Bull. Fac. Agric. Cairo Univ.,56
(2005):767- 778.

EVALUATING SOME EGYPTIAN COTTON GENOTYPES BY USING TWO STEPS OF RANDOMIZED COMPLETE BLOCK DESIGN AND CLUSTER ANALYSIS

(Received :2.5.2005)

By
H . A . Idris

Cotton Research Institute, Agriculture Research Center, Giza, Egypt.

ABSTRACT

The present investigation aimed to evaluate five long staple Egyptian cotton genotypes (*Gossypium barbadense* L.), *viz.* G.80, G.83 and G.90 (cultivars), while the others were hybrids (G.81 x G.83) and (G.89 x Pima S-6) with respect to yield and its components under some different environments. Two field experiments were carried out in two different locations during 2003 and 2004 seasons. Two steps using randomized complete block design and cluster analysis to evaluate and classify such genotypes. The various steps in the analysis considered each location as one replicate. Results indicated that no difference between the two ways of analyses (combined and two steps), with respect to locations or environments effects. On the other hand, the results showed that if the interaction between (genotypes x locations) or (environments) was significant for combined then the genotypes exhibited no significantce for the two steps. Results of cluster analysis in step (1) showed association between two hybrids stronger with respect to yield and its components during the two seasons. Cluster results in step (2) divided genotypes into two groups, *viz.* cultivars and hybrids. Hybrids were of lowest level of similarity than cultivars indicating that this group was more similar than the other group in yield and its components under different environments.

***Key words**: cluster analysis, Cotton, , environments locations,randomized complete block design,.*

1. INTRODUCTION

The future success of cotton cultivation in Egypt will depend on improving yield potential and yield stability of modern cultivars as well as production systems. Cultivars need to be widely adapted to perform consistently over environments. Researchers need a statistic that provides a measure of stability or consistency of performance across a range of environments, particularly one that reflects the contribution of each genotype to the total (genotype x environment) interaction.

The essence of randomized complete block design is that the experimental material is divided into groups, each of which constitutes a single trial or replication. At all stages of any experiment, the object is to keep the experimental errors within each group as small as is practicable. Thus, when the units are assigned to the successive groups, all units which go in the same group should be closely comparable. Similarly, during the course of the experiment, a uniform technique should be employed for all units in the same group. Any changes in the technique or in other conditions that may affect the results should be made between groups (Cochran and Cox, 1950).

Abd El- Bary (1999) found that locations and genotypes mean squares were highly significant different for boll weight, lint percentage, seed index and lint index. El Oraby (1998) and El Ameer (1999) evaluated Egyptian cotton genotypes under different environments. They reported that the mean squares of the genotypes with respect to boll weight, seed index and lint percentage differed significantly. Baker (2001) evaluated four Egyptian cotton genotypes, viz. G.80, G.83, G.85 and G.86, four Pima cotton, viz, P-S4, P-S6, P-S7 and Earlipima. He found significant variations due to environments and genotypes with respect to yield (seed and lint). Idris (2002) evaluated some Egyptian cotton cultivars under different locations. He found that both the first analysis (locations, cultivars and the interaction between them) and the second analysis (environments, cultivars and the interaction between them) mean squares were significanty different with respect to yield and its components.

A classification is a grouping based on relationship. Cluster analysis can be a very useful way to illustrate the structure present in a large data set. Any of a number of different methods of cluster analysis can be used to generate a classification. However, some grouping will result from a cluster analysis whether structure is present in the data or

not. Therefore, a classification should be evaluated to judge whether the associations indicated are real. A classification resulting from cluster analysis can be validated in terms of criteria that test the matrix and / or the dendrogram, each was the outcome of a major step in the clustering process. The matrix contains estimates of relationship between all possible pairs of objects to be categorized. These estimates are computed from raw data. Then, the estimates of relationship embodied in the matrix are used to assign the objects to the groups depicted in the dendrogram (Johnson and Wichern, 1998).

Geng and Bassey (1990) studied locations in cotton, evaluated a possible reclassification of cotton based on the characteristics of the testing locations. Idris (2002) classified ten Egyptian cotton cultivars with respect to yield and components under three different locations (governorates) Sharkia (L_1), Gharbia (L_2) and Dakhlia (L_3). Results from cluster analysis showed that associations between (G.45 and G.87) were stronger in (L_1) and (L_2). Associations between (G.85 and G.86) were stronger in (L_2) and (L_3), indicating that these cultivars were more similar in yield and its components under different environments.

