

Cloning and sequence analysis of oilseed rape (*Brassica napus*) *SHP2* gene

Xiao-Li TAN^{1,2,3,*}, Zong-Wei XIA^{2,3}, Li-Li ZHANG², Zhi-Yan ZHANG², Zhong-Jian GUO², and Cun-Kou QI^{1,*}

¹*Institute of Industrial Crop, Jiangsu Academy of Agricultural Sciences, Nanjing, 200014, P.R. China*

²*Institute of Life Sciences, Jiangsu University, Zhenjiang, 212013, P.R. China*

(Received September 8, 2008; Accepted April 13, 2009)

ABSTRACT. MADS-box family transcriptional factors SHATTERPROOF1 (SHP1) and SHATTERPROOF2 (SHP2) are required for fruit dehiscence by regulating the cell specification in valve margin in *Arabidopsis thaliana*. BLASTN search against *Brassica* database in TAIR using *AtSHP2* gene as a query, revealed several ESTs sharing high similarity to *AtSHP2*. Two highly homologous cDNAs of *SHP2* were obtained from *Brassica napus* by PCR and RACE, and they are named as *BnaA.SHP2.a* and *BnaC.SHP2.b*, respectively. In this work, we focused on *BnaA.SHP2.a*. Sequence alignment and phylogenetic analysis indicated that *BnaA.SHP2.a* is highly identical to *AtSHP2*. The southern blotting demonstrated that there are at least two copies of *BnaSHP2* in *B. napus* genome. Semi-quantitative RT-PCR showed that *BnaSHP2* was expressed in roots, floral buds, and pods, but not in stems, leaves, or seeds. *BnaSHP2* was localized in the nuclei of HzAM1 cells when fused with EGFP and expressed using HaSNPV bacmid system.

Keywords: *Arabidopsis*; *BnaSHP2* gene; *Brassica napus*; Fruit dehiscence; MADS-box family transcriptional factors; Sequence alignment.

INTRODUCTION

Many flowering plants produce dehiscent fruits, which open during maturation for seed dispersal. This fruit dehiscence or pod shatter for seed dispersal is a programmed cell-separation process. It exists in many important crops such as oilseed rape and soybean (Liljegren et al., 2000; Rajani and Sundaresan, 2001; Romkaew and Umezaki, 2006), which causes a huge problem in agriculture and thus becomes a significant research focus. At present, researches on pod shattering are primarily conducted in model plant *Arabidopsis thaliana*, and great progress has been made towards understanding the genetic mechanisms underlying the regulation of pod dehiscence (Ferrandiz, 2002).

Oilseed rape fruit has similar structure to *Arabidopsis* fruit which develops from a gynoeceium composed of two fused carpels, and grows to a silique after fertilization (Bowman et al., 1999). Two valves enclose seeds and connect septum in inner along the pod, and replum outside of the pod while fruit develop, and separate from the replum in the mature fruit to disperse the seeds (Din-

neny et al., 2005). Fruit dehiscence takes place under a series of co-ordinate events involving the prior development of Dehiscence zone (DZ) formed in valves margin (Rajani and Sundaresan, 2001). Dehiscence zone (DZ), known as the region at which valves separate from replum, consists of separation layer and lignified layer in valves margin. The separation layer is a strip of labile nonlignified cells sandwiched between a layer of relatively large parenchyma cells on the side of the replum and a group of lignified cells in the valve margin (Meakin and Roberts, 1990; Spence et al., 1996; Ferrandiz, 2002). The lignified layer of the valve margin is thought to act in conjunction with the lignified inner sub-epidermal layer of the valve (enb), creating tensions upon silique drying that detach the valves from the replum along the separation layer of which the middle lamella is broken down by hydrolytic enzymes weakening the cell wall adhesion (Spence et al., 1996). Recent studies in *Arabidopsis* have shown that DZ plays an important role in silique dehiscence process, and also demonstrated the importance of the transcriptional network in specifying the development of the DZ and the inner subepidermal layer (Wu et al., 2006).

Previous studies of *Arabidopsis* fruit development have identified a few transcriptional regulators involved in valve margin specification and valve development. Five transcriptional factors are identified as follows: INDEHISCENT (IND), SHATTERPROOF (SHP1), SHATTERPROOF2 (SHP2), ALCATRAZ (ALC) and FRUITFULL (FUL). IND, an atypical basic helix-loop-helix (bHLH)

³Xiao-Li Tan and Zong-Wei Xia contributed equally to this paper.

*Corresponding authors: E-mail: xltan@ujs.edu.cn (Xiao-Li Tan); E-mail: qck@jaas.ac.cn, Tel: +86-25-84390372 (Cun-Kou Qi).

family protein, promotes the differentiation of the separate layer and lignification of the lignified layer in DZ (Liljegren et al., 2004). *ALC*, which is also a member of bHLH family and expressed at similar locations as *IND*, seems to primarily regulate the formation of the separation layer (Rajani and Sundaresan, 2001). Two MADS box genes—*SHP1* and *SHP2* (formerly known as *AGL1* and *AGL5*, respectively) belong to C-function genes according to the ABC/ABCDE floral development model, and have been proved to be required for fruit dehiscence by regulating the cell specification in valve margin in *Arabidopsis* (Flanagan et al., 1996; Liljegren et al., 2000). Moreover, *SHP1* and *SHP2* are functionally redundant and only *shp1 shp2* double mutation causes the failure of pod shatter (Liljegren et al., 2000). Another MADS-box transcription factor—*FUL* is previously reported to be responsible for the valve cell development during the elongation of fruit (Gu et al., 1998; Liljegren et al., 2000), but later this role was shown to be largely indirect. *FUL* negatively regulates *SHP1*, *SHP2*, *IND* and *ALC* by restricting their expression in valve margin from valve side. In the *ful* mutant, *SHP1*, *SHP2*, *IND* and *ALC* are all ectopically expressed throughout the valves of the fruit (Ferrandiz et al., 2000; Liljegren et al., 2004). Moreover, constitutive expression of *FUL* with 35S promoter in *Arabidopsis* produces indehiscent siliques, and similar result is obtained in *Brassica juncea* (Østergaard et al., 2006), indicating that *FUL* expression is sufficient to repress these valve margin identity genes and convert the valve margin and outer replum cells into valve cells.

Oilseed rape (*Brassica napus*) is one of the most important oilseed crops, as it is now the main source for edible vegetable oil and future renewable energy. Therefore, to unravel the genetic mechanisms underlying the regulation of pod dehiscence in oilseed rape is of great significance in seeking proper approach to solve the problem caused by pod shattering. Since *B. napus* is closely related to *Arabidopsis* (Spence et al., 1996), it may be possible to apply our understanding of *Arabidopsis* fruit development to inhibit pod shatter in *B. napus*. However, many genes related to pod dehiscence have not been identified in oilseed rape. Till now, only the orthologue of *SHP1* in *B. napus*, called *BnaSHP1* has been reported. Fortunately, based on a broad bioinformatics platform provided by large-scale sequencing and continuously improved analysis system, more and more novel genes are being isolated at present. In this study, a new putative *SHP2* gene has been cloned from *B. napus*. We also analyzed its sequence and function.

