Int J Pharm Bio Sci 2015 July; 6(3): (B) 1390 - 1394

Research Article **Biotechnology**

International Journal of Pharma and Bio Sciences **ISSN 0975-6299**

GENETIC ANALYSIS OF SNPS IN CANDIDATE GENES ASSOCIATED WITH DROUGHT TOLERANCE IN *GOSSYPIUM HERBACEUM*.

ASHISH KUMAR GUPTA

Research Scholar, I.T.M University, Gwalior, Madhya Pradesh, India.

ABSTRACT

Drought tolerance is an important physiological phenomenon for any agricultural crop in the depletion of ground water scenario. Recent case-control transcriptome and association studies have identified a network of genes underlying with the complex trait. We aimed to determine whether some of the genes confer drought tolerance in the Levant cotton (*Gossypium herbaceum* L.) population and thus to establish an association of SNPs in candidate genes with drought tolerance severity in Levant cotton with respect to Upland and Egyptian cotton population. This study analyzed leaf DNA isolated from 20 genotypes of *G. herbaceum* and 10 genotypes of allotetraploid species. A comparative in *silico* analysis of microarray and transcriptome data of *G. herbaceum* and *G. hirsutum* concluded few putative genes like blue-copper binding protein (*BCB*), UDP-Glucosyltransferase (*ATUGT85A3*), DNA binding/transcription factor (*CAPRICE*), senescence-associated genes (*ATLEA5,SAG21*), lipid transfer protein 4 (*LTP4*), DNAJ heat shock protein, trehalose-phosphatase (*TPS6, ATTPS6*), zinc finger (CCCH-type) family protein, dark inducible 6 (DIN6), Ethylene forming enzyme (*EFE*), acid phosphatase/protein serine/threonine phosphatase (*PAP10*), those played a significant role during drought tolerance. Out of these genes, six with SNPs (accession number yet to be submitted)–snp55, snp298, snp306, snp365, snp365, snp560 and snp574– were genotyped using Snapshot kit (ABI, USA) in high-throughput 96-capillary DNA analyzer 3730xl (ABI, USA) with *G. herbaceum*, *G. hirsutum*, *G. barbadense* collected from different major cotton research stations in India. Our results suggest that a sequence variant on snp574 and snp55 are two important loci that confer high cross-species drought tolerance.

KEYWORDS: Association mapping, cultivars, drought tolerance, *Gossypium herbaceum* L., Single Nucleotide Polymorphism (SNPs).

ASHISH KUMAR GUPTA
Research Scholar, I.T.M University, Gwalior, Madhya Pradesh, India.
E-Mail:ashishkumar4605@gmail.com

***Corresponding author**

INTRODUCTION

Cotton is the dominant source of natural textile fiber, edible seed oil as well as cotton seed meal. Cotton's strength, absorbency, and capacity to be washed and dyed also make it adaptable to textile industries[1]. Cotton is a soft, fluffy, staple fiber that grows in a boll around the seeds of the cotton plant[2,3.] It is a shrub native to tropical and subtropical regions around the world, including the Americas, Pakistan, India and Africa. *Gossypium* is a genus of 45-50 species of shrubs in the mallow family Malvaceae. The four commercial species are *G. hirsutum* (90% of world production), *G. barbadense* (8%), *G. arboreum* and *G. herbaceum* (together, 2%). Fiber growth goes through 4 major developmental stages, from initiation, through elongation and secondary cell wall synthesis to maturation. A single-nucleotide polymorphism (SNP, pronounced *snip*) is a DNA sequence variation occurring when a single nucleotide — A, T, C, or G — in the genome (or other shared sequence) differs between members of a species (or between paired chromosomes in an individual)[4]. Ten primers selected for this work SNP574, SNP560, SNP377, SNP365, SNP322, SNP306, SNP298, SNP55, SNP710, and SNP474 are taken from genome and from Internal transcribed spacer (ITS) region of nuclear region of the genome.

MATERIALS & METHODS

Green leaves were collected and identified on the basis of their leaf morphology and floral structure from germplasm bank maintained at National Botanical Research Institute, Lucknow (Table-1). These leaves were used for DNA isolation, PCR amplification and sequencing was done. The 24 selected genotypes of four species of cotton are used for DNA isolation with the help of C-TAB method according to Jena et al,2006[5] ,after completing the DNA isolation, check the quality & quantity were checked on gel and nanodrop respectively. A value of 1.8 is optimum for best DNA preparation. A value of the ratio below 1.8 indicates the presence of protein in the preparation and a value above 1.8 indicates that our sample has RNA contamination.

