Establishment and analysis of the mRNA expression patterns of ABI3-like and storage protein genes during soybean seed development

Ruochen Du¹, Yonggang Qiao¹, Xiao Wang², Xinyun Lv², Jinsheng Wang¹*

¹College of Agriculture, ShanXi Agriculture University, Jin Zhong 030801, China; ²College of Life Science, ShanXi Agriculture University, Jin Zhong 030801, China

ABSTRACT

Abscisic acid insensitive 3 (ABI3) has a regulatory effect on the accumulation of storage proteins, as well as on the dormancy, quality, and drought tolerance of Arabidopsis seeds during development. The storage protein genes (SPGs) and relevant proteomics are already well understood; however, the mRNA expression patterns of the ABI3-like gene in soybean remain unknown. In this paper, the expression of ABI3-like and SPGs in soybean seeds was evaluated at the transcriptional level, which can provide a theoretical basis for soybean molecular breeding. Soybeans of the type ‘NC111-1’ were used as the qRT-PCR test material, in order to study the mRNA expression patterns of ABI3-like factors (GmABI3-3-8 and GmABI3-3-18), soybean SPGs (Gy3-1, Gy3-2, Gy3-3, Gy3-4, Gy3-5, Gy3-7, 7sx, 7sx’, 7sβ), and their correlation. The results demonstrate that the mRNA expression patterns of GmABI3-3-8 and GmABI3-3-18 include bimodal curves, with the two peak periods corresponding to the filling and dehydration stages of soybean, respectively. The mRNA expression patterns of all SPGs exhibited single peak curves. 7sx, Gy3-1, and 7sxβ play major roles, and the expression level of 7sx was significantly higher than that of other SPGs (P<0.05). The expression levels of GmABI3-3-8 mRNA and each SPG (excluding Gy3-7) mRNA were significantly linearly correlated (P<0.05, r>0.8) in the filling stage, while GmABI3-3-18 mRNA was highly expressed during the dehydration stage. These new details about the specific transcription factors and regulatory roles of each storage protein, as well as the specific metabolic pathway of ABI3, show that GmABI3-3-8 and GmABI3-3-18 play an important role in the development of soybean seeds. This new data will help facilitate the development of molecular breeding methods for the production of new high-protein soybean varieties.

Keywords: Soybean seed; GmABI3-3-8; GmABI3-3-18; Storage protein genes; qRT-PCR

INTRODUCTION

The abscisic acid insensitive 3 (ABI3) transcription factor is a member of the LAV family in the Arabidopsis B3 superfamily, and has regulatory effects on abscisic acid (ABA) responsive genes, as well as the dormancy, quality, and drought tolerance of developing Arabidopsis seeds (Suzuki et al., 2008; Sugliani et al., 2010; Zeng et al., 2013; Roscoe et al., 2015; Mao et al., 2015). The soybean ABI3-like gene GmABI3-3-8 is located on chromosome eight (Acc. No.: XM_003532261.2), and GmABI3-3-18 is located on chromosome eighteen (Acc. No.: XM_006602493.1). Soybean storage proteins are categorized according to their sedimentation coefficients into 11S and 7S types. The genes of the 11S subunit include Gy3-1, Gy3-2, Gy3-3, Gy3-4, Gy3-5, and Gy3-7, and the genes of the 7S subunit mainly include CG-alpha-1 (7sa), CG-alpha-1’ (7sa’), and CG-beta-1 (7sb) (Nielsen et al., 1997; Harada et al., 1989; Beilinson et al., 2002; Asakura et al., 2012). The expression level of Gy3-7 is very low (Beilinson et al., 2002). As a cis-acting element, the RY motif is commonly used as a seed-specific promoter. Most legume storage protein genes contain one or more RY repeating elements, which are the only common regulatory sequences (Nielsen et al., 1989; Li et al., 2011). Studies have shown that the binding of the ABI3 transcription factor with the RY motif can regulate the accumulation of storage proteins in Arabidopsis seeds, as well as the physiological processes of dormancy (Romanel et al., 2009; Sakata et al., 2010; Delahaie et al., 2013). Elimination of the RY repeating elements can affect the spatial structures of the associated genes, resulting in a decrease in the binding activity of the transcription factor and the upstream elements of the promoter (Lelievre et al., 1992). In 2015, a study of the ABI3 transcription factor in maize demonstrated that

