

## Isolation and characterization of sucrose phosphate synthase promoter from cotton (*Gossypium hirsutum* L.)

Nadia Iqbal<sup>1\*</sup>, Muhammad Asif<sup>1</sup>, Amara Masood<sup>1</sup>, Rubab Zahra Naqvi<sup>1</sup>, Asia Khattoon<sup>1</sup>, Aftab Bashir<sup>2</sup>

<sup>1</sup>National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan

<sup>2</sup>Department of Biological Sciences, F.C. College University, Lahore, Pakistan

\*Corresponding author: naaaadia7@gmail.com

### Abstract

Cotton fiber genes and promoters are of great importance in understanding fiber development mechanism as well as for improvement of fiber. Sucrose phosphate synthase gene (SPS) (insert abbreviation of gene) has found to express at higher rate in developing cotton fibers. It is an important enzyme that has major role in sucrose as well as cellulose synthesis. Upstream region of a SPS gene from cotton was retrieved through HTGS database and analyzed using bioinformatics tools. Sequence analysis identified various regulatory motifs including light, drought, heat responsiveness and MYB binding sites in the promoter. The SPS promoter was isolated from cotton genomic DNA (using what method) and fused to  $\beta$ -glucuronidase (GUS) gene in a plant expression vector. Transient GUS expression analysis in various cotton tissues showed that promoter was active in the fiber tissues. Full length 2 kb SPS promoter showed high expression in fibers during elongation and secondary cell wall synthesis stage. A 1.5 kb deletion fragment showed reduced expression in fibers. Our results suggest that cotton SPS promoter may be used to express genes specifically in fiber cells for improvement of cotton fiber quality traits.

**Keywords:** Promoter, Sucrose phosphate synthase, sequence analysis, transient GUS expression.

**Abbreviations:** SPS\_Sucrose phosphate synthase, HTGS\_High Throughput Genomic Sequences, PlantCARE\_Plant Cis Acting Regulatory Element, PlantPAN\_plant Promoter Analysis Navigator, GUS  $\beta$ -Glucuronidase.

### Introduction

Cotton is the major fiber crop of the world used in textile industry. Cotton fibers are seed outgrowths differentiated from epidermal cells of ovule. Cotton fiber development involves a large number of genes controlled by their upstream regulatory regions called promoters. Many fiber genes along with their promoters have been explored in developing fibers (Huang et al., 2013; Delany et al., 2007). Most of the fiber related genes and their promoters exhibit tissue or fiber developmental stage specific expression. Earlier, promoters of various lipid transfer proteins showed organ specific expression (Wu et al., 2007). A cotton arabinogalactan 6H6L promoter showed GUS expression in petiole and pedicle (Wu et al., 2009). On the other hand, a cotton protodermal factor promoter exhibited specific expression on ovular surface and elongating fibers (Deng et al., 2012). Promoters of some MYB transcription factor genes showed trichome as well as fiber specific expression in cotton (Machado et al., 2009). For genetic modification of cotton fiber, the use of such promoters is highly desirable. There are numerous binding sites in promoter for transcription factors that have crucial role in controlling genes expression. These conserved motifs are scattered over whole promoter region and have role in determining specific expression of promoter. These motifs along with core promoter elements can be identified through a number of available databases (Hieno et al., 2014).

Sucrose Phosphate synthase (SPS) is an enzyme found to be highly expressed in developing cotton fibers. Sucrose phosphate synthase regulates sucrose cycling in plant cells during secondary cell wall synthesis (Winter et al., 2000). It is the major enzyme for synthesis of sucrose-6-phosphate from UDP-glucose and fructose-6-phosphate (Lunn et al., 2003). Sucrose thus synthesized is broken down into UDP-glucose and fructose by sucrose synthase (susy). UDP glucose is used as substrate for cellulose synthesis which is the major component of cotton fiber (Amor et al., 1995). Increasing cellulose contents may help to increase fiber strength which is desirable for textile processing. In cotton and other plants, over expression of SPS might result in increasing plant biomass due to increased synthesis of cellulose. Transgenic cotton with over expressed SPS exhibited higher concentration of sucrose ratio as compared to starch (Haigler et al., 2007). Promoters of SPS gene have also been isolated and characterized in many plant species (Yonekura et al., 2013).

In the present study, we isolated and cloned a SPS promoter from cotton. Sequence analysis revealed various regulatory crucial for promoter activity and tissue specific expression. Promoter was cloned in plant expression vector and its expression was assayed in cotton tissues through transient GUS assay. The results showed 2 kb SPS promoter was able to express GUS gene in developing fibers. However, its 1.5 kb deletion mutant with some deleted cis-

acting motifs and transcription factor binding sites showed reduced activity. So, 2 kb SPS promoter may be used to express transgene specifically in cotton fibers. So, SPS promoter has potential to derive fiber specific gene expression for improvement of cotton fiber quality. In agriculture based country like Pakistan, there is a need to explore novel promoters to be used freely to generate transgenic crops. These promoters may also be used as components of multiple expression cassettes to avoid gene silencing caused by sequence homology of single or few promoters used to control expression of many genes.

