***Pak. J. Bot*., 41(1): 177-183, 2009.**

SINGLE NUCLEOTIDE POLYMORPHISM ANALYSIS OF MT-SHSP GENE OF *GOSSYPIUM ARBOREUM* AND ITS RELATIONSHIP WITH OTHER DIPLOID COTTON GENOMES, *G. HIRSUTUM* AND *ARABIDOPSIS THALIANA*

TAYYABA SHAHEEN, MUHAMMAD ASIF, YUSUF ZAFAR AND [*]MEHBOOB-UR-RAHMAN

National Institute for Biotechnology & Genetic Engineering (NIBGE)
PO Box 577 Jhang Road Faisalabad, Pakistan.

Abstract

Single nucleotide polymorphisms (SNPs) are the most popular DNA markers because of their abundance and consistency in the genomes. House keeping genes are conserved in nature across the genomes of different organisms. Study of variations in these conserved genes can reveal the hidden facts of evolution which can not be excavated with conventional DNA marker systems. In the present study, mitochondrial small heat shock protein gene (MT-sHSP) has been explored to find nucleotide variations within *Gossypium arboreum*, with the other diploid *Gossypium* genomes, *G. hirsutum* and also with *Arabidopsis thaliana*. A conserved region spanning 300bp was amplified and sequenced from two *G. arboreum* (A_2) genotypes, species of other diploid genomes belonging to A_1, C_1, E_1, D_4, D_6, D_9 genomes and tetraploid species *G. hirsutum* (AD). Sequence of the gene of *A. thaliana was* retrieved from Genbank. These sequences were aligned. Within *G. arboreum* genome one Indel was found while, 'C' genome showed the least nucleotide variations with the 'A' genome species (*G. arboreum)* as compared to other genomes. D genome species and *G. hirsutum* were closely related with each other. *A. thaliana* was most distantly related with other genomes. The present studies reveal that SNP markers could be identified in conserved regions where conventional markers are of little or no use. This study will lead to the better understanding of *G. arboreum* evolution and understanding how these variations can be utilized for the improvement of cotton genome.

Introduction

Single nucleotide polymorphisms (SNPs) are a single base change or small insertions and deletions in homologous DNA fragments. In human genome sequencing 10 to 30 million SNPs were found and were the most abundant source of polymorphisms (Collins *et al*., 1998) present both in coding and noncoding regions (Aerts *et al*., 2002). As a marker SNPs are preferred over other marker systems because they are more frequent, codominant in nature and are sometimes associated with morphological changes (Lindblad-Toh *et al*., 2000). Genomes of higher plants like barley (Kanazin *et al.*, 2002), maize (Tenaillon *et al*., 2001), soybean (Choi *et al*., 2007), sugar beet (Schneider *et al*., 2001), sunflower (Lai *et al*., 2005), rye (Varshney *et al*., 2007) and cotton (Lu *et al.*, 2005; Shaheen *et al*., 2006; Ahmad *et al*., 2007) have also been surveyed for SNPs discovery and characterization. Because SNPs are highly polymorphic, every gene should contain a few SNPs even among strains (Cho *et al*., 1999). MT-sHSP gene is an important gene which helps to tolerate heat shock. The MT-sHSP protects NADH: ubiquinone oxidoreductase of the electron transport chain during heat stress in plants (Herrman *et al*., 1994).

---

[*]Email: mehboob_pbd@yahoo.com
Phone: +92 41 2554378 Fax: +92 41 2651472

Table 1. Origin of cotton species used in study (Guo *et al*., 2006).

S. No.	Species name	Genome	Distribution
1.	*G. arboreum*	(A_2)	Old world
2.	*G. herbaceum africanum*	(A_1)	Africa
3.	*G. aridum*	(D_4)	Mexico
4.	*G. laxum*	(D_9)	Mexico
5.	*G. stocksii*	(E_1)	Arabian peninsula
6.	*G. gossypoides*	(D_6)	Mexico
7.	*G. sturtianum*	(C_1).	Australia
8.	*G. hirsutum*	(AD)	New world

SNP markers, combined with QTL data for phenotypic character, can provide a new system of breeding i.e., gene-mediated breeding instead of marker-assisted selection (Lange & Whittacker, 2001). Genetic improvement of cotton fiber and agricultural productivity will be enhanced by the availability of rapidly developing genetic resources and tools, including high-density genetic maps (Lacape *et al*., 2005). Cotton being economically important crop (Rahman *et al*., 2002; 2005) is being explored to understand its genome (Chen *et al*., 2007). In understanding cotton genome major problem is the huge genome size and occurrence of polyploidy in cotton. Polyploid genomes are more difficult to analyze for SNPs than diploids. The ratio of SNP alleles varies in polyploid genomes (Adams *et al*., 2003).

