



Genetic and epigenetic status of triple exotic consanguinity cotton introgression lines

S.P. He, J.L. Sun and X.M. Du

State Key Laboratory of Cotton Biology, Cotton Research Institute,
Chinese Academy of Agricultural Sciences, Anyang, Henan, China

Corresponding author: X.M. Du
E-mail: duxm@cricaas.com.cn

Genet. Mol. Res. 10 (4): 4063-4072 (2011)
Received April 26, 2011
Accepted August 30, 2011
Published October 3, 2011
DOI <http://dx.doi.org/10.4238/2011.October.3.1>

ABSTRACT. Introgression lines are some of the most important germplasm for breeding applications and other research conducted on cotton crops. The DNA methylation level among 10 introgression lines of cotton (*Gossypium hirsutum*) and three exotic parental species (*G. arboreum*, *G. thurberi* and *G. barbadense*) were assessed by methylation-sensitive amplified polymorphism (MSAP) technology. The methylation level in the introgression lines ranged from 33.3 to 51.5%. However, the lines PD0111 and PD0113 had the lowest methylation level (34.6 and 33.3%, respectively) due to demethylation of most non-coding sequences. Amplified fragment length polymorphism (AFLP) was used to evaluate the genetic polymorphism in the cotton introgression lines. A high degree of polymorphism was observed in all introgression lines (mean 47.2%) based on AFLP and MSAP analyses. This confirmed the effects of genetic improvement on cotton introgression lines. The low methylation varieties, PD0111 and PD0113 (introgression lines), clustered outside of the introgression lines based on MSAP data, which was incongruent with an AFLP-based dendrogram. This phenomenon could be caused by environmental changes or introgression of exotic DNA fragments.

Key words: *Gossypium*; Introgression lines; DNA methylation; Polymorphism

INTRODUCTION

Upland cotton (*Gossypium hirsutum* L.) is a new tetraploid species formed by two diploids a million years ago (Beasley, 1940; Wendel et al., 1992). Nowadays it has become the dominant fiber in production throughout the modern world. However, the genetic bottleneck and limited diversity in upland cotton has already been pointed out (Iqbal et al., 2001). To improve and broaden the genetic diversity of upland cotton, breeders have attempted for decades to transfer the elite traits of wild species into cultivars through interspecific hybridization and repeated backcrossing (Culp and Harrell, 1973). Over the last 50 years, many new cotton varieties and germplasm lines, called “introgression lines”, have been evolved. Most of the cotton introgression lines displayed elite traits in contrast to their parents. However, due to the complex genetic background, it is difficult to determine the diversity relationship among introgression lines through traditional morphological and agricultural trait analysis. The previous study based on SSR markers and agricultural traits showed that molecular markers were a reliable approach to investigate the genetic diversity of cotton introgression lines (Pang et al., 2006). Distinguished from genetics, epigenetics refers to the heritable changes in phenotype or level of gene expression without changes in DNA sequences (Bender, 2004). As the best-described epigenetic mechanism (Rapp and Wendel, 2005), DNA methylation was found to be involved in nearly all fundamental cellular processes and regulation (Bird, 2007). The study of epigenetic diversity in cotton introgression lines could provide an understanding of the elite trait introgression mechanism. Methylation-sensitive amplified polymorphism (MSAP) could widely investigate the DNA methylation status of CCGG sites of the genome. It has proven to be a highly efficient approach to study DNA methylation in *Arabidopsis* (Cevera et al., 2002), *Gossypium* (Keyte et al., 2006), *Brassica* (Salmon et al., 2008) and *Hordeum* (Li et al., 2008).

Based on the SSR evaluation of 160 cotton introgression lines (Pang et al., 2006), 10 elite cotton introgression lines were chosen with same exotic consanguinity. The objectives of this study were to 1) assess the cytosine methylation level in cotton introgression lines; 2) compare the relationship between genetic and epigenetic (DNA methylation) polymorphism; and 3) reveal what kinds of sequences are subject to DNA methylation polymorphism in cotton introgression lines.

