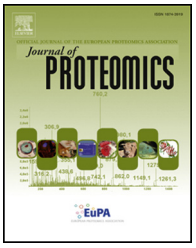
JOURNAL OF PROTEOMICS XX (2013) XXX–XXX

Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/jprot



Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation☆

Hui Liu[a,b], *Cuiping Wang*[a,b], *Setsuko Komatsu*[c], *Mingxia He*[d], *Gongshe Liu*[a,*], *Shihua Shen*[a,*]

[a]*Key Laboratory of Research and Development for Resource Plant, Institute of Botany, the Chinese Academy of Sciences, Beijing 100093, China*
[b]*University of Chinese Academy of Sciences, Beijing 100049, China*
[c]*National Institute of Crop Science, Tsukuba 305-8518, Japan*
[d]*Kunming Institute of Botany, the Chinese Academy of Sciences, Kunming 650201, China*

ARTICLE INFO

Article history:
Received 1 February 2013
Accepted 25 June 2013

Keywords:
Jatropha curcas
Proteomics
Seed development
Carbon flux
Lipid accumulation

ABSTRACT

To characterize the metabolic signatures of lipid accumulation in *Jatropha curcas* seeds, comparative proteomic technique was employed to profile protein changes during the seed development. Temporal changes in comparative proteome were examined using gels-based proteomic technique at six developmental stages for lipid accumulation. And 104 differentially expressed proteins were identified by MALDI-TOF/TOF tandem mass spectrometry. These protein species were classified into 10 functional categories, and the results demonstrated that protein species related to energy and metabolism were notably accumulated and involved in the carbon flux to lipid accumulation that occurs primarily from early to late stage in seed development. Glycolysis and oxidative pentose phosphate pathways were the major pathways of producing carbon flux, and the glucose-6-phosphate and triose-phosphate are the major carbon source for fatty acid synthesis. Lipid analysis revealed that fatty acid accumulation initiated 25 days after flowering at the late stage of seed development of *J. curcas*. Furthermore, $C_{16:0}$ was initially synthesized as the precursor for the elongation to $C_{18:1}$ and $C_{18:2}$ in the developing seeds of *J. curcas*. Together, the metabolic signatures on protein changes in seed development provide profound knowledge and perspective insights into understanding lipid network in *J. curcas*.

Biological significance

Due to the abundant oil content in seeds, *Jatropha curcas* seeds are being considered as the ideal materials for biodiesel. Although several studies had carried out the transcriptomic project to study the genes expression profiles in seed development of *J. curcas*, these ESTs hadn't been confirmed by qRT-PCR. Yet, the seed development of *J. curcas* had been described for a pool of developing seeds instead of being characterized systematically. Moreover, cellular metabolic events are also controlled by protein–protein interactions, posttranslational protein modifications, and enzymatic activities which cannot be described by transcriptional profiling approaches alone. In this study, within the overall objective of profiling differential protein expression in developing *J. curcas* seeds, we provide

* *Corresponding authors*. Tel.: +86 10 62836545.
E-mail address: shshen@ibcas.ac.cn (S. Shen).

1874-3919/$ – see front matter
http://dx.doi.org/10.1016/j.jprot.2013.06.030

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

a setting of physiological data with dynamic proteomic and qRT-PCR analysis to characterize the metabolic pathways and the relationship between mRNA and protein patterns from early stage to seed filling during the seed development of *J. curcas*. The construction of *J. curcas* seed development proteome profiles will significantly increase our understanding of the process of seed development and provide a foundation to examine the dynamic changes of the metabolic network during seed development process and certainly suggest some clues to improve the lipid content of *J. curcas* seeds.

1. Introduction

The seed development of higher plants generally proceeds through three different phases [1]. During this process, the embryo is formed as a result of the fusion of a sperm and an egg, while the endosperm is considered as part of the female gametophyte [2,3]. Starches, lipids and proteins were mainly biosynthesized in plastids for seed development [4,5]. For oil plants such as rapeseed (*Brassica napus*), lipids and proteins are the major storage products [6,7]. In many angiosperm seeds, lipids as an important form of carbon storage usually consist of triacylglycerols (TAGs) which mainly accumulate during the maturation phase of embryo and/or endosperm [6]. Studies show that seed filling processes are highly complex, as many genes are involved in a number of pathways and regulated precisely in each storage component [8]. It is important to understand the synchronized mechanisms responsible for lipid synthesis due to the nutritional and economical value of these storage components. However, the conventional techniques of biochemistry and molecular biology by isolating enzymes in pathways largely limit us to obtain the comprehensive information about seed development. Therefore, it is necessary to study the association of metabolic networks and seed development on a broad scope with application of advanced approaches. Proteomic analysis has been reported to become a powerful tool to profile the biological processes of plants [9,10]. This approach has been widely applied to study the protein changes during seed development for a variety of plants including barley (*Hordeum vulgare*), *Medicago truncatula*, soybean (*Glycine max*), wheat (*Triticum aestivum*), rapeseed, *Lotus japonicus* and castor (*Ricinus communis*) [8,11–16].

Jatropha curcas L. is a tropical or subtropical shrub which belongs to the Euphorbiaceae family [17]. A mature seed of *J. curcas* typically has a tiny embryo embedded in a thick endosperm which accounts for more than 90% of the total kernel weight [18]. Due to its unique biological characteristics of eximious growth and high yield seed, *J. curcas* has been indicated as a potential energy plant for biodiesel [17,19–23]. The lipid in *J. curcas* seed is mainly composed of unsaturated fatty acids including linoleic or oleic acid. *J. curcas* oil is more suitable for fuel purpose as compared with other vegetable oils because of its production and high fuel rate [24]. Another advantage of *J. curcas* oil as biodiesel is its great oxidation stability. For example, given the uncontrolled temperature and oxidation properties, biodiesel from Jatropha and Palm can be blended together to achieve an optimum mix for usage in Asia [25].

J. curcas was mostly reported for its economic value for biodiesel, but only a few studies focus on its lipid characteristics from the molecular perspective. Annarao et al. studied oil content and lipid profile in the seed development of *J. curcas*, and their results showed that lipid synthesis was detectable at nearly the third week after fertilization (WAF). The seeds actively synthesized TAG from the fourth WAF [26]. Additional studies were performed at mRNA level in *J. curcas*, as four cDNA clones encoding β-ketoacyl-acyl carrier protein (ACP) synthase III (KAS III), ACP-thioesterase (FATA/FATB), stearoyl-ACP desaturase (SAD) and acetyl-CoA carboxylase (ACCase) have been isolated. Quantitative Real-Time PCR (qRT-PCR) analysis revealed that all of these four genes were highly expressed in the late developing phase of seeds [27–30]. Yang et al. adopted ultrastructural observation and proteomic analysis of endosperm to illuminate oil mobilization in germinating seeds of *J. curcas*. Fifty protein species display remarkable changes in abundance, which are involved in β-oxidation, glyoxylate cycle, glycolysis, tricarboxylic acid (TCA) cycle and oxidative pentose phosphate pathways (OPPP) in the process of oil mobilization [31].

Due to tremendous advances in high-throughput technologies, a great body of information on *J. curcas* has been achieved with the application of genomic and transcriptomic sequencing. Costa et al. and Natarajan et al. employed the transcriptome approach to unveil the genes contributed to lipid accumulation in seed development. Except *hydroxyacyl-ACP dehydrase* (*HAD*), a considerable number of expressed sequence tags (ESTs) coding for most enzymes have been identified, and 28,794 non-redundant transcripts sequences are available from *J. curcas* at present [32–34]. Sato et al. sequenced the whole genome by using a combination of the conventional Sanger method and multiplex sequencing methods. Nineteen classes of Jatropha genes including 73 genes involved in the synthesis of TAGs were listed in the supplement [35]. These transcriptomic and genomic data are valuable for better understanding the lipid accumulation in the seed development of *J. curcas*. Although several studies had carried out the transcriptomic project to study the genes expression profiles in development of *J. curcas* [32–34], these ESTs hadn't been confirmed by qRT-PCR. Yet, the seed development of *J. curcas* had been described for a pool of developing seeds instead of being characterized systematically. Moreover, cellular metabolic events are also controlled by protein–protein interactions, posttranslational protein modifications, and enzymatic activities which cannot be described by transcriptional profiling approaches alone. The proteomes of *J. curcas* seeds have not yet been explored during seed development. In this study, within the overall objective of profiling differential protein expression in developing *J. curcas* seeds, we provide a setting of physiological data with dynamic proteomic and qRT-PCR analysis to characterize the metabolic

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

pathways and the relationship between mRNA and protein patterns from early stage to seed filling during the seed development of *J. curcas*. The construction of *J. curcas* seed development proteome profiles will significantly increase our understanding of the process of seed development and provide a foundation to examine the dynamic changes of the metabolic network during seed development process.

2. Materials and methods

2.1. Plant materials

Developing seeds of *J. curcas* were harvested randomly from 5 plants starting at 5 days after flowering (DAF) with 5-days intervals until 45 DAF in Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Yunnan Province (Supplemental Fig. S1). Three independent biological replicates of developing seeds at each stage were collected and divided into 2 parts, one part was prepared for paraffin fixation, the other part was snap-frozen immediately in liquid nitrogen and then stored at −80 °C for further use.

Fresh and dry weights of seeds were determined in 10 replicates randomly from three independent biological replicates for each stage. The fresh weight of each seed was assessed immediately after the shell was removed, and the dry weight was assessed after parch-dried for 48 h at 80 °C. Water content of seeds was calculated by subtracting the dry weight from the fresh weight.

2.2. Fixation, section preparation and staining

Fresh seeds in developing phases at 5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF and 30 DAF were immersed in fixative solution containing 0.3% chromic acid, 2% acetic acid and 10% formalin at room temperature. The seeds were fixed first by infiltrating under vacuum to remove air and then dehydrated in a series of ethanol solution (70%, 80%, 90% and 100%). After that, the seeds were embedded in Technovit 7100 (Haereus Kluver, Werheim, Germany) following the manufacturer's instructions. Thin sections in 12 μm thickness were obtained by employing Reichert Histostat 820 (AOReichert Scientific, Buffalo, NY, USA) and counterstained with Safranin-fast green dye. Tissue slides were subjected to observation under light microscopy.

2.3. Analysis of fatty acid composition

The parch-dried seeds in developing phases at 25 DAF, 30 DAF, DAF 35, 40 DAF and 45 DAF were ground into powder with pestle and mortar. The content of crude reserve proteins was calculated by following the instructions of the National food safety standard determination of protein in foods (GB/T 17377-2008), while the crude reserve lipid content and the methyl ester of the FAs in lipids were analyzed as described by Yang et al. [31].

2.4. Protein extraction

The protein samples of developing *J. curcas* seeds at 5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF and 30 DAF were chosen for proteomic analysis. Water soluble aqualous proteins of the developing seeds of *J. curcas* at six developmental stages were extracted using a modified method according to Shen et al. [36]. Briefly, for each developmental stage, 500–1000 mg fresh seeds were homogenized in 2 mL of the homogenization buffer containing 20 mM Tris–HCl (pH 7.5), 250 mM sucrose, 10 mM ethylene glycol tetraacetic acid, 1 mM phenylmethylsulfonyl fluoride, 5% 2-mercaptoethanol and 1% Triton X-100. The homogenate was collected into an Eppendorf tube and centrifuged at 10,000 *g* for 10 min at 4 °C. The supernatant was transferred to a fresh tube and precipitated by adding 10% cold trichloroacetic acid on ice for more than 30 min. The mixture was centrifuged at 15,000 *g* for 10 min at 4 °C, and the supernatant was discarded. After washed three times with acetone, the pellet was collected by centrifugation, air-dried and then suspended in sample buffer containing 7 M urea, 2 M thiourea, 4% 3[(cholamidopropyl) demethylammonio] -1-propane sulphonate, 2% pharmalyte, pH 3.5–10 (GE Healthcare Bio-Sciences, Little Chalf-ont, U.K.), and 1% dithiothreitol (DTT). Concentrations of protein samples for proteomic experiment were quantified according to Bradford method [37]. Albumin (A5503, Sigma) was used as a standard for protein quantification. Three independent biological and replicates were performed independently for each developmental stage of *J. curcas* seeds.

2.5. Two-dimensional electrophoresis

Two-dimensional electrophoresis (2-DE) was performed as previously described by Liu et al. [38] with minor modifications. Isoelectric focusing (IEF) was performed using immobilized pH gradient (IPG) strips, linear pH gradient 4–7, and length at 11 cm (GE Healthcare Bio-Science). Protein samples (about 500 μg) were loaded at the rehydration step. After 12 h passive rehydration, IEF was carried out on a Multiphor II electrophoresis device (GE healthcare Bio-Sciences) at 20 °C, in a setting of running parameters: 1 h at 300 V, 1 h at 600 V, 1 h at 1000 V, 1 h at 8000 V, finally followed by 32,000 Vh, all at 50 μA per strip. After IEF was finished, IPG strips were saturated with the equilibration buffer containing 6 M urea, 30% (v/v) glycerol, 2.5% (w/v) sodium dodecyl sulfate (SDS), 1% DTT, 50 mM Tris–HCl, pH 6.8 for 15 min at room temperature. Additional equilibration step was followed for 15 min in the same buffer except DTT was replaced with 2.5% iodoacetamide. The strips were sealed on the top of the 1 mm vertical second-dimension gel with 1% agarose in distilled water. SDS-polyacrylamide electrophoresis (SDS-PAGE) was carried out in the gradient 4% and 15% polyacrylamide gel with the running buffer contained 0.3% Tris, 1.44% glycine and 0.1% SDS. The electrophoresis was run in 25 mA and terminated when the bromophenol blue reached the bottom of the gel. The gels were detached and stained with Coomassie Brilliant Blue (CBB) R-250. Three independent technical replicates were performed at least for each developmental stage of seeds (5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF and 30 DAF), and a total of 54 CBB-stained 2-DE gel images were obtained.

