Differential expression analysis of transcripts related to oil metabolism in maturing seeds of *Jatropha curcas* L.

Divya Chandran · H. M. Sankararamasubramanian ·
M. Ashok Kumar · Ajay Parida

Received: 16 September 2013 / Revised: 14 December 2013 / Accepted: 15 December 2013 / Published online: 5 March 2014
© Prof. H.S. Srivastava Foundation for Science and Society 2014

Abstract *Jatropha curcas* has been widely studied at the molecular level due to its potential as an alternative source of fuel. Many of the reports till date on this plant have focussed mainly on genes contributing to the accumulation of oil in its seeds. A suppression subtractive hybridization strategy was employed to identify genes which are differentially expressed in the mid maturation stage of *J. curcas* seeds. Random expressed sequence tag sequencing of the cDNA subtraction library resulted in 385 contigs and 1,428 singletons, with 591 expressed sequence tags mapping for enzymes having catalytic roles in various metabolic pathways. Differences in transcript levels in early and mid-to-late maturation stages of seeds were also investigated using sequence information obtained from the cDNA subtraction library. Seven out of 12 transcripts having putative roles in central carbon metabolism were up regulated in early seed maturation stage while lipid metabolism related transcripts were detected at higher levels in the later stage of seed maturation. Interestingly, 4 of the transcripts revealed putative alternative splice variants that were specifically present or up regulated in the early or late maturation stage of the seeds. Transcript expression patterns from the current study using maturing seeds of *J. curcas* reveal a subtle balancing of oil accumulation and utilization, which may be influenced by their energy requirements.

Keywords Biodiesel · Suppression subtractive hybridization (SSH) · Seed Oil Metabolism · *Jatropha curcas*

Introduction

The quest for alternative sources of fuel has gained impetus in recent times due to the dwindling fossil fuel reserves and the environmental concerns associated with the use of fossil fuel. Green plants store energy in seeds in the form of carbohydrates, proteins and lipids, making them attractive targets for the exploration of renewable and sustainable sources of cleaner fuel. However, there are questions (Tenenbaum 2008) related to food security if food crops are to be suggested for the production of biofuel. Furthermore, the process of biofuel production should consume less water and be highly energy efficient, starting from the cultivation of related crops to fuel burning in engines (Costa et al. 2010). *Jatropha curcas* offers a viable alternative in this context as it is not used for edible purposes, is grown only as a fencing plant and is well-adapted to varying soil conditions. Seeds of *J. curcas* accumulate about 35–40 % of oil (Achten et al. 2008) in the form of triacylglycerol (TAG) and it varies according to the conditions of growth and agronomic practices (Openshaw 2000). In spite of its potential, the major bottleneck for the improvement of *J. curcas* for yield and seed quality traits is the narrow genetic diversity of the germplasm (Basha and Sujatha 2007, 2009; Popluechai et al. 2009; Wen et al. 2010; Yi et al. 2010). Molecular breeding to improve seed oil yield can circumvent the problem of narrow genetic base, if genes contributing to oil metabolism in the seeds are identified.

Over the last few years we have witnessed significant advances in metabolic engineering through the application of genomics and proteomics to gain a holistic understanding of metabolic pathways (Capell and Christou 2004). There are many examples to show the potential use of genomics to understand oil biosynthetic pathway and further transformation for increased oil content or composition in other oil seed crops especially, *Brassica* (Li et al. 2006). Metabolically specialised enzymes pertaining to oil biosynthesis in

D. Chandran · H. M. Sankararamasubramanian (✉) ·
M. A. Kumar · A. Parida
M S Swaminathan Research Foundation, 3rd Cross Road,
Institutional Area, Taramani, Chennai 600113, India
e-mail: justforscience@gmail.com

J. curcas can be identified easily through EST sequencing of cDNAs prepared from seeds because mRNA clones corresponding to these enzymes could be relatively abundant. Large scale single-pass sequencing of cDNAs from different plants has provided a large repository of genes whose tissue-specific expression can be evaluated for the cloning of genes in addition to the development of markers for map-based cloning (White et al. 2000). In *J. curcas*, volumes of genomic data that include the whole genome (Sato et al. 2011) and seed transcriptome information have been generated in the last few years (King et al. 2011). A total of 12,084 expressed sequence tags (ESTs), from developing *J. curcas* seeds were sequenced and they represented genes coding for diverse biological functions (Natarajan et al. 2010). Costa et al. (2010) reported EST sequencing based identification of genes related to oil accumulation and degradation using developing and germinating seeds of *J. curcas*. Global gene expression profiles of developing *J. curcas* seeds revealed key genes involved in the synthesis of storage lipids and proteins (Jiang et al. 2012).

