UDC 575: 633.15
DOI:10.2298/GENSR0803301B
Original scientific paper

GENOTYPE BY ENVIRONMENT INTERACTION IN MAIZE BREEDING

Milosav BABIĊ, Violeta ANĐELKOVIĊ and Vojka BABIĊ

Maize Research Institute, Slobodana Bajica 1, 11080 Zemun-Belgrade

Babić M, V. Anđelković, and V. Babić (2008): G*enotype by environment interaction in maize breeding* – Genetika, Vol. 40, No. 3, 303 -312.

Because proximity measures occur in pairs where both, similarity and dissimilarity measures exploit the same type of information, companion classification and ordination techniques can be applied. They complement each other in analysis of genotype by environment interaction (GxE) data. Choice of method, companion diagnostics and graphical presentation are required within each of methodologies.

By clustering of 12 genotypes into 5 groups, 96.26% of variability for genotypes contained in original data is kept. By applying same analysis for environments, 96.45% of variability contained in original data matrix is kept with grouping of 31 environments into 11 groups. Caused by genotypes and environments grouping 78.10% of GxE variability contained in original data matrix remained in analysis of such two-way reduced data. Based on shown results, it is not possible to define smaller growing regions. Clustering of environment can be useful not only for

Corresponding author: Milosav Babić, Maize Research Institute, Zemun Polje, Belgrade, Serbia, S. Bajića 1, 11185 Zemun Polje, Serbia

defining mega environments but also for smaller growing regions defining only in combination with some of additional analysis (AMMI, discrimination analysis, correspondent analysis etc.). In such kind of analysis experience of investigator would be of great importance. Choice of test sites for breeding programme can be made based on obtained grouping to a limited extent (rather for restructuring existing test sites network in order to obtain "better" information with same number of test sites then for its rationalization with number of test sites decreasing).

Key words: cluster, GxE interaction, maize, pattern analysis.

.

INTRODUCTION

From the very beginning of organized breeding based on science, Genotype by Environment interaction has been an important and challenging issue among plant breeders, geneticists and agronomists engaged in performance testing. Experiments in single environment (location or year) do not allow drawing a general conclusion regarding the tested genotypes. Breeders want to know how the genotype reacts in wide range of environments. So, Multi Environment Trials are essential for breeding process. From breeders stand point, Genotype by Environment interaction is noticeable (significant) when genotypes being evaluated rank differently in different environment (locations and/or years).

The genotype by environment interaction is an important aspect in both, plant breeding programs and the introduction of new maize hybrids. For dealing with Genotype by environment interaction, numerous models have been proposed. These include regression on the environmental mean (FINLAY and WILKINSON, 1963), Pattern analysis methods (BYTH *et al.,* 1976.), principal coordinate analysis (EISEMANN, 1981), canonical variate analysis (SEIF *et al*., 1979 and principal component analysis (GOODCHILD and BOYD, 1975; KEMPTON, 1984; GAUCH, 1988; ZOBEL *et al*., 1988. Each has proved successful in the analysis of univariate GE data in certain situations.

The data structure typically emerging from a yield trial is a two-way factorial of genotypes and environments. The awkward situation emerges that the environment main effect causes most of the variation in yields but is wholly irrelevant to selection. Cluster analysis has been used to group locations that discriminate among genotypes in a similar manner or to summarise patterns of genotypic performance across environments. Cluster analysis provides a method for classifying environments or genotypes (ABOU EL FITTOUH *et al.,* 1969, FOX and ROSIELLE, 1982.). ANDERBERG, 1973, gives a review of similarity – dissimilarity measures as well as descriptions for different classification strategies). MUNGOMERY *et al*. (1974) integrated both Principal component analysis and cluster analysis into "Pattern Analysis" that classifies on the basis of Genotype by Environments interaction and can focus on genotype, environments or both.

This precision facilitated formation by cluster analysis of more cohesive groups of genotypes and/or locations for biological interpretation of interactions than it occurred with unadjusted means.

