CHROMOSOMAL ASSIGNMENT OF SIX ALPHA-EXPANSIN GENES USING SNP MARKERS IN COTTON (GOSSYPIUM SPP.)

Chuanfu An
Department of Plant and Soil Science
Mississippi State, MS
Sukumar Saha and Johnie N. Jenkins
USDA-ARS, Crop Science Research Laboratory
Mississippi State, MS
Brian Scheffler
USDA-ARS-CGRU
Stoneville, MS
Thea A. Wilkins
Department of Plant Sciences
Davis, CA
David M. Stelly
Texas A&M University
College Station, TX

Abstract

Thirteen pairs of PCR primers were designed from the available sequence information of the six member of cotton *alpha-expansin* gene family for amplification of TM-1, HS46 and MARCABUCAG8US-1-88 (*Gossypium hirsutum*), 3-79 (*G. barbadense)*, *G. tomentosum*, *G. mustelinum* and two putative diploid ancestral species accessions (*G. arboreum* and *G. raimondii*) of tetraploid cotton. The PCR amplified fragments were cloned into pCR4-TOPO vector by the TA cloning method. Twelve colonies were selected from each of the amplified product for forward and reverse sequencing. Based on the sequence similarity with the two diploid species, sequences of the tetraploid cotton were putatively assigned first to A or D genome. Comparative analysis of sequences of *alpha-expansin* genes revealed the presence of 190 SNP sites from about 12 kb of sequences in six tetraploid cotton lines. The average rate of interspecies SNP per nucleotide between *G. hirsutum* and *G. barbadense*, *G hirsutum* and *G. mustelinum*, *G. hirsutum* and *G. tomentosum* were 1.0%, 0.66% and 1.2% respectively. An average frequency of one SNP per 311 bp was present at intraspecies level within *G. hirsutum*. About 58% of SNPs were present in the exon region. Fifty five suitable SNP primers were developed from sequence information for detection of gene specific SNP markers. ABI Prism SNaPshot^{TM} multiplex kit was used for SNP marker identification. We discovered chromosomal locations of 25 SNP markers specific to six *alpha-expansin* genes to the long arm (Lo) of chromosome 20, 10, 9, 3 and 1 respectively by deletion analysis method. Chromosomal locations were concordant with the genomic origin based on comparative sequences of the tetraploid species with the ancestral diploid species utilizing phylogenetic analysis methods.

Objective

Alpha-expansin genes play a very important role in cell wall extension by disrupting noncovalent bonds between the wall components during cotton fiber development. The objectives of this study were to: (1) discover SNP markers specific to six *alpha-expansin* genes using the amplified sequences information in tetraploid cotton species and (2) identify the chromosomal location of these SNP makers.

Material and Methods

Plant Material

Two putative diploid ancestral species (*G. arboreum* and *G. raimondii*) of the tetraploid cotton and four tetraploid species, TM-1, HS46 (RIL-1) and MARCABUCAG8US-1-88 (RIL-2) (*G. hirsutum*), 3-79 (*G. barbadense)*, *G. tomentosum*, and *G. mustelinum* were used to identify SNP markers. Interspecific monosomic or monotelodisomic substitution lines (BC_0F_1) of *G. barbadense* and *G. tomentosum* in TM-1, monosomic reciprocal translocation lines in *G. tomentosum* and euploid interspecific chromosome substitution lines (CS-B, BC_5S_1) of *G. barbadense* in TM-1 were used to identify chromosomal location of SNP markers.

Methods

Using available sequence information 13 pairs of PCR primers were designed to amplify six *alpha-expansin* genes in the eight diverse cotton lines (Harmer *et al.*, 2002). The amplified products were cloned into pCR4-TOPO vectors and 12 colonies from each of the amplified product were used for forward and reverse sequencing. The alignment of the sequences specific to the same PCR primer in 8 lines was performed with the SeqMan and MegAlign program of DNASTAR software (DNASTAR, Inc., Madison, Wisconsin, USA). The phylogenetic relationship of the diverse sequences were analyzed using UPGMA, NJ or ML method by PAUP (Swofford *et al*., 2003), PHYLIP 3.6 (Felsenstein, 2005) and Clustalx (Thompson *et al*., 1994). ABI Prism SNaPshot^{TM} multiplex kit and an ABI 3100 electrophoretic system (Applied Biosystems, Foster City, CA, USA) were used for SNP marker screening and chromosomal assignment. Chromosomal locations were discovered by deletion analysis method (Liu *et al*., 2000).

