African Journal of Biotechnology Vol. 11(31), pp. 7795-7802, 17 April, 2012
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.3641
ISSN 1684–5315 © 2012 Academic Journals

Full Length Research Paper

Characterization of the fertilization independent endosperm (*FIE*) gene from soybean

Yingkao Hu[1]*, Ya Gao[1], Hongliang Xu[1], Jinyue Sun[2], Yaxuan Li[1] and Yueming Yan[1]

[1]College of Life Sciences, Capital Normal University, Beijing, 100048, China.
[2]Plant Biotechnology Institute, National Research Council, Saskatoon, SK, S7N 0W9, Canada.

Accepted 24 January, 2012

Reproduction of angiosperm plants initiates from two fertilization events: an egg fusing with a sperm to form an embryo and a second sperm fusing with the central cell to generate an endosperm. The tryptophan-aspartate (WD) domain polycomb protein encoded by fertilization independent endosperm (*FIE*) gene, has been known as a repressor of hemeotic genes by interacting with other polycomb proteins, and suppresses endosperm development until fertilization. In this study, one *Glycine max FIE* (*GmFIE*) gene was cloned and its expression in different tissues, under cold and drought treatments, was analyzed using both bioinformatics and experimental methods. *GmFIE* showed high expression in reproductive tissues and was responsive to stress treatments, especially induced by cold. *GmFIE* overexpression lines of transgenic *Arabidopsis* were generated and analyzed. Delayed flowering was observed from most transgenic lines compared to that of wild type. Overexpression of *GmFIE* in *Arabidopsis* also leads to semi-fertile of the plants.

Key words: Polycomb proteins, fertilization independent endosperm (FIE), *Glycine max*, *Arabidopsis thaliana.*

INTRODUCTION

Dramatic gene expression pattern shifts happen during angiosperm plant reproduction, especially at fertilization when the zygotic process is initiated. *Arabidopsis* ovule, which generates the female gametophyte, consists of an egg cell, two synergid cells, three antipodal cells, and a central cell (Drews et al., 1998). Double fertilization is initiated when an entering pollen tube discharges two genetically identical sperm cells. Along with this process, four distinct developmental programs are activated and lead to the formation of the embryo, endosperm, seed coat, and mature fruit. The two products of fertilization, embryo and endosperm are formed through distinct developmental patterns. The embryo is the result of fertilization of the egg cell as it passes through five stages: globular, heart, torpedo, walking stick, early maturation, and maturation (Goldberg et al., 1994). The endosperm is generated by fertilization of the central cell, which then divides mitotically without cytokinesis to form a syncytium of nuclei that fills the expanding central cell. Subsequently, the endosperm cellularizes to produce large amounts of protein, starch, and lipid to support the development of the embryo (Lopes and Larkins, 1993).

However, the balance between embryo and endosperm development can be disturbed by mutation of the fertilization independent endosperm (*FIE*) gene. *FIE* encodes a tryptophan-aspartate (WD)-type protein, which has seven WD motifs (Ng et al., 1997) and is homologous to the *Drosophila* polycomb group (PcG) protein, extra sex comb (ESC) (Ohad et al., 1999), and mammal's embryonic ectoderm development (EED) (Korf et al., 1998). Both ESC and EED promote their interactions with other polycomb proteins and repress homeotic target gene expression (Simon and Tamkun, 2002). Similarly, *Arabidopsis* PcG protein MEDEA (MEA) and FIE interact with each other through the amino-terminal region of MEA, regulating endosperm and embryo development (Luo et al., 1999; Spillane et al. 2000; Yadegari et al., 2000; Kohler et al., 2003a, b). Mutations of *FIE* genes allow diploid endosperm

*Corresponding author. E-mail: yingkaohu@yahoo.com. Tel: +86-10-68901842. Fax: +86-10-68901857.

Abbreviations: **WD**, Tryptophan-aspartate; **FIE**, fertilization independent endosperm.

