GENETIC DIVERSITY ANALYSIS OF COTTON (Gossypium hirsutum L.) GENOTYPES USING ISSR MARKERS

Srushti Parkhiya¹, K. Gohel², Dr.D.R Mehta³

¹,²,³Department of Biotechnology, College of Agriculture, Junagadh Agricultural University, Junagadh, Gujarat, India

ABSTRACT

Cotton (Gossypium hirsutum L. 2n=52, family: Malvaceae) is one of the most important commercial fiber and oil yielding crop in India. Genetic variability and relationships among fifteen cotton genotypes were investigated using nine ISSR primers. A total of 86 bands were obtained from nine ISSR primers, out of which 54 showed polymorphism of 62.79 % with an average six bands per primer. The PIC ranged from 0.8616 to 0.9090 and genetic similarity varied from 0.60 to 0.917. The cluster analysis revealed two main clusters with highest 66 % similarity percentage. Cluster I consisted of only one genotype GJHV-460, while cluster II consisted of rest of the genotypes grouped together in their respective subclusters. Large numbers of single type of markers could be screened for genetic diversity in cotton genotypes and utilized diverse genotypes in crop improvement programme to enhance crop productivity of cotton.

Key words: ISSR, genetic diversity

I. INTRODUCTION

Cotton is one of the most important commercial fibre and oil yielding crops playing a key role in economic, political and social affairs of the world. Because of its worldwide economic importance, new cultivars are constantly being released in the world. It is cultivated in about 60 countries of the world but 10 countries viz., Russia, USA, China, India, Brazil, Pakistan, Turkey, Egypt, Mexico and Sudan account for about 85 % of the total production and in varying climatic condition.

India is the world’s second largest cotton producer after China, produced 365 lakh bales of cotton from an area of 117.78 lakh hectares with productivity of 518 kgs per hectare. Cotton is cultivated on large scale in the states of Maharashtra, Gujarat, Karnataka, Madhya Pradesh, Punjab, Rajasthan, Haryana, Tamil Nadu and Uttar Pradesh. Gujarat produced 93 lakh bales from an area of 24.97 lakh hectares with productivity of 633 kgs per hectare. (Anonymous, 2013a)
Worldwide, the textile industry depends largely on cotton (*Gossypium* sp.) fibers. Spinnable fibers are obtained from four cultivated species, *Gossypium hirsutum* L., known as Upland Cotton, Long Staple Cotton, or Mexican Cotton and *Gossypium barbadense* L., known as Sea Island Cotton, Extra Long Staple Cotton, American Pima, or Egyptian Cotton are tetraploids (2n = 4x = 52), whereas *Gossypium herbaceum* L., known as Levant Cotton and *Gossypium arboreum* L., known as Tree Cotton are diploids (2n = 2x = 26) (Wendel *et al.* 1992). Among these species, *G. hirsutum* and *G. barbadense* possess long and fine fibre.

Cotton is a warm season crop. It requires a mean annual temperature of over 16 °C and annual rainfall of at least 1200 mm well distributed throughout the growing season. Cotton can be grown on all types of soils, such as sandy loam, loam and clay loam soils. Black to medium black soils of Middle and South Gujarat regions are preferable for cotton cultivation.

Cotton is grown chiefly for its fibers, which are used in manufacturing of cloth for the mankind. It is also used for several other purposes like making threads, for mixing in other fibers and for extraction of oil from cotton seed. The oil content in the cotton seed ranges from 15 to 25 % depending on the varieties. Cotton seed oil contains 13.3 g of fat, 4.8 mg vitamin E, 3.4 mg vitamin K. The coefficient of digestibility of the oil is 97 % and nutritive value is about 9 Kilo calories per gram. Cotton seed cake after extraction of oil is good organic manure and contains about 41 % protein, 23 % fiber and many minerals like calcium, magnesium, phosphorus, potassium, sodium, sulfur, copper, iron, manganese, molybdenum and zinc (Anonymous, 2013b). The cotton seed provide oil and are also used as cattle feed. The stalk is use as fuel and the leaves falling on the ground increase organic matter in the soil. It sustains the country’s cotton textile industries. It provides around 70% of raw material for textile industry and 35-40 % of total foreign exchange (Bhaskaran and Ravikesavan, 2008).

