Electronic Journal of Plant Breeding

Research Note

Genetic diversity analysis in American cotton (*Gossypium hirsutum* L.) genotypes

Y. Satish

Regional Agricultural Research Station, Acharya N.G. Ranga Agricultural University, Lam, Guntur-522034
E-Mail: tej.satish1@gmail.com

Abstract
Genetic divergence studies were taken up using 55 genotypes for yield and its contributing traits using D^2 statistics in upland cotton. Based on D^2 values, the grouping was done and the 55 genotypes were placed in six groups and six principal components explaining 91.8 per cent variability. Results of the multivariate analysis indicated that maximum divergence among genotypes L1785, L1687, L1748, L1780, L1783, and L 604 was detected indicating that the crosses between these genotypes will produce highly heterotic hybrids.

Key words: Genetic divergence, D^2 statistic, Principal Component Analysis, American Cotton

Cotton is an important fibre crop having high commercial and export value with 70% of fibre consumption in the textile sector. Regarded as the king of fibre crops, it has a large contribution to the Indian economy. Cotton is a predominant fibre in the Indian textile industry. For the development of high heterotic hybrids with good yield potential, the genetic diversity among parents is of utmost importance. In this study, the genetic divergence among 55 genotypes was estimated using two methods of clustering based on principal component analysis and D^2 analysis. Mahalanobis' D^2 statistics, is an ideal tool to estimate the degree of divergence and also provides a degree of association between genetic diversity and geographic distribution (Mahalanobis, 1928) and Principal Component Analysis changes the inter dependent characters into a set of independent characters and reduces the dimensionality of the data structure (Banfield, 1978).

The experiment was taken up using 55 genotypes of cotton (*Gossypium hirsutum* L.) which were studied using randomized complete block design with two replications at RARS, Lam, Guntur in *kharif, 2019*-20. The genotypes were raised in 6 m length with row to row and plant to plant spacing was 105 cm x 60 cm. Data were recorded on five selected plants from each genotype in each replication for different yield and its contributing traits *viz*., plant height (cm), monopodia/plant, sympodia/plant, bolls/ plant, and seed cotton yield (kg/ha), while, data on plot basis was recorded for days to 50 per cent flowering, boll weight (g), seed index (g), lint index (g) and ginning out-turn (%). Genetic diversity was estimated by using Mahalanobis' D^2 statistics as per Rao (1952) and principal component analysis (PCA) as described by Jackson (1991).

Based on the results obtained by D^2 statistics, the 55 genotypes were grouped into six clusters Cluster I is the largest which is comprised of 18 genotypes, followed by clusters II, III, and IV which had 15, 14, and six genotypes, respectively. While in clusters V and VI only one genotype was found (**Table 1 and Fig.1**).

