INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
16–348/2016/18–6–1143–1152
DOI: 10.17957/IJAB/15.0218
http://www.fspublishers.org

Full Length Article

Differential Gene Regulation of Lipid Synthesis in the Developing Seeds of Two Biodiesel Tree Species, *Jatropha* and *Vernicia*

Dan Yang[‡], Huaiyun Zhang[‡], Kuan Peng, Lili Chen, Hanjie He, Xiaoxi Huang, Jieming Qin, Gongxiu He and Dangquan Zhang[*]

Key Laboratory of Cultivation and Protection for Non-Wood Forest Trees (Ministry of Education) & Hunan Provincial Key Laboratory of Forestry Biotechnology, Central South University of Forestry and Technology, Changsha 410004, China

[‡]These authors contributed equally to this work

[*]For correspondence: zhangdangquan@163.com

Abstract

The fatty acid compositions of *Jatropha* oil and *Vernicia* oil are strikingly different, which leads to a great difference in combustion performance, low temperature performance and oxidation stability. A comparative transcriptomic study was made in *Vernicia* and *Jatropha*, with a focus on the gene regulation of differential oil accumulation process. Transcriptome sequencing was conducted with seeds at the initial- and fast- stage of oil accumulation from both. More than 24 billion bases of cDNA sequence were obtained, with 49,583 and 45,414 high-quality unigenes identified for *Vernicia* and *Jatropha* seeds, respectively. Multiple comparative transcriptome approaches revealed a number of species-specific fatty acid desaturases (FAD2, FADX, FAH12 etc.) contributing to their differentiated fatty acid compositions in seeds of *Vernicia* and *Jatropha*. Meanwhile, the results suggested that DGAT majorly regulates TAG synthesis than PDAT in *Vernicia* seed, and PDAT may have more important role regulating TAG synthesis in *Jatropha* seed than in *Vernicia* seed. It was also implied that specific oleosins involving in oil bodies may have member bias and may affect lipid contents in seeds of *Vernicia* and *Jatropha*, as some of which were 30-50 fold up-regulated (with their RPKM values over 10,000 at fast-stage). Some important factors were identified and can differentially regulate lipid pathways in seeds of *Vernicia* and *Jatropha*. © 2016 Friends Science Publishers

Keywords: Biodiesel; Transcriptomic analysis; Lipid synthesis; Vernicia; Jatropha

Introduction

The increasing demand for diesel coupled with continuous air pollution concerns has stimulated the development for an ecologically sustainable and alternative renewable fuel source. Biodiesel production by oil-rich plants as the alternative to petroleum fuel is one of the most energy-rich and abundant forms of reduced carbon available from nature (Fairless, 2007; Durrett *et al.*, 2008). Most of the herbaceous oilseed crops grown today are annuals. Compared to these plants, the oil trees have deeper root systems, which help store more carbon and maintain soil quality. They have therefore been advocated as potentially more efficient ways of farming, especially on degraded soils and waste lands unsuitable for food crops.

Jatropha (*Jatropha curcas* L.) of the Euphorbiaceae family is a perennial poisonous shrub originated in Central America. Much of the interest in *J. curcas* has arisen due to its ability to grow on poor quality land. Using marginal land for *J. curcas* cultivation is therefore attractive since it would not displace food-producing crops. Current estimates suggest that there are now 2.5 million hectares of *J. curcas* planted in India and China alone (King *et al.*, 2009). Research about Jatropha biodiesel ranges from oil extraction to genomics analysis (Shah *et al.*, 2005; Sato *et al.*, 2011). The transcriptome, proteome and genome of *Jatropha* seeds have been analyzed to facilitate the understanding of the mechanisms of oil production (Yang *et al*., 2009; Natarajan *et al*., 2010; Jiang *et al*., 2012).

Vernicia (*Vernicia fordii* Hemsl.), belonging to the same family as *Jatropha*, is well known as Tung tree and native to Southern China with a subtropical climate. It is famous with the production of Tung oil which is believed to have originated in ancient China. The use of Tung oil for lamps and furniture in the 13th century appeared in Marco Polo's journey. Tung oil remains as a popular natural oil for finishing furniture, and for the waterproofing purpose of clothing and paper. Tung oil is also a valuable resource for biodiesel (Shang *et al.*, 2010). Despite of its importance, the study of the molecular mechanism of Tung oil synthesis is extremely limited. Little genome-wide studies had been reported for the Tung tree thus far.

The quality and performance of biodiesel depends on the chemical composition of the fatty acids present in the oil that biodiesel with high monounsaturated fatty acid content (oleate) has excellent characteristics with respect to ignition

To cite this paper: Yang, D., H. Zhang, K. Peng, L. Chen, H. He, X. Huang, J. Qin, G. He and D. Zhang, 2016. Differential gene regulation of lipid synthesis in the developing seeds of two biodiesel tree species, *Jatropha* and *Vernicia*. *Int. J. Agric. Biol*., 18: 1143–1152

quality, nitrogen oxides (NOx) emissions and fuel stability (Ramos *et al.*, 2009; Peng *et al.*, 2016). However, most plant oils used as biodiesel feedstock have a high level of polyunsaturated fatty acids (linoleate and linolenate acids) which impacts the biodiesel with poor cold-temperature performance and low oxidative stability. Improving the fuel characteristics of biodiesel can be achieved by altering the fatty acid composition, this has been a long-standing goal of academic researchers and the biotechnology industry (Durrett *et al.*, 2008; Graef *et al.*, 2009; Liu *et al.*, 2016).

The fatty acid compositions of Jatropha oil and Tung oil are strikingly different. The former is enriched in oleic acid (34.3–45.8%; 18:1), linoleic acid (29.0–44.2%; 18:2), palmitic acid (14.1–15.3%; 16:0) and stearic acid (3.7–9.8%; 18:0), while the latter contains approximately 80% eleostearic acid, an unusual conjugated fatty acid. Generally, woody oil plants store the lipid in the form of triacylglycerols (TAGs) in kernels, and their fatty acid synthesis pathways are similar (Ohlrogge and Browse, 1995). The first step involves the synthesis of fatty acids in plastids. The second step involves the modification of these fatty acids by enzymes located primarily in the endoplasmic reticulum (ER). The third step involves the packaging of the nascent fatty acids into TAGs, which subsequently accumulate in oil bodies that bud off from the ER.

Previous investigations have confirmed that, desaturation of oleic acid (18C:1) to linoleic acid (18C:2) or linolenate acid (18C:3) is catalyzed by fatty acid desaturase 2 (FAD2) and FAD3, respectively (Sperling and Heinz, 1993; Okuley *et al.*, 1994). Gene mutations of *FAD2* in peanuts (Wang *et al.*, 2011) and RNAi suppression of *FAD2-1* in soybean (Wagner *et al.*, 2011) showed significant variation in fatty acid composition. Meanwhile, overexpression of diacylglycerol acyltransferases (*DGATs*) has been shown to increase oil content in Arabidopsis (Jako *et al.*, 2001) and soybean (Lardizabal *et al.*, 2008). Interestingly, overexpression of soybean transcription factors *GmDOF4* and *GmDOF11* in transgenic Arabidopsis also resulted in increased oil content (Wang *et al.*, 2007). These findings suggest a common multi-level and multi-component controlled lipid synthesis in oil plants, and the differential lipid compositions in *Vernicia* and *Jatropha* of interest to us might have both genome and transcriptome origins.