At present cotton account for 44 % of all fibers used in the world. Such a crop has a pride of place in the farming of Gujarat’s Agriculture. The economy of the entire state is very much dependent on success or failure of this cash crop.

Several challenges have been overcome in cotton genomic research and now genetic linkage maps of cotton have been developed based on both intraspecific (*intra*-*hirsutum*) and interspecific (*Gossypium hirsutum* x *Gossypium barbadense*) population and the QTLs responsible for leaf shapes, plant trichomes, photoperiodism, stomatal conductance, disease resistance, yield and fibre quality traits have been mapped (Preetha and Raveendren, 2008).
Molecular investigations of germplasm are essential for their collection, conservation and its utilization in breeding programmes. The knowledge of genetic diversity in a crop species is fundamental to its improvement. DNA marker technology would provide a tool to the plant breeders to select desirable plants directly on the basis of genotype instead of phenotype. The use of molecular marker system offers significant advantages for species identification in that they are rapid, relatively cheap, independent of environmental conditions, eliminate the need to grow plants up to maturity. The use of molecular markers for the evaluation of genetic diversity is receiving much attention than morphological characterization. DNA markers are mainly classified into three classes based on their detection: 1) Hybridization based DNA markers 2) PCR based DNA markers 3) DNA sequence based DNA markers (Singh, 2003).

Tanksley (1983) listed five properties that distinguish molecular markers from morphological markers. These properties are (1) genotypes can be determined at the whole plant, tissue and/or cellular level, (2) a relatively larger number of naturally occurring alleles exists at many loci, (3) phenotypic neutrality i.e., deleterious effects are not usually associated with different alleles, (4) alleles at many loci are co-dominant, thus all possible genotypes can be distinguished and (5) few epistatic or pleiotrophic effects are observed.

ISSR analysis involves PCR amplification of genomic DNA using a single primer that targets the repeat per se, with 1–3 bases that anchor the primer at the 3’ or 5’ end. In addition to freedom from the necessity of obtaining flanking genomic sequence information, ISSR analysis is technically simpler than many other marker systems. The method provides highly reproducible results and generates abundant polymorphisms in many systems.

Keeping in view the above, the present investigation was planned to study molecular characterization of upload cotton (Gossypium hirsutum L.) genotypes through molecular markers with the following objectives. (1) To study molecular characterization of different cotton genotypes using molecular marker (ISSR); (2) To find out the phylogenetic relationship among different cotton genotypes and (3) To know the degree of genetic divergence among different cotton genotypes.

II. MATERIAL AND METHODS

2.1. Plant materials and DNA extraction

The experimental materials consisted of fifteen genotypes of cotton (Gossypium hirsutum) which were collected from Cotton Research Station, J.A.U., Junagadh.

2.2. Molecular analysis
Total genomic DNA was isolated from young leaves of different cotton plants grown in pots. DNA extraction was carried out by CTAB method as described by Doyle and Doyle (1987) with minor modifications. Leaf tissues were cut into small pieces, homogenized and digested with extraction buffer (pH= 8.0): 1 M Tris HCl, 0.5 M EDTA (Ethylene diaminetetraacetic acid), 5 M NaCl, 2 X CTAB, 4 % PVP and β-mercaptoethanol. After incubation at 65 °C in water bath for one hour with gentle swirling, the mixture was emulsified with an equal volume of Phenol: Chloroform: Isoamyl alcohol (25:24:1). Equal volume of ice-cold iso-propanol was added to precipitate DNA and pelleted by centrifugation. The pellets were washed with 70 % alcohol, air dried and resuspended in 100 µl of TE buffer (1 M Tris HCl, 0.5 M EDTA, pH 8.0) and finally treated with 1 µl of RNase. DNA was loaded into the sample spot of Nanodrop Spectrophotometer (Thermo Scientific, U.S.A.) and the concentration of DNA and absorbance at 260 nm and 280 nm were measured. The A$_{260}$/A$_{280}$ ratio was automatically calculated by the software.

