# GENOTYPE BY ENVIRONMENT INTERACTION AND INDIRECT SELECTION IN SUNFLOWER FOR ARGENTINA. I. MULTI-ATTRIBUTE TWO-MODE PATTERN ANALYSIS

Abelardo J. de la Vega, Advanta Semillas S.A.I.C., Ruta Nac. 33 Km 636, CC 294, (2600) Venado Tuerto, Argentina Fax: +54 3462 435231 ; e-mail: avega@waycom.com.ar

Scott C. Chapman, CSIRO Tropical Agriculture, 120 Meiers Rd., Indooroopilly, Qld 4068, Australia Fax: +61 7 32142288 ; e-mail: scott.chapman@tag.csiro.au

Antonio J. Hall, IFEVA, Facultad de Agronomía, Universidad de Buenos Aires, Av. San Martín 4453, (1417) Buenos Aires, Argentina Fax: +54 11 4514 8730 ; e-mail: hall@ifeva.edu.ar

## Summary

A reference set of 10 sunflower hybrids was evaluated in 21 environments of Argentina, to identify patterns of genotype × environment (G×E) interaction and opportunities for indirect selection. The locations covered subtropical (N) and temperate (C) growing environments. Three managed environments were included to assess whether they provide discriminations among genotypes relevant for the target population of environments and contribute to the understanding of the underlying causes of G×E interactions.

ANOVA showed that, for oil yield, the portion of the total sum of squares that accounted for  $G \times E$  interaction was three times larger than the portion explained by G. Pattern analysis revealed the existence of two mega-environments that corresponded well to the N and C locations. The yield gains under N environments would have been unlikely to occur if selection had been done in C environments. While the C environments discriminated among genotypes in a similar fashion, the N environments were more divergent over both years and locations. Cluster analysis revealed 3 genotypic groups: northern, central and broadly adapted. Late planting dates in a C location associated positively with the N environments, representing an opportunity for indirect selection. When photoperiod was extended to 15.5 h in these trials, genotypes exhibited responses similar to those of normal planting dates in C environments, indicating that photoperiod could be a central factor underlying the observed  $G \times E$  interactions. Pattern analyses of yield components revealed different patterns of discrimination among genotypes and may help focus the search on the physiological bases of the  $G \times E$  interaction.

# Introduction

Sunflower (Helianthus annuus L.) is grown in Argentina on almost four million hectares from Southern Buenos Aires (39°S) to Chaco (26°S). Effective identification of superior genotypes in multi-environment trials is generally complicated by the presence of genotype  $\times$ environment (G×E) interactions, whereby cultivar relative yields vary across different locations and seasons. These interactions reflect differences in genotype adaptation, which may be exploited by selection for broad or specific adaptation. To determine the scope for selection for broad adaptation (minimizing interactions) or specific adaptation (emphasizing favorable interactions), crop environments can be characterized in terms of the way they influence the relative performance of genotypes (Cooper et al., 1996). Pattern analysis (Williams, 1976) combines classification and ordination and has been extensively recommended for identifying genotypes that have similar patterns of response across environments, for identifying environments that are similar in the way in which they discriminate among genotypes, and to examine correlations between environments to exploit opportunities for indirect selection (DeLacy et al., 1996). In the present study, we used a reference set of sunflower genotypes to characterize some sunflower growing environments of Argentina, to define mega-environments, to evaluate the scope for achieving indirect response to selection for yield in some mega-environments from the results of managed environments, and to interpret changes in relative oil yield in terms of the changes in its components.

# **Materials and Methods**

A reference set of 10 sunflower single-cross hybrids (Table 1) was evaluated in 21 central (C), northern (N) and managed (M) environments of Argentina (Table 2). The locations were evenly distributed between subtropical (N) and temperate (C) growing environments. In three M environments (MV27, MV29 and MVI9), the normal October planting date for central locations was delayed until December and in MVI9 the day length was artificially extended to 15.5 h during the whole crop cycle. In each environment, a RCBD with 3 reps was used. Plot data of grain yield (GY, at 11% moisture), oil content (OC), oil yield (OY), 1000 grain weight (W1000) and grain number m<sup>-2</sup> (GN) were recorded for all trials. ANOVA was conducted using S-Plus<sup>1</sup> to examine the partitioning of sums of squares to G, E and G×E interaction. Before pattern analysis, the matrices of genotype means in each environment were environment-standardized (Fox and Rosielle, 1982). For classification, a hierarchical agglomerative clustering method (Williams, 1976) with incremental sum of squares (Ward, 1963) as the fusion criterion was applied to the matrices. The principal components (PCs) of the squared Euclidean distance matrix of each attribute were estimated using a singular value decomposition procedure (Gabriel, 1971). A biplot of the first two PCs for each attribute was constructed from this analysis (Gabriel, 1971). Pattern analyses were carried out using software developed by the University of Queensland<sup>2</sup>.