2.6. Imaging and data statistical analysis

The 2-DE profiles of each developmental stage (5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF and 30 DAF) of *J. curcas* seeds were

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

acquired by scanning the 2-DE gels with ImageScanner III (GE Healthcare Bio-Sciences). The image analysis was carried out by ImageMatser 2D-platinum v5.0 software (GE Healthcare Bio-Sciences). Protein spots were detected with a setting of following parameters: Smooth, 2; Min Area, 10 and Saliency, 5.0. The spot volumes were normalized as the proportion of the sum of total spots per gel. And a match analysis was performed in automatic mode, followed by manual editing to correct the mismatched and unmatched spots. After image analysis, normalized spot volumes were obtained for each gel, standard deviation (SD) and coefficient of variance (CV) were calculated. Because the high rate of missing spot values in the nature of proteomic experimental conditions and the known dependence between the volume and variance for individual spots, the structure of raw proteomic data is known to be extremely disturbing for multivariate statistics and clustering algorithms [39–41]. To solve this problem, variable spots were only considered if they were either consistent or absent in at least one stage as Sghaier-Hammami et al. described [42]. The "0" value means just below the level of detection in this study.

The 2-DE gels of developing *J. curcas* seeds at six stages were compared with each other. Differential protein spots were defined and chosen for MS analysis after applying the one-way ANOVA and Student's t-tests ($p < 0.05$) by employing the SPSS software. A multivariate analysis was performed over the whole set of spots and on those showing differences. Principal component analysis (PCA) was applied to the correlation matrix to reduce its dimensionality. Using un-rotated principal component (PC) scores, the relation between the different experimental stages was studied by determining the spots with the highest load on the variance. All statically analysis was performed under R environment (Version 3.0.1, Windows 32-bit). Samples were clustered employing Cluster 3.0 using the Correlation method over an average linkage dissimilarity matrix and plotted with Java Treeview 1.1.6 software.

2.7. *In-gel digestion and protein identification*

Protein spots with significant changes in the seed development of *J. curcas* were digested as previously described [31,36]. Briefly, protein spots were excised from the 2-DE gels and then destained with 100 mM NH_4HCO_3 in 30% acetonitrile (ACN) for 1 h at 40 °C. After removing the destaining buffer, the gel pieces were lyophilized and digested overnight with 30 ng trypsin at 37 °C. After that, peptides were extracted with 0.1% trifluoroacetic acid (TFA) in 60% ACN three times. Peptide extracts were pooled and lyophilized for the two types of MALDI-TOF/TOF tandem mass spectrometric analysis. A gel piece free of proteins was treated as above and used for a control to identify autoproteolysis products by trypsin.

2.7.1. *Protein identification with ultrafleXtreme™ mass spectrometry*

One μL of pooled extracts was spotted onto the AnchorChip™ MALDI target plate (Bruker Daltonics, Billerica, MA, USA). After the solution dried, one μL of matrix solution (1 mg/mL, a-cyano-4-hydroxycinnamic acid in 70% ACN containing 0.1% TFA) was applied to the same location. MS and MS/MS spectra were acquired using the ultrafleXtreme™ MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) operating in the positive reflection mode and externally calibrated using the peptide calibration kit (Bruker Daltonics). The mass accuracy and mass resolution of mass measurement were set as the default. The time-of-flight spectra were acquired with 400 laser shots per spectrum and recorded with a mass range from 700 to 4000 m/z. MS/MS spectra were obtained using 1500 laser shots per fragmentation spectrum. 15 strongest peaks of each MS spectra were selected as precursor ions for MS/MS fragmentation spectra acquisition, excluding trypsin autolytic peptides and other known background ions. In MS–MS positive ion mode, spectra were averaged, and valid peaks were selected when the ratio of signal-to-noise was >5.

For database mining, MS data were uploaded with Biotools software (Ver. 3.2 Bruker Daltonics) to Mascot for database searching on the Matrix Science (London, U.K.) public web site (http://www.matrixscience.com) and searched against NCBInr protein databases (version 20120107; 16831682 sequences and 5781564572 residues). Search parameters were set as trypsin cleavage, one missed cleavage allowed, carbamidomethylation set as fixed modification, oxidation of methionines allowed as variable modification, peptide mass tolerance set to 100 ppm, fragment tolerance set to ±0.5 Da. Only significant hits, as defined by Mascot probability analysis, were considered in this study.

2.7.2. *Protein identification with ABI 4800 Proteomics Analyzer*

MS and MS/MS spectra were obtained using the ABI 4800 Proteomics Analyzer MALDI-TOF/TOF (Applied Biosystems, Foster City, CA, USA) operating in a result-dependent acquisition mode. Peptide mass maps were acquired in positive ion reflector mode (20 kV accelerating voltage) with 1000 laser shots per spectrum. Monoisotopic peak masses were automatically determined within the mass range 800–4000 Da with a signal to noise ratio minimum set to 10 and a local noise window width of m/z 250. Up to five of the most intense ions with minimum signal to noise ratio: 50 were selected as precursors for MS/MS acquisition, excluding common trypsin autolysis peaks and matrix ion signals. In MS/MS-positive ion mode, spectra were averaged, collision energy was 2 kV, and default calibration was set. Monoisotopic peak masses were automatically determined with a signal to noise ratio minimum set to 5 and a local noise window width of m/z 250. The MS together with MS/MS spectra were searched against the UniprotKB/SwissProt database (v. 2009.03.03, release number:14.9/56.9) using the software GPS Explorer, version 3.6 (Applied Biosystems) and MASCOT version 2.1 with the parameter settings as described for ultrafleXtreme™.

2.8. *qRT-PCR analysis*

Similar to proteomic analysis, developing seeds at 5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF and 30 DAF were used for qRT-PCR analysis in the seed development of *J. curcas*. Total RNA was extracted from seeds in different developmental stages (5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF and 30 DAF) using the method as Logemann et al. described [43]. The first-strand cDNA was synthesized with Reverse Transcriptase M-MLV (TaKaRa) by following the manufacturer's instruction. qRT-PCR operation was carried out in an Mx3000P™

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

Real-Time PCR system (Stratagene, La Jolla, CA, USA). In order to validate the primers for the Actin and other listed genes, different cDNA fragments of seeds at each developmental stage were amplified respectively with the specific primers listed in Table 1, and PCR products were verified by agarose gel electrophoresis. The Actin gene in *J. curcas*, amplified the 145 bp product with Actin primers, was used to normalize the cDNA amounts. Each qRT-PCR was performed in a 10 μL reaction mix containing 1 μL of template cDNA, 5 μL of SYBR Green Realtime PCR Master Mix (TOYOBO), 1 μL of each primer and 2 μL double distilled water. Thermal cycling programs were set as followed: 95 °C for 1 min; 40 cycles of 95 °C for 15 s, 55 °C for 15 s, 72 °C for 15 s; and then 95 °C for 1 min, 55 °C for 30 s and 95 °C for 30 s for the dissociation stage. After cycles were completed, the success of qRT-PCR was determined by analysis of the melting curve and the Ct value. The relative expression ratios of cDNA were calculated based on qRT-PCR efficiencies and the Ct values of each sample versus the reference sample. The Ct is defined as the cycle at a statistically significant increase in fluorescence above the background. Individual cDNA levels were qualified by normalizing with actin in the same sample, and the final relative cDNA amounts of different genes represented the means of three replicates.

Table 1 – Primer sequences used in the experiments.

Primer		Sequence(5′-3′)	Product length (bp)
GPDH-P[1]	F	TTGCAGACCCTGATCTTGAT	126
	R	TAGCCTCTACACCTTCCTTTATCT	
PEPCase	F	CGAATCAAAGCAGAGGGTAT	105
	R	GAGCGGTAGATAGTGGAAGTG	
PD-E1-M[2]	F	TTGCTGTATCTGAAGCCGTTA	112
	R	AAGTATCTGCCCAACACCAT	
RuBisCo-SSU	F	GGCTAAGGAAGTTGAATACC	125
	R	TCCAGTAGCGTCCATCGTAG	
PD-E2-P[1]	F	GTTGGAGCATCCCAGCCTAC	109
	R	CAGCACCATAGATTACACGA	
ACCase	F	GAACAAACTTCTTGGTCGTG	149
	R	GAGGGATACAACTTAGCCAG	
KASI	F	GTCCCTTCCACAAGAAATCC	105
	R	GCTTTGGGGCTGTAACAGTA	
KASII	F	GCTACGAAGCAAAGGCGAGTG	149
	R	CTCGTTGGAAATTGGGCACA	
KASIII	F	GCAGGTTTTGGAGCTGGTCT	136
	R	AAAGAAAGGGAGGGCAAGAAT	
KAR-P[1]	F	TTTGACAGGGTCTTGGTTGG	121
	R	TCCAGGTAATTGGCCACGAT	
FATB	F	GCTGCTACTTCCTCGTTCTT	103
	R	GTTTTGATTTGATTCCTCCC	
FATA	F	GCTACCGTTGAGACTATCGC	123
	R	CCATATGAGATGCAGCTTCC	
SAD	F	GCCCTTCCCACTTATCAAAC	129
	R	AAGGTCGCCATGTCTATTCT	
DGAT1	F	TACCAGCCAAGTTATCCTCG	146
	R	AAGGGGTGTTGTGAATTCTG	
DGAT2	F	GGCAGAAGAAGAAAAGAAGA	138
	R	AAGTTGAAATGGATAGAGCC	
Oleosin-1	F	GGTGAGTGGGACGCTTCTTT	134
	R	CCGACAACCCAATGACCAAC	
Oleosin-14.3	F	GTTCTTGTTCCTGCGGTTAT	136
	R	CTGCTCCTGGTGGATGCTTT	
Oleosin-16.6	F	TCTTTAGCCCTGTTCTTGTC	122
	R	GAGGTATTTCAGAACCCACG	
Caleosin	F	ATGGAAGCGATTCTGGTGTT	116
	R	CAACTCGCTTGATGTTAGGG	
Actin	F	AATGTATGTCGCCATCCAGG	145
	R	GTCAAGACGGAGGATAGCAT	

P[1]: Plastid; M[2]: Mitochondrion; F: Forward; R: Reverse.
GPDH: glyceraldehyde 3-phosphate dehydrogenase; PEPCase: phosphoenolpyruvate carboxylase; RuBisCo-SSU: ribulose-1-5-bisphosphate carboxylase small subunit; PD: pyruvate Dehydrogenase; ACCase: acetyl-CoA carboxylase; KAS: beta-ketoacyl-ACP synthase; KAR: ketoacyl-ACP reductase; FAT: stearoyl acyl-ACP-thioesterase; SAD: stearoyl-ACP desaturase; GPAT: glycerol-3-phosphate acyltransferase; DGAT: acyl-CoA-diacylglycerol acyltransferase.

3. Results

3.1. *Characterization of the seed development in J. curcas*

The seed development in *J. curcas* features noticeable changes in morphology as the seed color switches from milky white to light black then to dark black, while seeds are initially succulent and gradually become callous over the development progress. At the end, the seed shell becomes so hard and requires tools to break into parts (Fig. 1A). As shown, the seed sizes alter greatly during this period. At first, seed length was first employed to characterize the seed development, and showed the range of 3 to 22 mm from 5 DAF to 45 DAF, respectively. The seed lengths initially increased from 5 DAF to 25 DAF followed a slight decline from 30 DAF to 40 DAF, and were generally paralleled in trend with a minor shrinkage until 45 DAF (Fig. 1A and B), suggesting that the seed progresses into a desiccation phase after 40 DAF. Based on these morphological changes, we chose the seeds at 5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF, 30 DAF, 35 DAF, 40 DAF and 45 DAF to determine their fresh weight, dry weight and water content (Fig. 1C). In agreement with previous reports [26], our results showed that fresh and dry weight increased from 5 DAF to 40 DAF, followed by a significant decrease at 45 DAF (Fig. 1C). As such, the entire procedure of the seed development in *J. curcas* was further divided into 3 phases. Stage I is for the histodifferentiation characterized by the water content as 90% of the seed fresh weight with the seed length ranged at 3–20 mm. Stage II is the transition from histodifferentiation to seed filling, at which the seed fresh and dry weight increased rapidly and the seed length reached the maximum as long as 22 mm at 25 DAF. Stage III is the desiccation stage as the seed length of 45 DAF decreased to 18 mm as well as the water content decreased to below 10% (Fig. 1B and C).