MATERIAL AND METHODS

Twelve maize hybrids were used in this study. Most of investigated hybrids are commercial hybrids widely grown (ZP 578, ZP 677, ZP 684, ZP 680, ZP 704, ZP 434, ZP 571, ZP 360 and ST 600). Remaining three hybrids are perspective ones registered by state variety commission. Macro trials (31 locations) are set up in order to obtain more dependable information's based on which is possible to make decision about new hybrids commercialization. Plot size was 0.1 ha. Planting and harvesting were mechinanized. During the harvesting samples for moisture content were taken. Grain moisture content was electronically tested on GAC II analyzer. Prior to harvesting obtained plant density

For data analysis, pattern analysis was performed, in order to obtain information's about new hybrids performance. Squared Euclidian distance was used as a measure of distance for Wards clustering method (Incremental Sums of squares). As performed analysis includes clustering of genotypes and/or environments, analysis of variance and PCA analysis of such reduced data, the results enable us not only making general decision based on average performance (grain yield) of hybrids, but also more specific decisions regarding target regions for specific hybrids. As check hybrids are, by definition, widely grown hybrids – widely adapted, based on obtained grouping it is also possible to have an idea about new hybrids stability or specific adaptation.

In order to make all environments equally important for classification of genotypes, prior to cluster analysis the data are environment standardized. From the same reason, prior to classification of environments the data are genotype standardized. Also, as, especially for higher fusion levels the distance between clusters are frequently overestimated, at abscissa of dendrograms instead of distance is fusion level.

RESULTS AND DISCUSSION

By clustering of 12 genotypes into 5 groups, 96.26% of genotypes variability contained in original data was kept. At the same time, 96.45% of variability for environments contained in original data matrix was kept by grouping of 31 environments into 11 groups. Caused by such grouping, 42.37% of G x E variability contained in original data remained in analysis of such two way reduced data (Table. 1.).

First steps of clustering resulted in clusters that are geographically close, while further clustering was based on precipitation, soil type and/or growing technology (Fig. 2.). This shows that such kind of classification can be useful for

defining mega environments. For smaller growing regions defining method should be used in combination with some of additional analysis (AMMI, discrimination analysis, correspondent analysis etc.). Choice of test sites for breeding programme can be made based on obtained grouping to a limited extent (rather for restructuring existing test sites network in order to obtain "better" information with same number of test sites then for its rationalization with number of test sites decreasing or for initial setting up of test sites network).

Table 1. Analysis of variance for the partition of the sum of squares for the G x E model for the two-way grouping model

SOURCE	DF	SSQ	MSQ	(%)
Genotypes	11	132.668	12.061	-
Between groups (genotypes)	4	127.712	31.928	96.26
Within groups (genotypes)	7	4.956	0.708	3.74
Environments	30	274.404	9.147	-
Between groups (environments)	10	264.665	26.466	96.45
Within groups (environments)	20	9.739	0.487	3.55
G x E interaction	330	208.329	0.631	-
Between G grp x between E grp	40	88.264	2.207	42.37
Within G grp x within E grp	140	39.594	0.283	19.01
Remainder of interaction	150	80.471	0.536	38.63