Maximum contribution for genetic divergence was by the seed cotton yield (kg/ha) (34.68 %) followed by lint index (23.37 %), boll weight (17.98 %), bolls/plant
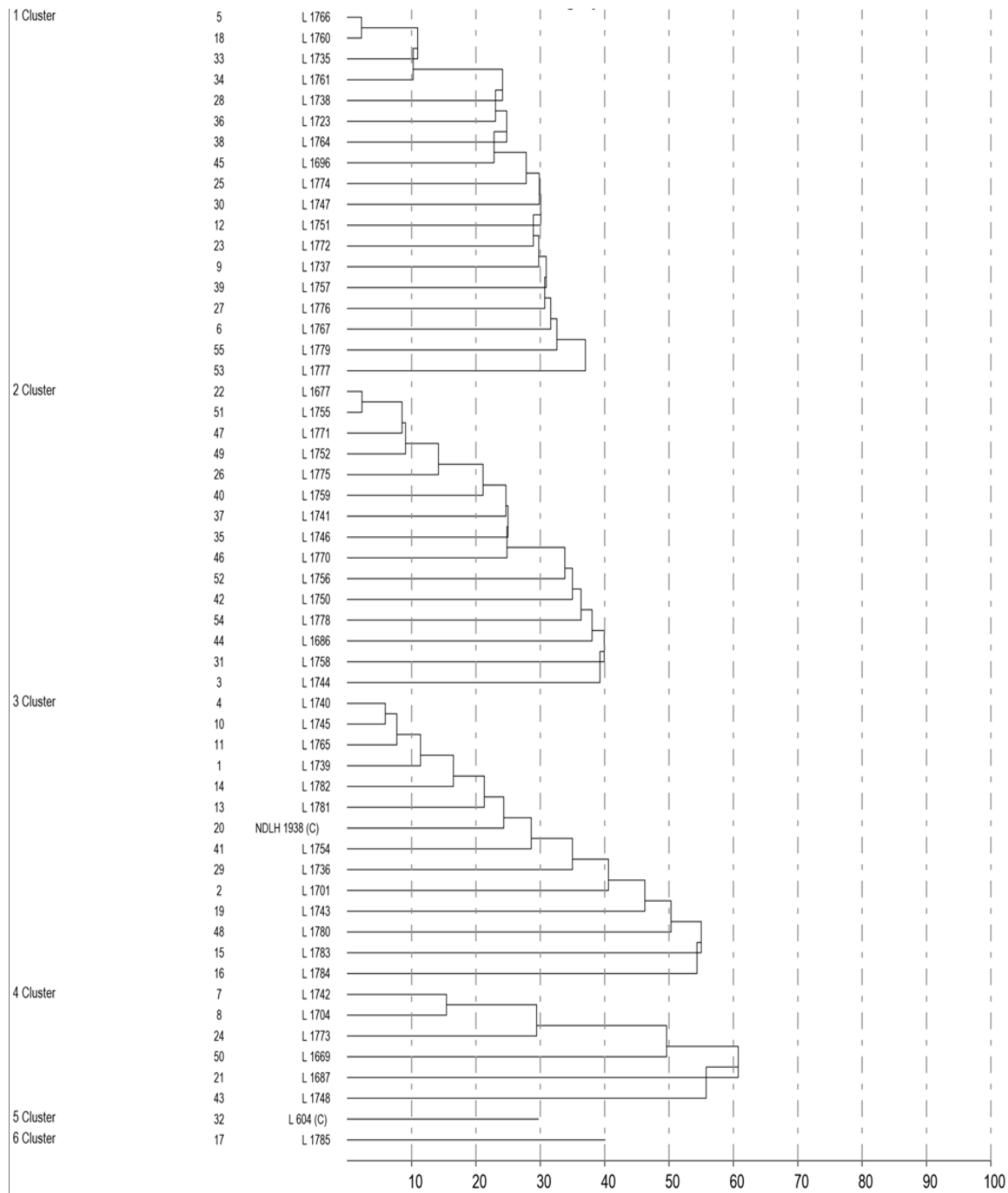
Fig. 1. Dendrogram showing relationship among 55 cotton genotypes in six clusters based on Mahalanobis' D² values

(5.52 %), ginning out-turn (5.25 %), plant height (3.64 %), seed index (3.5%), days to 50 per cent flowering (3.03 %), monopodia/plant (2.36 %) and sympodia/plant (0.67 %) (**Table 2 and Fig. 2**).

The intra cluster distance was maximum for cluster IV (62.70) followed by cluster III (43.88), cluster II (34.85), and cluster I (32.18), while, it was zero for cluster V and VI as they had single genotype (**Table 3**). In cluster IV, the

Table 1. Clustering pattern of 55 cotton genotypes by Tocher's method

Cluster number	Number of genotypes	Name of the genotype
I	18	L1766, L1760, L1735, L1761, L1738, L1723, L1764, L1696, L1774, L1747, L 1751, L1772, L1737, L1757, L1776, L1767, L1779 and L 1777
II	15	L1677, L1755, L1771, L1752, L1775, L1759, L1741, L1746, L1770, L1756, L1750, L1778, L1686, L1758 and L1744
III	14	L1740, L1745, L1765, L1739, L1782, L1781, NDLH1938, L1754, L1736, L1701, L1743, L1780, L1783 and L1784
IV	6	L1742, L1704, L1773, L1669, L1687 and L1748
V	1	L604
VI	1	L1785