3.2. *Cellular structure of endosperm from developing J. curcas seeds*

Since the TAG was mainly stored in the endosperm of *J. curcas* [31], we collected the endosperms of *J. curcas* at 5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF, 30 DAF, 35 DAF and 40 DAF to examine changes of their cellular structures. As shown in paraffin sections (Fig. 2), nucleus with a clear boundary was observed in cells of seeds at the initial stages (5 DAF, 10 DAF, 15 DAF and 20 DAF), indicating that they are at the histodifferentiation

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

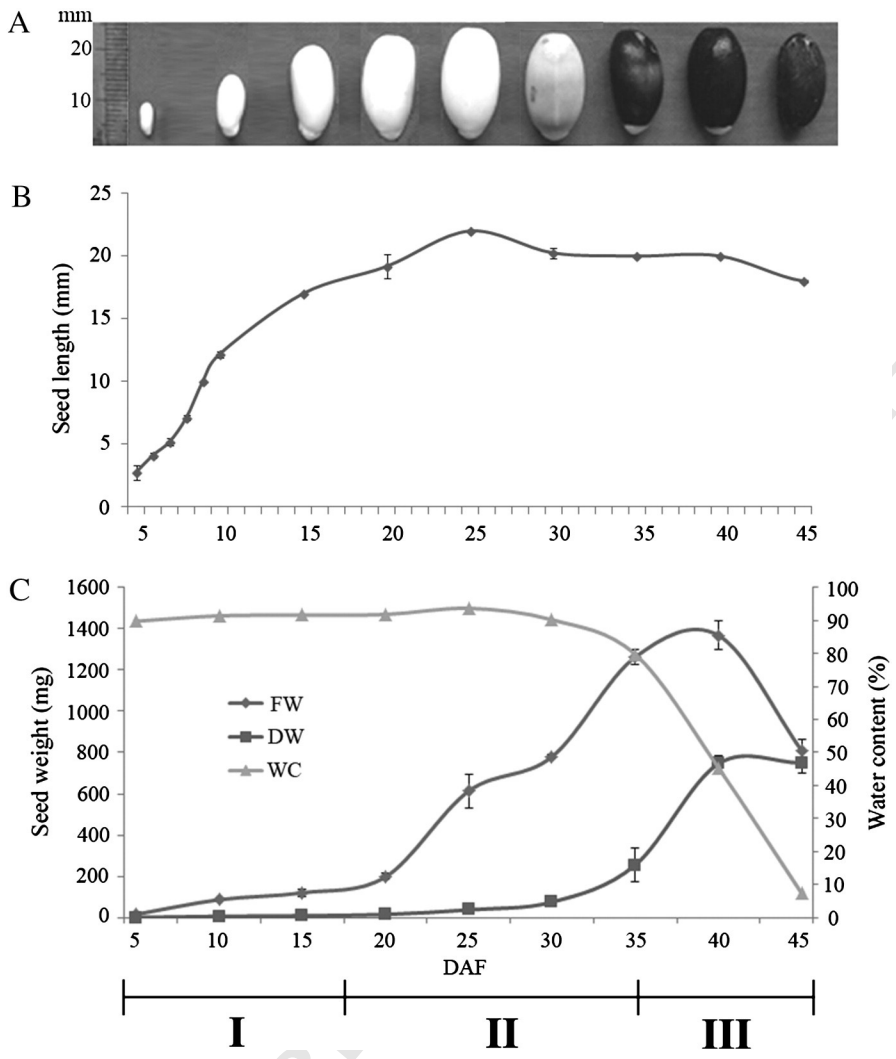


Fig. 1 – Characterization of *J. curcas* seed development. A) Images on the developing seeds of *J. curcas* from 5 DAF to DAF 45 with 5-day intervals. The bar (mm) indicating seed lengths was marked on the left. B) The profiles of seed length in the seed development of *J. curcas*. Error bars represent SD of 3 replicates. C) Water content, fresh weight and dry weight of developing seeds of *J. curcas*. FW: fresh weight, the line with rhombus; DW: dry weight, the line with square; WC: water content, the line with triangle. Average dry mass was subtracted from the average fresh mass to determine water content. Error bars represent SD for 10 replicates. Seeds at 5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF and 30 DAF were chosen for protein extract materials.

phase [1]. In addition, the dense cytoplasm and the endosperm array in a compact manner were found in cells, which showed the characterization of meristematic parenchyma cell. At 25 DAF, the cytoplasm density started to decrease, and a number of small oil bodies are accumulated in the endosperm cells at 35 DAF. Remarkably, large oil bodies were found in the connective endosperm cells at the 40 DAF in an abundant manner (Fig. 2), which suggested the end of the seed development of *J. curcas*.

3.3. *Component analysis and fatty acid (FA) composition*

We chose to determine the contents of reserve proteins and lipids in *J. curcas* seeds at the late development stages (Stage II and Stage III in Fig. 1) such as 25 DAF, 30 DAF, 35 DAF, 40 DAF and 45 DAF, because of their low contents at the early development phase (Stage I in Fig. 1). Our results showed that crude lipid and protein contents increased gradually in the late stages of the seed development (Fig. 3). Both the lipid and the protein content in seeds at 25 DAF were less than 5% of their dry weights, which was responding to the cellular results (Figs. 2 and 3). Beginning at 35 DAF, the accumulation rate of lipid was much faster than that of protein. The lipid and protein contents of the developing seeds reached the maximum at 45 DAF, but the protein content (15%) was much less than lipid content (40%) in seeds. Meanwhile, the FA compositions of the crude lipid at different DAFs were also analyzed (Table 2). The *J. curcas* seeds at 45 DAF contained 46.3% oleic acid ($C_{18:1n9c}$) and 32.92% linoleic acid ($C_{18:2n6c}$) as the highest level, and the next level for 13.64% palmitic acid ($C_{16:0}$) and 6.19% stearic acid ($C_{18:0}$). These results were similar with the previous reports [31,44]. Interestingly, the palmitic acid and the α-linolenic acid ($C_{18:3n3}$) both decreased from in developing seeds of *J. curcas* from 25 DAF to 45 DAF. Importantly, the accumulation of oleic acid and linolenic acid

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

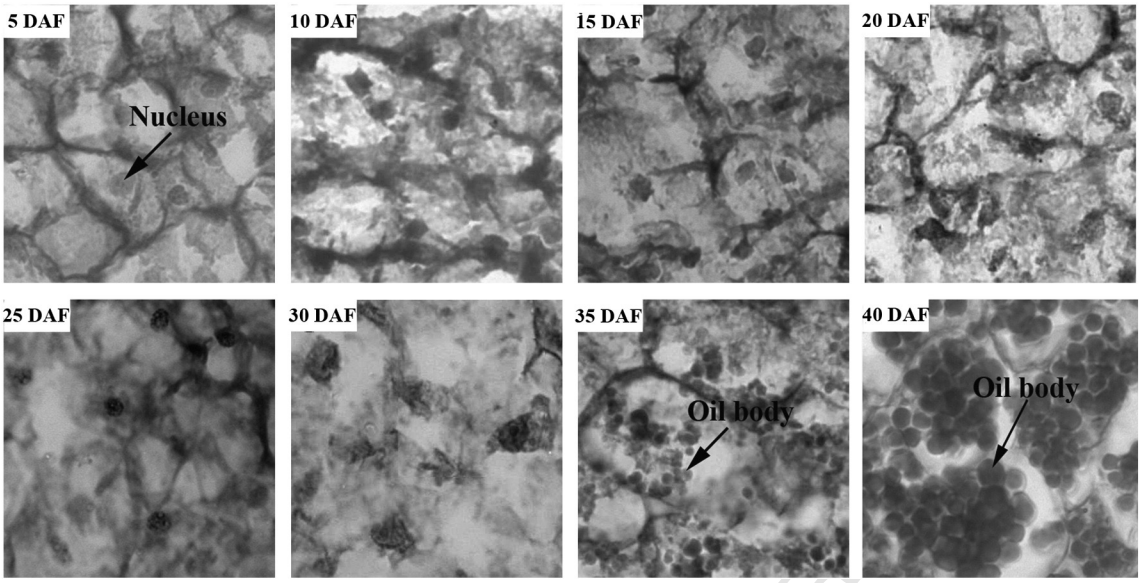


Fig. 2 – Cellular observation of endosperm in the seed development of *J. curcas*. Seeds from 5 DAF to 40 DAF with 5-day intervals were prepared for cellular observation.

showed its maximum abundance at 45 DAF and 35 DAF respectively while the palmitic acid reached their maximum abundance at 25 DAF, which suggested that $C_{16:0}$ was the precursor for elongation to $C_{18:1n9c}$ and $C_{18:2n6c}$, they are the most important fatty acids in *J. curcas* seeds. The decreasing content of α-linolenic acid might indicate the increase in lipid content was accompanied by a shift in lipid composition from phosphatidylcholine to TAG accumulation with seed development processing.

3.4. *Proteomic analysis of seed development*

The water soluble protein samples of different stages were performed by 2-DE in the range of linear pH 4–7. Protein profiles were established from 14.4 to 97.0 kDa with 527 (±6), 565 (±8), 590 (±8), 611 (±9), 617 (±8) and 702 (±10) replicable protein spots for developing *J. curcas* seeds at 5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF and 30 DAF respectively (Supplemental Fig. S2), the number of resolved spots increased from 5 DAF to 30 DAF gradually. With the development of *J. curcas* seeds process, greater protein diversity and higher content of specific spots could be visible in 2-DE profile of *J. curcas* as the same concentration of protein was loaded onto the gel. This result indicated that storage proteins in *J. curcas* had been accumulated in the late seed development.

By comparison of six stages of protein profiles with each other, 104 differential protein spots were statistically significant ($p < 0.05$), in which 24 differential spots were qualitative difference while 80 differential spots were quantitative difference, and mean, SD and CV of each differential protein spot were calculated (Supplemental Table S1). According to multivariate analysis, a general picture of the main variation and of interrelations between protein spots had been figured out. Firstly, by

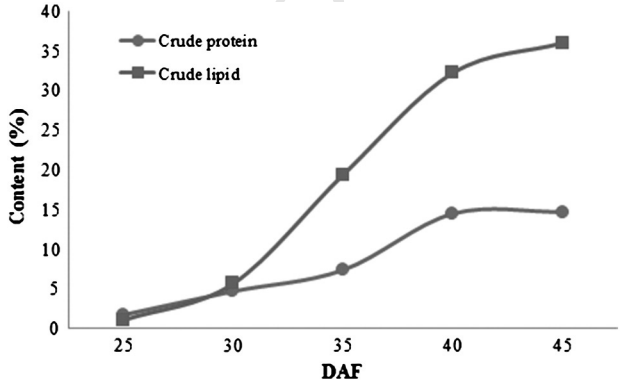


Fig. 3 – Crude lipid and protein content of *J. curcas* seeds in the late seed development. About 10 g dry masses sample at each developmental stage (25 DAF, 30 DAF, 35 DAF, 40 DAF and 45 DAF) was determined. Percentages were calculated using the average dry masses.

Table 2 – Changes of fatty acid and composition (%) for *J. curcas* seed oil in the seed development [a].

Sample	25 DAF	30 DAF	35 DAF	40 DAF	45 DAF
$C_{14:0}$ [b]	—	0.08%	0.04%	0.04%	0.05%
$C_{16:0}$	29.44%	20.23%	14.12%	13.8%	13.64%
$C_{16:1n7}$	—	1.69%	0.75%	0.59%	0.57%
C_{18}	2.54%	4.70%	6.09%	5.76%	6.19%
$C_{18:1n9c}$	3.89%	29.18%	38.98%	43.1%	46.30%
$C_{18:2n6c}$	36.86%	43.32%	39.72%	36.35%	32.92%
$C_{18:3n3}$	27.27%	0.58%	0.14%	0.16%	0.14%
$C_{20:0}$	—	0.18%	0.16%	0.16%	0.15%
$C_{20:1}$	—	0.04%	—	0.04%	0.04%

[a] Content of each fatty acid was calculated as the percentage that each fatty acid represented in the total measured fatty acids. The '—' indicates that data are undetectable.

[b] The numbers show the number of carbons and double bonds. For instance, C14:0 means 14 carbons and no double bond.

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

means of PCA analysis, a data reduction to the whole dataset and to the 104 differential spots was applied. Out of the potential PCs extracted, the first seventeen PCs accounted for 100% of the biological variability of each whole or differential spots dataset, respectively (Table 3). PCs 1–2 were significant in the two studied datasets (one-way ANOVA, >95% confidence) when PCs were tested for differences between groups. By employing these components in 3-D (plotting PC1, PC2 and PC3) representations, it was able to effectively separate the samples into their original groups (Supplemental Fig. S3), and the plot structure was slightly different between whole and differential spots datasets. Samples of 5 DAF, 10 DAF and 15 DAF were closely grouped in the plot of whole spot dataset, indicating similarity in the 2-DE spot map (Supplemental Fig. S3A). 25 DAF samples were grouped in the plot corresponding to the differential spots but not in the whole spot dataset, which was not strange. As Sghaier-Hammami et al. mentioned [42], this difference between groups indicate that some key spots with high loading over PCs and mutual to both samples were removed after univariate analysis.