Percentage of total sum of squares retained between groups 78.10

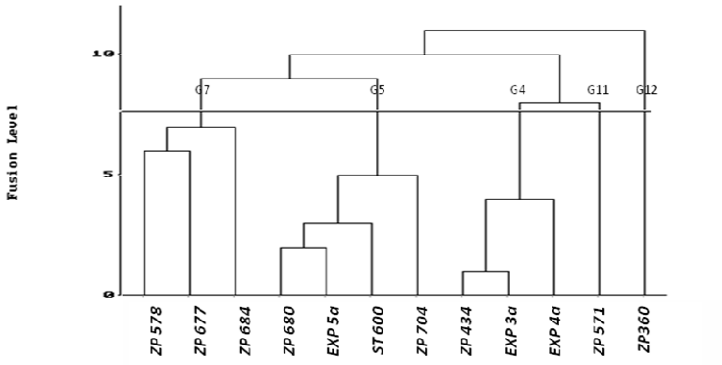

Figure 1. Dendrogram for genotypes

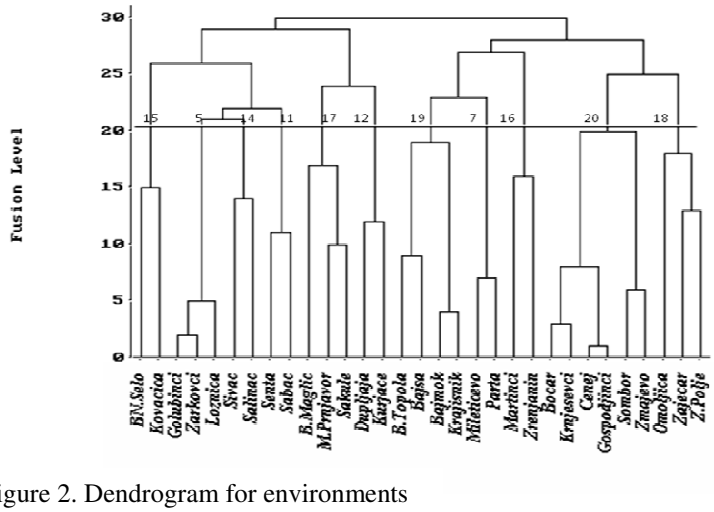

Figure 2. Dendrogram for environments

The first two vectors obtained by PCA cumulatively accounted 42.51% and 19.05% of the total information contained in original GE matrix (Fig. 4). Vector one was found to be closely associated (high negative correlation) to overall mean, but vector 2 had no obvious association with the overall mean or environment means. Second vector might be associated with deviation from regression. In such kind of analysis the main problem is to associate PCA vector (mathematical term) with some biological, biometrical or agronomical term.

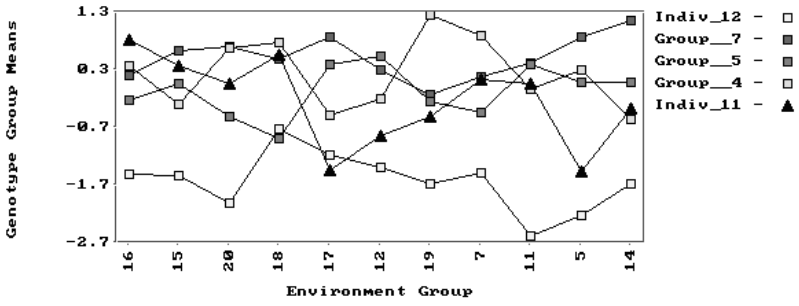

Figure 3. Performance plot

Genotype groups on biplot were quite distinct in space and their members were positioned close together. Group 7 had the highest grain yield in most environments (groups of environments), while grain yield of genotype 12 was the lowest. As group 7 include commercial hybrid ZP 684, high yielding, remaining two members could be of particular interest. This could be expected as all members of group are medium late and already commercial and widely grown. Group 4 is interesting because it is also high yielding group and includes most stable check hybrid, ZP 434. Also its members are positioned very close confirming their close relatedness in agronomic sense (Fig. 1. and Fig. 4.). It was expected regarding relatedness in genetic sense based on pedigree data. As this group represents medium early hybrids, two new hybrids can be interesting, as they posses increased level of drought tolerance as well as ZP 434. It is considered that in our condition out of five years two are with moderate and one with severe drought. It makes group 4 more interesting.

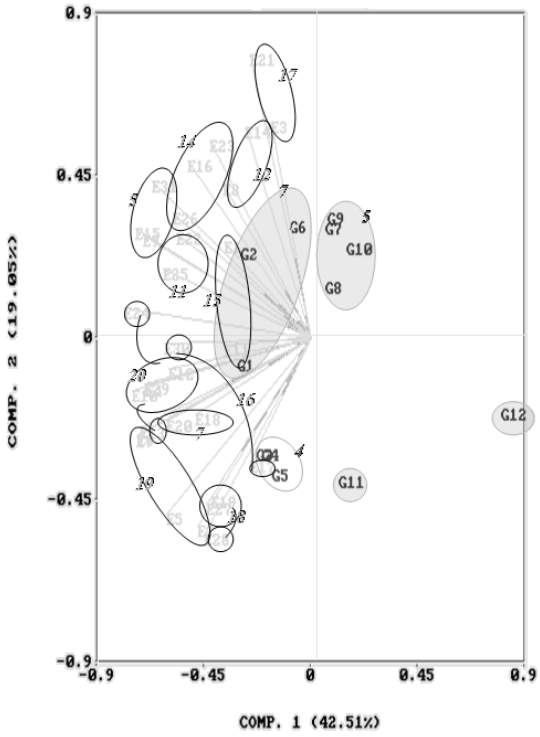

Figure 4. PCA1 – PCA2 Biplot

In order to visually simplify, and get clear idea about genotype (genotype group) performance across environments, it is very useful to draw performance plot (Fig. 3). Even from this plot it is possible to separate high, average and low yielding environments, and also get some idea of variability in performances for particular genotype (genotype group). It can be seen that from the same figure that groups 4 and 7 were the best yielding showing the smallest variations at the same time. Performance plot also gives some idea about discriminating power of individual environments (groups of environments). From the same figure is noticeable that hybrid 12 (ZP 571) could be pulled out from domestic market as it performed lowest yielding. Environment groups that showed best discriminating power were 5, 12, 15, 16 and 20. These groups could be designed as key environments. Table 2. Shows the percentage of between genotype groups by between environment group sums of squares retained for different grouping level. It helps to make decision about which number of groups for environments and genotypes should be retained in analysis which also depends on goal of research. As goal of research in presented investigation were to obtain more detailed information about new hybrids in order of their commercialization as well as classification of environments, we decided to reduce data to five genotype and eleven environment groups, still retaining 42.4% of between genotype groups by between environment group sums of squares. With different goal, only classification of environments, one way reduction of data should be recommended, classification only for environments.