Table 2. Contribution of different characters towards genetic divergence in 55 cotton genotypes

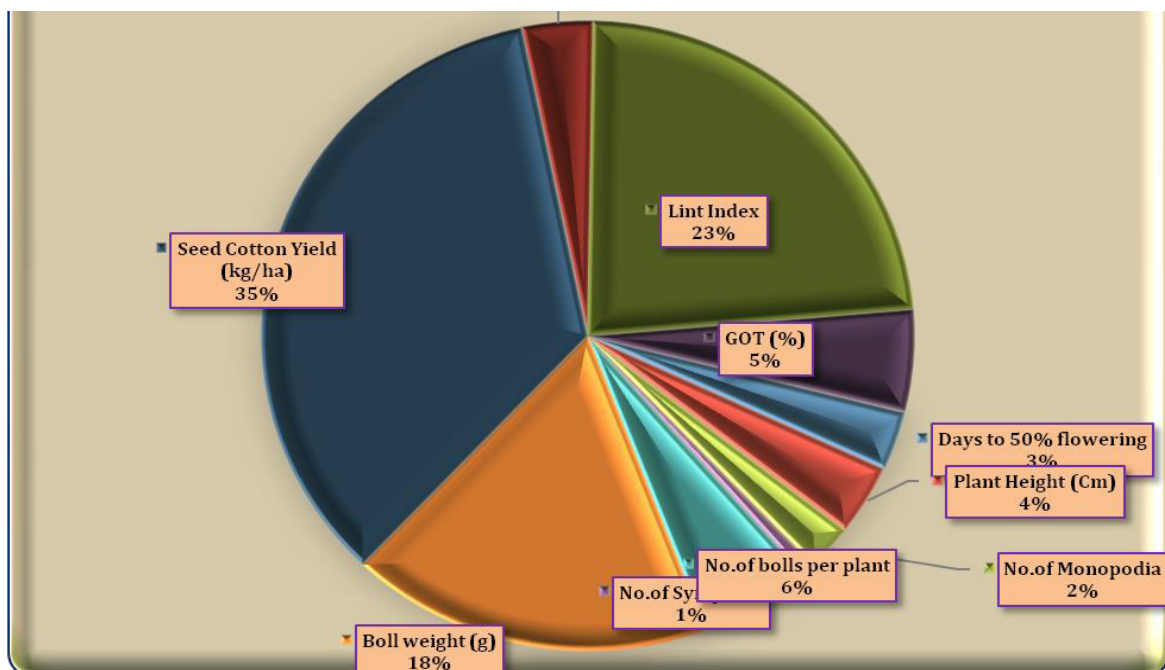| S. No.	Character	Contribution towards divergence (%)	Times ranked first
1	Days to 50 percent flowering	3.03	45
2	Plant height	3.64	54
3	Monopodia/plant	2.36	35
4	Sympodia plant	0.67	10
5	Bolls plant	5.52	82
6	Boll weight	17.98	267
7	Seed cotton yield	34.68	515
8	Seed index	3.5	52
9	Lint index	23.37	347
10	GOT	5.25	78

Fig. 2. Per cent contribution of different characters towards genetic divergence in 55 cotton genotypes

intra-cluster distance was maximum which indicated the presence of wide genetic diversity among the genotypes present within this cluster. The inter-cluster distance was maximum between cluster IV and VI (274.06) followed by cluster I and VI (202.66), cluster IV and V (137.42), cluster III and VI (136.05), cluster III and IV(119.38), cluster II and IV (110.64), cluster II and III (96.88) and cluster II and VI (93.22). The cluster mean values are presented in **Table 4**. This indicated that genetic diversity is wide between these clusters. Crosses made between the genotypes of distant clusters will be highly heterotic.

