Molecular Characterisation of *Jatropha curcas*; Towards an Understanding of its Potential as a Non-edible Oilseed-Based Source of Biodiesel

Siam Popluechai

Thesis submitted for the Degree of Doctor of Philosophy

Faculty of SAgE

Newcastle University

Newcastle upon Tyne

School of Biology

November 2010

Molecular Characterisation of *Jatropha curcas*; Towards an Understanding of its Potential as a Non-edible Oilseed-Based Source of Biodiesel

I certify that no part of the material offered has been previously submitted by me for a degree of other qualification in this or any other University.

Siam Popluechai

November 2010

ABSTRACT

Jatropha curcas, a non-edible oilseed, is becoming popular as a source of biodiesel. However, there are some limitations to using *J. curcas* such as the presence of carcinogenic substances, unpredictable yield, asynchronous flowering and plant height. Understanding the genetic control of these characteristics will help to find solutions. To accumulate genetic information for *J. curcas* plant breeding programmes, the genetic and phenotypic characterisation of *J. curcas* seeds obtained from different geographical locations was undertaken (Chapter 2). The results showed a narrow genetic diversity among accessions but apparent phenetic diversity. The screens employed ranged from phenotypic characterizations of seeds to molecular, biochemical, physiological and genotyping assessments. Seed dimensions and weight measurements provided data on seed density and kernel weight, both important parameters for assessing oilseed quality. Biochemical analysis provided data on oil properties relevant to biodiesel e.g. total oil content, oleic/linoleic acid ratio, saponification number, cetane number and gross energy value. Expression analysis of genes (*acc*, *sad*, *fad*, and *dgat*) essential for the quantity and quality of seed/oil were targeted as molecular screens. Understanding the molecular mechanism of lipid cycling in *J. curcas* seeds will guide future improvements in oil yield. However, this is still an area of active research because little information exists for this in *J. curcas.*

With regards to oil storage capacity of seeds, no information previously existed on the biology of oil bodies in *Jatropha curcas*. In this study, oil bodies from *J. curcas* were studied at the gene transcript and protein levels (Chapter 3). Three *Jatropha oleosin* genes were isolated. Sequence analysis showed that all three oleosin genes contained the proline knot conserved domain. The upstream regulatory element of one *Jatropha* oleosin gene was also isolated. *In silico* analysis of this upstream sequence indicated the presence of a number of seed specific regulatory motifs. Transcript analysis for the three oleosin genes revealed tissue-specific differences. Immunoassay by western blot analysis using antibodies against *Arabidopsis* oleosin and the *J. curcas* specific oleosin revealed a diversity of oleosins. Proteomic analysis of oil bodies revealed additional minor proteins including caleosin, steroleosin, aquaporin and curcin. Information from these studies is essential for varietal improvement of *Jatropha* and/or developing oleosin gene fusion technology for expression of novel compounds.

In this study 623 nucleotides upstream of the *JcOleosin3* gene have been isolated and characterized in transgenic *Arabidopsis* plant using *β-glucuronidase* (GUS) as the reporter gene (Chapter 4). Result showed that the *JcOleosin3* promoter directs

expression of the *β-glucuronidase* gene in seed and pollen but not in leaves, root, stem or flower. In addition, in transgenic *Arabidopsis*, the *JcOleosin3* promoter drives GUS activity at the early seedling stage; this was expected based on earlier reports of expression of *Jatropha* oleosins.

To understand the molecular basis of seed germination, proteomic studies were performed on resting and germinating seeds of *J. curcas*. Such studies using 2-DE and MALDI-TOF MS revealed the presence of patatin-like lipases in the germinating seeds but not in the resting seeds (Chapter 5). These proteins showed sequence similarity and conserved domains as found in the patatin lipases from *Solanum tuberosum* and *Arabidopsis thaliana*. Patatin lipases play a significant role in the mobilization and lipolysis of TAGs during the germination of oil seeds. Further, there is little knowledge on the triacylglycerides (TAGs) lipid cycling in *J. curcas*., Bioinformatic data-mining confirmed the TAGs cycling in *J. curcas* to include the expected stages of synthesis, storage and lipolysis of fatty acids. Based on the data-mining results for the presence of various genes for TAG storage, synthesis and utilization, a scheme is presented for TAG cycling in *J. curcas*. The scheme presented in this study provides essential information required for oil manipulation in *J. curcas*.