development, seed coat formation, fruit elongation, and even partial embryo development in the absence of fertilization (Ohad et al., 1996, 1999; Chaudhury et al., 1997). These facts indicate that the wild-type *FIE* gene is essential to suppress endosperm development until fertilization. Down regulation of FIE in *Hieracium* can lead to seed abortion or inhibit autonomous embryo and endosperm initiation (Katz et al., 2004; Rodrigues et al., 2008). Despite the aberration phenotypes caused by low expression levels of FIE, the plants can still reach maturity. In addition to its previously described roles in seed development, FIE has also been reported to repress the expression of MADS-box gene family members: *Minichromosome maintenance 1* (*MCM1*) genes in yeast, *Agamous* (*AG*) in *Arabidopsis*, *deficiens* (*DEF*) in Antirrhinum and serum response factor (*SRF*) in humans (Riechmann and Meyerowitz, 1997). As in flowering plants, FIE PcG complex plays a central role in regulating the transition of female gametophyte to sporophyte (Goodrich et al., 1997; Guitton et al., 2004; Kohler et al., 2003a; Ohad et al., 1996, 1999). Low expression levels of FIE lead to dramatic morphological aberrations such as loss of apical dominance, curled leaves, early flowering and homeotic conversion of leaves in *Arabidopsis*. It has been shown that the FIE PcG protein interacts with curly leaf (CLF), a SET domain PcG protein, which functions in the regulation of leaf and flower differentiation (Goodrich et al., 1997). Therefore, it is quite possible that FIE controls the gametophytic and sorophytic developmental programs by associating with different SET-domain proteins (Katz et al., 2004).

Much work has been done on the study of the mutants of *FIE* and interactions of FIE PcG proteins with other PcG proteins. Given the conservation of *FIE* among species, we focused our present study on the potential effects of overexpression of extraneous FIE from soybean in *Arabidopsis* to provide more insights into the function of FIE PcG. As one of the most important crops in the world, studying the mechanism of seed development in soybeans has significant scientific and economic value. In this study, a soybean *FIE* gene was cloned and over expressed in *Arabidopsis* to further illustrate its function. The expression of soybean FIE among different tissues and under different stress treatments was also explored.

MATERIAL AND METHODS

Plant growth and treatments

Soybean seeds (*Cultiva Zhonghuang 13*) were germinated in pots containing 1.5 kg soil collected from an experimental field of the Chinese Academy of Agricultural Sciences (Beijing, China). The five-leaf stage seedlings were transferred to natural conditions to grow from April 4th, 2010 until harvesting. Soybean cotyledons, epicotyls, and hypocotyls were collected from five-day old seedlings. Vegetative tissues such as leaves, stems, and roots were collected from four-week old seedlings. Flowers were collected when they were in full bloom. Developmental seeds were collected two weeks after flowering (Xu et al., 2010).

For drought treatment, four-week old seedlings were stressed in culture solution with 20% polyethylene glycol (PEG) 6000. Leaves of the stress-treated seedlings were collected at time intervals of 0, 2, 5, 8 and 12 h. For cold treatment, four-week old seedlings were put in a chamber at 4°C for 48 h. Leaf samples were collected at time interval of 0, 5, 8, 12, 24 and 48 h. All samples were immediately frozen with liquid nitrogen and stored at -80°C for later use.

Arabidopsis thaliana ecotype Columbia (Col-0) was used in this study. Seeds were surface-sterilized and germinated on MacConkey agar (MA) plates. After germination plants were transferred to pots of soil and grown in a chamber at 23°C under long day conditions (16/8 h light/dark cycle).

Cloning of *GmFIE*

Arabidopsis FIE (Genbank accession no: AAD23584) was used to search the soybean EST database using the TBLASTN program (http://blast.ncbi.nlm.nih.gov/Blast.cgi.) to retrieve similar ESTs. These segments were assembled into one contig by SeqMan program of DNAStar software. ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) was used to analyze the open reading frame of the contig sequence. Specific primers 5'-AGATGGTGGGTGAAACGGC-3' and 5'- CATTAACTTAGCTTGA-GGACGCAC-3' were designed according to the sequence of the contig.

Bioinformatics analysis of *GmFIE*

Soybean genome database phytozom v6.0 (http://www.phytozome.net/soybean) was searched to determine the location and number of copies of *GmFIE*, and retrieve corresponding genomic sequences as well as promoter region sequences. Software GSDS (http://gsds.cbi.pku.edu.cn/) was employed to analyze the structure of *GmFIE*. The online software compute pI/Mw tool (http://au.expasy.org/tools/pi_tool.html) was used to predict the molecular weight and *pI* for GmFIE. The online protein database prosite (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_prosite.html) was used to analyze the amino acids composition features of GmFIE. Meanwhile, online database PLACE (http://www.dna.affrc.go.jp/PLACE/index.html; (Higo et al., 1999)) was used to analyze the potential *cis*-elements of the *GmFIE* promoter. *In silico* expression analysis of *GmFIE* was carried out in the UniGene database (http://www.ncbi.nlm.nih.gov/unigene/) with *GmFIE* Genbank number: EU169385. Lastly, Bioedit was used for mulit-alignment of GmFIE with FIEs from tobacco, *Arabidopsis*, potato, rice, *Physcomitrella patens* and *Micromonas pusilla*.

