Electronic Journal of Plant Breeding, 8(1): 169-175 (March 2017)
ISSN 0975-928X
DOI: 10.5958/0975-928X.2017.00024.2

Research Article

Assessment of Genetic diversity in upland cotton (*Gossypium hirsutum* L.) germplasm accessions to improve their yield and fibre quality traits

D. Kavithamani and P. Amalabalu
Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore-641 003
E-mail: kavitharice1@gmail.com

(Received: 11 Nov 2015; Revised: 17 March 2017; Accepted: 22 March 2017)

Abstract
The improvement of any crop mainly depends upon the nature and magnitude of genetic variability present in the base population. The objective of this study was to assess the genetic diversity and relationship among the *G. hirsutum* accessions using multivariate Mahalanobis D^2 statistics. Fifty four *G. hirsutum* accessions of diverse origin were utilized in this study. In D^2 analysis, the 54 accessions were grouped into into 12 clusters. Grouping of accessions into different clusters was independent of the geographical origin of accessions and it also revealed that single plant yield contributed the maximum divergence of 63.5per cent. The highest cluster mean for number of bolls per plant was recorded by the cluster XI, for boll weight by cluster X, for single plant yield by cluster XII and for Fibre length cluster VI. The crosses between the genotypes of these clusters *viz.*, SVPR 4 (XI), TCH 1732 (X), Okra Narrow (XII), and TCH 1710 (VI) were expected to express high heterotic vigour. The molecular diversity analysis was also carried by using SSR markers. The markers NAU 4900, NAU 2591 were identified as very effective with polymorphism. The outcome of this study will be helpful in selection of parents with wide genetic distances which could be used as parents in hybridization programme and to develop mapping population to identify Quantitative Trait Loci linked to traits of agronomic importance in cotton.

Key words
Genetic divergence, *G. hirsutum*, inter cluster, intra cluster, multivariate analysis.

Introduction

Cotton (*Gossypium spp*.), king of fibres is the world's leading natural fibre crop and it is the corner stone of textile industries worldwide. It is providing directly and indirectly livelihood to more than 60 million people and accounting for about 30 per cent of India's export earnings. It is the number one commercial crop of India. The economy of many countries depends on production on utilization and export of cotton, mainly in apparel manufacturing. Gossypium is a large, diverse and economically viable genus, which include many diploid and tetraploid species. The extensive genetic variation is present in this genus is distributed among 50 species, of which four are cultivated, 44 are wild diploids and two are wild tetraploids (Percival and Kohel, 1990). India is the only country growing all the four species of cultivated cotton *Gossypium arboreum* and *G. herbaceum* (Asian cotton), *G. barbadense* (Egyptian cotton) and *G. hirsutum* (American upland cotton) besides hybrid cotton. Among the two cultivated tetraploid species, upland cotton((*Gossypium hirsutum)* in considered the most important one for its wide adaptability high yielding and better spinning ability as democrated by the release of number of stable varieties. Yield and fibre quality are the two important criteria which decides the sustainability of the variety or hybrid for commercial cultivation. Genetic diversity enables for long term sustainability and agricultural self reliance.

The presence of genetic diversity is important for improving any crop species. Genetic diversity is the basic portion of biological diversity and in the base of biological polymorphism and genus diversity. Often plant breeders limit their effort to a narrow range of adapted lines for genetic improvement, which may likely to produce economic gains in short term but may have enhanced genetic vulnerability to biotic insects and other abiotic stress. The precious evaluation of the genetic diversity of the excellent germplasm will provide a guide for choosing parents and predicting the degree of inheritance, variation and level of heterosis, which are essential for releasing the breeding goal. Usually breeders have been employing morphological markers for genetic diversity estimation and a number of morphological descriptors in various crops are in vogue for characterization purpose. To overcome the limitations associated with the morphological markers, a large number of molecular markers have come up in recent past. These molecular markers can increase the speed and precision of breeding. Molecular marker analysis is a modern technique, which discloses genetic difference at the DNA level in plants and is an effective tool for testing genetic diversity of germplasm in breeding programme. Modern and molecular technique can illumarate the individual differences and relationship many species at the DNA level.