Next-generation sequencing technology has provided unprecedented opportunities for efficient discovery of key genes and gene functions accounting for a biological trait, as well as deciphering novel mechanisms of gene/genome regulation, through sequencing and analysis of the cDNA libraries generated from a whole transcriptome or cDNA/gDNA libraries from specific populations of RNA/DNA of the experimental cells and tissues (Kahvejian *et al.*, 2008; Ball *et al.*, 2009; Xue *et al.*, 2009, 2013; Voineagu *et al.*, 2011; Xiao *et al.*, 2012). For the study of biological species without a reference genome, transcriptome sequencing offers an effective opportunity for simultaneous identification of genes encoded by a genome and their regulated expression.

In order to understand the causes of different fatty acid compositions of Jatropha oil and Tung oil at gene level, and improving their fuel characteristics, we used transcriptome sequencing strategy in seeds of *Jatropha* and *Vernicia*. We completely sequenced the polyadenylated mRNAs expressed in *Jatropha* and *Vernicia* seeds during the initial- and fast- phases of oil accumulation. We generated more than 24 billion bases of high-quality cDNA sequence and obtained 45,414 and 49,583 unigenes from the seed transcriptomes of *Jatropha* and *Vernicia*, respectively. The assembled, annotated transcriptome sequences and gene expression profiles provide useful information for the identification of genes involved in unsaturated fatty acid biosynthesis and regulation. Transcriptome analysis shows that the two oil trees express similar genes for oil synthesis, but differ significantly in their genes involving in lipid storage, highlighting the unrecognized contribution.

Materials and Methods

Sample Information

Seeds of three-years-old *Jatropha* or *Vernicia* trees were collected, with the former from a natural farmland in southern Guizhou province of China and the later from northwest Hunan province. Mature *Jatropha* female flowers were tagged and hand-pollinated during June 10–25, 2012 when their stigma became fully expanded, with the respective tagging dates assigned as 0 day after pollination (DAP). *Vernicia* seeds were similarly treated, with the sampling time set at since May 15, 2012. Seeds collected at various maturing periods were either dissected for seed weight measurement, lipid extraction, or frozen in liquid nitrogen and stored at -80°C for later RNA extraction. The seed oil content of developing seeds were extracted and measured as described by Hara and Radin (1978).

Total RNA Isolation and Library Preparation

The developing seeds of *Jatropha* and *Vernicia* at the initial stage (S1) and in the fast oil accumulation stage (S2) were picked for the transcriptome profiling experiments. Total RNA was extracted using CTAB method and treated with RQ1 DNase (promega) to remove DNA. The quality and quantity of the purified RNA were determined by measuring the absorbance at 260 nm/280 nm (A260/A280) using smartspec plus (Bio-Rad). RNA integrity was further verified by 1.5% Agrose gel electrophoresis.

For each sample, 10 μg of total RNA was used for RNA-seq library preparation. Polyadenylated mRNAs were purified and concentrated with oligo (dT)-conjugated magnetic beads (Invitrogen) before used for directional RNA-seq library preparation. Purified mRNAs were iron fragmented at 95°C followed by end repair and 5' adaptor

ligation. The reverse transcription was performed with RT primer harboring 3' adaptor sequence and randomized hexamer. The cDNAs were purified and amplified, and PCR products corresponding to 200–500 bps were purified, quantified and stored at -80°C until used for sequencing.

For high-throughput sequencing, the libraries were prepared following the manufacturer's instructions and applied to Illumina GAIIx system for 80nt single-end sequencing.

Analysis of Illumina Sequencing Results

Raw reads were first discarded if containing more than 2-N bases, then reads were processed by clipping adaptor and removing low quality bases, too short reads (less than 20nt) were also dropped. FASTX-Toolkit (Version 0.0.13) was used to get the clean reads (Blankenberg *et al.*, 2010).

Reads Assembly

Clean Reads from all samples were combined to perform the subsequent assembly, software Trinity was used to assemble these clean reads into unigenes with a minimum length of 200bp (Haas *et al.,* 2013).

Annotation and Functional Classification

Annotation of the assembled transcript sequences was performed using BLASTX algorithm and non-redundant protein database at NCBI and some other database, such as Nt, COG with an e-value cutoff of 1e-5 (Altschul *et al.,* 1990). The BlastX results were also used to assess the full-length nature of the contigs. The automated BlastX analysis was done using BLAST2GO to assign GO terms for the unigenes (Gotz *et al.*, 2008). The transcripts were classified under three GO terms such as molecular function, cellular process and biological process.

Clean Reads Alignment Statistics

Clean reads were aligned to the assembled unigenes by bowtie (Langmead *et al.*, 2009) with 2 mismatches. Based on gene annotations of the genome, aligned reads with more than one genome location were discarded as being ambiguous. Uniquely localized reads were used to calculate reads numbers and RPKM values (RPKM represents reads per kilo base and per million). Other statistical results, such as gene coverage and depth, reads distribution along unigenes, were also obtained.

Analysis of Differentially Expressed Genes

DEGs between the test sample and control sample were analyzed by using edgeR (Robinson *et al*., 2010). For each gene, the p-value was computed and the significance threshold to control FDR at a given value was calculated. The fold changes were also estimated within the edgeR statistical package. A gene with Pvalue lower than 0.01 or fold change over 2 is deemed as differentially expressed in our work.

Q-PCR Analysis

RT product (1:5 diluted, 2 μL) and SYBR Green I Master Mix (Roche) were used for qPCR experiments on an ABI7300 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). The specificity of PCR was checked by melting curve analysis. In every qPCR assay, GAPDH was used as the control for the significant bias of starting materials across samples.

Statistical Analysis and Figure Plot

Sample correlation analysis, cluster analysis and the figures of representation were obtained by edgeR software and MapMan (Thimm *et al.*, 2004).

Results

Similar Oil-accumulating Process in Developing Seeds of *Vernicia* and *Jatropha*

Seed development and oil accumulation in *Vernicia* seeds (Fig. 1A) seemed similar to a previous investigation on *Jatropha* seeds (Fig. 1B). Development from a pollinated *Vernicia* female flower to a mature seed took about 180 days, much longer than the reported 51 days for *Jatropha* seeds. Accordingly, four developmental stages from embryogenesis to seed dispersal (Fig. 1A) were defined for *Vernicia*. Seeds in the first stage (within 30 days after pollination, DAP) had a water content of ca. 90%. The second stage (30–90 DAP) was associated with a rapid increase in seed size, as well as a very low oil content and slightly changed water content. The third stage (90–150 DAP) came with very fast oil accumulation and significantly increased dry-weight of the kernel, while water content rapidly decreased. The last stage (after 150 DAP) seemed to be associated with nearly complete mature, and the accumulation of oil or dry material notably slowed down.