## Results

### Retrieval of SPS promoter

A cotton sucrose phosphate synthase gene was selected for promoter isolation and characterization. BLAST search of nucleotide sequence of SPS gene in HTGS database revealed that sequence matched with HTGS of *Gossypium hirsutum* clone ZMMBBb244L7 (Accession number 193940.1). There was no coding region in selected 2 kb sequence and it was named SPS promoter. Patent results showed that nucleotide sequence of SPS promoter had no significant similarity with any patented sequence except 19% homology with patent WO 2009143995 (Accession No. HC 2078821)

### Sequence analysis of SPS promoter

Transcription initiation site for SPS promoter started at A with probability of 68% at position of -1170 nucleotide in promoter (Fig. 1). Output of plantCARE and plantPAN identified several cis-acting motifs (Table 1 and 2). TATA box was identified at position of -25 bp upstream of transcription initiation site. Another motif of core promoter region, CAAT box which has role in enhancing transcriptional activity of promoter was located at region of -105 relative to transcription initiation site. The 5'UTR py-rich stretch, an important motif associated with high level of gene expression was also found in SPS promoter at -916 position. Furthermore, the SPS promoter had three endosperm specific elements named as Skn-1 motif and GCN4 motif. The SPS promoter also contained two MBS motifs those bind specifically with MYB transcription factors. A TC rich repeat (ATTCTCTAAC) related to stress and defense response have been found in SPS promoter. A consensus nucleotide sequence GCCACT named as CAT box involved in meristem specific expression is also found in SPS promoter at position -1120. The SPS promoter had two WI motifs (TTGACC) involved in fungal elicitor response; one of them was in reverse orientation. A cis-acting element CE3 (GACGCGTGTC) was also found in putative promoter involved in ABA (abscisic acid) and VP1 responsiveness. Several other motifs regulated by light were found in whole promoter region. These included AE box, I-box, F-box and G-box, GT-1, Box 4 and 4cl-CMA2a motif. There was a heat shock element with consensus sequence (AAAAAATTTC) found to have role in heat stress responsiveness. An ARE motif (TGGTTT) related to anaerobic responsiveness was located in SPS promoter.

Sequence analysis through plantPAN demonstrated that SPS promoter contained a number of binding sites for MYB proteins including MYB1AT (TAACCA), MYB2AT (CAGTTA), AtMYB2 (CTAACCA) and AtMYC2 (CACATG). There were 20 binding sites for ARRIAT (aataatACGTaggtgg), and 4 sites for MYBST1 (GGATA) having role in initiation and activation of transcription.

PlantPAN analysis also identified several other motifs including GATA box (activation of transcription) AGL3 (regulator of transcription), INRNTPSADB (initiation element of transcription). Two binding sites for an epidermal factor HDZIP2ATAT with sequence TAATAATTAA were also located in SPS promoter sequence. Promoter region also contained 15 GAREAT motifs for gibberillic acid responsiveness. Another unique binding site in SPS promoter was for TELBOXATEEF1AA (elongation factor related motif) with consensus sequence AGTCAG. There were binding sites for ABRELATERD1 (etioloating response), ANAEROICONSENSUS (anaerobic response) and WBOXATNPRI (salicylic acid response). There were also a number of motifs those may bind specifically with WBOXNTRCF3 (wound inducibility) and WBOXNTRCHNH48 (elicitor response). SPS promoter also contains binding sites for transcription factors related to sugar response including SREATMSD (GGATAA) and SURE1STPAT21 (TTTTCTATT).

### Cloning of SPS promoter

To characterize expression pattern, SPS promoter was isolated from cotton genomic DNA and cloned in plant expression vector. For amplification of 2 kb SPS promoter, the optimized annealing temperature was 54°C as shown in the (Fig. 2A). An internal PCR amplification to generate 1.5 kb deletion fragment was also carried out. Both 2 kb and 1.5 kb amplified promoter fragments were cloned in general purpose TA cloning vector. Recombinant clones were confirmed through restriction analysis using *SacI* and *ApaI* (Fig. 2B and 2C). Promoter fragments were further subcloned into expression vector pGR1 by replacing 35S promoter upstream of GUS gene followed by CaMV terminator. The resultant clones were confirmed by digestion with *SacI* and *HindIII* (Fig. 2D) and were named pGRSPS2 and pGRSPS1.5 (Fig. 3).

### Transient GUS expression

Strong GUS expression was detected in cotton fibers after bombardment with vector constructs having 2 kb SPS promoter. Non fiber tissues did not exhibited any GUS activity. However, very weak GUS expression was detected in cotton leaves (Fig. 4). The transient GUS assay of 1.5 kb SPS promoter revealed that GUS stain was detected in all fiber tissues (Fig. 5). No GUS stain was detected in cotton root, stem and leaves after bombardment. Positive control having 35S promoter showed expression in fiber tissues (Fig. 4A) while negative control (promoter less construct) exhibited no expression under same experimental conditions (Fig. 4B).

## Discussions

Fiber specific genes and promoters are of great importance in understanding fiber development mechanism as well as improvement of cotton through transgenic technology. We have reported the isolation and evaluation of SPS gene promoter from cotton genome. Although there is homology of SPS gene with other plant SPS genes but its regulatory region has no homology with previously reported SPS promoter sequences. The SPS consisted of various motifs crucial for promoter activity. There was one motif related to methyl Jesmonic acid responsiveness (CGTCA) in SPS promoter as in other JA-inducible genes promoter in *Arabidopsis*