Despite the availability of such comprehensive gene expression data, our knowledge on the regulation of oil metabolism in general and *J. curcas* in particular is still very limited. Seed oil content is governed by quantitative trait loci (QTLs) like any other economically important trait (Liu et al. 2011) and other contributing factors (Baud and Lepiniec 2010) including the environmental influence on the metabolism of oil in seeds. The pathway from the assimilation of photosynthates to the formation of oil in seeds could involve subtle regulation of genes that are sensitive to environmental cues, which eventually decides the fate of photosynthates. Moreover, the importance of utilisation of energy stored in lipids as the seed matures and therefore, possibly decreases in oil quantity also need to be studied. In other words, to increase oil yield in *J. curcas*, modulating genes downstream of final product of photosynthesis could be the right approach. In order to do this, expression levels of genes at specific stages of seed maturation in which photosynthates (source) are directed towards oil (sink) metabolism in combination with genes that regulate the utilisation of sink need to be clearly investigated. Here we report a suppression subtractive hybridization (SSH) based approach to identify differentially expressing genes in different maturation stages of *J. curcas* seeds that correspond to active oil metabolism. SSH is a powerful technique that enriches differentially expressing genes (Diatchenko et al. 1996) and has been used to study seed development related genes in *Phaseolus vulgaris* (Abid et al. 2012). In *Brassica napus*, differentially expressing genes in seeds of two near-isogenic lines having different oil content were investigated using SSH. However, SSH has not been exploited thus far to study the differentially expressed genes in maturing *J. curcas* seeds. We also monitored the transcript expression pattern of a few central carbon metabolism, lipid metabolism and seed development related genes represented

in the SSH library, in early and mid-to-late maturation stages of *J. curcas* seeds to understand transcript level regulation of oil metabolism.

Materials and methods

Seed material

Jatropha curcas fruits were collected from Kudankulam (8°11'0" N 77°41'0" E), Tamil Nadu, India. The fruits and seeds were segregated into four categories, based on morphological parameters like size and colour (Fig. 1a): early maturation stage (JC1), early-to-mid maturation stage (JC2), mid-to-late maturation stage (JC3) and late maturation stage (JC4).

RNA extraction

RNA extraction was done using different maturation stages of *J. curcas* seeds. Total RNA was extracted using a CTAB-guanidine hydrochloride method. Five grams of seeds were ground into a fine powder using liquid nitrogen. Thirty milliliter of hot CTAB (65 °C) and 100 µl of β-mercaptoethanol were added to the powder, mixed well and incubated at 65 °C for 15 min. To this 30 ml of 8 M guanidine hydrochloride was added, thoroughly mixed by shaking and centrifuged at 10,000 g for 15 min. The resultant supernatant was processed as described by Chang et al. (1993). The quality of RNA was verified using a UV-Visible spectrophotometer (Shimadzu UV-1601). mRNA purification was done using NucleoTrap mRNA purification kit (Machery-Nagel) with 5 mg of total RNA (Fig. 1b) following the manufacturer's protocol. DNase treatment of mRNA was performed according to the manufacturer's recommendations (Ambion TURBO DNase) before cDNA synthesis was carried out.

SSH library preparation and EST sequencing

Two micrograms of mRNA were used from each seed maturation stage (JC2 and JC3) as starting material for cDNA synthesis. cDNA subtraction was done using Clontech's Suppression Subtractive Hybridization (SSH) kit following the manufacturer's protocol (Diatchenko et al. 1996) with 33 and 21 cycles for primary and secondary PCR respectively. Forward subtraction was done using cDNA synthesised from seeds categorized as 'tester' from mid-to-late maturation stage (JC3) and as 'driver' from early-to-mid maturation (JC2) stage (Fig. 1c). cDNA library was first cloned into InstAclone T-vector (MBI Fermentas) and later used to transform chemically competent *E. coli* (OmniMAX 2 T1 Phage-Resistant, Invitrogen). The resultant library was plated in ampicillin selection medium, grown

Fig. 1 Suppression Subtractive Hybridization (SSH) library preparation from *J. curcas* seeds.

a Morphological categorization of *J. curcas* seed maturation stages (JC1- early, JC2-early-to-mid, JC3-mid-to-late and JC4-late); box indicates stages used in subtraction (JC3 subtracted from JC2). **b** Ethidium bromide stained formaldehyde agarose gel of total RNA from JC2 and JC3 seeds extracted by CTAB-guanidine hydrochloride method. **c** Unsubtracted (US) and subtracted (S) double stranded cDNAs. **d** Colony PCR of subtracted clones to verify insert size. **e** Plasmid dot blot of SSH library clones hybridized with forward subtracted cDNA probes. **f** Replica dot blot of (e) hybridized with reverse subtracted cDNA probes. Arrows indicate differentially expressed clones

overnight and colonies from individual plates were scrapped on to LB liquid medium and plasmid DNA was purified using alkali-lysis method (Birnboim and Doly 1979). One microlitre from different plasmid pools was used for retransformation of *E. coli* before isolating individual plasmids for EST sequencing (BigDyeTM Terminator v3.1 Cycle Sequencing Kit in 3130xl Genetic Analyzer, Applied Biosystems) of the inserts using M13 reverse primer. More than five thousand clones across plasmid pools were sequenced.