# Results

Environments accounted for 87.2% of the treatment sums of squares for OY (i.e. excluding residuals) and of the remaining sums of squares, the G×E interaction was 3.1 times that of the contribution of G (data not shown). For GY, the G×E interaction sum of squares was 7.9 times that of the G sum of squares. GN is the OY component where the contribution of G×E

<sup>&</sup>lt;sup>1</sup> S-Plus, MathSoft Inc., Seattle, Washington, USA

<sup>&</sup>lt;sup>2</sup> GEBEI software, available from I.H. DeLacy, Department of Agriculture, University of Qld., Brisbane, 4072, Australia. Web site: http://biometrics.ag.uq.edu.au/software.htm

sum of squares was largest in relation to G (1.7 times), while for W1000 and OC,  $G \times E$  interaction accounted for a less sum of squares than G (0.5 and 0.3 times, respectively).

Table 1. Agronomic characters and genotype grouping (by pattern analysis) of the sunflower hybrids evaluated in Argentina northern (N), central (C), and managed (M) environments (mean values for 21 trials). All genotypes are single cross hybrids, developed by Advanta Argentina sunflower breeding program, except Morgan 734 (Dow-Morgan Argentina). The grouping is that from hierarchical agglomerative clustering of oil yield

Trt.	Code	Status <sup>a</sup>	Maturity	Days to	Oil yield	Grain	Grain	1000	Oil	Grouping
				flowering <sup>b</sup>	$(\text{kg ha}^{-1})$	yield (kg	number	grain	(%)	
						ha <sup>-1</sup> )	m <sup>-2</sup>	weight		
								(g)		
1	Contiflor 15	С	ML	75	1117	2263	4929	43.9	48.2	2
2	Aguará	С	ML	79	1047	2231	5364	40.2	46.2	1
3	GV23105	E	ML	74	1023	2357	4190	53.9	42.9	1
4	GV25015	E	ML	79	905	2362	4777	47.3	37.5	1
5	GV25086	E	М	74	1199	2399	4616	49.3	49.1	3
6	TC 2001	С	ML	76	1089	2161	4551	44.4	49.0	2
7	GV23146	E	ML	78	1232	2503	5065	46.8	48.5	3
8	GV22510	E	ML	74	1125	2328	4563	49.1	47.4	3
9	Contiflor 9	С	М	73	1080	2204	4320	48.2	47.4	2
10	Morgan 734	С	М	73	1057	2367	3729	60.6	44.9	1

<sup>a</sup> C: commercial, E: experimental; <sup>b</sup>50% of the plants showing 50% anthesis

When clustered on OY (Table 1), the genotypes could be separated into three groups of different response patterns across environments (plots not shown here). Group 1 consisted of four entries (2, 3, 4, 10) that showed average to good performance in the N environments, but yields lower than the other groups in the C environments. Group 2 consisted of three hybrids (1, 6, 9) of relatively good performance in the C environments, but with yields that were lower than the other groups in the N environments and group 3 consisted of three broadly adapted hybrids (5, 7, 8), with relatively good performance across all environments. Genotype classification based on the other traits reveals different groupings to that of OY, indicating that different genotypic attributes could underlie the observed G×E interaction for OY. The classification of environments for OY is shown in Table 2. Environment groups 1 and 2 are constituted by N environments and the MV2 trials. Group 3 includes the two lowest yielding N environments, severely affected by rainfall excess. Group 4 includes two C environments and VI9. Group 5 is composed of four C environments, and group 6 includes four C environments and the two highest yielding N environments. The results of the ordination analysis are presented in biplots of the 1<sup>st</sup> and 2<sup>nd</sup> principal components (PC) (Figure 1). Entries that are close together are similar in performance across environments. For any particular environment, genotypes can be compared by projecting a perpendicular from the genotype markers to the environment vector, i.e. entries that are further along in the positive direction of the vector are higher yielding and vice versa (Kroonenberg, 1997). Acute angles between any two environmental vectors indicate positive associations, i.e. they influence the genotypic relative performance in a similar manner; 90° angles indicate no association; and angles greater than 90° indicate negative associations. For OY, the environment vectors covered a wide range of Euclidean space, indicating the existence of strong G×E interactions among the 21 evaluated environments (Figure 1A). The maximum angle among the vectors of C environments is smaller than 90°, suggesting that these environments are relatively similar in the manner they discriminate among genotypes. The maximum angle among the vectors of N environments is larger than 90°, indicating larger interactions within this region. The 1<sup>st</sup> PC contrasts the C and N environments, such that high genotype scores are associated with adaptation to the C environment. The 2<sup>nd</sup> PC was related to adaptation to N environments with the higher yielding entries in this environment type at the top of the diagram. Broadly adapted hybrids tend to be at the top right hand quadrant of the diagram.