RNA extraction, first-strand cDNA synthesis and quantitative real-time polymerase chain reaction (qRT-PCR) analysis

RNA extraction, first–strand cDNA synthesis, as well as qRT-PCR were carried out according to the methods described by Xu et al. (2010). Constitutively expressed soybean *CYP2* (*cyclophilin*) gene was used as internal reference for normalization with a pair of primers: 5'-CGGGACCAGTGTGCTTCTTCA-3' and 5'-CCCCTCCACTACAAAGGCTCG-3' (Jian et al., 2008). Specific primers for qRT-PCR analysis of *GmFIE* were used as follows: forward primer 5'- AAATACCCTGTTCCTGAGTGTG-3' and reverse primer 5'- CCCTTCCCTGTTACCCACT-3'. The PCR was performed on a CFX96 real-time PCR detection system (Bio-Rad) and the conditions were as follows: 95°C for 3 min, 45 cycles of 15

s at 95°C, 58°C for 15 s and 72°C for 20 s. Three replicates were used for each sample. Data were analyzed by using CFX Manager Software (Bio-Rad).

Generation of transgenic plants overexpressing GmFIE

GmFIE open reading frame was amplified by PCR with the following primers: Forward 5'- GGGTCTAGACGGGAATTCGATAAGATG-3' (added *Xba* I site underlined) and reverse 5'-ACACCCGGGCGCGAATTCACTAGTGATTC-3' (added *Xma* I site underlined). PCR products were digested with *Xba* I and *Xma* I, and inserted into the *Xba* I/*Sma* I (share the same site with *Xma* I) sites of the pBI121 vector under the control of the CaMV 35S promoter. The recombinant plasmid pBI121-GmFIE was used to transform *Agrobacterium tumefaciens* (strain GV3101) by freeze-thaw methods (Weigel and Glazebrook, 2006). *Arabidopsis* plants were transformed by the floral dip method described by Clough and Bent (1998). Transformed plants were grown under the long day condition described previously. T1 generation seeds were harvested and screened on MS medium containing 50 mg/L kanamycin. 60 independent transgenic plants were obtained and transferred to pots of soil. Two leaves of 20-day old plants were used for DNA extraction with plant genomic extraction kit (TIANGEN). Specific primers (forward 5'-AAACCTCCTCGGATTCCATTGCC-3' and reverse 5'-CCCATTGATTCCTCCAGCCACA-3') were designed to identify the transgenic plants. Primers were designed across the border of target gene *FIE* and the vector pBI121. Two transgenic lines F5-3 and F8-4 were selected in the T2 generation for subsequent experiments.

Characterization of *GmFIE* overexpression lines

Seeds of T2 generation and WT were germinated and grown under the long day conditions described previously. The flowering date for transgenic and WT lines were noted. Pods collected from the transgenic lines and WT were dissected under stereomicroscope (Stemi SV11, ZEISS). Seeds of transgenic lines were calculated and results were chi-square tested with SPSS16.0.

RESULTS

Cloning and bioinformatics analysis of *GmFIE* and its promoter region

Protein sequence of Arabidopsis *FIE* (AF129516) was used to search the soybean EST database and four ESTs were retrieved, namely: EV281299, BI424788, GR841855, and BW656508. After assembly by the SeqMan program of DNAStar software, one contig was generated. With specific primers, a full-length cDNA fragment was amplified from first-strand cDNA by RT-PCR. The amplified fragment was purified and cloned into the pMD18-T vector for sequencing. The full-length soybean *GmFIE* was 1487 bp which is exactly the same size as the *in silico* cloning result. The *GmFIE* cDNA encodes a protein (ABW23440) which contains 381 amino acid residues with an estimated molecular weight of 42.4992 kDa and pI 5.56. Right at 5' of the initiation codon, there was a termination codon. The UTR at the 3' had one typical tailing signal. This sequence was stored at GenBank with an accession number EU169385. Multiple alignment of *GmFIE* with six other *FIEs* from species tobacco, *Arabidopsis*, potato, rice, *P. patens*, and *M. pusilla* species showed that they have high similarity, being 95, 87, 93, 89, 85 and 62%, respectively. The aforementioned evidence clearly indicated that the cloned *GmFIE* cDNA sequence was complete. Genome search results showed that *GmFIE* has two copies, one is located on soybean chromosome 10 with locus number Glyma10g02690, and another is located on chromosome 2 with locus number Glyma02g17110. The copy on chromosome 10 is exactly as cloned in our experiments and has a 99% similarity to the copy on chromosome 2. Comparison of the coding sequence (CDS) and genomic sequences showed that *GmFIE* has 11 introns and 12 exons. Prosite database analysis showed that GmFIE has seven WD motifs which are conserved among plants.

