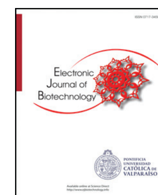
Electronic Journal of Biotechnology 32 (2018) 19–25

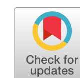
Contents lists available at ScienceDirect

Electronic Journal of Biotechnology



Research article

# Transcriptome profiling and digital gene expression analysis of genes associated with salinity resistance in peanut



Jiongming Sui [a,1], Pingping Jiang [b,1], Guilong Qin [b,1], Shupeng Gai [b], Dan Zhu [b], Lixian Qiao [a], Jingshan Wang [a,*]

[a] College of Agriculture, Qingdao Agricultural University/Dry-land Farming Technology Lab, Qingdao, 266109, China
[b] College of Life Science, Qingdao Agricultural University, Key Laboratory of Qingdao Major Crop Germplasm Resource Innovation and Application, Key Lab of Plant Biotechnology in Universities of Shandong, Qingdao 266109, China

ARTICLE INFO

*Article history:*
Received 22 March 2017
Accepted 5 December 2017
Available online 10 December 2017

*Keywords:*
Digital gene expression
Gene
Mutant
NaCl
Peanut (*Arachis hypogaea* L.)
RNA-seq
Salinity stress
Salinity tolerance
Soil salinity
Transcripts
Unigenes

ABSTRACT

*Background:* Soil salinity can significantly reduce crop production, but the molecular mechanism of salinity tolerance in peanut is poorly understood. A mutant (S1) with higher salinity resistance than its mutagenic parent HY22 (S3) was obtained. Transcriptome sequencing and digital gene expression (DGE) analysis were performed with leaves of S1 and S3 before and after plants were irrigated with 250 mM NaCl.
*Results:* A total of 107,725 comprehensive transcripts were assembled into 67,738 unigenes using TIGR Gene Indices clustering tools (TGICL). All unigenes were searched against the euKaryotic Ortholog Groups (KOG), gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases, and these unigenes were assigned to 26 functional KOG categories, 56 GO terms, 32 KEGG groups, respectively. In total 112 differentially expressed genes (DEGs) between S1 and S3 after salinity stress were screened, among them, 86 were responsive to salinity stress in S1 and/or S3. These 86 DEGs included genes that encoded the following kinds of proteins that are known to be involved in resistance to salinity stress: late embryogenesis abundant proteins (LEAs), major intrinsic proteins (MIPs) or aquaporins, metallothioneins (MTs), lipid transfer protein (LTP), calcineurin B-like protein-interacting protein kinases (CIPKs), 9-cis-epoxycarotenoid dioxygenase (NCED) and oleosins, etc. Of these 86 DEGs, 18 could not be matched with known proteins.
*Conclusion:* The results from this study will be useful for further research on the mechanism of salinity resistance and will provide a useful gene resource for the variety breeding of salinity resistance in peanut.

© 2017 Pontificia Universidad Católica de Valparaíso. Production and hosting by Elsevier B.V. All rights reserved. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

## 1. Introduction

Over the past few decades, there has been a dramatic increase in the salinization of arable land. As the land available for conventional agriculture becomes increasingly limited, plants grown on marginal soils will be exposed to higher levels of soil salinity. Soil salinity is a major abiotic stress responsible for reduced growth and yield of many crops [1]. Consequently, a better understanding of salt tolerance in crops is needed.

Peanut (*Arachis hypogaea* L.), which is an important oil-crop and protein production in the tropics and subtropics [2], is likely to face increased drought and salinity stresses in the near future [3]. Hence, genes responsible for resistance to drought and salinity stress in peanut need to be identified and studied. Unfortunately, little progress has been made in the study of salinity tolerance in peanut, in part because of the lack of peanut germplasm with high resistance to salinity stress. In our previous studies, we conducted in vitro mutagenesis (with pingyangmycin as the mutagen) and directed screening with a medium containing NaCl to generate mutants with salt tolerance [4]. One mutant (designated S1) with enhanced salinity tolerance was obtained. This mutant had a much higher germination rate than its mutagenic parent Huayu 22 (designated S3) in a 0.7% NaCl solution, and its self-pollinated offspring grew better than S3 in a saline-alkali field in Dongying City, China. Little is known, however, about the molecular mechanisms resulting in salt tolerance in peanut.

High-throughput RNA-sequencing (RNA-Seq) is a recent and effective technology for the analysis of gene expression, the discovery of novel transcripts, and the identification of differentially expressed genes (DEGs). This powerful technology makes it possible to study non-model organisms [5,6,7].

To investigate the molecular basis for the salinity-tolerance in peanut, we compared the transcriptome and digital gene expression (DGE) profiles in the leaves of S1 and its salinity-sensitive parent, S3,

* Corresponding author.
*E-mail address:* jswang319@126.com (J. Wang).
[1] These authors contribute equally to this study.
Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

https://doi.org/10.1016/j.ejbt.2017.12.002
0717-3458/© 2017 Pontificia Universidad Católica de Valparaíso. Production and hosting by Elsevier B.V. All rights reserved. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

before and after the application of a salinity-stress treatment. We identified the specific transcripts related to salinity-stress resistance in peanut, and we discussed the possible roles of the DEGs.

## 2. Materials and methods

### 2.1. Plant growth and stress treatments

The seeds of S1 (the mutant with enhanced salinity tolerance) and S3 (Huayu 22, the control) were grown in a growth chamber with a dark/light cycle of 8/16 h at 28°C for six weeks. Then, the seedlings of each genotype were irrigated with 250 mM NaCl for salinity stress under culture-room conditions according to our previous report [8]. At 0, 6, 12, 24, and 48 h after the seedlings were subjected to the NaCl solution, the leaves of the S1 and S3 seedlings were removed and placed in liquid nitrogen.