The objective of the present study was to evaluate and classify some Egyptian cotton genotypes under some different environments by using two steps of randomized complete block design and cluster analysis with respect to yield and its components.

2 . MATERIALS AND METHODS

Two field experiments were carried out in two different locations during 2003 and 2004 seasons. The experiments were conducted in Bini Sweif (L_1) and Assuit (L_2) Governorates. Five long staple Egyptian cotton genotypes (G), (*Gossypium barbadense* L.) were used. Three of them are cultivars (C), *viz.* G.80, G.83 and G.90. The other two genotypes are hybrids (H), *viz.*(G.81 x G.83) and (G.89 x Pima S-6).

A randomized complete block design with four replications was used in each experiment. Each plot consists of five rows. The row was 4 meters long, 65 cm apart, 20 cm between hills and two plants per hill.

The seed cotton yield was obtained from the three inner rows while the outer rows were used for sampling of yield components (25 bolls). Planting took place during the last week of March. All cultural practices were applied. In both seasons, yield (seed S.C.Y. and lint

L.C.Y) kentar/ faddan, (kentar equals 157.5 kg for seed cotton yield and 50 kg for lint cotton), (faddan equals 4200 m^2 (m = meter)), boll weigh (B.W) gm, lint percent % (L.P.) , seed index (S.I) gm, and lint index (L.I) gm were studied.

2. 1. Analysis of randomized complete block design

The various steps used in the analysis considered each location as one replicate. First step: analysis of each season to estimate the effect of replications (locations) and genotypes. Second step: analysis of both seasons and locations to estimate the effect of replications (environmental) and genotypes. Statistical analysis of the experiment was carried out according to Gomez and Gomez (1984). Homogeneity test of variances (Bartlett test) were used according to the procedures reported by Snedecor and Cochran (1976). The treatment means were compared by L.S.D. test as given by Steel and Torrie (1961).

Table (1): Comparison Between Different Analyses Combined and Two steps.

Combined Analysis		Two Steps Analysis	
Sources of Variance	**d.f**	**Sources of Variance**	**d.f**
Locations (L)	(l-1)	Replications (R)	(r-1)
Rep. /Locations	l (r-1)		
Genotypes (G)	(g-1)	Genotypes (G)	(g-1)
(G x L)	(g-1) (l-1)	Experimental error	(g-1) (r-1)
Experimental error	l (r-1) (g-1)	Sampling (k) error	r g (k – 1)

2 . 2 . Cluster Analysis

To classify five Egyptian cotton genotypes with respect to the behavior of yield and its components under different environments , the cluster analysis was used. Cluster analysis was carried out by the hierarchical cluster analysis procedure of the program SPSS for windows.

2.2.1.Hierachical Methods: these procedures are characterized by the construction of a hierarchy or tree-like structure. In some methods each point starts out as a unit (single-point) cluster. At the next level the two closest points are placed in a cluster. At the next level, a third point joins the first two or else a second two-point cluster is formed, based on various criterion rules for assignment. In application, hierarchical clustering is useful in determining if points are substitutable rather than mutually exclusive.

2.2.2.Single Linkage: the single linkage or minimum distance rule starts out by finding the two points with the minimum distance. These are placed in the first cluster. At the next stage, a third point joins the already - formed cluster of two if the minimum distance to any of the members of the cluster is smaller than the distance between the two closest unclustered points. Otherwise, the two closest unclustered points are placed in a cluster. The process continues until all points end up in one cluster. The distance between two clusters is defined as the shortest distance from a point in the first cluster that is closest to a point in the second. Matrix gives the distances between genotypes. The distance was calculated using the euclidean distance method. (Johnson and Wichern, 1998).

3 . RESULTS AND DISCUSSION

3 . 1 . Analysis of randomized complete block design

3 . 1. 1. Comparison between combined and step (1) analyses

No combined analysis for two traits boll weight and seed cotton yield during 2003 and 2004 season, respectively due to significant bartlett test. The analysis of variance combined and step (1) revealed the presence of significant variation due to locations, genotypes, the interaction between them (combined), replications and genotypes step (1) during two seasons in (2) Table. Both combined and step (1) analysis exhibiting significant variation due to locations was observed for yield (seed and lint) and its contributing variables in the two seasons except seed cotton yield for step (1) during 2004 season. These results indicated no difference between the two ways of analyses with respect to location effects.