In this study we have conducted an experiment to analyze frequency of SNPs in *G. arboreum* and its comparison with other diploid genomes of cotton *G. hirsutum* and *A. thaliana* as well.

Material and Methods

Isolation of total genomic DNA: Total genomic DNA was isolated from two local cotton varieties of G. *arboreum* (A_2) (var. Ravi and Entry-17) and diploid cotton species *G. herbaceum africanum* (A_1), *G. sturtianum* (C), *G. aridum* (D_4), *G. gossypoides* (D_6), *G. laxum* (D_9), *G. stocksii* (E_1) and *G. hirsutum* (AD) (Table 1) by a method used by Iqbal *et al*., (1997).

Primer designing and PCR amplification: Gene sequences were used to search homology in NCBI using Blast N tool to find the conserved regions. Primers were designed using primer 3 software on the basis of regions spanning the conserved regions. Polymerase chain reaction (PCR) was performed in a total volume of 20µl, using 2.5µl (15ng/µl) of cotton DNA, 10 x PCR buffer without $MgCl_2$ (10mM Tris-HCl, 50mM KCl, PH 8.3), 3mM $MgCl_2$, 0.1mM each of dATP, dGTP, dCTP and dTTP and 0.5 units of *Taq* DNA polymerase, 0.15 mM of each primer. *Taq* DNA polymerase together with 10 x PCR buffer, $MgCl_2$ and dNTPs were from MBI Fermentas. Polymerase chain reaction consisted of denaturation at 94°C for 1 min., 35 cycles of 94°C for 30 sec., 50°C for 30 sec., 72°C extension for 1 min., and final extension at 72°C for 10 min. PCR products were resolved on 2% agarose gel.

Sequencing of PCR product: Sequencing of PCR products was done on ABI automated DNA sequencer from Macrogen. Sequences were edited manually. Sequence of *A. thaliana* was retrieved from Genbank.

Single nucleotide polymorphism detection: DNA sequences obtained were aligned to detect nucleotide polymorphism using Megalign DNA star. Phylogenetic analysis was done with Megalign after alignment with ClutalW.

Results

Types and distribution of SNPs: Conserved region of MT-sHSP gene spanned from nucleotide position 138 to 440 in *G. arboreum*. On alignment of sequences obtained from *G. arboreum, G. herbaceum africanum, G. sturtianum, G. aridum, G. gossypoides, G. laxum, G. stocksii and G. hirsutum* along with sequence obtained from GeneBank for *Arabidopsis thaliana* SNPs were detected at 21 positions while Indels at 10 positions. Base substitutions include T/C, T/G, A/G, T/A, C/A, C/G and T/C/G. Nucleotide variations with *A. thaliana* were at 101 positions. Maximum number of base substitutions was T/C which are transitions (Fig. 1).

Phylogenetic analysis of diploid genomes: Phylogenetic relations portrayed by the SNPs study resulted into two clusters. One cluster included A genome species *G. arboreum* (A_2), *G. herbaceum* (A_1) and C genome species *G. sturtianum* (C_1). Both varieties of *G. arboreum* were closely related with each other. While second cluster included D genome species *G. aridum* (D_4), *G. gossypoides* (D_6), *G. laxum* (D_9), *G. stocksii* (E_1) and AD genome species *G. hirsutum* (AD). *A. thaliana* was at 20% distance from other genomes (Fig. 2).

Frequency of SNPs and Indels: In a genomic region of 300 bases, 21 SNPs were detected which indicate occurrence of 1SNP per 14 b and 10 indels were found which indicate occurrence of one indel per 30 bases.