### 2.2. Library construction and transcriptome sequencing

A total amount of 1.5-μg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) f and index codes were added to attribute sequences to each sample. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X). First strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H-). Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3′ ends of DNA fragments, NEBNext Adaptor with hairpin loop structure were ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 150–200 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then, 3-μl USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37°C for 15 min followed by 5 min at 95°C before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. At last, The RNA from each combination of seedling type (S1 or S3) and time after salt treatment was pooled and then analyzed with an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). Sequencing of the RNA was carried out by Novogene Bioinformatics Technology Co., Ltd. (Beijing, China) on an Illumina HiSeq2000 sequencer. For no reference genome, after the acquisition of clean reads, the clean reads needs to be spliced to obtain the reference sequence for subsequent analysis. All clean sequence read data were deposited in the NCBI SRA database (accession number SRR3114511), and then they were assembled into comprehensive unigenes using Trinity and TGICL [9].

### 2.3. Transcriptome functional annotation

The assignment of sequence orientations and functional annotations of the all-unigenes were determined by BLASTx against the following databases: the NCBI non-redundant (NR) protein database, the Swiss-Prot protein database with an E-value cut-off of $10^{-5}$, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database with an E-value of $10^{-10}$, and the euKaryotic Ortholog Groups of proteins (KOG) database with an E value of $10^{-3}$. The all-unigenes were assigned to Gene Ontology (GO) categories with an E-value cut-off of $10^{-6}$. In addition, unigenes were aligned with the NCBI nucleotide (NT) databases using BLASTn with an E value of $10^{-5}$.

### 2.4. Digital gene expression (DGE) sequencing and mapping

The RNA samples from S1 and S3 were labeled with the sampling times (0, 6, 12, 24, and 48 h after NaCl treatment) as follows: S1_0, S1_6, S1_12, S1_24, and S1_48 for S1, and S3_0, S3_6, S3_12, S3_24, and S3_48 for S3. Each combination of genotype and sampling time after salt treatment was represented by two replicate RNA samples. DGE sequencing was carried out with a single 50-bp end read for each reaction; all clean sequence read data were deposited in the NCBI SRA database (accession number SRR3204213 and SRR3204348). Then, all reads of each library were separately mapped onto the unigenes using the default parameters in SOAP, and the uniquely mapped reads were extracted for abundance quantification. Finally, unigene expression was normalized using the value of RPKM (reads per kilobase per million reads). Multiple comparisons were carried between the data sets of different samples.

Expression was compared both within each genotype and between the two genotypes. The comparison between S1 and S3 samples resulted in D series data sets, which represented the DEGs between S1 and S3 samples in response to salinity stress treatment; they were denoted as D_0, D_6, D_12, D_24 and D_48 with the sampling time point of 0, 6, 12, 24 and 48 h. Between the genotypes, expression was compared at 0, 6, 12, 24, and 48 h. If the level of expression was significantly different (the adjusted $P$-value < 0.05) in a comparison, the gene was considered to be differentially expressed. Within S1 and S3, expression was compared between each two sampling time of 0, 6, 12, 24, and 48 h; if the level of expression was significantly up- or down-regulated (the adjusted $P$-value < 0.05) in a comparison, this gene was proposed to be responsive to salinity stress. Pathways that were statistically significant (FDR ≤ 0.05) were enriched with KEGG.

### 2.5. Real-time PCR analysis

To determine whether the expression analyses were correct, we performed real-time PCR analysis on selected DEGs. Reverse transcription were performed using an Invitrogen SuperScript Reagent Kit. For real-time PCR, the SYBR® Premix Ex Taq™ (TAKARA) was used on a Bio-Rad CFX96 real-time PCR detection system (Bio-Rad, Hercules, CA). Gene expression was analyzed for S1 and S3 samples at 0, 6, 12, 24, and 48 h after application of the salinity-stress treatment. All reactions for each gene were performed in three biological replications with a 20-μl reaction volume. The relative expression level of each gene among samples was calculated using the $2^{-\triangle\triangle Ct}$ method with normalization to the internal reference *actin* gene from peanut. The parameters of the thermal cycle were 95°C for 30 s, followed by 40 cycles of 95°C for 10 s and 50–56°C for 25 s.

## 3. Results

### 3.1. Transcriptomic sequencing and de novo assembly

The transcriptomic analysis of pooled samples resulted in a total of 62,887,762 clean reads and 7.86 G clean base pairs. The comprehensive reads were assembled into transcripts using paired-end reads, resulting in 107,725 comprehensive transcripts. With the criteria of more than 50-bp overlap and 90% identity, the transcripts were further assembled into 67,738 unigenes using TGICL. The size of unigenes ranged from 201 to 18,360 bp with an average length of 766 bp; the N50 value was 1362 bp (Fig. S1).

### 3.2. Function annotation and classification

Predicted functions of these unigenes were obtained by searching against several protein databases. The number and percentage of the 67,738 unigenes that were annotated in the NR, NT, SwissProt, and PFAM databases are indicated in Table S1.

All unigenes were searched against the euKaryotic Ortholog Groups (KOG) database to divide ortholog clusters by phylogenetical relations. A total of 10,571 (15.61%) of the 67,738 unigenes were assigned to the 26 function categories (Fig. 1, Table S2). The top five categories were “General function prediction only” (1983, 18.76%), “Posttranslational modification, protein turnover, chaperones” (1369, 12.95%), “Signal transduction mechanisms” (938, 8.87%), “Translation, ribosomal structure and biogenesis” (647, 6.12%), and “Intracellular trafficking, secretion, and vesicular transport” (640, 6.05%).

We performed gene ontology (GO) analyses and characterized the sequence annotation of these unigenes. In total, 22,163 (32.71%) of the 67,738 unigenes were assigned GO pathways (Fig. 2, Table S3). These pathways were divided into 56 terms in three categories, and the top five clades were “Binding” (12,999, 58.65%), “Cellular process” (12,612, 56.90%), “Metabolic process” (12,235, 55.20%), “Catalytic activity” (10,728, 48.41%), and “Single-organism process” (9595, 43.29%) (Fig. 2).