Primer Designing

Designing of forward, reverse, extension primer for SNP using web based software PRIMER 3 (http://fokker.wi.mit.edu/Primer3/) with criterion of primer size in range of 18 to 27, Primer tm in between 55 to 65 °C , GC content above 40%. The primer design for EST-SNP and genomic SNP were used to carry PCR reaction[6]. Various PCR condition and different composition of PCR component were tried to optimize the PCR to obtain desirable size of amplicon. After various combinations following PCR condition were used to amplify the SNP. The PCR product was analysed by Agarose gel electrophoresis the easiest and common way of separating and analysing DNA.The amplicon product was purified by gel extraction kit (Quigen) according to the manufactures instruction. Sequencing was done in ABI-3730XL capillary sequencer. Sequencing was performed directionally. SNAP Shot was used for SNP detection. For the fragment analysis of SNP Gene Mapper Software was used and sequencing analysis was done by the Clustal-W Software.

RESULTS & DISCUSSION

The SNPs discovered in silico from genomic expressed sequence tag (EST) sequences available in the data base, or through sequencing or resequencing of the candidate genes/PCR product /whole genome in more than one genotype[7]. However with the availability of the microarrays, SNP platform have been developed, which allow genotyping of thousands of markers in parallels. Besides SNPs, some other novel marker systems, including single feature polymorphism, diversity array technology and restriction site associated DNA markers, have also been developed, where array based assays have been utilized to provide for the desired ultra-high throughput and low cost. These microarray based markers are the markers of choice for the future and are

already being used for construction of high density maps, quantitative trait loci (QTL) mapping (including expression QTLs) and genetic diversity analysis with the limited expense in term of times and money[8,9] We briefly describe the characteristics of these array based marker systems and review that has already been done involving the development and use of these markers. We identified SNPs in the sequenced PCR amplified products of conserved regions of ESTs of *G. arboreum*, *G. herbaceum*, *G. hirsutum*, *and G. barbadense*. As the primers were based on conserved EST regions and the amplified product size was same as expected from primers amplification. Hence, it can be concluded that the primers amplified only the exon regions. Development of new SNPs by resquencing of PCR ampicons with or without pre-screening has been reported. In most plant species, SNPs have been detected by comparisons of two accessions as in maize and soybean[9]. Similarly in cotton, a PCR based direct DNA sequencing technique and fragment analysis work used to identify SNPS in different drought related genes[10].

In silico analysis with NBRI In-house microarray, transcriptome data and public domain data for selection of drought related genes

50 genes, playing a vital role during drought condition reported in the public domain, were chosen. To validate those genes, we compared the same with the transcriptome data as well as microarray data of cotton. Eventually, eighteen important genes playing an important role during drought were selected. Thus, eighteen genes were selected with a comparative data of public domain, microarray data and transcriptome data

Screening of eighteen genes for putative SNPs and Checking of SNP site with respect to amplicon size

All eighteen genes sequences were deployed to design primers for SNP site. All designed SNP sites were amplified from the genomic DNA with respective forward and reverse primer for the required product of specific genes. Out of eighteen, six amplified products were having the amplicon sizes of expected product sizes, ten amplified product having larger amplicon sizes with expected product sizes and two with no amplification. This indicates that six SNP sites of respective genes were having no intron regions. Thus, the six SNP sites can be validated with genomic DNA. Ten SNP sites were having larger amplicon size by comparison with our expected size which indicates that these SNP sites were having intron region that increase the amplicon size. Two SNP sites were not amplified because there may be error in primer designing or there may be possibility of the start of intron region in one of the primer sites.

Aligning sequencing data for 2 SNPs

Two SNP sites were selected for validating the SNPs in sequencing approach. The products of two SNP site for respective genotypes were sequenced and the sequences were aligned using Clustal-W software for targeted SNP. These two SNP were validated in 24 genotype .The SNP site [T/C] was validated having 'T' allele in 19 genotypes of *G. herbaceum*, *G. arboreum* and *G. hirsurtum* and 'C' allele in 5 genotypes of *G. barbadense* and *G. hirsutum*. In other SNP site [G/A], 'G' allele was found in few genotypes where 'A'allele was found in many genotypes of cotton.