**Table 1.** PlantCARE analysis of cotton SPS promoter.

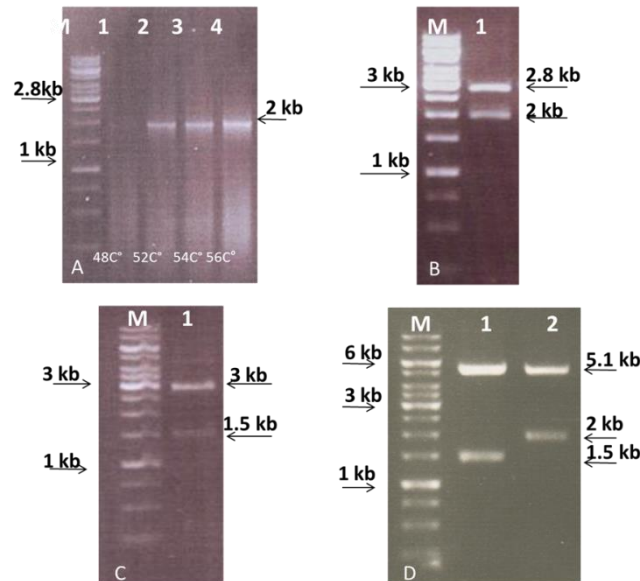
Motif	Organism in which reported	sequence	Putative function
4cl-CMA2a	<i>Petroselinum crispum</i>	TCATCACTAACAC	Light responsive
5'UTRPy-rich stretch	<i>Lycopersiconesculentum</i>	TTTCTTCTCT	Cis-acting element conferring high transcription
WI motif	<i>Arabidopsis</i>	TTGACC	Fungal response
AE-Box	<i>Arabidopsis</i>	AGAAACAA	Light responsiveness
Box 4	<i>Petroselinum crispum</i>	ATTAAT	Light responsiveness
ARE	<i>Zea mays</i>	TGGTTT	Anaerobic responsiveness
Box 1	<i>Pasiumsativum</i>	TTTCAAA	Light responsiveness
CAT-box	<i>Arabidopsis</i>	GCCACT	Meristem expression
CE3	<i>Oryza-sativa</i>	GACGCGTTGTC	ABA and VP1 responsiveness
G-box	<i>Arabidopsis</i>	CACGTG	Light responsiveness
G-Box	<i>Pisumsativum</i>	CACGTG	Light responsiveness
GT1	<i>Avena sativa</i>	GGTTAA	Light responsiveness
HSE	<i>Brassica oleracea</i>	AAAAATTTTC	Heat responsiveness
I-box	<i>Pisumsativum</i>	ATGATATGA	Light responsiveness
MBS	<i>Arabidopsis</i>	TAACGT	MYB binding sites
TATA-box	<i>Brassica napus/ Arabidopsis</i>	ATTATA/TATATATA	Core promoter element about -30 of transcription start site
TC-rich repeats	<i>Nicotianatabacum</i>	ATTCTCTAAC	Defense and stress responsiveness
Skn-1 motif	<i>Oryza-sativa</i>	GTCAT	Endosperm specific expression
GCN4 motif	<i>Hordeum vulgare</i>	TGTGTCA	Endosperm specific expression

-841 AAAAATATTACTTTTCTCAATTTTTTTAAGCAATAAATTGAGTGCAATATTGGCTTACC -779  
 -778 AATTATCATAACAATTTTATTCAAACCTTATAACCAAGTTTGATAATAATTTTTT -719  
 -718 TAATTGTTGGTTGAGTCTTAACCAATTAGTATAGGTATAATTATCAATGTAGAAGAAATG -659  
 -658 AAATTCGAGTACTGAAGCGCTTATCTCTTATTTATGGGTAGAGAGAGACTAAAGT -599  
 -598 AATTCGAGCATGTATAAAAAACCAATATGATCGAAATCTAAATAAAATTTATTTAA -539  
 -538 AAAATTCCTCTCTATTTTAAATTTTACCATTCAATGTTTTTTCTATTAGATGTTCT -479  
 -478 CTAATTTATTTGATTATTATATAGATTATCCGATATGTATGTTGTAATAATA -419  
 -418 ATTATACATGAGTTTGTGTTAATTCATGAGATGCTTCAAATATGACATATTTTA -359  
 -358 AATTAACCTCCAAATTAATAAAATTCATTCATTAAGGGATAAATCTCAAAT -299  
 -298 ATACATGAACCTTGATTTAATGTGAATTATATACATGAACCTTAATTTAGTGCAAT -239  
 -238 TACAGTGAACTCTAATTGTTGTTCAAATATATACCTTGAACATTAATTTGATTAAATC -179  
 -178 ATGCATATTTAAAGAAATAAATACATCAATTTATTTTATATTAGATAAAATATAATTATT -119  
 -118 TATGTATGCAATATATAACATAAGATGACATTATGACATAAATATGTGTAATAATTAC -59  
 -58 AAGAATTAAATAAAATCAAACCTTTATGTATAAAATACACAAAAATCAAATTCATGT 2  
 3 ACAAATTAACAGTTAAACCATAGTTCAAATATAGTTTGTAGATTACCCCATTCATAAAA 63  
 64 TGTAAGTATATTATCATTGATTTTTTTTGTAAATTAATCGTAAATTTGAGCGTTTAAT 124  
 125 GATGGCTTAAATTTGATGTAGATGCTAACTTAAACATATGAATAGAAACAGTATGATATG 184  
 185 AAAGTCCTTAAATTTCAAACCTTAAAGGTAAATTTGACCTGATTGCTCCACATGTTTA 244  
 245 AATTGGTATTCATATGTTGAGCTTAAATGTCAAACCTACAAATTTAATATTTTAGGATTC 304  
 305 AGTTATAGAAATTTAAATCTTAGGATAAATTTAGAGTTTGTCTATAGGTCATCTTGTGG 364  
 365 AATTAACCTCTATTATTTCAATATTTTACATAAATTAATAACCGAGGTCAAATTTG 424  
 425 ACTATGTTGCTTTACTCTGACACAATGTCGGTCTCAGGAGTAAAGAGAAAAGGTAA 484  
 485 AAATGTTTTTTTTCTCAAATAATGGTCAAATTAATTTAGATACCTTTACTATGTGT 544  
 545 AAATTTAGATACTCTAAAGTGTCTGTATTAGAGAAAAAATATTAGATTTTAATTTA 604  
 605 TAGAATTTGATGTTTAAAGTTAATTTATTAGTATAAATTTGATTTGTAAAAGAAATAG 664  
 665 GAAAAAGTGTTCAGAGTTAGTGTAGTAGTTTGAATAGAAACAATTATAAATAACA 724  
 725 ATTTAATAGCATAGTATTTTGGAAATAATTAATAAATTTTATAATTTTAAATTTAGTGG 784  
 785 CTAATAATATAAGTTATTATTAATTTAGTGTAGTTAGATCCATAAAAAACATTATACCTA 824  
 825 TTTCTTCATCAGTAAACAGTGTGGTCCCTCTAACCAAGAAACTAGCACACGCGTCAA 884  
 885 CCTCATCCCTGTATTTAAATCAAAGCTCCAAACCAACCAACCAAAATATAACAATA 944  
 945 TCTCATCAGGTTTATGATGATTCCGCTCTTCATCATCCGTTTCTAGCACCTAAT 1004  
 1005 GATACTATTATCATTAAATCAAATTTTTTTGAATATTGCTACACCGCATAGACTCT 1064  
 1065 GTAAATTCAGTCTCTCGGTTTCTTTATTTTAAATCTTTGTTATTTGTTGGTGGAT 1124  
 1125 TTTTTCAGTCAAAGAACTGTTATACCTGGGTCTAAATCTCATCC 1172