Finally, the Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed, and the biological pathways were identified. In total, 9821 (14.49%) of the 67,738 unigenes were assigned KEGG pathways. These pathways were divided into 32 groups in five categories. “Signal transduction” (971, 9.89%) was the largest group, followed by “Carbohydrate metabolism” (905, 9.21%), “Translation” (745, 7.59%), “Amino acid metabolism” (615, 6.26%), and “Folding, sorting and degradation” (607, 6.18%) (Fig. 3, Table S4).

Against the special databases, and an overall view of them was obtained. Among these unigenes, 11,536 (38.5%), 3968 (13.2%), 3951 (13.2%), 3497 (11.7%), and 902 (3.0%) were matched to genes from *Glycine max*, *Phaseolus vulgaris*, *Cicer arietinum*, *Medicago truncatula*, and *Arachis hypogaea*, respectively.

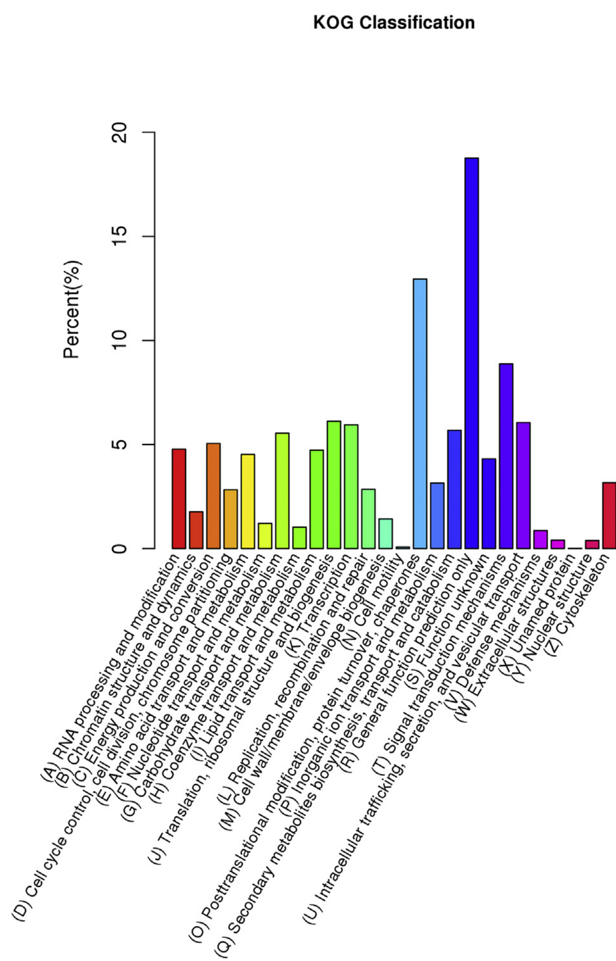



**Fig. 1.** The euKaryotic Ortholog Groups (KOG) annotation of putative proteins. All putative proteins showing significant homology to those in KOG database were functionally classified into 26 molecular families.

### 3.3. DGE library sequencing, mapping, and clustering analysis of samples

To investigate the gene expression patterns before and after the salinity-stress treatment, 20 DGE libraries were constructed and sequenced using Illumina deep sequencing technology. For each sample, from 10, 341, 937 to 14, 276, 045 clean reads and from 0.52 to 0.71 G clean base pairs were generated (Table S5).

The number of unigenes (FPKM > 0.3) ranged from 47,414 to 56,120 in S1 series data sets with an average of 51,914, while the number of unigenes (FPKM > 0.3) ranged from 48,041 to 54,105 in S3 series data sets with an average of 51,754. From 89.26 to 90.73% of the clean reads in each sample were mapped to our transcriptome reference database (Table S5).

To obtain a global view of gene expression in S1 and S3, we analyzed the gene expression profiling for the 10 samples from the two genotypes using clustering algorithms and Treeview. Most samples between S1 and S3 at the same time points showed very similar gene expression patterns (Fig. 4).

### 3.4. Identification of DEGs

After the multiple comparisons, the DEGs were identified under the criteria with the adjusted *P*-value under 0.05. These 10 data sets (one data set for each combination of genotype and time after treatment) represent the DEGs before and after salinity-stress treatment within each genotype. The comparison between S1 and S3 samples resulted in D data sets, which indicated the DEGs between S1 and S3 samples in response to the salinity-stress treatment.

The changes in DEGs with time after treatment in S1 and S3 were investigated. Relative to expression at 0 h, the following numbers of DEGs were responsive to salinity stress at 6, 12, 24, and 48 h in both S1 and S3: 131, 877, 38 and 26 (Fig. 5).

The number of down-regulated DEGs was markedly greater than the number of up-regulated DEGs at 12 and 24 h in S1 and at 12 h in S3 (Fig. 5). There were 9 DEGs that were common at the four times after salinity-stress treatment in S1, while 101 DEGs were common at the four times in S3 (Fig. 6a, b).

For the D series data sets, in which expression in S1 was compared with that in S3, the number of DEGs was greatest at 24 h (Fig. 6c), and the number of down-regulated DEGs was greater than the number of up-regulated DEGs at 12 h but was less than the number of up-regulated DEGs at 24 h (data not shown). There was little overlap in DEGs among the sampling times (Fig. 6c).