MATERIALS AND METHODS

Plant materials

Oilseed rape cultivar *Ningyou16* from the subgenera *B. napus* was used in this study. Roots, stems, leaves, floral buds and pods (35 days after anthesis) were harvested at the indicate stages and stored at -70°C for DNA and RNA isolation.

Total RNA extraction and reverse transcription

Total RNAs were extracted from *B. napus* roots, stems, leaves, floral buds and pods respectively using Trizol (Invetrogen) reagent according to its instruction. The total RNAs were treated with RNase-free DNase (Takara). The RNA precipitates were dissolved in 50 µl of 10 mM Tris/1 mM EDTA buffer (pH 8.5) and spectrophotometrically quantified at 260 nm. The integrity of the isolated RNA was checked by 1% agarose gel. Single stranded cDNAs for PCR were synthesized with M-MLV Reverse Transcriptase (Fermentas).

Cloning of the full-length cDNA of *BnaSHP2*

The cDNA sequence of *AtSHP2* (GenBank No NM_129844) was used as a query for a Washington University-Basic Local Alignment Search Tool (WU-BLAST) search against *Brassica* database in TAIR (<http://www.arabidopsis.org>), and several ESTs sequences sharing high similarity with *AtSHP2* were obtained. These ESTs were assembled to one cDNA fragment using DNASTar tool. The PCR primers were designed as follows: 5'-AGG GTT GAT ATA AAT GGA GGG-3' (*BnaSHP2-S*), 5'-TTA GAC TTT ACT TAA ACA AGT TGG A-3' (*BnaSHP2-A*). Using the first strand cDNA mix from floral buds as template, PCR was performed using the following thermal-cycling conditions: 94°C for 3 min, followed by 35 cycles, with each cycle at 94°C for 30 s, at 54°C for 30 s, and at 72°C for 1 min, followed by a final extension at 72°C for 10 min. PCR product was checked, and recovered using DNA Purification Kit (Shenergy Biocolor, Shanghai). The sub-cloning was done with pMD18-T vector (Takara) and *E. coli* strain DH5α.

3' RACE was used to clone 3' flanking sequence. A set of nested primers 3'-*BnaSHP2-1*(5'-CGC TGT TAA TCC TCC TTC CGT C-3') / 3'-*BnaSHP2-2*(5'-CCG CTC CAA GAA GCA TGA GAT G-3') were designed. The first-strand cDNA was synthesized from 1 µg of total RNA using a cDNA synthesis primer 3'-AAP (5'-TAC TAG TCG ACG CGT GGC CTT TTT TTT TTT T-3'). Using first-strand cDNA and primer 3'-*BnaSHP2-1* (5'-CGC TGT TAA TCC TCC TTC CGT C-3') as well as universal amplification primer 3'-AUAP (5'-TAC TAG TCG ACG CGT GGC C-3'), the first round PCR was performed in a total volume of 50 µl containing 1 µl cDNA, 10 pmol each of primer 3'-*BnaSHP2-1* and 3'-AUAP, 10 µmol dNTPs, 5 µl 10 × cDNA reaction buffer and 5U Taq polymerase. PCR reaction was carried out under the following conditions: 94°C for 3 min, 35 cycles of amplification (1 min at 94°C, 1 min at 60°C and 2 min at 72°C), and at 72°C for 10 min. Utilizing the first round PCR product and primers 3'-*BnaSHP2-2* (5'-CCG CTC CAA GAA GCA TGA GAT G-3') / 3'-AUAP (5'-TAC TAG TCG ACG CGT GGC C-3'), the second round PCR was performed similarly. The second PCR product was purified and sub-cloned into pMD18-T vector followed by sequencing.

5' RACE was also performed to obtain the upstream sequence. Two gene specific primers 5'-*BnaSHP2-1*

(5'-GTG GTT GGT TTT GGT CGG AGG AGT-3') and 5'-BnaSHP2-2 (5'-GTT GAT GTA TCA CAC TCG CTT CTT-3') were used to amplify the 5'-end of BnaSHP2 gene. The first-strand cDNA for 5' RACE was synthesized followed by tailing cDNA with oligo (C). The first round of PCR was performed with 5'-BnaSHP2-1 and 5'-AAP (5'-GGC CAC GCG TCG ACT AGT ACG GGI IGG GII GGG IIG-3'). The volume of PCR reaction mixture was 50 μ l, containing 1 μ l cDNA, 10 pmol each of primer 5'-BnaSHP2-1 and 5'-AUAP, 10 μ mol dNTPs, 5 μ l 10 \times cDNA reaction buffer and 5U Taq polymerase. PCR was carried out by denaturing cDNA at 94°C for 3 min followed by 35 cycles of amplification (1 min at 95°C, 1 min at 65°C and 2 min at 72°C) and by extension at 72°C for 10 min. The PCR product was used for nested PCR with the second round of amplification with 5'-BnaSHP2-2 and 5'-AUAP (5'-GGC CAC GCG TCG ACT AGT A-3'). PCR product was purified and sub-cloned into pMD18-T vector followed by sequencing.

Bioinformatic analysis

DNA and amino-acid sequences analysis was performed with DNASTar 5.0 software. BLAST was performed in TAIR (<http://www.arabidopsis.org>) and NCBI (<http://www.ncbi.nlm.nih.gov/>). ExPasy tools (<http://www.expasy.org/>) were used to analyze the protein profile of BnaSHP2. The sequences alignment was performed using the program Multalin with the default parameters (<http://prodes.toulouse.inra.fr/multalin/multalin.html>) (Corpet, 1988) and BLAST 2 in NCBI, and GeneDoc were used for editing sequence alignment. Neighbor-Joining Phylogenetic Tree with bootstrap values was constructed using Mega 3.1.

DNA extraction and southern blotting

Total genomic DNA was extracted from *B. napus* leaves for Southern blotting analysis using the cetyltrimethylammonium bromide (CTAB) method. Genomic DNA (10-15 μ g) was firstly digested overnight with *Bam*HI, *Eco*RI, *Eco*RV, *Hind*III, and *Xba*I respectively, then the digestion products were purified and resuspended in 30 μ l of TE solution. The purified DNA restriction fragments were separated through electrophoresis on a 0.8% agarose 1 \times TAE gel at 1 V cm⁻¹ for 20 h followed by alkaline (0.4 M NaOH) upward capillary transfer for 3 h onto a nylon membrane (Hybond N⁺) (Sambrook and Russell, 2001; Chen et al., 2009; Huang et al., 2009). The probe was PCR-amplified fragment from *B. napus* floral buds cDNA with BnaSHP2-EST-S (5'- TCT TCG ATC CAA GAT TAG TG -3') and BnaSHP2-EST-A (5'- AAC AAG TTG GAG AGG TGG TT -3') primers, and was labeled with [³²P]-dCTP (Random Primer DNA Labeling Kit Ver. 2.0, TaKaRa) according to the manufacturer's instructions. After a 16-hour hybridization at 65°C and washing, membranes were exposed to film (Kodak MS) at -70°C.

