Plant Science 181 (2011) 696–700

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Analysis of expressed sequence tags from biodiesel plant *Jatropha curcas* embryos at different developmental stages

Mao-Sheng Chen[a], Gui-Juan Wang[a], Ru-Ling Wang[a], Jun Wang[c], Song-Quan Song[b,*], Zeng-Fu Xu[a,**]

[a] *Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, 88 Xuefu Road, Kunming 650223, Yunnan, China*
[b] *Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China*
[c] *Beijing Genomics Institute at Shenzhen, Shenzhen 518000, China*

ARTICLE INFO

Article history:
Received 1 September 2010
Received in revised form 7 August 2011
Accepted 10 August 2011
Available online 31 August 2011

Keywords:
Bioenergy plant
Embryo
Expressed sequence tags
Lipid metabolism
Physic nut
Seed development

ABSTRACT

Jatropha curcas is considered a potential biodiesel feedstock plant whose seeds contain up to 40% oil. However, little is currently known about the seed biology of *Jatropha*. Therefore, it would be valuable to understand the mechanisms of development and lipid metabolism in *Jatropha* seeds. In the present study, three cDNA libraries were constructed with mRNA from *Jatropha* embryos at different stages of seed development. A total of 9844 expressed sequence tags (ESTs) were produced from these libraries, from which 1070 contigs and 3595 singletons were obtained. One hundred and seven unigenes were found to be differentially expressed in the three cDNA libraries of *Jatropha* embryos, indicating that these genes may play key roles in seed development. We have identified 59 and 61 unigenes that might be involved in the development and lipid metabolism in *Jatropha* seeds, respectively. Some of these genes may also play important roles in embryogenesis, morphogenesis, defense response and adaptive mechanisms in plants.

1. Introduction

Jatropha curcas L., belonging to the family Euphorbiaceae, is a drought-resistant oil plant that is widely distributed in tropical and subtropical areas [1,2]. *Jatropha* has been traditionally used as a living fence against domestic animals and for fertilizer, medicine production and control of soil erosion [1,3,4]. Unlike crops such as oil-seed rape, maize and soybean, *Jatropha* does not occupy farmland, and can produce non-edible seeds with high oil content (up to 40%) [5]. In recent years, the use of *Jatropha* as a biodiesel feedstock plant has greatly interested a large number of researchers. China, India and several African countries have initiated large-scale plantations of *Jatropha* [5,6]. *Jatropha* has small genome size (about 416 Mb) and few chromosome number ($2n = 22$) [7], and is amenable to genetic transformation [8–10]. Recently, *Jatropha* whole genome sequence was made publicly available [11], which make it suitable as a model plant for biodiesel feedstock research. As a wild plant, however, seed yield of *Jatropha* is poor and insufficient for the biodiesel industry [6,12,13]. Moreover, little is currently known about the genetic information and molecular biology of *Jatropha*. Therefore, it would be valuable to understand the mechanism of seed development and lipid metabolism in *Jatropha*, which would be helpful in using genetic engineering to develop new *Jatropha* cultivars.

Expressed sequence tag (EST) analysis can provide a convenient and efficient method for identification of genes expressed in specific tissues and cells, as well as allow the characterization of transcript expression levels [14]. Combined with breakthroughs in highly parallel designs for gene expression analysis, EST analysis can also provide valuable information for understanding the molecular basis of important agricultural traits in plants [15]. In this study, we sequenced the 5′-ends of about 10,000 cDNA clones randomly selected from three cDNA libraries derived from different developmental stages of *Jatropha* seeds. A large number of putative genes associated with signal transduction, synthesis of stored reserves, and metabolism and transport of amino acids, lipids, and proteins, have been identified. Therefore, this study provides a basis for understanding the mechanism of seed development and lipid metabolism in *Jatropha*. Moreover, the study can also provide a valuable resource for the cloning of new genes, annotation of genomic sequences and the development of molecular markers for gene mapping, polymorphism and marker-assisted selection breeding of *Jatropha*, which are prerequisites for the application of genetic engineering in *Jatropha* breeding [16].

* Corresponding author. Tel.: +86 10 6283 6484; fax: +86 10 6259 0833.
** Corresponding author. Tel.: +86 871 514 4423; fax: +86 871 516 0916.
E-mail addresses: sqsong@ibcas.ac.cn (S.-Q. Song), zfxu@xtbg.ac.cn, zengfu.xu@gmail.com (Z.-F. Xu).