Table 2. The percentage of between genotype group x between environment group sums of squares retained for different grouping levels

		Genotype group						
		2	3	4	5	6	7	8
Environment group	10	3.5	29.1	35.3	40.1	47.6	50.6	56.3
	11	3.7	29.6	35.9	42.4	50.0	53.2	59.2
	12	4.1	30.2	36.5	43.6	51.3	55.4	61.7
	13	4.5	30.9	37.4	44.6	52.3	57.1	63.6
	14	4.6	31.1	37.7	45.0	53.5	58.5	65.3
	15	4.9	31.4	38.0	45.6	54.5	59.6	66.5
	16	5.6	32.1	38.8	46.7	55.5	60.9	68.1
	17	5.6	32.4	39.1	47.5	56.5	61.2	69.1
	18	5.7	32.6	39.7	48.1	57.7	63.3	70.5
	19	5.7	32.6	40.0	48.5	58.2	63.9	71.1
	20	5.7	32.7	40.2	48.7	58.5	64.9	72.3
	21	5.7	32.7	40.2	48.8	58.6	65.2	72.6
	22	5.8	33.4	40.9	49.5	59.3	65.9	73.4
	23	5.9	33.5	41.1	49.7	59.6	66.8	74.4
	24	5.9	33.5	41.1	49.7	59.7	67.1	75.0
	25	5.9	33.6	41.1	50.2	60.2	67.8	75.8

Maize breeders in the sub-region have frequently reported the largest proportion of total variation in yield trials is attributable to environments (BADU APRAKU *et all,* 1995, FACOREDE and ADEYEMO, 1986.). Usually, the genotype and GxE interaction sources are relatively much smaller. Perhaps breeders should be more concerned about uniformity of the environments than on GxE interaction per se. However, the microenvironments and crop management practices may vary widely between locations among and within countries for the same ecological zone (BADU APRAKU *et all*., 2003.). As the same authors outlined, the differential response of genotypes to variable environmental conditions constitutes a major limitation to the identification of superior genotypes for narrow or wide adaptation. From this reason, proper analysis and accurate interpretation of GxE interaction enhance the value of genotypes in regional yield trials. ATLIN *et all.* (2000.) underlined that in small target region, it may be possible to exploit local adaptation to increase gain from selection. Grouping of locations in this investigation did not reveal clear small target regions. More precise classification of environments (locations) supposes to be possible base on data that include more years.

CONCLUSION

Applied analysis can be useful for defining mega environments. For smaller growing regions defining method should be used in combination with some of additional analysis (AMMI, discrimination analysis, correspondent analysis etc.).

Choice of test sites for breeding programme can be made based on obtained grouping to a limited extent (rather for restructuring existing test sites network in order to obtain "better" information with same number of test sites then for its rationalization with number of test sites decreasing).

Depending on goal of research, different resulting intra-group composition and/or inter-group relatedness could be interpreted in several ways: a) Influence of parentage and origin of selection on grouping formed; b) Resolution of the influence of physiological and environmental factors responsible for adaptation c) Prediction and/or inferring adaptation of entries that have been grouped with standard cultivars of known performance d) Identification of key environment that elicit the greatest variation in genotypic response, so that field testing may be limited to those key environments.

Significant Genotype by Environment interaction for quantitative traits, such as grain yield, reduces the usefulness of genotype means, over all environments, for selecting superior genotypes. When Genotype by Environment interaction is significant, its cause, nature and implications should be carefully considered in breeding programs.

ACKNOWLEDGEMENT

This research was supported by the Ministry of Science and Technology, Republic of Serbia, through Project TR 20014.

Received November 24th, 2008
Accepted December 15th, 2008

REFERENCES

ABOU EL FITTOUH, H.A., J.O. RAWLINGS and P.A. MILLER (1969): Classification of environments to control genotype by environment interactions with an application to coton. Crop Sci. *9*: 135-140.

ANDERBERG, M.R.C. (1973): Cluster analysis for applications. New York: Academic Press.

ATLIN G.N., R.J. BAKER, K.B. MCRAY and X. LU (2000): Selection responce in subdivided target regions. Crop Sci. *40*: 7-135.

