Article

Sampling and exploitation of genetic variation exist in locally adapted accessions using phenotypic and molecular markers for genetic improvement of cotton

Krishnasamy Thiyagu[1], Narayanan Manikanda Boopathi[2], Nagasamy Nadarajan[1], Ayyanar Gopikrishnan[2], Pandi Selvakumar[1], Santoshkumar Magadum[1] and Rajasekar Ravikesavan[1]

[1]Department of Cotton, Centre for Plant Breeding and Genetics

[2]Cotton Genomics Laboratory, Department of Plant Molecular Biology and Biotechnology, Centre for Plant Molecular Biology, Tamil Nadu Agricultural University, Coimbatore – 641 003, India

Geneconserve 10(40): 129-153

Received December 24, 2010

Accepted January 24, 2011

ABSTRACT

Inherent diversity and parenthood of germplasm play an important role in genetic improvement of cotton. The present investigation was conducted to assess the genetic divergence among 19 elite lines of cotton that are more frequently employed in regional cotton breeding programs using morphological (multivariate Mahalanobis D2 statistics) and molecular (cluster analysis using simple sequence repeats or SSR) markers. The genotypes were grouped into nine different groups based on 16 yield and fibre quality

traits and the inter-cluster distance ranged from 11.27 to 62.75. Sixty six SSR markers that span entire cotton genome have amplified a total of 161 alleles with an average of 2.44 alleles per microsatellite locus. Genetic similarity coefficient was estimated for all the Gossypium accessions which ranged from 0.49 to 1.00. The polymorphism information content (PIC) value of SSR ranged from as low as 0 to 0.776 with an average of 0.334. This study grouped the 19 Gossypium accessions into two main cluster i.e., G. hirsutum and G. barbadense accessions and identified genetically distinct cultivars. Cultivars identified in this study may constitute potentially important source of germplasm for future breeding programme as well as genetic dissection of complex agronomic traits.

Key words: cotton; dendrogram; molecular diversity; morphological diversity.

INTRODUCTION

Cotton (Gossypium spp.) is an important natural fibre crop of global importance and widely cultivated in 30.29 million ha (Mha) (http://www.fas.usda.gov/psdonline/psdreport.aspx). In India, cotton is being cultivated in ~10 Mha with an expected productivity of 488 kg ha-1 during 2009–2010 (http://cotcorp.gov.in/statistics.asp). The productivity of Indian cotton cultivars is very low when compared with world average productivity (755 kg ha-1), due to the key fact that most of the cotton crop are grown under rainfed condition. The demand for raw cotton is ever increasing and the Indian Cotton Mills Federation has projected the cotton requirement for 2025 at 60 million bales (Deshpande et al., 2008). Apart from these, the widespread use of high speed spinning technology in the textile mills has increased the demand for raw cotton fibre with higher strength and length. Hence, cotton fibre productivity and quality must be improved to remain competitive with synthetic fibres and to meet the needs of new spinning and weaving methods (Kohel, 1999). Therefore, one of the main objectives of cotton improvement program is to breed varieties and hybrid cultivars with high seed cotton yield and superior quality of fibre. Cotton hybrids are always being utilized to give an extra push to yield and quality potential.

Knowledge on genetic variation exists in the available breeding materials is a prerequisite for any crop improvement programme. Assessment of the extent and distribution of genetic variation in crop species and their relatives is essential in understanding pattern of diversity and evolutionary relationships among accessions that help to sample genetic resources in a more systematic fashion for conservation and plant improvement (Wu et al., 2007). Genetic similarity or genetic distance estimates among genotypes are helpful in selecting parental combinations for creating diverse segregating populations for genetic dissection of quantitative trait loci (QTL) followed by marker assisted selection (MAS) and classification of germplasm into heterotic groups for hybrid breeding programme (Ali et al., 2008).