Materials and methods

The study was conducted in the department of cotton, Tamil Nadu Agricultural University, Coimbatore. Fifty four *Gossypium hirsutum* cotton genotypes were planted in randomized block design with two replications. Uniform spacing of

DOI: 10.5958/0975-928X.2017.00024.2

90 x 45 cm and all the recommended field operations were carried out. In each replication five competitive plants were randomly selected and observations were recorded for 16 characters *viz*., days to 50 % flowering, plant height (cm), internode length (cm), number of sympodia / plant, number of ovules/ plant, number of bolls/ plant, boll weight(g), number of seeds/ plant, seed setting percentage, seed cotton yield/ plant, lint index, seed index, ginning outturn(%), 2.5% span length (mm), bundle strength (g/tex) and fibre fineness. The genetic divergence was worked out by using Mahalanobis D^2 statistic as described by Rao (1952). Based on the D^2 value, evaluated genotypes were grouped into different clusters by employing Tocher's method as outlined by Rao (1952).

The molecular diversity experiment was conducted in the marker assisted selection laboratory, Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore. Twenty four *G. hirsutum* accessions were involved in this SSR analysis (Table 1). These accessions were selected from the study material based on their diversified geographical origin and morphological characters. The Ntsyspc software package (Perrier and Jacquemoud-Collet, 2006) was used for Statistical analysis and construction of dendrogram using molecular data. Polymorphic Information Content (PIC) values were calculated for SSR markers in order to characterize the capacity of each primer to reveal or detect polymorphic loci among the genotypes. It is the sum total of polymorphism information content values of all the markers produced by a particular primer.

Results and discussion

The analysis of variance showed highly significant differences among genotypes for all the characters studied and infers existence of considerable genetic diversity among genotypes. Hence, further analysis was carried out for relative magnitude of D^2 values for all the characters and all the genotypes were grouped into 12 clusters, which are given in table 2 and depicted as dendrogram in fig. 1. Among the 12 clusters, the maximum number of accessions was grouped in cluster III (15 accessions), followed by cluster I (12), cluster VI (9), cluster II (7) and cluster IV (4) accessions and all other clusters were having only one accession. This result indicated that there is no parallel relationship between genetic diversity and geographical diversity which was supported by Kowsalya and Raveendran (1996). As Sambamurthy *et al.* (1995) said the *G. hirsutum* accessions taken for analysis from same origin (India) grouped into different clusters and some germplasm accessions from different origin formed single cluster. From this clustering pattern we can infer that in both *Gossypium* species clustering of genotypes is independent of eco geographical origin. So selection of parents for hybridization should not be done only based on geographic diversity while priority should be given for genetic diversity in relation to the characters. Singh and Singh (1984) and Singh and Gill (1984) revealed that the clustering pattern of genotypes from different sources clustered together indicate that there was no association between eco geographical distribution of genotypes and genetic divergence. The possible reason for grouping of genotypes of different states in one cluster could be the free exchange of germplasm among the breeders of different regions, or unidirectional selection practiced by breeder in tailoring the promising cultivars for different regions.

The inter and intra cluster distances of the obtained clusters are given in the table 3. The maximum inter cluster distance was higher than intra cluster distances except in some clusters. This indicated that the lack of considerable amount of genetic diversity among the accessions.The maximum intra cluster distance was found in cluster IV (11.74) followed by cluster III (11.6).