Cis-element analysis of *GmFIE* promoter region (2000 bp upstream of initiation condon) showed that a wide variety of elements existed in this region (Table 1). More significantly, many potential regulatory elements associated with stress-related responses were found, including 12 MYB transcription factor binding sites (all 12 are MYB core) and 16 MYC transcription binding sites.

Expression analysis of *GmFIE* in different tissues and under different treatments

In silico expression analysis showed that *GmFIE* was highly expressed in seed coats and somatic embryos with 93 transcripts per million and 153 transcripts per million respectively (Table 2). In roots, only 20 transcripts per million were detected. No expression in cotyledons, epicotyls, hypocotyls, flowers, leaves, stem, meristems nor pods was detected.

Tissue specific analysis with qRT-PCR showed that *GmFIE* was highly expressed in epicotyl, hypocotyl, cotyledon and seeds, and highest in cotyledon. Its expression in the other four chosen tissues, root, stem, leaf, and flower was significantly low (Figure 1).

Analysis of drought treatment showed that the pattern of *GmFIE* was slightly down-regulated in the first 8 h then sharply up-regulated in the next 4 h (Figure 2). The expression level for 12 h treatment was more than three times that for 0 h. For cold treatment, the expression patterns of *GmFIE* were first up-regulated then down-regulated, and the highest with the 5 h treatment -almost 50 times of that of the 0 h (Figure 2).

Features of *Arabidopsis GmFIE* overexpression lines

Among the 60 seedlings of six T2 transgenic lines, 49 showed delayed flowering averaged two days compared to that of WT. *Arabidopsis* pods from 13 plants of two randomly chosen transgenic lines: F5-3 and F8-4, as well

Table 1. Primary *cis*-regulatory elements and functions of *GmFIE*.

Elements type	Shared sequence	Element name	Element number	Function of element
Basic element	TATA	TATA box	11	Critical for accurate initiation of gene transcription
Upstream elements	GCCAATCT	CAAT box	25	Influence initiation rates of transcription
	TCTCTCTCT	CaMV35S	7	Enhance gene expression
	GATA	GATA	11	Core of a CuRE (copper-response element)
Stress responsive elements	CNGTTR (R = A/G)	MYBCORE	12	Binding site for all animal MYB and at least two plant MYB
	CANNTG	MYC	16	Binding site of ATMYC2, ICE1
	TGAC	WRKY710S	15	Core of TGAC-containing W-box

Table 2. Expression profile of *GmFIE* in different tissues and stages by *in silico* analysis.

Organ	Transcripts per million (TPM)	Gene EST/ total EST in pool
Cotyledon	0	0/30082
Epicotyl	0	0/2692
Hypocotyl	0	0/22641
Flower	0	0/16569
Leaf	0	0/45127
Root	22	1/45166
Stem	0	0/20640
Meristem	0	0/5891
Pod	0	0/8637
Seed coat	93	1/10690
Somatic embryo	153	2/13053

as WT were collected and dissected under stereomicroscope. In the pods of most transgenic lines unfertilized ovules were observed (Figure 3). Unfertilized ovule rates of transgenic lines were significantly higher than that of WT (Table 3). The chi-square test results for unfertilized ovules rates of transgenic line for F5-3 is X^2=74.186, df=1, p=0.000<0.01 and for F8-4, X^2=70.485, df=1, p=0.000<0.01.

DISCUSSION

Expression analysis of *GmFIE* in different tissues

It has been reported that *FIE* widely expresses in different tissues such as stem, cauline leaves and roots (Luo et al., 2000; Ohad et al., 1999). Interaction with different SET domain PcG proteins FIE is essential for the development control of plant shoot and leaf development (Katz et al., 2004). In this study, 11 different tissues or organs were chosen for *in silico* expression analysis, and eight for qRT-PCR analysis. *In silico* results showed that *GmFIE* in somatic embryo and seed coat have the highest expression level, while qRT-PCR showed that epicotyl, hypocotyl, cotyledon and seed have the highest expression of *GmFIE*. They both showed that *GmFIE* was highly expressed in tissues closely related with soybean reproduction. So, it seems that despite its possible roles in the control of the development of vegetative tissue, *FIE* mainly functions in plants reproduction process.