Traditionally, genetic diversity is assessed based on morphological features such as plant height, reproductive features, day length sensitivity and local adaptation and

biochemical markers such as isozymes in cotton. Through such characters only a part of the total genetic variation may be revealed. Further, it may not provide an accurate indication of the genetic divergence among cultivars or species due to environmental influence and developmental specific trait expression. However, morphological markers have their own role and commonly used in genetic diversity analysis despite their limitation. Advent of SSR have brought a new, user friendly and highly polymorphic class of genetic markers for cotton in addition to morphological markers. SSR are polymerase chain reaction (PCR) based markers, usually codominant, well dispersed throughout the genome and easily shared between labs via flanking primer sequence (Shaghai Maroof et al. 1994). Reddy et al. (2001) suggested that the total pool of SSR present in the cotton genome is sufficiently abundant to satisfy the requirements of extensive genome mapping and MAS.

SSR have been largely employed to study the extent of genetic diversity among cotton germplasm. Even though some of the studies clearly discriminate the evaluated germplasm (Bertini et al., 2006; Guo et al., 2007; Lacape et al., 2007; Boopathi et al., 2008) and phylogenic evolution of Gossypium species (Wu et al., 2007), few studies have revealed poor level of polymorphism within G. hirsutum genotypes (Liu et al., 2000; Rungis et al., 2005; Abdurakhmonov et al., 2009). Hence, to enhance the efficiency of discrimination power, deployment of both morphological and molecular markers in genetic diversity analysis will be an ideal strategy.

In the state of Tamil Nadu, India cotton is cultivated both under irrigated and rainfed condition. The textile industry of the state is mainly based on extra long stable with high bundle strength of cotton to the extent of 89%. All the four cultivated species of cotton, namely G. hirsutum, G. barbadense, G. arboreum and G. herbaceum, are grown in this state along with their intra- and inter-specific hybrids. Among the southern states of India which grow G. barbadense genotypes, the climate of Tamil Nadu is conducive for G. barbadense hybrids and varieties. Tamil Nadu produces only one sixth of its cotton requirements and the balance is met by purchases from up-country markets. Among the seven agroclimatic zones of Tamil Nadu, cotton was grown in four agroclimatic zones viz., Cauvery delta zone (Rice fallow / Short duration), Western zone (Winter irrigation), North western zone (Winter irrigation) and Southern zone (Summer irrigated / Rain fed / Winter rainfed), where yield increase by all means is stagnated for several years. Hence, there is a scope for increasing the productivity and production of cotton by developing improved varieties and hybrids. The high heterotic hybrids and high yielding cultivars could be produced by selecting the germplasm that are adapted to the given agroclimatic zones and having divergence for yield and fibre quality traits. Further, identification of phenotypically and genotypically contrasting parents could be useful in linkage map construction and genetic dissection of agronomically important traits via QTL mapping. This condition opens up new avenue in plant breeding which will increase the efficiency of breeding process via MAS.

Hence, the objective of the present investigation was analyzing the genetic diversity, both at morphological and molecular level, existing in the G. hirsutum and G.

barbadense accessions which would be helpful in selecting the diverse parents for heterotic breeding programmes and development of mapping population to genetically dissect yield and fibre quality traits.

MATERIALS AND METHODS

Plant materials

A total of 19 cotton genotypes (15 G. hirsutum and four G. barbadense) that are currently being cultivated and representing all the four agroclimatic zones in Tamil Nadu, India were subjected to diversity analysis. The pedigree, origin/source and special features of the chosen genotypes are given in Table 1.