Analysis of *BnaSHP2* expression

To determine the tissue distribution of *BnaSHP2* tran-

scripts, the first strand cDNA mix from roots, stems, leaves, floral buds and pods of *B. napus* (cv. *ningyou16*) were used as template for RT-PCR. BnaSHP2-S and BnaSHP2-A (shown in 1.3), were used as primers. *B. napus* β -Actin gene (GenBank No AF111812), amplified with primers β - Actin F (5'-ATG GCC GAT GGT GAG GAC ATT C-3') and β - Actin R (5'-GGT GCG ACC ACC TTG ATC TTC-3'), was used as a control in the experiments. The thermal cycling profile consists of initial denaturation at 95°C for 3 min and 35 cycles at 95°C for 30 s, at 54°C for 30s, and at 72°C for 1 min. The PCR reaction mixture contains 2.0 μ l of buffer (+Mg²⁺) (10 \times), 2 μ l of dNTP (2 mmol/L), 2 μ l of each primer (2 μ mol/L), 0.2 μ l enzyme (TianGen) and 11.8 μ l of cDNA templates and ddH₂O.

Sub-cellular localization of *BnaSHP2*

To generate a GFP fusion construct, the ORF of *BnaSHP2* without stop codon (TAA) was amplified from the floral cDNA with two primers: BnaSHP2-F 5'-GGA TCC ATG GAG GGT GGT GCG AGT G-3' and BnaSHP2-R 5'-TCT AGA AAC AAG TTG GAG AGG TGG T-3'. pMD18-T (TAKARA) was used for intermediate vector, and the PCR product of BnaSHP2 was cloned into pEGFP-N1 to generate *BnaSHP2-EGFP* fusion construct. The *BnaSHP2-EGFP* fusion gene was cloned into pFast-bac1 vector to produce the donor plasmid pFast-BnaSHP2-EGFP, then transformed into competent DH10B cells containing helper and HaSNPV bacmid (HaHZ8). Recombinant virus vHa-BnaSHP2-EGFP was identified by PCR and restriction enzymes analysis, and then was used to transfect HzAM1 cells.

HzAM1 cells (1 \times 10⁵) were grown on glass cover slips in Petri dishes and infected with vHa-BnaSHP2-EGFP at MOI of 10. At 24, 36, 48, 72 and 96 hours post infection (hpi), the cells were fixed and stained with Hoechst (Beyotime), then examined by a S2 Leica laser confocal scanning microscope. vHa-EGFP was used as GFP control.

RESULTS

Cloning of the full-length cDNA of *BnaSHP2*

AtSHP2 gene sequence was used as a query to search against *Brassica* database, and many ESTs sharing high similarity with *AtSHP2* were obtained. Among them, Eleven ESTs sharing 88%-90% similarity with *AtSHP2* were assembled to a single cDNA fragment which is 777 bp in length. A complete open reading frame (ORF) of the assembled cDNA was 750 bp, which is 9 bp longer than that of *AtSHP2* (741 bp). The fragment of expected size was amplified from floral buds cDNA of *B. napus* (cv. *Ningyou16*) by PCR, and was designated *BnaSHP2*. To further understand the 5'-untranslated region (5'-UTR) and 3'-untranslated region (3'-UTR) of *BnaSHP2*, RACE was performed. 3' RACE revealed three different lengths sequences, while 5' RACE detected more sequences (Figure 1). Four of 5'-UTR and three of 3'-UTR were

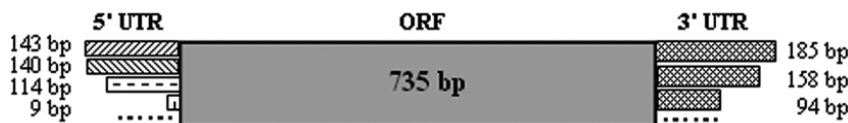


Figure 1. Schematic diagram of different transcripts of *BnaSHP2* obtained by RACE. The diverse 5' UTR, four cDNA sequences of which were sequenced, were described in front of ORF region, while three different lengths of 3' UTR were scaled out on the right of Figure. The possible other sequences of 5' UTR and 3' UTR were indicated with ".....".

sequenced, while other possible sequences are still not known. It is interesting that these four 5'-UTR sequences are quite divergent and three 3'-UTR sequences are identical. According to these sequences, four upstream primers and three downstream primers were designed and used to amplify the full length of cDNA of *BnaSHP2* by their cross-pairing, so several PCR products ranged from 838 bp to 1,063 bp were obtained, indicating that multiple transcripts existing in *BnaSHP2*. These transcripts shared an almost same ORF (735 bp) sequence except for fifteen nucleotides, and were divided into two groups (α and β). This may suggest that these transcripts of *BnaSHP2* may be derived from two paralogues. In addition, four genomic DNA clones from *B. rapa* (GenBank AC189588 and DU113592), *B. oleracea* (BZ061644) and *B. napus* (FP018823) are found in TAIR database of genomic DNA. The genomic DNA clones from *B. rapa* and *B. napus* are consistent with group α while that from *B. oleracea* is consistent with group β , indicating that transcripts in group α are derived from A genome while those in group β are derived from C genome. According to the standardization of *Brassica* gene naming provided by Østergaard and King (Østergaard and King, 2008), core sequences of two groups represent two genome types of *BnaSHP2* are named respectively as *BnaA.SHP2.a* and *BnaC.SHP2.b* (GenBank Acc. No. EU424342 and EU424343). Their putative amino acid sequences of two paralogues of *BnaSHP2* are almost identical except for one amino acid residue (S or N) at the position of 187aa (Figure 2). In this paper, we used *BnaA.SHP2.a* for subsequent analysis and experiments (if without specification).

ExPasy tool (<http://www.expasy.org/>) was used to analyze profiles of the *BnaSHP2* protein. The 735 bp ORF of *BnaSHP2* encoded a protein of 244 amino acids with a predicted molecular weight of 28 kDa and a calculated isoelectric point of 9.12. Total number of negatively charged residues (Asp + Glu) and positively charged residues (Arg + Lys) are 31 and 37, respectively. The *BnaSHP2* has three N-glycosylation sites (28 to 31 NTTN) (74 to 77 NNSV) (232 to 235 NSSD), eight Protein kinase C phosphorylation site (12 to 14 SSK) (13 to 15 SKK) (30 to 32 TNR) (64 to 66 STR) (76 to 78 SVR) (109 to 111 SSK) (155 to 157 SKK) (187 to 189 SER), two casein kinase II phosphorylation site (96 to 99 SVTE) (205 to 208 TVYE).