It is well known that with the increase in the genetic divergence, the heterosis will be increased. Therefore, to get highly heterotic hybrids, it is advisable to make crosses between the genotypes of distant clusters. These heterotic hybrids will give a wide range of segregants on which selection can be practiced.

Based on intra-and inter-cluster distances, it is desirable to make crosses between the genotypes of cluster IV (L 1742, L1704, L1773, L1669, L1687, and L1748) and cluster VI (L1785), between genotypes of cluster I (L1766, L1760, L1735, L1761, L1738, L1723, L1764, L1696, L1774, L1747, L 1751, L1772, L1737, L1757, L1776, L1767, L1779 and L 1777) and cluster VI (L1785), between the genotypes of cluster IV (L 1742, L1704, L1773, L1669, L1687 and L1748) and cluster V (L 604) and between the genotypes of cluster III (L1740, L1745, L1765, L1739, L1782, L1781, NDLH1938, L1754, L1736,

L1701, L1743, L1780, L1783 and L1784) and cluster VI (L1785) after confirmation of their combining ability.

The principal component analysis identified six principal components (PCs) and the contribution of these six PCs was 91.8 per cent of cumulative variance. The first principal component (PC_1) contributed maximum towards variability (29.49) followed by PC_2 (20.39), PC_3 (13.75), PC_4 (12.01), PC_5 (9.91), and PC_6 (6.21) (**Table 5**). The traits *viz.,* lint index, boll weight, seed cotton yield, seed index, bolls/plant, ginning outturn, monopodia/plant, and days to 50 per cent flowering contributed towards maximum divergence in PC1. The 2D graph (**Fig. 3**) showed more divergence between L1785, L1687, L1748, L1780, L1783, and L 604 indicating their use in future cotton breeding to develop highly heterotic hybrids.

Identification of parents for hybrid breeding programme is crucial in the exploitation of heterosis. For obtaining highly heterotic hybrids the question generally arises is regarding the ideal distance at the phenotypic level. Both methods of grouping revealed the single concept of non-correspondence of genetic divergence and geographical diversity in this study. Therefore, genetic diversity plays a more important role in the selection of genotypes than geographical diversity. Similar results were also reported by Pradeep *et al*. (2018), Anjani *et al*. (2018), Naik *et al*. (2016), Tulasi *et al.* (2014), and Kavithmani *et al.* (2013).

Table 3. Average intra-and inter-cluster D² values among six clusters in 55 cotton genotypes

Cluster number	I	II	III	IV	V	VI
I	**32.18**	75.85	79.87	54.28	72.10	202.66
II		**34.85**	96.88	110.64	89.16	93.22
III			**43.88**	119.38	67.09	136.05
IV				**62.70**	137.42	274.06
V					**0.00**	128.61
VI						**0.00**

Table 4. Mean values of six clusters estimated by Tocher's method from 55 genotypes of Cotton