Diversity analysis using morphological markers Plants were evaluated in randomized block design with three replications during winter 2008–2009 at Department of Cotton, Tamil Nadu Agricultural University, Coimbatore, India. The experimental materials were raised in two rows of 6 m length with the spacing of 90 × 60 cm. Recommended agronomic practices and need based plant protection measures were followed under irrigated condition to obtain good crop stand. Observations were recorded on days to flowering (DF), days to 50 per cent flowering (DFF), plant height (cm) (PH), number of sympodial branches per plant (NSy), number of bolls per plant (NB), boll weight (g) (BW), number of seeds per boll (NSB), seed cotton yield per plant (g) (SCY), ginning percentage (%) (GP), lint index (LI), seed index (SI), 2.5 per cent span length (mm) (SL), bundle strength (g/tex) (BS), fibre fineness (micronaire) (FF), uniformity ratio (UR) and elongation percentage (%) (EP). Data were collected from five randomly selected plants of each genotype per replication. The genetic divergence was worked out by using Mahalanobis D2 statistics described by Rao (1952). On the basis of D2 values, the genotypes were grouped into different clusters by employing Tocher's method (Rao, 1952).

Diversity analysis using SSR markers

A total of 66 SSRs that span all the chromosomes of cotton were used to assess the extent of genetic diversity among the selected cotton germplasm and SSR primer sequences were obtained from the cotton Data Base (cottonDB) (http://algodon.tamu.edu/ cgi-bin/searches/browser). Genomic DNA was extracted from fresh and young leaves of cotton plants (Zhang and Stewart, 2000). The isolated genomic DNA was quantified by NanoDropTM 1000 Spectrophotometer and the final concentration of DNA was adjusted to 20 ng/µL using low salt Tris-EDTA buffer. Further, the isolated DNA was checked for its intactness, homogeneity and purity by 0.8 per cent agarose gel electrophoresis.

PCR was performed as per Zhang and Stewart (2000) with slight modification: initial extended denaturation at 93°C for 12 minutes followed by 40 cycles of denaturation at 95°C for 15 seconds, primer annealing temperature at 48°C to 58°C (varies with melting temperature (Tm) of each SSR primer pair) for 30 seconds and elongation at 72°C for 1

minutes. The 40th cycle was followed by an extended primer extension step at 72℃ for 6 minutes and then being held at 4℃ until electrophoresis. The PCR products were resolved using 3% metaphor agarose gel electrophoresis and documented.

After visualizing the gel, amplified fragments of each SSR allele was binarily scored as ‘1’ and ‘0’; where ‘1’ indicated the presence of a specific allele and ‘0’ indicated its absence. Polymorphism information content (PIC) of SSR markers was calculated as per Anderson et al. (1993). Genetic diversity analysis was done using NTSYSpc ver. 2.02i (Rohlf, 2000). Genetic similarities (GS) between pairs of accessions were measured by the DICE similarity co-efficient based on the proportion of shared alleles with SIMQUAL module. Genetic distances (GD) between pairs of lines were estimated as GD = 1 GS. The clustering of accessions was done based on a similarity matrix using an unweighted pair group method with arithmetic average (UPGMA) algorithm following SAHN module. The clustering result was used to construct a dendrogram following TREE module (Ali et al., 2008).

RESULTS AND DISCUSSIONS

Genetic diversity is the base of biological and species diversity which is essential in the context of evolution. Genetic diversity and the knowledge on relationship between genotypes are of great importance for cotton breeding, since it provides an ample supply of allelic variation that can be used to create new favorable gene combinations. Hence, precise evaluation of the genetic divergence in the available germplasm will provide a guide for choosing parents, predicting the degree of inheritance, variation and level of heterosis etc., which is essential for realizing the breeding goal. Generally, diversity is measured by genetic distance or similarity, both of which indicate that there are either differences or similarities, respectively, at genetic level (Weir, 1990). Conventionally this has been done with morphological markers and recently the efficiency of genetic diversity analysis is accelerated by use of molecular markers such as SSRs. Combination of both morphological and molecular markers were employed in this study to capture the best features of both of these strategies in genetic diversity analysis.

Morphological diversity analysis

The analysis of variance using morphological traits revealed that mean squares due to genotypes were highly significant for all the traits studied and indicated existence of considerable genetic diversity among genotypes. Hence, further analysis was done to estimate the D2 values and on the basis of relative magnitude of D2 values, all the 19 genotypes were grouped into nine clusters (Table 2).