Transcriptome Landscape of the Developing Seeds from *Vernicia* and *Jatropha*

Vernicia and *Jatropha* seeds at two distinct developing stages (early oil accumulation stage S1, Vfo_90 DAP and Jcu_29 DAP; fast oil accumulation stage S2, Vfo_135 DAP and Jcu_37 DAP, Fig. 1C) were selected for further transcriptome investigation. To minimize the sampling error, for each selected representative, total RNAs from three seeds were independently extracted and then equally mixed for illumina library preparation.

Table 1: Assembly and annotation of transcriptome unigenes of *V. fordii* and *J. curcas* during seed development

Parameters	Vfo S1	Vfo S2	Jcu S1	Jcu S2
Total input reads (Cleaned)	54,297,859	60,500,795	56,929,935	51,003,303
Assembled reads	48,088,064	50,321,932	50,223,238	44,705,136
Transcripts	69,701		59,809	
Average transcript length (bp)	894.51		752.04	
Minimum transcript length (bp)	201		201	
Maximum transcript length (bp)	12,719		7,499	
N50	1,508		1,221	
Unigenes	49,583		45,414	
Total mapped reads	41,450,782	39,588,784	35,608,464	37,572,831
Unique mapped reads	41,064,099	38,096,184	33,291,521	36,963,391
Transcripts with Ath homologs	35,388		32,933	
Non-Ath transcripts with NR homologs	2,837		3,310	
Unigenes annotated by Ath	21,611		22,599	
Annotated Ath genes	13,333		13,320	
Unigenes with GO annotations	10,615		10,786	
Total GO terms	1,443		1,455	

Vfo: *Vernicia fordii*; Jcu: *Jatropha curcas*; Ath: *Arabidopsis thaliana*

As shown in Table 1, 50–60 million raw reads were obtained on a Hiseq2000 platform for these representatives, and removal of adaptor sequences and trimming bases yielded ~50 million high quality clean-reads. Trinity-derived transcriptome assembly assigned 69701 and 59809 transcripts for *Vernicia* and *Jatropha*, with their mean lengths as 895 and 752, respectively (NCBI's Sequence Read Archive (SRA) database GSE76386). These results implied a less complicate genome structure of *Vernicia*. Grouping highly similar transcripts into a single 'unigene' resulted in 49583 and 45414 unigenes, respectively defined for *Vernicia* and *Jatropha*, both with approximately 80% of their raw reads mappable in these defined unigenes. Finally, when annotated against *Arabidopsis* genes (TAIR10) with BlastX (e-value cut-off: 10^{-5}), 35,388 *Vernicia* transcripts and 32,933 *Jatropha* transcripts showed significant similarity with *Arabidopsis* genes, with 43.5% and 49.7% of *Vernicia* unigenes (21611 out of 49583) and *Jatropha* unigenes (22599 out of 45414), respectively annotated on 13333 and 13320 *Arabidopsis* genes.

Gene Ontology Annotation and Classification of Unigenes from *Vernicia and Jatropha*

The obtained *Vernicia* and *Jatropha* unigenes were translated into polypeptide sequence by ESTScan program and then analyzed by OrthoMCL program, with a Venn diagram obtained and shown in Fig. 2A. Subsequent WEGO analysis (Table 1 and Fig. 2B) showed that 11,148 and 11,293 out of annotated *Vernicia* unigenes (13333) and *Jatropha* unigenes (13320) could be categorized into 1443 and 1455 GO terms (p-value: 0.1), basically in concern with biological processes, cellular components and molecular functions.

GO analysis of the species-specific unigenes of the two (Fig. 2C) highlighted some non-orthologous gene displacements. For instance, some unigenes belonging to GO category 'antioxidant' (GO:0016209, GO:0004601)

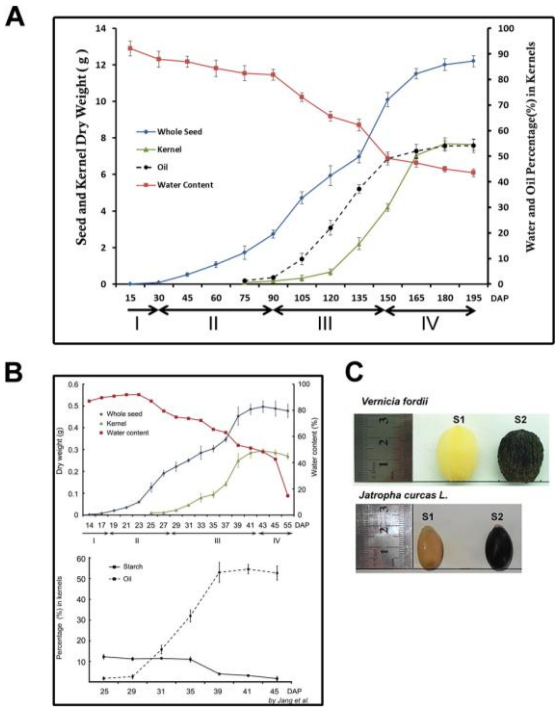

Fig. 1: Characterization of *Vernicia* (A) and *Jatropha* (B), seed development stages, appearance and size of *Vernicia* and *Jatropha* seeds at stage S1 and S2 (C)

and 'electron carrier' (GO:0009055) appeared only in *Vernicia*, which seemed very likely to provide a more reductive environment necessary for *Vernicia* seeds. A similar GO category 'nutrient reservoir' (GO: 0045735) was observed, which might be involved in the storage of some special nutrients in *Vernicia* seeds. Further GO analysis of lipid metabolism- and seed development-related unigenes (Fig. 2D) showed that 'Glycolipid transport' (GO:0046836) and 'steroid biosynthetic and metabolic process' (GO:0016126, GO:0016125) were