Discussion

In this study we observed 21 SNPs and 10 indels in 300 bases amplified from conserved regions of MT-sHSP gene among diploid genomes of cotton including *G. arboreum* (A_2), *G. herbaceum* (A_1), *G. sturtianum* (C), G. *aridum* (D_4), *G. gossypoides* (D_6), *G. laxum* (D_9), *G. stocksii* (E_1) and tetraploid species *G. hirsutum* (AD). Results depict frequency of occurrence of one SNPs per 14b and one Indel per 30b. Maximum type of variations are transitions (T/C), it is comparable with the previous results with data published for human genome and sugar beet where transitions were more frequent than transversions. (Schneider *et al*., 2001). The frequency of SNPs and Indels observed in this study is higher as compared to intraspecific frequency of occurrence of SNPs as in cotton 1 SNP/500b (Lu *et al*., 2005), 1 SNP/3.3 kb and 1 Indel /6.6 kb in *Arabidopsis* (Jander *et al*., 2002), 1 SNP/130 bp in sugar beet (Schneider *et al*., 2001), 1 SNP/60.8 bp alongwith 1 Indel/126 bp in maize (Ching *et al*., 2002) and 1 SNP/170 bp in rice (Yu *et al*., 2002). A high degree of interspecific occurrence of SNPs has been observed in earlier reports as compared to intraspecific variations in genomes (Shaheen *et al*., 2006; Ahmad *et al*., 2007).