Results of combined analysis showed significant variation due to the genotypes for yield and its component except seed cotton yield during 2003 season. Significant variation due to the interaction between genotypes x locations was also observed for lint percentage in the first season, boll weight in the second season and lint cotton yield during the two seasons.

Results of step (1) analysis exhibited significant variation due to the genotypes was for both the two traits seed index and lint index during the two seasons and lint percentage for the 2004 season. These results indicated that the genotypes revealed non significant differences with respect to the same traits (significant for combined) due to the interaction between genotypes x locations was significant. Both

combined and step (1) analyses, genotypes observed non significant differences with respect to seed cotton yield indicated that no difference between the two ways of analyses. Similar results were obtained by Abd El Bary (1999) who found that locations and genotypes mean squares were highly significant different for lint index. El Oraby (1998) and El Ameer (1999) reported that mean squares of genotypes with respect to boll weight, seed index and lint percentage were significant. Idris (2002) found significant variation due to locations, cultivars and the interaction between them for yield and its components.

The results in (Table 3) show mean performance of yield and its components for combined and step (1). G.80 had the highest value of seed index and lint index. In both seasons, it did not significantly differ from other genotypes with respect to seed index except G.81 x G.83 for step (1) while it significantly surpassed all genotypes except G.90 for combined. On the other hand, with respect to lint index it significantly exceeded the two hybrids for step (1) and all genotypes for combined except G.83 in 2003 season. In the second season, as for lint percentage G.80 did not significantly differ from other genotypes except G.89 x Pima S-6 for step (1) while it significantly surpassed all genotypes except G.83 for combined. These results indicated that the differences between the two ways of analyses are due to the differences of LSD values.

3. 1. 2. Comparison between Combined and Step (2) Analyses

The analysis of variance combined and step (2) showed the presence of significant variation due to environments, genotypes, the interaction between them (combined), replications and genotypes step (2) Table (4). Both combined and step (2) analyses exhibited that no difference between the two ways of analyses with respect to environmental effects. Significant variation due to environments was observed for yield (seed and lint) and its contributing variables. These results indicated that no difference was observed between the two ways of analyses with respect to environments effects.

Results of combined analysis revealed significant variation due to genotypes and the interaction between genotypes x environments were detected for all traits except lint percentage, seed index and lint index with respect to the interaction.

Table (2) : Mean Squares of Combined and Step (1) Analyses for Yield and its Components .

2003 Season							
Combined							
Traits / Sources of Variance	d.f	S.C.Y (K/F)	L.C.Y. (K/F)	B.W. (gm)	L. P. (%)	S. I. (gm)	L.I. (gm)
Locations (L)	1	158.34**	160.04**	----	25.06**	19.27**	2.25**
Rep. /Locations	6	3.89	4.98	----	0.223	0.124	0.095
Genotypes (G)	4	1.02	2.99*	----	3.21**	1.17**	1.06**
(G x L)	4	1.28	2.23*	----	1.45*	0.165	0.025
Experimental error	24	0.543	0.748	----	0.393	0.211	0.070
Step (1)							
Traits / Sources of Variance	d.f	S.C.Y (K/F)	L.C.Y. (K/F)	B.W. (gm)	L. P. (%)	S. I. (gm)	L.I. (gm)
Replications (R)	1	158.34**	160.04**	0.989*	25.06*	19.27**	2.25**
Genotypes (G)	4	1.02	2.99	0.128	3.21	1.17*	1.06**
Experimental error	4	1.28	2.23	0.089	1.45	0.165	0 025
Sampling (K) Error	30	1.21	1.60	0.016	0.358	0.194	0.075
2004 Season							
Combined							
Traits / Sources of Variance	d.f	S.C.Y (k / f)	L.C.Y. (k / f)	B.W. (gm)	L. P. (%)	S. I. (gm)	L.I. (gm)
Locations (L)	1	----	47.57**	0.710**	15.40**	8.81**	8.76**
Rep. /Locations	6	----	2.89	0.011	0.668	0.192	0.032
Genotypes (G)	4	----	6.00**	0.089**	9 15**	1.51**	1.92**
(G x L)	4	----	7.16**	0.081**	0.486	0.089	0.069
Experimental error	24	----	1.32	0 014	0.842	0.226	0.104
Step (1)							
Traits / Sources of Variance	d.f	S.C.Y (k / f)	L.C.Y. (k / f)	B.W. (gm)	L. P. (%)	S. I. (gm)	L.I. (gm)
Replications (R)	1	19.92	47.57**	0.710*	15.40**	8.81**	8.76**
Genotypes (G)	4	2.99	6.00	0.089	9.15**	1.51**	1.92**
Experimental error	4	8.49	7.16	0.081	0.485	0.089	0.069
Sampling (K) Error	30	1.04	1.63	0.013	0.807	0.219	0.090

----, not combined analysis due to bartlett test was significant. *,** Significant at the 0.05 and 0.01 levels , respectively.