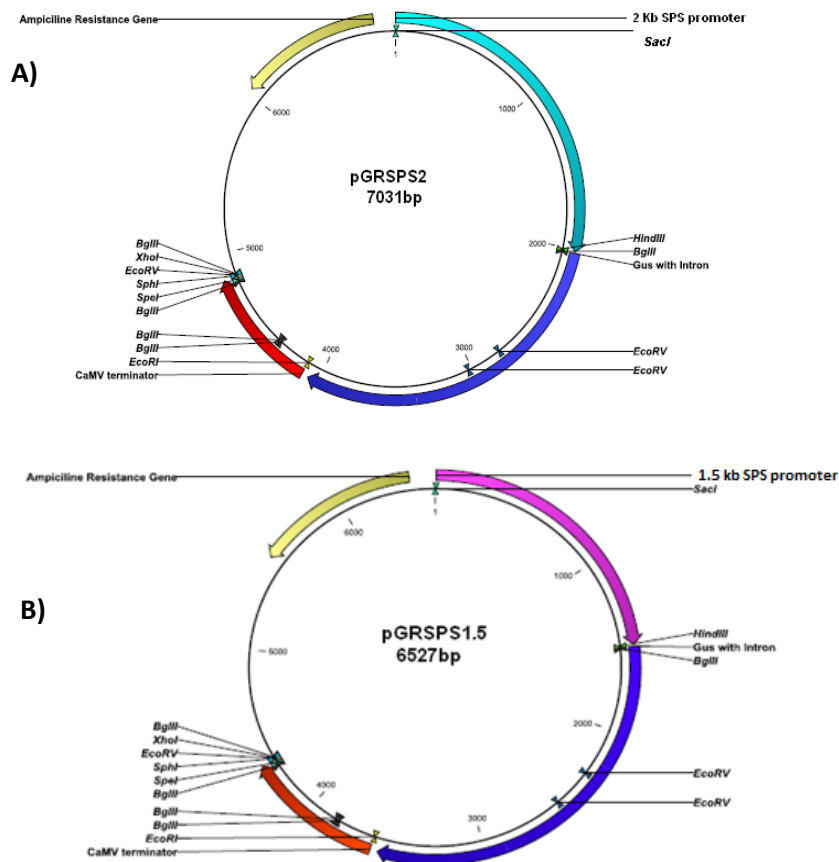
**Fig 1.** Nucleotide sequence of SPS promoter isolated from cotton. Translation initiation codon (ATG) of SPS gene is highlighted with blue color and transcription initiation site (+1) has been indicated with letter (A) and red color. TATA box is located at (–25) position indicated by light grey color while CAAT box is located at –326 position represented by green color. G-box and ABRE are highlighted with pink and yellow colors respectively.

**Table 2.** PlantPAN analysis of SPS promoter.

Motif	Organism in which reported	Sequence	Function
MYB1AT	<i>Arabidopsis</i>	TAACCA	Dehydration responsiveness
MYB2AT	<i>Arabidopsis</i>	CAGTTA	Dehydration responsiveness
AtMYB2	<i>Arabidopsis</i>	CTAACCA	Drought & ABA related responsiveness
AtMYC2	<i>Arabidopsis</i>	CACATG	ABA signalling
ARRIAT	<i>tobacco/ Arabidopsis</i>	aataatcACGTaggttg	Transcription activation
MYBST1	<i>Arabidopsis</i>	GGATA	Transcriptional activation
GATABOX	<i>Arabidopsis</i>	TATC	High transcription motif
AGL3	<i>Arabidopsis /pea</i>	CTATTTATGG	Regulator of transcription
INRNTPSADB	<i>Arabidopsis</i>	TTCAGTTC	Initiation element for transcription
HDZIP2ATATHB2	<i>tomato/Lycopersicon</i>	TAATAATTA	Epidermal differentiation
GAREAT	<i>petunia/Arabidopsis</i>	TTTGTTA	GA induced expression
TELOBOXATEEF1AA1	<i>Arabidopsis</i>	AAACCCTAA	Elongation factor related motif
GT1GMSCAM4	<i>Arabidopsis</i>	TTTTTC	Pathogen/heat responsiveness
ANAERO1CONSENSUS	<i>Arabidopsis</i>	AAACAAA	Anaerobic responsiveness
CCAATBOX1	<i>carrot/A. thaliana</i>	ATTGG	Heat responsiveness
ERELEE4	<i>Arabidopsis</i>	AATTCAAA	Ethylene responsive motif
WBOXNTPCHN48	<i>tobacco</i>	AGTCAG	Elicitor response
WBOXNTERF3	<i>tobacco</i>	TGACC	Wound inducible expression
SREATMSD	<i>potato</i>	GGATAA	Sugar responsive element
SURE1STPAT21	<i>Arabidopsis</i>	TTTTCTATT	Sucrose responsiveness

