	EC-178168, EC-178170, EC-198073, IC-625745, TSG-391, SCG-22, EC-178159, SCG-66, EC-276582, EC-198101, RCR-72, R-8, IC-640320, RGP-58-P_4-S_2, SCG 46, EC-399512, IC-601951, ARM-250B, GP_6-325, GP_6-965, GP_6-1001, RGP-60-P_2-S_1, RGP-84-P_2, RGP-89-P_2-S_3
	Total		150	

four principal components were significant (eigenvalue >1) and contributed to 96% of the total variation. Correlation coefficients were calculated between the original data of the characters being studied and the respective principal components.The results of PCA revealed that the first four components with Eigenvalues of greater than 7.47 accounted for about 96% of the total variability in 358 sunflower inbreds involving all eight traits. Plant height (0.989) contributed more to the variation in PC-I followed by the number of leaves per plant (0.113). The second principal component (PC-II) accounted for 87% of the total variation and was positively associated with days to 50% flowering (0.674) and days to maturity (0.654), whereas it was negatively associated with seed yield per plant (-0312). This means that genotypes with high PC-II values have fewer days to flowering and maturity and lower seed yield. The third principal component (PC-III) accounted for 93.0% of the total variation in the inbreds. Oil content (0.691) contributed the maximum to PC-III followed by seed yield per plant (0.946) while the head diameter (-0.112) contributed negatively. This means that genotypes with high values of PC-III have higher oil content and seed yield. Likewise, the fourth principal component accounted for 96.0% of the total variation. The major characteristics that highly contributed to the variations include 100 seed weight (0.982), while oil content (-0.717) and the number of leaves per plant (-0.705) negatively contributed to the variation. Therefore, based on statistical analysis, component breeding among selected genotypes would be beneficial or worthwhile in developing ideal sunflower genotypes. Maruthi Sankar et al. (2004) confirmed that PCA and regression analysis are interesting for selecting superior sunflower genotypes. Principal component analysis has already been employed in sunflower germplasm, which could reduce the number of characters that need to be examined (Nooryazdan et al. 2010).

Biplot display

The biplot obtained from the first two principal components almost confirmed the cluster analysis group. Fig. 2 shows how the 358 sunflower inbreds were distributed along with the Ist and IInd principal components based on measured morphological attributes. All the measured parameters, including days to 50% flowering, days to maturity, plant height, number of leaves per plant, head diameter, 100 seed weight, seed yield per plant and oil content, were distributed in the biplot according to their loading and mean values. The important consideration in the biplot diagram is the angles of the vectors. The correlation coefficient between any two characters can be approximated by the cosine of the angle between the vectors (Yan and Kang 2019). According to Yan 1 and Rajcan (2002), the correlation is positive if the angle of vectors between two characters is <90 and the correlation is negative if it is >90. If the angle is exactly 90 degrees, however, they are independent. There were three vectors of traits, *viz.,* 100 seed weight, number of leaves/plant and head diameter, that had a small angle with each other; similarly, two traits, i.e., days to 50% flowering and days to maturity, had a small angle of vector which meant that they had positive correlations. However, plant height had the opposite direction vector with the traits, revealing that this trait had a negative correlation. Therefore, the smaller angle between vectors indicates a greater positive correlation among related traits and *vice-versa*. Hossein et al. (2011) also reported similar results in F_1 sunflower hybrids during germination and early seedling growth stages. Kroonenberg (1995) concluded that the angle of vectors shows correlations between vectors and, therefore, among traits. Representative inbred lines with the best performance of identified characters may be chosen from the particular groups for hybridization programmes with other approved inbreds. This strategy will help to identify and to combine favorable inbred lines to obtain important traits in one inbred with a broad genetic base.

