Chapter 5
Isolation of Differentially Expressed Genes from Developing Seeds of a High-Protein Peanut Mutant and Its Wild Type Using Genefishing^{TM} Technology

Shu Tao Yu, Hong Bo Yu, Guo Qing Yu, Li Ren Zhao, Hong Xi Sun, Yue Yi Tang, Xiu Zhen Wang, Qi Wu, Quan Xi Sun and Chuan Tang Wang

Abstract Peanut is a good source of dietary protein. Raising protein content in peanut will not only fill the growing need for vegetable protein, but also in most cases lower oil content, which is good news to health-conscious populations. However, no attempt has been made to isolate genes related to protein content in peanut. In the present study, a total of 40 unique differentially expressed genes in developing seeds of high-protein peanut mutant (SDPM) and its normal-protein wild type (SDPW) at 46 or 49 days after flowering were isolated using Genefishing technology. Of them, 8 sequences were undescribed previously; the rest 32 were found to be significantly similar to the sequences in GenBank nr database. Three genes potentially related to protein content in peanut, viz., P2-2-2, P2-92-2 and P1-89-1-5 with high homology to thioredoxin h, arachin ahy-4 and abc transporter, respectively, were selected for further analysis. All the 3 genes validated by qRT-PCR showed differential expression between SDPM and SDPW, with relative expression ranging from 0.41–10.60. The detailed functions of the differentially expressed genes isolated from developing seeds in the present study in conditioning peanut seed protein content are yet to be validated by transgenic experiments.

Keywords Differentially expressed gene · Genefishing technology · Mutant · Peanut · Protein content

S.T. Yu (✉) · H.B. Yu · G.Q. Yu · L.R. Zhao · H.X. Sun
Sandy Land Amelioration and Utilization Research Institute of Liaoning, 123000 Fuxin, China
e-mail: shutaoyu@163.com

Y.Y. Tang · X.Z. Wang · Q. Wu · Q.X. Sun · C.T. Wang (✉)
Shandong Peanut Research Institute, Qingdao 266100, China
e-mail: chinapeanut@126.com

© Springer-Verlag Berlin Heidelberg 2015

T.-C. Zhang and M. Nakajima (eds.), *Advances in Applied Biotechnology*, Lecture Notes in Electrical Engineering 332, DOI 10.1007/978-3-662-45657-6_5

5.1 Introduction

The cultivated peanut, *Arachis hypogaea* L., as a major cash crop valued worldwide, is cultivated in more than eighty countries/regions. In western world, peanut is a main dietary protein source. In developing nations, even though most peanut produce is crushed for edible oil, the portion for food uses has been on a steady increase. Raising protein content in peanut seeds will not only fill the growing need for vegetable protein, but also in most cases lower oil content, which is good news to health-conscious populations. To lower oil content, defatting is generally useful; however, it cannot be applied to peanut for in-shell consumption.

In literature, Spanish- or Valencia-type peanut landraces/cultivars with high-protein content are not uncommon, and there are reports on heritability, heterosis and combinability for peanut protein content [1]. Recently, [2] analyzed the data from 2 environments, and concluded that protein content in peanut is conditioned by polygenes. Quantitative trait loci (QTLs) with lower than 15 % contribution to phenotypic variations have been reported from two independent research groups [3, 4]. But to the best of our knowledge, no attempt has been made to isolate genes related to this valuable trait.

The present study represented an effort aimed at isolation of candidate genes governing protein content in peanut through Genefishing technology using a large-podded peanut genotype and its high-protein chemical mutant.

5.2 Materials and Methods

5.2.1 Peanut Materials

SDPM, a peanut EMS (ethyl methane sulfonate) mutant with 28.67 % protein, developed at Shandong Peanut Research Institute (SPRI), Qingdao, China, and its wild type counterpart with 17.69 % protein (SDPW) were used in the present study. The high-protein mutant was first identified in near infra-red reflectance spectroscopy (NIRS) analysis of intact seeds [5] and further confirmed by nitrogen amount determination using an automated Kjeldahl Analyzer (model 2300 II, Foss, Sweden). A conversion factor of 5.46 was used to convert the amount of nitrogen to amount of protein [6].