|	Days to 50 percent flowering	Plant height (cm)	Monopodia /plant	Sympodia / plant	Bolls / plant	Boll weight (g)	Seed cotton yield (kg/ha)	Seed index	Lint index	Ginning outturn (%)
Cluster 1	52.22	135.58	1.01	20.15	55.64	4.10	1723.55	9.89	7.65	33.84
Cluster 2	51.80	141.07	1.09	21.96	55.10	3.52	1346.30	8.21	5.47	31.27
Cluster 3	58.61	152.71	1.08	21.30	57.71	3.88	1748.02	9.15	6.62	32.58
Cluster 4	50.67	158.50	1.18	21.55	58.25	4.41	2050.27	10.84	8.09	33.16
Cluster 5	57.50	97.50	1.00	14.95	37.00	3.85	1460.32	8.65	5.85	31.64
Cluster 6	57.00	156.50	0.60	20.75	51.00	2.50	1111.11	7.55	4.84	30.72
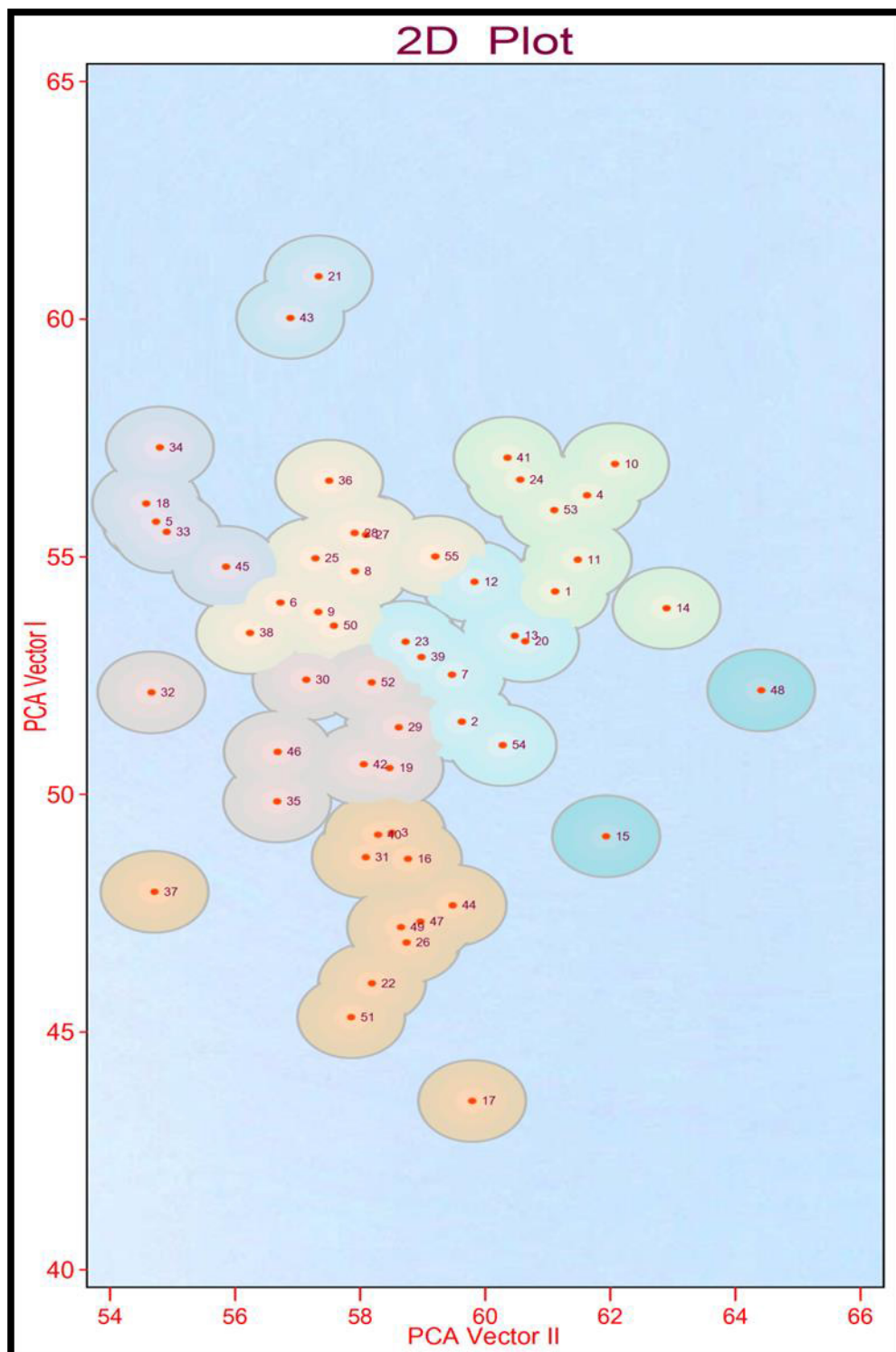
Fig. 3. Two dimensional graph showing relative position of 55 cotton genotypes based on PCA scores