Table (3) : Mean Performance of Yield and its Components for Combined and Step (1) Analyses .

2003 Season						
Traits / Genotypes	S.C.Y (k / f)	L.C.Y. (k / f)	B.W. (gm)	L. P. (%)	S. I. (gm)	L.I. (gm)
G. 80	11.36	13.28	2.73	39.03	10.28	6.45
G. 83	11.25	13.65	2.73	39.22	9.80	6.30
G.90	11.92	14 14	2.70	38.31	9.88	6.13
G.81 x G.83	10.97	12.70	2.57	38.37	9.13	5.67
G.89 x Pima S-6	11.16	12.75	2.44	37.63	9.42	5.66
L.S.D. combined at 5%	----	0.89	N	0.65	0.47	0.27
L.S.D. Step (1) at 5 %	----	----	----	----	1.13	0.44
2004 Season						
Traits / Genotypes	S.C.Y (k / f)	L.C.Y. (k / f)	B.W. (gm)	L. P. (%)	S. I. (gm)	L.I. (gm)
G. 80	13 26	16.04	2.78	39 82	11.21	7.41
G. 83	11.80	14.60	2.77	39.25	10.53	6.81
G.90	12.40	14.74	2.68	38.39	10 93	6.82
G.81 x G.83	11.76	13.83	2.52	38.09	10.06	6.21
G.89 x Pima S-6	12.13	14.03	2.65	37.05	10.66	6.28
L.S.D. combined at 5%	N	1.18	0.12	0 95	0.49	0.33
L.S.D. Step (1) at 5 %	----	----	----	1.94	0.83	0.73

N, not combined analysis due to bartlett test was significant. ----, Not significant at 5 %.

Table (4) : Mean Squares of Combined and Step (2) Analyses for Yield and its Components.

Traits		Combined S.C.Y	L.C.Y.	B.W.	L. P.	S. I.	L.I.
Sources of Variance	d.f	(k / f)	(k / f)	(gm)	(%)	(gm)	(gm)
Environments (E)	3	65.22**	81.26**	0.581**	13.49**	16.20**	6.60**
Rep. /Locations	12	2.62	3.94	0.009	0.446	0.158	0.063
Genotypes (G)	4	2.69*	6.21**	0.178**	11.32**	2.53**	2.84**
(G x E)	12	3.70**	4.06**	0.070**	0.992	0.136	0.076
Experimental error	48	0.749	1.03	0.016	0.618	0.218	0.087
		Step (2)					
Traits		S.C.Y	L.C.Y.	B.W.	L. P.	S. I.	L.I.
Sources of Variance	d.f	(k / f)	(k / f)	(gm)	(%)	(gm)	(gm)
Replications (R)	3	65.22**	81.26**	0.581**	13.49**	16.20**	6.60**
Genotypes (G)	4	2.69	6.21	0.178	11.32**	2.53**	2.84**
Experimental error	12	3.70	4.06	0.070	0.992	0.136	0.076
Sampling (K) Error	60	1.69	2.42	0.022	0.875	0.310	0.124

*,** Significant at the 0.05 and 0.01 levels , respectively.

Results of step (2) analysis exhibited significant variation due to the genotypes for three yield components, lint percentage, seed and lint index. These results indicate that if the interaction between genotypes x environments was significant for combined then the genotypes exhibited no significant effect for step (2). Baker (2001) and Idris (2002) found that significant variations due to environments, cultivars and the interaction between them were observed with respect to yield (seed and lint).

The results in (Table 5) show the mean performance of yield and its components for combined and step (2). G.80 had the highest value for lint percent, seed index and lint index. It did not significantly differ from other genotypes with respect to lint percent except 89 x Pima S-6 for step (2), while it significantly surpassed all genotypes excepted, G.83 for combined. As for seed index, it significantly exceed the two hybrids for step (2) and all genotypes for combined except G.90.

On the other hand, it significantly surpassed all genotypes with respect to lint index in the two way of analysis except G.83 for step (2) indicated that similar results were obtained by using the two way of analyses.