*Corresponding author:
Jinsheng Wang, College of Agriculture, Shan Xi Agriculture University, Jin Zhong 030801, China, Fax:+86 354 6287006.
Tel: +86 354 6287006, China, Fax: +86 354 6287006. Tel: +86 354 6287006, E-mail: edu_sxndwjs@126.com

Received: 08 January 2016; Revised: 02 February 2016; Accepted: 02 February 2016; Published Online: 14 February 2016
it also plays an important role during the filling period (Grimault et al., 2015). The storage proteins exhibited similar metabolic patterns during soybean proteomics research (Hajduch et al., 2005; Xu et al., 2015). The soybean genome has been sequenced, and the relevant proteomics research is complete; however, there have been no studies investigating the expression patterns of GmABI3-8 and GmABI3-18 mRNA and soybean SPGs mRNA. Data on the mRNA-level regulation of ABI3-like transcription factors in soybean seeds could provide a theoretical basis for soybean molecular breeding. Thus, the sample material used in this study consisted of soybean seeds in different post-flowering growth stages, in order to study the mRNA expression patterns of GmABI3-8, GmABI3-18, Gy1, Gy2, Gy3, Gy4, Gy5, Gy7, 7sa, 7sa', and 7sβ mRNA in the NCBI GenBank, the RT-PCR primers were designed using Primer premier 5.0 (Premier Co. Ltd., Canada) (Table 1). Primers were also designed for the CYP2 gene, which was used as a housekeeping gene (Jian et al., 2008) (Table 1). Then, 50 DAF cDNA was used as a template to detect specific primers, using the following amplification program: 94˚C for 5 min; 35 cycles of 94˚C for 30 s, 57˚C for 30 s, 72˚C for 30 s, and 72˚C for 5 min. The products were examined using 3% agarose gel electrophoresis, then sequenced.

**MATERIALS AND METHODS**

**Plant materials**
The ‘NC111-1’ (G. max ‘NC111-1’) soybeans were provided by the Life Science College of Shanxi Agricultural University. ‘NC111-1’ soybeans that were planted in Dongli Village, Taigu County, Shanxi Province were collected. Soybean flowering was recorded as day 0, and a plate was hung with the marked flowers. No less than 500 marked seeds were collected within the 10d after flowering (DAF), immediately frozen in liquid nitrogen, then stored for long-term preserved at -80˚C. The process was terminated when the soybean seeds had matured and reached the standard age for preservation (Fehret et al., 1971). This process was repeated twice from summer to autumn in 2013 and 2014, resulting in ten samplings per year.

**Physiological data determination**
The length, width, and thickness of the seeds from each sampling period were measured using an ultramicrometer (HANGGONG Inc., Hangzhou, China). The fresh weight of 100 seeds was measured using an electronic balance (Sartorius Inc., Beijing, China), while the dry weight was measured after drying the seeds to a constant weight, in order to calculate their water content. The protein content of the seeds was measured using the Kjeldahl nitrogen determination method (Jung et al., 2003), with a conversion coefficient of 5.71 (F=5.71). The protein quality within a single seed was calculated according to the dry weight and protein content. The test was repeated three times.

**RNA isolation and cDNA synthesis**
Total RNA was extracted using a Quick RNA Isolation Kit (HUAYUEYANG Bio Co. Ltd., Beijing, China) according to the manufacturer’s protocol. The total RNA concentration of each sample was quantified using an ND-1000 spectrophotometer (Nano-drop Technologies, Wilmington, DE, USA), after which the samples were stored at -80˚C. The RNA was detected via agarose gel electrophoresis. The cDNA were synthesized using a Transcriptor First Strand cDNA Synthesis Kit (Roche Bio Co. Ltd., Shanghai, China). The reverse transcription product was stored at -20˚C, after standardizing the concentration.