BADU-APRAKU, B. J.M. FAJEMISIN, A.O. DIALLO (1995): The performance of early and extra-early varieties across environmentsin west and central Africa. pp. 149-159, in: B.Badu-Apraku, M.O. Acoroda, M. Ouedraogo. F.M. Quin (Eds). Contributing to food self-sufficiency: Maize research and developement in West and Central Africa. proceedings of a Regional Maize Workshop, 29 may -2 june, IITA. Cotonou. Benin Republic.

BADU-APRAKU, B., F.J. ABAMU, A. MENKIR, M.A.B. FAKOREDE, K. OBENG-ATWI and C. THE (2003): Genotype by enwironment interactions in the regional early maize variety trials in West and Central Africa. Maydica *48*: 93-104.

BYTH D.E., R.L. EISEMANN and I.H. DELACY (1976): Two way pattern analysis of large data set to evaluate genotypic adaptation. Heredity *37*: 215-230.

EISEMANN R.L. (1981): Two methods of ordination and their application in analysing genotype-environment interactions. In interpretation of plant responce and adaptation to agricultural environments, D.E. Byth and V.E. Mungomery (Eds.). Queensland Branch, Australian Institute of agricultural Sciences, Brisbane, pp 293-307.

FINLAY K.W. and G.N. WILKINSON (1963): The analysis of adaptation in plant breeding programme. Aust. J. Agric. Res. *14*: 742-754.

FOX, P.N., and A.A. ROSIELLE (1982): Reducing the influence of environmental main effects on patern analysis of plant breeding environments. Euphutica *31*: 645-656.

GAUCH H.G. (1988): Model selection and validation for yield trials with interaction. Biometrics 44: 705-715.

GAUCH H.G. and R.W. ZOBEL (1988): Predictive and postdictive success of statistical analysis of yield trials. Theoretical and Applied genetics *76*: 1-10.

GOODCHILD N.A. and W.J.R. BOYD (1975): Regional and temporal variation in wheat yield in western Australia and their implications in plant breeding. Aust. J. Agric. Res. *26*:209-217.

KEMPTON R.A. (1984): The use of bi-plots in interpreting variety by environment interactions. Journal of Agricultural Science *103*: 123-135.

MUNGOMERY V.E., R. SHORTER and D.E.BYTH (1974): Genotype x Environment interactions and environmental adaptation. I. Pattern analysis – applications to soya bean populations. Aust. J. Agric. Res. 25: 59-72.

INTERAKCIJA GENOTIPA I SPOLJAŠNJE SREDINE U OPLEMENJIVANJU KUKURUZA

Milosav BABIĆ, Violeta ANĐELKOVIĆ i Vojka BABIĆ

Institut za kukuruz, Slobodana Bajica 1, 11080 Zemun-Beograd

I z v o d

Kako se mere sličnosti i različitosti uvek javljaju u paru, i kako obe koriste isti tip informacije, moguće je istovremeno primeniti združene tehnike klasifikacije i ordinacije. One dopunjuju jedna drugu u analizi interakcije genotipa i spoljašnje sredine. Izbor metoda, dijagnostike i grafičkog pretstavljanja se pojavljuje kao zahtev u okviru svake od metodologija.

Grupisanjem 12 genotipova u pet grupa, 96.26% varijabilnosti sadržane u originalnim podacima je žadržano. Primenom iste analize za spoljašnje sredine, 96.45% varijabilnosti je zadržano u analizi grupisanjem 31 spoljašnje sredine u 11 grupa. Usled ovakvog grupisanje 78.10% interakcijske varijabilnosti sadržane u originalnoj matrici je zadržano u analizi ovako dvostruko redukovanih podataka. Na osnovu iznetih rezultata nije moguce jasno definisanje manjih regiona gajenja. Grupisanje spoljašnjih sredina može biti korisno ne samo za definisanje velikih ragiona gajenja, već i manjih, samo u kombinaciji primenjene metodologije sa nekim dodatnim analizama (AMMI, Diskriminaciona analiza, korespodenciona analiza itd.). U takvim analizama, iskustvo istrazivača je od velikog značaja. Na osnovu iznetih rezultata, izbor test lokacija moze biti učinjen u ograničenoj meri (pre u smislu restruktuiranja mreže ogleda u cilju dobijanja pouzdanijih informacija nego njenoj racionalizaciji u smislu smanjenjabroja lokacija).

Primljeno 24. XI 2008.
Odobreno 15. XII..2008.