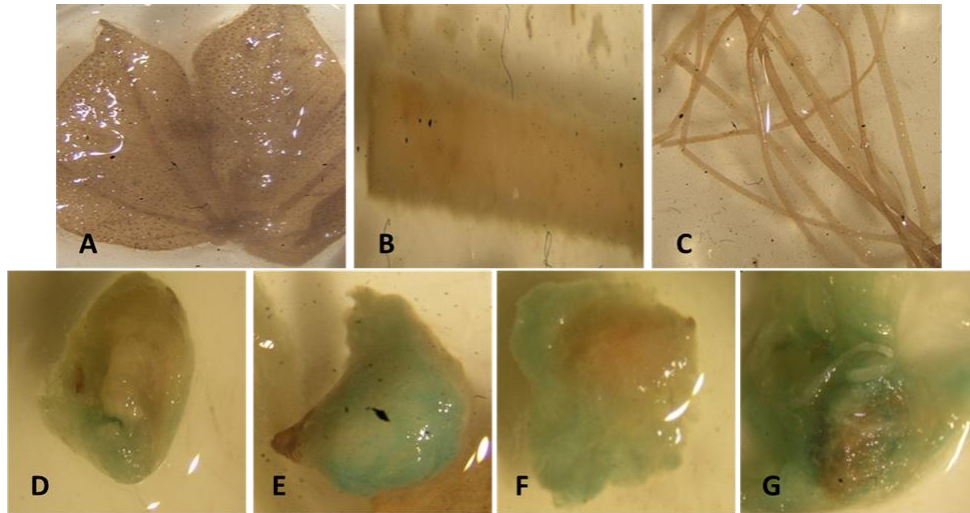


**Fig 2.** PCR amplification and cloning of SPS promoter in TA cloning vector and pGR1. A) M: 1Kbp DNA ladder, Lanes 1 to 4: optimization of annealing temperature for PCR amplification of SPSpromoter. B) M: 1Kb DNA ladder, Lane 1 SPSpromoter cloned in TA vector restricted with *SacI/ApaI* excising the 2 kb SPS promoter and 2.8 Kb vector backbone promoter. C) M: 1Kb DNA ladder, Lane 1: Deletion fragment of SPS promoter cloned in TA vector restricted with *SacI/ApaI* excising the 1.5kb SPS promoter and 2.8 Kb vector backbone. D) M: 1Kbp DNA ladder, 1 and 2 represent: SPS promoter cloned in pGR1 digested with *SacI/HindIII* and releasing 2 Kb full length promoter clone and 1.5 Kb deletion fragment along with 5.1 kb vector back bone.

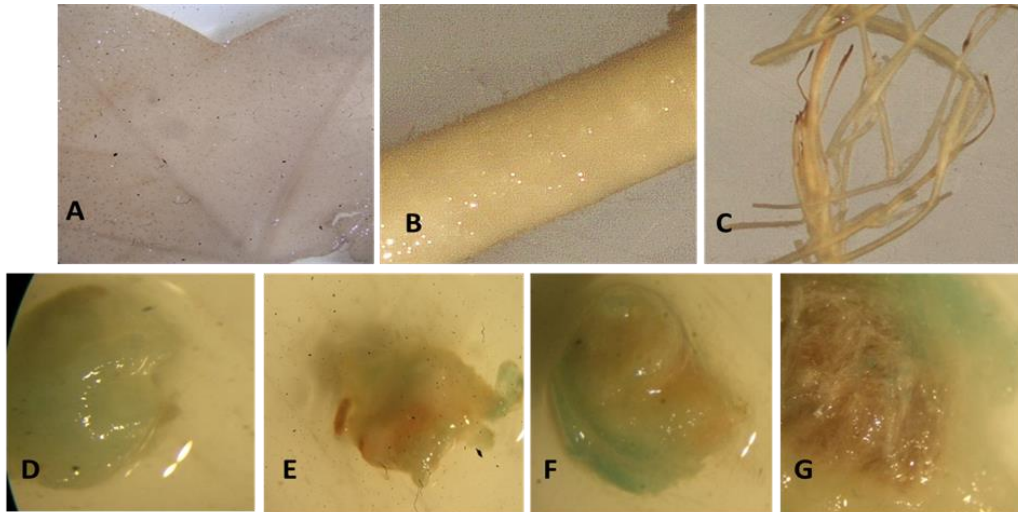


**Fig 3.** Plasmid drawings of vector constructs having promoters fused to GUS gene and CaMV terminator. A) Physical map of expression cassette having 2 kb SPS promoter (pGRSPS1.5). B) Physical map of expression cassette having 1.5 kb SPS promoter (pGRSPS2).