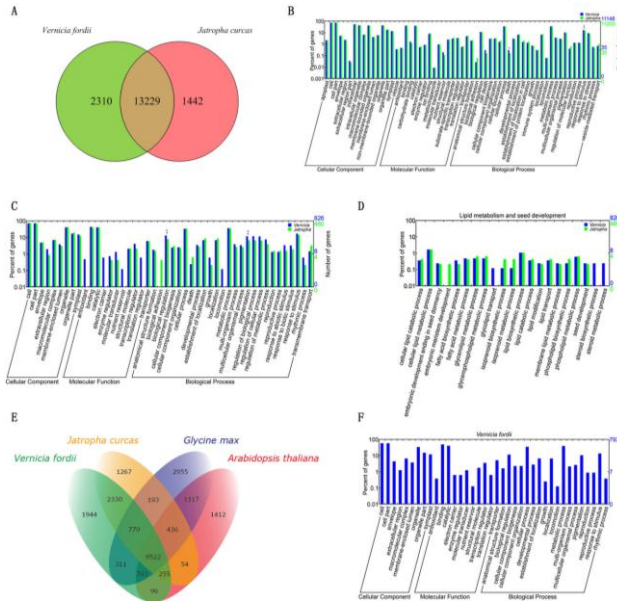

Fig. 2: Comparison and characterization of *Vernicia* and *Jatropha* unigenes found that they showed very similar transcriptome compositions but big differences in lipid storage and seed development. (A) Venn diagram showing the majority gene families were shared between *Vernicia* and *Jatropha*. (B) Gene Ontology classification of assembled unigenes also revealed similar transcriptome organization in terms of GO terms and their relative frequencies. The results are summarized in three main categories: Cellular component, Molecular function and Biological process. In total, 11,148 and 11,293 unigenes with BLAST matches to known proteins from *Vernicia* and *Jatropha* were assigned to gene ontology, respectively. (C) Gene Ontology classification of *Vernicia* and *Jatropha* specific unigenes. (D) Biological processes classification of *Vernicia* and *Jatropha* specific unigenes which were related to oil accumulation and seed development. (E) Comparison of the unigenes from Tung tree, *Jatropha* and soy bean and *Arabdopsis*. (F) Gene Ontology classification of *Vernicia* specific unigenes

solely detected in *Vernicia* seed. On the contrary, genes in the functional term of 'embryonic meristem development' (GO:0048508) were enriched only in *Jatropha*.

Venn diagram (Fig. 2E) of *Vernicia*, *Jatropha* as well as *Glycine max* and *Arabdopsis thaliana* (two model oil plants with their transcriptome data downloaded from Phytozome data base, http://www.phytozome.net/) showed that *Vernicia* was genetically close to *Jatropha*, with more unique unigenes enriched in *Vernicia* seeds (1944vs1267). All 1944 unique unigenes from *Vernicia* seeds was further analyzed by GO annotation and classification (Fig. 2F), and these unigenes sporadically distributed in diverse functional terms such as antioxidant, metabolic process, response to stimulus and biological regulation.

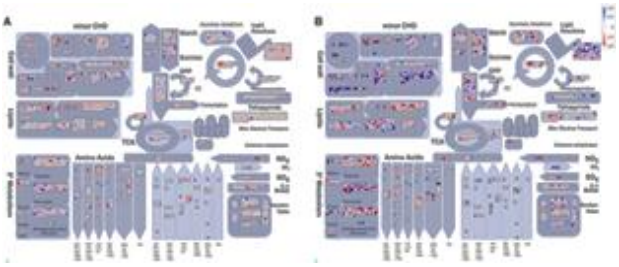

Fig. 3: MAPMAN diagrams showing differential expression patterns of metabolic pathway genes in stage S1 and S2. Ratio data were converted to a log base 2 scale and imported into MAPMAN. Blue boxes represent individual genes that are down-regulated in senescence, red boxes represent individual genes that are up-regulated. Intensity of the color indicates the relative level expression. (A) Expression patterns of metabolic pathway genes in *Vernicia*. (B) Expression patterns of metabolic pathway genes in *Jatropha*

General Gene Expression Profile of *Vernicia* and *Jatropha*

The expression levels of all identified *Vernicia* and *Jatropha* unigenes were analyzed by reads per kb per million reads (RPKM) method and edgeR program, with 8,590 up-regulated and 6,623 down-regulated unigenes determined for *Vernicia*, while 12,891 up-regulated and 10,127 down-regulated unigenes for *Jatropha*. Further analysis of these unigenes was carried out by MapMan software, and the results Additional analysis of the unigenes differentially expressed in the two developing stage of *Jatropha* and Vernicia was carried out by MapMan software, yielding similar results (Fig. 3). To validate the above expression profiles derived from RPKM analysis, a total of 20 unigenes were randomly selected for quantitative RT-PCR assays, and for all selected unigenes, RT-PCR results matched well with statistics from RPKM analysis (Supplementary Fig. S1).

K-means Clustering analysis was used to further group all unigenes into clusters based on their common expression patterns at the two stages. A total of 20 distinct clusters were obtained (Fig. 4), with genes in each cluster listed in Supplementary Dataset S1. Clusters 6, 9 and 12 contained unigenes up-regulated in *Vfo S2,* while unchanged in *Jatropha*. On the contrary, Clusters 2, 3, 10, 15 and 20 contained genes up-regulated in *Jcu_S2,* while unchanged in *Vernicia*. Interestingly, unigenes encoding 2 oleosins (comp26826_c0 and comp15571_c0) and 3 seed storage proteins (comp25375_c0, comp12120_c0 and comp15608_c0) in *Vernicia*, as well as unigenes encoding 1 oleosin (comp13078_c0) and 4 seed storage proteins (comp17433_c0, comp27581_c0, comp13072_c0 and comp13089_c0) in *Jatropha*, were identified with outstandingly high expression levels

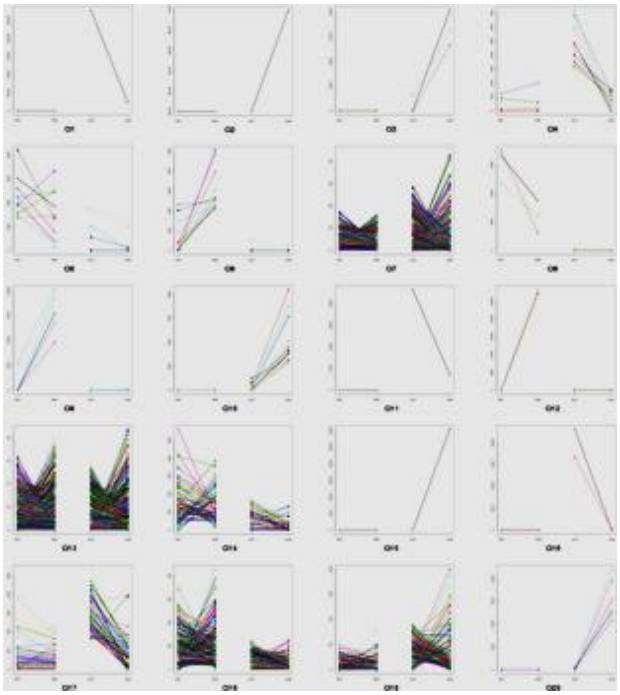

Fig. 4: Expression pattern clustering. Cluster 6, 9 and 12 members with expression pattern up at Vfo_S2 and no changes in *Jatropha*. These genes were storage proteins and specific extremely expressed in Vfo_S2. Clutster 2, 3, 10, 15 and 20 members with expression pattern up at Jca_S2 and no changes in *Vernicia*. These genes were seed development genes and specific extremely expressed in Jca_S2

(Supplementary Dataset S1). According to GO analysis, these unigenes are mainly involved in protein/lipid storage processes, which obviously count much at late maturation stage.

