Analysis of Genes Expressed during the Early Maturation of Sesame Seeds

Motonobu Yoshida, Akihiro Kishikawa, Takayuki Nakano, Eiji Tanesaka

Lab of Plant Breeding and Agrobiotechnology, Department of Agricultural Science
Kinki University
Nakamachi 3327-204, Nara 631-8505, Japan
Tel: 81-742 43 5245, Fax: 81-742 43 1155
yoshida_m@nara.kindai.ac.jp

Abstract

Sesame seeds of cultivated species, *Sesamum indicum,* contain abundant oil, in particular, large amounts of unsaturated fatty acid oleic and linoleic acids. Total amounts of oleic and linoleic acids in sesame seeds keep constant among varieties of *S. indicum*. However, the regulation of the mechanism is unknown. Therefore, this study attempted to clarify these points through transformation of *Arabidopsis thaliana* with novel genes in sesame seeds. Oil contents increase rapidly in sesame seeds within 4 weeks after flowering. A full-length cDNA library prepared from sesame seed of 1 to 3 weeks oldwas substracted with cDNAs from plantlets of 4 weeks after germination. The results showed that the cDNA library was expressed specifically during the early maturation.The 1,545 cDNA clones were sequenced. Of these, 30.3% of the clones were responsible for protein destination, 16.7% of the clones for the metabolism, and 6.9% of the clones for the cell growth and division. Among them, 13 genes for a transcription factor were identified, four were identified as a transcription factor involved in ethylene signaling. In addition, nine genes: osmotin-like protein, expansin-like protein, aquaporin-like protein, MADS-box protein 4, novel putative uncharacterized protein, ethylene response factor 2, ethylene responsive element binding protein 3, AP2 containing transcription factor and lipid transfer protein were analyzed by overexpression of *A. thaliana.* The *A. thaliana* overexpression strain for novel putative uncharacterized protein and aquaporin-like protein, respectively showed the increase of unsaturated fatty acids. The amount of oleic acid per seed grain in a strain for novel protein was 1.7 fold, and that of linoleic acid 1.5 fold, respectively.

Keywords: Sesamum indicum*, fatty acids, full-length cDNA, Arabidopsis transformation*

Introduction

Sesame seeds of cultivated species, *Sesamum indicum,* contain abundant oil, in particular, large amounts of unsaturated fatty acid oleic and linoleic acids. These unsaturated fatty acids lower the cholesterol level in the body (Satchithanandam *et al*., 1993). Because of their effects and antioxidant substances, attention has been paid to sesame seeds as health food (Budowski and Markley, 1951). Total amounts of oleic and linoleic acids in sesame seeds are constant among varieties of *S. indicum*. However, the regulation of the mechanism is unknown. It is necessary to clarify these points to breed novel varieties with high content of unsaturated fatty acids. Oil contents of sesame seeds increase rapidly within 4 weeks day after flowering (DAF). Therefore, this study attempted to analyze the function of novel genes expressed specifically during the early maturation of sesame seeds through transformation of *Arabidopsis thaliana* with novel genes of sesame seeds.

Materials and Methods

Plant materials and growth conditions

Sesamum indicum seeds were obtained from Toyama University. *Arabidopsis thaliana* plants were grown under continuous light at 22-23^{o}C and used for transformation.