**Primer specificity detection**
Based on the submitted sequences of soybean GmABI3-8, GmABI3-18, Gy1, Gy2, Gy3, Gy4, Gy5, Gy7, 7sa, 7sa', and 7sβ mRNA in the NCBI GenBank, the RT-PCR primers were designed using Primer premier 5.0 (Premier Co. Ltd., Canada) (Table 1). Primers were also designed for the CYP2 gene, which was used as a housekeeping gene (Jian et al., 2008) (Table 1). Then, 50 DAF cDNA was used as a template to detect specific primers, using the following amplification program: 94˚C for 5 min; 35 cycles of 94˚C for 30 s, 57˚C for 30 s, 72˚C for 30 s, and 72˚C for 5 min. The products were examined using 3% agarose gel electrophoresis, then sequenced.

**RNA quality and primer specificity detection**
The RNA agarose gel electrophoresis results (Fig. 1a) show the RNA to be intact during each sampling period. Conventional RT-PCR detection (Fig. 1b) demonstrated that there are no primer dimers. The sequencing results were checked using BLAST, and the homology was found to be over 95%.

**Quantitative real-time PCR**
Quantitative real-time PCR (qRT-PCR) was carried out according to the protocol provided with the SYBR Premix Ex TaqTM II Kit (TaKaRa Bio Co. Ltd., Dalian, China). qRT-PCR was carried out using Bio-Rad CFX 3.1 (Bio-Rad Bio Co. Ltd., CA, USA), and water, rather than the template, was used as a negative control. The process was carried out three times per gene, using the following amplification program: 95˚C for 3min; 40 cycles of 95˚C for 10s, 57˚C for 30s, and 72˚C for 20s; 70˚C for 5s; 95˚C for 5s. Then, the CFX Manager 3.1 software program (Bio-Rad Bio Co. Ltd., CA, USA) was used to analyze the melting curve. The ΔΔCt to the housekeeping gene (Pfaffl., 2001) was used for relative quantitation.

**Data analysis**
Photographs were taken with FinePix S9500 Digital Camera (FUJIFILM Co. Ltd., Shanghai, China). The mean and standard error of the data were calculated using Microsoft Excel 2003 (Microsoft Co. Ltd., Beijing, China). The significance, correlation, and regression of the data
Du, et al.: mRNA expression patterns of storage protein genes


were analyzed using SPSS 19.0 (IBM Co. Ltd., NY, USA). The data generated above were converted to graphs using SigmaPlot 12.5 (Systat Software Inc., UK). The results were expressed as the mean ± standard error (X±SE).

RESULTS

Physiological data analysis and stage division
The physiological indices of the soybean seeds and pods in each period were summarized. The results (Fig. 2) show that pod length varies significantly over time, reaching its maximum at 70 DAF; small changes were observed in the width and the thickness; the trends in seed length, width, and height were similar, changing significantly after 30 DAF, and reaching their maximum at 70 DAF before decreasing; seed fresh weight reached its maximum at 70 DAF and then decreased, while the dry weight continued to increase; the water content of the seeds exhibited an overall downward trend, with a trough centered at 20 DAF; seed protein content varied within 42.31±5.25%, the maximum and minimum values being recorded at 20 DAF and 50 DAF, respectively; the single seed protein mass variation curve formed an “S” shape. According to the changes in the indexes above, the whole development period of the seeds was divided into two main periods: a seed filling stage (0-70 DAF), and a dehydration stage (70-100 DAF).

The mRNA expression profiles of GmABI3-8 and GmABI3-18
The GmABI3-8 and GmABI3-18 qRT-PCR results (Fig. 3) showed that GmABI3-8 and GmABI3-18 mRNA varied along a bimodal curve during the entire period of seed growth and development. The expression level of GmABI3-8 was low at 10DAF, 20DAF, and 30DAF, then increased suddenly, peaking at 40DAF before decreasing.