On average, the angle between C and N environments tends to be slightly larger than 90°, which suggests that these tend to be two different mega-environments. This difference seems to be repeatable over years according to the discrimination patterns of the environments that were evaluated in more than one season.

Table 2. Attributes and yield components for northern (N), central (C), and managed (M) environments where the reference set of genotypes was evaluated. Managed environments were: V2- Venado Tuerto, December planting, and VI- Venado Tuerto, December planting with artificially extended photoperiod to 15.5 h during the whole crop cycle. Agronomic traits are means for 10 hybrids. Grouping according to hierarchical agglomerative clustering of oil yield

Location	Region	Trial code	Latitude	Sowing	Oil yield	Grain	Grain	1000	Oil	Grouping
		(location,		date	(kg ha <sup>-1</sup> )	yield (kg	number	grain	(%)	
		year)				$ha^{-1}$ )	$m^{-2}$	weight		
								(g)		
Las Breñas	Ν	LB7	27.1 S	02/10/96	441	1077	1868	54.4	41.3	1
Orán	Ν	OR7	23.0 S	18/09/96	1290	2998	6243	45.9	43.0	2
Reconquista	Ν	RE7	29.1 S	11/09/96	573	1213	3164	36.6	47.1	2
V. Tuerto	С	VT7	33.7 S	28/10/96	1608	3349	6547	50.5	48.1	4
V. Tuerto	Μ	V27	33.7 S	14/12/96	1058	2565	5696	43.5	41.1	1
9 de Julio	С	9J7	35.5 S	25/10/96	1125	2398	5207	44.3	46.9	5
Reconquista	Ν	RE8	29.1 S	28/08/97	337	736	2255	33.1	45.9	3
Margarita	Ν	MA8	29.7 S	26/08/97	346	794	2748	30.8	43.8	3
V. Tuerto	С	VT8	33.7 S	29/10/97	1007	2141	4558	47.4	46.8	5
9 de Julio	С	9J8	35.5 S	24/10/97	764	1535	3493	44.4	49.4	6
Daireaux	С	DX8	36.6 S	23/10/97	1037	2191	4260	51.8	47.0	4
Orán	Ν	OR9	23.0 S	24/09/98	534	1356	3347	39.7	39.2	2
Reconquista	Ν	RE9	29.1 S	24/08/98	1366	2713	5002	53.7	50.4	6
Margarita	Ν	MA9	29.7 S	05/09/98	1602	3124	5478	56.4	51.2	6
Paraná	Ν	PA9	31.7 S	30/10/98	1206	2516	4989	49.9	47.9	1
V. Tuerto	С	VT9	33.7 S	22/10/98	1992	4118	6259	64.9	48.4	5
V. Tuerto	Μ	V29	33.7 S	19/12/98	950	2332	5321	45.1	40.8	2
V. Tuerto	М	VI9	33.7 S	19/12/98	709	1768	4505	36.8	39.4	4
9 de Julio	С	9J9	35.5 S	07/10/98	1628	3210	5501	61.4	50.6	6
Daireaux	С	DX9	36.6 S	10/10/98	1782	3519	6123	56.7	50.6	6
Sampacho	С	SA9	33.3 S	01/10/98	1449	3009	4600	67.6	48.2	5

For GY, the environment vectors covered the total range of Euclidean space, showing on average a strong negative association between C and N environments in the 1<sup>st</sup> PC (Figure 1B). For GN, W1000, and OC, the 1<sup>st</sup> PC largely reflected the average standardized value of the hybrids across the environments (Figures 1C to 1E). GN, W1000, and OC 2<sup>nd</sup> PCs appear to be related to  $G \times E$  interaction for these traits, with most of the N environments with negative values and most of the C environments with positive values for these PCs (Figures 1C to 1E).

When normal October planting dates in a C location (Venado Tuerto) were delayed to December (V27 and V29), they associate positively with the N environments for OY and its components. When photoperiod was extended to 15.5 h in this environment (trial VI9), genotypes reverted to responses similar to those of normal planting date C environments (Figures 1A to 1E).

## Discussion

The relative contributions of G and G×E interaction to the total sum of squares for oil yield in this study indicate that it is difficult to achieve an indirect response to selection over all of the target population of environments for sunflower in Argentina from selection in a few environments, ignoring the observed G×E interactions. The large and regional nature of the observed G×E interactions require testing strategies structured to accommodate their effects by selecting for broad or specific adaptation.