Structure analysis of *BnaSHP2* and homologous alignment

Using SMART program (Simple Modular Architec-

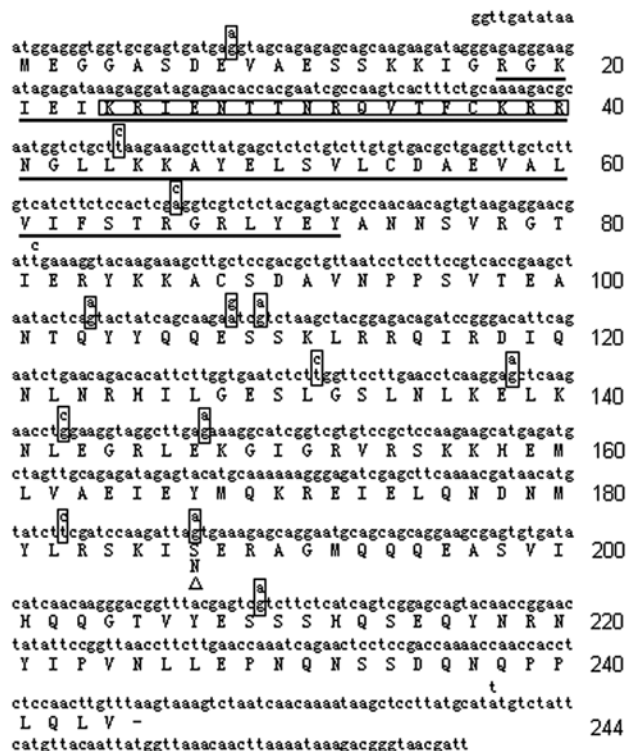


Figure 2. Core cDNA sequence of the *BnaSHP2* gene and deduced amino acid sequence from it the core nucleotide sequence of *BnaA.SHP2.a* is shown, and only the nucleotides in *BnaC.SHP2.b* that are different from *BnaA.SHP2.a* are shown in little boxes upon the *BnaA.SHP2.a* sequence. The amino acid sequence of *BnaA.SHP2.a* was shown beneath the nucleotide sequence. "—" indicates the different amino acid residue between *BnaA.SHP2.a* and *BnaC.SHP2.b*. MADS-box domain signature region of *BnaSHP2* is underlined. A bipartite nuclear localization signal was predicted and marked with rectangular box.

ture Research Tool: <http://smart.embl-heidelberg.de>), A MADS-box domain (amino acid 18-72) and a K-box domain (amino acid 102-192) as DNA-binding and dimerization regions were identified in the *BnaSHP2*, indicating that the protein belongs to MADS-box transcription factor family. By comparing *BnaSHP2* with other MADS-box proteins, the conservative sequence was primarily restricted to MADS-box and K-box regions. Other regions showed a high degree of variation among the MADS-box proteins compared (data not show). Sequence alignment of the MADS-box and K-box domains from *BnaSHP2* and related MADS-box proteins from yeasts, plants, and humans indicated that the MADS-box is the most conservative region in all MADS-box members' proteins,

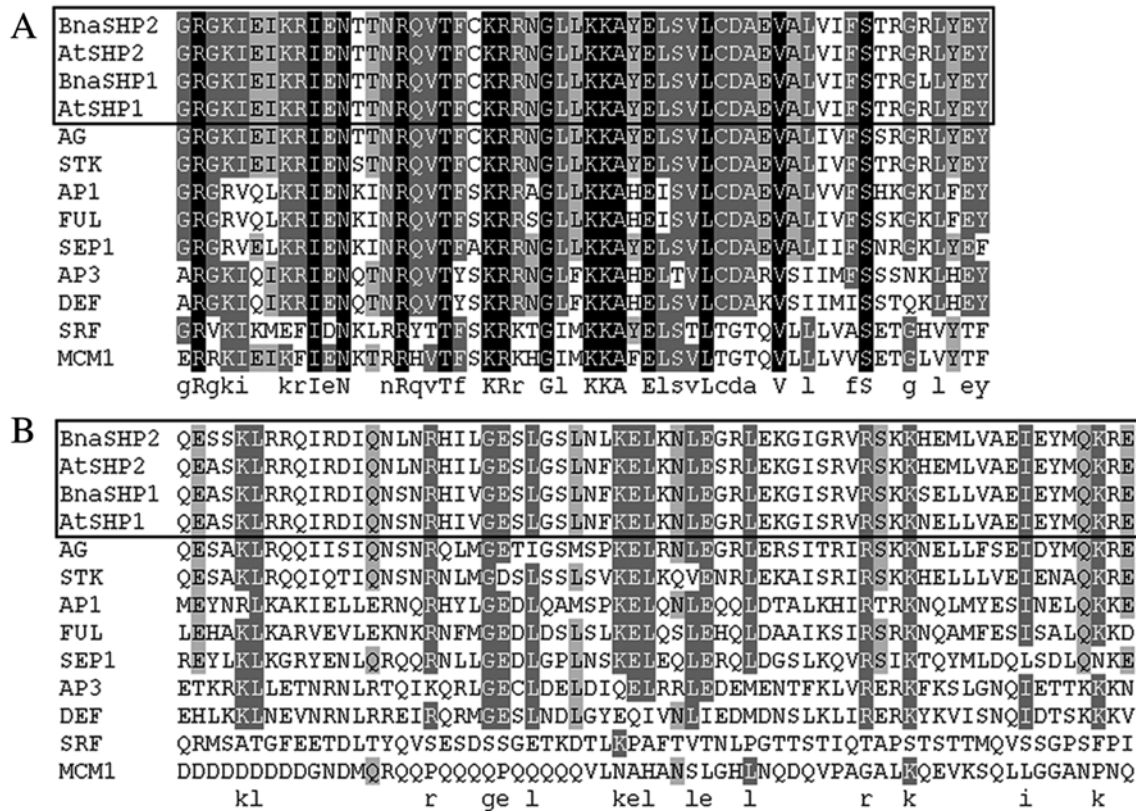


Figure 3. Multiple sequences alignment of MADS-box (A) and K-box (B) regions from BnaSHP2 and other MADS-box family proteins from plants, yeasts and humans. The proteins aligned are: AtSHP1 (SHATTERPROOF1) (NP_191437); AtSHP2 (SHATTERPROOF2) (NP_565986); AP1 (APETALA1) (NP_177074); AP3 (APETALA 3) (NP_191002); AG (AGAMOUS) (NP_567569); STK (SEEDSTICK) (NP_192734); SEP1 (SEPALLATA1) (NP_568322); FUL (FRUITFULL) (NP_568929) from *Arabidopsis*; BnaSHP1 (AAK62033); BnaSHP2 (this paper) from *B. napus*; DEF(DEFICIENS) (CAA36268) from *Antirrhinum majus*; SRF (Serum Response Factor) (XP_001136889) from humans; MCM1(MiniChromosome Maintenance1) (CAA32389) from Yeast. More similar MADS-box and K-box regions from BnaSHP2, AtSHP2, BnaSHP1 and AtSHP1 were marked with rectangular box. The shading of the alignment presents identical residues in black, conserved residues in dark gray and similar residues in light gray. Identical or conserved residues among these MADS family proteins were presented at bottom of alignment. The sequences were aligned using the program Multalin with the default parameters (<http://prodes.toulouse.inra.fr/multalin/multalin.html>) (Corpet, 1988), and Clustalx and GeneDoc were used for editing sequence alignment.

while K-box existing in most MADS-box proteins except for SRF and MCM1 was less conservative than MADS-box region. Moreover, the sequence similarity shared between BnaSHP2 and AtSHP2, or BnaSHP1 and AtSHP1 is not restricted to the two conserved regions but extends throughout the entire length of the proteins (92%, data not shown). In addition, BnaSHP2 shared 80% and 84% identity with BnaSHP1 and AtSHP1 respectively (rectangular box in Figure 3). However, FUL, also a MADS-box transcription factor required for pod dehiscence in *Arabidopsis*, only shared 31% identity with BnaSHP2 (Figure 3).