Table 1: Primers used in this study

<table>
<thead>
<tr>
<th>Name of primers</th>
<th>Acc. no.</th>
<th>Sequences of primers (5ʹ-3ʹ)</th>
<th>Product (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GmABI3-8</td>
<td>XM_003532261.2</td>
<td>F: TTTCTTGACCGTGCCTCCTATC R: CACCAATTGCTATTGTCCTTC</td>
<td>91</td>
</tr>
<tr>
<td>GmABI3-18</td>
<td>XM_006602493.1</td>
<td>F: CAGCAAGGGAGAACAGAAGAG R: CATCGTCTACTCCAACACCATC</td>
<td>55</td>
</tr>
<tr>
<td>Gy1</td>
<td>NM_001248898</td>
<td>F: TTGGTGAACGCATTACCAAGAG R: GAGGGGAACGAGGAACTGGA</td>
<td>108</td>
</tr>
<tr>
<td>Gy2</td>
<td>D00216</td>
<td>F: CGGGTTGTTCCCTAGCCTAATC R: CTCCTGAAAGGGGTGCTAATC</td>
<td>79</td>
</tr>
<tr>
<td>Gy3</td>
<td>AB030494</td>
<td>F: CGTGGACAGGAGATAGGAGG</td>
<td>R: TCCTGAGGTGCTTGTGCTG</td>
</tr>
<tr>
<td>Gy4</td>
<td>NM_001251079</td>
<td>F: AGAGGATGCGAGACAAGAAG R: GAGGGGCTTGGAGGGTACTAATG</td>
<td>110</td>
</tr>
<tr>
<td>Gy5</td>
<td>NM_001249747</td>
<td>F: CCCCCCTCCTACCCCTCCTC R: GCGTAGTTGTAGGTCCAGAAG</td>
<td>90</td>
</tr>
<tr>
<td>Gy7</td>
<td>NM_001248425</td>
<td>F: ACATGATGCAGCCAACCAAGAG</td>
<td>R: GCCTACTGGTCCCTGCTGCAGAAG</td>
</tr>
<tr>
<td>7sx</td>
<td>M26128</td>
<td>F: CAGGGAGGACACACCTTTTGA</td>
<td>R: CTGGGAAACTCCTGCTGGGATTAC</td>
</tr>
<tr>
<td>7sx'</td>
<td>AB234094</td>
<td>F: ACCATGCTGACGCTGATTAC R: CGGCCAGATTTGAGGGTATGAA</td>
<td>80</td>
</tr>
<tr>
<td>7sf</td>
<td>NM_001249943</td>
<td>F: GTCGAAACCTACCTTACACCT</td>
<td>R: CTCCCTGCATTGCTTTTCTATCT</td>
</tr>
</tbody>
</table>

Fig 1. Agarose gel electrophoresis Agarose gel electrophoresis of RNA: lane M contains the 2000 bp DNA marker (Real-Times Biotechnology Co. Ltd., Beijing, China), lanes 1 to 10 contain 10 DAF to 100 DAF. B.) Agarose gel electrophoresis of RT-PCR amplification; lane M contains 5000 bp DNA marker (Real-Times Biotechnology Co. Ltd., Beijing, China), lanes 1 to 12 contain Gy1, Gy2, Gy3, Gy4, Gy5, Gy7, 7sx, 7sx’, 7sf, GmABI3-8, GmABI3-18, and CYP2 RNA, respectively.
This was followed by a second peak at 80DAF, after which the expression level decreased again. The expression level of GmABI3-18 was low at 10DAF, 20DAF, and 30DAF, then increased suddenly, peaking at 50DAF before decreasing. This was followed by a second peak at 80DAF, after which expression levels decreased again. The expression of GmABI3-18 was higher than that of GmABI3-8 after 70DAF.
The expression patterns of SPG mRNAs