The 104 differential protein were successfully identified by MALDI TOF/TOF tandem MS analysis (Fig. 4 and Supplemental Table S2). All identified protein species were classified into 10 functional classes that were originally established by Bevan et al. in the *Arabidopsis* genome project [45]. Protein species associated with metabolism, energy, destination and storage, and disease defense were most abundant than other functional classes (Fig. 5 and Supplemental Table S2). Thirty-three identified protein species related to carbon flux to lipid accumulation were chosen for analysis of the metabolic network in the seed development of *J. curcas*, and enlarged images were displayed in Fig. 6 (Figs. 4, 6 and Table 4). They were clustered employing Pearson distance-based dissimilarity matrix (Fig. 7), and "0" value was replaced by a minimal value, 10^{-9} to indicate the biological significance. The abundance profiles of protein species involved in glycolysis including Enolase (spot B1), Fructokinase (spots A14, B18 and B19), Aldolase (spot F16), Phosphoglycerate kinase (spot C22), and 3-phosphoglycerate dehydrogenase (spot C6) were in contrast with those involved in plastidial glycolysis including Enolase (spot C11) and Phosphoglycerate kinase (spot E12). UDP-Glc pyrophoshorylase (UGPase) (spot D8), 6-phosphogluconate dehydrogenase (spot E3), 6-phosphogluconolactonase (spot F30) and RuBisCo large subunit (spot F6) shared similar abundance, which was different with other OPPP enzymes, Ribose-5-phosphate isomerase (spot A23) and Transaldolase (spots C20 and F15). All protein species involved in TCA cycle containing ATP synthase subunit beta (spots D6 and F5), Succinyl CoA ligase beta subunit (spots D10 and D13) were down accumulated except ATP synthase subunit beta (spot D7). Interestingly, protein species participated in the FA synthesis, including KAR (spots B26 and F31), SAD (spot C25), Dihydrolipoyl dehydrogenase (spot F7) and KAS I (spot D9) were significantly up-accumulated except acetyl-CoA biotin carboxylase (spots B7 and C14), the rate-limiting enzyme which catalyzes ATP-dependent malonyl-CoA formation (Figs. 6, 7 and Table 4). Overall, 16 in 33 protein species (spots B26, F31, C11, C25, F34, F7, A14, D9, B18, D8, E3, F16, E12, F6, D7 and F30) were clustered into one major group while other protein species (spots B7, B15, E14, B19, B25, C14, C20, C22, C6, D6, B1, A16, A23, F15, F5, D10 and D13) were clustered into another major group (Fig. 7).

3.5. *qRT-PCR analysis of key genes in seed development*

Combined with the results of proteomic analysis, 19 genes were chosen to confirm their mRNA expression. Five of them were involved in the metabolism of carbon flux including *PD* in plastid and mitochondrion, *GDPH* in plastid, *PEPCase* in cytosol, and *RuBisCo*. The genes related to FA synthesis were *ACCase*, *KASI*, *KASII*, *KASIII*, *KAR*, *FATB*, *FATA* and *SAD*. Additionally, *DGAT1*, *DGAT2*, 3 isoforms of *oleosins* and *caleosin*, were chosen for our analysis due to their important role in lipid formation (Fig. 8). According to the cluster analysis, it revealed that the expressions of *PDE2*, *PEPCase*, *ACCase*, *KASI*, *KASII*, *KASIII*, *KAR* and *FATB* showed the similar expression

Table 3 – PCs calculated from the 466 spots matched between samples and from the 104 differential spot set.

Component	Eigenvalues (466 spots)			Eigenvalues (104 spots)		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
PC1	190.258	46.301	46.301	127.4334	68.781	68.781
PC2	134.3464	23.086	69.387	64.49922	17.62	86.401
PC3	66.4698	5.651	75.039	30.28382	3.884	90.286
PC4	60.69402	4.712	79.75	27.01839	3.092	93.377
PC5	55.67925	3.965	83.716	20.34491	1.753	95.131
PC6	49.15833	3.091	86.807	17.9899	1.371	96.501
PC7	45.13632	2.606	89.413	13.10066	0.727	97.228
PC8	39.53315	1.999	91.412	12.22479	0.633	97.861
PC9	35.77831	1.637	93.049	9.617031	0.392	98.253
PC10	35.02229	1.569	94.618	9.129335	0.353	98.606
PC11	31.42418	1.263	95.881	8.582107	0.312	98.918
PC12	29.78093	1.134	97.016	7.680731	0.25	99.168
PC13	24.80844	0.787	97.803	7.547759	0.241	99.409
PC14	23.61958	0.714	98.516	6.45729	0.177	99.586
PC15	21.86971	0.612	99.128	6.177163	0.162	99.747
PC16	18.81193	0.453	99.581	5.867839	0.146	99.893
PC17	18.10389	0.419	100	5.022614	0.107	100

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

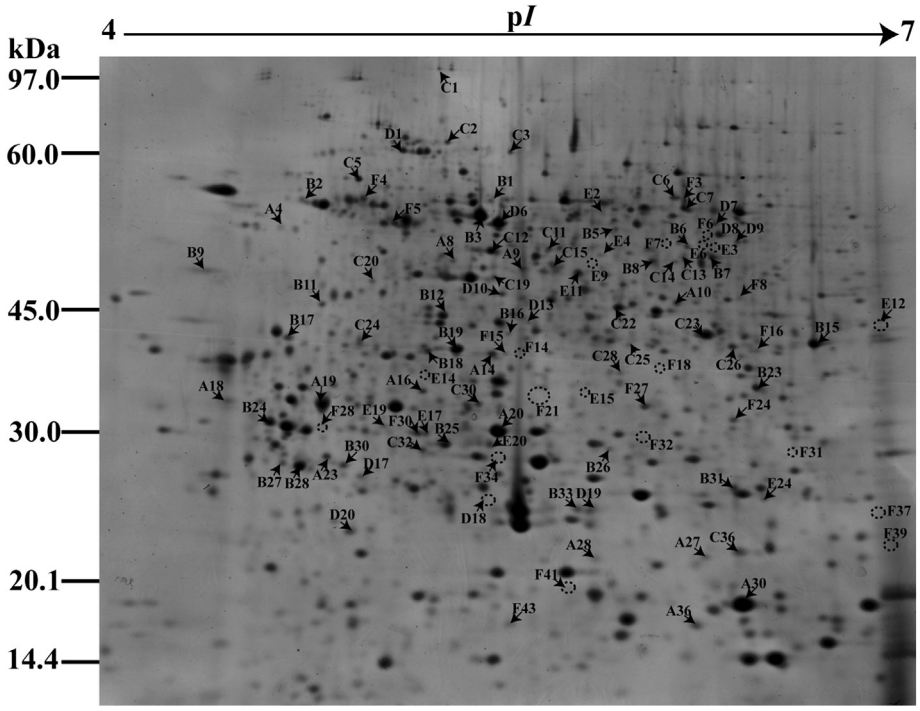


Fig. 4 – 2-DE master gel of *J. curcas* seeds in development process. Arrows indicate the 104 protein species with significantly variance identified by MALDI/TOF-TOF. The A, B, C, D, E, F letters of the protein spot number stand for the protein spots from the gel of seeds at 5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF and 30 DAF. The Mr is given on the left, while the pI is given at the top. The circle with dotted line indicates the spot is absent in the 2-DE master gel.

among these six seed developmental stages. Meanwhile, *GPDH-P*, *PDE1-M*, *RuBisCo-SSU*, *FATA* and *SAD* reached the highest expression level at 25 DAF. *DGAT1* and *DGAT2* are the key enzymes for TAG synthesis in endoplasmic reticulum (ER), both of which showed the highest expression level at 30 DAF. Moreover, the four genes encoding for two main oil body

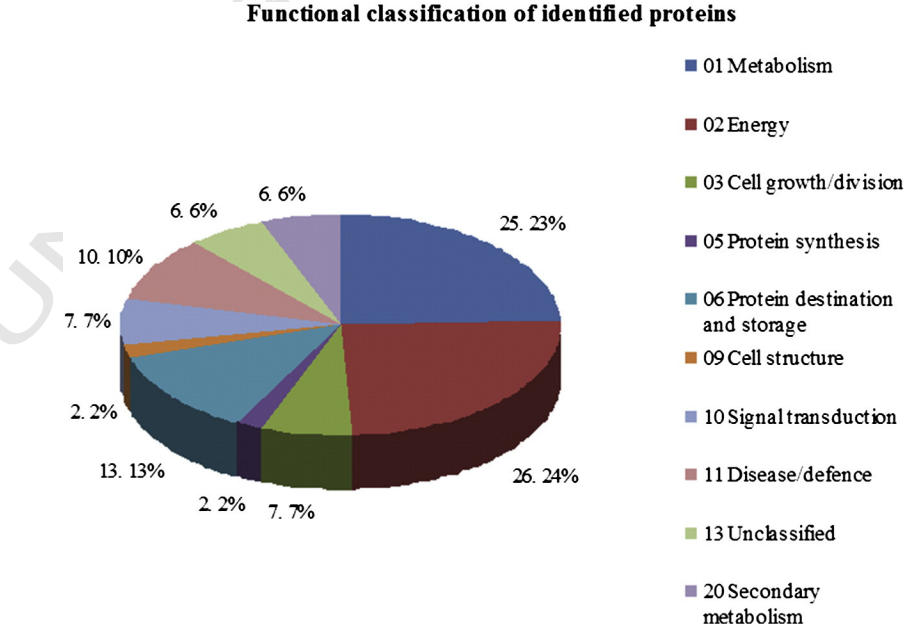


Fig. 5 – Functional classification of 104 identified differential proteins in the seed development of *J. curcas* according to criteria described by Bevan et al. (1998).

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

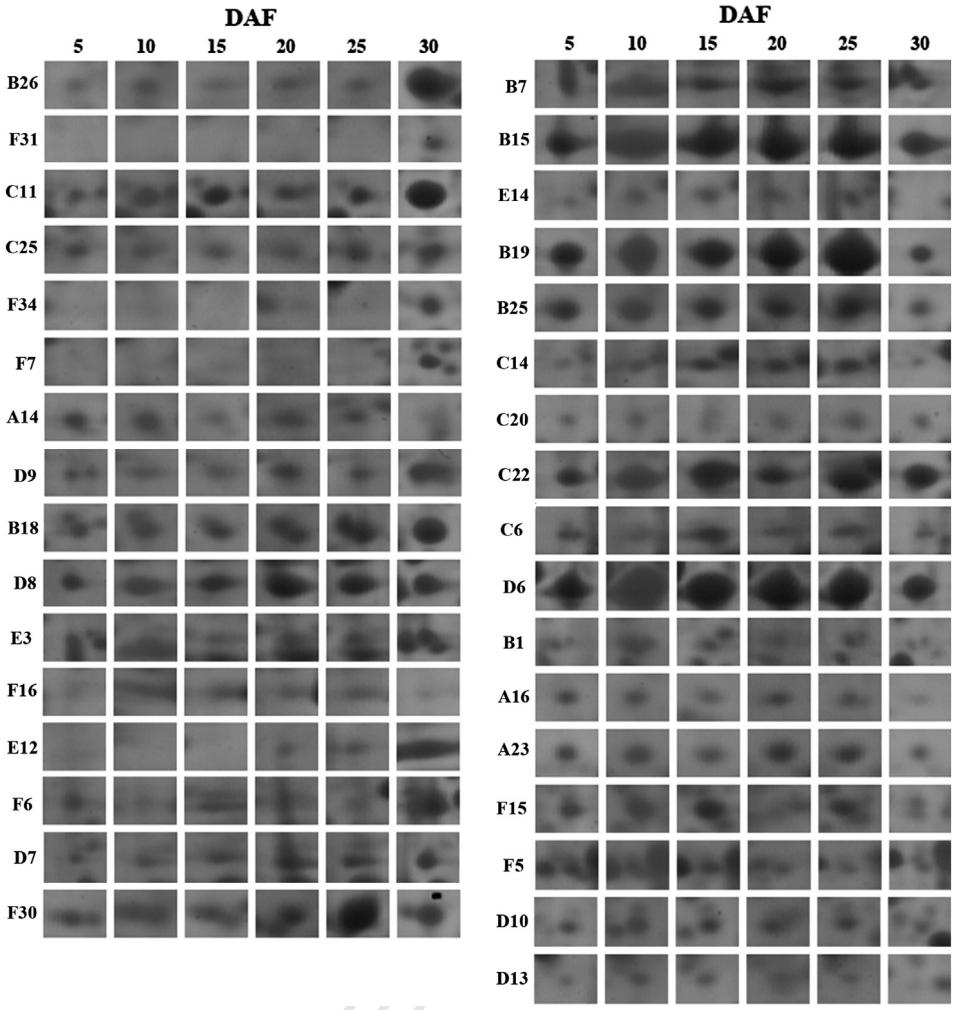


Fig. 6 – Enlarged images to show the distribution of the 33 differentially protein spots involved in the metabolism from carbon flux to lipid accumulation.

proteins *oleosin* and *caleosin* showed the expression profiles of *DGAT1* and *DGAT2* in the same manner (Fig. 9).