Phylogenetic tree view of *BnaSHP2*

The phylogenetic tree was constructed to show the relationship among MADS-box proteins from yeasts, plants and humans. The famous ABCDE model which is proposed for annotating the pattern of floral development involves many MADS-box proteins such as AtSHP1 and AtSHP2 (Coen and Meyerowitz, 1991; Theissen and

Saedler, 2001; Ferrario et al., 2003). According to this model, these MADS-box proteins can all be divided into classes A, B, C/D and E except for MCM1 and SRF which are from yeast and human respectively (Figure 4). DEF from *Antirrhinum majus* and AP3 from *Arabidopsis* are the members of DEF clade of euAP3 subgroup of class B, while AG belongs to the euAG clade of AG subgroup of class C/D together with FAR. It is obvious that BnaSHP2 belongs to the PLE clade of AG subgroup of class C/D proteins as well as AtSHP1, AtSHP2, BnaSHP1 and PLE. Moreover, BnaSHP2 possesses two conserved motifs in C terminus, AG I motif and AG II motif, which are typical motifs for this subfamily (Figure 4). High bootstrap values of 98 and 100 shared between BnaSHP2 and AtSHP2, BnaSHP1 and AtSHP1 respectively indicated that they are highly homologous. Thus, our analysis indicates that BnaSHP2 and AtSHP2 are orthologous. Since AtSHP1 and AtSHP2 belong to class C-function genes involved in specifying the development of floral organ (Pinyopich et

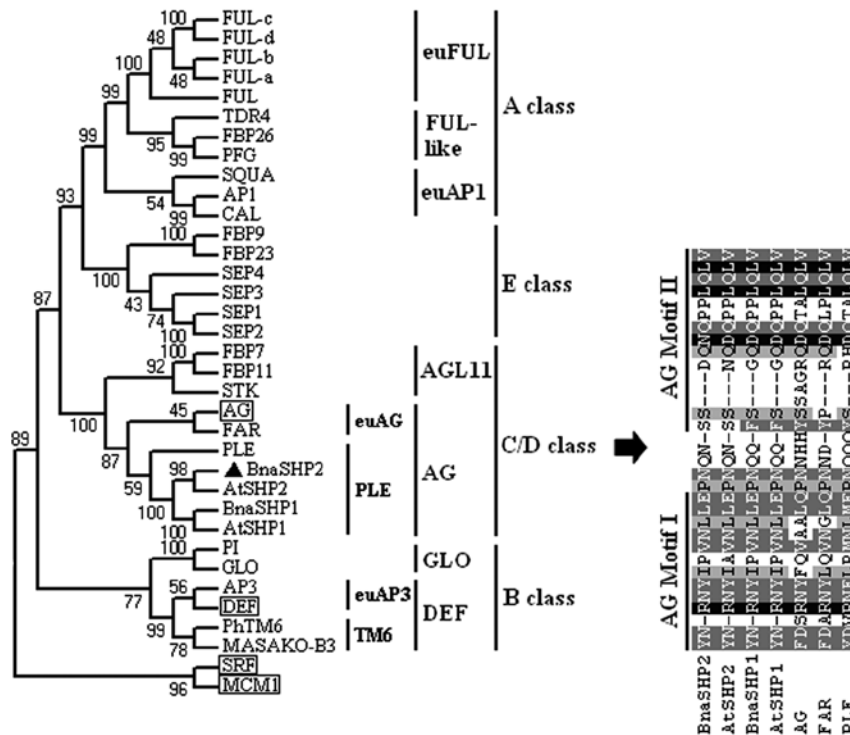


Figure 4. Phylogenetic tree of the MADS-box family proteins from different species including *BnaSHP2*. The analysis was performed with CLUSTALX. Classifications of MADS-box proteins were shown at the right side. *BnaSHP2* was indicated with “—” before. Four primarily MADS-box proteins (MCM1, AG, DEF, SRF) are marked with little rectangular frames. Typical motifs of C/D class, AG I motif and AG II motif, were aligned in their members including *BnaSHP2*. Sequences used for the analysis were obtained from accession numbers: CAL (CAULIFLOWER) (NP_564243); PI (PISTILLATA) (NP_197524); SEP2 (SEPALATA2) (NP_186880); SEP3 (SEPALATA3) (NP_564214); SEP4 (SEPALATA4) (NP_178466) from *Arabidopsis thaliana*; FUL-a (CAD47849); FUL-b (CAD47850); FUL-c (CAD47851) and FUL-d (AJ505844) from *Brassica oleracea*; FAR (FARINELLI) (CAB42988); GLO (GLOBOSA) (CAA48725); PLE (PLENA) (AAB25101); SQUA (CAA45228) from *Antirrhinum majus*; FBP7 (Floral Binding Protein number 7) (CAA57311); FBP9 (AAK21249); FBP11 (CAA57445); FBP23 (AAK21254); FBP26 (AAF19164); PhTM6 (AAS46017); PFG (AAF19721) from *Petunia hybrida*; TDR4 (CAA43169) from *Solanum lycopersicum*; MASAKO-B3 (BAB63261) from *Rosa rugosa*; AtSHP1; AtSHP2; AP1; AP3; AG; STK; SEP1; FUL; *BnaSHP1*; *BnaSHP2*; DEF; SRF; MCM1 referring to Figure 2.

al., 2003), and especially in specifying the different cell types in valve margin (Liljegren et al., 2000). Therefore, due to their orthologous relationships, *BnaSHP1* and *BnaSHP2* may share same functions as *AtSHP1* and *AtSHP2*.

Extron/intron and genomic structure of *BnaSHP2*

The exon/intron structure is important for verifying whether a novel gene is orthologue of a known gene among different species. The alignment of cDNA sequence of *BnaX.SHP2* (*BnaX.SHP2* represents either *BnaA.SHP2.a* or *BnaC.SHP2.b*) against genomic DNA database in TAIR revealed a genomic DNA clone of *B. rapa* subsp. *pekinensis*, KBrH012N11 (AC189588), which shows seven regions with highly identical sequences to cDNA of *BnaX.SHP2*, indicating that there are seven exons in *BnaSHP2*. The exon/intron structures of *AtSHP2* and *BnaSHP2* are showed in Figure 5A. Although the divergence of 5'-URT of *BnaSHP2* cDNAs causes uncertainty of first exon, the whole exon/intron

structures between *AtSHP2* and *BnaSHP2* are similar. To determine the gene copy number of *BnaSHP2* gene in *B. napus* genome, southern blotting was performed. *Bam*HI, *Eco*RI, *Eco*RV, *Hind*III, and *Xba*I were used to digest the genome of *B. napus* respectively. PCR product of *BnaSHP2-EST* was used as probe. The blotting result showed that only one band was detected in lane *Xba*I, two bands in lane *Bam*HI, *Eco*RI, *Eco*RV, and three bands in lane *Hind*III (Figure 5B). In addition, sequence analysis showed that there is no restriction site of *Bam*HI, *Eco*RI, *Eco*RV and *Xba*I in *BnaSHP2-EST*, while there are two restriction sites of *Hind*III in it. So *BnaSHP2* was likely to be a two-copy gene in *B. napus* genome.