SNAP-Shot analysis for 4 SNPs

PCR product used for single base extension using extension primer with SNP shot kit. All PCR products of respective genotypes were extended a single base and analysed in Genemapper software for the extended base. Detection of allelic variation is based on fluorescence variation of nucleotide at target SNP site (SNAP Shot)[11].

- For SNP 574 [A/G], among three *G. arboreum,* two were having 'A'allele and one having 'G' allele .In *G. barbadensel* and *G. hirsutum*, all three genotypes were homozygotes for allele 'G' .Five genotypes of *G. herbaceum* were found be homozygous for allele [G] while four were homozygous for [A] and other 2 genotypes were found to be heterozygous for both alleles.(Table-2)

• For SNP 306 [AG/GA], all genotypes were having homozygous for allele 'AG'

• For SNP 55 [T/C], all genotypes of G. arboreum, G. barbadense and G. hirsutum were homozygous for allele 'T' while in G .herbaceum, six were heterozygous for both alles and four genotypes were homozygous for 'T', one for homozygous for 'C'.(Table-3)

• For SNP 298 [A/G], all genotypes were found to be homozygous for 'G'

Table-1

Twenty-four studied genotypes of four species of cotton

Sr. No	**Species**	**Taxonomical Details Of Genotype**
1	*Gossypium hirsutum*	*G.hirsutum-771*
2		*G. hirsutum-AS2*
3		*G. hirsutum-LRA*
4	*Gossypium barbadense*	*G. barbadense-B1*
5		*G. barbadense-B2*
6		*G. barbadense-SUVIN*
7	*Gossypium herbaceum*	*G. herbaceum-VAGAD*
8		*G. herbaceum-RAHS14*
9		*G. herbaceum-IPS187*
10		*G. herbaceum-RAHS127*
11		*G. herbaceum-GUJCOT21*
12		*G. herbaceum-221-547*
13		*G.herbaceum-H-17*
14		*G. herbaceum-RAHS132*
15		*G.herbaceum-AH41*
16		*G. herbaceum-GH-18-2LC*
17		*G.herbaceum-AK235*
18		*G. herbaceum-221-568*
19		*G.herbaceum-AH-7-GP*
20		*G. herbaceum-DB312*
21		*G.herbaceum-221-573*
22	*Gossypium arboreum*	*G. arboreum-551*
23		*G.arboreum-575*
24		*G. arboreum-579*

Table-2

Detected SNP 574 (gene) of 24 genotypes of cotton after Snap-Shot Analysis

Sample Name	**Marker**	**Allele 1**	**Allele 2**	**Inferences**
*G.arboreum*_551	SNP_574		A	Homozygous for 'A'
*G.arboreum*_575	SNP_574	G		Homozygous for 'G'
*G.arboreum*_579	SNP_574		A	Homozygous for 'A'
*G.barbadense*_B1	SNP_574	G		Homozygous for 'G'
*G.barbadense*_B2	SNP_574	G		Homozygous for 'G'
*G.barbadense*_SUVIN	SNP_574	G		Homozygous for 'G'
*G.herbaceum*_221-547	SNP_574	G		Homozygous for 'G'
*G.herbaceum*_221-568	SNP_574	G		Homozygous for 'G'
*G.herbaceum*_221-573	SNP_574	G		Homozygous for 'G'
*G.herbaceum*_AH7GP	SNP_574		A	Homozygous for 'A'
*G.herbaceum*_DB312	SNP_574	G	A	Heterozygous for 'G/A'
*G.herbaceum*_GUJCOT 21	SNP_574	G		Homozygous for 'G'
*G.herbaceum*_H17	SNP_574		A	Homozygous for 'A'
*G.herbaceum*_IPS187	SNP_574	G	A	Heterozygous for 'G/A'
*G.herbaceum*_RAHS127	SNP_574		A	Homozygous for 'A'
*G.herbaceum*_RAHS14	SNP_574		A	Homozygous for 'A'
*G.herbaceum*_VAGAD	SNP_574	G		Homozygous for 'G'
*G.hirsutum*_771	SNP_574	G		Homozygous for 'G'
*G.hirsutum*_AS2	SNP_574	G		Homozygous for 'G'
*G.hirsutum*_LRA	SNP_574	G		Homozygous for 'G'

Table-3
Detected of SNP 55 (gene) of 24 genotypes of cotton after Snap-Shot Analysis