Transcript Analysis of Lipid Pathways Related Unigenes from *Vernicia* and *Jatropha*

For both *Vernicia* and *Jatropha*, over 100 unigenes (Supplementary Dataset S2) were identified by GO annotation as involved in lipid metabolism (Fig. 5A). Comparison of these enzymes and/or protein complexes transcript levels (Fig. 5B) and changes between stage S1 and S2 (Fig. 5C), while indicated similar expression patterns, also revealed some exceptions between *Vernicia* and *Jatropha* seeds.

For the fatty acid synthesis step, there were at least 14 homologous pairs of enzymes and/or protein complexes responsible for pyruvate to fatty acids conversion. Comparative analysis of their expression levels at stage S1 and S2 showed that in both cases, acyl carrier protein (ACP) and stearoyl-ACP desaturase (*SAD*) were obviously up regulated (less than 2 fold at most), while pyruvate dehydrogenase complex (*PDHC*) and acetyl-CoA

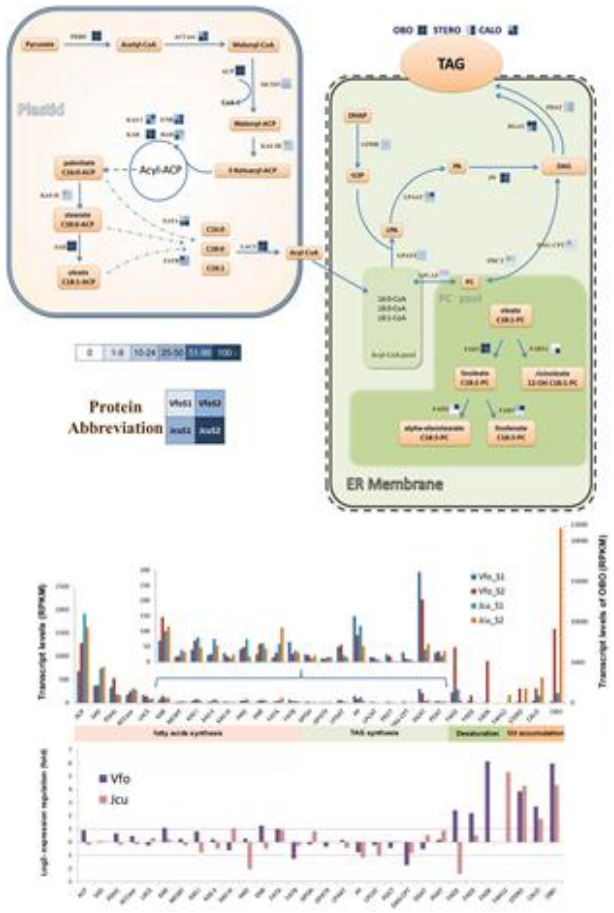

Fig. 5: The lipid-related genes were under transcriptional regulation in *Vernicia* and *Jatropha* during the seed maturation. (A) Expression patterns of lipid pathway genes in *Vernicia* and *Jatropha* during the seed maturation. The comparison of these enzymes and/or protein complexes transcript levels (B) and changes between S1 and S2 (C)

carboxylase (*ACCase*) involved in the initial synthesis exhibited much higher expression levels than those involved in elongation stage such as malonyl-CoA: ACP malonyl transferase (*MCMT*) and ketoacyl-ACP synthase (*KAS*). *FATA/FATB* thioesterases responsible for the production of 16- or 18-carbon fatty acid were differentially regulated, with *FATA* up-regulated in S2 and quite opposite for *FATB* in both *Vernicia* and *Jatropha*.

For TAG assembly (Kennedy pathway), 9 pairs of homologous enzymes were identified. Most of these enzymes encountered little change in their expression levels, except for phosphatidate phosphatase (*PP*) and diacylglycerol acyltransferase (*DGAT*). Twenty unigenes within nine PP isoforms (the largest gene families in TAG assembly pathway) were identified in our work, yet the average RPKM values of all these unigenes from both the two species remained less than 20, implying a relatively inactive involvement in the S1 to S2 switch.

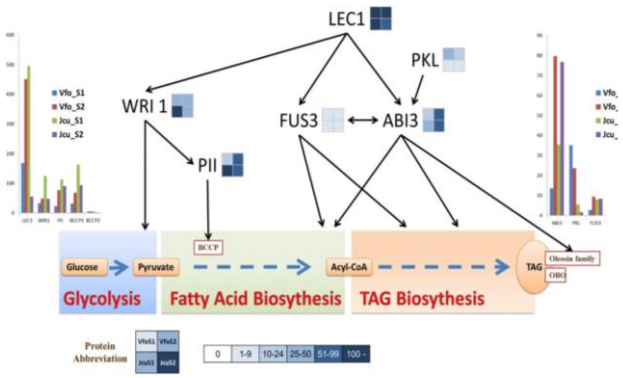

Fig. 6: The potential regulatory factors expression levels in stage S1 and S2 of *Vernicia* and *Jatropha*, respectively. Vfo: *Vernicia fordii*; Jca: *Jatropha curcas*

In comparison, enzymes in concern with fatty acid modification and oil accumulation encountered the most significant regulation, with oleate desaturase (FAD2), linoleate desaturase (FAD3), oleateΔ12-hydroxylase (FAH12, only expressed in *Jatropha*) and Δ12 fatty acid conjugase (FADX, only expressed in *Vernicia*) identified for this category. Most of these transcripts were over 4-fold up-regulated (in particular, over 70 times for *FADX* gene in *Vernicia*), except for *FAD3* in *Jatropha* (almost unchanged). Notably, *FAD2* from *Jatropha* was the only one down-regulated at S2 stage, which might be related to the preference over oleic acid in *Jatropha*. On the other hand, all fatty acid storage-related transcripts (Fig. 5C) were substantially up-regulated at S2 stage. In particular, oleosin genes showed the highest RPKM value among all transcripts in lipid metabolic pathway (near 10,000 in Vfo_S2 and over 20,000 in Jcu_S2), which might be related to their active involvement in the storage of certain kinds of lipids.

Besides, at least 8 regulatory factor genes playing important roles in lipid synthesis were identified, and their expression patterns looked quite different (Fig. 6). For instance, the expression of Leafy Cotyledon 1 (*LEC1*) was the highest in both *Vernicia* and *Jatropha*, while ABA insensitive 4 (*ABI4*), *LEC2* and *FUSCA 3* (*FUS3*) exhibited very low expression levels (Fig. 6 and Supplementary Dataset S2). The expression level of *PICKLE* (*PKL*) was down-regulated during S1 to S2 switch in both *Vernicia* and *Jatropha*, and consistently, ABI3 transcript under its inhibition showed significant up-regulation during this switch. Interestingly, *LEC1*, *PII* and *WRINKLED1* (*WRI1*) as a group, showed quite opposite trends in their expression level during S1 to S2 switch in *Vernicia* (coincidently up-regulated) and *Jatropha* (down-regulated).