Bioinformatics and qRT-PCR analysis of *GmFIE* revealing its new function in stress response

Thanks to the genome sequencing of soybeans and powerful online bioinformatics analysis tools, a relatively comprehensive analysis of *GmFIE* was carried out in this study. Located on chromosome 10 with 12 exons, the complex structure indicated complex and important roles of *GmFIE*. Promoter region analysis showed 12 MYB transcription factor binding sites and 16 MYC

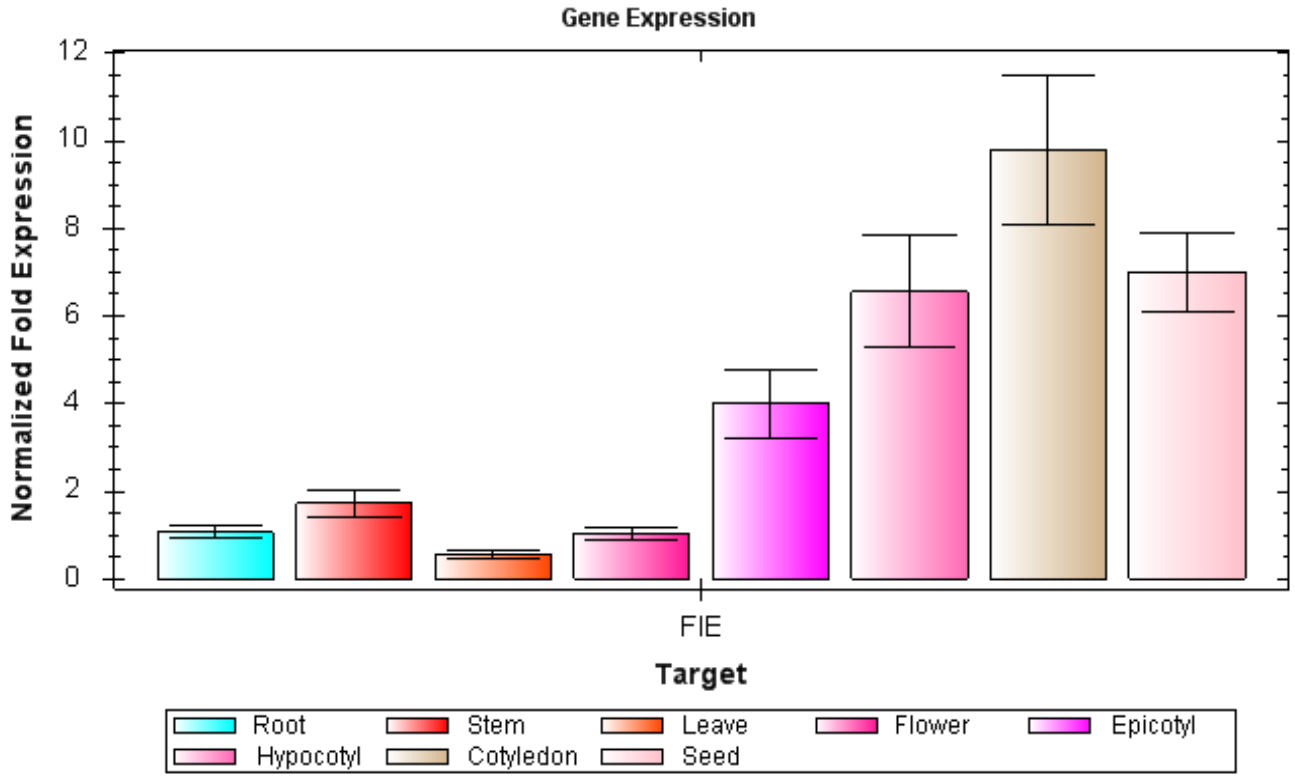

Figure 1. Transcript level of *GmFIE* in 8 different tissues and organs. The x-axis is for different tissues or organs. The bars above gene name indicate different tissues or organs. The order from left to right is: root, stem, leaves, flower, epicotyl, hypocotyl, cotyledon, and seed. The y-axis indicates the expression level after normalization to reference gene *CYP2*.

transcription binding sites (Chinnusamy et al., 2003; Hartmann et al., 2005; Urao et al., 1993) distributed in this area. Then high density of *cis*-elements indicated that *GmFIE* could be a highly stress responsive gene. This was confirmed by the qRT-PCR results of soybean drought and cold treatments. No such reports concerning FIE have been reported before. However, since a dynamic change happens during fertilization and seeds development, it is no surprise to find that *GmFIE* is stress responsive (Xu et al., 2010).

Overexpression of *GmFIE* in *Arabidopsis* confers evolution conservation and significant functions of FIE polycomb proteins

Introduction of an extraneous *FIE* gene to different plant species has been used in lots of research works in order to illustrate its function. It has been reported that the defects caused by *FIE* mutants can be rescued or partially rescued by the *FIE* gene from other species. One example in case is the target deletion of *FIE* in *P. patens*, which causes gametophore meristems over proliferation resulting in the moss not reaching the reproductive phase. This defect can be partially rescued by the *FIE* gene of *A. thaliana*. Meanwhile, the *A. thaliana FIE* mutant's gametophytic lesion can also be partially complemented by introduction of *PpFIE* (Mosquna et al., 2009). This fact indicates that the function of FIE proteins have been highly conserved during evolution and that the FIE protein from one species can function well in another species.