5.2.2 Peanut Cultivation and Seed Sampling

Peanut seeds were sown in field under polythene film mulch and routine agronomic practices were followed as per the description by Wan et al. [7]. Plant population was 150,000 hills per ha, with two plants per hill. Flowers on the first and second

node of the cotyledonary branches were tagged and pods were harvested 46 and 49 days after flowering (DAF). Seeds were then stored in liquid nitrogen.

5.2.3 RNA Isolation

Total RNA was extracted from developing seeds of SDPM and SDPW using RNAprep pure Plant Kit (Tiangen, Beijing, China) following manufacturer's instructions. RNA concentration and integrity were determined by spectrophotometry and relative intensity of brightness of GelRed (Biotium, CA, USA) stained bands resolved on a 1.2 % agarose gel [8].

5.2.4 Cloning and Sequence Analysis of DEGs

Differentially expressed genes (DEGs) from developing seeds of SDPM or SDPW harvested at 46 DAF and 49 DAF were identified using Genefishing™ DEG Premix Kit (Seegene, Korea). RNase-free water was added to the mixture of 3 μg of total RNA and 2 μl of 10 μM dT-ACP1 to a total volume of 9.5 μl. The mixture was incubated at 72 °C for 10 min., cooled on ice for 2 min., centrifuged briefly, and then 4 μl of 5 × RT buffer, 5 μl of 2 mM dNTP mix, and 0.5 μl of RNase inhibitor (40 U/μ1) (Tiangen, Beijing, China) along with 200 U of M-MLV reverse transcriptase (TaKaRa, Japan) were added. Reverse transcription was conducted at 42 °C for 90 min, followed by incubation at 70 °C for 10 min to terminate the reaction. First strand cDNA products were then diluted with 80 μl of DNase-free water and directly used in subsequent isolation and analysis of differentially expressed genes from peanut using Genefishing PCR. PCR mixture (20 μl) contained 50 ng of first strand cDNA, 0.5 μM arbitrary ACP, 0.5 μM dT-ACP2 and 2 × SeeAmp ACP Master-mix. PCR program was 94 °C for 5 min, 50 °C for 3 min and 72 °C for 1 min, followed by 40 cycles of 94 °C for 40 s, 65 °C for 40 s and 72 °C for 40 s, and a final extension of 72 °C for 5 min. PCR products were separated on a 2 % agarose gel, stained with Gelred and visualized under UV light. Amplicons of interest from treated samples were cloned into a pGM-T vector (Tiangen, Beijing, China), and sequenced by Genscript Inc., Nanjing, China. Resultant sequences, after trimmed to remove poor quality reads and vector sequences, were assembled with DNAStar (DNASTAR Inc., London, UK) package. Transcript annotation and functional assignment were carried out with BLAST2GO (http://blast2go.org).

Table 5.1 List of realtime PCR primers

Primer ID	Forward primer(5′–3′)	Reverse primer(5′–3′)
p-P2-2-2	ACTCATAGAATGGCAGAGG	TGTTTCGTCTTTCGTTTCC
p-P2-92-2	GGCATTCAAGACAGACTCA	CGGTGGAACGAAGAACTT
p-P1-89-1-5	AAGTTGCGGATAGGAAGAT	GGATGGAGACGAAGAAGATA
β-actin	TTGGAATGGGTCAGAAGGATGC	AGTGGTGCCTCAGTAAGAAGC

5.2.5 qRT-PCR

Realtime PCR primers were designed using Beacon Designer 7 (Premier Biosoft International, CA, USA) (Table 5.1). qRT-PCR was performed in a Lightcycler 2.0 PCR machine (Roche, USA) as per Tang et al. (2011), with 3 replications for each reaction. β-actin gene was utilized as an internal control. Fold changes in RNA transcripts were calculated using the $2^{-\Delta\Delta Ct}$ method.

5.3 Results and Discussion

Of 36 primers (ACPs) provided with the Genefishing kit, 14 produced bands that could clearly differentiate SDPM and SDPW (Fig. 5.1). Cloning, sequencing and assembly of the DEGs resulted in a total of 40 unique sequences (Table 5.2). There were 8 sequences undescribed previously; the rest 32 DEGs were found to be significantly similar to the sequences in GenBank nr database, of which 12 and 20 DEGs were from SDPW and SDPM, respectively.