Results

(1) Developed an improved method of SNP marker detection in tetraploid cotton. Based on the sequence similarity with the two diploid species, sequences of the tetraploid cotton were putatively assigned first to A or D genome. SNP markers of six *alpha-expanisns* genes were discovered by aligning sequences from eight different cotton lines (Figure 1). A total 190 SNPs were identified from about 12 kb of sequences of six tetraploid cotton lines. About 58% of them were present in the exon regions. The average rate of interspecies SNP per nucleotide between *G. hirsutum* and *G. barbadense*, *G hirsutum* and *G. mustelinum*, *G. hirsutum* and *G. tomentosum* were 1.0%, 0.66% and 1.2% respectively. An average frequency of 1 SNP per 311 bp (0.3%) was occurred within *G. hirsutum* species (Figure 1 and 2).
(2) Chromosomal locations of 6 different *alpha-expansin* genes were discovered (Table 1 and Figure 3). The results were confirmed further by more than one SNP marker from different regions of the same gene sequence (Table 1).
(3) Genomic origins based on chromosomal locations of different *alpha-expansin* genes were concordant with the sequence similarity of the diploid ancestral species with the tetraploid cotton species analyzed by phylogenetic analysis methods. For example, *alpha-expansin-1* and *alpha-expansin-2* are on homeologous chromosomes 10 and 20 (Figure 4, Lacape *et al*., 2005).

Conclusion

In summary, we identified chromosomal locations of 25 different SNP markers specific to six member of cotton *alpha-expansin* gene family to the long arm (Lo) of chromosome 20, 10, 9, 3 and 1 respectively using deletion analysis method.

Reference

Harmer *et al*. (2002) Characterisation of six *α-expansin* genes in *Gossypium hirsutun* (upland cotton). *Mol Genet Genomics* **268**: 1-9
Lacape *et al.* (2005) QTL analysis of cotton fiber quality using multiple *Gossypium hirsutum* × *Gossypium barbadense* backcross generations. Crop Science **45**: 123-140
Liu *et al*. (2000) Chromosomal assignment of microsatellite loci in cotton. *The Journal of Heredity* **91**: 326-332

Acknowledgements

We are grateful for the technical help from Mr. Douglas A. Dollar and the suggestions from Dr. Z. T. Burive and Ms. Yufang Guo.

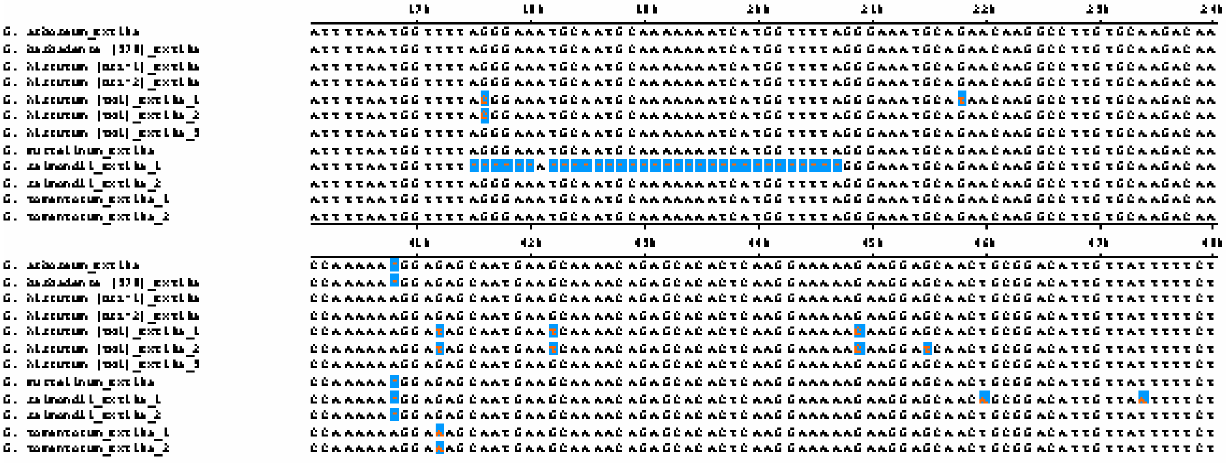

Figure 1. Alignment of the sequences of *alpha-expansin-4* gene fragments showing the presence of deletion or insertion and SNP sites