High frequency of occurrence of SNPs makes them potent markers to develop a dense genetic map of cotton. Molecular markers being used in cotton like RAPD, SSRs, AFLP, and SCARs have their own limitations (Ayeh *et al*., 2008). Molecular markers reveal a low degree of polymorphisms intraspecifically (Zhang *et al*., 2008).

```
    - - - - - - - - - - - - - - - T T T A - - C C T T C A C T A A C A A - T - - A A A C T T G - T A C A Majority
                                10                      20                        30
7   - - - - - - - - - - - T C C A T T T A - - C C T T C A C T A A C A A - T - - A A A C T T G - T A C A A2 (G. arboreum
138 C C T T A T T C A C A . . . . . . . . - - . . . . . . . . . . . . . - . - - . . . . . . . - . . . . A2 (G. arboreum
2   - - - - - - - - - - - - . N . . . . . - - . . . . . . . . . . . . . - . - - . . . . . . . - . . . . A1(G. herbaceum
4   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - . . - . - - . . . N . . . - . . . . C1 (G. sturtian
2   - - - - - - - - - - - - - - - - - - - - - - . C . T . . . . . . . . - . - - . . . . . . . - . . . . D4 (G. aridum)
2   - - - - - - - - - - - - - - - - - - - - - - - - . T . . . . . . . . - . - - . . . . . . . - . . . . D6 (G. gossypoi
1   - - - - - - - - - - - - - - - - - - - - - . . . . . . . . . . . . . - . - - . . . . . . . G . . . . D9 (G. laxum)
11  - - - - - - - - - - - - - - - . . . . - - . . . . . . . G . . . . . G . - - . . . . . . A - C . . . E1 (G. stocksii
149 C C T C A T T C A C A . . . . . . . . - - . . . . . . . . . . . . . - C - - . . . . . . . - . . . . AD (G. hirsutum
149 A T T C A C T C A C T . A . T . C C . T T . . A . . T . A . . . . . T C C A . . . T A A A A C . . T Arabidopsis

    C T A C T C A A C G G T - G A C T T G A - T A C A C A T C C T T C C - - T C T C C T C C - T C C T T Majority
    40                  50                      60                  70                        80
40  C T A C T C A A C G G T - G A C T T G A - T A C A C A T C C T T C C - - T C T C C T C C - T C C T T A2 (G. arboreum
182 . . . . . . . . . . . . - . . . . . . . - . . . . . . . . . . . . . - - . . . . . . . . - . . . . . A2 (G. arboreum
34  . . . . . . . . . . . . - . . . . . . . - . . . . . . . . . . . . . - - . . . . . . . . - . . . . . A1(G. herbaceum
18  . . . . . . N . . . . A - . . . . . . . - . . . . . . . . . . . . . - - . . . . . . . . A . N . . . C1 (G. sturtian
26  . . . . . . . . . . . . - . . . . . . . - . . . . . . . . . . . . . - - . . . . A . . . - . . . . . D4 (G. aridum)
24  . . . . . . . . . . . . - . . . . . . . - . . . . . . . . . . . . . - - . . . . A . . . - . . . . . D6 (G. gossypoi
27  . . . . . . . . . . . . - . . . . . . . - . . . . . . . . . . . . . - - . . . . A . . . - . . . . . D9 (G. laxum)
41  . . . . . . . . . . . N T . . . . . . . A . . . . . . . . . . . T . C T . . . . . . . . N . . . . . E1 (G. stocksii
193 . . . . . . . . . . . . - . . . . . . . - . . . . . . . . . . . . . - - . . . . A . . . - . . . . . AD (G. hirsutum
199 . . . G . . . . . . T . T - . T G . . . - C G A . T . . T G . . A . G - - . . . A . . . - - T . . . Arabidopsis

    - C A C T T T - A G G C A C A A C C A C - T T T C A A C A C C C C A T T T T T C A T C T C A G C C T Majority
                90                    100                   110                 120                 130
85  - C A C T T T - A G G C A C A A C C A C - T T T C A A C A C T C C A T T T T T C A T C T C A G C C T A2 (G. arboreum
227 - . . . . . . - . . . . . . . . . . . . - . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A2 (G. arboreum
79  - . . . . . . - . . . . . . . . . . . . C . . . . . . . . . C . . . . . . . . . . . . . . . . . . . A1(G. herbaceum
64  - . . . . . . - . . . . . G . . . . . . - . . . . . . . . . G . . . . . . . . . . . . . . . . . . . C1 (G. sturtian
71  - . . . . . . - . . . N . . . . . . . . N . . . . . . . . . C . . . . . . . . . . . . . . . . . . . D4 (G. aridum)
69  - . . . . . . - . . . . . . . . . . . . - . . . . . . . . . C . . . . . . . . . . . . . . . . . . . D6 (G. gossypoi
72  - . . . . . . - . . . . . . . . . . . . - . . . . . . . . . C . . . . . . . . . . . . . . . . . . . D9 (G. laxum)
91  T . . . N . . T . . . . . . . . . . . . N . . . . . . . . . C . . . . . . . . . . . . . . . . . . . E1 (G. stocksii
238 - . . . . . . - . . . . . . . . . . . . - . . . . . . . . . C . . . . . . . . . . . . . . . . . . . AD (G. hirsutum
243 T A . T C . . T G . A - . T C . . . . . - . . . . . . . . . A . . . . . C . . . . . T . . C . . . . Arabidopsis

    T T A T C T C A T C C A C C T T A T A C A T G G T G G G G G G - T A G G T C A A G C C T G C T C G A Majority
                    140                 150                 160                   170                 180
132 T T A T C T C A T C C A C C T T A T A C A T G G T G G G G G G - T A G G T C A A G C C T G C T C G A A2 (G. arboreum
274 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . G . . . . . . . . . . . . . . . . . . A2 (G. arboreum
127 . . . . . . . . . . . . . . . . . C . . . . . . . . . . . . . - . . . . . . . . . . . . . . . . . . A1(G. herbaceum
111 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . - . . . . . . . . . . . . . . . . . . C1 (G. sturtian
119 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . - . . . . . . . . . . . . . . . . . . D4 (G. aridum)
116 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . - . . . . . . . . . . . . . . . . . . D6 (G. gossypoi
119 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . - . . . . . . . . . . . . . . . . . . D9 (G. laxum)
141 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . - . . . . . . . . . . . . . . . . . . E1 (G. stocksii
285 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . - . . . . . . . . . . . . . . . . . . AD (G. hirsutum
291 . A . . . . . . . . A G T . . . G . . T . C T T . C T C C . . - . . A C . . . . T . . . A . . . . T Arabidopsis

    G T A C C G T C T C C C T C C C C C C T C T T C T T C T T C T T C T T C - - - A T C C T C C C A A T Majority
                      190                 200                 210                       220
181 G T A C C G T A T C C T T C C C C C C T C T T C T T C T T C T T C T T C - - - G T C C T C C C A A T A2 (G. arboreum
324 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . - - - . . . . . . . . . . . A2 (G. arboreum
176 . . . . . . . . . . . C . . . . . . . . . . . . . . . . . . . . . . . . - - - . . N . . . . . . . . A1(G. herbaceum
160 . . . N N . . . . . . C . . . . . . . . N . . . . . . N . . . . . . . N G A C A - - - . . . . . . . C1 (G. sturtian
168 . . . . . . . C . . . C . . . . . . . . . . . . . . . . . . . . . . . . T T C A C . . . . . . . . . D4 (G. aridum)
165 . . . . . . . C . . . C . . . . . . . . . . . . . . . . . . . . . . . . - - - A . . N . . . N . . . D6 (G. gossypoi
168 . . . . . . . C . . . C . . . . . . . . . . . . . . . . . . . . . . . . - - - A C . . . . . . . . . D9 (G. laxum)
190 . . . . . . . C . . . C . . . . . . . . . . . . . . . . . . . . . . . - - - C A . . . . . . . . . . E1 (G. stocksii
334 . . . . . . . C . . . C . . . . . . . . . . . . . . . . . . . . . - - - - - - A C . . . . . . . . . AD (G. hirsutum
340 A A . . . - - - . . . G . . . A T . T C . A G A A A . A . . . . . . - - - - - - C . . . . . - - - . Arabidopsis

    C T T T A C C C C C T T C T C C T C G G A T G A T G A G T G T G T T C T G T T C C A C - - - - - - - Majority
        230                 240                 250                 260
228 C T T T A C C C C C T T C T C C T C G G A T A A T G A G T G T G T T N - - - - - - - - - - - - - - - A2 (G. arboreum
371 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . C T G T T C C A C C G A G A T C A2 (G. arboreum
223 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . C T G T T C C A N - - - - - - - A1(G. herbaceum
207 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . C T G T T C C A N - - - - - - - C1 (G. sturtian
218 . . . . . . . . . . . . . . . . . . . . . . G . . . . . . . . . . . C T G T T C N A N - - - - - - - D4 (G. aridum)
212 N A . . . N T A . . . . . . T . . .                                                                  D6 (G. gossypoi
215 . . . . . . . . . . . . . . . . . . . . . . G . . . . . . . . . . . C T G T T C C A C - - - - - - - D9 (G. laxum)
237 . . . . . . . . . . . . . . . . . . . . . . G . . . . . . . . . . . C T G T T C - - - - - - - - - - E1 (G. stocksii
378 . . . . . . . . . . . . . . . . . . . . . . G . . . . . . . . . . . C T G T T C C A C C G A G A T C AD (G. hirsutum
378 . . G . T T . . . . . . . . . . . . T A . . C . C C . A . . . . . . C T G T T C C A A A G C C A A T Arabidopsis
```

