

GENETIC DIVERGENCE AMONG OPEN-POLLINATED POPULATIONS OF SUNFLOWER (*Helianthus annuus* L.)

M. del P. Alvarez, N. Mancuso, E. Frutos.

Instituto Nacional de Tecnología Agropecuaria - EEA Pergamino, CC 31, 2700 Pergamino (BA).
República Argentina

ABSTRACT

Genetic variability is essential to achieve success in sunflower breeding. Traditional varieties and populations are important sources of genetic variation. The objective of this study was to evaluate genetic variability in 20 populations from different geographical origins, and estimate their potential usefulness in sunflower breeding. Three experiments were sowed at INTA Pergamino (Argentina) during 1991/92 and 1992/93. Seventeen traits were measured in randomized complete-block designs with four replications. Analysis of variance, estimation of genetic parameters (heritability, genetic coefficient of variation and genetic advance by selection), Principal Components and Cluster Analysis were performed. All characters presented significant differences ($P < 0.01$) between genotypes. Days to flowering, flowering length, number of leaves, plant height, total number of achenes, grain yield and oil percentage should have better possibilities of success in a selection program according to genetic coefficient of variation and genetic advance, for this group of materials. Principal Components Analysis showed that days to flowering, days to maturity, number of leaves, plant height, total number of achenes, grain yield, 1000-achene weight and protein percentage had the best discriminatory power. Cluster Analysis was performed with the first five principal components (86 % of total variability), determining three groups of similarity which generally agreed with geographical origins. The group formed mainly by Russian genotypes had higher means for grain yield and oil percentage.

KEY WORDS: sunflower - genetic divergence - open-pollinated populations - genetic parameters - Multivariate Analysis.

INTRODUCTION

Genetic variability is essential to achieve success in genetic breeding. Sunflower (*Helianthus annuus* L.) breeding has to keep pace with the importance of this crop in oil production and increasing needs of food production. In order to do so it is necessary to satisfy short-term requirements through the utilization of a little proportion of inheritable genetic diversity, besides meeting long-term needs, for which is essential to generate and maintain a broad genetic variability.

Diffusion of sunflower hybrid cultivars and progressive replacement of traditional varieties, has started to erode genetic reserves that are important sources of variability for breeding. Sala et al (1991) reported that several populations and traditional commercial varieties from Argentina, have disappeared or are about to disappear.

The objective of this research was to study genetic divergence of twenty open-

pollinated populations of sunflower and estimate their potential usefulness in sunflower breeding.

MATERIALS AND METHODS

Twenty open-pollinated populations of sunflower from different geographical origins were evaluated: 1) Balcarce 1, 2) Antilco, 3) Co P5, 4) Aguaribay, 5) KLM, 6) PGRL, 7) Cordobés, 8) Guayacán, 9) Negro Bellocq, 18) Co P6 y 20) Co P1 (Argentina); 11) Chernianka II, 12) Peredovik, 16) Chakinskij, 17) VNIIMK 8883 y 19) Armavirskij (Russia); 13) T-407, 14) T-219 y 15) T-907 (Turkey) e 10) Iregi Szurke Csiko (Hungary). Evaluations were performed on field experiments at EEA INTA Pergamino (Buenos Aires, Argentina) during 1991/92 and 1992/93 in three environments determined by one sowing date in the first year and two sowing dates in the second. The following traits were measured: days to flowering (E-F), flowering length (FL), days from flowering to maturity (F-M), days of total growing period (GP), number of leaves (NL), leaf width (LW), stem diameter (SD), plant height (PH), head diameter (HD), total number of aquenes (SN), grain yield (GY), 1000-aquene weight (SW), filled aquene percentage (FS), aquene oil percentage (OP), oleic acid percentage (OLP), linoleic acid percentage (LIP) and protein percentage (PP).

A randomized complete block design with four replications was used. Experimental plot was formed by four rows of five metres long. Data were tested for aditivity, normal distribution and variance homogeneity, according to analysis of variance assumptions (Montgomery, 1991).

Analyses of variance were done individually for each environment and combined accross environments. From expected mean squares of combined analyses, components of genotypic variance, genotype-environment interaction variance and experimental error variance were estimated. Broad sense heritability (H^2) was obtained from these components, as reported by Hanson (1963). Genetic coefficient of variation (GCV) and genetic advance by selection (GA) were estimated according to Burton (1952) and Fehr (1987), respectively. GA was obtained for a selection intensity of 20 %.