In summary, using multivariate analysis, the present study assessed the genetic diversity of 358 publicly available inbreds from India and USDA-ARS. There is considerable genetic diversity among the exotic as well as Indian inbred lines, including maintainer and restorer lines for the majority of the traits. The information about the genetic diversity and relationships among inbred lines is not only helpful for germplasm conservation and inbred/parental line identification but also facilitates the selection of parental lines from a huge number of parental lines or inbreds for hybrid breeding in crops, including sunflower (Meena et al. 2013). Thirteen inbreds for early flowering (<50 days) and 7 inbred lines (COSF-6B, COSF-7B, COSF-3B, RGP-93-P_1-S_3, GP6-1116, R-83, RCR-72) for high oil content (>36%)were identified under Indian conditions. Thirty-six inbreds for short stature (<90 cm) and 24 inbreds for higher number of leaves/plant (>30) were identified (Table 7). Intercrossing the inbreds from different clusters may generate large variability and it is expected to introduce high yield and transgressive segregants for population improvement programs. Crossing among genetically divergent parents like CMS-248B, CMS-249B, COSF-4B, COSF-6B and COSF-7B of Cluster VII with fertility restorer parents like RGP-50-P_2, RGP-52-P_3, RGP-61-P_5-S_1, RGP-50-P_2-S_1, RGP-58-P_4-S_2, RGP-19-P_1-S_3, RGP-30-P_1-S_1, RGP-50-P_1-S_1 and RGP-60-P_4 of Cluster Ior female lines *viz.,* HA-228B, HA-248B, CMS-243B of cluster I with male line RGP-60-P_2-S_2 of Cluster VIII are expected to give rise the hybrids having better heterotic potential. These inbred lines can be exploited for the development of potential heterotic combinations with diverse genetic sources. The USDA lines, which are agronomically superior in

Indian climatic conditions, serve as a new source of diversity and can accelerate hybrid breeding programs in India and other sunflower-growing nations.

Supplementary material

Supplementary Table S1 with details of material is provided, www.isgpb.org

Authors' contribution

Conceptualization of research (HPM, MS); Designing of experiments (MS, HPM); Contribution of experimental materials (HPM, MS, MYD, MSU, KV, RS, MKG, SN, SBS, AM, SDN, EU, KJ); Execution of field/lab experiments and data collection (HPM, MS, PY); Analysis of data and interpretation (MYD, BSY); Preparation of the manuscript (HPM, MS).

Acknowledgments

The authors thank Dr. Laura Fredrick Marek, USDA-ARS-MWA-PIRU, Iowa State University, USA, for providing the seed materials. The funding from ICAR for conducting the experiments is duly acknowledged.

References

Babych V.O., Kuchuk M.V., Popov V.N., PariiYa.F.,Parii M.F. and Symonenko Yu. V. 2021. The use of molecular markers for acceleration of the selection process while developing sunflower (*Helianthus annuus* L.) maintainer lines. Indian J Genet Plant Breed., **81**(4): 582-585.

Balalić I., Zorić M., Branković G., Terzić S. and Crnobarac J. 2012. Interpretation of hybrid × sowing date interaction for oil content and oil yield in sunflower. Field Crops Res., **137**: 70-77.

Bhandari H.R., Bhanu A.N., Srivastava K., Singh M.N. and Hemantaranjan A. 2017. Assessment of genetic diversity in crop plants-an overview. Adv. Plants Agric Res., **7**(3): 279-286.

Core Team 2013. R: a language and environment for statistical computing (Software). http://www.r-project.org/. Accessed 26 Dec 2017.

Elizondo-Barron J. 1991. A factor analysis of plant variables related to yield in sunflower under water stress conditions. Helia, **14**(15): 55-64.

Forleo M.B., Palmieri N., Suardi A., Coaloa D. and Pari L. 2018. The eco-efficiency of rapeseed and sunflower cultivation in Italy. Joining environmental and economic assessment. J Clean Prod., **172**: 3138-3153.

Hossein Z.T., Esin S. and Kamil H. 2011. Principal components analysis of some F_1 sunflower hybrids at germination and early seedling growth stage. Atatürk University J. Agric. Facul., **42**(2): 103-109.

Husson F., Le S. and Pages J. 2009. Analyse de donnees avec R. Presses Universitaires de Rennes, Rennes Cedex.

Jeffers J.N.R. 1967. Two case studies in the application of principal component analysis. J. Appl. Stat., **16:** 225-236.

Jobson J.D. 1992. Applied multivariate data analysis. Volume II: Categorical and Multivariate Methods. New York: Springer-Verlag, USA.

Johnson H.W., Robinson H.F. and Comstock R.E. 1955. Estimates of genetic and environmental variability in soybean. Agron. J., **47**: 314-318.

Kaya Y., Topau A., Gonulau E. and Arisoy R. 2002.Factor analysis of yield traits in genotypes of durum wheat (*Triticum durum* L.). Indian J. Agric. Sci., **72**(5): 301-303.