The inter cluster distance ranged from (10.62) to 48.53. The maximum inter cluster distance (48.53) was observed between cluster VIII and XII followed by XI and XII (45.46), which indicated that the genotypes in these clusters were more diverse between them than other clusters. The cross between the genotypes of these two divergent clusters will give heterotic hybrids with high yield and fibre quality and also the resultant segregants from the crosses would be of promising nature. Minimum relative inter cluster distance was observed between V and XII (10.6) suggesting the genotypes of these clusters were not genetically much diverse. The highest intra cluster distance was observed in the cluster IV which infers that the genotypes in these clusters had an accountable amount of difference within themselves. Genetic diversity study also provided information on the characters that contributed maximum to the total divergence among the genotypes. The relative contribution of 16 traits to the total genetic diversity is furnished in the table 4. The percentage of contribution of each characters to genetic diversity varied from 0 (no. of seeds/ boll) to 63.6 (single plant yield). Among the characters studied, single plant yield (63.6) contributed more towards the total genetic diversity followed by 2.5 % span length (26.6). Selection based on these characters will be effective in improving the single plant yield in cotton. Rajarathinam *et al*., 1994 also reported that the considerable contribution of 2.5 % span length to total genetic divergence. The characters plant height and number of seeds per boll contribute minimum to the genetic divergence. The cluster mean values for different characters can also be considered for creations highest possible variability in the yield components. The superior

DOI: 10.5958/0975-928X.2017.00024.2

cluster was selected based on overall scoring of cluster mean. By selecting the accessions for hybridization based on the cluster mean will lead to crop improvement.

The mean values for all sixteen traits in respect to 12 clusters was estimated and given in the table 5. Based on the cluster mean values, the cluster XI recorded highest mean values for the characters days to 50% flowering (59), plant height (134.5), internode length (6), number of sympodia per plant (25.5), number of bolls per plant (38.5). But through overall scoring of cluster mean cluster X was observed as superior elevate with a score of 54. Cluster XI include only one accession that is SVPR 4 and the cluster X contained TCH 1732. From this study it was concluded that these accessions (TCH 1732 and SVPR 4) from the cluster X and XI was indented as best genotypes for hybridization programme to serve as parents. The highest mean for number of bolls per plant was recorded by the cluster XI, for boll weight by cluster X, for single plant yield by cluster XII and for fibre length by the cluster VI. The lowest value for single plant yield was exhibited by the clusters VIII. From this result it was concluded that the crosses between SVPR 4(XI), TCH 1732(X), Okra Narrow (XII) and TCH 1710(VI) were expected to express high heterosis followed by high yielding segregants with desirable characters.

Genetic diversity using Molecular Markers: Successful breeding program depends on the complete knowledge and understanding of genetic diversity in and among genetic resources of available germplasm which enable plant breeders to choose parental sources that will generate diverse population for selection. Germplasm is an asset for breeders who involved in cotton improvement. The Lack of genetic diversity is implicated in the slowing of progress in developing new cotton cultivars with improved yield and quality potential as well as stress resistance. In order to widen the cotton genetic base this may be accomplished by collection of available germplasm or developing intra/inter specific hybrids. D^2 statistics helps in the selection of genetically divergent parents for their exploitation in the hybridization programme. Morphological markers are routinely used to study the genetic diversity. In general, low levels of genetic diversity have been found in modern cotton cultivars, due to narrow genetic base of upland cotton germplasm used in breeding (Meredith, 2000). But to overcome the limitations associated with the morphological markers, a large number of molecular markers have come up in recent past. These molecular markers can increase the speed and precision of breeding in selection of diverse parents to widen the breeding gene pool. DNA-based methods have been employed in studies of cotton genetic diversity and in genetic improvement of the crop.

A total of 22 primer pairs were screened to identify the polymorphic markers that would discriminate the genetic relationship of 24 *G. hirsutum* accessions. Out of this 22 primer pairs ten primers were found to be polymorphic and they amplified a total of 30 alleles. The number of bands amplified for SSR primers ranged from one to two. The maximum number of amplified product (2) was observed in the profiles of primers NAU 3207, NAU 3194, NAU 3632, NAU 4900, NAU 2591, NAU 3479, MUSS 432, whereas the remaining primers showed minimum number of amplified product (1). The average number of alleles generated per primer was 1.3. It is higher than that reported by Liu *et al*., (2000) and Lacape *et al.* (2007) and lower than Venkatesan (2008). The PIC value was the highest for the SSR primer NAU 3479 (0.73) and lowest for the primer NAU 3052 (0.37). The SSR markers which resulted high PIC value could be inferred as more informative. Hence the primers, NAU 4900 and NAU 2591 were found to be the most informative as they contain high PIC value. It was concluded that the primers of which mentioned above were efficient in genetic diversity analysis in cotton.