A maximum of four genotypes were included in cluster V followed by three genotypes in cluster I. The remaining clusters viz., II, III, IV, VI, VII had two genotypes each while clusters VIII and IX had only one genotype each. The pattern of grouping into different clusters was at random. These results are in confirmation with the earlier studies on Gopinath et al. (2009) and Satish et al. (2009). The geographical distribution of

genotypes is not the only factor that causes morphological genetic diversity. Genetic diversity may be due to the outcome of several other factors like natural and artificial selection, exchange of breeding material, genetic drift and environmental variation. Therefore, selection of parents for hybridization programme should be based on genetic rather than geographical diversity. For example, though the MCU and TCH prefixed lines were developed from same geographical location (Table 1), they were grouped into different clusters along with other genotypes that were developed at different geographical locations. Therefore, the selection of parents for hybridization programme should be based on genetic diversity rather geographic diversity.

The average values of intra- and inter-cluster D2 and distance (D) values are presented in Table 3. The maximum inter-cluster distance was observed between clusters V and VIII (62.75) followed by clusters V and IX (56.20) and least inter-cluster distance was observed between clusters II and III (11.27). Based upon these results, crosses may be attempted between genotypes of cluster V (KC 3, SVPR 2, SVPR 3 and TCH 1705) and cluster VIII (Suvin) and / or cluster IX (TNB 1) to obtain promising recombinants in cotton. The maximum intra-cluster distance values were observed in the cluster V (21.67) followed by cluster VII (19.66) and cluster I (15.31) while minimum intra-cluster distance value was nil in clusters VIII and IX as each cluster consisted of only one genotype. The high intra-cluster distance in cluster V indicated the presence of wide genetic diversity among KC 3, SVPR 2, SVPR 3 and TCH 1705. Therefore, selection and hybridization of genotypes with high D2 value falling in two different clusters (cluster V and VIII and cluster V and IX) would produce potential hybrids and desirable segregants in further final generations. Use of such genetically distant genotypes as parents to get the most promising breeding material has already been reported (Amudha et al., 1997; Sandhu and Boparai, 1997).

Endang et al. (1971) stated that the clustering pattern could be utilized to choose parents for hybridization programme since this can generate maximum possible variability due to highest mean value in genetically distinct clusters for important yield and fibre quality characters. By the study of cluster means among nine clusters for yield and fibre quality traits (Table 4), it was understandable that considerable differences were noticed between the clusters. The lowest mean values for days to 50% flowering is recorded in cluster V (61.17) thus providing its desirability for earliness and the highest mean value for number of sympodial branches per plant is in cluster IX (18.49); for number of bolls per plant is in cluster VI (29.16); for seed cotton yield per plant is in cluster VI (96.55). The genotypes that fall under these clusters can be used for direct adoption or hybridization to improve the seed cotton yield and other yield attributes. The genotypes that fall under clusters IV recorded highest mean value for ginning percentage (37.45) and for lint index (6.42). The highest mean value for fibre quality traits were recorded in cluster VIII for 2.5 per cent span length (37.47) and bundle strength (27.30). Improving cotton fibre length and strength are the main breeding targets to meet out the growing demand by modernized textile mills. Hence, Suvin which grouped in cluster VIII can be utilized as the potential parent in hybridization programme for improving the fibre quality.

In summary, the cluster V (KC 3, SVPR 2, SVPR 3 and TCH 1705) and the cluster VIII (Suvin) showed higher genetic distance and divergence for different seed cotton yield and fibre quality traits and hence the hybrids or segregants derived from these genotypes could be useful for simultaneous improvement of both yield and fibre quality traits.