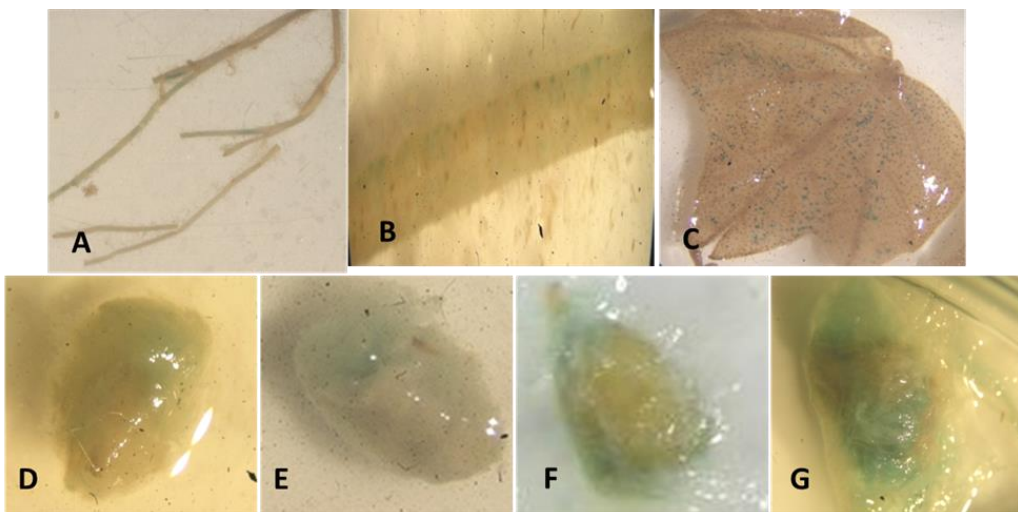




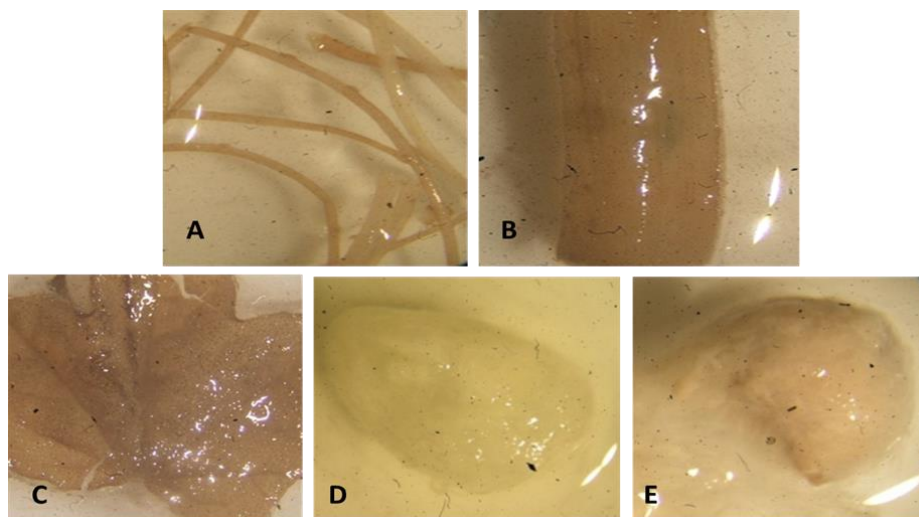
**Fig 4.** GUS expression of 2 kb SPS promoter in cotton tissues . No GUS expression was detected in leaf (A) , stem (B) and roots (C) . Strong GUS stain was observed in 5 DPA (D) , 10 DPA (E) , 15 DPA (F) , 20 DPA fibers (G) .



**Fig 5.** GUS expression of 1.5 kb SPS promoter in cotton tissues . No GUS expression was detected in leaf (A) , stem (B) and root (C) . Weak GUS stain was observed in 5 DPA (D) , 10 DPA (E) , 15 DPA (F) , 20 DPA fibers (G).



**Fig 6.** Transient GUS expression of 35 S promoter in cotton tissues. GUS expression was detected in all tissues including roots (A), stem (B) and leaf (C) , 5 DPA (D) , 10 DPA (E) , 15 DPA (F) , 20 DPA fibers (G).



**Fig 7.** Transient GUS expression of promoterless construct in cotton tissues . No GUS expression was detected in tissues including roots (A) , stem (B) and leaf (C) , 5 DPA fibers (D) , 15 DPA fibers (E).

(Guerineau et al., 2003). The SPS promoter had two WI motifs with consensus sequence (TTGACC) one of them is in reverse orientation. This motif is found to have role in fungal elicitor responsiveness (Rushton et al., 1996). Heat shock element (AAAAAATTTTC) located in SPS promoter found to have role in heat stress response (Nover et al., 2001). 5'UTRpy-rich stretch an important motif associated with high level of gene expression was also found in SPS promoter at -916 position (Daraselia et al., 1996). A TC repeat region was also found in SPS promoter. This motif is involved in fungus attack response as reported previously (Daiz-D-leon et al., 1993). In some light regulated gene promoters these motifs act as positive or negative modulators of transcription. For example, GT-1 motif act as negative regulator while I-box act as activator of transcription. GT-1 motif has been found to have interaction with GT-1 like transcription factors found to have role in pathogen and salt stress response (Zhou et al., 1999). On the other hand, G-box is protein binding sequence found in the Rbcs promoter induced by environment stimuli (Mary et al., 1992). The sequence analysis revealed that cotton SPS promoter have various binding sites for MYB transcription factors which have previously found to have role in fiber morphogenesis. MYB are the largest family of transcription factors. MYB proteins have role in trichome development in *Arabidopsis* and fiber development in cotton (Pu et al., 2008).

A consensus nucleotide sequence GCCACT named as CAT box involved in meristem specific expression is also found in SPS promoter as previously reported (Manevski et al., 2000). A cis-acting motif CE3 (GACGCGTGTC) was found in SPS promoter involved in ABA (abscisic acid) and VP1 responsiveness. These motifs found to interact with ABRE motifs required for ABA and VP1 related expression (Hobo et al., 1999). A unique motif in SPS promoter was HDZIP2ATATAHB2 interacts with *athb2* and reported to be involved in morphological changes (Ohgishi et al., 2001). This cis-acting motif was found to present in cotton PROTODERMALFACTOR1 gene promoter that derives fiber specific expression during initiation and early elongation phase (Deng et al., 2012). Presence of various light regulated motifs and pathogenic responsive sequences also depicted that this promoter may be induced by light. The results further suggest that SPS promoter may be used for methyl jasmonic acid and abscisic acid induced

expression. Promoters of SPS gene have been isolated and characterized in other plant species previously. A rice SPS promoter is reported to be controlled by light and circadian rhythm (Yonekura et al., 2013).

Experimental characterization of promoter revealed that it derived expression of reported gene at high rate in fibers. The GUS expression pattern after biolistic transformation depicted that promoter was active in the fiber tissues. Cotton tissues were also bombarded with 35S promoter as positive control and promoter less construct as -ve control to avoid false results (Fig. 6 and 7). Bombardment of full length SPS promoter cassette (pGRSPS2) exhibited strong GUS activity in elongating fibers (Fig. 4). Deletion mutant of SPS promoter (pGRSP1.5) showed weak GUS stain in fiber tissues (Fig. 5). Several motifs deleted in 1.5 kb SPS fragment included ARRIAT, HDZIP2ATATAHB2, 5' UTR and GATA box. These motifs are reported to be crucial for high transcriptional activity. Additionally, in 1.5 kb deletion fragment regulatory motif INTRPSADB1 (associated with initiation of transcription) was deleted (Table 2). Deletion of 500 bp might have resulted in the removal of important motifs associated with high transcription activity which reduced promoter activity drastically. The decreased activity of 1.5 kb promoter might also be associated with negative regulatory elements in this region. Sequence analysis showed that 500 bp region away from TSS harbors TCA elements, TC rich repeats and several other motifs. Deletion of these elements might have resulted in decrease of the promoter activity to great extent. This suggests that regulatory elements present away from core promoter also play crucial role in modulating gene expression. The results also suggest that there might be some interaction between regulatory elements present in whole promoter region. Some of element may exert positive or negative influence on gene expression depending upon position and interaction with other promoter elements.