Principal Components Analysis (PCA) and Cluster Analysis (CA) were performed. PCA was carried out on the correlation matrix of original characters, obtained from means of genotypes across replications and environments. CA was calculated by a nonhierarchical method, by means of FASTCLUS procedure (SAS, 1990). This procedure uses a method called nearest centroid sorting, performing a disjoint cluster analysis on the basis of Euclidean distances. The first five principal components axes (PC) were used as a measure of dissimilarity. Clustering structure visualized on PC plots, "cubic clustering criterion" and pseudo-F statistic (SAS, 1990), and iteration during clustering process were the criteria considered in order to find out the accurate number of groups the material is divided into.

RESULTS

All variables fitted analysis of variance assumptions, with exception of filled aqunes percentage that had to be transformed as "arcsine of proportion". Combined analysis of variance showed that genotypes were significantly different ($P < 0.01$) for all characters studied. Genotype-environment interaction mean square was highly significant in most cases, only filled aqune and protein percentages did not had significant interaction with environment. Table 1 displays estimated genetic parameters.

Table 1 - Broad sense heritability, genetic coefficient of variation and genetic advance by selection, for seventeen characters in twenty sunflower populations.

TRAITS	H ²	G.C.V..	G.A.(mean %)
E-F	93.51	8.45	11.05
FL	67.12	11.16	12.36
F-M	81.26	7.85	9.57
GP	87.07	4.51	5.69
NL	90.12	12.49	16.00
LW	60.23	3.89	4.08
SD	64.50	6.34	6.87
PH	84.82	10.55	13.12
HD	62.32	5.02	5.36
AN	71.25	14.22	16.08
GY	73.62	17.19	19.85
AW	43.02	5.52	4.91
FA	71.57	7.62	8.70
OP	93.99	12.37	16.20
OLP	53.30	3.85	3.81
LIP	53.20	2.21	2.17
PP	80.81	5.03	6.11

Table 2 shows ACP results: eigenvalues, proportion of total variation and cumulative proportion of variation for the first five PC.

Table 2 - Eigenvalues, proportion of total variation and cumulative proportion of variation for the first five PC.

	Eigenvalue	Proportion	Cumulative
PRIN 1	4.96	0.310	0.310
PRIN 2	3.69	0.230	0.540
PRIN 3	2.56	0.160	0.700
PRIN 4	1.55	0.097	0.797
PRIN 5	0.97	0.061	0.858

Table 3 exhibits the first five PC coefficients for every trait included in the analysis.