Table (5) : Mean Performance of Yield and its Components for Combined and Step (2) Analyses.

Traits	S.C.Y	L.C.Y	B.W.	L. P.	S. I.	L.I.
Genotypes	(k / f)	(k / f)	(gm)	(%)	(gm)	(gm)
G. 80	12.31	14.66	2.75	39.42	10.65	6.93
G. 83	11.54	14.13	2.75	39.23	10.16	6.56
G.90	12.16	14.44	2.69	38.35	10.40	6.47
G.81 x G.83	11.36	13.27	2.54	38.23	9.59	5.94
G.89 x Pima S-6	11.65	13.39	2.55	37.34	10.04	5.97
L.S.D. combined at 5%	0.62	0.72	0.09	0.56	0.33	0.21
L.S.D. Step (2) at 5 %	----	----	---	1.53	0.57	0.42

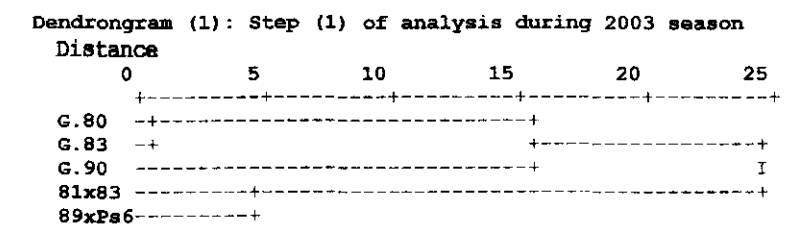
---- : Not significant at 5 %.

3 . 2. Cluster Analysis

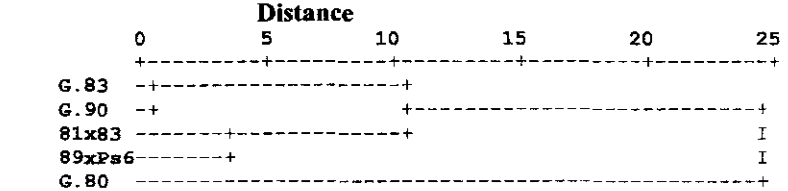
Results in (Table 6) show euclidean distance among five genotypes evaluated under different locations and environments with for the two steps, respectively. Dendrograms (1 and 2) illustrate thes results cluster in step (1) of analysis. Association among the two hybrids were stronger with respect to yield and its components during the two seasons, due to the euclidean distance between them was lowest. This result showed that the two hybrids were more similar in yield behavior during the two seasons. Results of step (1) analysis exhibited G.90 and G.80 differed than other genotypes with respect to yield and its components in the first and second seasons, respectively. Dendrogram (3) illustrates the results on cluster in step two of analysis. Genotypes are divided into two groups cultivars and hybrids. Hybrids were of the lowest level of similarity indicating that this group was more similar than the other group in yield and its components under different environments. Idris (2002) found that associations between (G.85 and G.86) was stronger in Gharbia (L_2) and Dakhlia (L_3) indicated that these cultivars were more similar in yield and its components under different environments.

Table (6) : Euclidean Distance Among Genotypes with respect to Yield and its Components for the two Steps analysis.

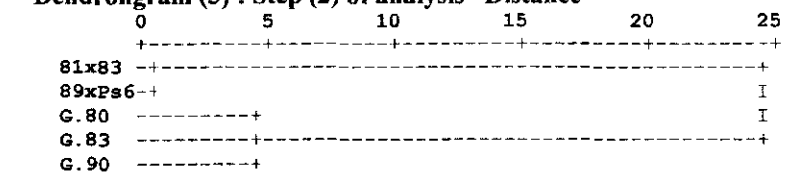
Step (1)				
2003 Season				
Genotypes	**G. 80**	**G.90**	**G.90**	**G.81 x G.83**
G. 83	.662			
G.90	1.36	1.25		
G.81 x G.83	1.70	1.61	1.94	
G.89 x Pima S-6	1.93	2.00	1.86	829
2004 Season				
Genotypes	**G. 80**	**G.90**	**G.90**	**G.81 x G.83**
G. 83	2.31			
G.90	2.22	1.14		
G.81 x G.83	3.60	1.61	1.58	
G.89 x Pima S-6	3.82	2.36	1.66	1.28
Step (2)				
Genotypes	**G. 80**	**G.90**	**G.90**	**G.81 x G.83**
G. 83	1.13			
G.90	1.22	1.15		
G.81 x G.83	2.53	1.59	1.73	
G.89 x Pima S-6	2.78	2.13	1.67	1.05



Dendrongram (2): Step (1) of analysis during 2004 season



Dendrongram (3) : Step (2) of analysis



4. REFERENCES

Abd-El Bary, A. M. (1999). Inheritance of quantitive traits of Egyptian cotton (*Gossypium barbadense L.*). M. Sc. Thesis, Mansoura Univ., Egypt.