Molecular diversity analysis

Molecular marker analysis discloses genetic differences at the DNA level in plants and is an effective tool for testing genetic diversity of germplasm (Jia, 1996; Xie et al., 1998). Molecular based genetic diversity analysis also has potential for assessing changes in genetic diversity over time and space (Duwick, 1984). A total of 66 SSRs were amplified in 19 Gossypium spp., and 51 SSRs were found to be polymorphic (77.27%) among G. hirsutum and G. barbadense genotypes. A sum of 161 alleles were amplified by 66 SSRs with an average of 2.44 alleles per microsatellite locus, similar to Bertini et al. (2006) who had noticed an average of 2.13 alleles per locus. However, use of converted G. hirsutum backcross derived accessions from various wild race stocks by Liu et al. (2000) helped them to report an average of five SSR alleles per locus which may possibly be due to inclusion of wild species in their analysis.

The PIC value calculated to estimate the informativeness of each primer pair was ranged from as low as zero (for example, BNL2921) to 0.776 (BNL2960) with an average of 0.334. Similar kind of PIC values were observed by Liu et al., (2000) and slightly higher PIC values was noticed by Bertini et al. (2006). Genetic similarity co-efficient was estimated for each pair of the 19 Gossypium accessions which ranged from 0.49 to 1.

The cluster analysis grouped the 19 Gossypium accessions into two main clusters viz., I and II. As expected, Cluster I included all the 15 G. hirsutum accessions and cluster II included all the four G. barbadense accessions (Fig. 1). In both the clusters, genotypes that were developed at same geographic regions were grouped as sub clusters.

For example, the genotypes (MCU and TCH lines) were grouped in clusters Ia, Ib, Ic and Id which were evolved at Cotton Breeding Station, Tamil Nadu Agricultural University, Coimbatore; The KC lines (KC 2 and KC 3), which were developed from Agricultural Research Station, Kovilpatti, using at least one parent as common in their parentage, were clustered in Ie and SVPR 3 developed from Cotton Research Station, Srivilliputtur, was clustered in If. In contrast, the sub cluster Ia contained the genotypes that were originated both from Srivilliputtur (SVPR 2) and Coimbatore (MCU 7, TCH 1705, MCU 9, MCU 12 and TCH 1218). This again indicated that the genetic diversity of cotton is not completely related to geographic distribution. This may be attributed to either occurrence of similar genetic variation independently in the different geographic regions or artificial transfer of the accessions from one region to another region (Huang et al. 2002). G. barbadense accessions Barbados and EC 101786 were found in sub cluster IIb, while Suvin and TNB 1 was found in sub clusters

IIa and IIc, respectively. These G. barbadense accessions have been developed from different sources or origins (Table 1) and hence they might be grouped into different groups. Based on genetic similarity co-efficient analysis estimated by SSR markers, it was clear as that of morphological trait based genetic diversity that the genotypes KC 2, KC 3, SVPR 3 and SVPR 2 showed lower genetic similarity (~50%) with Suvin when compared with other groups.

Conclusion

Narrow genetic pool exists in present cotton cultivars is detrimental to future breeding program. The genetic relationship information generated in this study among Gossypium spp., will be helpful in future regional cotton breeding programme to maintain genetic diversity and improve fibre productivity and quality. Both morphological and molecular marker based genetic diversity analysis clearly indicated that the G. hirsutum genotypes viz., SVPR 2, SVPR 3, KC 2, KC 3 and TCH 1705 and G. barbadense genotype viz., Suvin showed higher genetic distance between these species. This denoted that recombinants developed from these genotypes would give higher heterotic percentage for yield and fibre quality traits. Further, based on higher inter cluster distance and lower genetic similarity co-efficient, this study has also identified pairs of cultivars SVPR 2 / Suvin and KC 3 / Suvin which could be used as parents to create mapping populations. Two types of mapping populations, recombinant inbred lines (RIL) and back cross (BC) progenies are being developed using these parents for genetic dissection of fibre quality traits. Interestingly, KC 3 and SVPR 2 are well adapted to water limited environments and hence these mapping populations can also serve as a potential genetic material for mapping quantitative trait loci and further marker-aided breeding for drought-prone environments, besides fiber quality traits in cotton.