The qRT-PCR results of each SPG (Fig. 4) indicate that the expression of Gy1, Gy2, Gy3, Gy4, Gy5, 7sa, 7a’, and 7β mRNA occurred along unimodal curves. The expression levels of Gy1, Gy2, Gy3, Gy4, Gy5, 7sa, and 7a’ mRNA were the lowest overall at 20DAF, while that of 7β mRNA was the lowest overall at 30DAF. There was a sudden significant increase at 40DAF, and a significant decrease after the period of high expression between 40 and 70DAF. The expression levels of Gy1, Gy2, Gy3, Gy4, Gy5, 7sa’, and 7β mRNA peaked at 50DAF, and 7a’ mRNA reached its maximum at 40DAF. The qRT-PCR results for Gy7 mRNA (Fig. 5) show that Gy7 expression varies in a manner similar to those of GmABI3-8 and GmABI3-18, with the two peaks appearing at 50DAF and 90DAF, respectively. Each SPG relative expression level from the main expression period (40-70DAF) was aggregated. The results (Fig. 6a) indicate that the most highly expressed mRNA during each period was as follows: 7sa at 40DAF and 60DAF, Gy1 at 50DAF, and 7β at 70DAF. The total expression level of each SPG over the whole development period (10-100DAF) (Fig. 6b) from high to low was, in turn: 7sa, Gy1, 7β, Gy4, 7sa’, Gy5, Gy2, Gy3, and Gy7. The expression level of 7sa was significantly higher than that of other SPGs, while that of Gy7 was significantly lower than that of other SPGs (P<0.05).

Correlation and regression analysis of GmABI3-8, GmABI3-18, and SPGs

A correlation analysis of GmABI3-8, GmABI3-18, and each SPG mRNA expression level during the seed filling stage (Table 2) showed that GmABI3-8 was significantly linearly correlated with Gy1, Gy2, Gy3, Gy4, Gy5, 7sa, 7a’, and 7β (P<0.05, r>0.8); GmABI3-18 was significantly linearly correlated with Gy7 (P<0.05, r>0.8); ABI3-like was significantly linearly correlated with Gy3, Gy4, Gy5, Gy7, 7sa’, and 7β (P<0.05, r>0.8).

![Fig 3. GmABI3-8 and GmABI3-18mRNA relative expression pattern GmABI3-8 mRNA relative expression pattern. B.) GmABI3-18 mRNA relative expression pattern. Each value is the average of 3 datasets; SE is denoted by error bars.](image-url)

**Table 2: ABI3-like and SPGs correlation analysis**

<table>
<thead>
<tr>
<th></th>
<th>GmABI3-8</th>
<th>GmABI3-18</th>
<th>ABI3-like*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CC</td>
<td>P value</td>
<td>Re</td>
</tr>
<tr>
<td>Gy1</td>
<td>0.903</td>
<td>0.005</td>
<td>y=1276.876x-0.819</td>
</tr>
<tr>
<td>Gy2</td>
<td>0.858</td>
<td>0.014</td>
<td>y=1924.350x-0.236</td>
</tr>
<tr>
<td>Gy3</td>
<td>0.951</td>
<td>0.001</td>
<td>y=224.631x-0.142</td>
</tr>
<tr>
<td>Gy4</td>
<td>0.957</td>
<td>0.001</td>
<td>y=81.054x-0.493</td>
</tr>
<tr>
<td>Gy6</td>
<td>0.964</td>
<td>0.000</td>
<td>y=122.676x+0.000</td>
</tr>
<tr>
<td>Gy7</td>
<td>0.773</td>
<td>0.041</td>
<td>y=0.369x+0.000</td>
</tr>
<tr>
<td>7sa</td>
<td>0.955</td>
<td>0.001</td>
<td>y=1920.670x-1.256</td>
</tr>
<tr>
<td>7sa’</td>
<td>0.969</td>
<td>0.000</td>
<td>y=467.295x-0.276</td>
</tr>
<tr>
<td>7β</td>
<td>0.837</td>
<td>0.019</td>
<td>y=859.882x-0.061</td>
</tr>
</tbody>
</table>

*ABI3-like: The sum of GmABI3-8 and GmABI3-18 mRNA relative expression. CC: Correlation coefficient. Re: Regression equation