### 3.5. Annotation and KEGG enrichment analyses of DEGs based on D series data sets (comparison of S1 vs. S3)

For functional annotation of DEGs, the KEGG enrichment analyses were mainly referenced. For D series data sets (S1vs S3 at the paired time points), only 18 DEGs were detected at 0 h, i.e., before salinity stress was applied (Fig. 6c), indicating that the genetic background is very similar between S1 and S3. Only one and three DEGs were screened after salt stress of 6 and 48 h, but they were not found in KEGG enrichment analysis. Eight differentially expressed DEGs at 12 h were assigned KEGG pathways, which all were down-regulated, according to the corrected *P*-value; the main concerned pathways were shown in Fig. 7a, they were “Photosynthesis-antenna proteins”, “Glycerolipid metabolism”, “Glycerophospholipid metabolism, “Glycosphingolipid biosynthesis-globo series, “Circadian rhythm - plant”. In the comparison of gene expression in S1 vs. S3 at 24 h, 11 down-regulated DEGs and three up-regulated DEGs were assigned KEGG pathways, according to the corrected *P*-value; the main pathways

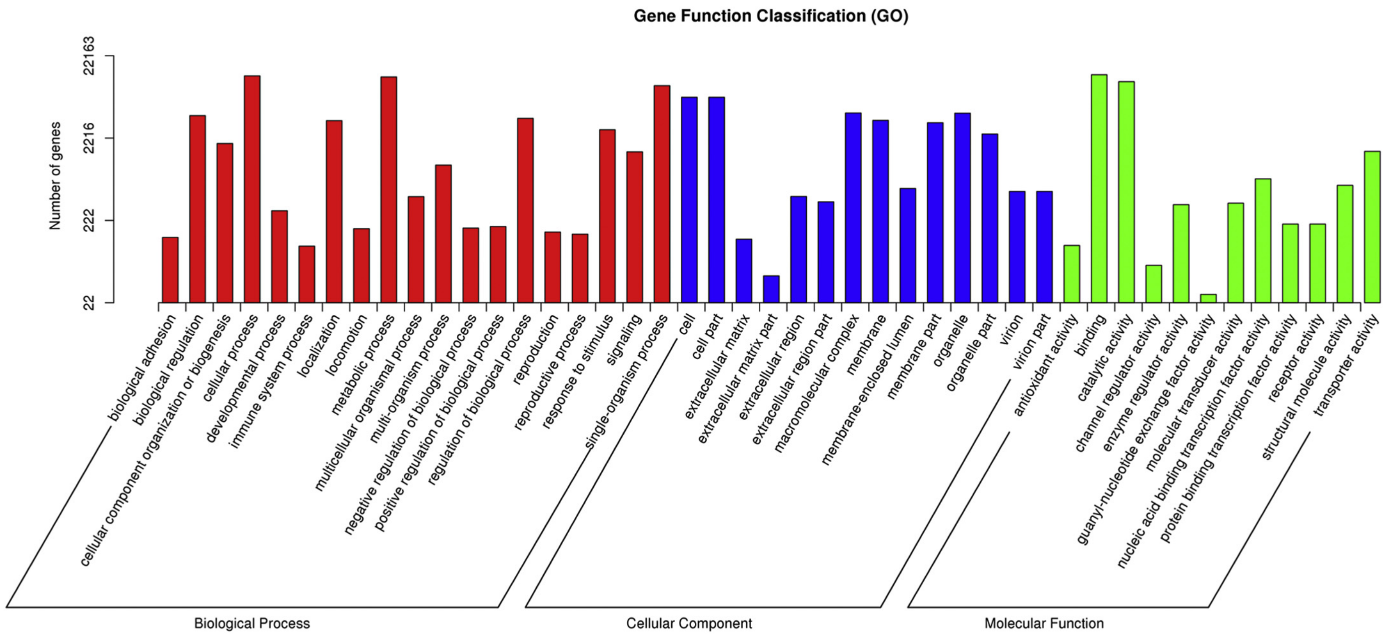



**Fig. 2.** Histogram of gene ontology (GO) classification. The genes were assigned to three main categories: biological process, cellular component, and molecular function.

are indicated in Fig. 7b. Among the down-regulated DEGs, two (c26640_g1 and c32118_g1) were involved in many pathways including "Circadian rhythm", "Hypertrophic cardiomyopathy", "Regulation of autophagy", "Adipocytokine signaling pathway", "mTOR signaling pathway", "FoxO signaling pathway", "Oxytocin signaling pathway", "AMPK signaling pathway", "Non-alcoholic fatty liver disease (NAFLD)", "Insulin signaling pathway", and "PI3K-Akt signaling pathway".

### *3.6. Screening of salinity-responsive genes in D data sets*

To further elucidate the genes responsible for salinity resistance in S1, 112 genes whose expression was significantly up- or down-regulated in S1 relative to S3 at 6, 12, 24 and 48 h after salinity-stress treatment in the D data sets were identified (Table S6). The expression patterns of these 112 DEGs were investigated with respect to expression before and