Fig. 1. Alignment of cotton genomes and *A. thaliana* to detect SNPs and Indels.
Similar nucleotides are presented as dotted lines and nucleotides which differ are shown.

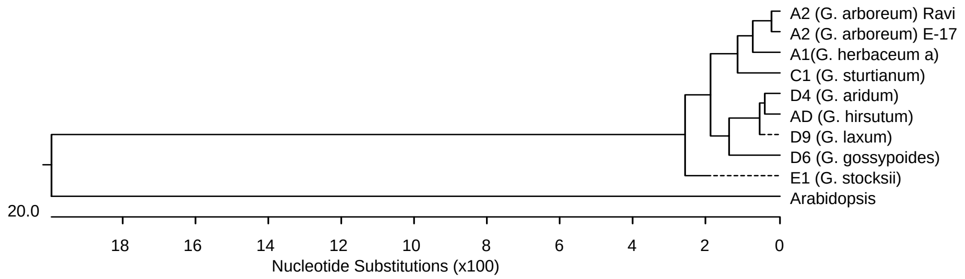

Fig. 2. Phylogenetic assessment of cotton genomes and *Arabidopsis thaliana.*

Results portrayed with phylogenetic assessment are comparable with the results obtained with other DNA markers (Guo *et al*., 2006; Wu *et al*., 2007). Both varieties of *G. arboreum* are very closely related and *G. herbaceum* is also clustered with them, C genome species is also in close proximity. Clustering of A genome species into one cluster and D genome species into separate cluster is comparable with previous studies (Wu *et al*., 2007). Parental lineages of AD genome belong to A and D genomes (Guo *et al*., 2005), however AD genome species has shown more affiliation with D genome species in phylogenetic analysis. Grouping of C genome species with A genome reflect its more genetic affiliation with A genome as compared to D genome. A high degree of similarity (80%) between *Arabidopsis* and cotton strengthen the concept of same ancestry of *Arabidopsis* and cotton which has been proposed in earlier studies (Rong *et al*., 2005).