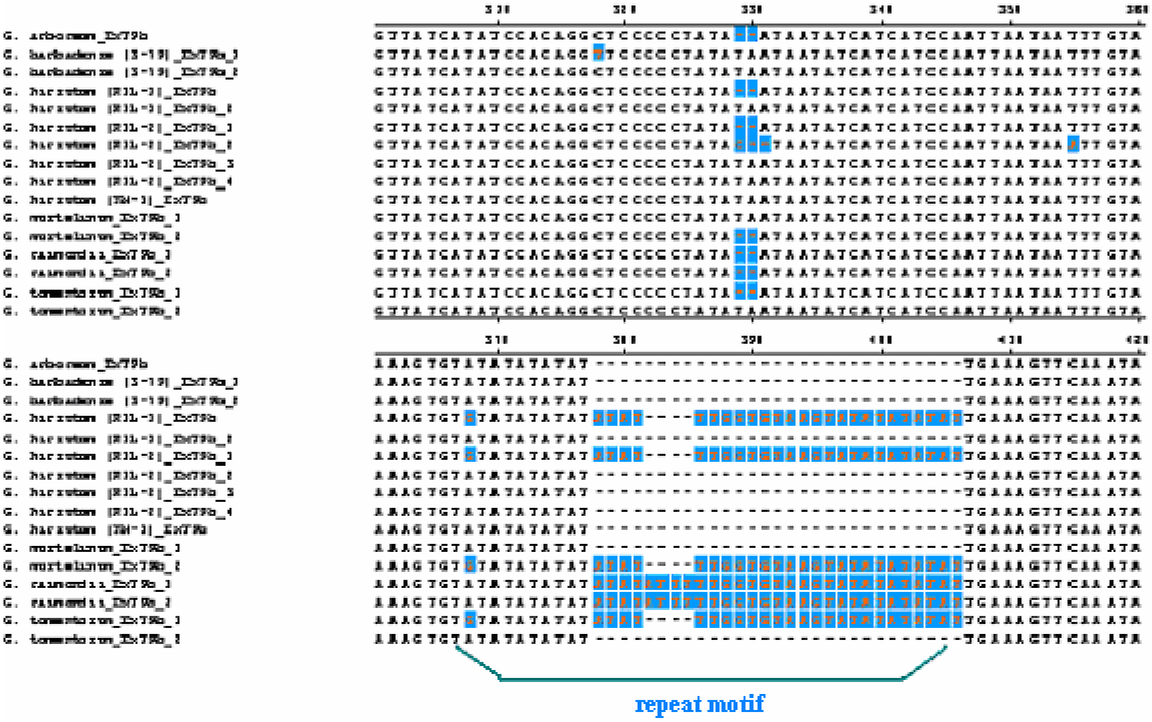

Figure 2. Alignment of the sequences of *alpha-expansin-5* showing the presence of simple sequence repeat (TA) motif and SNP sites