According to previous reports, loss-of-function mutations of *FIE* allowed endosperm development in the absence of fertilization (Ohad et al., 1996, 1999; Chaudhury et al., 1997), and reduced level of FIE protein in *Arabidopsis* caused pleiotropic aberrant phenotypes (Katz et al., 2004), indicating that the FIE polycomb complex is an essential component to suppress the floral program in the early stages of plant development (Kinoshita et al., 2001). Also, PcG proteins, especially FIE plays a major role in maintaining the repression of homeobox gene beyond their appropriate and temporal and spatial expression boundaries (Simon and Tamkun, 2002). Though not only that, an extraneous FIE proteins can function as well as endogenesis ones. Here we hope to know what will happen if extraneous FIE protein is overexpressed in *Arabidopsis*. In the current work, *GmFIE* was introduced into *Arabidopsis* and the transgenic lines and two significant phenotype changes were observed. First, over half of the transgenic lines have observed flowering date delay compared to that of

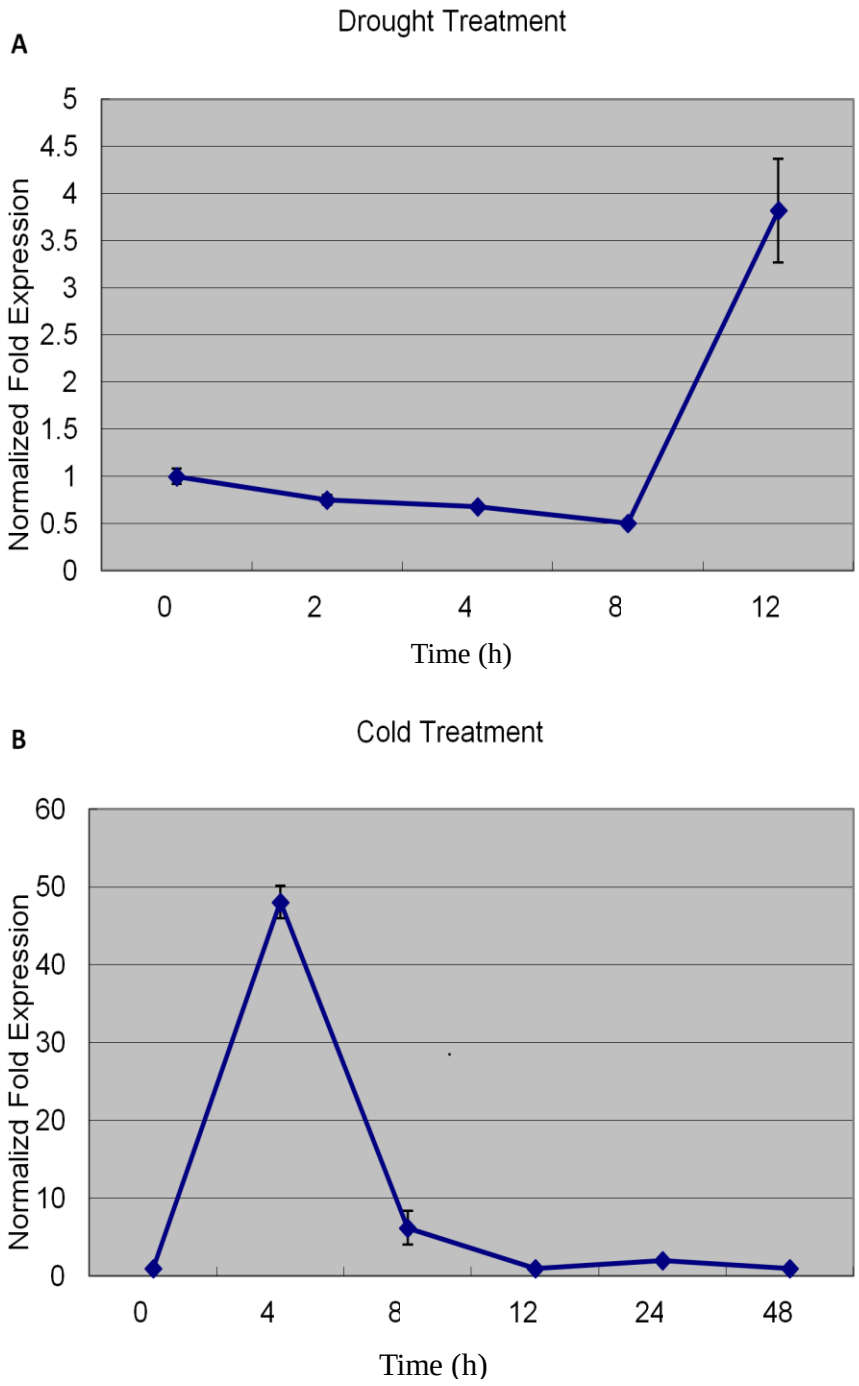

Figure 2. Expression profile of *GmFIE* under drought and cold treatments. The x-axis shows time courses of drought (A) and cold (B) treatments. The bars from left to right indicate time courses of drought treatment 0, 2, 5, 8 and 12-h intervals (A) and cold treatment 0, 5, 8, 12, 24 and 48-h intervals (B). The y-axis shows expression levels after normalization to internal control gene *CYP2*.

WT. While low level expression of FIE in *Arabidopsis* can lead to early flowering (Katz et al., 2004), the delayed flowering in the study provide another evidence that FIE PcG protein functioned in suppressing of critical aspect of early plant reproduction.