Three of the unique DEGs potentially related to protein content in peanut, viz., P2-2-2, P2-92-2 and P1-89-1-5 with high homology to *thioredoxin h*, *arachin ahy-4* and *abc transporter*, respectively, were selected for further analysis. The annotation results were shown in Table 5.3. Relative expression of the 3 genes between high-protein EMS mutant (SDPM) and normal-protein wild type (SDPW) in developing

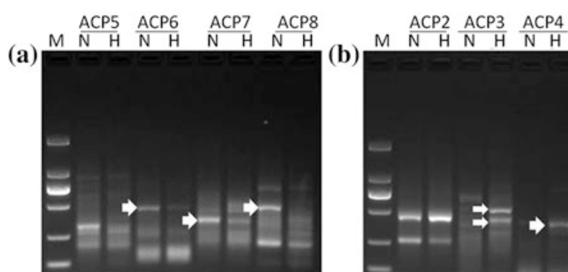

Fig. 5.1 Differentially expressed genes between SDPW (N) and SDPM (H) as indicated by *arrows* in developing seeds harvested at 46 DAF (**a**) or 49 DAF (**b**)

Table 5.2 BLAST* results of 40 unique DEGs

Seq. name	ACP primer	Seq. description	Seq. length (bp)	Min. eValue	Mean similarity (%)
C1-2-1-2	ACP2	Unnamed protein product [Vitis vinifera]	395	1.25E-20	68.50
C1-2-2-4	ACP2	High mobility group family	779	1.23E-52	85.80
C1-22-2	ACP22	—NA—	247		
C1-28-1	ACP28	—NA—	426		
C1-3-1-2	ACP3	Type 2 metallothionein	527	1.60E-39	87.55
C1-3-1-4	ACP3	Type 2 metallothionein	519	8.05E-39	87.60
C1-3-2-1	ACP2	Arachin ahy-4	427	1.24E-16	90.20
C1-3-2-5	ACP3	—NA—	477		
C1-8-1	ACP8	Sterile alpha motif domain-containing protein	488	7.38E-36	81.20
C1-83-3	ACP83	Phage tail tape measure lambda family	752	2.17E-125	98.75
C1-83-5	ACP83	Cyclophilin	379	4.16E-26	97.40
C1-85-1	ACP85	*Arachis hypogaea* seed storage protein SSP1 mRNA, partial cds	511	1.00E-141	98
C1-92-1	ACP92	*Arachis hypogaea* iso-Ara h3 mRNA, complete cds	719	0	96
C1-92-2	ACP92	Arachin ahy-4	556	9.35E-49	91.5
C2-28-1	ACP28	—NA—	375		
C2-92-4	ACP92	Arachin ahy-4	461	8.06E-39	91.95
P2-2-2	ACP2	Thioredoxin h	348	5.91E-12	87.35
P1-22-5	ACP22	DNA-directed rna polymerases and iii subunit rpabc2-like	487	1.83E-39	97.70
P1-3-2	ACP3	60 s ribosomal protein	625	6.14E-70	95.50
P1-36-1-1	ACP36	Endothelial differentiation-Related factor 1	345	2.71E-25	93.30
P1-36-2-3	ACP36	—NA—	130		
P1-36-3-1	ACP36	—NA—	130		
P1-7-1	ACP7	Gdp-mannose-dehydratase	454	1.14E-54	96.15
P1-7-2	ACP7	Sterile alpha motif domain-containing protein	477	6.63E-36	81.20
P1-7-3	ACP7	Germin-like protein	477	2.57E-71	84.50
P1-81-1	ACP81	Uncharacterized protein loc100804883	419	9.45E-09	73.50
P1-81-3	ACP81	—NA—	374		
P1-84-2	ACP84	Aminotransferase-like protein	382	2.63E-21	65.65
P1-84-3	ACP27	Protein	532	4.94E-66	93.45
P1-84-4	ACP28	Photosystem II reaction center protein m	330	1.80E-10	75.00
P1-84-5	ACP29	Aminotransferase-like protein	385	1.49E-21	65.20
p1-89-1-1	ACP89	Stf2-like protein	776	3.50E-43	78.60
P1-89-1-5	ACP89	Abc transporter	952	1.59E-149	94.95

(continued)