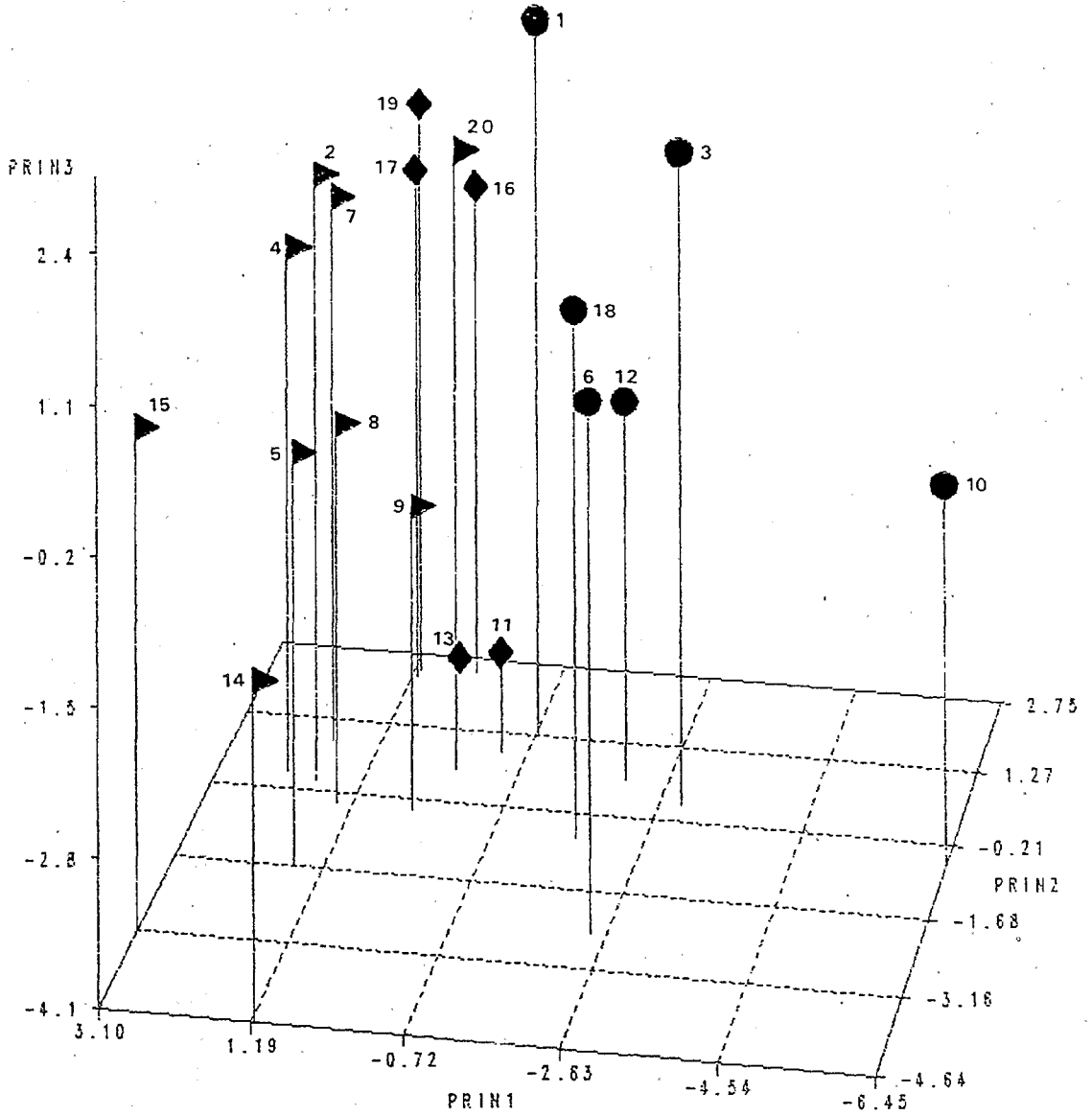
Table 3 - Eigenvectors for the first five PC.

TRAITS	PRIN 1	PRIN 2	PRIN 3	PRIN 4	PRIN 5
E-F	0.23	-0.42	-0.09	-0.04	-0.10
FL	0.02	0.19	0.38	0.25	0.18
F-M	-0.02	0.10	0.33	0.56	0.25
GP	0.22	-0.32	0.21	0.36	0.10
NL	0.30	-0.34	-0.003	-0.006	-0.24
LW	0.22	0.08	-0.22	-0.17	0.74
SD	0.24	-0.2	-0.31	0.37	-0.11
PH	0.39	-0.12	-0.07	-0.03	-0.01
HD	0.25	0.24	-0.35	0.02	0.21
AN	0.41	0.13	0.10	0.01	0.01
GY	0.34	0.30	0.03	-0.03	-0.11
AW	0.03	0.36	-0.36	0.09	-0.24
FA	0.17	0.28	0.27	-0.30	-0.02
OP	0.16	0.28	0.37	-0.19	-0.27
LIP	0.19	-0.26	0.24	-0.42	0.19
PP	-0.34	-0.08	-0.04	-0.16	0.17

Criteria mentioned to determine which is the appropriate number of groups by AC, agreed on the presence of three similarity groups: cluster 1: Co P5, PGRL, Iregi Szurke Csiko, Peredovik, Co P6 and Balcarce 1; cluster 2: Chernianka, Chakinskij, Armavirskij, VNIIMK 8883 and T-407; cluster 3 Co P1, Negro Bellocq, Guayacán, KLM, Antilco, Cordobés, Aguaribay, T-219 and T-907. Cluster 1 registered mean values of 97 days for total growing period, 1.41 m for plant height, 1800 kg/ha for grain yield and 35.7 % for oil content. Cluster 2 presented mean values of 94 days for total growing period, 1.60 m for plant height, 2860 kg/ha for grain yield and 38.8 % for oil content. The third cluster had means of 103 days for total growing period, 1.73 m for plant height, 2450 kg/ha for grain yield and 36.4 % for oil content. Grain yield values corresponded to a stand of 45000 plants/ha.