Discussion

Regulation of lipid producing in oil trees is of great interest to researchers and in-depth recognition of their biosynthesis logic may hopefully help us developing genetically engineered bio-systems to solve future energy problems. Here we analyzed and compared the developing seeds of two *Euphorbiaceae* plants, *Vernicia* and *Jatropha*, with a focus on the correlation between their transcriptome profiles and final lipid compositions. Our results disclosed a number of desaturases, transferases, regulators and preservers important for lipid synthesis in these two plants and they are discussed below:

Vernicia and *Jatropha* seed oil dramatically differ in their fatty acid compositions, with α-eleostearic acid (C_{18}:3, 70%) being the overwhelming component for the former while oleic acid (C_{18}:1, 30%) and linoleic acid (C_{18}:2, 50%) for the latter. Previous work has proved successive fatty acid desaturations in various oil plants, including the stearic acid to oleic acid conversion catalyzed by SAD, subsequent linoleic acid formation by FAD2, as well as α-linoleic acid or α-eleostearic by FAD3 and FADX, respectively.

In *Jatropha* seeds, our results showed a steadily high expression of *SAD* for the two selected stages, while *FAD2* was significantly down-regulated from an initial high level at the fast accumulation stage. These findings perfectly match a previous study, in which oleic acid content increased since 29 DAP, while linoleic acid content increased constantly and then declined after 35 DAP. A later down-regulation of *FAD2* may well explain the relatively high content of oleic acid observed in *Jatropha* seeds. In *Vernicia* seeds, all the fatty acid desaturase genes (*FAD2*, *FADX* and *FAD3*) were up-regulated. A six-fold up-regulation was observed for *FADX* over *FAD3* during the fast accumulation stage, although both were at a high expression level. Such a bias might make α-eleostearic accumulation almost the only outcome and cause its overwhelming amount in mature *Vernicia* seeds.

In most oil plants, the final DAG to TAG conversion is catalyzed by Diacylglycerol acyltransferase (DGAT) or phospholipid:diacylglycerol acyltransferase (PDAT), with acyl-CoA or phospholipid as their respective acyl donors (Kennedy, 1961; Dahlqvist *et al.*, 2000). Consistently, 3 DGAT unigenes and 2 PDAT unigenes were identified from both *Vernicia* and *Jatropha* in our work. In *Vernicia* seeds, *DGAT1* unigene expression seemed to be at a low level for the two selected stages, while both *DGAT2* and *DGAT3* were maintained at a much higher level, with significant up-regulation (5.8 fold) observed for *DGAT2* and slight down regulation (2.7 fold) for *DGAT3* at the fast stage. Compared with *DGATs*, the 2 *PDAT* unigenes were all along maintained at a relatively low expression level. These findings suggested that in *Vernicia* seeds, TAG synthesis is mainly catalyzed by DGAT, with DGAT3 presumably responsible for an initial maturation, while DGAT2 counts more latter. In *Jatropha* seeds, on the contrary, no significant change was observed for the expression levels of all *DGAT* and *PDAT* unigenes. Meanwhile, expression levels of *PDAT* unigenes were slightly higher than those in *Vernicia* seed. These results indicated TAG synthesis in *Jatropha* seed

might be regulated coordinately by PDATs and DGATs.

Oleosins are vital in seed tissue for controlling oil body structure and lipid accumulation (Jolivet *et al.*, 2004; Siloto *et al.*, 2006). Inhibition of the major oleosin gene (18 kD) expression in *Arabidopsis* resulted in unusually larger oil bodies, disruption of storage organelles and significant decrease in the amount of lipids (Siloto *et al*., 2006). On the contrary, overexpression of soybean oleosin gene (24 kD) in transgenic rice seed significantly improved the lipid content with massive smaller oil bodies (Cao *et al.*, 2014). In our work, oleosins were identified from both *Vernicia* and *Jatropha*, each with 6 distinctive members. Oleosin unigenes (comp26826_c0 and comp15571_c0 in Vernicia, comp27587_c0 and comp13078_c0 in *Jatropha*) were extraordinarily up-regulated at the fast accumulation stage, with RPKM values over 10,000 for *Jatropha* type. Significant up-regulations of these genes, except for comp13078_c0 from *Jatropha*, were also confirmed by K-means Clustering method. These results implied directing roles of oleosins over TAGs with different fatty acid compositions, their preferences for specific unsaturated fatty acids (Liu *et al.*, 2011) might result in the selective accumulation of certain TAGs.

The LEC1 plays key roles in regulating embryo development. *LEC1* overexpression causes increased fatty acid content in *Arabidopsis* (Mu *et al.*, 2008), Maize (Shen *et al.*, 2010) or *Brassica napus* (Tan *et al.*, 2011). In *Arabidopsis*, overexpression of LEC1 resulted in over 58% genes of lipid pathway were up-regulated (Mu *et al.*, 2008). LEC1 acted as positive regulators upstream of WRI1, FUS3 and ABI3, which controlled the expression of genes involved in fatty acid and TAG synthesis (To *et al.*, 2006). WRI1 can regulate the steps transferring pyruvate into fatty acid synthesis (Shen *et al.*, 2010). PII can regulate ACCase activity by interacting with BCCP subunits of heteromeric HtACCase (Feria Bourrellier *et al*., 2010), which was regulated by WRI1 (Baud *et al.,* 2010). ABI3 and FUS3 were key regulatory factors involved in TAG synthesis (Yamamoto *et al.*, 2010). In the regulatory system "LEC1-WRI1-PII", *LEC1* had a significant down-regulation in fast oil accumulation stage of *Jatropha* seeds, and a series of downstream factors and genes (e.g., *WRI1*, *PII*, *BCCP*, *FAD2*) displayed the same trends, which was also observed in previous research (Jiang *et al.*, 2012). However, the contrary situation was observed in *Vernicia* seed, and both these genes were distinctly up-regulated. The oppositely regulated in these two seeds might origin from their different growth cycles as well as lipid compositions. For Vernicia seeds with an absolute majority of eleostearic acid content and a much longer maturation process, these genes and factors might have to remain active until a set maturity is achieved. In the regulatory system "LEC1-ABI3", up-regulation of ABI3 in both *Jatropha* and *Vernicia* at fast accumulation stage seemed very likely concerned with ABI3's crucial roles in TAG accumulation and oil body stability (Crowe *et al.*, 2000; Monke *et al.,* 2012). Meanwhile, *ABI3* was negatively regulated by PKL both in seeds of *Vernicia* and *Jatropha*, which accords with its *ABI3*-repressing activity (Perruc *et al.*, 2007). The expression change of PKL unigene was contrary to ABI3 both in seed of *Vernicia* and *Jatropha*. FUS3 might act as a negligible regulator here, though previous investigations on other plants proposed an inter-regulation with ABI3.

Conclusion

Among 49583 and 45414 unigenes respectively identified for *Vernicia* and *Jatropha*, 15,213 and 23,018 unigenes exhibited ≥2-fold expression changes for the two selected developing stages. Further analysis showed that the different fatty acid compositions and lipid contents observed in *Vernicia* and *Jatropha* seeds might origin from the differential expression of a series of key genes (such as FAD2, FADX, DGAT, oleosins etc.) and regulatory factors (such as LEC1, WRI1 and ABI3). These results would be helpful in understanding the regulation of oil accumulation in *Vernicia* and *Jatropha* seeds, further elucidations on molecular regulatory mechanisms of these genes, as well as transgenic oil-crops, may hopefully provide new bio-diesel resources in the future.