0168-9452/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.plantsci.2011.08.003

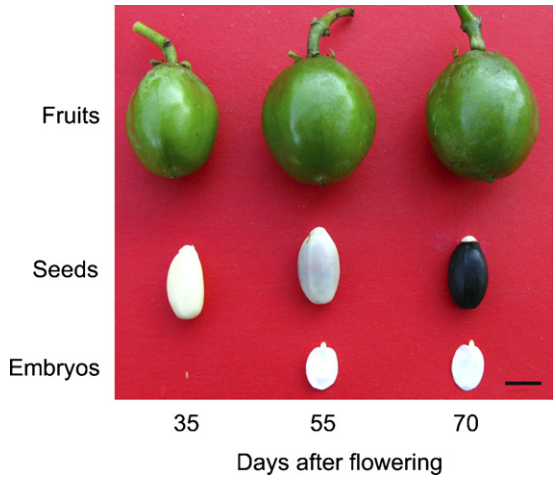

Fig. 1. *Jatropha curcas* fruits, seeds and embryos at different developmental stages. Seeds were harvested at 35–55, 56–70 and 71–95 days after flowering (DAF). Total RNA was extracted from embryos to construct cDNA libraries. Scale bar = 1.0 cm.

2. Materials and methods

2.1. cDNA library construction and sequencing

Jatropha seeds were harvested at 35–55, 56–70 and 71–95 days after flowering (DAF), which represent the early, middle and late stage of seed development, respectively (Fig. 1). Embryos were excised from seeds at the three development stages. mRNAs were purified with Oligotex mRNA Isolation Kits (Qiagen, Valencia, CA) from total RNA extracted from embryos. cDNAs were synthesized with the SuperScript II-RT system (Invitrogen). Three cDNA libraries, early stage library (35–55 DAF, library I), middle stage library (56–70 DAF, library II) and late stage library (71–95 DAF, library III), were generated with the plasmid pBluescript II SK (+). cDNA colonies in each library were picked randomly, and sequenced once from the 5′-end of each clone using an automated DNA sequencer (GE MegaBase 1000 sequencers) at the Beijing Genomics Institute.

2.2. Sequence assembly and analysis

After removing ribosomal RNA, poly(A), vector and low-quality sequences, high-quality ESTs (length of sequence ≥ 100 bp and phred quality value ≥ 20) were assembled into clusters using the PHRED/PHRAP/CONSED software package [17,18]. The 4665 unigenes were searched against the NCBI non-redundant nucleotide databases (NT) using the blastn program with an *E*-value ≤ 1.0E−05, and the NCBI non-redundant protein database (NR) and the SWISS-PROT database using the blastx program with an *E*-value ≤ 1.0E−05, and Kyoto Encyclopedia of Genes and Genomes database (KEGG, http://www.genome.jp/kegg/) using the blastx program with an *E*-value ≤ 1.0E−10 [19,20]. Functional classification of the ESTs was further examined according to the NCBI Clusters of Orthologous Groups of Proteins (COGs) database using the blastx program with an *E*-value ≤ 1.0E−10 [21].

2.3. EST expression profiling

Data analysis on gene expression profiles in three cDNA libraries was performed using R statistics with Bonferonni correction at the significance threshold of 1.85E−05 using the web tool IDEG6 [22,23]

Table 1
Summary of the number of unigenes in three cDNA libraries from *Jatropha curcas* embryos.

Library	Unigenes	Redundancy (%)	Mean length (bp)
I. Early stage (35–55 DAF)	2295	32.5	444.4
II. Middle stage (56–70 DAF)	1646	55.3	476.2
III. Late stage (71–95 DAF)	1512	46.7	536.4
Total	4665	52.6	496.2

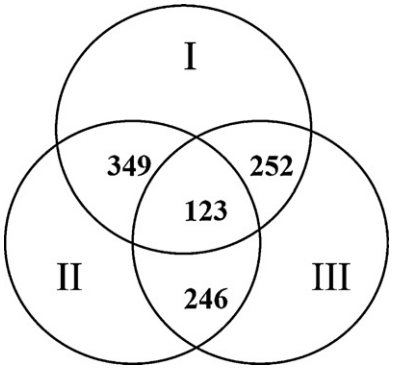

Fig. 2. Overlapping of unigenes detected in three libraries. A total of 2295, 1646, and 1512 unigenes were present in libraries I, II and III, respectively. Among them, 123 unigenes were common in the three libraries. Libraries I and II shared 349 unigenes, libraries II and III shared 252, and libraries III and I shared 246.