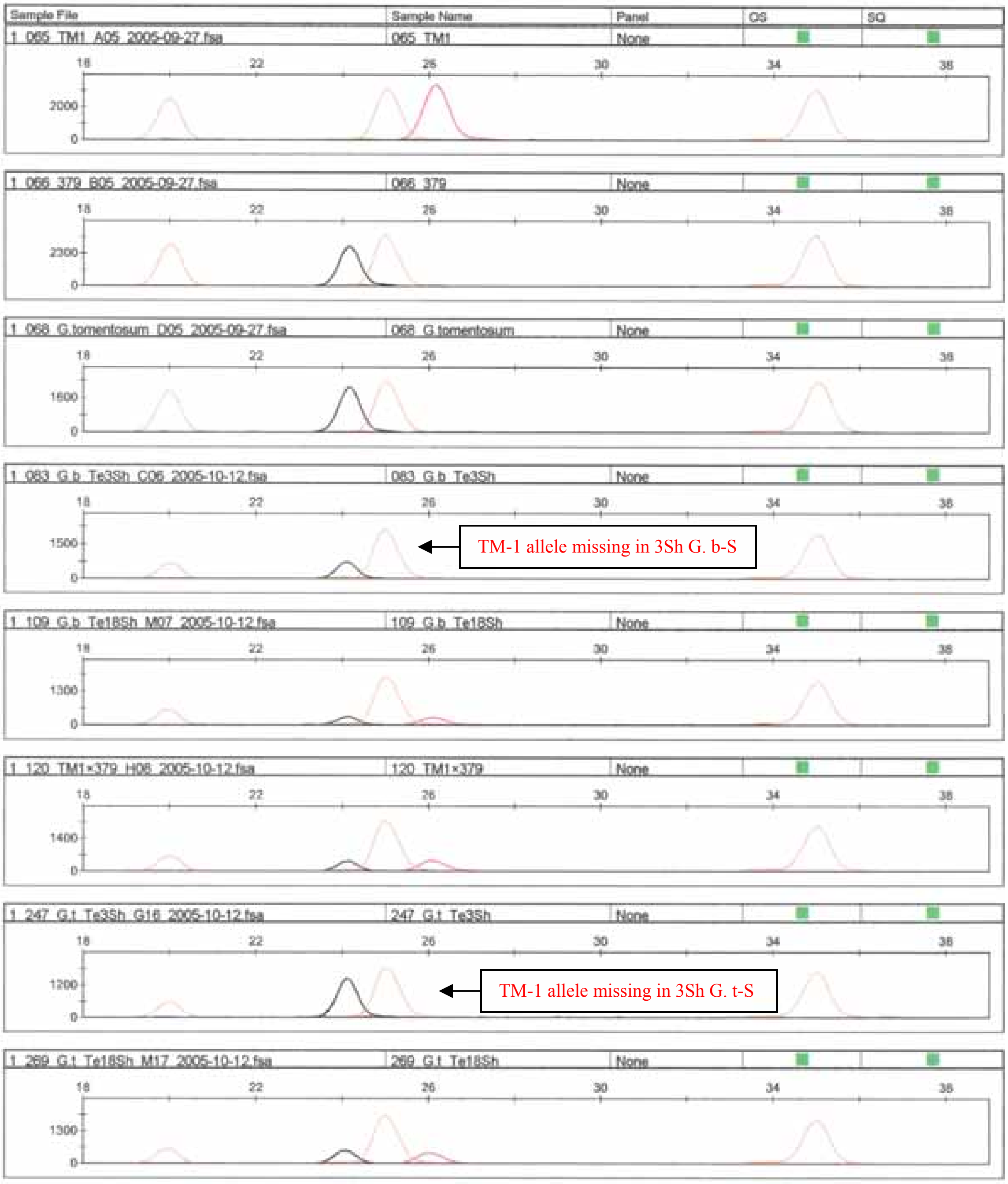

Figure 3. Chromosomal location of *alpha-expansin-6* gene using ExT8aGbmt_96_R SNP marker
The results showed that both aneuploid substitution line missing the long arm of chromosome three (3Sh aneuploid substation line) of *G. barbadense* and *G. tomentosum*, respectively, does not have TM-1 specific SNP allele suggesting this SNP marker is located on the long arm of chromosome 3. Here, different bases showed different color; X axis represented DNA fragment size and Y axis showed the amount of the product.