MATERIAL AND METHODS

Plant material and DNA isolation

A total of 14 germplasm lines, including 10 introgression lines (*G. hirsutum*) with *G. thurberi*, *G. arboreum* and *G. barbadense* consanguinity, and 3 exotic parental species (*G. thurberi*, *G. arboreum* and *G. barbadense*) were analyzed. TM-1 (*G. hirsutum*) of cultivated upland cotton was used as a check line (Table 1). Plants were grown in the experimental field of the Cotton Research Institute, Chinese Academy of Agricultural Sciences (CRI, CAAS), Anyang, China, during the year of 2007 and 2008. All plants were maintained with great care and strictly selfed under natural growing conditions. Genomic DNA was extracted from seedling leaves following the method described by Paterson et al. (1993).

Table 1. List of germplasm used in this study.

No.	Germplasm Name	Description
1	<i>G. arboreum</i>	Exotic parent
2	<i>G. thurberi</i>	Exotic parent
3	<i>G. barbadense</i>	Exotic parent
4	TM-1	<i>G. hirsutum</i> check line
5	F	Introgression lines
6	FJA	Introgression lines
7	AC239	Introgression lines
8	AC241	Introgression lines
9	PD3246	Introgression lines
10	PD3249	Introgression lines
11	PD2164	Introgression lines
12	PD0109	Introgression lines
13	PD0111	Introgression lines
14	PD0113	Introgression lines

Protocols of AFLP and MSAP

Standard AFLP analysis with minor modification for higher yield of PCR production was performed to assess the genetic variation of introgression lines (Vos et al., 1995). For digestion, 500 ng purified genomic DNA was prepared. A 30 μ L mixture containing 3 μ L 10X buffer Tango and 5 U *EcoRI* (MBI Fermentas, Burlington, Canada) was digested at 37°C for 3 h. Then 5 U *MseI* was added to the mixture and incubated at 65°C for 3 h. In the digestion mixture, 1 μ L *EcoRI* adaptor, 2 μ L *MseI* adaptor, 1.5 μ L 10X ligase buffer (Promega, Fitchburg, USA) and 5 U T4 ligase were added and ligated at 16°C overnight. After that the DNA was completely digested. The digested DNA was diluted into fivefold solution for preamplification. The preamplification volume of 50 μ L included 5 μ L diluted digested DNA template, 2 U Taq polymerase, 1 U Taq polymerase buffer, 50 ng of each primer *EcoRI*+A and *MseI*+C, and 0.1 mM of each dNTP. PCR reactions were carried out on the thermocycler (Biometra, Goettingen, Germany) for 22 cycles of 30 s at 94°C, 30 s at 56°C, and 1 min at 72°C. Then 3 μ L PCR production was electrophoresed on 1.2% agarose gel to verify the DNA smear in the range 100-200 bp.

Preamplification production was diluted 20-fold for selective amplification. This step was performed in a total of 15 μ L volume containing 3.8 μ L diluted preamplification production, 50 ng of each selective primer, 2 U Taq polymerase, 1 U Taq polymerase buffer and 0.1 mM of each dNTP. Touchdown PCR was used for selective amplification. This program comprised 12 cycles at 94°C for 30 s, 65°C for 30 s decreasing 0.7°C per cycle, 72°C for 1 min and then 22 cycles at 94°C for 30s, 72°C for 1 min and final amplification at 72°C for 10 min. A 5 μ L volume of denatured PCR product was separated on 8% denaturing polyacrylamide gel for 1.5 h at 75 W. The gel was stained by the silver stain method.

The MSAP procedure was used according to the protocol of Xu et al. (2000). The reaction components and procedures were the same as for the AFLP protocol, except for a substitution of the pair of isoschizomers *HpaII* and *MspI* for *EcoRI*. All the adapters and primer sequences used in this work are shown in Supplementary Table 1.

Gel pattern scoring and analysis

The presence or absence of the AFLP band on the gel was scored as 1 or 0, respectively.

A matrix of data based on the Jaccard (1908) coefficient of similarity was produced. The MSAP bands for polymorphism study were scored by a criterion described by Keyte et al. (2006) with several modifications. The methylation status of one DNA was determined by two gel lanes due to a single DNA being digested by two restriction enzyme combinations. The presence of a band in the *HpaII* lane and the absence in the *MspI* lane (H/M = +/-), or presence of a band in the *MspI* lane and the absence in the *HpaII* lane (H/M = -/+) was scored as 1, which represented methylated status at this site. The presence of bands in both lanes simultaneously (H/M = +/+) was scored as 0, which represented unmethylated status at this site; Absence of a band in both lanes was scored as?, which represented the unknown status at this site (Table 2). All dendrograms were constructed by using the NTSYSpc 2.2n software (Rohlf, 2009; Exeter Software).