Sequence analysis

Vector sequences were removed using *VecScreen* (<http://www.ncbi.nlm.nih.gov/VecScreen/>) and ESTs were submitted to NCBI's EST database. ESTpiper (Tang et al. 2009) was used

to perform contig analysis. Homology search, annotation and pathway mapping of the ESTs were carried out using Blast2GO (Conesa et al. 2005). EST sequence data (3513) was submitted to NCBI's dbEST database. The sequences have been published with the accession number starting from JK610133.1 to JK613645.1.

Differential screening of SSH library

Clones for which the EST information was already obtained through *blastX* analysis were randomly selected for the verification of differential expression. Fifty nanograms of plasmid DNAs representing these ESTs were denatured at 95 °C for 5 min and dot-blotted in duplicates. Forward and reverse subtracted cDNAs were freshly prepared with JC2 cDNA serving as 'tester'

and JC3 cDNA serving as the ‘driver’ for the reverse subtraction. Fifty nanograms of the subtracted cDNAs were α -p32-dCTP labelled using *Rediprime II* DNA labelling Kit (GE Healthcare), denatured at 95 °C and hybridized separately against duplicate blots of forward subtracted library clones. Hybridization was performed at 60 °C overnight and the blots were washed twice in 2XSSC with 0.1 % SDS solution at 60 °C for 10 min and 0.5XSSC with 0.1 % SDS solution once at 60 °C for 10 min before autoradiography.

Semi-quantitative reverse transcription (RT) PCR

A total of 21 genes putatively representing central carbon metabolism (CCM), lipid metabolism and seed development were selected from the SSH library after differential screening to study the differences in transcript expression in the early maturation (JC1) and mid-to-late maturation (JC3) stages of *J. curcas* seeds. Actin from *J. curcas* was used as an endogenous control. Primers for the 21 selected genes were designed using sequence information of ESTs given in Table 1. Regions in the sequence that showed high identities to *Ricinus communis* were considered for primer design. Total RNA was extracted from stages JC1 and JC3 as described earlier. DNase treatment was done using MBI Fermentas kit and later purified by NucleoSpin RNA Cleanup Kit (Machery – Nagel) as per the manufacturer’s instructions. RNA integrity was checked on a formaldehyde gel and the concentration and purity of RNA was checked using Nanodrop-2000 (Thermo Scientific). Two micrograms of total RNA were used as template to synthesize first strand cDNA using Takara’s BluePrint RT-PCR Kit (Cat #RR714A). PCRs specific for individual genes were performed using Ex-*Taq* HS from the same kit, following the same cycling parameters suggested by the manufacturer. The number of PCR cycles and annealing temperatures varied according to genes mentioned in Table 1. To exclude the possibility of DNA contamination, genomic DNA isolated from *J. curcas* leaves and first strand cDNA of JC1 and JC3 were amplified using actin primers. The PCR products were visualised on 1.4 or 2 % TAE agarose gels.

Results

SSH library

In order to identify differentially expressing genes pertaining to oil metabolism in different seed maturation stages of *J. curcas*, a PCR based cDNA subtraction strategy was followed. JC2 and JC3 stages, based on our morphological observation and earlier reports (Jiang et al. 2012) corresponded to early- to- mid maturation and mid- to- late maturation stages respectively. Total RNA purified from both these stages using

a modified CTAB method was of good quality (Fig. 1b) when observed on a formaldehyde agarose gel (1 %). Distinct size differences (Fig. 1c) between unsubtracted and subtracted cDNAs along with results from differential screening (Fig. 1e and f) indicate the efficiency of subtraction. Insert size of the subtracted clones ranged between 200 bp to 1 kb (Fig. 1d).

EST analysis and pathway mapping

Random sequencing of ESTs from a plasmid library resulted in 3513 quality reads (NCBI-dbEST accession numbers JK613645.1 to JK610133.1). Contig analysis using CAP3 (Huang and Madan 1999) program through EST piper (Tang et al. 2009) resulted in 385 contigs and 1,428 singletons with 2,085 ESTs forming the contigs. Senescence associated proteins formed the biggest contig. Functional annotation and pathway mapping of the ESTs were performed using Blast2GO program (Conesa et al. 2005). ESTs were categorized based on biological processes in which they were involved (Fig. 2), molecular functions (Fig. 2) and as cellular components (Fig. 2). Catalytic activity (26.03 %) and binding (28.57 %) constituted major molecular functions while metabolic and cellular (34.14 %) processes constituted major biological processes represented by the ESTs. cDNAs related to cytoplasmic (36.51 %), organellar (28.68 %) and membrane (19.36 %) proteins were the major cellular components. Metabolic pathway mapping of ESTs using KEGG (Kanehisa and Goto 2000) through Blast2GO resulted in 591 sequences that performed catalytic roles in different metabolic pathways. Twenty four percent of the ESTs mapped to the central carbon metabolism. Aminoacid metabolism was represented by 19 % of the ESTs while fatty acid and lipid metabolism related ESTs comprised of 9 % of the ESTs mapped (Fig. 3).