Analysis of *BnaSHP2* expression

RT-PCR analysis was used to investigate tissue-distribution of *BnaSHP2* in *B. napus*. Our data showed that the expression of *BnaSHP2* was mainly in root, floral buds and pods, and most strongly in floral buds (Figure 6). And they were not detected in stems, leaves and seeds. This expression profile of *BnaSHP2* was similar to that of

AtSHP2, which is expressed preferentially in floral buds (Ma et al., 1991). Microarray data showed it is mainly expressed in flowers and siliques, and weakly in roots, stems, and leaves (Duarte et al., 2006)

Sub-cellular localization of *BnaSHP2*

To analyze the subcellular localization of *BnaSHP2*, Motif Scan (http://myhits.isb-sib.ch/cgi-bin/motif_scan) was used and a putative bipartite nuclear localization signal (amino acid 24-40) (Figure 2 amino acid residues in rectangular box) was found, suggesting that *BnaSHP2* may localize in nucleus. To investigate whether *BnaSHP2* protein is indeed targeted to the nucleus, a construct of *BnaSHP2-EGFP* fusion gene was made and inserted into HaHZ8 vector to form a recombinant baculovirus, which subsequently infected HzAM1 cells, meanwhile EGFP gene alone was used as control. The HzAM1 cells infected with vHa-*BnaSHP2-EGFP* and vHa-EGFP respectively were examined by fluorescence microscope at 24, 48, 72 and 96 hpi. Green fluorescent protein (GFP) signal was only detected in the nucleus of HzAM1 cells containing *BnaSHP2-EGFP* fusion gene, however, the control construct containing *EGFP* alone showed homogeneous fluorescence in the cytoplasm and nucleus (Figure 7), indicating that *BnaSHP2* is a nuclear protein.

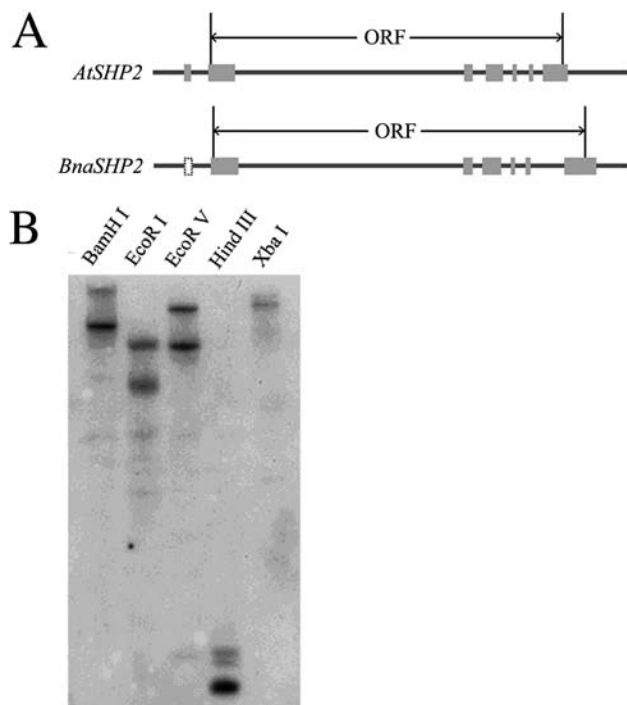


Figure 5. (A) Extron/intron structure of *BnaSHP2* and *AtSHP2*. The introns were shown as line, exons as grey boxes and the first uncertain exon of *BnaSHP2* as dashed line-box. (B) Southern blotting of *BnaSHP2* gene in *B. napus* genome. Genomic DNA was digested with BamHI (line 1), EcoRI (line 2), EcoRV (line 3), HindIII (line 4) and XbaI (line 5) respectively, followed by hybridization with the PCR product of *BnaSHP2* as the probe.

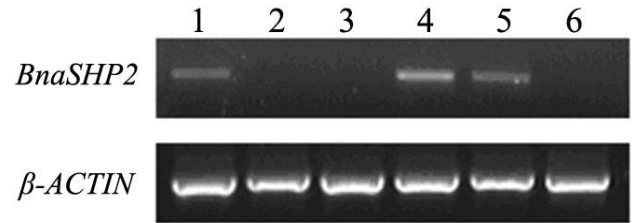


Figure 6. Expression profile of *BnaSHP2* was performed by RT-PCR using β -*ACTIN* as control. 1: roots; 2: stems; 3: leaves; 4: floral buds; 5: pods; 6: seeds.

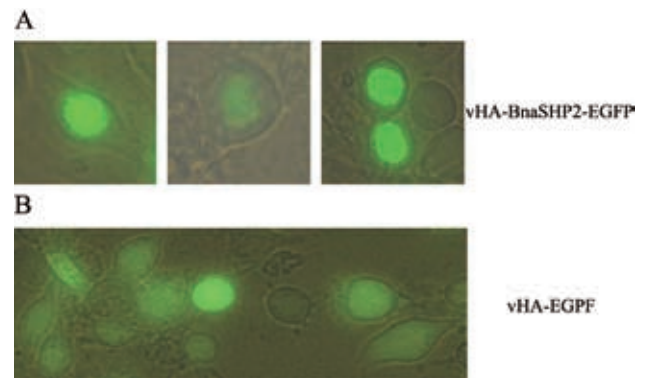


Figure 7. Subcellular localization of *BnaSHP2* in HzAM1 cells. HzAM1 cells were infected with vHa-*BnaSHP2-EGFP* (A) and vHa-EGFP (B) respectively at MOI of 10 and observed at 24 hpi.

DISCUSSION

In this study, the cDNA of *BnaSHP2* was isolated from *B. napus*. Interestingly, our RACE data showed several transcripts with diverse 5' and 3'-end sequences existing in *BnaSHP2*, suggesting there may exist multiple transcription initiation and poly (A) addition sites. Sequence and phylogenetic analysis indicated that *BnaSHP2* belongs to PLE clade of C class MADS box gene, and is orthologous to *AtSHP2*. The expression profile of *BnaSHP2* showed that its transcripts mainly exist in roots and floral tissues in *B. napus*. Similarly to *AtSHP2*, *BnaSHP2* primarily expressed in floral buds and fruits (Ma et al., 1991). This is consistent to the famous ABCDE model of floral organ identity, of which C class MADS box genes specify stamens together with class B genes and alone determine carpels identity (Coen and Meyerowitz, 1991; Weigel and Meyerowitz, 1994; Colombo et al., 1995; Pelaz et al., 2000; Pelaz et al., 2001; Pinyopich et al., 2003). Many studies demonstrated that function of MADS box genes are not restricted to flower organ development, some MADS box genes were detected to be involved in the initiation of flowering, determination of meristem identity (Weigel, 1995), embryonic development (Perry et al., 1999), root formation (Alvarez-Buylla et al., 2000), development of vascular tissue, and seed and fruit formation (Buchner and Boutin, 1998; Gu et al., 1998; Liljegren et al., 1998). On the other hand, *BnaSHP2*

expressed in roots as well as flower and pod while the expression of *AtSHP2* was also detected in root using microarray experiment (Duarte et al., 2006), indicating that *SHP2* is not unique to development of flowers and siliques, but may be also involved in roots development. Whether *SHP2* is related to the specification of cell type in roots, as in flowers and pods, still needs further studies.