Secondly, unfertilized ovules rates of transgenic lines were significantly higher than WT. Many possible reasons may contribute to the phenomenon. Possibly, the over-expressed protein may act in a dominant-negative manner to interfere with the function of the endogenous

Table 3. Quantity statistics of *Arabidopsis* seeds for transgenic lines and WT.

Plant	Normal seed number	Unfertilized ovule	Total number
WT	553	22	575
F5-3	424	111	535
F8-4	428	108	536

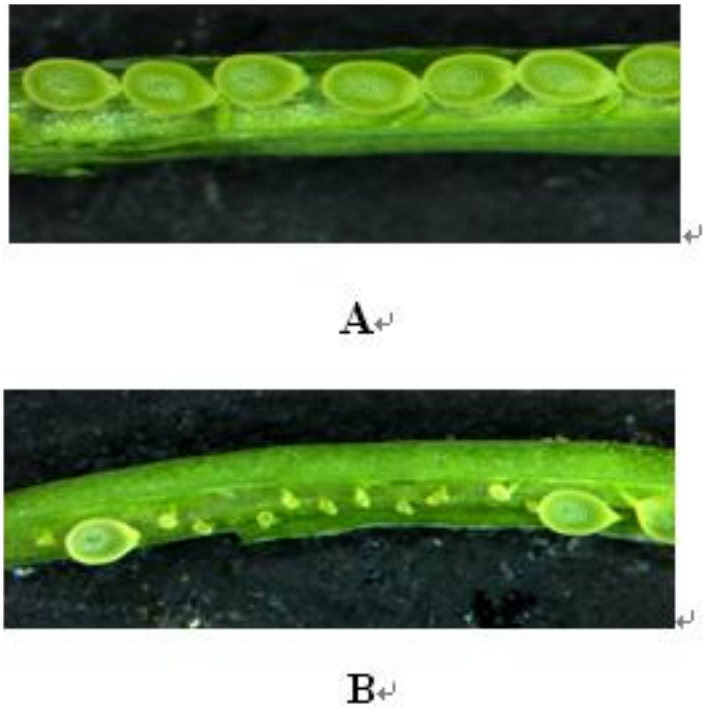

Figure 3. Morphological character of seeds of transgenic *GmFIE* plant of T3 generation and WT (×1.6). A, WT; B, transgenic *GmFIE* plant of T2 generation.

Arabidopsis FIE protein, by forming inactive PcG complexes or competing for the other components of the complex and thus reducing the number of complete complexes present. There is also the possibility that the over expressed protein suppressed the expression of endogenous Arabidopsis *FIE* gene. The actual mechanism behind the phenomenon needs more evidence from further studies.

Acknowledgements

Authors would like to thank the National Natural Science Foundation of China (30971783) for financial support.

REFERENCES

Chaudhury AM, Luo M, Miller C, Craig S, Dennis ES, Peacock WJ (1997). Fertilization independent seed development in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA, 94: 4223-4228.

Chinnusamy V, Ohta M, Kanrar S, Lee BH, Hong X, Agarwal M, Zhu JK (2003). ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. Gene Dev. 17: 1043-1054.

Clough SJ, Bent AF (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. 16: 735-743.

Drews GN, Lee D, Chistensen CA (1998). Genetic analysis of female gametophyte development and function. Plant Cell, 10: 5-17.

Goldberg RB, De Paiva G, Yadegari R (1994). Plant embryogenesis: zygote to seed. Science, 266: 605-614.

Goodrich J, Puangsomlee P, Martin M, Long D, Megerowitz EM, Coupland G (1997). A polycomb-group gene regulates homeotic gene expression in *Arabidopsis*. Nature, 386: 44-51.