Kroonenberg P.M. 1995. Introduction to biplots for G × E tables. Department of Mathematics, Research Report 51. University of Queensland, Australia.

Laude T.P. and Carena M.J. 2015. Genetic diversity and heterotic grouping of tropical and temperate maize populations adapted to the Northern US Corn Belt. Euphytica, **204**: 661-677.

Marinkovic R., Dozet B. and Vasic D. 2003. Oplemenjivanjesuncokreta. Monografija. Skolskaknjiga, Novi Sad.

Maruthi Sankar G.R., Vanaja M., Maruthi V., Raghuram Reddy P. and Narasimha Murthy D. 2004. Selection of consistent plant traits for sunflower growth using principal component analysis. Helia, **27**(41): 113-122.

Meena H.P., Sujatha M. and Varaprasad K.S. 2013. Achievements and bottlenecks of heterosis breeding of sunflower (*Helianthus annuus* L.) in India. Indian J. Genet. Plant Breed., **73**: 123-130.

Mohammadi S.A. and Prasanna B.M. 2003.Analysis of genetic diversity in crop plants-salient statistical tools and considerations. Crop Sci., **43**: 1235-1248.

Mollasadeghi V., Imani A.A., Shahryari R. and Khayatnezhad M. 2011.Classifying bread wheat genotypes by multivariable statistical analysis to achieve high yield under after anthesis drought. Middle-East J. Scientific Res., **7**: 217-220.

Murthy B.R. and Arunachalam V. 1966. The nature of genetic divergence in relation to breeding system in crop plants. Indian J. Genet. Plant Breed., **26A:** 188-189.

Nooryazdan H., Serieys H., Bacilieri R., David J. and Berville A. 2010. Structure of wild annual sunflower (*Helianthus annuus* L.) accessions based on agro-morphological traits. Genet Resour. Crop Evol., **57**: 27-39.

Panse V.G. and Sukhatme P.V. 1964.Statistical methods for agricultural workers, 2nd edn. ICAR, New Delhi.

Praduman Yadav and Murthy I.Y.L.N. 2016.Calibration of NMR spectroscopy for accurate estimation of oil content in sunflower, safflower and castor seeds. Curr. Sci., **110**(1): 73-76.

Radić V., Mrđa J., Terzić S., Dedić B., Dimitrijević A., Balalić I. and Miladinović D. 2013. Correlations and path analyses of yield and other sunflower seed characters. Genetika, **45**(2): 459-466.

Razi H. and Assad M.T. 1999. Comparison of selection criteria in normal and limited irrigation in sunflower. Euphytica, **105**: 83-90.

Riaz A., Iqbal M.S., Fiaz S., Chachar S., Amir R.M. and Riaz B. 2020. Multivariate analysis of superior *Helianthus annuus* L. genotypes related to metric traits. Sains Malays., **49**(3): 461-470.

Sahari K., Khalil K., Silva J.A.T., Nadia A., Salah R. and Fayçal B.J. 2013. Assessment of diversity of phenologically and morphologically related traits among adapted populations of sunflower (*Helianthus annuus* L.). Helia, **36**(58): 29-40.

Sefaoglu F. 2023. Investigation on the genotype and environmental relationships on yield and its components in sunflower (*Helianthus annuus* L.) in eastern region of Türkiye. Indian J. Genet. Plant Breed., **83**(1) : 77-87.

Shojaei S.H., Ansarifard I., Mostafavi K., Bihamta M.R. and Zabet M. 2022. GT biplot analysis for yield and related traits in some sunflower (*Helianthus annuus* L.) genotypes. J. Agric. Food

Res., **10**: 100370.

Sincik M. and Goksoy A.T. 2014. Investigation of Correlation between Traits and Path Analysis of Confectionary Sunflower wvGenotypes. Not. Bot. Horti. Agrobot., **42**(1): 227-231.

Singh S. and Pawar I.S. 2005. Theory and Application of Biometrical Genetics. India: CBS Publishers.

Sujatha M., Sudhakara Babu S.N., Lakshmi Prayaga, Meena H.P., Dudhe M.Y., Santha Lakshmi Prasad M., Satya Srinivas P., Sakthivel K., Satish Kumar G.D., Jawahar Lal J., Praduman Yadav and Boopathi T. 2022. 50 Years of Research Achievements in AICRP-Sunflower, 2022. ICAR-Indian Institute of Oilseeds Research, Hyderabad 500 030. P 182.