In conclusion SNPs markers are more frequent in nature and comparable with other potent markers. They can be very helpful in making dense genetic maps and understanding evolutionary mechanisms.

Acknowledgements

We are thankful to Central Cotton Research Institute (CCRI), Multan, Pakistan for providing leaf samples of different species of cotton. We are also thankful to Higher Education Commission Pakistan for providing funds for the present studies through the HEC Project "Finding Single Nucleotide Polymorphisms in cotton genome." under Presidential Young Innovative Programme and indigenous PhD fellowship scheme.

References

Adams, K.L., R. Cronn, R. Percifield and J.F. Wendel. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc. Natl. Acad. Sci., USA,* 100: 4649-4654.

Aerts, J., Y. Wetzels, N. Cohen and J. Aerssens. 2002. Data mining of public SNP database for the selection of intragenic SNPs. *Hum Mutat*., 20: 162-173.

Ahmad, S., M. Ashraf, T. Zhang, N. Islam, T. Shaheen and M. Rahman. 2007. Identifying genetic variation in *Gossypium* L., based on single nucleotide polymorphism. *Pak. J. Bot.,* 39: 1245-1250.

Ayeh, K.O. 2008. Expressed sequence tags (ESTs) and single nucleotide polymorphisms (SNPs): Emerging molecular marker tools for improving agronomic traits in plant biotechnology. *Af. J. Biotech*., 7(4): 331-341.

Chen, Z.J., B.E. Scheffler, E. Dennis, B.A. Triplett, T. Zhang, W. Guo, X. Chen, D.M. Stelly, P.D. Rabinowicz and C.D. Town. 2007. Toward sequencing cotton (*Gossypium*) genomes. *Plant Physiol*., 145(4): 1303-1310.

Ching, A., K.S. Caldwell, M. Jung, M. Dolan, O.S. Smith, S. Tingey, M. Morgante and A.J. Rafalski. 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet*., 3: 19.

Cho, R.J., M. Mindrinos, D.R. Richards, R.J. Sapolsky, M. Anderson, E. Drenkard, J. Dewdney, T.L. Reuber, M. Stammers, N. Federspiel, A. Theologies, W.H. Yang, E. Hubbel, M. Au, E.Y. Chung, D. Lshkary, B. Lemieux, C. Dean, R.J. Lipshutz, F.M. Ausubel, R.W. Davis and P.J. Oefner. 1999. Genome wide mapping with biallelic markers in *Arabidopsis thaliana*. *Nature Genet*., 23: 203-207.

Choi, I.Y., D.L. Hyten, L.K. Matukumalli, Q. Song, J.M. Chaky, C.V. Quigley, K. Chase, K.G. Lark, R.S. Reiter, M.S. Yoon, E.Y. Hwang, S.I. Yi, N.D. Young, R.C. Shoemaker, C.P. van Tassell, J.E. Specht and P.B. Cregan. 2007. A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genetics*, 176: 685-696.

Collin, F.S., L.D. Brooks and A. Charkravarti. 1998. A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res*., 8: 1229-1231.

Guo, W., D. Fang, W. Yu and T. Zhang. 2005. Sequence Divergence of microsatellites and phylogeny analysis in tetraploid cotton species and their putative diploid ancestors. *J. Integr. Plant Biol*., Formerly *Acta Botanica Sinica.*, 47(12): 1418-1430.

Guo, W., W. Wang, B. Zhou and T. Zhang. 2006. Cross species transferability of *G. arboreum*-derived EST-SSRs in the diploid species of *Gossypium*. *Theor. Appl. Genet*., 112: 1573-1581.

Herrmarm, J.M., R.A. Stuart, E.A. Craig and W. Neupert. 1994. Mitochondrial heat shock protein 70, a molecular chaperone for proteins encoded by mitochondrial DNA. *J. Cell. Bio.*, 127(4): 893-902.

Iqbal, M.J., N. Aziz, N.A. Saeed, Y. Zafar and K.A. Mailk. 1997. Genetic diversity of some elite cotton varieties by RAPD analysis. *Theor. Appl. Genet*., 94: 139-144.