3. Results and discussion

3.1. Sequencing and analysis of Jatropha ESTs

Three cDNA libraries were constructed from *Jatropha* embryos at different developmental stages of seeds (Fig. 1). From each cDNA library, 3000–4000 clones were sequenced, from which a total of 10,913 5′-end sequences were generated. After trimming low-quality and vector sequences and removing ribosomal RNA sequences, 9844 high-quality ESTs with a minimum of ≥100 bp, phred quality value ≥ 20 and average length of 496 bp were obtained. All sequences have been deposited in DDBJ/EMBL/GenBank (FM887038–FM896881). The 9844 ESTs were assembled into 4665 unigenes containing 1070 contigs and 3595 singletons using PHRED/PHRAP/CONSED software, which represent putative transcripts that vary during seed development in *Jatropha* (Additional file 1). Libraries I, II and III contains 2295, 1646 and 1512 unigenes, respectively (Table 1). There was an 8.86% overlap between libraries I and II, a 6.62% overlap between I and III, and a 7.79% overlap between II and III, respectively (Table 1; Fig. 2). Surprisingly, there was a 2.64% overlap among three libraries and only 123 unigenes were common (Fig. 2). These results show that the development of *Jatropha* embryo is a complex process with progressive changes in physiology and morphology, involving a large number of genes that are specifically expressed during different developmental stages. In addition, because only 3000–4000 clones have been sequenced from each library, some common genes expressed in the three libraries have not been identified.

3.2. Annotation and functional classification of Jatropha unigenes

A homology search by using the blastx program revealed that 65%, 40.8% and 59.5% of the 4665 *Jatropha* unigenes had significant matches with sequences in the NCBI non-redundant protein database (NR), SWISS-PROT protein database and the Kyoto Encyclopedia of Genes and Genomes database (KEGG), respectively. By using the blastn program, about 97.9% of these unigenes were

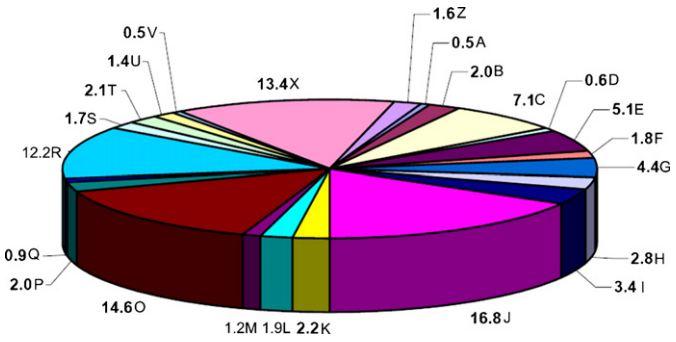

Fig. 3. Functional categories and percentage of unigenes from *Jatropha curcas* embryos. A total of 1002 unigenes were classified into 23 functional categories according to their predicted gene products using the Cluster of Orthologs (COG) with an *E*-value ≤ 1.0E−10. A: RNA processing and modification. B: Chromatin structure and dynamics. C: Energy production and conversion. D: Cell cycle control, cell division, chromosome partitioning. E: Amino acid transport and metabolism. F: Nucleotide transport and metabolism. G: Carbohydrate transport and metabolism. H: Coenzyme transport and metabolism. I: Lipid transport and metabolism. J: Translation, ribosomal structure and biogenesis. K: Transcription. L: Replication, recombination and repair. M: Cell wall/membrane/envelope biogenesis. O: Post-translational modification, protein turnover, chaperones. P: Inorganic ion transport and metabolism. Q: Secondary metabolites biosynthesis, transport and catabolism. R: General function prediction only. S: Function unknown. T: Signal transduction mechanisms. U: Intracellular trafficking, secretion, and vesicular transport. V: Defense mechanisms. X: Unassigned. Z: Cytoskeleton.