Baker K. M. (2001). Comparative performance of Pima and Egyptian cotton under local environmental conditions. M.Sc. Thesis, Fac., of Agric., Cairo Univ., Egypt.

Cochran W.G. and Cox G. (1950). Experimental Designs. John Wiley and Sons, New York., U.S.A.

El Ameer, M. A. (1999). Genetical studies on types of gene effects on some crosses of Egyptian cotton (*Gossypium barbadense L.*). M. Sc. Thesis, Mansoura Univ., Egypt.

El Oraby, S.H. (1998). Studied on yield of some Egyptian cotton varieties under different environments. M.Sc. Thesis, Kafr El Sheikh, Tanta Univ., Egypt.

Geng S. H. and Bassett D. M. (1990). Quantification and classification of locational effects on cotton cultivar testing programs. Agron. J. 82 : 514 –518.

Gomez K. A. and Gomez A. A. (1984). Statistical Procedures for Agricultural Research. Wiley, New York, U.S.A..

Idris, H. A. (2002). Models for evaluating Egyptian cotton responses to environmental variables. Ph.D. Thesis, Fac., of Agric., Cairo Univ., Egypt

Johnson R. A, and Wichern D. W. (1998). Applied Multivariate Statistical Analysis. Prentice Hall, N. J.

Snedecor G. W. and Cochran W. G. (1976). Statistical Methods. Iowa State Univ. Press, Ames.

SPSS for Windows Release 8.0 (1997). Statistical Package Software.

Steel R. G. and Torrie J. H. (1961). Principles and Procedures of Statistics. McGraw. Hill. Book Co., Inc.

تقييم بعض التراكيب الوراثية من القطن المصري باستخدام خطوتين من تحليل القطاعات كاملة العشوائية والتحليل العنقودي

حاتم أحمد إدريس

معهد بحوث القطن – مركز البحوث الزراعية – الجيزة

ملخص

تم تقييم المحصول ومكوناته لخمس تراكيب وراثية من القطن المصري تتبع طبقة الأقطان طويلة التيلة ، ثلاثة منها أصناف تجارية هى جيزة ٨٠ ، جيزة ٣ ، جيزة ٩٠ واثنان هجن مبشرة (جيزة ٨١ x جيزة ٨٣) ، (جيزة ٨٩ x بيما إس٦) فى بنى سويف و أسيوط خلال موسمي ٢٠٠٣ ، ٢٠٠٤.

تم استخدام خطوتين من تحليل القطاعات كاملة العشوائية تعتمد على أن يمثل كل موقع بمكرر واحد كما يلى : الخطوة الأولى : تحليل كل موسم على حده تجميعا للموقعين لدراسة تأثير التفاعل بين المواقع والتراكيب الوراثية. الخطوة الثانية : تحليل الموقعين و السنتين تجميعا لدراسة تأثير التفاعل بين البيئات (المواقع والمواسم) والتراكيب الوراثية.

وقد أظهرت النتائج مايلى :

أولا تحليل القطاعات كاملة العشوائية

لا يوجد فرق بين تحليل القطاعات كاملة العشوائية و التحليل التجميعي بالنسبة للتأثير المستقل لكل من المواقع والبيئات. كما أوضحت النتائج أنه اذا كان بين التراكيب الوراثية وكل من المواقع والبيئات فرق معنوي فى التحليل التجميعي أظهرت التراكيب الوراثية فروق غير معنوية فى تحليل القطاعات الكاملة العشوائية.

ثانيا التحليل العنقودي

أظهرت نتائج الخطوة الأولى من التحليل أن الهجن كانت متماثلة فى الموسمين.

قسمت التراكيب الوراثية الى مجموعتين المجموعة الأولى وتشمل الأصناف،و الثانية (الهجن) فى الخطوة الثانية من التحليل.

المجلة العلميـــة لكلية الزراعة – جامعة القاهرة – المجلد (٥٦) العدد الرابع (أكتوبر ٢٠٠٥): ٧٦٧-٧٧٨.