Table 5.2 (continued)

Seq. name	ACP primer	Seq. description	Seq. length (bp)	Min. eValue	Mean similarity (%)
P1-89-2-5	ACP89	—NA—	137		
P1-92-1	ACP92	Arachin ahy-4	555	8.78E-49	91.50
P2-3-3	ACP3	Type 2 metallothionein	471	1.35E-38	86.45
P2-3-5	ACP3	Kda class i heat shock protein	366	2.27E-16	97.20
p2-90-3	ACP90	Rar1 protein	382	1.38E-48	95.20
P2-92-2	ACP92	Arachin ahy-4	556	9.35E-49	91.5
p2-84-3	ACP84	*Arachis hypogaea* oleosin 18.5 gene, promoter region and complete cds	203	3.00E-49	96

* BLAST = basic local alignment search tool

Table 5.3 Annotation of three DEGs using BLAST2GO

Seq. name	Seq. description	#Hits	Min. e value	Mean similarity (%)	#GOs	GOs
P2-2-2	Thioredoxin h	20	6.07E-12	87.35	7	P: cell redox homeostasis; F: electron carrier activity; P: glycerol ether metabolic process; F: protein disulfide oxidoreductase activity; C: cytoplasm; P: electron transport chain; P: transport
P2-92-2	Arachin ahy-4	20	4.77E-49	90.95	1	F: nutrient reservoir activity
P1-89-1-5	Abc transporter	20	1.70E-149	94.95	10	F: ATPase activity; F: ATP binding; F: hydrolase activity; F: nucleotide binding; F: nucleoside-triphosphatase activity; F: phosphonate transmembrane-transporting ATPase activity; P: ATP catabolic process; F: transporter activity; C:mitochondrion; C: plastid

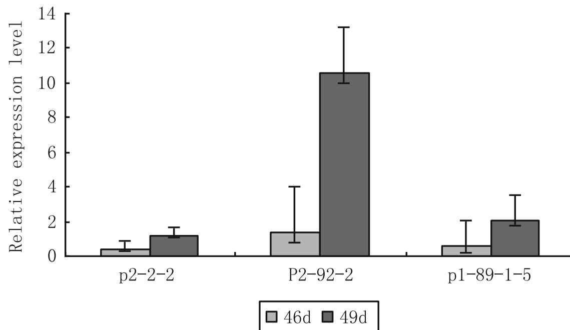

Fig. 5.2 Relative expression of 3 genes between the high-protein EMS mutant (SDPM) and the normal-protein wild type (SDPW) in developing seeds harvested at 46DAF(*46d*) or 49DAF(*49d*), respectively. Relative expression of SDPM was computed based on the corresponding gene expression of SDPW. The error bars indicate standard deviation of mean

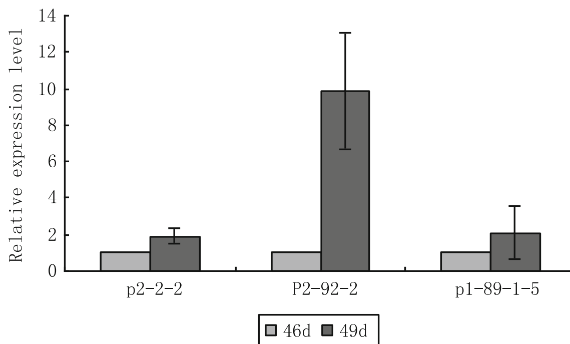

Fig. 5.3 Relative expression of the 3 genes in developing peanut seeds of SDPM harvested at 46 DAF (*46d*) and 49DAF. Error bar indicating standard deviation of mean. Relative expression at 49 DAF (*49d*) was computed based on the corresponding gene expression at 46 DAF

seeds harvested at 46 DAF or 49 DAF was illustrated in Fig. 5.2. Notably, expression of P2-92-2 (*arachin ahy-4*) at 49 DAF in SDPM was 10.60 times as high as in SDPW. Expression of the same 3 genes in seeds of SDPM harvested at 49 DAF relative to the expression at 46 DAF was shown in Fig. 5.3. P2-92-2 exhibited a marked increase in expression at 49 DAF as compared with that at 46 DAF.