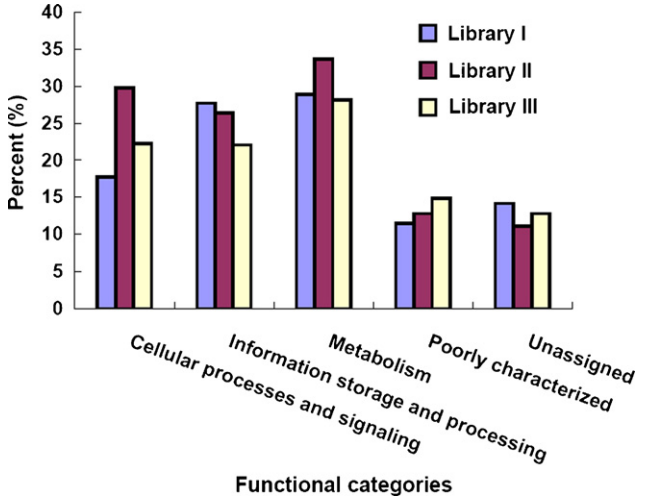

Fig. 4. Functional categories and percentage of unigenes from *Jatropha curcas* embryos at different developmental stages. A total of 1002 unigenes were classified into 5 major functional categories using the Cluster of Orthologs (COG) with an *E*-value ≤ 1.0E−10. Library I represents the early stage of seed development (35–55 DAF); Library II represents the middle stage of seed development (56–70 DAF); Library III represents the late stage of seed development (71–95 DAF).

found to have similarity with sequences in the present data sets of *Jatropha* genome and ESTs [11,24,25]. A total of 1002 unigenes were classified into 23 functional categories according to their predicted gene products using the Cluster of Orthologs (COG) (Fig. 3; Additional file 2). In our functional classification, the "translation, ribosomal structure and biogenesis" and "posttranslational modification, protein turnover, chaperones" categories were the most prominent transcripts (16.8% and 14.6%, respectively) in *Jatropha* embryos (Fig. 3). However, categories with no concrete assignment, such as "function unknown", "general function prediction only" and "unassigned", accounted for a large fraction of transcripts (27.0%). In addition, the percentage of unigenes in the "metabolism" category was the highest for each library (28.9%, 33.7% and 28.1%, respectively) (Fig. 4). Moreover, genes involved in "cellular processes and signaling" and "metabolism" and those involved in "information storage and processing" are preferentially expressed during the middle and early stages of seed development, respectively (Fig. 4).

3.3. Analysis of differentially expressed genes

To analyze differentially expressed genes among three cDNA libraries from *Jatropha* embryos, an R statistic was calculated for all 1070 unigenes (EST number ≥ 2) using the Identification of Differentially Expressed Genes 6 test statistics (IDEG6) with Bonferroni correction at the significance threshold of 1.85E−05 [22]. From this analysis, 107 unigenes were found to be differentially expressed in *Jatropha* seeds (Additional file 3). Among them, unigenes encoding putative ribosomal protein L2, NADH–plastoquinone oxidoreductase subunit and gibberellin-responsive protein were abundant in library I. Unigenes encoding putative seed storage proteins (SSPs), late embryogenesis abundant proteins (LEAs), heat shock proteins (HSPs), oleosins, early methionine labeled (Em) proteins, GTP-binding nuclear protein Ran-3, cold acclimation-induced proteins and AWPM-19-like membrane family proteins and those encoding senescence-associated proteins and cytochrome P450 TBPs were abundant in libraries II and III, respectively. A number of unigenes related to seed maturation, such as encoding SSP, LEA and HSP, were highly expressed in library II, indicating that *Jatropha* seeds during 56–70 DAF reached the mature stage. Moreover, unigenes related to senescence showed high expression in library III, suggesting that the metabolic activities of seeds are decreasing and preparing for dormancy. Interestingly, there were 9 unigenes with no match in the databases, and 7 unigenes encoding unknown proteins were specifically expressed in library I and/or II. We speculate that these genes may reveal the existence of novel cellular processes or biochemical pathways [26], and may play important roles during *Jatropha* seed development.

3.4. Seed development-related genes

Seed development is a complex process in which the embryo and endosperm develop, storage products accumulate and desiccation tolerance forms, leading to seed dormancy. All these processes require the concerted action of several signaling pathways related to genetic programs, hormonal and metabolic signals [27]. Therefore, identifying the genes involved in seed development and their function is the key to understand the developmental mechanisms of *Jatropha* seeds.