Preparation of a full-lengh cDNA library and cDNA materials

Total RNAs were prepared from *S. indicum* seeds of 1 to 3 weeks old. A full-length cDNA library was prepared using pCMVFL3 vector as described previously (Uenishi *et al*., 2004). Total RNAs from young plants were extracted four weeks after germination. After purification of poly $(A)^+$RNA, cDNA materials were synthesized with cDNA synthesis kit M-MLV version (TaKaRa). A full-lengh cDNA library expressed specifically during the early maturation was prepared by subtraction of the following methods. A full-lengh cDNA library from *S. indicum* seeds of 1 to 3 weeks old was amplified with SP6 primer, T7 primer and the PCR programme which was 16 cycles of 98^{o}C for 10 sec, 56^{o}C for 5 sec and 72^{o}C for 3 min, followed by 72^{o}C for 3 min. They were used as a tester cDNA in subtraction. The subtraction was carried out with DsDD cDNA subtraction kit (Wako) using amplified cDNAs from a full-length cDNA library as a tester cDNA and cDNA materials from young plants of four weeks after germination as a driver cDNA. Subtracted cDNA clones were amplified with the PCR programme which was 35 cycles of 98^{o}C for 10 sec, 56^{o}C for 5 sec and 72^{o}C for 3 min, followed by 72^{o}C for 7 min. After removing cDNA clones of the lower molecular weight with Chroma spin^{TM}-200 (Clontech), they were selected in size with Chroma spin^{TM}-400, and pooled into four groups (fraction No. 4, fractions No. 5-7, fractions No. 8-14, fractions No. 15-30). Each pooled cDNA clone was amplified with *EX-taq* polymerase (TaKaRa), SP6 primer, T7 primer and the PCR programme which was 25 cycles of 98^{o}C for 20 sec, 50^{o}C for 30 sec and 72^{o}C for 3 min, followed by 72^{o}C for 5 min. They were ligated into pCR8/GW/TOPO vector (Invitrogen).

Sequence of cDNA clone

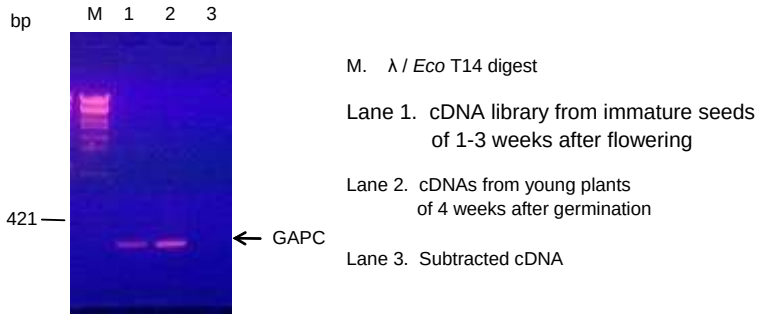
The cDNA clones from three groups (fraction No. 1-4, fraction No. 5-7, fraction No. 8-14) were sequenced with ABI Prism® 3100-Avant Genetic Analyzer (Applied Biosystems) using BigDye® terminator v3.1 cycle sequencing kit (Applied Biosystems). Homology analysis was carried out with BLASTX (DDBJ; http://www.ddbj.nig.ac.jp/welcome-j. html).

***Transformation of* Arabidopsis thaliana**

The cDNA clones ligated into pCR8/GW/TOPO vector were transferred into pBI-OX-GW (Inplanta Innovations Inc.) by using Gateway System (Invitrogen). Each cDNA clone transferred into pBI-OX-GW was used for the transformation of *Arabidopsis thaliana* which was done as described in a previous paper (Clough and Bent, 1998), with minor modification.

Results and Discussion

To obtain cDNA clones expressed specifically during the early maturation of *S. indicum* seeds, the subtraction was carried out using amplified cDNAs from a full-length cDNA library from *S. indicum* seeds of 1 to 3 weeks old as a tester cDNA and cDNA materials from young plants of four weeks after germination as a driver cDNA. The subtraction efficiency was confirmed with the PCR programme which was 22 cycles of 98^{o}C for 20 sec, 50^{o}C for 30 sec and 72^{o}C for 1 min, followed by 72^{o}C for 3 min. The primers were GAPC-F primer (5'-CCAACGCTAGCTGCACCAC-3') and GAPC-R primer (5'-AGGTCAACAACTGAGACATC-3') of the gene for cytosolic glyceraldehyde-3- phosphate dehydrogenase (gapC) which is one of house keeping genes (Figure 1). Signals were observed in both a full-length cDNA library from *S. indicum* seeds of 1 to 3 weeks old and cDNA materials from young plants of four weeks after germination, while no signals were observed in the subtracted cDNA library, suggesting that subtraction was efficiently performed.