Table 2. Different gel status.

Sequence status	Gel status		Site status	Scored as																
	<i>EcoRI</i> + <i>MspI</i>	<i>EcoRI</i> + <i>HpaII</i>																		
<table border="1" style="display: inline-table; margin-right: 10px;"> <tr><td>C</td><td>C</td><td>G</td><td>G</td></tr> <tr><td>G</td><td>G</td><td>C</td><td>C</td></tr> </table>	C	C	G	G	G	G	C	C	■		Methylated	1								
C	C	G	G																	
G	G	C	C																	
<table border="1" style="display: inline-table; margin-right: 10px;"> <tr><td>C</td><td>C</td><td>G</td><td>G</td></tr> <tr><td>G</td><td>G</td><td>C</td><td>C</td></tr> </table> or <table border="1" style="display: inline-table; margin-left: 10px;"> <tr><td>C</td><td>C</td><td>G</td><td>G</td></tr> <tr><td>G</td><td>G</td><td>C</td><td>C</td></tr> </table>	C	C	G	G	G	G	C	C	C	C	G	G	G	G	C	C	■	■	Methylated	1
C	C	G	G																	
G	G	C	C																	
C	C	G	G																	
G	G	C	C																	
<table border="1" style="display: inline-table; margin-right: 10px;"> <tr><td>C</td><td>C</td><td>G</td><td>G</td></tr> <tr><td>G</td><td>G</td><td>C</td><td>C</td></tr> </table>	C	C	G	G	G	G	C	C	■	■	Unmethylated	0								
C	C	G	G																	
G	G	C	C																	
<table border="1" style="display: inline-table; margin-right: 10px;"> <tr><td>C</td><td>C</td><td>G</td><td>G</td></tr> <tr><td>G</td><td>G</td><td>C</td><td>C</td></tr> </table> or no cutting site	C	C	G	G	G	G	C	C			Unknown	?								
C	C	G	G																	
G	G	C	C																	

The black boxes indicate methylated cytosine.

Obtaining sequences of MSAP fragments

To evaluate more detailed information on DNA methylation polymorphism fragments, eight reliable and repeatable MSAP fragments were recovered for sequencing. The fragments were cut carefully from gel and mashed in a 1.5 mL centrifuge tube. Then 20 μ L ddH₂O was added and heated in boiling water for 10 min. 5 μ L of this solution was used as a template for reamplification with the previous primer combinations following the protocol of selective amplification. The fragments were then separated on 0.8% agarose, extracted and ligated into the T-vector (Promega). The products were sequenced by Shanghai Sangon Biological Engineering Technology & Services Co Ltd. Finally, the sequences were BLASTed against the nr database of NCBI.

RESULTS

Methylation level among different introgression lines

Both *HpaII* and *MspI* recognize the same sequence 5-CCGG-3, but have differential sensitivity to DNA methylation. *HpaII* is inactive when any of the two cytosines is fully methylated, but cuts the hemi-methylated sequence; whereas, *MspI* cuts 5-C^mCCGG-3, but not 5-^mCCGG-3. Therefore we could distinguish the different status of DNA methylation in the sequence CCGG by the distinct band pattern on the gel. A total of 4478 legible and reproduc-

ible bands were generated using 26 primer-pair combinations in the MSAP analysis. All bands were divided into non-methylated CCGG sites and methylated CCGG sites. The methylated CCGG sites were further divided into fully methylated (internal C) and hemi-methylated sites (external C) (Table 3). The two diploid species, *G. thurberi* and *G. arboreum*, had higher levels of DNA methylation than all tetraploids, reaching 60.3 and 61.5%, respectively. Whereas two upland cotton lines (PD0111 and PD0113) had shown relatively lower methylation levels than other tetraploid cotton, 34.6 and 33.3%, respectively. The methylation levels ranged from 46.8 to 51.5% in the other introgression lines. There were more fully methylated sites detected at the internal cytosine as compared to hemi-methylated sites at external cytosine in all accessions except PD0111 and PD0113. The average of fully methylated and hemi-methylated sites was 27.0 and 22.1%, respectively in all accessions. The percentage of fully methylated sites was nearly twice that of hemi-methylated sites in both *G. thurberi* and *G. arboreum*; however, in PD0111 and PD0113 the percentage of fully methylated sites was nearly half that of hemi-methylated sites. This implies that the lower methylation levels in PD0111 and PD0113 were caused by the decreasing full methylation sites. Compared to tetraploid cotton, the methylation level of diploid cotton was higher and resulted from increasing full methylation.