The binary data from the polymorphic primers were used for computing Jaccard's similarity indices. The similarity co efficient value ranged from 0.3 to 1. The maximum similarity percentage (1.0) was observed between MCU 7 and MCU 5,Okra Narrow and MCU 7, Okra Narrow and MCU 7, RB 488 and CNH 2124. The least similarity percentage was found between Giza 1461 and RAC 9544 (0.3). From this investigation it was suggested that the germplasm accessions which had least similarity percentage between them can be used as a potential parents in hybridization programme to get heterotic F_1 hybrids. Highest similarity of 100% was observed. Hence to improve desirable traits in superior varieties these accessions can be used as parents for hybridization programme. It was concluded that this 100% similarity between these germplasm accessions may be due to their common parent background. Similar results were obtained already by Hirut *et al.* (2007). The dendrogram was constructed using NTSYS pc 2.02 i program based on the similarity index values. The dendrogram resulting from the UPGMA cluster analysis is depicted in the figure 2. The 24 accessions were grouped into four clusters (Table 6) of which cluster I had 13 accessions in it and the cluster II had only one accession (CNH 152). These clusters were identified at 0.70 similarity coefficient. These 24 accessions used were of same origin (India) but they grouped into four clusters. This was supported by the study of Kowsalya and Raveendran (1996). This genetic diversity evaluation concluded that all the characters under study showed significant differences among the accessions. Potential donor with highest mean value can be exploited for

DOI: 10.5958/0975-928X.2017.00024.2

further improvement in breeding programme. The genotypes from the diversified clusters can be used as parents for hybrid and varietal development programme. From the molecular study, the genetically diversified genotypes like Giza 1461 and RAC 9544 which were identified can also be used for hybridization programme.

References

Hirut Kebede, Gloria Burow, Raviprakash, G. Dani and Randy D. Allen. 2007. A-genome cotton as a source of genetic variability for Upland cotton (*Gossypium hirsutum*). *Genetic Resources and Crop Evolution*, **54**: 885–895.

Kowsalya, R. and T.S. Raveendran, 1996. Genetic variability and D^2 analysis in upland cotton. *Crop Res.*, **12**(1): 36-42.

Lacape, J., Dessauw, D., Rajab, M., Noyer, J.L.A.E., and Hau, B. 2007. Microsatellite diversity in tetraploid *Gossypium* germplasm: assembling a highly informative genotyping set of cotton SSRs. *Mol. Breed.*, **19**: 45–58.

Liu.S, Cantrell, R.R., McCarty, J.C. and Stewart, J. McD. 2000. Simple sequence repeat based assessment of genetic diversity in cotton race stock accessions. *Crop Sci*., **40**: 1459-1469.

Mahalanobis, P.C. 1936. On the generalized distance in statistics. *Proc. Natl. Acad. Sci.*, India **2**: 49-55.

Meredith, W.R.Jr. 2000. Cotton yield progress- why has it reached a plateau? *Better Crops*, **84**: 6-9.

Percival, A.V. and Kohel, R.J. 1990. Distribution, collection and evaluation of *Gossypium*. *Adv. Agron.*, **44**: 225-256.

Perrier, X. and Jacquemoud-Collet, J.P. 2006. *DARwin Software*. http://darwin.cirad.fr/darwin

Rajarathinam, S., Nadarajan, N. and Sukanya Subramanian. 1994. Genetic variability and association analysis in cotton (*Gossypium sp*). *J. Indian Soc. Cotton Improv.*, **18**(1): 54-59.

Rao, C.R. 1952. Advanced statistical methods in Biometrical Research, Widely and Sons, New York.

Sambamurthy, J.S.V., Reddy, D.M. and Reddy, K.H.G. 1995. Studies on the nature of genetic divergence in upland cotton. (*Gossypium hirsutum* L.). *Ann. Agric. Res.*, **16**: 307 -310.

Singh, T.H and Gill, S.S. 1984. Genetic diversity in upland cotton under different environment. *Indian J. Genet. Plant Breed.*, **44**: 506-513.