Tousi Mojarad M. and Bihamta M.R. 2007. Evaluation of grain yield and other quantitative attributes in bread wheat via factor analysis. J. Agric. Sci., **17**: 97-107.

Velkov V.N. 1980. Relation between yield and some characters in sunflower.GeneticaSelektsiya, **13b**(5): 329-338.

Yan W. and Kang M.S. 2019. GGE-Biplot Analysis: A Graphical Tool for Breeders, Geneticists, and Agronomists. Boca Raton: CRC Press, Boca Raton, FL, USA ISBN: 978-0-8493-1338-4.Pp 288.

Yan W. and Rajcan I. 2002. Biplot analysis of test sites and trait relations of soybean in Ontario. Crop Sci., **42**(1): 11-20.

Yildirim M.B., Caliskan C.F. and Cagirgan M.I. 1989. Factor analysis of various traits in potato. J. Agric. Faculty, **5**(1): 93-106.

Supplementary Table S1. Distribution of 358 sunflower inbreds in different clusters and their origin

Cluster	Number of genotypes	Genotypes	Origin or source
Cluster I	33	SCG-56, TSG-210, TSG-297, TSG-347, EC-399512, EC-276189, EC-585633, SCG-4, RGP-50-P_2, RGP-52-P_3, RGP-61-P_5-S_1, SCG-46, HA-228B, HA-248B, IC-623023, SS-2033, ID-32, GMU-520-1, CMS-243B (BL), HOCL-8, RCR-72, GPR-70, GP_6-18, COSFV-5, RGP-50-P_2-S_1, RGP-58-P_4-S_2, RGP-19-P_1-S_3, RGP-30-P_1-S_1, RGP-50-P_1-S_1, RGP-60-P_4, GMU-802, GMU-960, GMU-802	USDA, USA India (Coimbatore, Hyderabad, Ludhiana)
Cluster II	100	TSG-257, TSG-271, TSG-348, TSG-366, TSG-369, EC-93901, EC-93615, EC-512739, EC-175717, EC-198075, EC-198089, EC-198101, EC-49442, EC-494357, EC-318766, EC-276582, EC-178159, EC-242765, EC-210309, EC-276173, TSG-76, SCG-11, SCG-31, SCG-22, TSG-104, SCG-331, SCG-66, ID-33, HA-303B, HA-112B, HA-234B, HA-232B, EC-198052, RHA-378, RHA-379, EC-625783, EC-191897, EC-198052, EC-601933, EC-602060, RGP-58-P_4-S_1, RGP-58-P_4-S_4,HA-249B, RGP-58-P_4-S_3, NDI-6, NDI-7, NDI-11, CPI-1, CPI-3, GP_6-1254, CPI-12, NDLB-8, IC-601951, IC-625745, IC-640320, GP_6-1116, GMU-200, LSF-8, 99RT, LTRR-341, R_3Brown (Br), R_3 Black (Br), R-8, R-393, AKSF-6R, GP_6-127, GMU-1062, GP-1044-100, DRSF-113, GPR-135, GPR-85, COSF-12B, COSF-15B, ARM-250B, PSGRM-129, GP_6-325, GP_6-951, R-127-1, GP_6-1075, GPB-33, GPB-4, GPB-53, GPB-52, GPB-35, GPB-3, RCR-62, RHA-274, EC-601933, GMU-336, GMU-313, GMU-351, GMU-242, GMU-311, GMU-370, GMU-575, GMU-377, GMU-1085, GPB-100, GP_6-847, GP_6-18, RHA-6D1 ©	USDA, USA India (Akola, Bengaluru, Coimbatore, Latur, Hyderabad, Nandyal)
Cluster III	15	SCG-23, SCG-46, TSG-284, EC-512700, EC-399490, SCG-52, EC-191897, RHA-1-1, CMS-38B, CMS-7B, RCR-428, GPR-125, RGP-93-P1-S1, GMU-1190, GP_6-861	USDA, USA India (Bengaluru, Hyderabad, Raichur)