GAPC : Cytosolic glyceraldehyde-3- phosphate dehydrogenase

Figure 1. Confirmation of subtraction efficiency.

The subtracted cDNA library was amplified and separated in size into four groups (fraction No. 4, fractions No. 5-7, fractions No. 8-14, fractions No. 15-30). Fraction No. 15-30 mainly contains cDNA clones of less than 400 bp. Therefore, 1,545 cDNA clones were sequenced from fractions No. 4-14. Properties of cDNA clones in fraction No. 4, fractions No. 5-7 and fractions No. 8-14 were summarized as the following; the average size 1,043.7 bp in fraction 4; 996.7 bp in fractions 5-7; 706.4 bp in fractions 8-14. The 1,545 cDNA clones were categorized and summarized with classification of each function (Table 1). The cDNA clones responsible for protein destination formed the largest percentage of 30.52%. Next, the cDNA clones responsible for metabolism formed the second largest percentage of 16.36%, the cDNA clones for cell growth, division were 7.01%, the cDNA clones for transcription and RNA processing were 4.48%. Unknown and novel cDNA clones were 30.39%. Comparison of our results with results by Suh *et al*. (2003) showed that subtraction was effectively performed from the fact that the occupation percentage of house keeping genes was low, and that of unknown novel genes was high at the present experiment. Of them, functional analysis of nine cDNA clones: osmotin-like protein, expansin-like protein, aquaporin-like protein, MADS-box protein 4, novel putative uncharacterized protein, ethylene response factor 2, ethylene responsive element binding protein 3, AP2 containing transcription factor and lipid transfer protein was carried out by overexpression of *A. thaliana.* Southern blot analysis in transformed *A. thaliana* using each DIG-labeled sesame gene as a probe showed that each sesame gene was indeed introduced into transformed *A. thaliana* (data not shown). The *A. thaliana* overexpression strain for novel putative uncharacterized protein and aquaporin-like protein showed the increase of unsaturated fatty acids in *A. thaliana* seeds. The amount of oleic acid per seed grain in the overexpression strain for novel putative uncharacterized protein was 1.7 fold, and that of linoleic acid 1.5 fold. In a strain for aquaporin-like protein, the amount of oleic acid was 1.4 fold, and that of linoleic acid 1.5 fold.

Tabel 1. Functional category of identified genes

Function	%
Protein destination	30.52
Metabolism	16.36
Cell growth, division	7.01
Transcription and RNA processing	4.48
Cell resue, defense, senescence and death	4.16
Transport	2.27
Signal transduction	2.14
Energy generation	1.10
Protein synthesis	1.04
Intracellular trafficking	0.32
Cellular organization/biogenesis	0.19
unknown	30.39

References

Budowski P. and K.S. Markley. 1951. The chemical and physiological properties of sesame oil. *Chem. Rev.* 48: 125-151.

Clough S.J. and A.F. Bent. 1998. Floral dip: a simplified method for *Agrobacterium*- mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16: 735-743.

Suh M.C., M.J. Kim, C-G Hur, J.M. Bae, Y.I. Park, C-H. Chung, C-W. Kang and J.B.
Ohlrogge. 2003. Comparative analysis of expressed sequence tags from
Sesamum indicum and *Arabidopsis thaliana* developing seeds. *Plant Mol. Biol.* 52: 1107-1123.

Satchithanandam S., M. Reicks, R.J. Calvert, M.M. Cassidy and D. Kritchevsky. 1993. Coconut oil and sesame oil affect lymphatic absorption of cholesterol and fatty acids in rats. *J. Nutr.* 123: 1852-1858.

Uenishi H., T. Eguchi, K. Suzuki, T. Sawazaki, D. Toki, H. Sinkai, N. Okumura, N. Hamashima and T. Awata. 2004. PEDE (Pig EST Data Explorer): construction of a database for ESTs derived from porcine full-length cDNA libraries. *Nucleic Acids Res*. 32: 484-488.

-- back to Table of Content --