Singh, V.V. and Singh, A.K. 1984. Studies on genetic diversity in upland cotton. *ISCI journal*, pp. 41-45.

Venkatesan, S. 2008. Genetic diversity analysis through morphological and molecular markers in cotton (*Gossypium* spp.) M.Sc., (Agri.). Thesis, T.N.A.U. Coimbatore (Unpublished).

DOI: 10.5958/0975-928X.2017.00024.2

Table 1. List of *G. hirsutum* used for constructing the genotypic dendrogram

Sl. No.	Entries	Sl. No.	Entries	Sl. No.	Entries
1	MCU 5	9	CCH 526612	17	RAC 9544
2	MCU 7	10	CA 52	18	F-1946
3	PSCL VII	11	RHC 1694	19	L-752
4	Orka Narrow	12	HS 258	20	GISV 201
5	CCH 2117	13	ICMF 118	21	MHIS 7
6	ARB 2001	14	LH 1961	22	Giza 1461
7	GIHV 370	15	CNH 2124	23	Anjali
8	CNH 152	16	RB 488	24	Sumangala

Table 2. Clustering pattern of *G. hirsutum* accessions using D^2 analysis

Cluster number	Number of accessions	Accessions/genotypes
I	12	SVPR 2, KC 3, G-cot 16, CCH 2117, Pusa 953, ARB 2001, HS 258, CNH 2124, L 752, GISV 201, MHIS 5, Giza 1461
II	7	MCU 12, TCH 1715, TCH 1716, TCH 1744, LH 1961, RB 488, Sumangala
III	15	MCU 9, LRA 5166, KC 2, CCH 314, PSCL VII, CCH 727, AKH 8683, RS 2186, CSH 51, GIHV 370, CCH 526612, CA 52, ICMF 118, SCS 102, Anjali
IV	4	MCU 7, SVPR 3, RH 61694, MHIS 7
V	1	Sara 2
VI	9	MCU 5, MCU 13, Surabhi, TCH1710, tch1608, TCH 1734, RAC 9740, CMH 152, F 1946
VII	1	NDLH 1588
VIII	1	TCH 1705
IX	1	RAC 9944
X	1	TCH 1732
XI	1	SVPR 4
XII	1	Okra Narrow

Table 3. Intra and inter cluster distances between different groups of *G. hirsutum* accessions

Clusters	1	2	3	4	5	6	7	8	9	10	11	12
1	9.0	13.2	17.6	13.0	18.4	16.9	24.2	28.5	25.3	21.3	25.5	23.5
2		9.9	25.1	17.5	14.0	20.1	31.0	34.3	31.4	13.2	31.7	17.8
3			11.6	21.6	32.5	16.3	14.0	18.5	16.6	34.9	16.1	37.6
4				11.7	17.5	24.1	28.7	34.9	31.1	24.3	31.4	23.1
5					0.0	29.8	39.2	44.4	40.4	11.8	40.9	10.6
6						11.4	18.0	18.9	18.1	28.4	17.4	33.7
7							0.0	15.0	12.3	40.3	10.7	44.9
8								0.0	14.2	43.6	11.8	48.5
9									0.0	40.6	11.7	44.4
10										0.0	40.7	13.1
11											0.0	45.5
12												0.0

DOI: 10.5958/0975-928X.2017.00024.2

Table 4. Percentage contribution of different characters to the total genetic divergence

Characters	Number of first rank	Percentage contribution
Days to 50% flowering	7	0.5
Plant height (cm)	0	0.0
Inter node length (cm)	1	0.1
Number of sympodia / plant	10	0.7
Number bolls / plant	7	0.5
Boll weight (g)	1	0.1
Number of seeds / boll	0	0.0
Number of ovules / flower	13	0.9
Seed setting percentage	2	0.1
Single plant yield (g)	910	63.6
Lint index	4	0.3
Seed index	7	0.5
Ginning out turn (%)	3	0.2
2.5 % span length (mm)	380	26.6
Bundle strength (g / tex)	16	1.1
Micronaire value	70	4.9
Total	**1431**	**100.0**