Acknowledgements

I wish to express my sincere gratitude and appreciation to Dr Ajay Kohli for his supervision, encouragement, guidance, friendship throughout my study. I also would like to express my sincere gratitude to Professor Angharad M. Gatehouse for her supervision, support and encouragement throughout my PhD study.

I would like to express my gratitude to Dr Diego Breviario (IBBA-CNR, Italy) for his valuable guidance, support and collaboration.

I am also grateful to Dr. Thierry Chardot (INRA, France) and Dr. Maria Manuela R. Costa (Portugal) for their collaboration on oleosin work and their permission for using our publication as a part of my thesis

I would like to express my sincere thanks to Professor Anthony G. O'Donnell for his initial support and guidance for my PhD study.

I would like to thank the following authors who allowed me to use our publications as a part of my thesis: Professor John K. Syers, Dr Sujatha Mulpuri, Dr Harinder P.S. Makkar, Dr Attipalli R Reddy, Dr Enrico Palchetti, Dr Marine Froissard, Dr Pascale Jolivet, Mr Manish Raorane, Miss Udhaya Kannan, Mr. Ibrahim Khalil Adam, Mr João Alexandre Pereira Raimundo.

I would like to thank Dr Silvia Gianni for her help and valuable guidance throughout my work at IBBA-CNR. I would like to thank Dr Prapassorn D. Eungwanichayapant (Mae Fah Luang University, Thailand) for financially supporting part of the work on oleosin genes expression carried out in Thailand.

I also appreciate the help and pleasant working atmosphere from members of the Plant-Insect Interaction Laboratory, Newcastle University. I would like to thank members of Ecotub Laboratory, Milan for their help during my stay at IBBA-CNR.

I also wish to thank Mr Nikolaos Tsakirpaloglou, Mr Manish Raorane, and members of the Thai society in Newcastle for their sincere friendship.

I am greatly indebted to the Royal Thai Government who provided me with a PhD scholarship.

Finally, I wish to express my special thanks to my family and friends for their love, understanding, support and encouragement. I dedicate this thesis to my mother whom I love the most.

PUBLICATIONS RELEVANT TO THIS THESIS

Popluechai S, Raorane M, Syers KJ, O'Donnell AG. Kohli A. 2008**.** Research needs to make Jatropha a viable alternate oilseed crop. *Proceedings of the International Workshop on Feasibility of Non-Edible Oil Seed Crops for Biofuel Production,* May 25 – 27, 2007, Mae Fah Luang University, Chiang Rai, Thailand. Eds. Syers KJ and Wood D. MFLU publication, p 56 - 60

Popluechai S., Chardot T., Gatehouse AMR., Kohli A. Characterization of *Jatropha curcas* oil bodies using a systems biology approach. *Jatropha International Congress 2008,* 17-18 December 2008, Grand Copthorne Waterfront Hotel, Singapore

Popluechai, Siam, Breviario, Diego, Mulpuri, Sujatha, Makkar, Harinder, Raorane, Manish, Reddy, Attipalli, Palchetti, Enrico, Gatehouse, Angharad, Syers, Keith, O'Donnell, Anthony, and Kohli, Ajay. Narrow genetic and apparent phenetic diversity in *Jatropha curcas*: initial success with generating low phorbol ester interspecific hybrids. Available from Nature Precedings <http://hdl.handle.net/10101 /npre.2009.2782.1> (2009)

Kohli A, Raorane M, **Popluechai S,** Kannan U, Syers KJ, O'Donnell AG. 2009. *Jatropha curcas* as a novel, non-edible oilseed plant for biodiesel. In: Ferry N, Gatehouse AMR, eds. *Environmental impact of genetically modified novel crops*. Chapter 14 CAB International, London, UK. p 294 -322

Siam Popluechai, Marine Froissard, Pascale Jolivet, Diego Breviario, Angharad MR Gatehouse, Anthony G O'Donnell, Thierry Chardot, Ajay Kohli. (2010) Understanding *Jatropha curcas* seed oil storage components: characterization of seed oilbody oleosins. Plant Physiology and Biochemistry (accepted with minor modification, 05/10/10)

Siam Popluechai, João Alexandre Pereira Raimundo, Angharad M.R. Gatehouse, Maria Manuela R. Costa, Ajay Kohli (2010) *Jatropha* seed-specific, *JcOleosin3* gene, promoter in transgenic *Arabidopsis*. Manuscript under preparation