Semi-quantitative RT-PCR

Expression levels of 21 transcripts representing different levels of seed maturation and metabolism were monitored in early maturation stage (JC1) and mid-to-late maturation stage (JC3). Nine transcripts were up regulated in JC3 while 7 were up regulated in JC1 (Fig. 4). Three transcripts each were specific to JC1 and JC3 respectively. Transcript level for actin (GenBank accession HM044307.1) remained same in both the stages. Interestingly, four transcripts revealed the presence of an additional transcript in one or both the stages of seed maturation (Fig. 4f, h, l and t).

Expression of genes related to seed development

Seed maturation protein (GenBank accession JK610508) also revealed the presence of two transcripts in JC3. The

Table 1 List of genes used in semi quantitative RT-PCR and the PCR conditions

Gene	Genbank accession No.	Primer sequences (F- Forward & R- Reverse)	Annealing temperature in °C (Number of PCR Cycles)	Expected product size (bp)
Sucrose synthase -1	JK612395	F 5'- CGTGCTAAAGGATCGCAACAAGCC - 3' R 5'- CCCCATTCTCACTCGGTTCATCTG - 3'	58 (24)	272
Sucrose synthase -2	JK610705	F 5'- CGATGGCTAGGCTGGATCGG - 3' R 5'- CCAGCTACCACAACAAGGTTTCGC - 3'	58 (27)	103
UDP-glycosyltransferase	JK611904	F 5'- CCATTCTTGAAAAGCCTCATGCAGTTTG - 3' R 5'- GGAAGGAAGGAAGGCCAATGACAG - 3'	55 (33)	189
Pyruvate kinase	JK613331	F 5'- GTCAATCTTCTGGCGTTGTGGTG - 3' R 5'- CCACGGGAACCATGAATGAGTC - 3'	55 (27)	263
Citrate synthase	JK611304	F 5'- CCTATCTCGACTGCACGAGTTC - 3' R 5'- GCACAGCCTCATTGCTCCAC - 3'	55 (33)	182
Succinate dehydrogenase	JK611506	F 5'- CATGCAGGAAGGGATATGTGGG - 3' R 5'- CTGTATCCCTGCCATCCTCAGG - 3'	55 (33)	235
Malate dehydrogenase	JK610690	F 5'- ATGGAGGAGGGTTGCAGTCTG - 3' R 5'- ACAGCACCTGTGTCCATATGGC - 3'	55 (30)	216
Malate synthase	JK613240	F 5'- GGTTCAGTTGGCATGACCCAAC - 3' R 5'- ATAAGTCCTGGGTGTGCTGCC - 3'	55 (33)	225
Isocitrate lyase	JK611473	F 5'- GTTGCAAGAACTGATGCTGTGCGC - 3' R 5'- GTTGCGCCATTGCGAGCC - 3'	55 (30)	199
Phosphoenol pyruvate carboxykinase [Pep carboxy kinase]	JK611690	F 5'- GGAGTTCTCCACCTGTGAGC - 3' R 5'- CCATAGCTGCCACCTGACCAG - 3'	55 (30)	251
Fructose 1,6 bisphosphatase	JK612263	F 5'- GATCACGCAGCCGATGCTTACC - 3' R 5'- GGATGCATGTTGCCCCACTGC - 3'	55 (30)	286
Aldose 1- epimerase	JK611352	F 5'- CTGGATCATGGAAGGAAACGACATATG - 3' R 5'- CAGAACTGGAAGGCACAACCTGAGA - 3'	55 (33)	230
Beta-ketoacyl- <i>acp</i> synthase iii [Kas III]	JK610957	F 5'- GCTCTTCAGATGGCAGAGGTTGAAC - 3' R 5'- CCTCCCCTAATGTGACAAGCAGC - 3'	58 (33)	203
Oleosin 2	JK612128	F 5'- GCCCTGTTCTTGTCCCAGCTG - 3' R 5'- CCCAGCTCTCTGACTTCAACCC - 3'	55 (30)	300
Glycerol-3-phosphate 3-phosphatidyl transferase [G3p-3phosphatidyl transferase]	JK611945	F 5'- GTGTTGCGAGATGTTGCCCTTGTTG - 3' R 5'- CTGTCGTGGCAGCCACTAACC - 3'	55 (30)	259
Lipid transfer protein	JK611370	F 5'- GCTTCAGCTTCCTTCTGCTTGTCAG - 3' R 5'- AATGGCCGCAGCAATCGC - 3'	55 (27)	226
Lipid binding protein	JK613519	F 5'- GCATTTGCAAGCATGTCACACCG - 3' R 5'- CATGTGACTCCCACTTGGTTCC - 3'	58 (33)	120
NAC transcription factor	JK611375	F 5'- CACTACAGAAGGTCGAAAGTTGG - 3' R 5'- GAGGAGGAACAAGAACCATTGTTGC - 3'	55 (30)	252
α -Expansin 6	JK611284	F 5'- GGTGCAGGAGATGTTACAGCAG - 3' R 5'- GACCAAAGGACCATCCAGCAGG - 3'	55 (30)	193
Seed maturation protein	JK610508	F 5'- GGCGAGCAGCAGTTTCCTC - 3' R 5'- GTGAAGGCCACAGTTGCTCC - 3'	55 (33)	206
Senescence associated protein	JK612019	F 5'- CGAACCGTGAAAGCGTGGTC - 3' R 5'- GTCTCACGACGGTCTAATCCCAG - 3'	55 (30)	234
Actin	HM044307.1	F 5'- GAGCAAGGAGATCACTGCACTTGC - 3' R 5'- GGACCAGACTCATCACTCGCCC - 3'	58 (35)	162