Acknowledgement

The research work was supported by grants from the National Natural Science Foundation of China (31170640, 31172257).

References

Altschul, S.F., W. Gish, W. Miller, E.W. Myers and D.J. Lipman, 1990. Basic local alignment search tool. *J. Mol. Biol.,* 215: 403–410

Ball, M.P., J.B. Li, Y. Gao, J.H. Lee, E.M. LeProust, I.H. Park, B. Xie, G.Q. Daley and G.M. Church, 2009. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat. Biotechnol.,* 27: 361–368

Baud, S., A.B. Feria Bourrellier, M. Azzopardi, A. Berger, J. Dechorgnat, F. Daniel-Vedele, L. Lepiniec, M. Miquel, C. Rochat, M. Hodges and S. Ferrario-Mery, 2010. PII is induced by WRINKLED1 and fine-tunes fatty acid composition in seeds of *Arabidopsis thaliana*. *Plant J.*, 64: 291–303

Blankenberg, D., A. Gordon, G. Von Kuster, N. Coraor, J. Taylor, A. Nekrutenko and T. Galaxy, 2010. Manipulation of FASTQ data with Galaxy. *Bioinformatics*, 26: 1783–1785

Cao, H., L. Zhang, X. Tan, H. Long and J.M. Shockey, 2014. Identification, classification and differential expression of oleosin genes in tung tree (*Vernicia fordii*). *PloS One*, 9: e88409

Crowe, A.J., M. Abenes, A. Plant and M.M. Moloney, 2000. The seed-specific transactivator, ABI3, induces oleosin gene expression. *Plant Sci.,* 151: 171–181

Dahlqvist, A., U. Stahl, M. Lenman, A. Banas, M. Lee, L. Sandager, H. Ronne and S. Stymne, 2000. Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. *Proc. Natl Acad. Sci. USA*, 97: 6487–6492

Durrett, T.P., C. Benning and J. Ohlrogge, 2008. Plant triacylglycerols as feedstocks for the production of biofuels. *Plant J.,* 54: 593–607

Fairless, D., 2007. Biofuel: the little shrub that could--maybe. *Nature,* 449: 652–655

Feria Bourrellier, A.B., B. Valot, A. Guillot, F. Ambard-Bretteville, J. Vidal and M. Hodges, 2010. Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate-regulated by interaction of PII with the biotin carboxyl carrier subunit. *Proc. Natl Acad. Sci. USA*, 107: 502–507

Gotz, S., J.M. Garcia-Gomez, J. Terol, T.D. Williams, S.H. Nagaraj, M.J. Nueda, M. Robles, M. Talon, J. Dopazo and A. Conesa, 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucl. Acids Res.*, 36: 3420–3435

Graef, G., B.J. LaVallee, P. Tenopir, M. Tat, B. Schweiger, A.J. Kinney, J.H. Van Gerpen and T.E. Clemente, 2009. A high-oleic-acid and low-palmitic-acid soybean: agronomic performance and evaluation as a feedstock for biodiesel. *Plant Biotechnol. J.,* 7: 411–421

Haas, B.J., A. Papanicolaou, M. Yassour, M. Grabherr, P.D. Blood, J. Bowden, M.B. Couger, D. Eccles, B. Li, M. Lieber, M.D. Macmanes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C.N. Dewey, R. Henschel, R.D. Leduc, N. Friedman and A. Regev, 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protocols*, 8: 1494–1512

Hara, A. and N.S. Radin, 1978. Lipid extraction of tissues with a low-toxicity solvent. *Anal. Biochem*., 90: 420–426

Jako, C., A. Kumar, Y. Wei, J. Zou, D.L. Barton, E.M. Giblin, P.S. Covello and D.C. Taylor, 2001. Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiol.,* 126: 861–874

Jiang, H., P. Wu, S. Zhang, C. Song, Y. Chen, M. Li, Y. Jia, X. Fang, F. Chen and G. Wu, 2012. Global analysis of gene expression profiles in developing physic nut (*Jatropha curcas* L.) seeds. *PloS One,* 7: e36522

Jolivet, P., E. Roux, S. D'Andrea, M. Davanture, L. Negroni, M. Zivy and T. Chardot, 2004. Protein composition of oil bodies in *Arabidopsis thaliana* ecotype WS. *Plant Physiol. Biochem.,* 42: 501–509

Kahvejian, A., J. Quackenbush and J.F. Thompson, 2008. What would you do if you could sequence everything? *Nat. Biotechnol.*, 26: 1125–1133

Kennedy, E.P., 1961. Biosynthesis of complex lipids. *Fed. Proceed.*, 20: 934–940

King, A.J., W. He, J.A. Cuevas, M. Freudenberger, D. Ramiaramanana and I.A. Graham, 2009. Potential of Jatropha curcas as a source of renewable oil and animal feed. *J. Exp. Bot*., 60: 2897–2905

Langmead, B., C. Trapnell, M. Pop and S.L. Salzberg, 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.,* 10: R25

Lardizabal, K., R. Effertz, C. Levering, J. Mai, M.C. Pedroso, T. Jury, E. Aasen, K. Gruys and K. Bennett, 2008. Expression of Umbelopsis ramanniana DGAT2A in seed increases oil in soybean. *Plant Physiol.*, 148: 89–96

Liu, P., C.M. Wang, L. Li, F. Sun, P. Liu and G.H. Yue, 2011. Mapping QTLs for oil traits and eQTLs for oleosin genes in jatropha. *BMC Plant Biol*., 11: 132

Liu, Z., Peng, M. Motahari-Nezhad, S. Shahraki and M. Beheshti, 2016. Circulating fluidized bed gasification of biomass for flexible end-use of syngas: a micro and nano scale study for production of bio-methanol. *J. Cleaner Production*, 129: 249–255

Monke, G., M. Seifert, J. Keilwagen, M. Mohr, I. Grosse, U. Hahnel, A. Junker, B. Weisshaar, U. Conrad, H. Baumlein and L. Altschmied, 2012. Toward the identification and regulation of the Arabidopsis thaliana ABI3 regulon. *Nucl. Acids Res*., 40: 8240–8254

Mu, J., H. Tan, Q. Zheng, F. Fu, Y. Liang, J. Zhang, X. Yang, T. Wang, K. Chong, X.J. Wang and J. Zuo, 2008. LEAFY COTYLEDON1 is a key regulator of fatty acid biosynthesis in *Arabidopsis*. *Plant Physiol.*, 148: 1042–1054

Natarajan, P., D. Kanagasabapathy, G. Gunadayalan, J. Panchalingam, N. Shree, P.A. Sugantham, K.K. Singh and P. Madasamy, 2010. Gene discovery from Jatropha curcas by sequencing of ESTs from normalized and full-length enriched cDNA library from developing seeds. *BMC Genomics*, 11: 606