4. Discussions

The economic value of *J. curcas* largely depends on the lipid content in seeds. Hence, our discussions mainly focus on the pathways leading to de novo FA synthesis and lipid accumulation. Different from other organisms, plants produce FAs from acetyl-CoA through a common pathway in plastids [46]. In this study, the identified protein annotations less likely provide the accurate information about the intracellular localization was caused by lacking of comprehensive protein sequence databases. The NCBI (National Center for Biotechnology Information) non-redundant protein database contained only 1060 entries for *J. curcas* (Till to December 2012). Therefore, we may postulate protein localizations according to metabolisms involved in when the subcellular information is absent. Our results in this study, together with reports by other references [8,32,47], successfully elucidate a comprehensive network of the metabolic activities related to lipid accumulation (Fig. 10).

4.1. *Differential protein species involved in sugar release*

Sucrose is the main storage carbohydrate in plants [48]. Converting sucrose into UDP-G and Fru by Suc synthase is the initial release of sugar for glycolysis. UDP-G can be reversibly catalyzed by UGPase and then generate G-1-P. This is an important step of the Suc synthase pathway by which carbon enters the cytosolic hexose phosphate pool for carbon flux [49,50]. In this study, UGPase (spot D8) and Frk (spot B18) showed similar pattern to that of glycolytic enzyme Aldolase (spot F16) in cytoplasm and OPPP enzymes 6-phosphgluconate dehydrogenase (spot E3) and 6-phosphogluconolactonase (spot F30) in plastid (Figs. 6 and 7). Their increase is related to the production of F-6-P, G-6-P and ribulose-5-phosphate (Ru-5-P) (Fig. 10). It reveals that G-6-P is an important shuttle metabolite between cytoplasm and plastid in *J. curcas* seeds. At the same time, Pyrophosphate (ppi) is generated in these reactions and then converted to pi quickly by the catalysis of the acid PPase in the cytoplasm [51,52]. Ppi acts as the substrate of the forward reaction of cytosolic enzymes such as UGPase, on the other hand, ppi is also the substrate of the forward reaction of ppi-dependent phosphofructokinase (PFP) at the entry point of glycolysis [53]. We identified two acid PPase (spots A16 and B25)

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

Table 4 – Identification of differential protein species involved in lipid accumulation in the seed development of *J. curcas* by MALDI-ToF/TOF analysis.

Spot no.	Accession no.	Name	Organism	L[a]	Theor. Mr/pI	Exp. Mr/pI	PS[b]	P[c]	CS[d]
01.06 Lipid and sterol									
B26	gi|255634733	3-ketoacyl-ACP reductase	*Glycine max*	P	26766/5.97	28/5.81064	50	1	1
F31	gi|119791	3-ketoacyl-ACP reductase	*Cuphea lanceolata*	C	33310/9.28	29/6.22837	134	1	1
C25	gi|134943	Acyl-ACP desaturase	*Brassica napus*	C	45490/5.75	40/5.7922	72	1	1
B7	gi|238837063	Heteromeric acetyl-CoA biotin carboxylase	*Jatropha curcas*	P	58795/7.60	51/5.9922	400	5	5
C14	gi|238837063	Heteromeric acetyl-CoA biotin carboxylase	*Jatropha curcas*	P	58795/7.60	50/5.91702	309	4	4
D9	gi|116248667	Beta-ketoacyl-ACP synthase I	*Jatropha curcas*	P	50255/6.86	54/6.05145	168	2	2
F34	gi|6478218	Caleosin	*Sesamum indicum*	OB	27788.7/5.59	27/5.50729	112	2	2
F7	gi|449460949	Dihydrolipoyl dehydrogenase-like	*Cucumis sativus*	—	59713/6.73	54/5.91364	344	3	3
02.01 Glycolysis									
C6	gi|449445906	3-phosphoglycerate dehydrogenase	*Cucumis sativus*		63349/6.32	59/5.92553	207	3	3
F16	gi|218157	Aldolase	*Oryza sativa*	CP	39151.2/6.56	40/6.1782	139	5	5
B1	gi|255539693	Enolase	*Ricinus communis*	—	48149/5.71	59/5.54965	150	2	2
C11	gi|255539693	Enolase	*R. communis*	—	48149/5.71	52/5.63475	88	2	2
B18	gi|255585331	Fructokinase	*R. communis*	—	35591/5.18	39/5.34539	82	1	1
B19	gi|357454485	Fructokinase	*Medicago truncatula*	—	35386/5.20	39/5.41489	377	3	3
A14	gi|356550378	Fructokinase-2-like	*Glycine max*	—	35546/5.29	38/5.49858	296	3	3
C22	gi|255544582	Phosphoglycerate kinase	*R. communis*	—	42530/5.65	45/5.78794	285	3	3
E12	gi|1346698	Phosphoglycerate kinase	*Spinacia oleracea*	C	45658/5.83	44/6.40081	104	1	1
02.02 Gluconeogenesis									
A16	gi|255554527	Inorganic pyrophosphatase	*R. communis*	—	34062/5.91	34/5.33688	249	5	5
B25	gi|255561090	Inorganic pyrophosphatase	*R. communis*	—	24034/5.60	28/5.38652	443	6	4
B15	gi|225438145	Malate dehydrogenase	*Vitis vinifera*	CP	35881/6.18	41/6.23901	373	4	3
E14	gi|225438145	Malate dehydrogenase	*Vitis vinifera*	CP	35881/6.18	41/5.43025	71	1	1
D8	gi|255635072	UGPase	*Glycine max*	—	51576/5.41	54/6.01631	416	3	3
02.07 Pentose phosphate									
F30	gi|255554349	6-phosphogluconolactonase	*R. communis*	—	28115/5.44	30/5.25799	193	3	3
E3	gi|255537671	6-phosphogluconate dehydrogenase	*R. communis*	—	54511/6.25	53/5.98811	420	7	7
A23	gi|18654317	Ribose-5-phosphate isomerase	*Spinacia oleracea*	C	30958/6.54	28/5.08318	251	3	2
C20	gi|224075445	Transaldolase-like proteins	*Populus trichocarpa*	—	47693/6.38	48/5.2078	498	4	4
F15	gi|295687233	Transaldolase	*Gossypium hirsutum*	—	43059/5.78	41/5.60496	122	2	2
02.10 TCA cycle									
D10	gi|21593189	Succinyl CoA ligase beta subunit	*Arabidopsis thaliana*	M	45247/6.04	46/5.53223	426	5	5
D13	gi|11272036	Succinate CoA ligase ADP forming	*Arabidopsis thaliana*	M	45316/6.29	40/5.53903	271	4	4
D6	gi|231586	ATP synthase subunit beta	*Hevea brasiliensis*	M	60335/5.95	62/6.29438	1001	8	8
D7	gi|114421	ATP synthase subunit beta	*Nicotiana plumbaginifolia*	M	59933/5.95	61/4.81755	220	4	4
F5	gi|231586	ATP synthase subunit beta	*Hevea brasiliensis*	M	60335.4/5.95	56/5.15411	580	15	15
02.30 Photosynthesis									
F6	gi|1770216	RuBisCo large subunit	*Didymosalpinx norae*	C	52985/6.47	55/5.9762	349	5	5

Functional categories were decided by the system in Bevan et al. (1998). All the protein identities were from searching in NCBInr database, UniprotKB/SwissProt database and EST_Viridiplantae database. All the peptide charge is +1. The false discovered rate is less than 5%.

[a] L: subcellular location; C: Chloroplast; CP: Cytoplasm; M: Mitochondria; P: Plastid; V: Vacuole; —: No site information.

[b] PS, Protein score.

[c] P, the matched peptides.

[d] CS, the number of covered peptides.

had different MW and pI values in the 2-DE gels were annotated as the same protein function (Fig. 4 and Table 4), suggesting that they might be present as different isoforms of the same protein family. That is the reason that they share different abundance profiles, and the abundance profile of one PPase (spot B25) is closer to UGPase than another PPase (spot A16) in the seed development (Fig. 7). We speculate that former PPase participates in the reaction of UGPase, and latter one participates in the reaction of PFP instead of the traditional ATP-dependent phosphofructokinase (PFK). Besides, among the genes involved in cytosol glycolysis in developing *J. curcas*, the expression of *PFK* at 14 DAF, 19 DAF and 25 DAF were the lowest compared with that of other glycolytic genes [54]. Considering that FPK protein species was not identified in our proteomic analysis, these results notion may indicate that the abundance of PFK is low at the early seed development of *J. curcas*, which is similar in castor

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

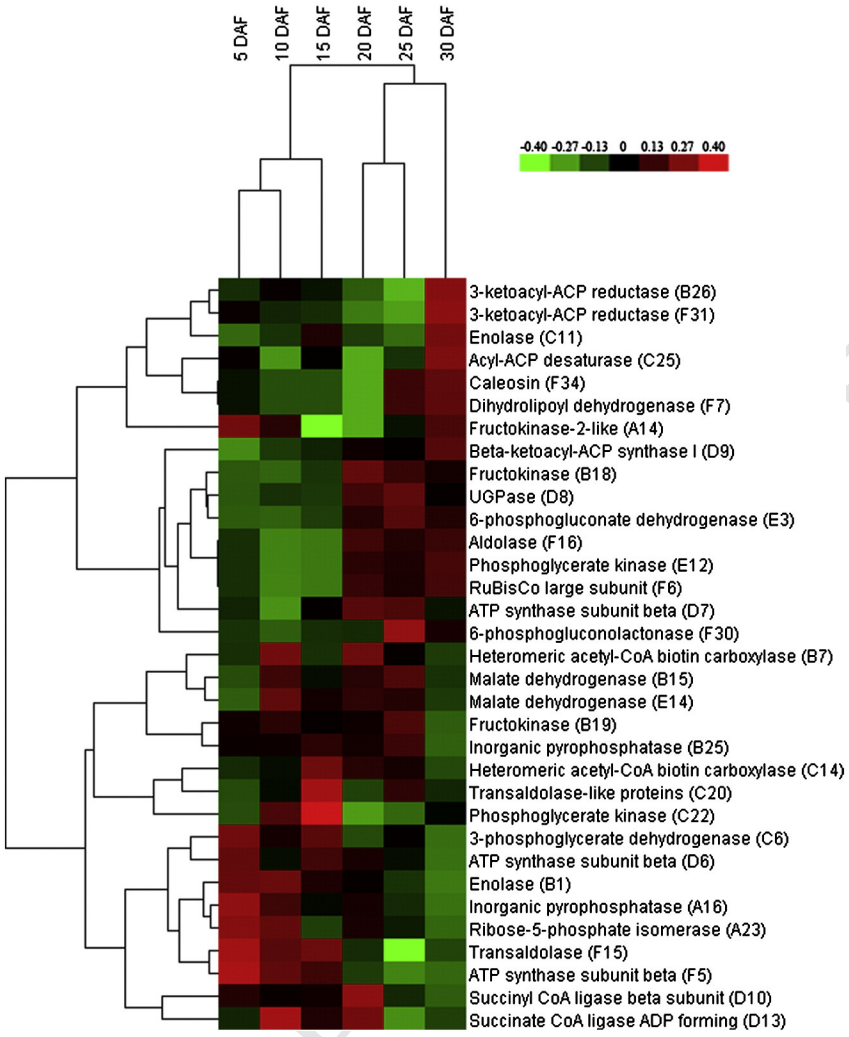


Fig. 7 – Cluster analysis of abundance profiles of 33 protein species related to metabolism form carbon flux to lipid accumulation. This heat map was plotted by employing Cluster 3.0 and Java Tree View. The values of protein spots were log-transformed and "0" value was replaced by a minimal value, 10^{-9} to indicate the biological significance. The protein spot number in brackets is corresponding to the description in Figs. 4, 6 and Table 4.

but different from rapeseed [8,16]. And another possibility is that FPK in developing *J. curcas* seeds may not function as traditional role in cytosolic glycolysis at the early stage.

On the other hand, the other two Frk (spots A14 and B19) with different abundance profiles also involved in the reaction described above (Figs. 6, 7 and 10), according to the observation from the 2-DE gels (Fig. 4), the spots A14, B18 and B19 corresponding to Frk have nearly the same MW with minor p*I* shift, indicating the post translational modification. Few literature reports that Frk-2 is studied in oilseed crops, and at least two isoenzymes expressed in the early fruit development of tomato [55]. In addition, the expression level of *Frk-2* mRNA plays a critical role in numbers of flower and fruit in tomato [56]. Moreover, Frk-2 showed its highest abundance at the early seed development of *J. curcas* (Fig. 6), which is consistent with the expression of mRNA *Frk-2* in tomato. We postulate that Frk-2 may execute similar biological functions in developing seeds of *J. curcas* as in tomato. This characteristic may provide us valuable information to further improve *J. curcas* for suitable biodiesel materials.

4.2. *Differential protein species related to carbon flux to FA synthesis*

Glycolysis, OPPP and TCA cycle produce sufficient carbon sources and energy to meet the requirement for FA synthesis. Therefore, in this study we investigated the contributions of G-6-P, triose-P, pyruvate and malate to FA synthesis by comparing metabolic enzymatic steps.