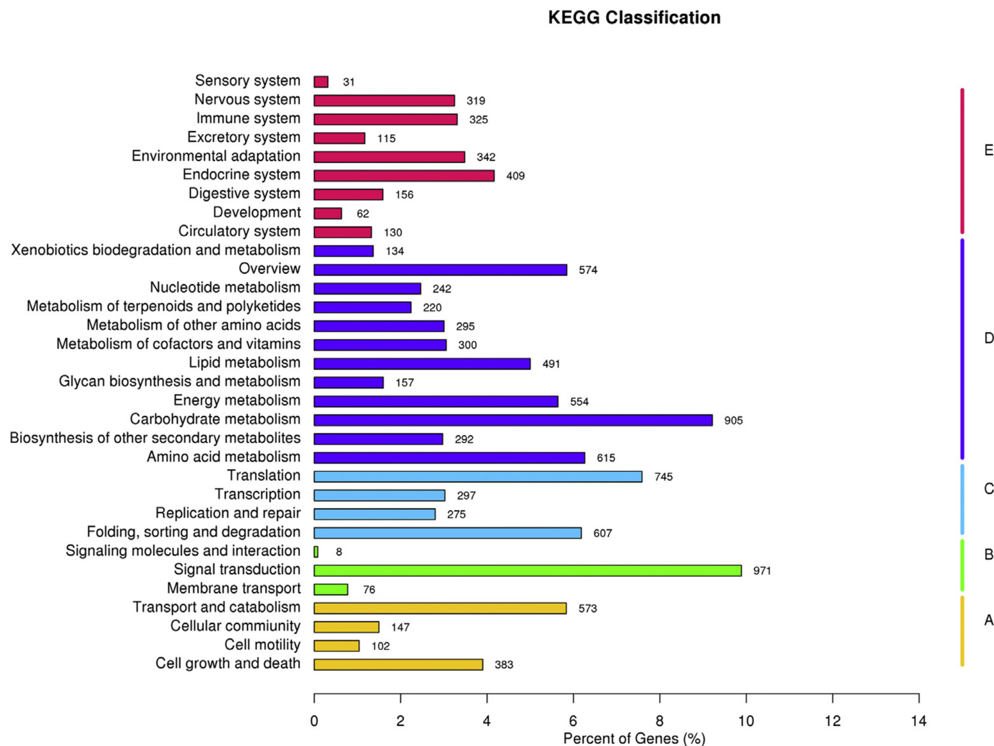



**Fig. 3.** Pathway assignment based on Kyoto Encyclopedia of Genes and Genomes (KEGG). (A) Cellular Processes; (B) Environmental Information Processing; (C) Genetic Information Processing; (D) Metabolism; and (E) Organismal Systems.

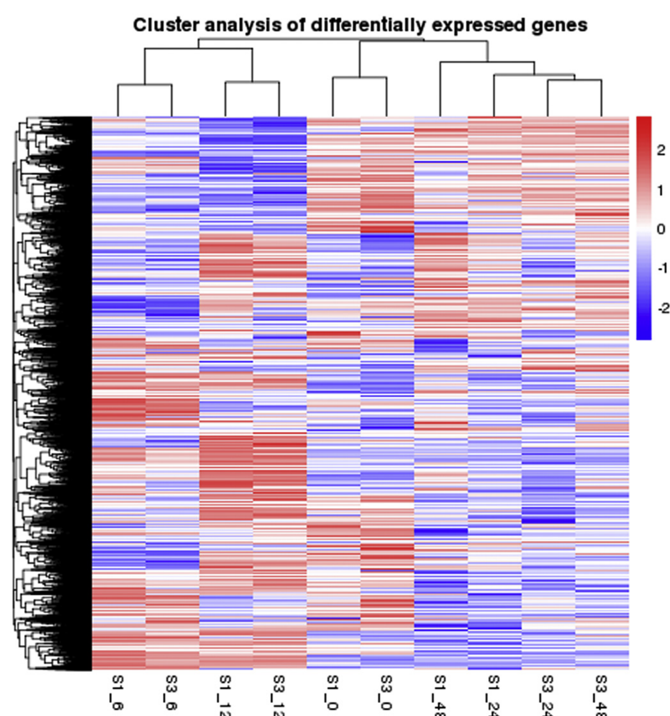



**Fig. 4.** Hierarchical clustering analysis of salinity-induced changes in gene expression in leaves of the salinity-tolerant, mutant peanut (S1) and the salinity-sensitive peanut (S3, Huayu 22). On the X axis, S1 and S3 refer to the two genotypes and the following number refers to 5 time points (from 0 to 48 h) before and after salinity-stress treatment. Each value of gene expression is an average of two replications. Red bar indicates upregulation, and blue bar indicates downregulation.

after salinity-stress treatment. Among the 112 DEGs, 86 were responsive to salinity-stress treatment in S1 and/or S3 (Table S7): 23 were responsive to salinity-stress treatment in both S1 and S3; 16 were responsive to salinity-stress treatment only in S1. In S3, 47 DEGs that were responsive to salinity stress treatment were detected with significant difference. Whereas 23 DEGs were significantly down regulated at 24 h, but still keep a high expression level in S1. The 86 DEGs above were mainly assigned to a few GO groups including oxidation–reduction process (5 DEGs), metabolic process (4 DEGs),

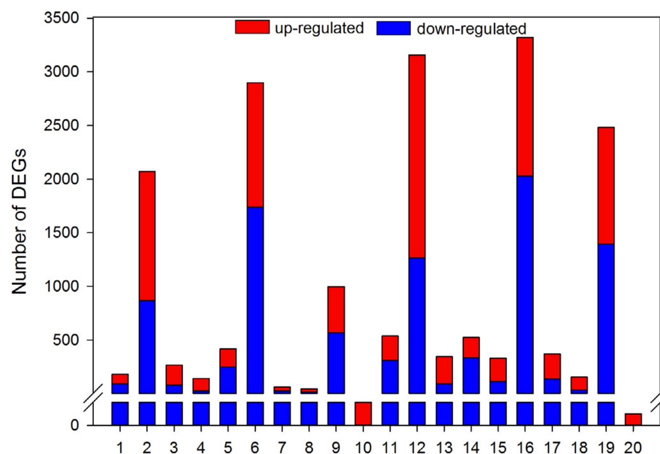



**Fig. 5.** Numbers of DEGs in the salinity-resistant mutant (S1) and in the salinity-sensitive parent (S3) before treatment (0 h) and 6 to 48 h after salinity-stress treatment. Data sets 1–10 represents S1_6 vs S1_0, S1_12 vs S1_0, S1_12 vs S1_6, S1_24 vs S1_0, S1_24 vs S1_6, S1_24 vs S1_12, S1_48 vs S1_0, S1_48 vs S1_6, S1_48 vs S1_12, S1_48 vs S1_24; data sets 11–20 represents S3_6 vs S3_0, S3_12 vs S3_0, S3_12 vs S3_6, S3_24 vs S3_0, S3_24 vs S3_6, S3_24 vs S3_12, S3_48 vs S3_0, S3_48 vs S3_6, S3_48 vs S3_12, S3_48 vs S3_24. Red numbers indicate up-regulation and blue numbers indicate down-regulation.

transport (3 DEGs), response to stress (2 DEGs), embryo development (2 DEGs), photosynthesis (2 DEGs), light harvesting (2 DEGs), pathogenesis (2 DEGs), proteolysis (2 DEGs), transcription, DNA-templated(2 DEGs), and sucrose metabolic process (2 DEGs). The 86 DEGs were mainly assigned to 13 GO groups including Cytochrome P450 CYP2 subfamily (3 DEGs), Aquaporin, major intrinsic protein family (2 DEGs), Cysteine proteinase Cathepsin L (2 DEGs), and Serine/threonine protein kinase (2 DEGs) (Table S7).