The PCR reactions for ISSR were carried according to method given by Rana et al., (2006) with required modifications. The ISSR primers were listed in Table 1. The amplified products of ISSR were analyzed using 1.5 % agarose gel in TBE buffer.

In order to score and preserve banding pattern, photograph of the gel was taken in a Gel Documentation System, under UV trans-illuminator. The presence of each band was scored as ‘1’ and its absence as ‘0’. The data matrix was read by NTSYS-pc version 2.02 (Numerical Taxonomy and Multivariate Analysis System for Personal Computers, Exeter Software) developed by Rohlf (2000) and analyzed by the SIMQUAL (similarity for qualitative data) program with Jaccard’s similarity coefficient. The SIMQUAL is a program for computing a variety of similarity and dissimilarity coefficients for qualitative data. The qualitative nature of the absence (0) or presence (1) state of ISSR marker was used as the basis for similarity analysis among various cotton genotypes. A matrix of 0 and 1 act as the input, and the output is a matrix of similarity or dissimilarity coefficients. The resultant similarity matrix was entered into SAHN (sequential, agglomerative, hierarchical, and nested clustering method) clustering program, a tree matrix was produced and a dendrogram was constructed using UPGMA (unweighted pair-group method with arithmetic averages). The assumption underlying the use of UPGMA clustering is the equal rate of evolution along all the dendrogram branches. The dendrogram of publication quality were produced from the output tree file of SAHN by TREE (tree display) program in graphics mode.

Clustering methods create clusters of the data, no matter whether there are true clusters in the data or not, so a check was made for the existence of true clusters. This was done by using the tree
matrix produced by SAHN to calculate the cophenetic values of similarity or dissimilarity by the program COPH (cophenetic values). The cophenetic value matrix was compared with the original tree matrix for goodness of fit of the cluster analysis to the data. This type of cophenetic correlation was done by the MXCOMP (matrix comparison) program (Rohlf, 2000). The program MXCOMP plots the cophenetic value matrix against the original tree matrix, and computes the Cophenetic correlation coefficient (r) and the Mantel test statistic (Z).

The mantel test statistic $Z = \sum X_{ij} Y_{ij}$, where, $X_{ij}$ is the off-diagonal elements of cophenetic value matrix, $Y_{ij}$ is the off-diagonal elements of original tree matrix and n is the number of elements of the matrix. As the cophenetic correlation coefficient is positively correlated to the Mantel test statistic and in standardized units, it is easier to use it as a measure of goodness of fit for a cluster analysis than the Mantel test statistic. The degree of fit can be referred as follows (Rohlf, 2000).

<table>
<thead>
<tr>
<th>Level</th>
<th>Degree of Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r \geq 0.9$</td>
<td>Very good fit</td>
</tr>
<tr>
<td>$0.8 &lt; r &lt; 0.9$</td>
<td>Good fit</td>
</tr>
<tr>
<td>$0.7 &lt; r &lt; 0.8$</td>
<td>Poor fit</td>
</tr>
<tr>
<td>$r &lt; 0.7$</td>
<td>Very poor fit</td>
</tr>
</tbody>
</table>

The PIC value for each locus was calculated on the basis of allele frequency (Anderson et al., 1993). $\text{PIC}_i = 1 - \sum_{j=1}^{n} P_{ij}^2$, where $P_{ij}$ is the frequency of $j^{th}$ allele for $i^{th}$ marker and summation extends over n alleles.

III. RESULTS AND DISCUSSION

3.1. Polymorphism as detected by ISSR analysis

All fifteen ISSR primers screened, in which nine ISSR primers amplified a total of 86 bands out of which 54 bands/alleles were polymorphic with an average of six bands per primer (Table 1). Out of 54 polymorphic bands, 50 bands were shared polymorphic, while four unique polymorphic bands were observed in four genotypes. i.e. one band in genotype G. Cot-12 by ISSR 2 (674 bp); one band in genotype BS-27 by UBC-807 (478 bp); one band in genotype G. Cot-18 by UBC-849 (1567 bp); and one band in genotype GJHV-460 by 844A (703 bp). From the data it was observed that 37.21 % monomorphic bands and 62.79 % of polymorphic bands were observed. UBC 807 primer produced the highest 12 bands (alleles), while ISSR 4 produced 11 bands; followed by ISSR 2, UBC 832 and 17899B each produced 10 bands. The primer 844A produced 9 bands and ISSR 1,
ISSR PRIMER 2 and UBC 849 each produced 8 bands. The average percentage polymorphism of 62.79% per primer was recorded in all the nine ISSR primers (Table 1).