**Table 3: ABI3-like and SPGs correlation analysis**

<table>
<thead>
<tr>
<th></th>
<th>GmABI3-8</th>
<th>GmABI3-18</th>
<th>ABI3-like</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CC</td>
<td>P value</td>
<td>Re</td>
</tr>
<tr>
<td>11S</td>
<td>0.934</td>
<td>0.002</td>
<td>y=3127.956x+1.967</td>
</tr>
<tr>
<td>7S</td>
<td>0.966</td>
<td>0.000</td>
<td>y=3247.846x+1.266</td>
</tr>
<tr>
<td>All genes</td>
<td>0.966</td>
<td>0.000</td>
<td>y=8375.802x-3.560</td>
</tr>
</tbody>
</table>

CC: Correlation coefficient. Re: Regression equation
Fig 4. mRNA relative expression pattern of each SPG Figure A to H illustrate the mRNA relative expression patterns of Gy1, Gy2, Gy3, Gy4, Gy5, Gy7, 7sα, 7sα', 7sβ, respectively. Each value is the average of 3 datasets; SE is denoted by error bars.
Correlation analyses of *GmABI3*-8, *GmABI3*-18, and 11S protein genes, 7S protein genes, and all SPGs mRNA expression levels during the seed filling stage (Table 3) showed that *GmABI3*-8 was significantly linearly correlated with 11S protein genes, 7S protein genes, and all SPGs (*P*<0.05, *r*>0.8); *ABI3*-like was significantly linearly correlated with 11S protein genes, 7S protein genes, and all SPGs (*P*<0.05, *r*>0.8).

The regression analyses (Fig 7, Tables 2 and 3) demonstrate that the order of the slopes of regression equations of *GmABI3*-8 with each SPG from high to low is: 7*α*, G*γ*1, 7*β*, G*γ*4, 7*α′*, G*γ*5, G*γ*2, G*γ*3, and G*γ*7. The order of the slopes of the regression equations of *GmABI3*-18 with each SPG from high to low is: 7*α*, G*γ*1, 7*β*, G*γ*4, 7*α′*, G*γ*2, G*γ*5, G*γ*3, and G*γ*7. The order of the slopes of the regression equations of *ABI3*-like with each SPG from high to low is: 7*α*, G*γ*1, 7*β*, G*γ*4, 7*α′*, G*γ*5, G*γ*2, G*γ*3, and G*γ*7. The order of the slopes of the regression equations of *GmABI3*-8 with 11S protein genes, 7S protein genes, and all SPGs from high to low is: all SPGs, 7S protein genes, and 11S protein genes. The order of the slopes of the regression equations of *GmABI3*-18 with 11S protein genes, 7S protein genes and all SPGs from high to low is: all SPGs, 7S protein genes, and 11S protein genes. The order of the slopes of the regression equations of *ABI3*-like with 11S protein genes, 7S protein genes, and all SPGs from high to low is: all SPGs, 7S protein genes, and 11S protein genes.

**DISCUSSION**

**Regularity of soybean growth and metabolism**

The process of soybean seed growth and development was divided into 4 stages (Meinke et al., 1981), based on changes in pod and seed size: morphogenesis and cell division, cell enlargement, seed maturation, and finally a desiccation and dormancy stage (Meinke et al., 1981). According to the changes in the indexes of our physiological data, the entire process of seed development was divided into two main periods: the seed filling stage (0-70 DAF), and the dehydration stage (70-100 DAF). This division is more suitable for studies investigating the changes in the storage protein. In this study, the seed pods grew more rapidly than the soybean seeds themselves. The dehydration process begins after the seeds completely fill the pod; in addition, the pods dehydrate more slowly than the seeds. These indicate that seed growth of seed is not limited to the interior space of the pod. On the contrary, it is likely that the seeds filling the pod trigger the dehydration process. The transcription factors and signaling pathways involved in this process are currently the subject of a great deal of research (Adams et al., 1983; Belamkar et al., 2014). The proteomics research conducted by Hajduch et al. (2005) on soybean seeds in the seed filling stage discovered that the storage protein content increased continuously; however, there is no available data about the dehydration phase. In this study, the total protein content did not exhibit any expected gradual upward trends. On the contrary, it peaked at 20DAF, while the protein content per seed continued to increase. This indicates that during the
Du, et al.: mRNA expression patterns of storage protein genes