### 3.7. Experimental verification of DEGs

To determine whether our RNA-Seq identification of salinity-responsive genes in peanut was reliable, we used real-time PCR to monitor the expression pattern of 10 candidate DEGs at five times before and after the salinity-stress treatment (0, 6, 12, 24, and 48 h) for both S1 and S3. These candidate DEGs included genes that are known to be related to stress response in other plant species. The primers of selected genes are listed in Table S8. Their expression levels as determined by real-time PCR were correlated with the results of the RNA-seq (Fig. S2, Table S9).

## 4. Discussion

The transcriptome and DGE analysis in this study identified some genes that were previously reported to be stress-related. LEA proteins are known to be involved in protecting higher plants from damage caused by environmental stresses [10]. Metallothioneins (MTs) are cysteine-rich, low weight proteins that are responsive to many biotic and abiotic stresses including those induced by metal ions, physical damage, virus infection, and thermal shock [11,12,13]. Overexpression of *MT* genes in plants can sharply increase the chlorophyll content, reduce plant membrane lipid peroxidation, and protect the cell membrane system [14]. For example, Dundar et al. [15] found that an olive MT had the capability of effectively binding toxic heavy metal; it played an important role in metal homeostasis. Lipid transfer protein (LTP) is the most abundant of the waxy proteins; these proteins are involved in wax synthesis of cell wall [16,17]. Cameron et al. [18] found that increases in the expression of *LTP* genes were synchronous with the increase in wax synthesis after treatment with a stress such as drought, salinity, or low temperature. Overexpression of the *LTP* gene in plants sharply increased the resistance to stress [19]. Major intrinsic proteins (MIPs), or aquaporins, allow for the passive transport of water and other small, uncharged polar molecules across membrane lipid bilayers. MIPs are believed to be important for water transport under stressful environmental condition [20,21]. Calcineurin B-like protein-interacting protein kinases (CIPKs) belong to a $Ca^{2+}$-mediated CBL-CIPK network that responds to stress. For example, AtCIPK24/AtSOS2 can interact with AtCBL4/AtSOS to affect the $Na^{+}/H^{+}$ antiporter AtSOS1 and enhance tolerance to salt stress. AtCIPK24/AtSOS2 is also involved in reactive oxygen species (ROS) signaling and scavenging [22].

Some genes that were differentially expressed in S1 vs. S3 were screened to be responsive to salinity stress. These included genes that encoded LEAs (c27965_g1, c44015_g1, c32599_g1, and c28981_g1), LTPs (c45293_g1 and c23310_g1), MIPs (c25645_g1 and c35628_g1), CIPKs (c32118_g1), MTs (c17889_g1 and c66314_g1), lectins (c33018_g1 and c33018_g2), a Cys peroxiredoxin (c28163_g1), and a key enzyme in ABA biosynthesis, 9-cis-epoxycarotenoid dioxygenase (NCED, c38237_g1). Among these genes, three of the *LEA* genes were significantly up-regulated in S1 relative to S3 at 24 h (the log2 (fold change) > 2). Genes encoding MT (c17889_g1) and Cys peroxiredoxin (c28163_g1) were greatly up-regulated in S1 relative to S3 genotype at 24 h (the Log2 (fold change) > 6).

Other DEGs that may be related to stress were also detected in our study of S1 and S3, and these included genes encoding the two-component response regulator (c20830_g1) and cytochrome P450

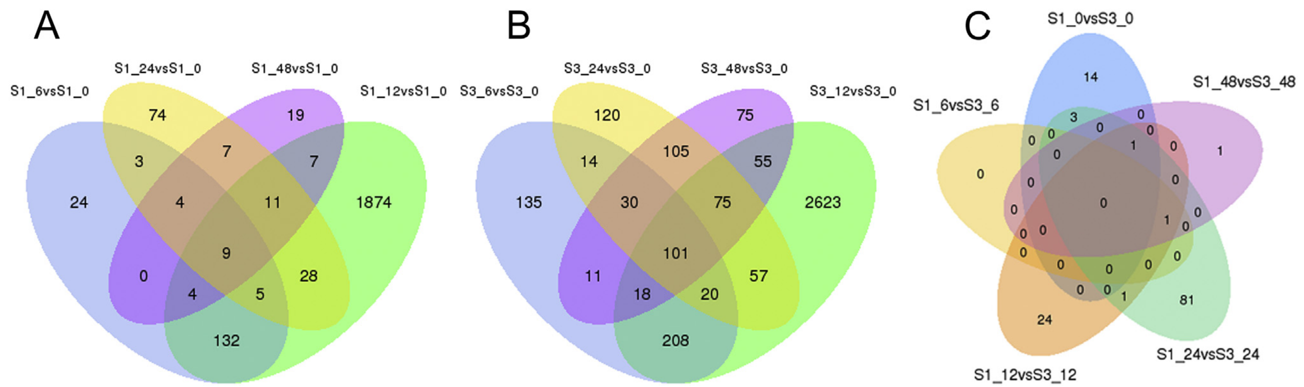



**Fig. 6.** Venn diagrams indicating the numbers of differentially expressed genes (DEGs) in the data sets of the S1 series (after vs. before salinity-stress treatment, (A)), the S3 series (after vs. before salinity-stress treatment, (B)), and the D series (after salinity stress treatment in S1 vs. S3, (C)). The numbers of DEGs exclusively expressed in one sample are shown in the non-overlapping regions. The numbers of DEGs with a common tendency of expression change between the two treatments are shown in the overlapping regions.