Similar findings were also reported by Dongre et al. (2004) and Hussein et al. (2006) which support the present results. Dongre et al. (2004) examined 19 ISSR primers which generated 90 markers in which 12 ISSR primers were polymorphic which produced 49 markers and average percentage polymorphism was recorded 54.44% per primer. Hussein et al. (2006) used twelve ISSR primers to estimate the genetic relationships among twenty-one cotton accessions and produced 125 amplicons with 49.6% polymorphism.

In our study, the highest polymorphism was obtained with primer ISSR 2 (90%), followed by 844A (77.77%), while primers ISSR 1 and ISSR PRIMER 2 gave the lowest (25%) polymorphism (Fig. 4.4 and Table 4.4). Sharaf et al. (2009) studied ten ISSR primers which amplified 70 fragments, 62 of them were polymorphic (88.5%). The polymorphism was obtained with 844A (75%). This finding supported the results obtained in the present study.

The amplified ISSR fragments were in the range of 117 bp to 2571 bp. The largest fragment of 2517 bp and the smallest fragment of 117 bp were amplified by 17899B and ISSR 1, respectively. Similar findings were also reported by Dongre et al. (2004) and Noormohammadi et al. (2013). Dongre et al. (2004) worked with 19 ISSR primers which generated 90 markers in which 12 ISSR primers were polymorphic and produced 49 markers in cotton with size ranged between 100 bp and 1444 bp. Noormohammadi et al. (2013) observed that 17 of 20 homo-ISSR and hetero-ISSR primers produced 206 reproducible fragments. The sizes of fragments obtained were ranged 250 bp to 2600 bp.

The polymorphic information content (PIC) was calculated for each primer (Table 1). The polymorphic information content ranged between 0.8616 and 0.9090. The highest PIC value of 0.9090 was noticed for ISSR 4, while lowest PIC value of 0.8616 was noticed for UBC 832 with an average of 0.8841 per primer. ISSR primer index (IPI) ranged from 6.98 to 10.84 with an average of 8.46 per primer. The highest IPI value was obtained by UBC 807 and lowest was obtained by ISSR PRIMER 2 (Table 1).

3.2. Genetic relationship among cotton genotypes

Similarity index and cluster analysis was done by Jaccard’s coefficient and UPGMA using NTSYSpc-2.02i software, respectively. The dendrogram was generated by using Jaccard’s coefficient values (Fig. 2) to estimate the genetic similarity among cotton genotypes. Jaccard’s coefficient of similarity between 15 cotton genotypes ranged from 60% to 91.7%. Dongre et al.
(2004) reported that Jaccard’s coefficient of similarity between 25 cotton germplasms ranged from 60 % to 95 %, which supported the results obtained in the present study.

Fifteen cotton genotypes were grouped into two main clusters I and II with an average similarity of 66 % (Fig. 2). The cluster I consisted of fourteen genotypes and these were divided into two subclusters IA and IB. The subcluster IA was grouped into thirteen genotypes with shared similarity of 91.7 %. The subcluster IA was grouped into two sub subcluster IAi and IAii. The sub subcluster IAi was grouped into twelve genotypes viz., G. Cot-18, GBHV-148, GBHV-170, GJHV-503, BC-68-2, BS-27, 76-IH-20, H-1316, G. Cot-12, LRA-5166, GISV-254 and GIHV-95/145. GBHV-148 and GBHV-170 shared similarity of 90 %; BC-68-2 and BS-27 (87.2 %); LRA-5166 and GISV-254 (86.7 %); G. Cot-18 and GJHV-503 (83.8 %); GBHV-170 and H-1316 (81.3 %); G. Cot-18 and GISV-254 (72.3 %) and MR-786 and GJHV-460 (60 %). The sub subcluster IAii consisted of single genotype MR-786. The subcluster IB also consisted of single genotype BS-279. The cluster II consisted of single solitary genotype GJHV-460. The cluster analysis showed the highest (91.7 %) similarity between the genotypes GISV-254 and GTHV-95/145 and lowest (60 %) similarity between MR-786 and GJHV-460.