Table 3. The percentage of non-methylated and methylated sites.

Accession	Total number of scored sites	Non-methylated CCGG sites (%)	Methylated CCGG sites (%)		
			Fully methylated sites (internal C)	Hemi-methylated sites (external C)	Total methylated sites (methylation level)
<i>G. thurberi</i>	179	39.7	43.6	16.7	60.3
<i>G. arboreum</i>	187	38.5	38.5	23.0	61.5
<i>G. barbadense</i>	333	53.2	26.7	20.1	46.8
TM-1	348	52.0	25.9	22.1	48.0
F	345	49.1	27.2	23.2	50.4
FJA	340	50.6	26.5	22.9	49.4
AC239	346	50.9	26.9	22.2	49.1
AC241	346	50.2	27.5	22.3	49.8
PD2164	347	48.7	28.8	22.5	51.3
PD3246	345	48.7	27.5	23.8	51.3
PD3249	346	49.7	27.2	23.1	50.3
PD0109	342	48.5	28.4	23.1	51.5
PD0111	338	65.4	12.1	22.5	34.6
PD0113	336	66.7	11.0	22.3	33.3
Average		50.9	27.0	22.1	49.1

The DNA sequence and methylation polymorphism in cotton introgression lines

For the comparison of genetic and epigenetic relationships of all accessions based on AFLP and MSAP markers, two dendrograms were constructed based on the Nei and Li (1979) coefficient by the NTSYSpc software. The dendrogram constructed with AFLP data reflected the nucleotide polymorphism in the introgression lines; the similarity coefficient ranged from 0.40 to 0.86 in all upland cottons (Table 4). Two diploid species (*G. arboreum* and *G. thurberi*) were first separated from the group, and then *G. barbadense*. The F line was excluded from other upland cotton varieties with its own branch. Unexpectedly, TM-1 was clustered with PD0109 as a subgroup in the upland cotton group; all other introgression lines composed another group (Figure 1A). A dendrogram based on MSAP data reflected the epigenetic diversity, and the similarity

coefficient ranged from 0.03 to 0.99 (Table 4). PD0113 and PD0111 were first clustered together as an individual group which presented significant dissimilarity with other upland cotton, whereas the other introgression lines were clustered together and distinguished well from diploid cotton and *G. barbadense*, even TM-1 was well separated from introgression lines (Figure 1B).

Table 4. AFLP- and MSAP-based similarities of selected accessions.

	<i>G. thu</i>	<i>G. arb</i>	<i>G. bar</i>	TM-1	F	FJA	AC239	AC241	PD2164	PD3246	PD3249	PD0109	PD0111	PD0113
<i>G. thu</i>		0.66	0.54	0.46	0.61	0.51	0.54	0.51	0.41	0.63	0.61	0.49	0.59	0.54
<i>G. arb</i>	0.71		0.59	0.46	0.56	0.46	0.44	0.46	0.37	0.54	0.51	0.54	0.44	0.44
<i>G. bar</i>	0.76	0.77		0.54	0.63	0.44	0.66	0.59	0.49	0.32	0.54	0.56	0.56	0.56
TM-1	0.84	0.80	0.89		0.51	0.66	0.73	0.66	0.71	0.49	0.56	0.73	0.68	0.63
F	0.84	0.79	0.87	0.88		0.61	0.59	0.71	0.56	0.59	0.71	0.59	0.59	0.63
FJA	0.89	0.79	0.90	0.89	0.92		0.63	0.80	0.76	0.63	0.71	0.59	0.73	0.78
AC239	0.85	0.82	0.88	0.92	0.93	0.94		0.68	0.73	0.41	0.63	0.61	0.76	0.66
AC241	0.86	0.80	0.88	0.91	0.92	0.93	0.99		0.85	0.68	0.80	0.59	0.68	0.83
PD2164	0.88	0.78	0.91	0.90	0.92	0.93	0.97	0.98		0.68	0.76	0.59	0.68	0.83
PD3246	0.79	0.76	0.85	0.88	0.95	0.92	0.93	0.94	0.93		0.73	0.56	0.61	0.66
PD3249	0.83	0.86	0.87	0.91	0.93	0.91	0.97	0.97	0.95	0.95		0.54	0.68	0.78
PD0109	0.88	0.76	0.88	0.88	0.94	0.92	0.93	0.94	0.96	0.97	0.95		0.71	0.56
PD0111	0.30	0.20	0.22	0.21	0.17	0.18	0.20	0.25	0.24	0.19	0.20	0.19		0.71
PD0113	0.14	0.12	0.08	0.07	0.03	0.03	0.06	0.06	0.09	0.03	0.06	0.06	0.50	