Several promoters of fiber related genes have shown specific expression in fiber tissue. Although constitutively expressed promoters have great application in biotechnology but they cannot be always used for transgene expression. Tissue specific promoters like newly isolated SPS promoter have great potential application for use to express genes in fiber tissues. Cotton fiber is mostly made up of cellulose which is synthesized through irreversible process during

secondary cell wall synthesis stage. SPS has role in regulation of cellulose synthesis by synthesizing sucrose which releases UDP glucose for cellulose synthesis, its promoter may be used to express fiber genes specifically during secondary cell wall synthesis in fibers. As fiber development in cotton is a complex process, studying the expression behavior of upstream regions of fiber genes may help to explore their role in this process. Regeneration in cotton is a time consuming process, biolistic transformation of intact plant organs is useful approach to study tissue specific expression of cotton promoters. Use of BT medium containing 10mg/l ascorbic acid prevents cotton tissues browning and improved bombardment results many times. Promoters are used on the basis desired expression but most of them are patented and cannot be used freely for commercialization of crops. In agriculture based country like Pakistan, there is need to explore novel promoters so that they may be used freely to generate transgenic crops. These novel promoters may also be used as components of multiple expression cassettes to avoid gene silencing caused by sequence homology of single or few promoters used to control expression of many genes. In the present study, a novel SPS promoter is isolated in cotton and analyzed through bioinformatics tools. The putative promoter region was analyzed for expression of GUS reporter gene through biolistic transformation. Further characterization of promoter revealed that it derived expression of reported gene at high rate in fibers. Therefore, it could be used to target fiber specific expression to alter fiber character.

## Materials and Methods

### Plant material

Cotton seed (CIM-707) were surface sterilized with 1% SDS and 0.1% HgCl<sub>2</sub> and washed with excessive sterile water. These seeds were grown in glass jars containing BT medium (Beasley and Ting, 1973) at 28 ± °C and 16/8hr day and night length. Cotton roots, leaves and stem were taken from invitro grown plants while fibers of different developmental ages were taken from greeen house grown plants.

### Retrieval and in silico analysis of SPS promoter

Nucleotide sequence of cotton sucrose phosphate synthase gene (JQ043231.1) was blast searched in HTGS. The sequence that matched with HTGS (Accession number 193940.1) of cotton was picked; its translation start codon and exact coding frame were identified through <http://www.expasy.ch>. Non-coding region identified through BLASTp was searched in data base to determine whether it was already patented or not. Finally, about 2 kb region was selected for promoter isolation and evaluation. Transcription start site for SPS promoter was determined using promoter prediction tool at ([www.fruitfly.org](http://www.fruitfly.org)). The selected nucleotide sequence was analyzed for presence of regulatory elements through plantCARE and transcription factors binding sites using plantPAN.

### Cloning of SPS promoter in TA cloning vector

For isolation of SPS promoter through PCR amplification, genomic DNA of fresh cotton leaves was used as template. A forward primer 5'TATGAGCTCGAGTGCAATATTGGCTTACC 3'with *SacI* site (underlined) and 5'GACAAGCTTGACTGAAAAAATCCACCAAAC 3' as reverse primer having *hindIII* site (both under lined) were

used. Annealing temperature, genomic DNA and Mg<sup>2+</sup> concentrations were optimized prior to cloning. PCR product was cloned into TA cloning vector. For amplification of 1.5 kb deletion fragment 5'GATGAGCTCTCAATTCATAAAGGGATAAATC 3'forward primers having *SacI* site and same reverses primer with *HindIII* site were used for amplification of 2 kb promoter. Deletion fragment of 1.5 kb was also cloned into TA cloning vector.

### Plant expression vector construction

A Plant expression vector pGR1 had 35S promoter fused to GUS gene followed by CaMV terminator. From pGR1 vector, 35S promoter was excised using *SacI* and *HindIII* enzymes. 2 kb SPS promoter was digested with *SacI* and *HindIII* enzymes and cloned into pGR1 to generate expression cassette. A 1.5 kb promoter fragment was also cloned in pGR1 upstream of GUS gene. For comparative analysis, a construct without promoter was also generated to be used as negative control. Vector pGR1 with GUS gene downstream of 35S promoter was used as positive control in bombardment experiments.

### Transient GUS assay

Histochemical GUS assay for 2 kb SPS promoter and its 1.5 kb fragment was conducted in cotton tissues (Jefferson et al.,1987). For each bombardment,500µl gold particles of 1 micron size were used. Microcarrier gold particles were washed with ethanol followed by single washing with sterile water. These gold particles were coated with Plasmid DNA (1µg/µl) of each construct. Cotton tissues to be bombarded were placed on petriplates containing BT medium in such way to expose maximum surface area for bombardment. All explants were bombarded at 9 cm target distance and 27 mmHg pressure with 1100 psi rupture disks. Plasmids having 35S promoter and promoterless construct were also bombarded using same conditions. Bombarded tissues were placed at 28°C for 24 hrs. All tissues were submerged in staining buffer containing 0.1M X-Gluc and incubated at 37 °C in dark till appearance of blue colour. Histochemically stained tissues were decanted and 70% ethanol was added to stop reaction as well as for removal of chlorophyll from green tissues. Tissues were photographed with digital camera attached with microscope.

## Conclusions

In conclusion, the above study contributes to understand the role of SPS gene in fiber development. Various regulatory motifs necessary for promoter activity have been found in upstream region of SPS gene. Cotton SPS promoter has found to highly active in the fiber tissues. So, it may be used to express gene products specifically in developing fibers. It may also be used as component of multiple expression cassettes to express beneficial genes in cotton to improve fiber quality traits without IPR issues. These novel promoters may also be used as components of multiple expression cassettes to avoid gene silencing caused by sequence homology of single or few promoters used to control expression of many genes.