Many MADS-box genes involved in flower development have redundant function, which was considered to be caused by gene duplication events during molecule evolution such as *AtSHP1* and *AtSHP2* (Kramer et al., 2004; Causier et al., 2005; Zahn et al., 2006). These genes groups derived from gene duplication and subsequent sequence divergence and intron loss constitute the multi-genes family during evolution (Becker and Theissen, 2003; Irish, 2003). Therefore, *BnaSHP1* and *BnaSHP2*, the orthologues of *AtSHP1* and *AtSHP2* in *B. napus*, should be redundant in function. And they may form similar homodimer or heterodimer with other transcription factors and further recognize and bind similar targets DNA sequence to regulate downstream target genes expressions as Nurrish and Treisman described (Nurrish and Treisman, 1995).

Baculovirus expression system (Bac to Bac) is very useful for expression of gene in eukaryotic cells. In this study, we use this expression system to expressed BnaSHP2-EGFP fusion protein in HzMA1 cells, and the result showed that BnaSHP2 mainly localized in nuclei, which confirmed the characteristic of BnaSHP2 as a transcriptional factor.

In the regulation network of fruit development in *Arabidopsis*, *SHP1*, *SHP2*, *IND* and *ALC* are valve margin identity genes which contribute to margin differentiation, while *FUL* is valve identity gene which restricts the expression of *SHP1*, *SHP2*, *IND* and *ALC* from valve to margin. So when *SHP1*, *SHP2*, *IND* and *ALC* ectopically expressed throughout the valves in *ful* mutant, the pod can not normally elongate and tear at the maturation (Ferrandiz, 2002; Liljegren et al., 2004). In contrast, 35S::FUL transgenic plants pods have a severe phenotype of non-valve margin for its inhibition in the functions of all four genes (*SHP1*, *SHP2*, *IND* and *ALC*), pod can not shatter at maturation (Ferrandiz et al., 2000; Liljegren et al., 2004). Previous study have shown that the ectopic expression of *FUL* gene by 35S promoter in *B. juncea* can produce indehiscent fruit (Østergaard et al., 2006), however, the strong nature of indehiscent phenotype brings difficulty in shattering by combine harvester, suggesting that less severe phenotype of indehiscence will be better and necessary to breeder. In fact, three margin identity genes do not work in a simply linear cascade of gene activity; instead they formed a nonlinear regulatory network. SHP is likely an upstream gene to regulate IND and ALC in this regulatory network (Ferrandiz et al., 2000; Liljegren et al., 2004). SHP, IND and ALC are only redundantly specific for the valve margin and not involved in regulating valves development, so they are ideal

candidates for research and the application in breeding new lines which are suitable for mechanized harvest.

To further understand the characteristics of *BnaSHP2*, RNAi or dominant-negative mutants can be used to inhibit the expression of endogenous *BnaSHP2* in *B. napus* to observe their effects on pod shattering.

Acknowledgments. This work was supported by Jiangsu Academy of Agricultural Sciences, Jiangsu Provincial Postdoctoral Foundation (0601015B) and Talent Foundation of Jiangsu University (05JDG003).

LITERATURE CITED

- Alvarez-Buylla, E.R., S.J. Liljegren, S. Pelaz, S.E. Gold, C. Burgeff, G.S. Ditta, F. Vergara-Silva, and M.F. Yanofsky. 2000. MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. *Plant J.* **24**: 457-466.
- Becker, A. and G. Theissen. 2003. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol. Phylogenet. Evol.* **29**: 464-489.
- Bowman, J.L., S.F. Baum, Y. Eshed, J. Putterill, and J. Alvarez. 1999. Molecular genetics of gynoecium development in *Arabidopsis*. *Curr. Top. Dev. Biol.* **45**: 155-205.
- Buchner, P. and J.P. Boutin. 1998. A MADS box transcription factor of the AP1/AGL9 subfamily is also expressed in the seed coat of pea (*Pisum sativum*) during development. *Plant Mol. Biol.* **38**: 1253-1255.
- Causier, B., R. Castillo, J. Zhou, R. Ingram, Y. Xue, Z. Schwarz-Sommer, and B. Davies. 2005. Evolution in action: following function in duplicated floral homeotic genes. *Curr. Biol.* **15**: 1508-1512.
- Chen, H.J., G.J. Huang, W.S. Chen, C.T. Su, W.C. Hou, and Y.H. Lin. 2009. Molecular cloning and expression of a sweet potato cysteine protease SPCP1 from senescent leaves. *Bot. Stud.* **50**: 159-170.
- Coen, E.S. and E.M. Meyerowitz. 1991. The war of the whorls: genetic interactions controlling flower development. *Nature* **353**: 31-37.
- Colombo, L., J. Franken, E. Koetje, J. van Went, H.J. Dons, G.C. Angenent, and A.J. van Tunen. 1995. The petunia MADS box gene *FBP11* determines ovule identity. *Plant Cell* **7**: 1859-1868.
- Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**: 10881-10890.
- Dinneny, J.R., D. Weigel, and M.F. Yanofsky. 2005. A genetic framework for fruit patterning in *Arabidopsis thaliana*. *Development* **132**: 4687-4696.
- Duarte, J.M., L. Cui, P.K. Wall, Q. Zhang, X. Zhang, J. Leebens-Mack, H. Ma, N. Altman, and C.W. Depamphilis. 2006. Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of *Arabidopsis*. *Mol. Biol. Evol.* **23**: 469-478.