G-6-P is produced normally through 2 potential pathways. Once synthesized, G-6-P either continues to involve in the glycolysis or is imported into plastid and converted into 6-phosphate-gluconolactone (6-PGL) following the continuous irreversible reactions. This process is catalyzed by 6-phosphogluconolactonase (spot F30) and 6-phosphogluconate dehydrogenase (spots E3). The latter protein species is a rate-limiting enzyme in the OPPP, which is the main pathway for generation NADPH and reducing power in most organisms. These abundance profiles suggest that the abundant cytosolic UGPase may play a critical role in plastidial OPPP, because its profile is the same as the composite abundance profiles of key

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

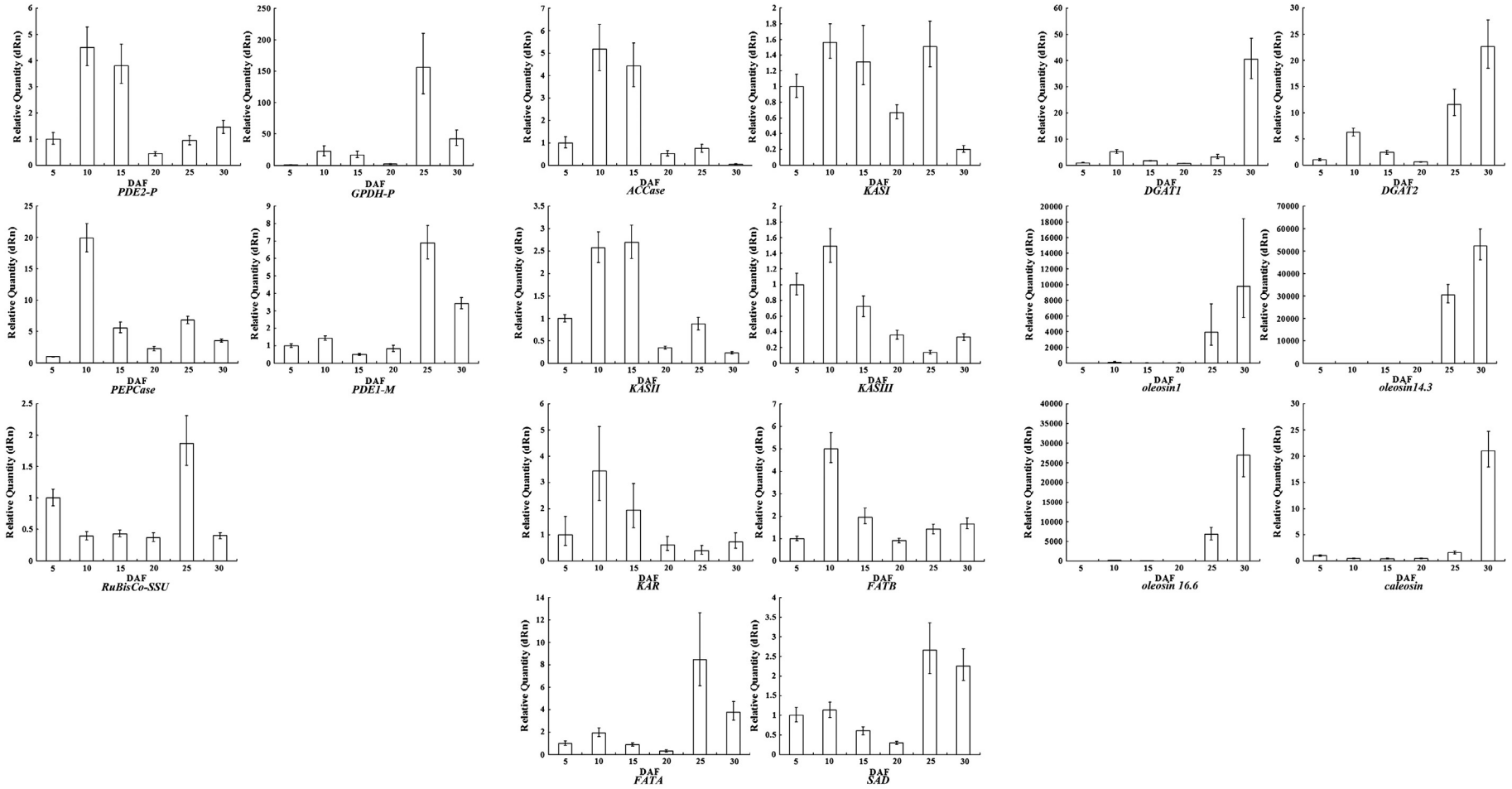


Fig. 8 – Temporal expression profiles of genes involved in metabolic network of lipid accumulation in the seed development of *J. curcas*. qRT-PCR was performed with cDNA isolated from developing seeds at 5 DAF, 10 DAF, 15 DAF, 20 DAF, 25 DAF and 30 DAF (from left to right). Relative expression ratio of each sample is compared with the sample at 5 DAF. Actin gene from *J. curcas* was used as an internal control. The final relative cDNA amounts of genes are means of 3 replicates. The relative expression ratios of genes were significantly different at $p < 0.05$. The abbreviation of gene name is corresponding to the description in Table 1.

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

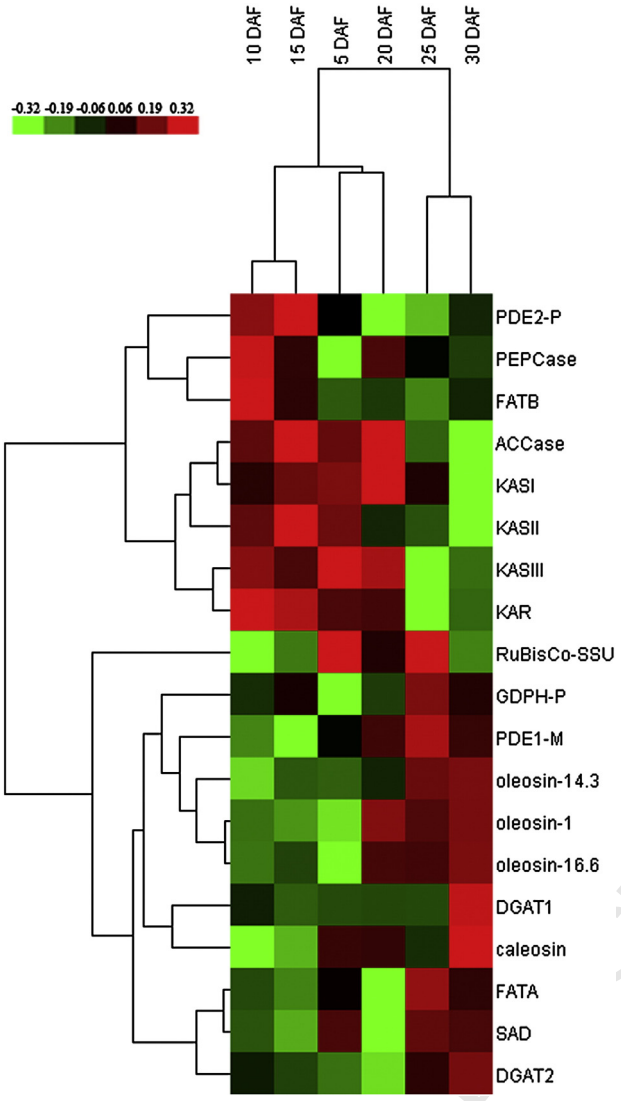


Fig. 9 – Cluster analysis of expression profiles of 19 mRNA involved in metabolic network of lipid accumulation. This heat map was plotted by employing Cluster 3.0 and Java Tree View. The gene name is corresponding to the description in Fig. 8 and Table 1.

OPPP enzymes. The results of this study, combined with our observations, strongly suggest that biochemical processes mentioned above were prepared at the appropriate time to provide energy and building blocks for the biosynthesis of fatty acids, besides, combined with the result of cellular observation (Fig. 2), hexose metabolism seems to be associated with meristematic activity in developing *J. curcas* seeds.

Nevertheless, one ribose-5-phosphate isomerase (spot A23) and two isoforms of transaldolase (spots C20 and F15) were down-accumulated and showed different abundance profiles with spots F30 and E3 (Fig. 7), which may indicate that the important product Ru-5-P in OPPP participates in another irregular pathway. Once Ru-5-P was converted to ribulose bisphosphate (RuBP), which is able to capture the carbon dioxide (CO_2) by the catalysis of RuBisCo. RuBisCo was recently found playing an essential role in the recycle of CO_2 released by plastidial pyruvate dehydrogenase complex in order to maintain the efficiency of oil production in embryo [57], the RuBisCo captured the CO_2 and directly mediated the synthesis of 3-phosphoglycerate instead of the traditional glycolysis with the participation of the GPDH and PK [58]. In this study, we indeed found that the RuBisCo large subunit (LSU) (spot F6) showed the similar pattern with Dihydrolipoyl dehydrogenase (spot F7), which is one of the component of pyruvate dehydrogenase complex, besides, the expression of *RuBisCo* small subunit (SSU) mRNA also increased dramatically as spots F7 and F6 presented (Figs. 6, 7, 8 and 9). These results suggest that RuBisCo may mediate the irregular pathway described above in seed development of *J. curcas*. In addition, one study had reported that the activation of RuBisCo without Calvin cycle improved the carbon efficiency during the oil formation in developing embryos of rapeseed, more than 20% acetyl-CoA could be used for fatty acid synthesis in this metabolic conversion compared with glycolysis [59]. Therefore, we compared the RuBisCo abundances in *J. curcas* with that in rapeseed and castor, and the result revealed that the abundance of RuBisCo is prominent in rapeseed and *J. curcas*, perhaps this difference is the reason of the low carbon efficiency and subsequently reduced oil in castor. These results of proteomic studies that focus on seed development of oil crop, may be supposed that RuBisCo may play specific and important roles to provide enough carbon source with high efficiency at the seed development of oil crops.

Through the metabolic process, a series of metabolites were generated in cytosolic and plastidial glycolysis. Starting from the triose-Ps, 3-phosphoglycerate dehydrogenase (spot C6), Phosphoglycerate kinase (spot C22) and Enolase (spot B1) participate in the cytosolic glycolysis, and show opposite pattern to Phosphoglycerate kinase (spot E12), Enolase (spot C11) involved in plastidial glycolysis (Figs. 6, 7 and 10). Especially, the abundance of enolase can barely be observed at 30 DAF (Fig. 6). In order to understand the role of triose-Ps in cytosolic and plastidial glycolysis, *GPDH* mRNA was analyzed by qRT-PCR, and data showed the expression profile of *GPDH* mRNA was at maximum in plastid at 25 DAF. The results indicate that in the seed development of *J. curcas*: 1) G-6-P and DHAP as glycolytic intermediates are the important metabolites transported into plastid to participate in the plastidial glycolysis; 2) most steps of OPPP take place in plastids; 3) glycolytic PEP and pyruvate are rarely used as carbon source in the late seed development of *J. curcas* (Fig. 10).

Subsequently, glycolytic intermediates, such as PEP, oxalacetic acid (OAA), malate and pyruvate may indirectly lead to similar components through TCA cycle and malate synthesis. Cytosolic malate dehydrogenase (MDH) (spots B14 and B15) was identified in this study, and showed the same MW and different p*I* values in the 2-DE gels were annotated as the same protein species (Fig. 4 and Table 4), suggesting that they might be present as different modifications of the same gene product. In higher plants, there are three MDH isoforms that are widely distributed enzymes in cytoplasm, mitochondria and microbodies. They play important roles in many key metabolic pathways containing the TCA cycle, gluconeogenesis and facilitation of exchange of metabolites between cytoplasm and subcellular organelles [60,61]. The favorite substrate of MDH is OAA, and the MDH abundance profiles in this study are consistent with Heteromeric acetyl-

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

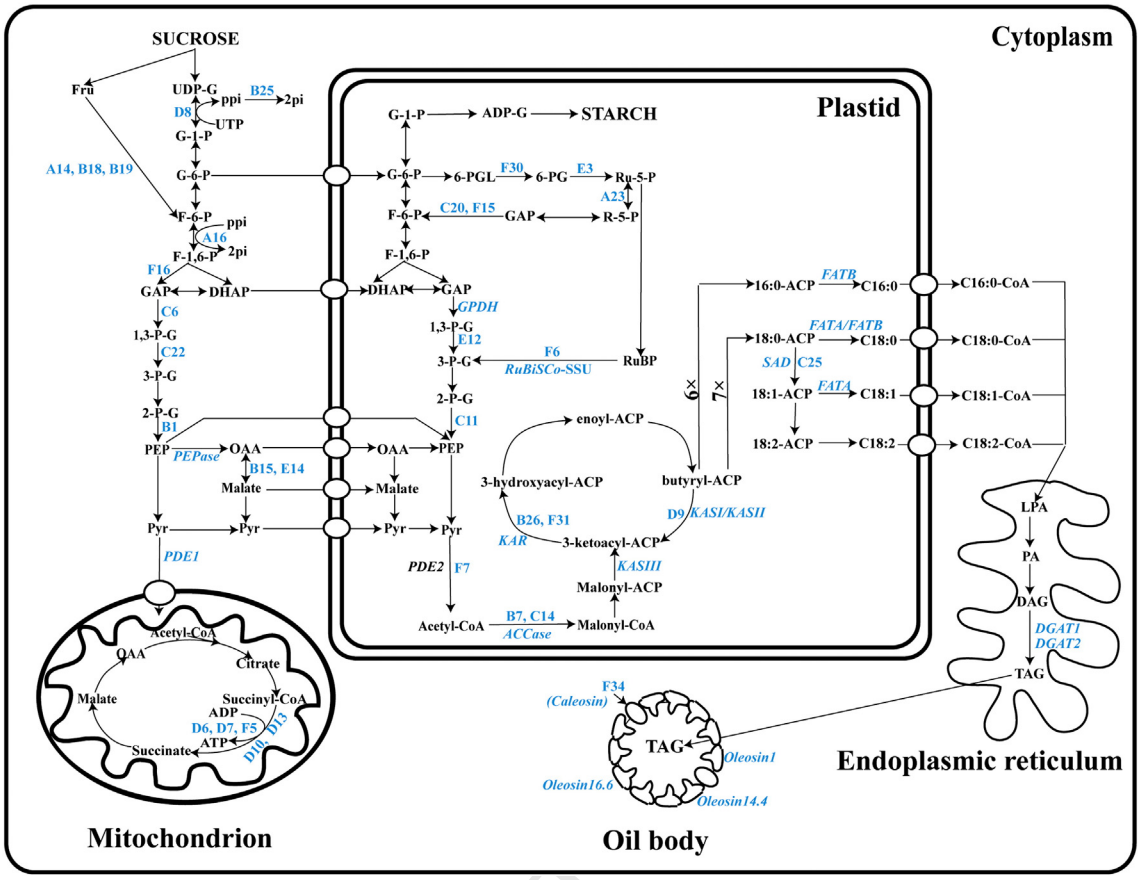


Fig. 10 – Metabolic network from carbon flux to lipid accumulation during seed developing process of *J. curcas*. The identified protein spots and detected key genes linked pathways from carbon flux to lipid accumulation are highlighted by blue color. Protein spot numbers are corresponding to the numbers described in Table 4, Figs. 4, 6 and 7. Italic type indicates gene confirmed by qRT-PCR analysis. Gene names are corresponding to the description in Figs. 8 and 9. GAP: glyceraldyhyde-3-phosphater; RuBP: Ribulose-1, 5-bisphosphate.