In this study, we identified 59 seed development-related unigenes (Additional file 4) with significant similarity to *Arabidopsis* genes [28]. Mutations of most of these genes cause defects in embryo development, whereas mutations of the last three genes lead to morphologically normal seeds with albino, pale green or fuscate color, indicating that they are essential to seed pigmentation [28]. In addition, a unigene (FM895270) encoding a putative invertase was identified, which irreversibly cleave sucrose into glucose and fructose, and is a key metabolic enzyme involved in plant life cycle and adaptation of plants to various environmental conditions [29]. FM887520, a homolog of At2g33880, was annotated as the WUSCHEL-related homeobox 9 (*STIP/WOX9*), which can promote cell proliferation and prevent premature differentiation in meristematic tissues during post-embryonic development [30,31]. In addition, Contig252 was a homolog of At1g15750 encoding a transcriptional co-repressor TOPLESS (TPL), which mediates auxin-dependent transcriptional repression by interacting with IAA12/BODENLOS (IAA12/BDL) through an ethylene response factor (ERF)-associated amphiphilic repression (EAR) motif during embryogenesis [32]. Although *ABSCISIC ACID INSENSITIVE 3* (*ABI3*), *FUSCA3* (*FUS3*) and *LEAFY COTYLEDONS* (*LEC1* and *LEC2*) are key transcriptional regulators during seed maturation of the model plant

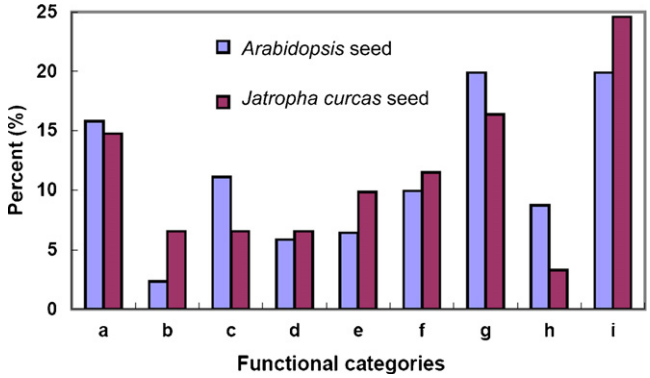

Fig. 5. Comparison of genes related to functional categories of lipid metabolism in seeds between *Arabidopsis* and *Jatropha curcas*. Similar to *Arabidopsis* seeds, lipid metabolism is a predominant function for *Jatropha* seed genes based on the percentage of expressed genes responsible for this function. a: Synthesis of fatty acids in plastids. b: Synthesis of membrane lipids in plastids. c: Synthesis of membrane lipids in the endomembrane system. d: Metabolism of acyl-lipids in mitochondria. e: Synthesis and storage of oil. f: Degradation of storage lipids and straight fatty acids. g: Lipid signaling. h: Fatty acid elongation and wax and cutin metabolism. i: Miscellaneous.

Arabidopsis [33], only FM889563, a homolog of *ABI3*, was identified in *Jatropha* embryos. So, we hypothesize that the transcripts of *FUS3*, *LEC1* and *LEC2* are very low in abundance in *Jatropha* embryos, and was not detectable in our libraries.

The above results show that many genes involved in the seed development of *Arabidopsis* are also present in *Jatropha* seeds, and, to some extent, indicate that the pattern of seed development is conserved in plants. Although a large number of genes related to seed development of *Jatropha* have been obtained, their functional analysis is based only on sequence similarity to the genes whose function have been determined in *Arabidopsis* [28], and their definitive function should be determined by further experiments.