Table 5. Cluster mean for the characters in the germplasm accessions of *G. hirsutum*

Clusters	DFF	PH	INL	NSP	NBP	BW	NSB	NOF	SSP	SPY	LI	SI	GO	2.5%SL	BS	MV
1	57.1	100.7	5.3	18.6	24.0	4.4	27.2	30.9	88.6	63.7	5.9	9.5	38.5	27.5	20.0	4.2
2	55.9	114.4	5.6	18.9	22.6	4.6	27.4	30.4	90.5	74.2	5.9	9.6	38.8	30.0	21.5	4.3
3	55.8	107.2	5.3	20.3	23.6	4.2	26.1	29.7	88.4	45.5	5.4	9.3	37.0	27.2	19.5	4.3
4	55.0	100.8	5.3	19.6	23.5	4.1	27.6	30.1	92.1	68.3	5.0	9.0	36.0	24.8	18.6	4.8
5	58.5	116.5	5.5	20.5	21.7	3.7	28.5	31.5	91.5	87.0	5.6	8.9	38.5	27.1	19.3	4.6
6	57.9	112.1	5.6	19.2	22.7	4.6	27.4	30.9	89.1	51.9	5.5	9.7	36.0	31.3	22.0	4.2
7	56.0	128.0	5.5	23.5	23.0	4.4	35.5	38.0	93.0	37.3	5.2	9.8	34.6	29.2	18.5	5.1
8	58.0	117.5	4.0	21.0	23.0	4.2	28.0	29.5	95.0	31.7	6.5	8.4	44.0	30.8	22.9	4.1
9	59.0	126.0	6.0	18.5	24.0	4.3	29.0	32.5	89.5	33.7	7.9	16.3	34.2	29.7	19.4	4.4
10	55.0	103.5	5.5	20.0	23.2	5.2	32.0	33.0	97.0	87.4	7.1	10.6	40.1	31.3	22.6	4.5
11	59.0	134.5	6.0	25.5	38.5	2.4	29.5	31.5	93.5	34.2	5.4	10.6	33.7	29.9	20.2	4.6
12	58.0	120.0	6.0	17.5	25.7	4.2	29.0	35.0	83.0	91.6	5.6	13.2	31.2	27.6	22.6	4.3

DFF – days to 50 % flowering; PH- plant height; INL – internode length; NSP – number of sympodia/plant; NBP – number of bolls/plant; BW – boll weight; NSB – number of seeds/boll; NOF – number of ovule/flower; SSP – seed setting per centage; SPY – single plant yield; LI – lint index; SI – seed index; GO – ginning outturn; 2.5%SL – 2.5% span length; BS – bundle strength; MV – micronaire value

Table 6. Clustering pattern of G. *hirsutum* accessions using SSR marker

Cluster No.	No. of accessions	Accessions
I	13	MCU 5, MCU 4, Okra Narrow, RHC 1694, Anjali, PSCL VII, L-752, GISV 201, LH 1961, MHIS 7, Giza 1461, CCH 526612, CCH 2117
II	1	CNH 152
III	4	ARB 2001, HS 258, RAC 9544, Sumangala
IV	6	GIHV 370, CNH 2124, RB 488, CA 152, F 1946, ICMF 118

Electronic Journal of Plant Breeding, 8(1): 169-175 (March 2017)
ISSN 0975-928X
DOI: 10.5958/0975-928X.2017.00024.2