Table 5. Eigen values, proportion of the total variance represented by first six principal components, cumulative per cent variance and component loading of different characters in cotton

	PC$_1$	PC$_2$	PC$_3$	PC$_4$	PC$_5$	PC$_6$
Eigen Value (Root)	2.9498	2.0391	1.3757	1.2017	0.9914	0.6220
% Var. Exp.	29.4984	20.3913	13.7572	12.0171	9.9141	6.2199
Cum. Var. Exp.	29.4984	49.8897	63.6469	75.6639	85.5780	91.7979
Days to 50 per cent flowering	0.0413	0.2704	0.6522	0.2445	0.2651	0.2693
Plant height	-0.0202	0.5477	-0.2536	0.2583	-0.2566	0.1258
Monopodia/plant	0.0635	-0.1854	0.2997	-0.6250	-0.5093	-0.0185
Sympodia/plant	-0.1475	0.4600	-0.2857	-0.3040	-0.2198	0.4987
Bolls/plant	0.1738	0.4683	0.4360	-0.2115	-0.1220	-0.2784
Boll weight	0.4976	-0.0703	-0.3048	0.0512	-0.1538	-0.1273
Seed cotton yield	0.4834	0.2517	-0.0662	0.1272	-0.1126	-0.4177
Seed index	0.4215	-0.2607	0.1045	0.0883	-0.1843	0.5622
Lint index	0.5114	-0.0039	0.0002	-0.0292	0.2331	0.2798
GOT	0.1463	0.1596	-0.2015	-0.5642	0.6455	-0.0018

The genotypes L1687, L1748, L1780, L1783, and L 604 showed maximum inter-cluster distance in Mahalanobis'D^2 analysis, principal component analysis and also had better *per se* performance for the traits *viz*., sympodia/plant, bolls/plant, boll weight, seed index, lint index, and seed cotton yield. Hence, these genotypes can be used in breeding programmes to generate high heterotic hybrids.

REFERENCES

Anjani, A., Padma, V., Ramana, J. V. and Satish, Y. 2018. Evaluation of genetic diversity in American cotton (*Gossypium hirsutum* L.). *Int. J. Cur. Microbiol. Appl.Sci.***7** : 3905-3911

Banfield, C.F. 1978. Principal component analysis for genstat. *J. Stat. and Computer Simulation*. **6**: 211-222. [Cross Ref]

Jackson, J. E. 1991 A User's Guide to Principal Components. John Wiley and Sons Inc., New York. [Cross Ref]

Kavithamani, D., Balu, P.A and Rajarathinam. 2013 Assessment of genetic diversity in *Gossypium hirsutum* L. *J. Cot. Res. and Devl,.***27** : 20-25

Mahalanobis, P.C. 1928. A statistical study at Chinese head measurement. *J. Asiatic Soc. Bengal,***25**: 301-307.

Mansingh Naik, B., Satish, Y. and Ratna Babu, D. 2016. Genetic diversity analysis in American Cotton (*Gossypium hirsutum* L.). *Electronic Journal of Plant Breeding*, **7**(4): 1002-1006. [Cross Ref]

Pradeep, S., Satish, Y. and Ratna Babu, D. 2018. Estimation of genetic diversity in upland cotton (*Gossypium hirsutum* L.) genotypes. *Electronic Journal of Plant Breeding,* **9** : 90 - 96. [Cross Ref]

Rao, C.R. 1952. Advanced Statistical Methods in Biometric Research. John Wiley and Sons Inc., New York. PP.236-272.

Tulasi, J., Ahaamed, M.L., Murthy, J.S.V.S and Rani,Y.A. 2014. Multivariate analysis in upland cotton (*Gossypium hirsutum* L.). *J. Cot. Res. and Devl,.***28** : 191-194.