shorter transcript was specific to JC3 while the longer transcript was present in both JC1 and JC3 but up regulated in JC1 (Fig. 4t). Transcripts for expansin (GenBank accession JK611284) and senescence

associated protein (GenBank accession JK612019) showed higher levels of expression in JC3 (Fig. 4s and u). cDNA product for NAC transcription factor transcript (GenBank accession JK611375) was detected

Fig. 2 Categorization of EST sequences using Blast2Go (Conesa et al. 2005) at gene ontology level 2

at very high levels in JC1 and at very low levels in JC3 (Fig. 4r).

Expression of genes related to central carbon metabolism

Sucrose metabolism

Levels of transcripts representing two sucrose synthase ESTs (GenBank accession JK612395 and JK610705) were different in the two seed maturation stages. Expression of both these transcripts was up regulated in JC1 (Fig. 4a and b). However, the differential expression of one of them was detected early at 24 cycles as compared to the other that was detected only at 27 cycles. Expression of UDP-glycosyl transferase (GenBank accession JK611904) mRNA was also up regulated in JC1 (Fig. 4c).

Glucose metabolism

The longer transcript representing the aldose 1 epimerase (GenBank accession JK611352) was up regulated in JC1 while the shorter transcript of the same was up regulated in JC3 (Fig. 4l). Transcript for the glycolytic enzyme pyruvate kinase (GenBank accession JK613331) was up regulated in JC3 (Fig. 4d). Expression pattern for mRNA representing the Krebs cycle enzyme succinate dehydrogenase revealed two differently sized bands, each specific for JC1 and JC3 (Fig. 4f). The level of shorter transcript expressed only in JC3 was higher as compared to that of the longer transcript observed in JC1. Two more transcripts representing Krebs cycle enzymes malate dehydrogenase (GenBank accession JK610690) and citrate synthase (GenBank accession JK611304) showed up regulation in JC1 up to 30 and 33 cycles of PCR respectively (Fig. 4g and e). Transcript levels for

Fig. 3 Categorization of ESTs based on KEGG (Kanehisa and Goto 2000) mapping

four the gluconeogenesis related enzymes were also followed in JC1 and JC3 maturation stages. Level of fructose 1, 6 biphosphatase (GenBank accession JK612263) transcript was more in JC3 (Fig. 4k). However, the expression of transcripts for isocitrate lyase (GenBank accession JK611473)

and phosphoenolpyruvate carboxykinase (GenBank accession JK611690) was highly up regulated only in JC3 (Fig. 4i and j). Two bands were observed for malate synthase (GenBank accession JK613240) transcript in JC3 out of which the shorter transcript was specific to JC3 (Fig. 4h). The longer

Fig. 4 Semi-quantitative reverse transcription PCR of genes in early maturation stage (JC1) and mid-to-late maturation stage (JC3) of *J. curcas* seeds. **a–i** Genes related to central carbon metabolism, **m–q** genes related to lipid metabolism, **r–u** genes related to seed development. Genes and corresponding accession numbers are detailed in Table 1. **v** Amplification product at the end of 35 PCR cycles for *J. curcas* actin using leaf DNA (LD) as template and first strand cDNA as template for JC1 and JC3 (to test for DNA contamination). Representative figure of three independent experiments

fragment was up regulated only in the mid-to-late maturation stage or JC3 (Fig. 4h).

Expression of genes related to lipid metabolism

All the transcripts that were monitored for their expression levels linked to lipid metabolism namely beta ketoacyl-ACP synthase III (GenBank accession JK610957); oleosin2 (GenBank accession JK612128); glycerol 3-phosphate 3-phosphatidyl transferase (GenBank accession JK611945); lipid transfer protein (GenBank accession JK611370) and lipid binding protein (GenBank accession JK613519) were up regulated in JC3 (Fig. 4m–q).