Fig 7. Equations of regression analyses of SPGs and ABI-like A.) Regression equations of GmABI-8 and each SPG; B.) Regression equations of GmABI-18 and each SPG; C.) Regression equations of ABI-like and each SPG; D.) Regression equations of GmABI-8 and 11S protein genes, 7S protein genes, and all SPGs; E.) Regression equations of GmABI-18 with 11S protein genes, 7S protein genes, and all SPGs; F.) Regression equations of ABI-like and 11S protein genes, 7S protein genes, and all SPGs. Each value is the average of 3 datasets; SE is denoted by error bars.

first stage of the seed development process, the expressed proteins were primarily those associated with growth and development. Although the protein percentage did not change significantly throughout the development period, the type and content of the proteins changed significantly. The seed fresh weight reached its maximum at 70DAF, after which the dry weight did not change significantly. Meanwhile the water content decreased, resulting in a significant decrease in seed weight. Therefore, using seeds at 70DAF (the end of the filling stage) for animal feed and other production applications is a more efficient use of resources.

SPGs expression patterns
The expression of the global soybean gene was detected by Asakura et al. (2012). Using a DNA microarray analysis method, they found that the expression of Gy1, Gy2, Gy3, Gy4, 7sa, 7sa', and 7sβ gradually increased during the seed filling stage (Asakura et al., 2012). The results obtained by Asakura et al. (2012) do not include changes in Gy5 and Gy7 expression, or changes taking place during the dehydration stage. In this study, the SPGs mRNA was found to be more highly expressed during the entire development period. The results reveal single-peak curves, and the expression levels were low at 10-30DAF, indicating that the storage proteins were almost no synthesized during this period. Studies show that the primary physiological activity of the soybean seeds during this period was morphogenesis; that is, most of the proteins expressed were metabolic proteins (Meinke et al., 1981). The mRNA of SPGs (excepting Gy7) was very highly expressed at 40-70DAF, which shows that
this is a period of peak storage protein synthesis in soybean seeds. This confirms the results of proteomics studies that used analyses of mRNA to elucidate the changes in storage protein content that occur during soybean seed development; that is, the changes in the storage protein content are consistent with the changes in SPGs mRNA expression levels (Hajduch et al., 2005). During this period of development, the most highly expressed mRNA was 7α at 40DAF and 60DAF, Gγ1 at 50DAF, and 7β at 70DAF, indicating that identity of the dominant expressed SPG changed overtime. Over the entire development period, the total expression level of 7α mRNA was significantly higher than that of the other SPGs, indicating that 7α was the most active SPG. In addition, Gγ1 and 7β mRNAs were also highly expressed. Therefore, 7α, Gγ1, and 7β were the three most highly expressed SPGs. The study showed that synthesis of soybean seed storage substances decreased significantly during the dehydration stage (Jin et al., 2014). In this study, the expression of SPGs decreased significantly after 70DAF, followed by small changes at relatively low levels, indicating that storage protein synthesis was mostly completed prior to the start of the dehydration stage. This indicates that little protein synthesis and accumulation took place during the dehydration stage. In the study carried out by Bellinson V et al. (2012), Gγ7 was detected midway through soybean seed development. This study describes the expression of Gγ7 throughout seed development: the expression of Gγ7 was significantly lower than that of other SPGs, and its expression levels varied differently from those of the other SPGs. However, Gγ7 displayed a metabolic pattern similar to that of ABI3-like, both in its expression level and in its expression trend. It has been proposed to have more complex functions, as it seems to act as a storage protein, but differs significantly from the storage protein gene and seems similar to ABI3 transcription factors; however, more research is required to fully elucidate its true function. In summary, the expression of SPGs is time-specific, with metabolic patterns differing among development stages, with clear differences in the expression levels.