Cluster IV	09	SCG-59, TSG-269, TSG-376, SCG-103, HA-302B, NDI-12, CPI-2, R-83, GP_6-126	USDA, USA India (Nandyal)
Cluster V	163	TSG-92, TSG-101, TSG-110, SCG-59, ID-34, TSG-114, TS-238, TSG-274, TSG-275, TS-283, TSG-23, EC-85818, SCG-36, EC-276274, EC-601803, EC-178168, EC-178170, EC-276465, RHA-834, EC-210237, EC-198073, ID-31, ID-97, TSG-283, RGP-61-P_5-S_1, IC-625774, IC-438395, Morden, GP_6-109, R-272, RHA-138-2, ID-1-1, GMU-1080, GMU-420, PM-155, R-127-1, R-77-2, R-630, AK-345-1, PM-134, GMU-1186-1-4, GMU-1080, GP-53-18, GMU-370, GMU-1006, GPR-81, GPR-93, GPR-122, HOCL-6, RCR-72, GPR-70, GPR-78, CMS-104B, CMS-110B, CMS-246B, COSF-3B, COSF-14B, TSG-240, EC-512700, EC-601807, GP_6-11, GP_6-883, GP_6-916, GP_6-965, GP_6-967, EC-101241, EC-113042, GP_6-228, GP_6-1063, GP_6-1060, RGM-41, RGM-49, GPN-45, EC-10751, GP-1072, GP_6-792, GP_6-139, GP_6-1102, GP_6-1026, GP_6-286, GP_6-160, GP_6-219, GMU-444, GMU-490, GP_6-887, P-107R, IL-85, IL-84, IL-82, IL-81, GPB-15, LTRR-83-273, GPB-13, GPB-31, GPB-43, GPB-29, GPB-19, EC-601964, RHA-284, RHA-298, GPB-1, CSFI-8002, CSFI-99, 150R, PM-66, PM-81, PM-82, RGP-21-P_4-S_{1-3}, RGP-89-P_2-S_1, CMS-302B, CSFI-13022, CSFI-13023, CMS-335B, GPB-23-2, GPB-26-1, GPB-372, GPB-42-1,GPB-42-4, GPB-85-3, RGP-18-P_3-S_2, RGP-18-P_2, RGP-11-P_1-S_1, RGP-21-P_5, RGP-29-P_2-S_1, RGP-28, RGP-25-P_1-S_1, RGP-21-P_4-S_{1-2}, RGP-21-P_2-S_3, RGP-21-P_6-S_2, RGP-21-P_2-S_2, RGP-21-P_4-S_2, EC-601832, RGP-21-P_4-S_1, RGP-21-P_3-S_1, RGP-32-P_1-S_1, RGP-50-P_2-S_2, RGP-50-P_1-S_3, RGP-50-P_1-S_1, RGP-50-P_3, RGP-50-P_2, RGP-50-P_1, RGP-58-P_4-S_3, RGP-60-P_2, RGP-60-P_2-S_1, RGP- 84-P_2, RGP-89-P_2-S_2, RGP-92-P_1-S_2, RGP-93-P_1-S_3, GMU-46, GMU-377, GMU-510, GMU-189, GMU-333, GMU-477, GMU-759, GMU-802, GMU-306, GMU-379, GMU-357, GP_6-906, ARM-243B©, LTRR-341©, DRSF-113©	USDA, USA India (Akola, Bengaluru, Coimbatore, Hyderabad, Latur, Raichur)
Cluster VI	01	TSG-277	USDA, USA
Cluster VII	36	EC-68413, EC-494384, EC-279309, TSG-102, TSG-260, TSG-339, TSG-391, GPR-87, CMS-248B, CMS-249B, COSF-4B, COSF-6B, COSF-7B, LDM-02, GP_6-1001, GPB-16, LTRR-822, RGP-50-P_1-S_4, RGP-50-P_2-S_2, RGP-58-P_3, RGP-58-P_3-S_1, RGP-58-P_4-S_1, RGP-60-P_3, RGP-61-P_5-S_1, RGP-87-P_3, RGP-89-P_2-S_3, RGP-96, RGP-95-P_1-S_1, EC-601762, GMU-389, GMU-498, GMU-479, GMU-768, GMU-363, GMU-349, GMU-307	USDA, USA India (Bengaluru, Coimbatore, Hyderabad, Latur)
Cluster VIII	01	RGP-60-P_2-S_2	India (Hyderabad)