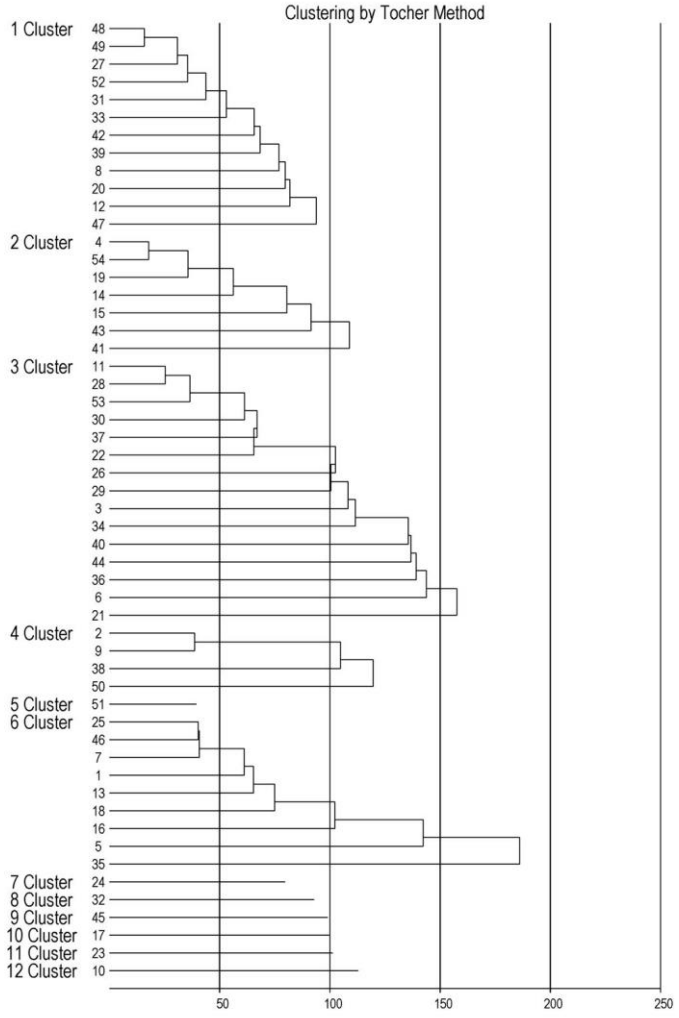

Fig. 1. Grouping of accessions in *G. hirsutum*

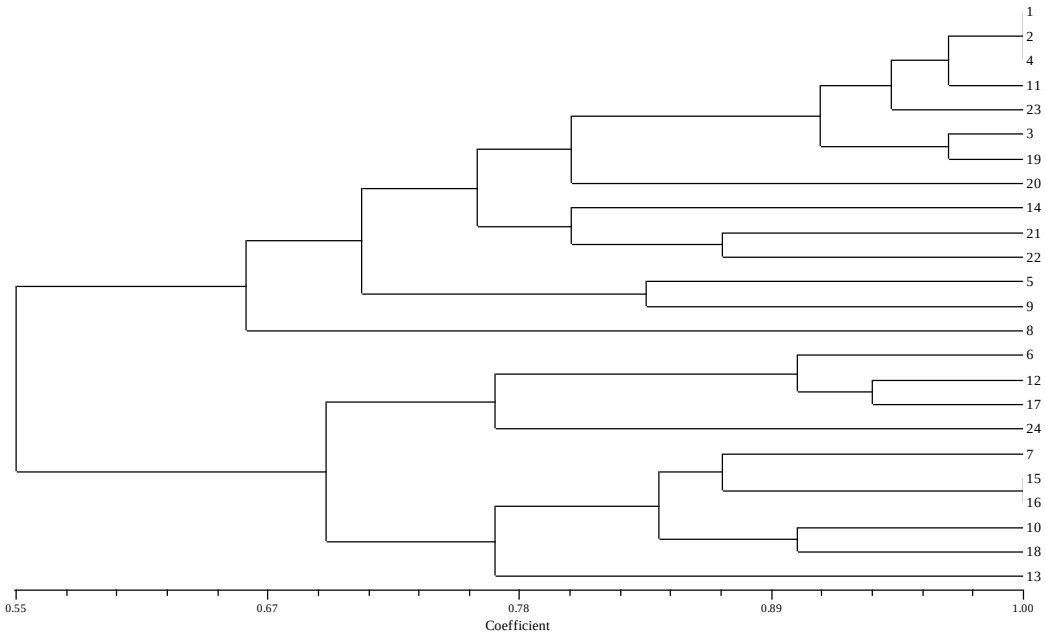

Fig. 2. Dendrogram generated using SSR data for *G. hirsutum*