Figure 1. Biplot of the 1<sup>st</sup> and the 2<sup>nd</sup> principal components for oil yield (A), grain yield (B), grain number (C), grain weight (D), and oil content (E) of 10 sunflower hybrids grown in 21 environments. Genotypes are represented by points and environments are represented by vectors. Same entry markers indicate genotype groups with members of a similar response pattern at the 3-group level for oil yield: N adapted hybrids (2, 3, 4, 10), C adapated hybrids (1, 6, 9), broadly adapted hybrids (5, 7, 8). Environment symbols are: northern environments (black triangles), central environments (open triangles), managed environments (open circles).

Pattern analysis showed extreme differences in the discrimination effects of N and C environments on hybrid performance (they are largely orthogonal Fig. 1A), indicating the existence of two mega-environments, i.e. yield gains under N environments would have been unlikely to occur if selection had been done only in C environments and vice versa. The relatively good performance of the hybrids of genotype group 3 across both C and N

environments suggests that exploiting G×E interaction through simultaneous selection for both types of environments is possible. But with no subdivision, only broad adaptation can be exploited. The nature of the observed G×E interactions for sunflower oil yield in Argentina, their repetitiveness, and the value of this seed market would justify the strategy of dividing the target population of environments into two mega-environments (C and N) and targeting hybrids separately for each one. Some hybrids of contrasting response pattern when considering yield components individually, showed similar patterns of relative performance across environments for OY. This suggests the existence of different specific genotype responses to specific environmental challenges within the same genotype groups, and would imply that more than one ideotype could be formulated for each mega-environment. Most of the sunflower breeding stations in Argentina are located in the C mega-environment. The value of the seed market of the sunflower N mega-environment is high enough to develop specifically adapted hybrids to it, but does not justify the establishment of a breeding station in situ. The use of managed-environments is an attempt to establish a selection regime which provides discrimination among genotypes relevant to the target production system (Cooper et al., 1995). When normal October planting dates in Venado Tuerto were delayed to December (V27 and V29) they associated positively with the N environments. When photoperiod was extended to 15.5 h in these trials, genotypes reverted to responses similar to those of the normal planting date C environments. This strongly suggests that photoperiod is the environmental challenge underlying the observed G×E interactions (perhaps via grain number) that is driving the association between the December planting C managedenvironments and the N environments. This situation represents an opportunity to exploit indirect responses to selection. One way of incorporating this knowledge would be to conduct early-stage screening of large numbers of hybrids during a late-season planting at Venado Tuerto. Selected hybrids could then go into testing in the N region that was structured to ensure adequate sampling of environments, in particular the year-to-year variation.

### Acknowledgments

We would like to thank Advanta Semillas (Ariel Lorenzo, Ricardo Siciliano, Alan Scott) for making this research possible. We also thank Aldo Martínez, Sergio Solián, Carlos Ghanem, Daniel Kennedy and the others technicians of the Sunflower Breeding Program of Advanta Argentina, for collaborating in the field experiments, and Joanne Walker, for her assistance in the pattern analyses.

### References

Cooper, M., Woodruff, D.R., Eisemann, R.L., Brennan, P.S., and DeLacy. I.H., 1995. A selection strategy to accommodate genotype-by-environment interaction for grain yield of wheat: managed-environments for selection among genotypes. Theor. Appl. Genet. 90: 492-502.

Cooper, M., Brennan, P.S., and Sheppard, J.A., 1996. A strategy for yield improvement of wheat which accommodates large genotype by environment interactions. In: M. Cooper, and G.L. Hammer, (Editors), Plant Adaptation and Crop Improvement. CAB INTERNATIONAL 1996, pp. 487-512.

DeLacy, I.H., Basford, K.E., Cooper, M., Bull, J.K., and McLaren, C.G., 1996. Analysis of multi-environment trials – An historical perspective. In: M. Cooper, and G.L. Hammer, (Editors), Plant Adaptation and Crop Improvement. CAB INTERNATIONAL 1996, pp. 39-124.

Fox, P.N., and Rosielle, A.A., 1982. Reducing the influence of environmental main-effects on pattern analysis of plant breeding environments. Euphytica 31: 645-656.

Gabriel, K.R., 1971. The biplot-graphical display of matrices with applications to principal component analysis. Biometrika 58: 453-467.

Kroonenberg, P.M., 1997. Introduction to biplots for G×E tables. Research Report #51. Centre for Statistics. The University of Queensland, Brisbane, Qld 4072 Australia.

Ward, J.H., 1963. Hierarchical grouping to optimise an objective function. J. Amer. Stat. Assoc. 58: 236-244. Williams, W.T., 1976. Pattern analysis in agricultural science. Elsevier Scientific Publishing Company, Amsterdam.