Sample Name	Marker	Allele 1	Allele 2	Inferences
*G.arboreum*_551	SNP_55	T		Homozygous for 'T'
*G.arboreum*_575	SNP_55	T		Homozygous for 'T'
*G.arboreum*_579	SNP_55	T		Homozygous for 'T'
*G.barbadense*_B1	SNP_55	T		Homozygous for 'T'
*G.barbadense*_B2	SNP_55	T		Homozygous for 'T'
*G.barbadense*_SUVIN	SNP_55	T		Homozygous for 'T'
*G.herbaceum*_221-547	SNP_55		C	Homozygous for 'C'
*G.herbaceum*_221-568	SNP_55	T	C	Heterozygous for 'T/C'
*G.herbaceum*_221-573	SNP_55	T		Homozygous for 'T'
*G.herbaceum*_AH7GP	SNP_55	T	C	Heterozygous for 'T/C'
*G.herbaceum*_DB312	SNP_55	T		Homozygous for 'T'
*G.herbaceum*_GUJCOT 21	SNP_55	T		Homozygous for 'T'
*G.herbaceum*_H17	SNP_55	T	C	Heterozygous for 'T/C'
*G.herbaceum*_IPS187	SNP_55	T	C	Heterozygous for 'T/C'
*G.herbaceum*_RAHS127	SNP_55	T		Homozygous for 'T'
*G.herbaceum*_RAHS14	SNP_55	T	C	Heterozygous for 'T/C'
*G.herbaceum*_VAGAD	SNP_55	T	C	Heterozygous for 'T/C'
*G.hirsutum*_771	SNP_55	T		Homozygous for 'T'
*G.hirsutum*_AS2	SNP_55	T		Homozygous for 'T'
*G.hirsutum*_LRA	SNP_55	T		Homozygous for 'T'

REFERENCES

1. Saleem, M.F., M.F. Bilal, M. Awais, M.Q. Shahid and S.A. Anjum. Effect of nitrogen on seed cotton yield and fiber qualities of cotton (Gossypium hirsutum L.) cultivars. The Journal of Animal and Plant Sciences, 20(1): 23-27. (2010)
2. Smith, C.W. Cotton (Gossypium hirsutum, L.). In: Crop Production, Evolution, History and Technology, C.W. Smith, (Ed.), John Wiley and Sons, Inc., New York, USA, pp: 287-349. 4(1995)
3. Mehasen, S.A.S., S.G. Gebaly and O.A. Seoudi,. Effectiveness of organic and inorganic fertilization in presence of some growth regulators on productivity and quality of Egyptian cotton. Asian Journal of Biological Sciences, 5: 171-182 (2012)
4. AYEH, Kwadwo Owusu. Expressed sequence tags (ESTs) and single nucleotide polymorphisms (SNPs): Emerging molecular marker tools for improving agronomic traits in plant biotechnology. *African Journal of Biotechnology* vol. 7, no. 4, p. 331-341. (February 2008)
5. Jena et al.C-TAB method for DNA extraction, Rapid and cheaper protocol for DNA extraction for molecular approaches .http://shodhganga.inflibnet.ac.in/.(2006)
6. Payeh K. Expressed sequence tags (ESTs) and single nucleotide polymorphisms (SNPs): Emerging molecular marker tools for improving agronomic traits in plant biotechnology. *African Journal of Biotechnology*,7:4,331-341 (2008)
7. 7. Barbazuk W., Patrick S. SNP mining from maize 454 EST sequences. *CSH Protocols, 2007:7, (2007)*.
8. Choi Ik-Young., Hyten D L A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genetics*, 176:1,685-696 (2007).
9. Frelichowski JE., Palmer MB., Main D. Cotton genome mapping with new microsatellites from Acala 'Maxxa' BAC-ends. *Molecular Genetics and Genomics*; 275:479–491. doi: 10.1007/s00438-006-0106-z.*(* 2006)
10. Michael M. Neff, Joseph D. Neff, Joanne Chory and Alan E. Pepper dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana*genetics,The plant journal 14:387-392 (1998)
11. Sutanto A., Hermanto C., Sukma D., dan Sudarsono. Pengembangan Marka SNAP Berbasis, Resistance Gene Analogue Pada Tanaman Pisang (Musa spp.) (Development of SNAP Marker Based On Resistance Gene Analogue Genomic Sequences in Banana (Musa spp.), Jurnal Hortikultura, 23(4):300-309, 2013.