(c35292_g1 and c37768_g1). Genes encoding the universal stress protein (c34894_g1), cysteine proteinase precursor (c35409_g1), defensin precursor (c24576_g1), tumor-related protein (c30721_g1), and NAD-dependent epimerase/dehydratase (c37242_g1) were significantly up-regulated at 6 h and down-regulated at 12 and 24 h in S3 but not in S1. Especially at 24 h, significant increase was present (the Log2 (fold change) ration >5). In contrast, a gene encoding a xylem cysteine proteinase (c36154_g1), was significantly down-regulated in S1 but not in S3 (Table S7).

*Oleosin* genes are mainly expressed in seeds, where they help determine the oil content [23]. The promoters of these genes have key cis-regulatory elements including the ABA-responsive element, 1egeumin box, and ACGT motif, which are responsive to stress and plant hormones, such as abscisic acid and jasmonic acid [24,25]. Interestingly, four *oleosin* genes (c29774_g1, c28303_g1, c33419_g1, and c30836_g1) in our research were significantly up-regulated in S1 leaves and could also be involved in stress resistance in peanut.

The functional classification in the current study showed that some unknown genes also responded to salinity stress. These unknown genes, which included those that had no hits or low identity with the protein database and those that matched with unclassified and unknown proteins, represented 18 of the 86 DEGs. The involvement of these genes in salinity stress responses in peanut should be investigated.

## 5. Conclusions

RNA-Seq is an effective method for gene discovery. Using transcriptome sequencing and digital gene expression analysis, we identified some genes whose expression in response to salinity stress differed between the salinity-resistant genotype (S1) and the salinity-sensitive genotype (S3) of peanut. These DEGs included genes that are known to be related to salinity resistance, such as *lea*, *expansin*, *nsLTP*, *NCED*, *TIP*, and *oleosin* genes, and some unknown DEGs that might be related to salinity resistance. The information from this study can serve as a useful gene resource for the breeding of salinity resistance and provide an important reference for continuing research on the mechanism of salinity resistance in peanut.

https://doi.org/10.1016/j.ejbt.2017.12.002

## Conflict of interests

The authors declare that they have no competing interests.

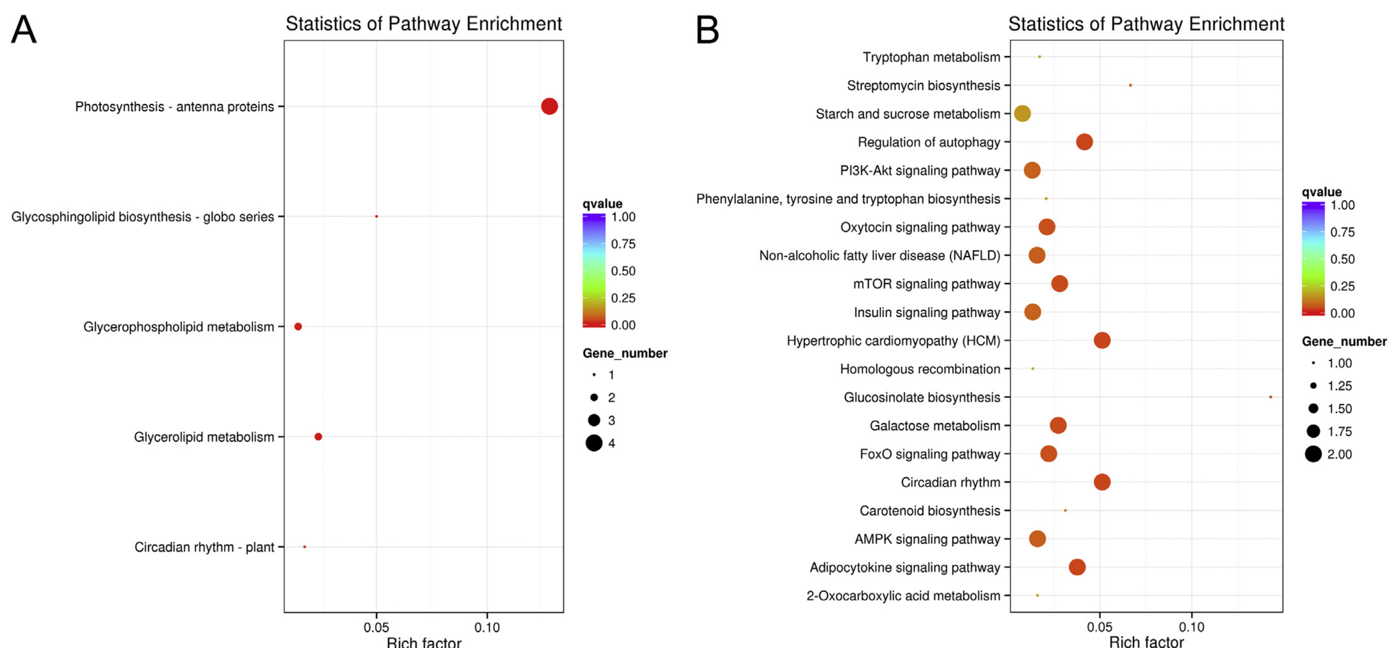



**Fig. 7.** KEGG pathway enrichment analyses of DEGs in D_12 (A) and D_24 (B) series data sets (S1 relative to S3 at 12 and 24 h).

## Financial support

We thank Prof. Bruce Jaffee for critical reading of the manuscript. This work was financially supported by the National Natural Science Foundation of the China (Grant No. 31471524, No. 31571705, and No. 31301356), and the Shandong Province Science and Technology Development Plan Project (Grant No. 2014GNC110002).

## References

[1] Govind G, Vokkaliga ThammeGowda H, Jayaker Kalaiarasi P, et al. Identification and functional validation of a unique set of drought induced genes preferentially expressed in response to gradual water stress in peanut. Mol Genet Genomics 2009;281(6):591–605. https://doi.org/10.1007/s00438-009-0432-z.

[2] Pandey MK, Monyo E, Ozias-Akins P, et al. Advances in *Arachis* genomics for peanut improvement. Biotechnol Adv 2012;30(3):639–51. https://doi.org/10.1016/j.biotechadv.2011.11.001.