- Ferrandiz, C. 2002. Regulation of fruit dehiscence in *Arabidopsis*. *J. Exp. Bot.* **53**: 2031-2038.
- Ferrandiz, C., S.J. Liljegren, and M.F. Yanofsky. 2000. Negative regulation of the *SHATTERPROOF* genes by *FRUITFULL* during *Arabidopsis* fruit development. *Science* **289**: 436-438.
- Ferrario, S., R.G. Immink, A. Shchennikova, J. Busscher-Lange, and G.C. Angenent. 2003. The MADS box gene *FBP2* is required for *SEPALLATA* function in petunia. *Plant Cell* **15**: 914-925.
- Flanagan, C.A., Y. Hu, and H. Ma. 1996. Specific expression of the *AGL1* MADS-box gene suggests regulatory functions in *Arabidopsis* gynoecium and ovule development. *Plant J.* **10**: 343-353.
- Gu, Q., C. Ferrandiz, M.F. Yanofsky, and R. Martienssen. 1998. The *FRUITFULL* MADS-box gene mediates cell differentiation during *Arabidopsis* fruit development. *Development* **125**: 1509-1517.
- Huang, G.J., S.S. Huang, H.J. Chen, Y.S. Chang, S.J. Chang, H.Y. Chang, P.C. Hsieh, M.J. Chang, Y.C. Lin, and Y.H. Lin. 2009. Cloning and expression of aspartic proteinase cDNA from sweet potato storage roots. *Bot. Stud.* **50**: 149-158.
- Irish, V.F. 2003. The evolution of floral homeotic gene function. *Bioessays* **25**: 637-646.
- Kramer, E.M., M.A. Jaramillo, and V.S. Di Stilio. 2004. Patterns of gene duplication and functional evolution during the diversification of the *AGAMOUS* subfamily of MADS box genes in angiosperms. *Genetics* **166**: 1011-1023.
- Liljegren, S.J., A.H. Roeder, S.A. Kempin, K. Gremski, L. Østergaard, S. Guimil, D.K. Reyes, and M.F. Yanofsky. 2004. Control of fruit patterning in *Arabidopsis* by *INDEHISCENT*. *Cell* **116**: 843-853.
- Liljegren, S.J., C. Ferrandiz, R. Alvarez-Buylla, S. Pelaz, and M.F. Yanofsky. 1998. *Arabidopsis* MADS-box genes involved in fruit dehiscence. *Flowering Newsl.* **25**: 9-19.
- Liljegren, S.J., G.S. Ditta, Y. Eshed, B. Savidge, J.L. Bowman, and M.F. Yanofsky. 2000. *SHATTERPROOF* MADS-box genes control seed dispersal in *Arabidopsis*. *Nature* **404**: 766-770.
- Ma, H., M.F. Yanofsky, and E.M. Meyerowitz. 1991. *AGL1-AGL6*, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. *Genes Dev.* **5**: 484-495.
- Meakin, P.J. and J.A. Roberts. 1990. Dehiscence of fruit in oilseed rape. II. The role of cell wall degrading enzymes. *J. Exp. Bot.* **41**: 1003-1011.
- Nurrish, S.J. and R. Treisman. 1995. DNA binding specificity determinants in MADS-box transcription factors. *Mol. Cell Biol.* **15**: 4076-4085.
- Østergaard, L. and G.J. King. 2008. Standardized gene nomenclature for the *Brassica* genus. *Plant Methods* **4**: 10.
- Østergaard, L., S.A. Kempin, D. Bies, H.J. Klee, and M.F. Yanofsky. 2006. Pod shatter-resistant *Brassica* fruit produced by ectopic expression of the *FRUITFULL* gene. *Plant Biotechnol. J.* **4**: 45-51.
- Pelaz, S., G.S. Ditta, E. Baumann, E. Wisman, and M.F. Yanofsky. 2000. B and C floral organ identity functions require *SEPALLATA* MADS-box genes. *Nature* **405**: 200-203.
- Pelaz, S., R. Tapia-Lopez, E.R. Alvarez-Buylla, and M.F. Yanofsky. 2001. Conversion of leaves into petals in *Arabidopsis*. *Curr. Biol.* **11**: 182-184.
- Perry, S.E., M.D. Lehti, and D.E. Fernandez. 1999. The MADS-domain protein *AGAMOUS*-like 15 accumulates in embryonic tissues with diverse origins. *Plant Physiol.* **120**: 121-130.
- Pinyopich, A., G.S. Ditta, B. Savidge, S.J. Liljegren, E. Baumann, E. Wisman, and M.F. Yanofsky. 2003. Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* **424**: 85-88.
- Rajani, S. and V. Sundaresan. 2001. The *Arabidopsis* *myc/bHLH* gene *ALCATRAZ* enables cell separation in fruit dehiscence. *Curr. Biol.* **11**: 1914-1922.
- Romkaew, J. and T. Umezaki. 2006. Pod dehiscence in soybean: assessing methods and varietal difference. *Plant Prod. Sci.* **9**: 373-382.
- Sambrook, J. and D.W. Russell. 2001. *Molecular Cloning: a Laboratory Manual*. CSHL press, New York, pp. 639-662.
- Spence, J., Y. Vercher, P. Gates, and N. Harris. 1996. 'Pod shatter' in *Arabidopsis thaliana*, *Brassica napus* and *B. juncea*. *J. Microscopy* **181**: 195-203.
- Theissen, G. and H. Saedler. 2001. Plant biology: Floral quartets. *Nature* **409**: 469-471.
- Weigel, D. 1995. The genetics of flower development: from floral induction to ovule morphogenesis. *Annu. Rev. Genet.* **29**: 19-39.
- Weigel, D. and E.M. Meyerowitz. 1994. The ABCs of floral homeotic genes. *Cell* **78**: 203-209.
- Wu, H., A. Mori, X. Jiang, Y. Wang, and M. Yang. 2006. The *INDEHISCENT* protein regulates unequal cell divisions in *Arabidopsis* fruit. *Planta* **224**: 971-979.
- Zahn, L.M., J.H. Leebens-Mack, J.M. Arrington, Y. Hu, L.L. Landherr, C.W. de Pamphilis, A. Becker, G. Theissen, and H. Ma. 2006. Conservation and divergence in the *AGAMOUS* subfamily of MADS-box genes: evidence of independent sub- and neofunctionalization events. *Evol. Dev.* **8**: 30-45.

甘藍型油菜中與角果開裂有關的基因 *SHP2* 的選殖及分析

譚小力^{1,2} 夏宗偉² 張麗麗² 張志燕² 郭忠建² 戚存扣¹

¹ 中國南京江蘇省農業科學院 經濟作物研究所

² 中國鎮江江蘇大學 生命科學研究院

擬南芥中的 SHATTERPROOF1 (SHP1) and SHATTERPROOF2 (SHP2) 屬 MADS-box 家族轉錄因數，它們通過調控角果皮邊沿細胞的分化，來控制角果的開裂。用擬南芥 *SHP2* 的序列在油菜的序列資料庫中比對，得到一批序列相似性較高的 EST，拼接成一個含有開放閱讀框的 cDNA，並在甘藍型油菜中得到選殖，命名為 *BnaSHP2*，但其 5' 和 3' 的非翻譯區序列不全。通過 RACE，得到了兩個轉錄序列，分別命名為 *BnaA.SHP2.a* 和 *BnaC.SHP2.b*。本研究對其中一條序列 *BnaA.SHP2.a* 做進一步分析。*BnaSHP2* 蛋白和其他 MADS-box 轉錄因數序列比對和系統分析表明，油菜 *BnaSHP2* 和擬南芥 *AtSHP2* 序列具有很高的相似性。Southern 雜交結果顯示 *BnaSHP2* 基因在油菜基因組中至少存在兩個同源基因；半定量反轉錄 PCR 結果說明 *BnaSHP2* 主要在根、花蕾和角果中表達，而在莖、葉和種子中不表達。把 *BnaSHP2* 和綠色螢光蛋白融合 (EGFP)，通過杆狀病毒系統，在棉鈴蟲細胞中表達，發現 *BnaSHP2* 只在細胞核中表達，證實了它作為轉錄因數的核定位特性。

關鍵詞：擬南芥；*BnSHP2* 基因；甘藍型油菜；角果開裂；MADS-box 家族轉錄因數；序列比對。