## Acknowledgements

The research grant for present study was provided by the Ministry of Food and Agriculture (Min. FA), Government of Pakistan and Higher Education Commission, Pakistan.

## References

- Amor Y, Haigler CH, Johnson S, Wainscott M, Delmer DP (1995) A membrane associated form of Sus and its potential role in synthesis of cellulose and callose in plants. *Proc Natl Acad Sci USA*. 92: 9353-9357.
- Beasley CA, Ting IP (1973) Effects of plant growth substances on in vitro fiber development from fertilized cotton ovules. *Am J Bot*. 60: 130-139.
- Daiz-De-leon F, Klotz KL, Legrimini LM (1993) Nucleotide sequence of the tobacco (*Nicotiana tabacum*) onion peroxide gene. *Plant Physiol*. 101:1117-1118.
- Daraselia, ND, Tarchevskaya S, Notita TO (1996) The promoter of tomato 3-hydroxyl-3-methyl glutaryl co-enzyme reductase gene 2 has unusual regulatory elements that direct high-level of expression. *Plant Physiol*. 112:727-733.
- Delaney SK, OrfordSJ, Harris MM, Timmis JN (2007) The fiber specificity of the cotton FSItp4 gene promoter is regulated by an AT-rich promoter region and the AT-hook transcription factor GhAT1. *Plant Cell Physiol*. 48: 1426-1437.
- Deng F, TuL, Tan J, Li Y, Nie Y, Zhang X (2012) Gbpdf1 is involved in cotton fiber initiation via the core cis-element HDZIP2ATATHB21. *Plant Physiol*. 158:890-904.
- Guerineau F, Bendia M, Zhou DX (2003) A jasmonate responsive element in the *thaliana* VSP1 promoter. *J Exp Bot*. 54:1153-1162.
- Haigler CH, Singh B, Zhang D, Hwang S, Wu C, Cai WX, Hozain M, Kang W, Kiedaisch B, Strauss RE, Hequet EF, Wyatt BG, Jividen GM, Holaday AS (2007) Transgenic cotton over-producing spinach sucrose phosphate synthase showed enhanced leaf sucrose synthesis and improved fiber quality under controlled environmental conditions. *Plant Mol Biol*. 63(6): 815-32.
- Hieno A, Naznin HA, Hyakumachi M, Sakurai T, Tokizawa M, Koyama H, Sato N, Nishiyama T, Hasebe M, Zimmer AD, Dang D, Reski R, Rensing S, Obokata J, Yamamoto YY. (2014) ppdb: plant promoter database version 3.0. *Nucleic Acids Res*. 42, 1188-1192.
- Hobo T, Asada M, Kowyama Y, Hattori T (1999) ACGT containing abscisic acid response element (ABRE) and coupling element 3 (CE3) are functionally equivalent. *Plant J*. 19: 679-689.
- Huang GQ, Gong SY, Xu WL, Li W, Li P, Zhang CJ, Li DD, Zheng Y, Li FG, Li XB (2013) A fasciclin-like arabinogalactan protein, GhFLA1, is involved in fiber initiation and elongation of cotton. *Plant Physiol*. 161(3):1278-1290.
- Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J*. 6:3901-3907.
- Lunn JE, MacRae E (2003) New complexities in the synthesis of sucrose. *Curr Opin Plant Biol*. 6: 208-214.
- Machado A, Wu Y, Yang Y, Llewellyn DJ, Dennis ES (2009) The MYB transcription factor GhMYB25 regulates early fibre and trichome development. *Plant J*. 59:52-62.
- Mary E, Foster R, Chua NH (1992) Sequences flanking the hexameric G-box core CACAGTG affect the specificity of protein binding. *Plant cell*. 4:485-496.
- Manevski A, Bertoni G, Bardet C., Tremousaygue D, Lescure B (2000) In synergy with various cis-acting elements, plant interstitial telomere motifs regulate gene expression in *Arabidopsis* root meristems. *FEBS Lett*. 483 (1): 43-46.
- Nover L, Bharti K, Döring KP, Mishra SK, GAnguli A, Scharf KD (2001) *Arabidopsis* and the heat shock transcription factors world: how many heat stress transcription factors do we need?. *Cell stress chap*. 6:177-189.
- Ohgishi M, Oka A, Morelli G, Ruberti I, Aoyama T (2001) Negative autoregulation of the *Arabidopsis* homeobox gene ATHB-2. *Plant J*. 25: 389-398.
- Pu L, Li Q, Fan X, Yang W, Y. Xue Y (2008) The R2R3MYB transcription factors GhMYB109 is required for cotton fiber development. *Genetics*. 180(2): 811-820.
- Rushton PJ, Torres JT, Parniske M, Wernert P, Hahlbrock K, Somssich IE (1996) Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of Parsley PR1 gene. *EMBO J*. 15:5690-5700.
- Winter H, Huber SC (2000) Regulation of sucrose metabolism in higher plants: localization and regulation of activity of key enzymes. *CRC Crit Rev Plant Sci*. 19: 31-67.
- Wu Y, Xu W, Huang G, Gong S, Li J, Qin Y, Li X (2009) Expression and localization of GhH6L, a putative classical arabinogalactan protein in cotton (*Gossypium hirsutum*). *Acta Biochem Biophys Sin*: 1-9.
- Wu AM, Lu SY, Liu JY (2007) Functional analysis of cotton Glucuronosyl transferase promoter in transgenic tobaccos. *Cell Res*. 17: 174 -183.
- Yonekura MN, Aoki T, Hirose K, Onai M, Ishiura M, Okamura M, Ohsugi R, Ohto C (2013) The promoter activities of sucrose phosphate synthase genes in rice, OsSPS1 and OsSPS11 are controlled by light and circadian clock, but not by sucrose. *Front Plant Sci*. 4: 31.
- Zhou DX (1999) Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. *Trends in Plant Sci*. 4: 210-214.