Ohlrogge, J. and J. Browse, 1995. Lipid biosynthesis. *Plant Cell,* 7: 957–970

Okuley, J., J. Lightner, K. Feldmann, N. Yadav, E. Lark and J. Browse, 1994. Arabidopsis FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell*, 6: 147–158

Peng, W., Z. Lin, L. Wang, J. Chang, F. Gu and X. Zhu, 2016. Molecular characteristics of Illicium verum extractives to activate acquired immune response. *Saudi J. Biol. Sci.*, 23: 348–352

Perruc, E., N. Kinoshita and L. Lopez-Molina, 2007. The role of chromatin-remodeling factor PKL in balancing osmotic stress responses during Arabidopsis seed germination. *Plant J. Cell Mol. Biol.*, 52: 927–936

Ramos, M.J., C.M. Fernandez, A. Casas, L. Rodriguez and A. Perez, 2009. Influence of fatty acid composition of raw materials on biodiesel properties. *Bioresour. Technol.*, 100: 261–268

Robinson, M.D., D.J. McCarthy and G.K. Smyth, 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26: 139–140

Sato, S., H. Hirakawa, S. Isobe, E. Fukai, A. Watanabe, M. Kato, K. Kawashima, C. Minami, A. Muraki, N. Nakazaki, C. Takahashi, S. Nakayama, Y. Kishida, M. Kohara, M. Yamada, H. Tsuruoka, S. Sasamoto, S. Tabata, T. Aizu, A. Toyoda, T. Shin-i, Y. Minakuchi, Y. Kohara, A. Fujiyama, S. Tsuchimoto, S. Kajiyama, E. Makigano, N. Ohmido, N. Shibagaki, J.A. Cartagena, N. Wada, T. Kohinata, A. Atefeh, S. Yuasa, S. Matsunaga and K. Fukui, 2011. Sequence analysis of the genome of an oil-bearing tree, *Jatropha curcas* L. *DNA Res.,* 18: 65–76

Shah, S., A. Sharma and M.N. Gupta, 2005. Extraction of oil from Jatropha curcas L. seed kernels by combination of ultrasonication and aqueous enzymatic oil extraction. *Bioresour. Technol.,* 96: 121–123

Shang, Q., W. Jiang, H. Lu and B. Liang, 2010. Properties of Tung oil biodiesel and its blends with 0# diesel. *Bioresour. Technol.,* 101: 826–828

Shen, B., W.B. Allen, P. Zheng, C. Li, K. Glassman, J. Ranch, D. Nubel and M.C. Tarczynski, 2010. Expression of ZmLEC1 and ZmWRI1 increases seed oil production in maize. *Plant Physiol.*, 153: 980–987

Siloto, R.M., K. Findlay, A. Lopez-Villalobos, E.C. Yeung, C.L. Nykiforuk and M.M. Moloney, 2006. The accumulation of oleosins determines the size of seed oilbodies in Arabidopsis. *The Plant Cell*, 18: 1961–1974

Sperling, P. and E. Heinz, 1993. Isomeric sn-1-octadecenyl and sn-2-octadecenyl analogues of lysophosphatidylcholine as substrates for acylation and desaturation by plant microsomal membranes. *Eur. J. Biochem./FEBS*, 213: 965–971

Tan, H., X. Yang, F. Zhang, X. Zheng, C. Qu, J. Mu, F. Fu, J. Li, R. Guan, H. Zhang, G. Wang and J. Zuo, 2011. Enhanced seed oil production in canola by conditional expression of Brassica napus LEAFY COTYLEDON1 and LEC1-LIKE in developing seeds. *Plant Physiol.*, 156: 1577–1588

Thimm, O., O. Blasing, Y. Gibon, A. Nagel, S. Meyer, P. Kruger, J. Selbig, L.A. Muller, S.Y. Rhee and M. Stitt, 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant J. Cell Mol. Biol.*, 37: 914–939

To, A., C. Valon, G. Savino, J. Guilleminot, M. Devic, J. Giraudat and F. Parcy, 2006. A network of local and redundant gene regulation governs Arabidopsis seed maturation. *The Plant Cell*, 18: 1642–1651

Voineagu, I., X. Wang, P. Johnston, J.K. Lowe, Y. Tian, S. Horvath, J. Mill, R.M. Cantor, B.J. Blencowe and D.H. Geschwind, 2011. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature*, 474: 380–384

Wagner, N., A. Mroczka, P.D. Roberts, W. Schreckengost and T. Voelker, 2011. RNAi trigger fragment truncation attenuates soybean FAD2-1 transcript suppression and yields intermediate oil phenotypes. *Plant Biotechnol. J.*, 9: 723–728

Wang, H.W., B. Zhang, Y.J. Hao, J. Huang, A.G. Tian, Y. Liao, J.S. Zhang and S.Y. Chen, 2007. The soybean Dof-type transcription factor genes, GmDof4 and GmDof11, enhance lipid content in the seeds of transgenic Arabidopsis plants. *Plant J. Cell Mol. Biol.,* 52: 716–729

Wang, M.L., N.A. Barkley, Z. Chen and R.N. Pittman, 2011. FAD2 gene mutations significantly alter fatty acid profiles in cultivated peanuts (*Arachis hypogaea*). *Biochem. Genet.*, 49: 748–759

Xiao, R., P. Tang, B. Yang, J. Huang, Y. Zhou, C. Shao, H. Li, H. Sun, Y. Zhang and X.D. Fu, 2012. Nuclear matrix factor hnRNP U/SAF-A

exerts a global control of alternative splicing by regulating U2 snRNP maturation. *Mol. Cell*, 45: 656–668

Xue, Y., K. Ouyang, J. Huang, Y. Zhou, H. Ouyang, H. Li, G. Wang, Q. Wu, C. Wei, Y. Bi, L. Jiang, Z. Cai, H. Sun, K. Zhang, Y. Zhang, J. Chen and X.D. Fu, 2013. Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. *Cell*, 152: 82–96

Xue, Y., Y. Zhou, T. Wu, T. Zhu, X. Ji, Y.S. Kwon, C. Zhang, G. Yeo, D.L. Black, H. Sun, X.D. Fu and Y. Zhang, 2009. Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Mol. Cell*, 36: 996–1006

Yamamoto, A., Y. Kagaya, H. Usui, T. Hobo, S Takeda and T. Hattori, 2010. Diverse roles and mechanisms of gene regulation by the Arabidopsis seed maturation master regulator FUS3 revealed by microarray analysis. *Plant Cell Physiol.*, 51: 2031–2046

Yang, M.F., Y.J. Liu, Y. Liu, H. Chen, F. Chen and S.H. Shen, 2009. Proteomic analysis of oil mobilization in seed germination and postgermination development of *Jatropha curcas*. *J. Proteome Res.,* 8: 1441–1451

(Received 13 July 2016; Accepted 08 August 2016)