AFLP-based similarity coefficients are shown in the upper matrix; MSAP-based similarity coefficient are shown in the lower matrix.

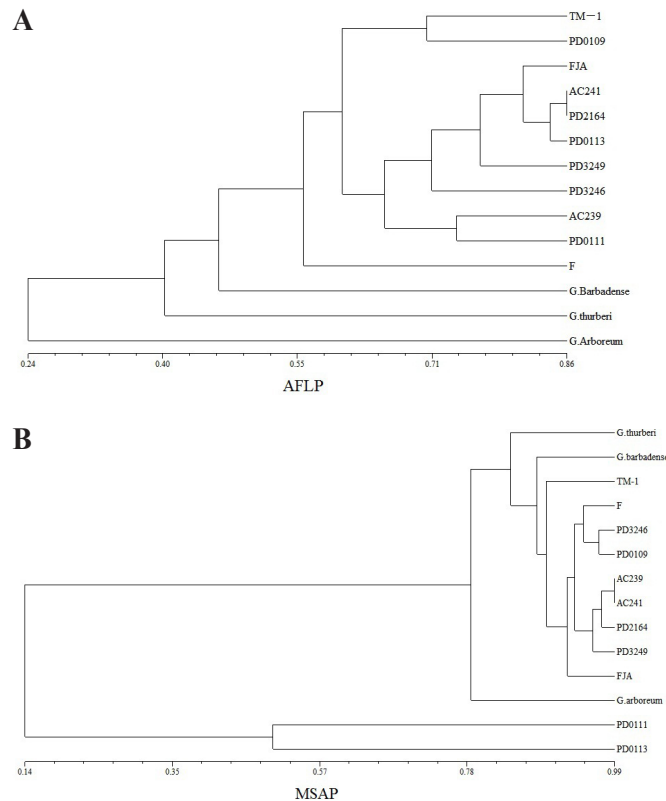


Figure 1. Dendrogram of introgression lines and their parents based on AFLP (A) and MSAP (B).

Sequence information of MSAP fragments

Legible and repeatable fragments which showed polymorphism in MSAP were sequenced, and the 1/2 similar sequences were identified at NCBI. Two fragments from PD0113 and PD0111 were highly homologous with a *gypsy*-type retrotransposon (Z3-245 and Z11-164), respectively (Table 5). This showed the great difference in genetics and epigenetics of PD0113 and PD0111 from other introgression lines, and also from upland cotton, due to the effect of the *gypsy*-type retrotransposon during the multiple crosses and backcrosses among the species.

Table 5. The BLAST search results for sequences of DNA methylation pattern variation in this study.

Sequence ID	Primer combination	Sample*	Sequence similarity	Fragment size (bp)
Z3-245	E-ACA+H/M-TCAC	PD0113/H	<i>G. hirsutum gypsy</i> -type retroelement, partial sequence (AY395704.1a)	245
Z4-160	E-ACA+H/M-TCAA	AC239/H	No similarity	160
Z5-220	E-ACA+H/M-TCAA	PD0111/H	Drypetes madagascariensis isolate RS_PHY16 PHYC (PHYC) gene, exon 1 and partial cds (AY579845.1)	220
Z6-148	E-ACA+H/M-TCAA	<i>G. barbadense</i> /M	No similarity	148
Z10-212	E-AAA+H/M-TCAG	PD3246/H	No similarity	212
Z11-164	E-AAG+H/M-TCAA	PD0111/H	<i>G. hirsutum gypsy</i> -type retroelement, partial sequence (AY395704.1)	164
Z12-89	E-AAG+H/M-TCAC	TM-1/H	IPPGHZ0085 Cotton fiber and embryo Lambda Zap Express Library <i>G. hirsutum</i> cDNA 5- similar to (AL022223) receptor protein kinase-like protein (<i>Arabidopsis thaliana</i>) mRNA sequence (AW587497.1)	89
Z23-246	E-AGT+H/M-TCAA	PD0109/H	No similarity	246

*M indicates that DNA was digested by *EcoRI*+*MspI*; H indicates that DNA was digested by *EcoRI*+*HpaII*.