ACKNOWLEDGEMENT

This work was funded by Department of Biotechnology, Ministry of Science and Technology, Government of India under Program Support for Research and Development in Agricultural Biotechnology at Tamil Nadu Agricultural University.

REFERENCES

Abdurakhmonov IY, Saha S, Jenkins JN, Buriev ZT, Shermatov SE, Scheffler BE, Pepper AE, Yu JZ, Kohel RJ, Abdukarimov A (2009). Linkage disequilibrium based association mapping of fibre quality traits in G. hirsutum L. variety germplam. Genetica 136: 401–417.

Ali ML, Rajewski JF, Baenziger PS, Gill KS, Eskridge KM, Dweikat M (2008). Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. Mol. Breed. 21: 497–509.

Amudha K, Raveendran TS, Krishnadoss D (1997). Genetic diversity in coloured linted

cotton varieties. Madras Agric. J. 84: 334–337.

Anderson JA, Churchill GA, Autrique JE, Sorrellis ME, Tanksley SD (1993). Optimizing parental selection for genetic linkage maps. Genome 36: 181–186.

Bertini HC, Schuster I, Sediyama T, Barros EGD, Moreira MA (2006). Characterization and genetic diversity analysis of cotton cultivars using microsatellites. Genet. Mol. Biol. 29: 321–329.

Boopathi NM, Gopikrishnan A, Selvam NJ, Ravikesavan R, Iyanar K, Muthuraman S, Saravanan N (2008). Genetic diversity assessment of G. barbadense accessions to widen cotton (Gossypium spp.) gene pool for improved fibre quality. J. Cotton Res. Dev. 22: 135–138.

Deshpande NP, Dheva NG, Potdukhe NR, Parmar JN, Wadodkar MB (2008). Combining ability study for yield and morphological traits in interspecific crosses between G. hirsutum L. and G. barbadense L. Intl. J. Plant Sci. 3: 158–162.

Duwick DN (1984). Genetic diversity in major farm crops on the farm and reserve. Econ. Bot. 32: 161–178.

Endang S, Andani S, Nasoetion H (1971). Multivariate classification of some rice (Oryza sativa) varieties and strain on yield components. In: Rice, Canon Newsl. 20: 26–34.

Gopinath M, Rajamani S, Naik RK, Rao CHM (2009). Genetic divergence for lint characters for upland cotton (G. hirusutum L.). J. Cotton Res. Dev. 23: 18–22.

Guo WZ, Sang ZQ, Zhou BL, Zhang TZ (2007). Genetic relationships of D-genome species based on two types of EST-SSR markers derived from G. arboreum and G. raimondii. Plant Sci. 172: 808–814.

Huang XQ, Borner A, Roder MS, Ganal MW (2002). Assessing genetic diversity of wheat (Triticum aestivum L.,) germplasm using microsatellite markers. Theor. Mol. Biol. 105: 699–707.

Jia JZ (1996). Molecular germplasm diagnostics and molecular marker assisted breeding. Scientia Agricultura Sinica 29: 1–10 (in Chinese with an English abstract).

Kohel RJ (1999). Cotton germplasm resources and the potential for improved fibre production and quality. In: Basra AS (eds), Cotton fibres, The Haworth press, New York, pp. 167–182.

Lacape JM, Dessauw D, Rajab M, Noyer JL, 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, Saha S, Stelly D, Burr B, Cantrell RG (2000). Chromosomal assignment of

microsatellite loci in cotton. J. Hered. 91: 326–332.

Rao CR (1952). Advanced statistical methods in biometrical research. John Wiley and Sons, New York, pp. 236–272.

Reddy OUK, Pepper AE, Abdurakhmonov I, Saha S, Jenkins JN, Brooks T, Bolek Y, El-Zik KM (2001). New dinucleotide and trinucleotide microsatellite marker resources for cotton genome research. J. Cotton Sci. 5: 103–113.

Rohlf FJ (2000). NTSYS-pc: Numerical taxonomy and multivariate analysis system. Exeter publishing, Setauket. 119.