**Effect of GmABI3-8 and GmABI3-18 on SPGs**

The expression of GmABI3-8 and GmABI3-18 mRNA exhibits double-peaked curves. During the seed filling stage, the early peak of GmABI3-8 mRNA occurred earlier than that of GmABI3-18. This suggests that the impact of the GmABI3-8 transcription factor on SPGs took effect earlier than that of GmABI3-18, indicating that after soybeans had experienced polyploidization during evolution, the metabolic patterns of the related genes featured certain differences in timing, but the overall trend was unchanged. The correlation analysis of mRNA expression during seed filling showed that the expression of GmABI3-8 and GmABI3-18 has a strong linear correlation with the mRNA expression of the SPGs. This indicates that the GmABI3-8 and GmABI3-18 transcription factors influence and regulate storage protein synthesis. The expression of GmABI3-8 was more significantly linearly correlated with the expression of SPGs than that of GmABI3-18. This indicates that GmABI3-8 had a more significant regulatory effect on SPGs. The results of the regression analysis show that GmABI3-8 and GmABI3-18 had positive regulatory effects on SPGs, and that their effects differed. Therefore, by further studying the mechanisms of GmABI3-8 and SPGs expression, the protein content of soybean seeds can be increased using molecular biology techniques, improving germplasm resources. The study conducted by Gagete et al. (2009), focusing on a functional analysis of the isoforms of an ABI3-like factor of Pisum sativum that is generated by alternative splicing, suggests that the common RY repeating elements of legume seed SPGs are very likely to be the target of ABI3-like type transcription factors. The regulatory network of ABI3 transcription factors has been established in research on Arabidopsis (Reidt al., 2000; Ruuska., 2002; Mönke et al., 2004). The results of this study can preliminarily prove the relationship between ABI3-like and SPGs at the mRNA level; however, whether GmABI3-8 and GmABI3-18 transcription factors are related to RY repeating elements, and whether they can regulate the expression of related genes must still be determined using techniques such as yeast hybridization.

**Functional prediction of GmABI3-8 and GmABI3-18**

Research by Sugliani et al. (2010) revealed that ABI3 transcription factors are also involved in the dehydration and dormancy processes of the Arabidopsis seeds, but did not demonstrate this to be true in other plants. In this study, the expression of GmABI3-8 and GmABI3-18 mRNA reached its maximum at 80DAF during desiccation and dormancy, increasing significantly compared to the previous data, while the expression of SPGs did not increase correspondingly. This indicates that the GmABI3-8 and GmABI3-18 transcription factors are mainly involved in the dehydration process, and do not promote the synthesis of storage proteins. However, the specific metabolic pathways of soybean seeds during this period require further study. The second increase in GmABI3-18 mRNA occurred earlier than that of GmABI3-8, and the expression level was more than twice that of the latter, indicating that GmABI3-18 transcription factors very likely play a major role in the processes of the dehydration stage. Interestingly, combining the metabolic patterns of GmABI3-8 during the seed filling period revealed that it was very likely that the relevant genes differ both in expression localization and timing, as well as functionality after soybean polyploidization, which is of great significance to research on the gene's evolution.
CONCLUSION

Over the course of soybean seed development, *GmABI3-8* and *GmABI3-18* play an important role in the accumulation of storage protein and the process of dehydration. These specific biological pathways still require further study.

ACKNOWLEDGEMENTS

This study was financially supported by the National 948 Project of China (2012-Z35).

Author contribution

Study concepts and design: R. D.; Material preparation: R. D. Y. Q. X. L.; Literature research: R. D. X. W.; Date analysis/interpretation, Manuscript Preparation and definition of intellectual content: R. D. Y. Q.; Manuscript editing R. D. J.W.; Manuscript revision/review J. W.; Manuscript final version approval R. D. Y. Q. X. W. X. L. J. W.; All authors read and approved the manuscript.

REFERENCES


Du, et al.: mRNA expression patterns of storage protein genes