Figure 1 combines PCA and CA results. CA, represented by different symbols, accounted for 86 % of total variability and supplements spatial distribution accomplished by the first three PC axes, which show 70 % of variation between populations. Thus, where two materials seem close but are assigned to different clusters, the complementary variance not shown in the threedimensional plot has caused the assignment to different groups according to CA.

Figure 1 – Asociation among populations, based on PCA and CA



Cluster 1: circles Cluster 2: diamonds Cluster 3: flags
Numbers correspond to the designation of populations

DISCUSSION

Analysis of variance revealed the presence of a broad variability in the set of sunflower populations studied. Days to flowering, growing period, number of leaves, plant height and oil percentage had higher values of H^2 (Table 1). This parameter established the proportion of total variability determined by genetic effects (Fehr, 1987), for the collection of genotypes considered; and is a guide in the election of the most efficient breeding method for these traits. GCV and GA specified that days to flowering, days to maturity, number of leaves, plant height, total number of achenes, grain yield, 1000-achene weight and protein percentage should have better possibilities of success in a selection program for the group of populations studied.

Percentage of oleic acid was highly correlated (0.99 **) with percentage of linoleic acid. Consequently, the percentage of oleic acid was not considered on PCA. Both characters inclusion disturbs the results of PCA, as it was based on the correlation matrix. Total variability could be reduced to the first five PC, as the fifth PC presented an eigenvalue of 0.7, and 86 % of total variation (Table 2) was accumulated by these five PC axes (Jolliffe, 1986). Coefficients of most important variables in each PC, are highlighted on Table 3. The first PC was mainly explained by overall size of plants and grain production and it expressed the contrast between these traits and protein percentage. Traits with higher absolute value coefficients in the two first PC axes were those which mainly expressed variability among genotypes. CA made it possible to identify three similarity groups that generally agreed with the geographical origin of populations. Cluster 1 was the most heterogeneous in the origin of populations and presented in average lesser values of grain yield and oil percentage. Cluster 2, formed mainly by Russian materials, had in average a short growing period, and high grain yield and oil percentage. The third group was mostly represented by Argentine types that had in average the longest growing period, the tallest plants, and medium means of grain yield and oil percentage.

CONCLUSIONS

- * The existence of broad variability was determined in the germplasm studied.
- * Days to flowering, days to maturity, number of leaves, plant height, total number of achenes, grain yield, 1000-achene weight and protein percentage, should have better possibilities of success in a selection program for the group of populations studied.
- * Days to flowering, days to maturity, number of leaves, plant height, total number of achenes, grain yield, 1000-achene weight and protein percentage explained most variability among materials.
- * Three groups of similarity that generally agreed with geographical origin of materials were identified. Therefore, the genetic divergence was explained by the origin of populations.

* Cluster formed mainly by Russian genotypes presented higher values of grain yield and oil percentage, so it turned out to be the most helpful one in a breeding program for these traits.

REFERENCES

- Burton, G. W. (1952). "Quantitative Inheritance in Grasses". Proc. 6th. Int. Grassld. Cong., 1: 277-283.
- Fehr, W. R. 1987. "Principles of Cultivar Development". Iowa State University. 536 p.
- Hanson, W. D. 1963. "Heredability". En Hanson W. D. and Robinson H. ed. Statistical Genetics and Plant Breeding". Washington D.C., National Academy of Sciences. Pp.125.
- Jolliffe, I. 1986. "Principal Components Analysis". Springer Verlag ed. New York. 485 p.
- Montgomery, D. C. 1991. "Design and Analysis of Experiments". Second Edition. J. Wiley and Sons. New York. 535 p.
- Sala, C. A.; V. R. Pereyra; R. H. Rodriguez y M.E. Bazzalo. 1991. "Colección de germoplasma de *Helianthus* L. de la Estación Experimental Agropecuaria del INTA de Balcarce". Evolución. Primera Reunión Nacional de Oleaginosos. Rosario.
- SAS Institute Inc., 1990. SAS/STAT User's guide, Version 6, Cary Inc.: SAS Institute Inc.