TABLE OF CONTENT

STATEMENT OF ORIGINALITY……..……i
ABSTRACT…….……ii
ACKNOWLEDGEMENTS…….....iv
PUBLICATIONS RELEVANT TO THIS THESIS…….…....v
TABLE OF CONTENTS……..vi
LIST OF FIGURES……...….x
LIST OF TABLES……...xiv

Chapter 1: Literature review: *Jatropha curcas* as a novel, non-edible oilseed plant for biodiesel

1.1 Summary……..…1
1.2 The need for renewable sources of energy……..….2
1.2.1 Depletion of non-renewable resources…….....2
1.2.2 Environmental damage……..…..3
1.2.3 Rising prices…….…...4
1.3 The case for renewables……...5
1.4 The need for replacement transport fuels…….5
1.4.1 Why the transport sector……..5
1.4.2 The options……...6
1.5 Why Biodiesel……...……7
1.6 Oils for food or fuel……12
1.6.1 Non-edible oilseed plants for biodiesel……..12
1.7 The case for *Jatropha curcas*.....13
1.7.1 Seed oil and biodiesel characteristics of *J. curcas*…….16
1.7.2 Potential *J. curcas* co-products……...18
1.7.3 Rural, socio-economic and ecological sustainability……..20
1.8 Research and development issues for *Jatropha*……...22
1.8.1 Levels of toxins and co-carcinogens……...22
1.8.2 Limited genetic variability……..24
1.8.3 Seed germination……28
1.8.4 Reproduction…….…..28
1.8.5 Propagation…….……31
1.8.6 Biotic stress assessment and management…….….32
1.8.7 Agronomic requirements…….……33

1.8.8 Gene mining and expression analysis……34
1.9 Conclusions……..36
1.10 Problem hypothesis and objectives……..37
1.11 References……...39

Chapter 2: Narrow genetic and apparent phonetic diversity in *Jatropha curcas*: initial success with generating low phorbol ester inter-specific hybrids

2.1 Abstract……...59
2.2 Introduction……60
2.3 Materials and Methods……63
2.4 Results……68
Genetic diversity in *Jatropha*……..68
RAPDs, AFLP, cTBP
Seed phorbol ester content……..70
Phenotypic variability ……70
2.5 Discussion……...87
2.6 References……...92
2.7 Appendix……..…98

Chapter 3: Understanding *Jatropha curcas* seed oil storage components: characterization of seed oilbody oleosins

3.1 Abstract……99
3.2 Introduction……100
3.3 Materials and Methods……...103
3.4 Results……109
Isolation and characterization of *J. curcas* oleosins at the gene level…..109
Characterization of *J. curcas* oleosins at the transcript level……110
Characterization of *J. curcas* oleosins at the protein level……112
Oleosins are responsible for surface properties of *Jatropha* oil bodies…114
Lipid content in *Jatropha* oil bodies……115
Diversity of oleosins in *Jatropha* sp……115
Relationship of oleosin expression to oil yield and fatty acid profile…..117
3.5 Discussion……..128
3.6 References……..135
3.7 Supplemental data……..142

3.8 Appendixes……147

Chapter 4: Characterization of the *Jatropha, JcOleosin3* promoter in transgenic *Arabidopsis*

4.1 Abstract……..163
4.2 Introduction……164
4.3 Materials and methods……..…….165
4.4 Results……167
Isolation and analysis of the 5' regulatory regions of JcOleosin3 gene...167
Characterization of transgenic *Arabidopsis*……..167
Tissue-specific accumulation of GUS driven by the *JcOleosin3* promoter…….168
Accumulation of GUS driven by the *JcOleosin3* promoter during transgenic seedling……168
4.5 Discussion……...…173
4.6 References……...…175
4.7 Supplemental data……...180

Chapter 5: Identify patatin-like lipase proteins and presentation for Triacylglycerides cycling scheme in *Jatropha curcas*

5.1 Abstract……...182
5.2 Introduction…….183
5.3 Materials and methods……..……..186
5.4 Results…….188
Identification of Patatin-like lipase using proteomic analysis……188
J. curcas genes involved in TAGs biosynthesis, storage and lipolysis……..194
5.5 Discussion……...196
5.6 Conclusions and future studies……202
5.7 References……...204