CoA biotin carboxylase (spot B7) which is an obvious contrast to those abundance profiles (down accumulated) of glycolytic protein species (spots C6 and B1), ATP synthase subunit beta (spots D6 and F5) and Succinyl CoA ligase beta subunit (spots D10 and D13) involved in TCA cycle (Figs. 7 and 10), which were able to convert succinyl-CoA to succinate with a substrate level phosphorylation of ADP producing ATP [62,63]. Especially, ATP synthase corresponding to spot D6 is one of high abundance protein spots from 2-DE profiles of *J. curcas* developing seeds, and its abundance at 5 DAF is nearly 3 times higher than at 30 DAF (Figs. 4 and 6). This result reveals that TCA cycle is active in the early seed development of *J. curcas*. According to their MW, pI and abundance in developing *J. curcas*, it seems that the proteins with higher molecular mass and abundance are more easily modified possibly attributing not only to more amino acids in these proteins, but also the conservation of the constitutive proteins that are required for the maintenance of basic cellular function. Besides, mRNA *PEPase* showed the different expression with *PDE1*, and reached maximum at the early seed development (Figs. 8 and 9), which indicates that another metabolism provides the major source of OAA for MDH instead of the cytosolic glycolysis. Therefore, we postulate that TCA cycle would provide citrate from the mitochondrion to cytosol in the late seed development of *J. curcas*. Although we failed to identify ATP-citrate lyase which catalyzes the ATP and CoA dependent cleavage of citrate to form OAA and acetyl-CoA, considering the fact that fatty acids were synthesized at the late seed development of *J. curcas* (Table 2) and acetyl-CoA should be prepared for the elongation of fatty acids, and together with the result of proteomic analysis of castor, a plant in Euphorbiaceae family as *J. curcas* [16], we suppose that cytosolic MDH may play a complementary role in the metabolite exchanges between cytoplasm and plastid in developing seeds of *J. curcas*.

4.3. *Differential protein species involved in FA synthesis*

All identified differential protein species (spots B26, F31, C25 and F7) involved in fatty acid synthesis were obviously up-accumulated in the late seed development of *J. curcas* except heteromeric acetyl-CoA biotin carboxylase protein species (spots B7 and C14) (Figs. 7 and 10), which catalyzed the conversion of acetyl-CoA into malonyl-CoA and was essential to regulate the rate of fatty acid synthesis [46,64]. This result suggests that carbon source has been greatly conversed into acetyl-CoA for the preparation of fatty acid synthesis in the late seed development of *J. curcas*. Abundance profiles of protein species involved in remaining reactions toward fatty acid synthesis were highly elevated at 30 DAF.

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

As the initiation enzyme of FA chain elongation, KASIII is responsible for the condensation reaction of malonyl-ACP and acetyl-ACP, KASI and KASII are the condensing enzymes for the elongation of the carbon chain from C_4 to C_{18} [65,66]. The expression profiles of *KASI*, *KASII* and *KASIII* are similar with each other but different from the abundance profile of KASI (Figs. 7, 8 and 9), which is possible that the gene expression is much earlier than protein expression or protein species change its chemical structure as a consequence of posttranslational modifications. Two KAR isoforms (spots B26 and F31) share similar abundance trends with KASI (Fig. 7). In particular, one of KAR corresponding to spot F31 newly appeared in the 2-DE gels at 30 DAF (Fig. 6 and Supplemental Fig. S2), combined with the expression profile of *KAR* is similar with *KASI*, *KASII* and *KASIII* (Fig. 9), these results reveal the consistency of condensation and reduction reaction.

FATB and *FATA*, two distinct thioesterase genes, are playing an essential role in the chain determination of fatty acid synthesis and in the channeling of carbon flux in higher plants. *FATB* encodes thioesterases preferring acyl-ACPs having saturated acyl groups, and *FATA* encodes the 18:1-ACP thioesterase [67]. In the seed development of *J. curcas*, mRNA expression of *FATA* occurred later than *FATB* (Fig. 8), consistent with the presumption of the ubiquitous 18: l-ACP thioesterase is a derivative of a 16:0 thioesterase, and the fatty acid component analysis indicated that the major fatty acids in *J. curcas* seed were the oleic, linoleic, palmitic and stearic fatty acid, in agreement with the previous report [44]. We found that the component of $C_{16:0}$ reached their maximum abundance at 25 DAF (Table 2) and decreased gradually at the late seed development of *J. curcas*. These results allow us to speculate that $C_{16:0}$ is firstly synthesized at the early seed development as the precursor for elongation to $C_{18:1}$ and $C_{18:2}$. SAD (spot C25) catalyzes the initial desaturation reaction in fatty acid biosynthesis and also plays an important role in determining the ratio of total saturated to unsaturated fatty acids in plants [68–70], and the expression profiles of *SAD* and *FATA* were clustered as the same group (Figs. 6, 8 and 9), the gradually increasing SAD abundance and $C_{18:1n9c}$ content during seed filling of *J. curcas* (Fig. 6 and Table 2) strongly support our theory on fatty acid elongation in *J. curcas* seeds. At the same time, it also suggests a low catalytic turnover of SAD as well as the highest stability of $C_{18:1}$ export from the plastids [25,71,72]. This result is different from the fatty acid biosynthesis in rapeseed, in which $C_{18:1}$ is the precursor for elongation to prominent $C_{22:1}$ [8]. Furthermore, reports showed that stearate levels were dramatically increased in rapeseeds of transgenic plants when *SAD* was knocked down in developing rapeseed embryos [73]. With the same genetic engineering techniques, it is possible to engineer specialized seed oil composition in *J. curcas*.

4.4. *Initiation of lipid accumulation in seed development*

As important carbon storage in many angiosperm seeds, TAG is usually synthesized by successive catalysis of a serial of enzymes in ER [74]. DGAT is the rate-limiting enzyme that acts in the final step of TAG synthesis for storage lipid accumulation in plants. DGAT1 and DGAT2, the two different main enzymes responsible for TAG synthesis, have been identified in several eukaryotic species. These two enzymes do not share high similarities on DNA or protein sequence, and play different roles in TAG synthesis in various tissues [75]. In the developing seed of *J. curcas*, we had not detected any differential protein species for DGAT; however, *DGAT1* and *DGAT2* were detected according to qRT-PCR analysis. The expressions of these two genes increased dramatically at the experimental stages and reached the maximum at 30 DAF (Fig. 7). These results suggest that lipid accumulation does not initiate before 30 DAF in the developing seeds of *J. curcas*. The synthesized TAG is usually stored in the small intracellular organelles called oil bodies at the late stage of seed development [6,76]. The major structural proteins of oil bodies are oleosin and caleosin [77], and their appearance is a prerequisite for subsequent lipid accumulation [78,79]. Based on their abundance and structures, caleosin as a less abundant protein plays limited biological role in the formation or degradation of seed oil bodies as compared with the structural role of relatively abundant oleosins [80]. Therefore, the oil body protein caleosin (spot F34) accumulated remarkably at 30 DAF (Figs. 4, 6 and Table 3), which may indicate the initiation of lipid accumulation. Additionally, expressions of *caleosin* and three *oleosins* isoforms had also been confirmed at the mRNA level. The expression of *oleosins* all upregulated in the seed development of *J. curcas* and showed similar pattern with *DGAT1*, *caleosin*, and *DGAT2* (Figs. 8 and 9). These results indicate that the lipid has accumulated in the *J. curcas* seeds at the late development. Importantly, the lipid content was less than 5% in developing seeds of *J. curcas* at 25 DAF, and lipids accumulated greatly in the endosperm of *J. curcas* and accounted for nearly 40% (Figs. 2 and 3) at the end stages of seed development of *J. curcas*. This notion on the regulation of mRNA and protein levels may offer mechanistic explanations of the cellular observation and lipid measurement.

5. Conclusions

This study performed a systematic analysis of metabolic processes from the perspective of protein abundance in the developing seed of *J. curcas*. Seed development is a complex process. We cannot expect to fully understand this process based on proteomics only, but the new information acquired here can support a platform to apply molecular methods to explore new protein species for seed development. Our results highlighted that several pathways, such as cytosolic and plastidial glycolysis, OPPP, TCA cycle, fatty acid synthesis, TAG synthesis and lipid formation were actively involved in the metabolic network from carbon flux to lipid accumulation in the seed development of *J. curcas*. These biochemical cascades were probably interacted by the collaboration of organelles in the developing seeds of *J. curcas*, such as plastid, mitochondrion, endoplasmic reticulum and oil body. In addition, signal transduction, ascorbate-glutathione system, disease responses, storage proteins and other unidentified protein species were also involved in this complex processes. Thus, further investigations of the seed proteome during the developmental process are necessary. The analysis of metabolic pathways with these data of identified protein species in the study provide the basis to understand the metabolic network of lipid accumulation in other woody oilseed plants, and certainly suggest some clues to improve the lipid content of *J. curcas* seeds.

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jprot.2013.06.030.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (31270653), and the Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-YW-G-035).