Jander, G., S.R. Norris, S.D. Rounsley, D.F. Bus, I.M. Levin and R.L. Last. 2002. *Arabidopsis* map-based cloning in the postgenome era. *Plant Physiol*., 129: 440-450.

Kanazin, V., H. Talbert, D. See, P. DeCamp, E. Nevo and D. Blake. 2002 . Discovery and assay of single nucleotide polymorphism in barley (*Hordeum vulgare*). *Plant Mol. Biol*., 48: 529-537.

Lacape, J.M. and T.B. Nguyen. 2005. Mapping quantitative trait loci associated with leaf and stem pubescence in cotton. *Heredity*, 96: 441-444.

Lai, Z., K. Livingstone, Y. Zou, S.A. Church, S.J. Knapp, J. Andrews and L.H. Rieseberg. 2005. Identification and mapping of SNPs from ESTs in sunflower. *Theor. Appl. Genet*., 111: 1532-1544.

Lange, C. and J.C. Wittacher. 2001. On prediction of genetic values in marker-assisted selection. *Genetics*, 159: 1375-1381.

Lindblad-Toh, K., E. Winchester, M.J. Daly, D.G. Wang, J.N. Hirschhorn, J.P. Laviolette, K. Ardlie, D.E. Reich, E. Robinson, P. Sklar, N. Shah, D. Homas, J.B. Fan, T. Gingeras, J. Warrington, N. Patil, T.J. Hudson and E.S. Lander. 2000. Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nature Genet.*, 24: 381-386.

Lu, Y., J. Curtiss, J. Zhang, R.G. Percy and R.G. Cantrell. 2005. Discovery of single nucleotide polymorphisms in selected fiber genes in cultivated tetraploid cotton. Beltwide Cotton Conference, *National Cotton Council of America, Memphis*, TN.

Rahman, M., M. Asif, I. Ullah, K.A. Malik and Y. Zafar. 2005. Overview of cotton genomic studies in Pakistan. *Plant & Animal Genome Conference XIII. San Diego, CA. USA.*

Rahman, M., D. Hussain and Y. Zafar. 2002. Estimation of genetic divergence among elite cotton cultivars—Genotypes by DNA fingerprinting technology. *Crop Sci.*, 42: 2137-2144.

Rong, J., J.E. Bowers, S.R. Schulze, V.N. Waghmare, C.J. Rogers, G.J. Pierce, H. Zhang, J.C. Estill and A.H. Paterson. 2005. Comparative genomics of *Gossypium* and *Arabidopsis*: unraveling the consequences of both ancient and recent polyploidy. *Genome Res.*, 15: 1198-1210.

Schneider, K., B. Weisshaar, D.C. Borchardt and F. Salamini. 2001. SNPs frequency and allelic haplotypes structure of *Beta vulgaris* expressed genes. *Mol Breed.*, 8: 63-74.

Shaheen, T., M. Rahman and Y. Zafar. 2006. Chloroplast RPS_8 gene of cotton reveals the conserved nature through out plant taxa. *Pak. J. Bot.*, 38: 1467-1476.

Tenaillon, M.I., M.C. Sawkin, A.D. Long, R.L. Gaut, J.F. Doebley and B.S. Gaut. 2001. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *Mays* L.). *Proc. Natl. Acad. Sci*., 98: 9161-9166.

Varshney, R.K., U. Beier, E.K. Khlestkina, R. Kota, V. Korzun, A. Graner and A. Borner. 2007. Single nucleotide polymorphisms in rye (*Secale cereale* L.): discovery, frequency, and applications for genome mapping and diversity studies. *Theor. Appl Genet*., 114: 1105-1116.

Wu, Y-X., M.K. Daud, L. Chen and S.J. Zhu 2007. Phylogenetic diversity and relationship among *Gossypium* germplasm using SSRs markers. *Plant Sys. Evol.*, 268: 199-208.

Yu**,** J., S. Hu, J. Wang, G.K. Wong, S. Li., B. Liu, Y. Deng, L. Dai, Y. Zhou and X. Zhang. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Sci.*, 296**:** 79-92.

Zhang, H., Y. Li, B. Wang and P.W. Chee. 2008. Recent Advances in Cotton Genomics *Inter J Plant Genom*., doi:10.1155/2008/74230.

(Received for publication 30 June 2008)