3.5. Lipid metabolism-related genes

Seed oils, composed mainly of triacylglycerols (TAGs), are not only a significant source of fatty acids for human nutrition, but also important material for industrial products [34]. Understanding the mechanism of lipid metabolism and identifying genes involved in this process are prerequisites for using genetic engineering to improve seed oil quantity and quality. In this study, we identified 61 unigenes from our libraries that show sequence similarity with *Arabidopsis* genes that are associated with lipid metabolism (Additional file 5) [35]. Among them, 24.6%, 16.4% and 14.8% of genes were related to the miscellaneous, lipid signaling and synthesis of fatty acids in plastids categories, respectively, in *Jatropha* seeds, similar to those of *Arabidopsis* seeds (19.9%, 19.9% and 15.8, respectively) (Fig. 5). It should be noted that unigenes encoding putative oleosins were the most abundant in *Jatropha* embryos. In *Arabidopsis*, oleosins could limit the coalescence of oil bodies during the cytoplasm compression caused by seed desiccation. Oleosin gene suppression would result in an aberrant phenotype of embryo cells, which causes the disruption of storage organelles, thus altering the accumulation of lipids and proteins and causing a delay in germination [36]. Moreover, unigenes encoding putative lipid transfer proteins (LTPs) were also abundant in *Jatropha* embryos. They are thought to participate in the cutin formation of plants, embryogenesis, symbiosis, defense reactions, and response to environmental stimuli [37]. We have found 24 unigenes involved in lipid metabolism pathway in the KEGG (Additional file 6). They are present in the fatty acid biosynthesis (ko00061), fatty acid metabolism (ko00071), synthesis and degradation of ketone bodies (ko00072), biosynthesis of steroids (ko00100), bile acid biosynthesis (ko00120), glycerolipid metabolism (ko00561), glycerphospholipid metabolism (ko00564), ether lipid metabolism (ko00565), arachidonic acid metabolism (ko00590) and sphingolipid metabolism (ko00600), respectively (Additional file 2; Additional file 6). Moreover, FM895690, a homolog of At1g15080, encoded a putative lipid phosphate phosphatase 2 (AtLPP2), which functions as a negative regulator upstream of ABI4 and its T-DNA mutants show hypersensitivity to ABA and significant accumulation of phosphatidic acid during germination [38]. FM887047 was annotated as a putative CTP:phosphorylethanolamine cytidylyltransferase (PECT), which is considered to be the rating-limiting enzyme in the cytidyldiphosphate ethanolamine pathway in phosphatidylethanolamine (PE) biosynthesis [39]. In *Arabidopsis*, its mutations display embryogenesis and morphogenesis defects and related to the lethality of *Arabidopsis* embryo at preglobular stages [40].

From the statement above, we can conclude that genes involved in lipid metabolism not only are essential for oil synthesis but also play important roles in embryogenesis, morphogenesis, defense response and adaptive mechanisms in plants. Therefore, while using genetic engineering to improve oil content and/or other useful components of *Jatropha* seeds, it is necessary to consider that normal plant growth may be affected. So far, all 694 genes that have been identified in *Arabidopsis* have the predicted function in lipid metabolism, including 171 genes expressed in seeds [35]. However, only 61 of these genes are present in *Jatropha* embryos, indicating that a majority of genes involved in lipid metabolism remain undiscovered. We believe that the main reasons for this are as follows: (1) 171 *Arabidopsis* genes were identified from whole seeds, whereas the unigenes of *Jatropha* identified in this study were only from embryos. It is possible that many genes involved in lipid metabolism are particularly expressed in the endosperm and seed coat, rather than in embryos. (2) Due to the differences of species, a number of genes from *Jatropha* embryos cannot be annotated using *Arabidopsis* databases, causing some genes to remain unidentified. (3) As mentioned above, some genes were missed in our libraries because they were present at low transcription level. Future studies are needed to identify the genes expressed in the endosperm because they may play more important roles in lipid metabolism than embryonic genes.

4. Conclusion

Understanding the mechanisms of development and lipid metabolism in *Jatropha* seeds and identifying the genes involved in these processes would ultimately be necessary for successful genetic engineering of *Jatropha*. In this study, a large number of ESTs were generated from cDNA libraries of *Jatropha* embryos at different developmental stages. Among them, some genes involved in seed development and lipid metabolism were identified that may play important roles in embryogenesis, morphogenesis, defense reactions and adaptation to various environmental conditions. This is very useful information for improving the breeding of *Jatropha*. However, to better elucidate the mechanisms of lipid accumulation in *Jatropha* seeds, *Jatropha* endosperm cDNA libraries should be constructed to obtain more information on seed development and lipid metabolism.

Acknowledgements

This work was supported by the Knowledge Innovation Programs of the Chinese Academy of Sciences (grant no. KSCX2-YW-Z-0723) and the Top Science and Technology Talents Scheme of Yunnan Province (grant no. 2009CI123).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.plantsci.2011.08.003.