REFERENCES

[1] Bewley JD, Black M. Seeds: physiology of development and germination. New York: Plenum Press; 1994 .
[2] Chaudhury AM, Koltunow A, Payne T, Luo M, Tucker MR, Dennis ES, et al. Control of early seed development. Annu Rev Cell Dev Biol 2001;17:677–99.
[3] Kohler C, Makarevich G. Epigenetic mechanisms governing seed development in plants. EMBO Rep 2006;7:1223–7.
[4] Kang F, Rawsthorne S. Starch and fatty acid synthesis in plastids from developing embryos of oilseed rape (*Brassica napus* L.). Plant J 1994;6:795–805.
[5] Rawsthorne S. Carbon flux and fatty acid synthesis in plants. Prog Lipid Res 2002;41:182–96.
[6] Voelker T, Kinney AJ. Variations in the biosynthesis of seed-storage lipids. Annu Rev Plant Physiol Plant Mol Biol 2001;52:335–61.
[7] Shewry PR, Napier JA, Tatham AS. Seed storage proteins: structures and biosynthesis. Plant Cell 1995;7:945–56.
[8] Hajduch M, Casteel JE, Hurrelmeyer KE, Song Z, Agrawal GK, Thelen JJ. Proteomic analysis of seed filling in *Brassica napus*. Developmental characterization of metabolic isozymes using high-resolution two-dimensional gel electrophoresis. Plant Physiol 2006;141:32–46.
[9] Canovas FM, Dumas-Gaudot E, Recorbet G, Jorrin J, Mock HP, Rossignol M. Plant proteome analysis. Proteomics 2004;4: 285–98.
[10] Miernyk JA, Hajduch M. Seed proteomics. J Proteomics 2011;74:389–400.
[11] Dam S, Laursen BS, Ornfelt JH, Jochimsen B, Staerfeldt HH, Friis C, et al. The proteome of seed development in the model legume *Lotus japonicus*. Plant Physiol 2009;149:1325–40.
[12] Finnie C, Melchior S, Roepstorff P, Svensson B. Proteome analysis of grain filling and seed maturation in barley. Plant Physiol 2002;129:1308–19.
[13] Gallardo K, Le Signor C, Vandekerckhove J, Thompson RD, Burstin J. Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. Plant Physiol 2003;133:664–82.
[14] Hajduch M, Ganapathy A, Stein JW, Thelen JJ. A systematic proteomic study of seed filling in soybean. Establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. Plant Physiol 2005;137:1397–419.
[15] Vensel WH, Tanaka CK, Cai N, Wong JH, Buchanan BB, Hurkman WJ. Developmental changes in the metabolic protein profiles of wheat endosperm. Proteomics 2005;5: 1594–611.
[16] Houston NL, Hajduch M, Thelen JJ. Quantitative proteomics of seed filling in castor: comparison with soybean and rapeseed reveals differences between photosynthetic and nonphotosynthetic seed metabolism. Plant Physiol 2009;151: 857–68.
[17] Fairless D. Biofuel: the little shrub that could–maybe. Nature 2007;449:652–5.
[18] Liu H, Liu YJ, Yang MF, Shen SH. A comparative analysis of embryo and endosperm proteome from seeds of *Jatropha curcas*. J Integr Plant Biol 2009;51:850–7.
[19] Kandpal JB, Madan M. *Jatropha curcas*: a renewable source of energy for meeting future energy needs. Renew Energy 1995;6:159–60.
[20] King AJ, He W, Cuevas JA, Freudenberger M, Ramiaramanana D, Graham IA. Potential of *Jatropha curcas* as a source of renewable oil and animal feed. J Exp Bot 2009;60: 2897–905.
[21] Openshaw K. A review of *Jatropha curcas*: an oil plant of unfulfilled promise. Biomass Bioenergy 2000;19:1–15.
[22] Durrett TP, Benning C, Ohlroqqe J. Plant triacylglycerols as feedstocks for the production of biofuels. Plant J 2008;54: 593–607.
[23] Yang CY, Fang Z, Li B, Long YF. Review and prospects of *Jatropha* biodiesel industry in China. Renew Sustain Energy Rev 2012;16:2178–90.
[24] Pramanik A. Properties and use of *Jatropha curcas* oil and diesel fuel blends in compression ignition engine. Renew Energy 2003;28:239–48.
[25] Sarin R, Sharma M, Sinharay S, Malhotra RK. *Jatropha*-Palm biodiesel blends: an optimum mix for Asia. Fuel 2007;86:1365–71.
[26] Annarao S, Sidhu OP, Roy R, Tuli R, Khetrapal CL. Lipid profiling of developing *Jatropha curcas* L. seeds using (1)H NMR spectroscopy. Bioresour Technol 2008;99:9032–5.
[27] Li J, Li MR, Wu PZ, Tian CE, Jiang HW, Wu GJ. Molecular cloning and expression analysis of a gene encoding a putative beta-ketoacyl-acyl carrier protein (ACP) synthase III (KAS III) from *Jatropha curcas*. Tree Physiol 2008;28:921–7.
[28] Wu PZ, Li J, Wei Q, Zeng L, Chen YP, Li MR, et al. Cloning and functional characterization of an acyl–acyl carrier protein thioesterase (*JcFATB1*) from *Jatropha curcas*. Tree Physiol 2009;29:1299–305.
[29] Luo T, Peng SM, Deng WY, Ma DW, Xu Y, Xiao M, et al. Characterization of a new stearoyl-acyl carrier protein desaturase gene from *Jatropha curcas*. Biotechnol Lett 2006;28: 657–62.
[30] Gu K, Chiam H, Tian D, Yin Z. Molecular cloning and expression of heteromeric *ACCase* subunit genes from *Jatropha curcas*. Plant Sci 2011;180:642–9.
[31] Yang MF, Liu YJ, Liu Y, Chen H, Chen F, Shen SH. Proteomic analysis of oil mobilization in seed germination and postgermination development of *Jatropha curcas*. J Proteome Res 2009;8:1441–51.
[32] Costa GG, Cardoso KC, Del Bem LE, Lima AC, Cunha MA, de Campos-Leite L, et al. Transcriptome analysis of the oil-rich seed of the bioenergy crop *Jatropha curcas* L. BMC Genomics 2010;11:462.
[33] Natarajan P, Parani M. De novo assembly and transcriptome analysis of five major tissues of *Jatropha curcas* L. using GS FLX titanium platform of 454 pyrosequencing. BMC Genomics 2011;12:191.
[34] Natarajan P, Kanagasabapathy D, Gunadayalan G, Panchalingam J, Shree N, Sugantham PA, et al. Gene discovery from *Jatropha curcas* by sequencing of ESTs from normalized and full-length enriched cDNA library from developing seeds. BMC Genomics 2010;11:606.
[35] Sato S, Hirakawa H, Isobe S, Fukai E, Watanabe A, Kato M, et al. Sequence analysis of the genome of an oil-bearing tree, *Jatropha curcas* L. DNA Res 2010;18:65–76.
[36] Shen S, Jing Y, Kuang T. Proteomics approach to identify wound-response related proteins from rice leaf sheath. Proteomics 2003;3:527–35.
[37] Ramagli LS. Quantifying protein in 2-D PAGE solubilization buffers. Methods Mol Biol 1999;112:99–103.
[38] Liu H, Yang Z, Yang M, Shen S. The differential proteome of endosperm and embryo from mature seed of *Jatropha curcas*. Plant Sci 2011;181:660–6.

[39] Curto M, Valledor L, Nauarrete C, Gutierrez D, Sychrova H, Ramos J, et al. 2-DE based proteomic analysis of Saccharomyces cerevisiae wild and K+ transport-affected mutant (trk1,2) strains at the growth exponential and stationary phases. J Proteomics 2010;73:2316–35.
[40] Meunier B, Dumas E, Piec I, Bechet D, Hebraud M, Hocquette JF. Assessment of hierarchical clustering methodologies for proteomic data mining. J Proteome Res 2007;6:358–66.
[41] Valledor L, Jorrin J. Back to the basics: maximizing the information obtained by quantitative two dimensional gel electrophoresis analyses by an appropriate experimental design and statistical analyses. J Proteomics 2011;74:1–18.
[42] Sghaier-Hammami B, Valledor L, Drira N, Jorrin-Novo JV. Proteomic analysis of the development and germination of date palm (*Phoenix dactylifera* L.) zygotic embryos. Proteomics 2009;9:2543–54.
[43] Logemann J, Schell J, Willmitzer L. Improved method for the isolation of RNA from plant tissues. Anal Biochem 1987;163: 16–20.
[44] Akbar E, Yaakob Z, Kamarudin SK, Ismail M, Salimon J. Characteristic and composition of *Jatropha curcas* oil seed from Malaysia and its potential as biodiesel feedstock. Eur J Sci Res 2009;29:396–403.
[45] Bevan M, Bancroft I, Bent E, Love K, Goodman H, Dean C, et al. Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. Nature 1998;391:485–8.
[46] Ohlrogge JB, Jaworski JG. Regulation of fatty acid synthesis. Annu Rev Plant Physiol Plant Mol Biol 1997;48:109–36.
[47] White JA, Todd J, Newman T, Focks N, Girke T, de Ilarduya OM, et al. A new set of *Arabidopsis* expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil. Plant Physiol 2000;124:1582–94.
[48] Schwender J, Ohlrogge J, Shachar-Hill Y. Understanding flux in plant metabolic networks. Curr Opin Plant Biol 2004;7:309–17.
[49] Kleczkowski LA, Geisler M, Ciereszko I, Johansson H. UDP-glucose pyrophosphorylase. An old protein with new tricks. Plant Physiol 2004;134:912–8.
[50] Plaxton WC, Podesta FE. The functional organization and control of plant respiration. Crit Rev Plant Sci 2006;25:159–98.
[51] Harold FM. Inorganic polyphosphates in biology: structure, metabolism, and function. Bacteriol Rev 1966;30:772–94.
[52] ap Rees T, Green JH, Wilson PM. Pyrophosphate:fructose 6-phosphate 1-phosphotransferase and glycolysis in non-photosynthetic tissues of higher plants. Biochem J 1985;227:299–304.
[53] Plaxton WC. The organization and regulation of plant glycolysis. Annu Rev Plant Physiol Plant Mol Biol 1996;47:185–214.
[54] Jiang H, Wu P, Zhang S, Song C, Chen Y, Li M, et al. Global analysis of gene expression profiles in developing physic nut (*Jatropha curcas* L.) seeds. PLoS One 2012;7:e36522.
[55] Kanayama Y, Dai N, Granot D, Petreikov M, Schaffer A, Bennett AB. Divergent fructokinase genes are differentially expressed in tomato. Plant Physiol 1997;113:1379–84.
[56] Odanaka S, Bennett AB, Kanayama Y. Distinct physiological roles of fructokinase isozymes revealed by gene-specific suppression of *Frk1* and *Frk2* expression in tomato. Plant Physiol 2002;129:1119–26.
[57] Baud S, Lepiniec L. Physiological and developmental regulation of seed oil production. Prog Lipid Res 2010;49:235–49.
[58] Ruuska SA, Ohlrogge JB. The capacity of green oilseeds to utilize photosynthesis to drive biosynthetic processes. Plant Physiol 2004;136:2700–9.
[59] Schwender J, Goffman F, Ohlrogge JB, Shachar-Hill Y. Rubisco without the Calvin cycle improves the carbon effiency of developing green seeds. Nature 2004;432:779–82.
[60] Nicholls DJ, Miller J, Scawen MD, Clarke AR, Holbrook JJ, Atkinson T, et al. The importance of arginine 102 for the substrate specificity of *Escherichia coli* malate dehydrogenase. Biochem Biophys Res Commun 1992;189:1057–62.
[61] Musrati RA, Kollarova M, Mernik N, Mikulasova D. Malate dehydrogenase: distribution, function and properties. Gen Physiol Biophys 1998;17:193–210.
[62] Johnson JD, Mehus JG, Tews K, Milavetz BI, Lambeth DO. Genetic evidence for the expression of ATP- and GTP-specific succinyl-CoA synthetases in multicellular eucaryotes. J Biol Chem 1998;273:27580–6.
[63] Walker JE, Saraste M, Runswick MJ, Gay NJ. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J 1982;1:945–51.
[64] Tong L. Acetyl-coenzyme A, carboxylase: crucial metabolic enzyme and attractive target for drug discovery. Cell Mol Life Sci 2005;62:1784–803.
[65] Wu GZ, Xue HW. *Arabidopsis* beta-ketoacyl-[acyl carrier protein] synthase I is crucial for fatty acid synthesis and plays a role in chloroplast division and embryo development. Plant Cell 2010;22:3726–44.
[66] Clough RC, Matthis AL, Barnum SR, Jaworski JG. Purification and characterization of 3-ketoacyl-acyl carrier protein synthase III from spinach. A condensing enzyme utilizing acetyl-coenzyme A to initiate fatty acid synthesis. J Biol Chem 1992;267:20992–8.
[67] Jones A, Davies HM, Voelker TA. Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant Cell 1995;7:359–71.
[68] Shanklin J, Somerville C. Stearoyl-acyl-carrier-protein desaturase from higher plants is structurally unrelated to the animal and fungal homologs. Proc Natl Acad Sci U S A 1991;88:2510–4.
[69] McKeon TA, Stumpf PK. Purification and characterization of the stearoyl-acyl carrier protein desaturase and the acyl–acyl carrier protein thioesterase from maturing seeds of safflower. J Biol Chem 1982;257:12141–7.
[70] Thompson GA, Scherer DE, Foxall-Van Aken S, Kenny JW, Young HL, Shintani DK, et al. Primary structures of the precursor and mature forms of stearoyl-acyl carrier protein desaturase from safflower embryos and requirement of ferredoxin for enzyme activity. Proc Natl Acad Sci U S A 1991;88:2578–82.
[71] Knothe G. Dependence of biodiesel fuel properties on the structure of fatty acid alkyl esters. Fuel Process Technol 2005;86:1059–70.
[72] Jain S, Sharma MP. Oxidation stability of blends of Jatropha biodiesel with diesel. Fuel 2011;90:3014–20.
[73] Knutzon DS, Thompson GA, Radke SE, Johnson WB, Knauf VC, Kridl JC. Modification of *Brassica* seed oil by antisense expression of a stearoyl-acyl carrier protein desaturase gene. Proc Natl Acad Sci U S A 1992;89:2624–8.
[74] Thelen JJ, Ohlrogge JB. Metabolic engineering of fatty acid biosynthesis in plants. Metab Eng 2002;4:12–21.
[75] Turchetto-Zolet AC, Maraschin FS, de Morais GL, Cagliari A, Andrade CM, Margis-Pinheiro M, et al. Evolutionary view of acyl-CoA diacylglycerol acyltransferase (DGAT), a key enzyme in neutral lipid biosynthesis. BMC Evol Biol 2011;11:263.
[76] Tzen JT, Huang AH. Surface structure and properties of plant seed oil bodies. J Cell Biol 1992;117:327–35.
[77] Shimada TL, Hara-Nishimura I. Oil-body-membrane proteins and their physiological functions in plants. Biol Pharm Bull 2010;33:360–3.
[78] Frandsen GI, Mundy J, Tzen JT. Oil bodies and their associated proteins, oleosin and caleosin. Physiol Plant 2001;112:301–7.
[79] Naested H, Frandsen GI, Jauh GY, Hernandez-Pinzon I, Nielsen HB, Murphy DJ, et al. Caleosins: Ca^{2+}-binding proteins associated with lipid bodies. Plant Mol Biol 2000;44:463–76.
[80] Jiang PL, Tzen JT. Caleosin serves as the major structural protein as efficient as oleosin on the surface of seed oil bodies. Plant Signal Behav 2010;5:447–9.

Please cite this article as: Liu H, et al, Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.06.030