[3] Guo BZ, Chen XP, Dang P, et al. Peanut gene expression profiling in developing seeds at different reproduction stages during *Aspergillus parasiticus* infection. BMC Dev Biol 2008;8:12. https://doi.org/10.1186/1471-213X-8-12.

[4] Zhao MX, Sun HY, Ji RR, et al. *In vitro* mutagenesis and directed screening for salt-tolerant mutants in peanut. Euphytica 2013;193(1):89–99. https://doi.org/10.1007/s10681-013-0956-5.

[5] Tombuloglu G, Tombuloglu H, Sakcali MS, et al. High-throughput transcriptome analysis of barley (*Hordeum vulgare*) exposed to excessive boron. Gene 2015; 557(1):71–81. https://doi.org/10.1016/j.gene.2014.12.012.

[6] Türktaş M, Yucebilgili Kurtoğlu K, Dorado G, et al. Sequencing of plant genomes–a review. Turk J Agric For 2015;39:361–76. https://doi.org/10.3906/tar-1409-93.

[7] Clouse JW, Adhikary D, Page JT, et al. The Amaranth genome: genome, transcriptome, and physical map assembly. Plant Genome 2016:9(1). https://doi.org/10.3835/plantgenome2015.07.0062.

[8] Song L, Li R, Xiang XH, et al. Overexpression of stress-inducible small GTP-binding protein AhRab7 (AhRabG3f) in peanut (*Arachis hypogaea* L.) enhances abiotic stress tolerance. J Food Agric Environ 2012;10(3–4):888–94.

[9] Grabherr MG, Haas BJ, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 2011;29:644–52. https://doi.org/10.1038/nbt.1883.

[10] Hundertmark M, Hincha DK. LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. BMC Genomics 2008;9:118. https://doi.org/10.1186/1471-2164-9-118.

[11] Zhao Y, Wu YR, Li Y, et al. OsMT1a, a type 1 metallothionein, plays the pivotal role in zinc homeostasis and drought tolerance in rice. Plant Mol Biol 2009;70(1–2): 219–29. https://doi.org/10.1007/s11103-009-9466-1.

[12] Mir G, Domenech J, Huguet G, et al. A plant type 2 metallothionein (MT) from cork tissue responds to oxidative stress. J Exp Bot 2004;55(408):2483–93. https://doi.org/10.1093/jxb/erh254.

[13] Kim SH, Lee HS, Song WY, et al. Chloroplast-targeted BrMT1 (*Brassica rapa* Type-1 metallothionein) enhances resistance to cadmium and ROS in transgenic *Arabidopsis* plants. J Plant Biol 2007;50(1):1–7. https://doi.org/10.1007/BF03030592.

[14] Huang GY, Wang YS. Expression and characterization analysis of type 2 metallothionein from grey mangrove species (*Avicennia marina*) in response to metal stress. Aquat Toxicol 2010;99(1):86–92. https://doi.org/10.1016/j.aquatox.2010.04.004.

[15] Dundar E, Sonmez GD, Unver T. Isolation, molecular characterization and functional analysis of *OeMT2*, an olive metallothionein with a bioremediation potential. Mol Genet Genomics 2015;290(1):187–99. https://doi.org/10.1007/s00438-014-0908-3.

[16] DeBono A, Yeats TH, Rose JKC, et al. *Arabidopsis* LTPG is a glycosylphosphatidylinositol- anchored lipid transfer protein required for export of lipids to the plant surface. Plant Cell 2009;21(4):1230–8. https://doi.org/10.1105/tpc.108.064451.

[17] Lee SB, Go YS, Bae HJ, et al. Disruption of glycosylphosphatidylinositol-anchored lipid transfer protein gene altered cuticular lipid composition, increased plastoglobules, and enhanced susceptibility to infection by the fungal pathogen *Alternaria brassicicola*. Plant Physiol 2009;150(1):42–54. https://doi.org/10.1104/pp.109.137745.

[18] Cameron KD, Teece MA, Smart LB. Increased accumulation of cuticular wax and expression of lipid transfer protein in response to periodic drying events in leaves of tree tobacco. Plant Physiol 2006;140(1):176–83. https://doi.org/10.1104/pp.105.069724.

[19] Jung HW, Kim KD, Hwang BK. Identification of pathogen-responsive regions in the promoter of a pepper lipid transfer protein gene (*CALTPI*) and the enhanced resistance of the *CALTPI* transgenic *Arabidopsis* against pathogen and environmental stresses. Planta 2005;221(3):361–73. https://doi.org/10.1007/s00425-004-1461-9.

[20] Weig A, Deswarte C, Chrispeels MJ. The major intrinsic protein family of *Arabidopsis* has 23 members that from three distinct groups with functional aquaporins in each group. Plant Physiol 1997;114(4):1347–57. https://doi.org/10.1104/pp.114.4.1347.

[21] Horie T, Kaneko T, Sugimoto G, et al. Mechanisms of water transport mediated by PIP aquaporins and their regulation via phosphorylation events under salinity stress in barley root. Plant Cell Physiol 2011;52(4):663–75. https://doi.org/10.1093/pcp/pcr027.

[22] Deng XM, Hu W, Wei SY, et al. *TaCIPK29*, a CBL-interacting protein kinase gene from wheat, confers salt stress tolerance in transgenic tobacco. PLoS One 2013; 8(7):e69881. https://doi.org/10.1371/journal.pone.0069881.

[23] Hu Z, Wang X, Zhan G, et al. Unusually large oilbodies are highly correlated with lower oil content in *Brassica napus*. Plant Cell Rep 2009;28(4):541–9. https://doi.org/10.1007/s00299-008-0654-2.

[24] van Rooijen CJ, Moloney MM. Plant seed oil-bodies as carriers for foreign proteins. Biotechnology 1995;13:72–7. https://doi.org/10.1038/nbt0195-72.

[25] Liu YH, Jia SR. Progress in the establishment of a plant seed oil body expression system. J Agric Biotechnol 2003;11(5):531–7.