Rungis D, Llewellyn D, Dennis ES, Lyon BR (2005). Simple sequence repeat (SSR) markers reveal low levels of polymorphism between cotton (Gossypium hirsutum L.) cultivars. Australian J. Agric. Res. 56: 301–307.

Sandhu B, Boparai S (1997). Genetic divergence in Gossypium arboreum L. Indian J. Genet. 54: 461–465.

Satish Y, Jain PP, Chhabra BS (2009). Studies on genetic diversity in upland cotton (G. hirsutum L.). J. Cotton Res. Dev. 23: 18–22.

Shaghai Maroof MA, Biyashev RM, Yang GP, Zhang Q, Allard RW (1994). Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations and population dynamics. Proc. Natl. Acad. Sci. USA, 91: 5466–5470.

Weir BS (1990). Genetic data analysis: methods for discrete population genetic data. Sinauer Associates, Inc. publishers. Sunderland, Massachusetts. 377.

Wu YX, Daud MK, Chen L, Zhu SJ (2007). Phylogenetic diversity and relationship among Gossypium germplasm using SSRs markers. Pl. Syst. Evol. 268: 199–208.

Xie J, Cai Z, Liu XH, Li FH, Cao HL, Luan YC (1998). Application of biotechnology on evaluation of genetic diversity of germplasm. Crops (Supplement): 71–76 (in Chinese).

Zhang J, Stewart JM (2000). Economical and rapid method for extracting cotton genomic DNA. J. Cotton Sci. 4: 193–201.

Table 1. Details on parentage and special features of genotypes that were used in this study

S. No.	Genotypes	Parentage	Originated Sourcea	Special features
Cauvery delta zone				
1	MCU 7	X – ray mutant of L 1143EE	CBS, Coimbatore	Suitable for rice fallow and earliness
2	SVPR 3	LH 900 × 1301 DD	CRS,	Suitable for rice fallow and drought

			Srivilliputtur	tolerant
Western zone and North western zone				
3	MCU 9	MCU 8 × MCU 5	CBS, Coimbatore	Long staple and suitable for winter irrigation
4	TCH 1218	MCU 9 × TCH 92-7	CBS, Coimbatore	Long staple with good combiner
5	MCU 5	Multiple cross derivative	CBS, Coimbatore	Long staple and suitable for winter irrigation
6	TCH 1569	TCH 1002 × MCU 5	CBS, Coimbatore	Good adaptability
7	TCH 1627	T 13 × M 12	CBS, Coimbatore	High yield and fibre quality traits
8	MCU 13	Multiple cross derivative	CBS, Coimbatore	Short duration and suitable for winter irrigation
9	MCU 12	LRA 5166 × MCU 11	CBS, Coimbatore	Short duration and suitable for winter irrigation
10	TCH 1705	Selection from AKH 2053	CBS, Coimbatore	Drought tolerant
11	TCH 1608	TCH 1002 × SRT 1	CBS, Coimbatore	Good adaptability over varied environments
12	TCH 1710	TCH 1629 × HLS 72	CBS, Coimbatore	Long staple
13	Suvin	Sujatha × SIV 135	CICR, Coimbatore	Extra long staple
14	TNB 1	Giza 7 × SB 1085-8	CBS, Coimbatore	Shorter duration, drought tolerant and long staple
15	Barbados	Exotic germplasm selection	Egypt	High yield with extra long stable
16	EC 101786	Exotic germplasm selection	Egypt	High yield with extra long stable
Southern zone				
17	SVPR 2	TSD 22 × JR 36	CRS, Srivilliputtur	Drought tolerant and suitable for summer irrigation
18	KC 2	MCU 10 × KC 1	ARS, Kovilpatti	Jassid resistant and suitable for winter irrigation
19	KC 3	TKH 497 × KC 1	ARS, Kovilpatti	Suitable for rainfed cultivation

Table 2. Distribution of 19 cotton genotypes in different clusters based on distance (D2) value for 16 morphological characters