Through NIRS-aided selection of mutagenized populations, SPRI scientists were able to identify peanut quality mutants, providing materials for the present study [5]. In contrast to randomly selected peanut materials with different protein content, mutant and wild types have similar genetic backgrounds, thereby precluding a large number of genes unrelated to the target trait when transcriptional profiling strategy is used.

Thus far, there have been few studies on genes related to protein content in plants [9]. Two reports have shown that *PEPC* (encoding phosphor-enolpyruvate carboxylase) and *VfAAP* (encoding *Vicia faba* amino acid permease) were genes conditioning protein content in legume seeds. As compared with untransformed control, transgenic bean seeds with the overexpression construct of *PEPC* accumulated up to 20 % more protein per gram seed dry weight [10]. Overexpression of the transporter gene *VfAAP* in pea resulted in 43 % increase in total globulins production in seeds [11].

In the present study, through comparison of gene transcription in developing seeds of a high-protein mutant and its normal-protein wild type peanut genotype at the stage of rapid seed protein accumulation (46 DAF and 49 DAF), totally 40 DEGs were isolated using Genefishing^{TM} technology. All the 3 DEGs further validated by qRT-PCR showed differential expression between the high-protein mutant and the peanut wild type, with relative expression ranging from 0.41–10.60. As indicated by BLAST2GO analysis, the 3 genes may have protein disulfide oxidoreductase activity, nutrient reservoir activity or ATPase/transporter activity, respectively. Further studies are still needed to validate the differential expression of the rest genes by qRT-PCR and to investigate the detailed and exact functions of confirmed DEGs through transgenic experiments.

Acknowledgments We wish to express our sincere thanks to the financial support from China Agricultural Research System (CARS-14), Liaoning Natural Science Foundation Project for Fostering Talent (2014027029), Key Project of Liaoning Provincial Science and Technology Department (2011201021) and China National Science and Technology Support Project (2012BAD36B00).

References

1. Yu SL, Wang CT, Yang QL et al (2011) Peanut genetics and breeding in China. Shanghai Science and Technology Press, Shanghai, pp 18–24
2. Zhang XY, Han SY, Xu J et al (2011) Genetic analysis of protein using major gene plus polygene methods in peanut (*Arachis hypogaea* L.). Chinese J Oil Crop Sci 33(2):118–122
3. Liang X, Zhou G, Hong X et al (2009) Overview of research progress on peanut (*Arachis hypogaea* L.) host resistance to aflatoxin contamination and genomics at the guangdong academy of agricultural sciences. Peanut Sci. 36:29–34
4. Sarvamangala C, Gowda MVC, Varshney RK (2011) Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (*Arachis hypogaea* L.). Field Crops Res 122:49–59
5. Wang CT, Zhang JC, Tang YY et al (2013) Peanut genetic improvement. Shanghai Science and Technology Press, Shanghai, p 531
6. Wang CT, Tang YY, Wang XZ et al (2011) Evaluation of groundnut genotypes from China for nutritional quality. J SAT Agric. Res. vol. 9. Available at: http://ejournal.icrisat.org/Volume9/Groundnut/Evaluation.pdf. (Accessed 31 May 2013)
7. Wan SB (ed) (2003) China peanut cultivation science. Shanghai Science & Technology Press, Shanghai pp 408–430

8. Tang YY, Wang CT, Yang GP et al (2011) Identification of chilling-responsive transcripts in peanut (*Arachis hypogaea L.*). Electronic J. Biotechnol. 14(5). DOI: 10.2225/vol14-issue5-fulltext-5
9. Weber H, Sreenivasulu N, Weschke W (2010) Molecular physiology of seed maturation and seed storage protein biosynthesis. In: Pua EC, Davey MR (eds) Plant developmental biology–biotechnological perspectives, Springer, Berlin Heidelberg, pp 83–104
10. Rolletschek H, Borisjuk L, Radchuk R et al (2004) Seed-specific expression of a bacterial phosphoenolpyruvate carboxylase in Vicia narbonensis increases protein content and improves carbon economy. Plant Biotechnol J 2(3):211–219
11. Rolletschek H, Hosein F, Miranda M et al (2005) Ectopic expression of an amino acid transporter (VfAAP1) in seeds of Vicia narbonensis and pea increases storage proteins. Plant Physiol 137(4):1236–1249