To test the goodness of fit of the clustering of ISSR data, matrixes of cophenetic values were also computed using the program COPH (Fig. 3). In the present study, the Mantel test statistic-Z was normalized and degree of goodness of fit for a cluster analysis (Matrix correlation: \( r = 0.8744 \)) was found to fall under the category of “good fit” as categorized by Rohlf (2000). A similar finding was also reported by Abdi et al. (2008). They estimated the ISSR based dendrogram, cophenet correlation using Jaccard’s similarity coefficient was as \( r = 0.84 \).
### Table 1: Size, number of amplified bands, per cent polymorphism and PIC obtained by ISSR primers

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>ISSR primers</th>
<th>Allele/Band Size (bp)</th>
<th>Total No. of allele (A)</th>
<th>Poly-morphic Bands (B)</th>
<th>Monomorphic bands</th>
<th>% Polymorphism (B/A)</th>
<th>PIC value</th>
<th>IPI (PIC×A)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S</td>
<td>U</td>
<td>Total Bands (T)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>ISSR 1</td>
<td>117-502</td>
<td>8</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>ISSR 2</td>
<td>240-1601</td>
<td>10</td>
<td>8</td>
<td>1</td>
<td>9</td>
<td>1</td>
<td>90</td>
</tr>
<tr>
<td>3</td>
<td>ISSR 4</td>
<td>182-778</td>
<td>11</td>
<td>8</td>
<td>0</td>
<td>8</td>
<td>3</td>
<td>72.72</td>
</tr>
<tr>
<td>4</td>
<td>ISSR PRIMER 2</td>
<td>160-1234</td>
<td>8</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>UBC 807</td>
<td>185-1968</td>
<td>12</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>58.3</td>
</tr>
<tr>
<td>6</td>
<td>UBC 832</td>
<td>204-1482</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>7</td>
<td>UBC 849</td>
<td>290-1602</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>75</td>
</tr>
<tr>
<td>8</td>
<td>844A</td>
<td>566-1936</td>
<td>9</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td>2</td>
<td>77.77</td>
</tr>
<tr>
<td>9</td>
<td>17899B</td>
<td>307-2571</td>
<td>10</td>
<td>7</td>
<td>1</td>
<td>8</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>86</strong></td>
<td><strong>50</strong></td>
<td><strong>4</strong></td>
<td><strong>54</strong></td>
<td><strong>32</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>AVERAGE</strong></td>
<td></td>
<td><strong>-</strong></td>
<td><strong>-</strong></td>
<td><strong>-</strong></td>
<td><strong>6</strong></td>
<td><strong>-</strong></td>
<td><strong>62.79</strong></td>
</tr>
</tbody>
</table>

S = Shared; U = Unique; T = Total Polymorphic Bands; PIC = Polymorphism information content; IPI = ISSR Primer Index
International Journal of Applied And Pure Science and Agriculture (IJAPSA)
Volume 01, Issue 01, [December - 2014]
Fig. 1: Properties of polymorphic and monomorphic bands amplified by the ISSR primers

Fig. 2: Dendrogram depicting the genetic relationship among 15 cotton genotypes based on ISSR markers
Fig. 3: Cophenetic values against Jaccard’s similarity coefficients from ISSR data of cotton genotypes
IV. Conclusions

The conclusions drawn from the present investigation are as under:

1. The formation of several subclusters within cluster I suggested the presence of moderate genetic diversity among the fifteen cotton genotypes studied. The geographical diversity was not associated with genetic diversity.

2. Nine ISSR primers selected in the present study gave 62.79 % polymorphism. The highest polymorphism percentage (90 %) was observed with primer ISSR 2. Therefore, this ISSR primer can be used further for diversity study in different cotton genotypes.

Bibliography