Clusters	Genotypes	Number of genotypes
I	KC 2, MCU 12 and TCH 1218	3
II	MCU 5 and TCH 1569	2
III	MCU 7 and MCU 13	2
IV	TCH 1608 and TCH 1627	2
V	KC 3, SVPR 2, SVPR 3 and TCH 1705	4
VI	Barbados and EC 101786	2
VII	MCU 9 and TCH 1710	2
VIII	Suvin	1

IX	TNB 1	1

Table 3. Average intra and inter cluster divergence (D2) values of nine clusters for 19 cotton genotypes based on within and between cluster distances for 16 morphological characters

Clustersa	I	II	III	IV	V	VI	VII	VIII	IX
I	234.24 (15.31)	236.41 (15.38)	152.29 (12.34)	223.88 (14.96)	610.54 (24.71)	1752.64 (41.87)	404.09 (20.10)	2213.16 (47.04)	1988.48 (44.59)
II		77.36 (8.80)	127.04 (11.27)	351.46 (18.75)	672.66 (25.94)	1168.85 (34.19)	222.65 (14.92)	1737.40 (41.68)	1798.41 (42.41)
III			84.39 (9.19)	183.62 (13.55)	519.99 (22.80)	1559.91 (39.50)	261.33 (16.17)	2132.84 (46.18)	2095.24 (45.77)
IV				104.92 (10.24)	387.33 (19.68)	2385.90 (48.85)	574.89 (23.98)	3120.15 (55.86)	2591.62 (50.91)
V					469.65 (21.67)	2795.34 (52.87)	1022.92 (31.98)	3937.86 (62.75)	3158.39 (56.20)
VI						137.95 (11.75)	1118.17 (33.44)	499.98 (22.36)	1262.56 (35.53)
VII							386.57 (19.66)	1516.87 (38.95)	2136.21 (46.22)
VIII								0.00 (0.00)	1510.72 (38.87)

a Figures in parentheses are distance (D)

Table 4. Cluster-wise mean values for 16 morphological characters in 19 cotton genotypes

Characters \ Clusters	I	II	III	IV	V	VI	VII	VIII	IX
DF	56.44	60.33	57.17	59.17	56.08	59.50	59.17	60.00	61.00
DFF	62.33	63.83	63.00	64.00	61.17	63.67	64.67	65.67	65.00
PH (cm)	103.45	96.81	96.59	103.25	98.43	87.68	83.75	88.27	149.76
NSy	17.15	16.28	16.51	16.27	17.61	18.25	15.10	14.96	18.49
NB	19.85	21.68	20.12	19.49	24.83	29.16	20.71	17.13	22.96
BW (g)	4.74	4.13	4.18	4.95	3.97	4.07	4.50	4.22	4.56
SPB	27.16	26.63	27.47	30.57	28.27	23.77	28.70	23.27	22.87
SCY (g)	74.81	69.63	63.76	79.12	78.66	96.55	72.55	52.27	84.70
GP (%)	35.41	34.22	36.73	37.45	35.15	33.18	34.28	29.73	31.79
LI	6.15	4.83	5.48	6.42	4.87	6.03	5.77	4.42	5.86
SI	11.18	9.73	9.72	10.69	8.92	11.87	11.00	12.07	12.62
SL (mm)	28.96	29.32	29.37	27.87	26.23	34.80	31.88	37.47	34.70
BS (g/tex)	20.48	21.55	20.30	19.13	18.33	26.23	21.98	27.30	25.87
FF	4.59	4.05	4.32	4.90	4.93	4.37	4.07	4.07	4.17
UR (%)	50.33	49.67	47.00	48.67	50.58	47.83	46.50	48.33	48.00
EP (%)	5.97	6.43	5.75	6.07	6.29	6.97	5.22	5.73	6.97

Figure 1. Dendrogram of 15 *G. hirsutum* and four *G. barbadense* cultivars revealed by cluster analysis of genetic similarity estimates generated by DICE co-efficient based on 66 SSRs.