Discussion

Seed maturation is characterized by the accumulation of storage compounds in tissues, acquisition of dormancy and desiccation tolerance by the embryo, leading to a developmental endpoint in seed coat and endosperm where storage reserves are directed towards the process of germination (Baud and Lepiniec 2010). Since seed oil metabolism is not restricted to the embryo or cotyledonous tissue and is intricately linked to the developmental and metabolic states of other tissues in seeds (Bourgis et al. 2011), we used whole seeds belonging to two mid maturation stages for cDNA subtraction and early and mid-to-late maturation stages to monitor changes in transcript expression. A large number of ESTs were sequenced to obtain a good representation of differential transcripts, since the original library that contained about 2,000 clones was stored as plasmid pools and later used for retransformation to isolate individual clones. SSH library prepared in this study revealed a good representation of seed specific transcripts, particularly to those of metabolism and cellular processes (Fig. 2) underlining the importance of both in contributing to overall seed development (Weber et al. 2008; Wobus and Weber 1999), which in turn determines the accumulation and fate of oil.

Transcripts related to seed development

Expression of transcripts related to growth and development in early and mid-to-late seed maturation stages were monitored to understand the association of seed development and oil metabolism. Expansins (Chen and Bradford 2000) and seed maturation proteins (Rosenberg and Rinne 1988) have been found to influence seed development and the progression of a seed to seedling stage. Higher levels of expansin transcript in JC3 in *J. curcas* seeds may be linked to cell expansion and increase in seed size because in wheat (Lizana et al. 2010) expression of expansin genes correlated with grain size. Seed maturation protein transcript revealed two bands in the RT-

PCR in JC3. Since DNA contamination of the RNA samples was ruled out due to the absence of longer actin transcript (Fig. 4v), detection of two bands in either of the stages (JC1 or JC3) or both for four transcripts tested in the present study may be attributed to the presence of alternatively spliced variants. Alternative splicing of mRNA is an important component of post-transcriptional regulation in eukaryotes and can lead to tissue-specific expression patterns (Sugliani et al. 2010) and seed development (Penfield et al. 2010). Presence of a shorter transcript for seed maturation protein only in JC3 and up regulation of a longer transcript in JC1 clearly suggest the importance of the two putative splice variants in seed development and oil metabolism in *J. curcas* seeds. A complex network of transcription factors control gene expression programs necessary for seed maturation (Baud and Lepiniec 2010). In this context, detection of very high transcript levels for putative NAC family transcription factor in JC1 indicates the importance of the same in regulating early embryogenesis as reported in *Arabidopsis* (Kunieda et al. 2008). Higher representation of transcripts related to senescence associated proteins may be attributed to the abundance of this transcript (Fig. 4u) in later stages of seed maturation (JC3) selected for subtraction in this study. A higher level of its transcript expression observed in JC3 as compared to JC1 (Fig. 4u) also suggests such a possibility. Our results in the present study are similar to that observed in soybean seed development (Jones and Vodkin 2013), where senescence associated proteins were found to be abundant.

Transcripts related to central carbon metabolism

We wanted to study the transcript expression pattern of a few central carbon metabolism (CCM) genes as they represented 24 % of the ESTs sequenced from the SSH library. They are also critical for plant growth and development in addition to being responsible for the production of accessible energy and the creation of primary building blocks of other metabolisms (Zhang et al. 2010). Up regulation of both the putative sucrose synthase transcripts identified in this study during early maturation (JC1 stage) suggests the importance of sucrose synthase (King et al. 1997; Li et al. 2006) in providing carbon for oil biosynthesis. Higher transcript level for the putative UDP-glycosyl transferase only in JC1 also indicates a key role for the corresponding enzyme in transferring (Bowles et al. 2005) sugar moieties from activated sugars to acceptor molecules like proteins, lipids and secondary metabolites, particularly during early seed maturation stages. Expression of seven out of twelve genes putatively involved in central carbon metabolism was up regulated in mid-to-late maturation stage (JC3). Interestingly, some of them also had putative splice variants that were either up regulated or

present only in JC1, as in the case of succinate dehydrogenase (Fig. 4f) and aldose-1-epimerase (Fig. 4l). Contrary to earlier reports in *J. curcas* (Costa et al. 2010; King et al. 2011; Natarajan et al. 2010; Jiang et al. 2012; Xu et al. 2011; Chen et al. 2011) which focussed primarily on accumulation of oil, our results suggest a complex regulation of oil metabolism, possibly involving a delicate balance between diversion of carbon to lipid biosynthesis and catabolism of storage reserves to meet energy demands during seed maturation in *J. curcas*. For example, up regulation of transcripts for glyoxylate cycle specific enzymes like malate synthase and gluconeogenesis related enzymes like fructose 1, 6 bisphosphatase and phosphoenolpyruvate carboxykinase in JC3 suggests a possible breakdown of lipids for energy. However, it is also possible that there is a constant withdrawal of carbon from the Krebs cycle which is anaplerotically met by the two glyoxylate cycle enzymes (Eastmond and Graham 2001). Up regulation of gluconeogenesis related transcripts may be a result of protein or oil utilisation to generate energy for seed development or maturation as was observed in *Brassica napus* (Chia et al. 2005) embryos where catabolism was active in embryos during and after the main period of oil accumulation.