Guitton AE, Page DR, Chambrier P, Lionnet C, Faure JE, Grossniklaus U, Berger F (2004). Identification of new members of Fertilization Independent Seed Polycomb Group pathway involved in the control of seed development in *Arabidopsis thaliana*. Development, 131: 2971-2981.

Hartmann U, Sagasser M, Mehrtens F, Stracke R, Weisshaar B (2005). Defferential combinatorial interactions of *cis*-acting elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. Plant Mol. Biol. 57: 155-171.

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999). Plant cis-acting regulatory DNA elements (PLACE) database: Nucleic Acids Res. 27:

297-300.
Jian B, Liu B, Bi Y, Hou W, Wu C, Han T (2008). Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Mol. Biol. 9: p. 59.
Katz A, Oliva M, Mosquna A, Hakim O, Ohad N (2004). FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development. Plant J. 37: 707-719.
Kinoshita T, Harada JJ, Goldberg RB, Fischer RL (2001). Polycomb repression of flowering during early plant development. Proc. Natl. Acad. Sci. USA. 98: 14156-14161.
Kohler C, Hennig L, Bouveret R, Gheyselinck J, Grossniklaus U, Gruissem W (2003a). Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. EMBO J. 22: 4804-4814.
Kohler C, Hennig L, Spillane C, Pien S, Gruissem W, Grossniklaus U (2003b). The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene *PHERES1*. Gene Dev. 17: 1540-1553.
Korf I, Fan Y, Strome S (1998). The polycomb group in Caenorhabditis elegans and maternal control of germline development. Development, 125: 2469-2478.
Lopes MA, Larkins BA (1993). Endosperm origin, development, and function. Plant Cell, 5: 1383-1399.
Luo M, Bilodeau P, Dennis ES, Peacock WJ, Chaudhury A (2000). Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing *Arabidopsis* seeds. Proc. Natl. Acad. Sci. USA. 97: 10637-10642.
Luo M, Bilodeau P, Koltunow A, Dennis ES, Peacock WJ, Chaudhury AM (1999). Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA. 96:296-301.
Mosquna A, Katz A, Decker EL, Rensing SA, Reski R, Ohad N (2009). Regulation of stem cell maintenance by the Polycomb protein FIE has been conserved during land plant evolution. Development, 136: 2433-2444.
Ng J, Li R, Morgan K, Simon J (1997). Evolutionary conservation and predicted structure of the *Drosophila* extra sex combs repressor protein. Mol. Cell Biol. 17: 6663-6672.
Ohad N, Margossian L, Hsu YC, Williams C, Repetti P, Fischer RL (1996). A mutation that allows endosperm development without fertilization. Proc. Natl. Acad. Sci. USA. 93: 5319-5324.
Ohad N, Yadegari R, Margossian L, Hannon M, Michaeli D, Harada JJ, Goldberg RB, Fischer RL (1999). Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. Plant Cell, 11: 407-416.
Riechmann JL, Meyerowitz EM (1997) MADS domain proteins in plant development. Biol. Chem. 378: 1079-1101.
Rodrigues JCM, Tucker MR, Johnson SD, Hrmora M, Koltunow AM (2008). Sexual and Apomictic Seed Formation in *Hieracium* Requires the Plant Polycomb-group Gene Fertilization Independent Endosperm. Plant Cell, 20: 2372-2386.
Simon JA, Tamkun JW (2002). Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes. Curr. Opin. Genet. Dev. 12: 210-218.
Spillane C, MacDougall C, Stock C, Kohler C, Vielle-Calzada JP, Nunes SM, Grossniklaus U, Goodrich J (2000). Interaction of the *Arabidopsis* polycomb group proteins FIE and MEA mediates their common phenotypes. Curr. Biol. 10: 1535-1538.
Urao T, Yamaguchi-Shinozaki K, Urao S, Shinozaki K (1993). An Arabidopsis myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. Plant Cell, 5: 1529-1539.
Weigel D, Glazebrook J (2006). Transformation of *Agrobacterium* using the freeze-thaw method. Cold Spring Harb. Protoc. doi:10.1101/pdb.prot4666.
Xu H, Li Y, Yan Y, Wang K, Gao Y, Hu Y (2010). Genome-scale identification of Soybean BURP domain-containing genes and their expression under stress treatments. BMC Plant Biol. 10: 197.
Yadegari R, Kinoshita T, Lotan O, Cohen G, Katz A, Choi Y, Nakashima K, Harada JJ, Goldberg RB, Fisher RL, Ohad N (2000). Mutations in the *FIE* and *MEA* genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms. Plant Cell 12: 2367-2382.