DISCUSSION

DNA methylation is considered a very important regulation mechanism for gene expression. Cytosine methylation levels in plants displayed dynamic variation, such as tissue-specific (Zhang et al., 2007) and developmental stage-specific variation (Ruiz-García et al., 2005). In the present study, the DNA methylation level of *Gossypium* represented variation not only in different species, but also among the introgression lines. Although most of the introgression lines displayed stable DNA methylation levels, PD0111 and PD0113 exhibited unexpectedly lower levels than others. Evidence from earlier studies shows that changes in DNA methylation levels can lead to the abnormal development of plants (Finnegan et al., 1996; Kakutani et al., 1996; Fieldes et al., 2005). However, no obvious phenotypic variation was detected among the three introgression lines (PD0109, PD0111 and PD0113) which were selected from the same line (data not shown). This implies that DNA methylation alternation activated the non-coding sequences or impacted the gene expression mechanism which did not appear in the phenotype. Further sequencing results of the methylation polymorphism fragments confirmed that most of the DNA methylation alteration happened at the repetitive sequences or non-coding sequences. However, the movement of transposable elements induced by DNA methylation alteration played an important role in the regulation of gene expression (Jin et al., 2008; Xu et al., 2009). The potential impacts, especially methylation alternation of transposable elements should be investigated on PD0111 and PD0113. Therefore, the methyla-

tion details shown in Table 3 demonstrate that DNA methylation was decreased in PD0111 and PD0113 due to decreasing fully methylated sites. On the contrary, hemi-methylation of external cytosines displayed high instability in three rice introgression lines (Dong et al., 2006). The underlying mechanism is still unknown; nevertheless, extensive variation of DNA methylation status exists in different species and different varieties in the same species.

Upland cotton was considered to lack polymorphism in DNA sequences (Iqbal et al., 2001). The high genetic diversity shown in this work indicated that introgression lines displayed a higher polymorphism level than common cotton varieties. This evidence clearly shows the expanding effect of cotton diversity by introgression. Both our result and another study on tetraploid cotton (Keyte et al., 2006) show wide DNA methylation diversity in cotton. Although both AFLP and MSAP could distinguish *Gossypium* species, there were some differences in the clusters. In the MSAP-based dendrogram, PD0111 and PD0113 failed to cluster with other introgression lines. Our previous results demonstrated that this might be caused by their abnormal DNA methylation status. A similar phenomenon was reported in a previous study on *Brassica* (Salmon et al., 2008). This could be explained by the following two possibilities: First, in some species, epigenetic variation is more environmentally sensitive than genetic (Lira-Medeiros et al., 2010). The experimental fields were changed in 2007, which led to wide DNA methylation variation due to different environments. Second, epigenetic variation can be caused by “introgression”. This has been proven extensively in many species after interspecific hybridization (Shaked et al., 2001; Lukens et al., 2006; Dong et al., 2006). Therefore, the formation of excellent traits in introgression lines was not only the result of genetic impact but also due to epigenetics. This could be helpful to explore the potential abilities of germplasm in breeding.

ACKNOWLEDGMENTS

Research supported by the National Key Technology R & D Program of China (#2006BAD13B04), and the Crop Germplasm Preservation Project of Agricultural Ministry (#NB09-2130135-36).