Transcripts related to lipid metabolism

Storage lipid accumulation generally occurs at the mid-to-late seed developmental stage (Hills 2004) in oilseeds, which is also the case in *J. curcas* (Xu et al. 2011). In the current study where JC3 corresponds to mid-to-late maturation stage, all the transcripts associated with lipid biosynthesis were up regulated (Fig. 4m–q), confirming earlier observations (Costa et al. 2010; King et al. 2011; Jiang et al. 2012; Xu et al. 2011; Chen et al. 2011) made in the same plant. Our results from transcript expression of lipid metabolism related genes also suggest that accumulation of oil could take place simultaneously with its utilisation as the seed matures in *J. curcas*.

Conclusion

Differential expression analysis of seed-specific transcripts in our study clearly indicates the complexity of oil metabolism in *J. curcas*, where regulation of oil biosynthesis may be linked to its breakdown to meet the energy requirements of a developing seed. Therefore, it is necessary to further investigate the functional role of the genes identified in the present study in an oil metabolic network to improve oil quantity in *J. curcas* seeds through molecular breeding methods.

Acknowledgments We acknowledge the Department of Biotechnology, Government of India for funding the project and providing Senior Research Fellowship to Divya Chandran.

References

- Abid G, Sassi K, Muhovski Y et al (2012) Identification and analysis of differentially expressed genes during seed development using suppression subtractive hybridization (SSH) in *Phaseolus vulgaris*. *Plant Mol Biol Rep* 30:719–730
- Achten WMJ, Verchot L, Franken YJ et al (2008) *Jatropha* bio-diesel production and use. *Biomass Bioenergy* 32:1063–1084
- Basha SD, Sujatha M (2007) Inter and intra-population variability of *Jatropha curcas* (L.) characterized by RAPD and ISSR markers and development of population-specific SCAR markers. *Euphytica* 156:375–386
- Basha SD, Sujatha M (2009) Genetic analysis of *Jatropha* species and interspecific hybrids of *Jatropha curcas* using nuclear and organelle specific markers. *Euphytica* 168:197–214. doi:10.1007/s10681-009-9900-0
- Baud S, Lepiniec L (2010) Physiological and developmental regulation of seed oil production. *Prog Lipid Res* 49:235–249. doi:10.1016/j.plipres.2010.01.001
- Bimboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* 7(6): 1513–1523
- Bourgis F, Kilaru A, Cao X et al (2011) Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. *Proc Natl Acad Sci U S A* 108: 12527–12532. doi:10.1073/pnas.1106502108
- Bowles D, Isayenkova J, Lim EK, Poppenberger B (2005) Glycosyltransferases: managers of small molecules. *Curr Opin Plant Biol* 8(3):254–263
- Capell T, Christou P (2004) Progress in plant metabolic engineering. *Curr Opin Biotechnol* 15:148–154. doi:10.1016/j.copbio.2004.01.009
- Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Mol Biol Rep* 11:113–116. doi: 10.1007/BF02670468
- Chen F, Bradford KJ (2000) Expression of an expansin is associated with endosperm weakening during tomato seed germination. *Plant Physiol* 124:1265–1274
- Chen M-S, Wang G-J, Wang R-L et al (2011) Analysis of expressed sequence tags from biodiesel plant *Jatropha curcas* embryos at different developmental stages. *Plant Sci* 181:696–700. doi:10.1016/j.plantsci.2011.08.003
- Chia TYP, Pike MJ, Rawsthorne S (2005) Storage oil breakdown during embryo development of *Brassica napus* (L.). *J Exp Bot* 56:1285–1296. doi:10.1093/jxb/eri129
- Conesa A, Götz S, García-Gómez JM et al (2005) Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676. doi:10.1093/bioinformatics/bti610
- Costa GGL, Cardoso KC, Del Bem LE V et al (2010) Transcriptome analysis of the oil-rich seed of the bioenergy crop *Jatropha curcas* L. *BMC Genomics* 11:462. doi:10.1186/1471-2164-11-462
- Diatchenko L, Laut YC, Campbell AP et al (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci U S A* 93:6025–6030
- Eastmond PJ, Graham IA (2001) Re-examining the role of the glyoxylate cycle in oilseeds. *Trends Plant Sci* 6:72–78
- Hills MJ (2004) Control of storage-product synthesis in seeds. *Curr Opin Plant Biol* 7(3):302–308

- Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. *Genome Res* 9:868–877
- Jiang H, Wu P, Zhang S et al (2012) Global analysis of gene expression profiles in developing physic nut (*Jatropha curcas* L.) seeds. *PLoS ONE* 7:e36522
- Jones SI, Vodkin LO (2013) Using RNA-Seq to profile soybean seed development from fertilization to maturity. *PLoS ONE* 8:e59270. doi:10.1371/journal.pone.0059270
- Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30
- King SP, Lunn JE, Furbank RT (1997) Carbohydrate content and enzyme metabolism in developing canola siliques. *Plant Physiol* 114:153–160
- King AJ, Li Y, Graham IA (2011) Profiling the developing *Jatropha curcas* L. Seed transcriptome by pyrosequencing. *Bioenerg Res* 4: 211–221
- Kunieda T, Mitsuda N, Ohme-Takagi M et al (2008) NAC family proteins NARS1/NAC2 and NARS2/NAM in the outer integument regulate embryogenesis in *Arabidopsis*. *Plant Cell* 20:2631–2642. doi:10.1105/tpc.108.060160
- Li R-J, Wang H-Z, Mao H et al (2006) Identification of differentially expressed genes in seeds of two near-isogenic *Brassica napus* lines with different oil content. *Planta* 224:952–962. doi:10.1007/s00425-006-0266-4
- Liu P, Wang CM, Li L et al (2011) Mapping QTLs for oil traits and eQTLs for oleosin genes in *Jatropha*. *BMC Plant Biol* 11:132. doi:10.1186/1471-2229-11-132
- Lizana XC, Riegel R, Gomez LD et al (2010) Expansins expression is associated with grain size dynamics in wheat (*Triticum aestivum* L.). *J Exp Bot* 61:1147–1157. doi:10.1093/jxb/erp380
- Natarajan P, Kanagasabapathy D, Gunadayalan G et al (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. doi:10.1186/1471-2164-11-606
- Openshaw K (2000) A review of *Jatropha curcas*: an oil plant of unfulfilled promise. *Biomass Bioenergy* 19:1–15. doi:10.1016/S0961-9534(00)00019-2
- Penfield S, Josse E-M, Halliday KJ (2010) A role for an alternative splice variant of PIF6 in the control of *Arabidopsis* primary seed dormancy. *Plant Mol Biol* 73:89–95. doi:10.1007/s11103-009-9571-1
- Popluechai S, Raorane M, Mulpuri S et al (2009) Narrow genetic and apparent phenetic diversity in *Jatropha curcas*: initial success with generating low phorbol ester interspecific hybrids. *Nat Proc*. doi:10.101/npre.2009.2782.1
- Rosenberg LA, Rinne RW (1988) Protein Synthesis during natural and precocious soybean seed (*Glycine max* [L.] Merr.) maturation. *Plant Physiol* 87:474–478
- Sato S, Hirakawa H, Isobe S et al (2011) Sequence analysis of the genome of an oil-bearing tree, *Jatropha curcas* L. *DNA Res* 18:65–76. doi:10.1093/dnares/dsq030
- Sugliani M, Brambilla V, Clerckx EJM et al (2010) The conserved splicing factor SUA controls alternative splicing of the developmental regulator ABI3 in *Arabidopsis*. *Plant Cell* 22:1936–1946. doi:10.1105/tpc.110.074674
- Tang Z, Choi J-H, Hemmerich C et al (2009) ESTPiper-A web-based analysis pipeline for expressed sequence tags. *BMC Genomics* 10: 174. doi:10.1186/1471-2164-10-174
- Tenenbaum DJ (2008) Food vs. fuel: Diversion of crops could cause more hunger. *Environ Health Perspect* 116(6):A254–A257
- Weber H, Heim U, Golombek S et al (2008) Assimilate uptake and the regulation of seed development. *Seed Sci Res* 8:331–345. doi:10.1017/S0960258500004268
- Wen M, Wang H, Xia Z et al (2010) Development of EST-SSR and genomic-SSR markers to assess genetic diversity in *Jatropha Curcas* L. *BMC Res Notes* 3:42. doi:10.1186/1756-0500-3-42
- White JA, Todd J, Newman T et al (2000) A new set of *Arabidopsis* expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil. *Plant Physiol* 124:1582–1594
- Wobus U, Weber H (1999) Seed maturation: genetic programmes and control signals. *Curr Opin Plant Biol* 2:33–38
- Xu R, Wang R, Liu A (2011) Expression profiles of genes involved in fatty acid and triacylglycerol synthesis in developing seeds of *Jatropha* (*Jatropha curcas* L.). *Biomass Bioenergy* 35:1683–1692. doi:10.1016/j.biombioe.2011.01.001
- Yi C, Zhang S, Liu X et al (2010) Does epigenetic polymorphism contribute to phenotypic variances in *Jatropha curcas* L.? *BMC Plant Biol* 10:259
- Zhang N, Gur A, Gibon Y et al (2010) Genetic analysis of central carbon metabolism unveils an amino acid substitution that alters maize NAD-dependent isocitrate dehydrogenase activity. *PLoS ONE* 5: e9991. doi:10.1371/journal.pone.0009991