References

[1] K. Openshaw, A review of *Jatropha curcas*: an oil plant of unfulfilled promise, Biomass Bioenergy 19 (2000) 1–15.
[2] A. Kumar, S. Sharma, An evaluation of multipurpose oil seed crop for industrial uses (*Jatropha curcas* L.): a review, Ind. Crops Prod. 28 (2008) 1–10.
[3] P.A. Wender, J.M. Kee, J.M. Warrington, Practical synthesis of prostratin, DPP, and their analogs, adjuvant leads against latent HIV, Science 320 (2008) 649.
[4] M. Debnath, P.S. Bisen, *Jatropha curcas* L., a multipurpose stress resistant plant with a potential for ethnomedicine and renewable energy, Curr. Pharm. Biotechnol. 9 (2008) 288–306.
[5] D. Fairless, The little shrub that could – maybe, Nature 449 (2007) 652–655.
[6] K. Sanderson, Wonder weed plans fail to flourish, Nature 461 (2009) 328–329.
[7] C.R. Carvalho, W.R. Clarindo, M.M. Pra a, F.S. Araújo, N. Carels, Genome size, base composition and karyotype of *Jatropha curcas* L., an important biofuel plant, Plant Sci. 174 (2008) 613–617.
[8] M. Li, H. Li, H. Jiang, X. Pan, G. Wu, Establishment of an *Agrobacteriuim*-mediated cotyledon disc transformation method for *Jatropha curcas*, Plant Cell Tissue Organ Cult. 92 (2008) 173–181.
[9] J.L. Pan, Q.T. Fu, Z.F. Xu, *Agrobacterium tumefaciens*-mediated transformation of biofuel plant *Jatropha curcas* using kanamycin selection, Afr. J. Biotechnol. 9 (2010) 6477–6481.
[10] M. Joshi, A. Mishra, B. Jha, Efficient genetic transformation of *Jatropha curcas* L. by microprojectile bombardment using embryo axes, Ind. Crops Prod. 33 (2011) 67–77.
[11] S. Sato, H. Hirakawa, S. Isobe, E. Fukai, A. Watanabe, M. Kato, K. Kawashima, C. Minami, A. Muraki, N. Nakazaki, Sequence analysis of the genome of an oil-bearing tree, *Jatropha curcas* L, DNA Res. 18 (2011) 65–76.
[12] B.N. Divakara, H.D. Upadhyaya, S.P. Wani, C.L.L. Gowda, Biology and genetic improvement of *Jatropha curcas* L.: a review, Appl. Energy 87 (2010) 732–742.
[13] B.-Z. Pan, Z.-F. Xu, Benzyladenine treatment significantly increases the seed yield of the biofuel plant *Jatropha curcas*, J. Plant Growth Regul. 30 (2011) 166–174.
[14] S. Audic, J.M. Claverie, The significance of digital gene expression profiles, Genome Res. 7 (1997) 986–995.
[15] D.J. Duggan, M. Bittner, Y. Chen, P. Meltzer, J.M. Trent, Expression profiling using cDNA microarrays, Nat. Genet. 21 (1999) 10–14.
[16] J. Gressel, Transgenics are imperative for biofuel crops, Plant Sci. 174 (2008) 246–263.
[17] B. Ewing, L. Hillier, M.C. Wendl, P. Green, Base-calling of automated sequencer traces using phred. I. Accuracy assessment, Genome Res. 8 (1998) 175–185.
[18] J. Burke, D. Davison, W. Hide, d2_cluster: a validated method for clustering EST and full-length cDNA sequences, Genome Res. 9 (1999) 1135–1142.
[19] D.A. Benson, I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, B.A. Rapp, D.L. Wheeler, GenBank, Nucleic Acids Res. 30 (2002) 17–20.
[20] R. Apweiler, M. Biswas, W. Fleischmann, A. Kanapin, Y. Karavidopoulou, P. Kersey, E.V. Kriventseva, V. Mittard, N. Mulder, I. Phan, Proteome analysis database: online application of InterPro and CluSTr for the functional classification of proteins in whole genomes, Nucleic Acids Res. 29 (2001) 44–48.
[21] R.L. Tatusov, N.D. Fedorova, J.D. Jackson, A.R. Jacobs, B. Kiryutin, E.V. Koonin, D.M. Krylov, R. Mazumder, S.L. Mekhedov, A.N. Nikolskaya, B.S. Rao, S. Smirnov, A.V. Sverdlov, S. Vasudevan, Y.I. Wolf, J.J. Yin, D.A. Natale, The COG database: an updated version includes eukaryotes, BMC Bioinform. 4 (2003) 41.
[22] C. Romualdi, S. Bortoluzzi, F. D'Alessi, G.A. Danieli, IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments, Physiol. Genomics 12 (2003) 159–162.