REFERENCES

- Beasley JO (1940). The origin of American tetraploid *Gossypium* species. *Am. Nat.* 74: 285-286.
- Bender J (2004). DNA methylation and epigenetics. *Annu. Rev. Plant Biol.* 55: 41-68.
- Bird A (2007). Perceptions of epigenetics. *Nature* 447: 396-398.
- Cervera MT, Ruiz-Garcia L and Martinez-Zapater JM (2002). Analysis of DNA methylation in *Arabidopsis thaliana* based on methylation-sensitive AFLP markers. *Mol. Genet. Genomics* 268: 543-552.
- Culp W and Harrell DC (1973). Breeding methods for improving yield and fiber quality of upland cotton (*Gossypium hirsutum* L.). *Crop Sci.* 13: 686-689.
- Dong ZY, Wang YM, Zhang ZJ, Shen Y, et al. (2006). Extent and pattern of DNA methylation alteration in rice lines derived from introgressive hybridization of rice and *Zizania latifolia* Griseb. *Theor. Appl. Genet.* 113: 196-205.
- Fieldes MA, Schaeffer SM, Krech MJ and Brown JC (2005). DNA hypomethylation in 5-azacytidine-induced early-flowering lines of flax. *Theor. Appl. Genet.* 111: 136-149.
- Finnegan EJ, Peacock WJ and Dennis ES (1996). Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. *Proc. Natl. Acad. Sci. U. S. A.* 93: 8449-8454.
- Iqbal MJ, Reddy OUK, El-Zik KM and Pepper AE (2001). A genetic bottleneck in the “evolution under domestication” of upland cotton *Gossypium hirsutum* L. examined using DNA fingerprinting. *Theor. Appl. Genet.* 103: 547-554.
- Jaccard P (1908). Nouvelles recherches sur la distribution florale. *Bull. Soc. Vaud. Sci. Nat.* 44: 270.

- Jin H, Hu W, Wei Z, Wan L, et al. (2008). Alterations in cytosine methylation and species-specific transcription induced by interspecific hybridization between *Oryza sativa* and *O. officinalis*. *Theor. Appl. Genet.* 117: 1271-1279.
- Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, et al. (1996). Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. *Proc. Natl. Acad. Sci. U. S. A.* 93: 12406-12411.
- Keyte AL, Percifield R, Liu B and Wendel JF (2006). Intraspecific DNA methylation polymorphism in cotton (*Gossypium hirsutum* L.). *J. Hered.* 97: 444-450.
- Li Y, Shan X, Liu X, Hu L, et al. (2008). Utility of the methylation-sensitive amplified polymorphism (MSAP) marker for detection of DNA methylation polymorphism and epigenetic population structure in a wild barley species (*Hordeum brevisubulatum*). *Ecol. Res.* 23: 927-930.
- Lira-Medeiros CF, Parisod C, Fernandes RA, Mata CS, et al. (2010). Epigenetic variation in mangrove plants occurring in contrasting natural environment. *PLoS. One.* 5: e10326.
- Lukens LN, Pires CJ, Leon E, Vogelzang R, et al. (2006). Patterns of sequence loss and cytosine methylation within a population of newly resynthesized *Brassica napus* allopolyploids. *Plant Physiol.* 140: 336-348.
- Pang C, Du X and Ma Z (2006). Evaluation of the introgressed lines and screening for elite germplasm in *Gossypium*. *Chin. Sci. Bull.* 51: 304-312.
- Paterson AH, Brubaker CL and Wendel JF (1993). A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol. Biol. Rep.* 11: 122-127.
- Rapp RA and Wendel JF (2005). Epigenetics and plant evolution. *New Phytol.* 168: 81-91.
- Rohlf FJ (2009). NTSYSpc: Numerical Taxonomy System. Ver. 2.21c. Exeter Software: Setauket, New York.
- Ruiz-Garcia L, Cervera MT and Martinez-Zapater JM (2005). DNA methylation increases throughout *Arabidopsis* development. *Planta* 222: 301-306.
- Salmon A, Clotault J, Jenczewski E, Chable V, et al. (2008). *Brassica oleracea* displays a high level of DNA methylation polymorphism. *Plant Sci.* 174: 61-70.
- Shaked H, Kashkush K, Ozkan H, Feldman M, et al. (2001). Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* 13: 1749-1759.
- Vos P, Hogers R, Bleeker M, Reijans M, et al. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23: 4407-4414.
- Wendel JF, Brubaker CL and Percival AE (1992). Genetic diversity in *Gossypium hirsutum* and the origin of upland cotton. *Am. J. Bot.* 79: 1291-1310.
- Xu M, Li XQ and Korban SS (2000). AFLP-Based detection of DNA methylation. *Plant Mol. Biol. Rep.* 18: 361-368.
- Xu Y, Zhong L, Wu X, Fang X, et al. (2009). Rapid alterations of gene expression and cytosine methylation in newly synthesized *Brassica napus* allopolyploids. *Planta* 229: 471-483.
- Zhang MS, Yan HY, Zhao N, Lin XY, et al. (2007). Endosperm-specific hypomethylation, and meiotic inheritance and variation of DNA methylation level and pattern in sorghum (*Sorghum bicolor* L.) inter-strain hybrids. *Theor. Appl. Genet.* 115: 195-207.