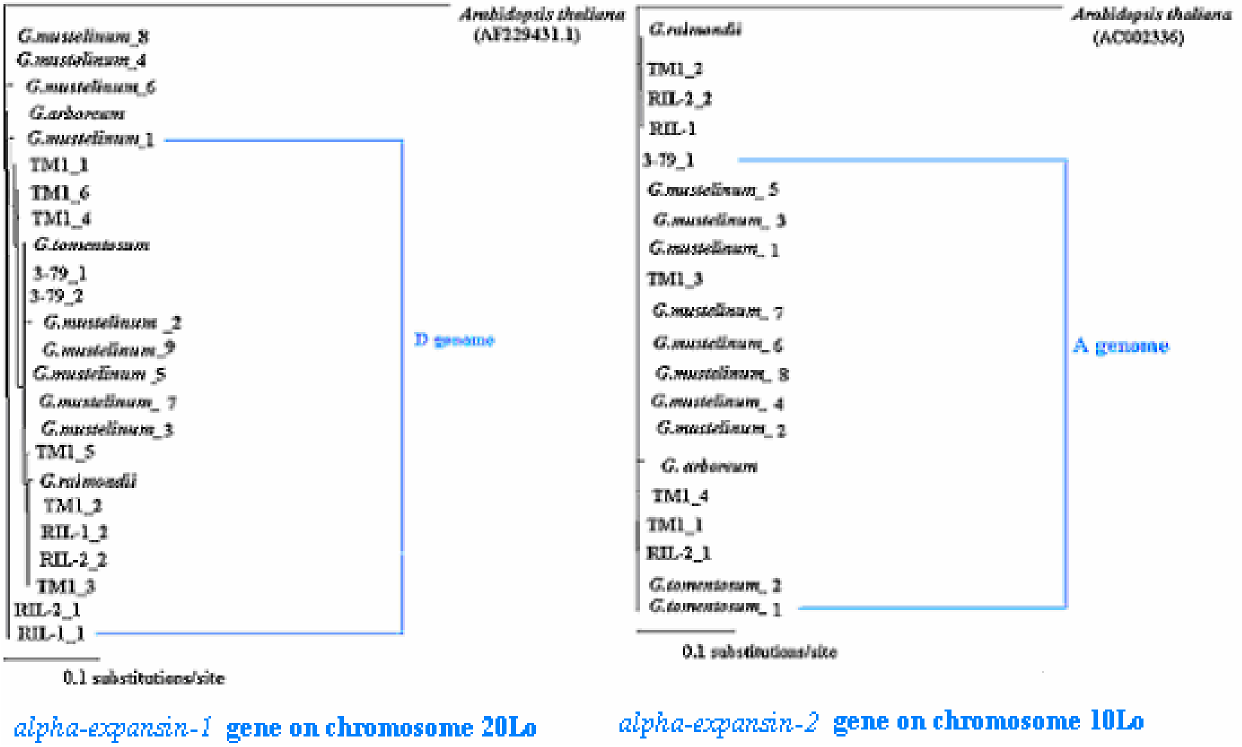

Figure 4. Phylogenetic tree based on sequence analysis of *alpha-expansin-1* and *alpha-expansin-2* by maximum likelihood method supported the homeologous relationship of the respected chromosomes in the tetraploid cotton species (Note that all sequences are different and the same number does not represent similar sequence)

Table 1. Chromosomal locations of six *alpha-expansin* genes specific SNP markers

Gene	**PCR primer pair**	**SNP primer**	**Chromosome locations**		
			Aneuploid *G. barbadense*	**Aneuploid *G. tomentosum***	**Euploid CS-B**
alpha-expansin-1	ExT13b	ExT13bGbmt_193_F	Te20Lo	Te20Lo	*
	ExT12a	ExT12aGbmt_378_R	Te20Lo	Te20Lo	*
alpha-expansin-2	ExT12b	ExT12bGbt_58_F	Te10Lo	Te10Lo	10
		ExT12bGbt_59_R	Te10Lo	Te10Lo	10
		ExT12bGbt_93_F	Te10Lo	Te10Lo	10
		ExT12bGbt_108_R	Te10Lo	Te10Lo	10
		ExT12bGbt_175_F	Te10Lo	Te10Lo	10
		ExT12bGb_182_R	♦	♦	10
		ExT12bGbt_312_F	Te10Lo	Te10Lo	10
		ExT12bGbmt_345_R	Te10Lo	Te10Lo	10
		ExT12bGbt_415_F	Te10Lo	Te10Lo	10
		ExT12bGbmt_422_R	Te10Lo	Te10Lo	10
		ExT12bGbt_508_R	♦	NTN10-19[§]	10
alpha-expansin-3	ExT11a	ExT11aGb_489_F	Te10Lo	♦	10
	ExT11b	ExT11bGb_372_F	♦	♦	10
alpha-expansin-4	ExT10a	ExT10aGbmt_65_F	Te9Lo	Te9Lo	9
		ExT10aGbmt_147_R	Te9Lo	♦	9
		ExT10aGbm_412_F	Te9Lo	♦	9
	ExT10b	ExT10bGbmt_244_F	Te9Lo	Te9Lo	9
alpha-expansin-5	ExT9b	ExT9bGmt_205_R	♦	1	*
	ExT9c	ExT9cGbt_444_F	Te1Lo	♦	1
alpha-expansin-6	ExT8a	ExT8aGb_77_F	Te3Lo	♦	*
		ExT8aGm_89_F	Te3Lo	♦	*
		ExT8aGbmt_96_R	Te3Lo	Te3Lo	*
		ExT8aGb_156_R	Te3Lo	♦	*

*The backcrosses euploid CS-B line was not available.

♦Monomorphic parents, chromosomal location could not be identified by this series of substitution stock.

§The duplicate deficient lines developed from reciprocal translocation were named as NTN. For example, NTN 10-19 line is missing some part of chromosome 10 and 19.