[23] D.J. Stekel, Y. Git, F. Falciani, The comparison of gene expression from multiple cDNA libraries, Genome Res. 10 (2000) 2055–2061.
[24] G.G.L. Costa, K.C. Cardoso, L.E.V. Del Bem, A.C. Lima, M.A.S. Cunha, L. de Campos-Leite, R. Vicentini, F. Papes, R.C. Moreira, J.A. Yunes, F.A.P. Campos, M.J. Da Silva, Transcriptome analysis of the oil-rich seed of the bioenergy crop *Jatropha curcas* L., BMC Genomics 11 (2010) 462.
[25] P. Natarajan, D. Kanagasabapathy, G. Gunadayalan, J. Panchalingam, N. Shree, P.A. Sugantham, K.K. Singh, P. Madasamy, Gene discovery from *Jatropha curcas* by sequencing of ESTs from normalized and full-length enriched cDNA library from developing seeds, BMC Genomics 11 (2010) 606.
[26] I. Tzafrir, R. Pena-Muralla, A. Dickerman, M. Berg, R. Rogers, S. Hutchens, T.C. Sweeney, J. McElver, G. Aux, D. Patton, D. Meinke, Identification of genes required for embryo development in *Arabidopsis*, Plant Physiol. 135 (2004) 1206–1220.
[27] L. Gutierrez, O. Van Wuytswinkel, M. Castelain, C. Bellini, Combined networks regulating seed maturation, Trends Plant Sci. 12 (2007) 294–300.
[28] D. Meinke, R. Muralla, C. Sweeney, A. Dickerman, Identifying essential genes in *Arabidopsis thaliana*, Trends Plant Sci. 13 (2008) 483–491.
[29] T. Roitsch, M.C. González, Function and regulation of plant invertases: sweet sensations, Trends Plant Sci. 9 (2004) 606–613.
[30] X. Wu, T. Dabi, D. Weigel, Requirement of homeobox gene *STIMPY/WOX9* for *Arabidopsis* meristem growth and maintenance, Curr. Biol. 15 (2005) 436–440.
[31] X. Wu, J. Chory, D. Weigel, Combinations of *WOX* activities regulate tissue proliferation during *Arabidopsis* embryonic development, Dev. Biol. 309 (2007) 306–316.
[32] J.A. Long, C. Ohno, Z.R. Smith, E.M. Meyerowitz, TOPLESS regulates apical embryonic fate in *Arabidopsis*, Science 312 (2006) 1520–1523.
[33] M. Santos-Mendoza, B. Dubreucq, S. Baud, F. Parcy, M. Caboche, L. Lepiniec, Deciphering gene regulatory networks that control seed development and maturation in *Arabidopsis*, Plant J. 54 (2008) 608–620.
[34] D.H. Hobbs, J.E. Flintham, M.J. Hills, Genetic control of storage oil synthesis in seeds of *Arabidopsis*, Plant Physiol. 136 (2004) 3341–3349.
[35] F. Beisson, A.J.K. Koo, S. Ruuska, J. Schwender, M. Pollard, J.J. Thelen, T. Paddock, J.J. Salas, L. Savage, A. Milcamps, V.B. Mhaske, Y.H. Cho, J.B. Ohlrogge, *Arabidopsis* genes involved in acyl lipid metabolism. A 2003 census of the candidates, a study of the distribution of expressed sequence tags in organs, and a web-based database, Plant Physiol. 132 (2003) 681–697.
[36] R.M.P. Siloto, K. Findlay, A. Lopez-Villalobos, E.C. Yeung, C.L. Nykiforuk, M.M. Moloney, The accumulation of oleosins determines the size of seed oilbodies in *Arabidopsis*, Plant Cell 18 (2006) 1961–1974.
[37] J.C. Kader, Lipid-transfer proteins in plants, Annu. Rev. Plant Physiol. Plant Mol. Biol. 47 (1996) 627–654.
[38] T. Katagiri, K. Ishiyama, T. Kato, S. Tabata, M. Kobayashi, K. Shinozaki, An important role of phosphatidic acid in ABA signaling during germination in *Arabidopsis thaliana*, Plant J. 43 (2005) 107–117.
[39] F.Q. Tang, T.S. Moore, Enzymes of the primary phosphatidylethanolamine biosynthetic pathway in postgermination castor bean endosperm – developmental profiles and partial purification of the mitochondrial CTP: ethanolaminephosphate cytidylyltransferase, Plant Physiol. 115 (1997) 1589–1597.
[40] J. Mizoi, M. Nakamura, I. Nishida, Defects in CTP:phosphorylethanolamine cytidylyltransferase affect embryonic and postembryonic development in *Arabidopsis*, Plant Cell 18 (2006) 3370–3385.