SUPPLEMENTARY MATERIAL

Supplementary Table 1. All primers used in this study.

AFLP	
Adapters	Sequence (5'-3')
<i>Eco</i> RI-1 adapter	5-CTCGTAGACTGCGTACC-3
<i>Eco</i> RI-2 adapter	5-AATTGGTACGCAGTCCAT-3
<i>Mse</i> I-1 adapter	5-GACGATGAGTCCTGAG-3
<i>Mse</i> I-2 adapter	5-TACTCAGGACTCAT-3
Pre-selective primers	
E+A	5-GACTGCGTACCAATTCA-3
M+C	5-GATGAGTCCTGAGTAAC-3
Selective primers	
M-CAA	5-GATGAGTCCTGAGTAACAA-3
M-CAC	5-GATGAGTCCTGAGTAACAC-3
M-CAG	5-GATGAGTCCTGAGTAACAG-3
M-CTA	5-GATGAGTCCTGAGTAAC-3
M-CAT	5-GATGAGTCCTGAGTAACAT-3
M-CCA	5-GATGAGTCCTGAGTAACCA-3
E-AGG	5-GACTGCGTACCAATTCAGG-3
E-ACG	5-GACTGCGTACCAATTCACG-3
E-AAC	5-GACTGCGTACCAATTC AAC-3
E-ACA	5-GACTGCGTACCAATTCACA-3
E-AGC	5-GACTGCGTACCAATTCAGC-3
E-ACT	5-GACTGCGTACCAATTC ACT-3
E-AAG	5-GACTGCGTACCAATTC AAG-3
E-ACC	5-GACTGCGTACCAATTC ACC-3
E-AAA	5-GACTGCGTACCAATTC AAA-3
E-AAT	5-GACTGCGTACCAATTC AAT-3
E-ATA	5-GACTGCGTACCAATTC ATA-3
E-ATT	5-GACTGCGTACCAATTC ATT-3
E-ATC	5-GACTGCGTACCAATTC ATC-3
E-ATG	5-GACTGCGTACCAATTC ATG-3
E-AGA	5-GACTGCGTACCAATTC AGA-3
E-AGT	5-GACTGCGTACCAATTC AGT-3
MSAP	
Adapters and primers	
<i>Eco</i> RI-1 adapter	5-CTCGTAGACTGCGTACC-3
<i>Eco</i> RI-2 adapter	5-AATTGGTACGCAGTC-3
<i>Hpa</i> II/ <i>Msp</i> I-1 adapter	5-GATCATGAGTCCTGCT-3
<i>Hpa</i> II/ <i>Msp</i> I-2 adapter	5-CGAGCAGGACTCATGA-3
Pre-selective primer	
E+A	5-GACTGCGTACCAATTCA-3
H/M+0	5-ATCATGAGTCCTGCTCGG-3
Selective primer	
E-AGG	5-GACTGCGTACCAATTCAGG-3
E-AAC	5-GACTGCGTACCAATTC AAC-3
E-ACA	5-GACTGCGTACCAATTCACA-3
E-AGC	5-GACTGCGTACCAATTCAGC-3
E-ACT	5-GACTGCGTACCAATTC ACT-3
E-AAG	5-GACTGCGTACCAATTC AAG-3
E-ACC	5-GACTGCGTACCAATTC ACC-3
E-AAA	5-GACTGCGTACCAATTC AAA-3
E-AAT	5-GACTGCGTACCAATTC AAT-3
E-AGT	5-GACTGCGTACCAATTC AGT-3
H/M+0+TCAC	5-ATCATGAGTCCTGCTCGGTCAC-3
H/M+0+TCAA	5-ATCATGAGTCCTGCTCGGTC AA-3
H/M+0+TCAT	5-ATCATGAGTCCTGCTCGGTC AT-3
H/M+0+TCAG	5-ATCATGAGTCCTGCTCGGTC AG-3