Chapter 6: General discussion and conclusions

6.1 Genotype and phenotype analysis of *J. curcas* ……..208

6.2 Molecular biology of *Jatropha* oilbodies and olesoin……210

6.3 *J. curcas* preferentially seed targeted promoter...……..…….211

6.4 *J. curcas* patatin-like lipases and presentation for TAGs cycling scheme of *J. curcas*……..……...213

6.5 Conclusions…….213

6.6 Future prospective……...214

6.7 References…….……..216

LIST OF FIGURES

Chapter 1

Figure 1 World marketed energy consumption 1980-2030……2

Figure 2 Basic steps in biodiesel production……...6

Figure 3 The summation strengths of *J. curcas*……15

Figure 4 : (A) Young *J. curcas* plant with flowers and developing seed pods.
(B) J. curcas inflorescence containing both male staminate flowers (M)
and female pistillate flowers (F). (King *et al.,* 2009)……...…..16

Figure 5 (C) Cross-section of a *J. curcas* seed pod containing three developing
seeds. (D) Mature seeds of *J. curcas.* (King *et al.,* 2009)……...……..16

Figure 6 UPGMA analysis of RAPD conducted on 17 *J. curcas* and
one *J. podagrica* accession…….28

Figure 7 Pests and natural enemies of *J. curcas*…….....33

Chapter 2

Figure 1 UPGMA analysis of RAPD conducted on 17 *J. curcas* and
one *J. podagrica* accession…….73

Figure 2A Representative gel showing DNA amplicons obtained using
two separate AFLP primer sets (I and II) on six *J. curcas* accessions
(lane 1 to 6) from Thailand and one (lane 7) from Mexico
(Non-toxic; NTMA).
2B UPGMA analysis of the AFLP results with the 32 primer pairs…….74

Figure 3 cTBP-mediated detection of monomorphism in *J. curcas*……75

Figure 4 cTBP-mediated genotyping of *J. curcas*. cTBP analysis for intron I (a)
and II (b) for 16 *J. curcas* accessions from 6 countries…….....76

Figure 5 UPGMA of cTBP analysis on different species of *Jatropha*……77

Figure 6 cTBP-mediated genotyping of the inter-specific hybrids and parents…....78

Figure 7 Relationship between FA biosynthesis gene transcripts and FA profile in the seeds of *J. curcas*. ……...…79

Figure 8 Classification of *J. curcas* accessions based on seed/oil properties listed
in Table S2……80

Figure 9 Relationship between FA biosynthesis gene transcripts and TAG
accumulation in the seeds of *J. curcas*……..81

Appendix 1 Combinatorial Tubulin Based Polymorphism (cTBP) for β-tubulin gene family introns analysis conducted on 17 *J. curcas* and one *J. podagrica* accession……...98

Chapter 3

Figure 1 Partial sequence alignment of gDNA and cDNA of *JcOleosin 3* showing the position and sequence of the intron in bold letters…….121

Figure 2 Expression of *J. curcas oleosins* in different tissues using quantitative real-time PCR (qPCR)…….122

Figure 3 (a) Accumulation of *oleosin* and *dgat* mRNA during *J. curcas* fruit development……...123

(b) Accumulation of mRNA of *oleosins* during *J. curcas* seedling development……...…123

Figure 4 (a) Amino acid sequence alignment of *J. curcas* oleosins with those known from other plants to show the conserved sequence from the proline knot……..124

(b) Optimal alignment of the protein region containing the 18 residue insertion at the C-terminal distinguishing the H- and L-form oleosins……..124

(c) Phylogenetic tree generated from the multiple oleosin sequence alignment clearly distinguishes the L- and the H-form oleosins……...124

Figure 5 Protein profiles of chloroform/methanol (11/7) seed extract of *J. curcas* accessions from India (JI) and Indonesia (JIO)……...125

Figure 6 Immunological detection of oleosins using antiserum raised against *A. thaliana* oleosin S2 (anti rS2N)……..126

Figure 7 Alignment of the *JcOleosin3* intron sequences obtained from various *Jatropha* accessions……127

Supplemental Figure 1 Surface properties of *Jatropha* seed oil bodies……144

Supplemental Figure 2 (a) SDS-PAGE and (b) immunoassay of alkaline phosphatase treatment on *JcOloesin3* using Shrimp Alkaline Phosphatase (SAP) at different time intervals (0.5, 1, 2, and 3 hour)……..145
Supplemental Figure 3 Immunoassay for *JcOleosin3* in *J. curcas* accessions……...146
Appendix 1 (A) Summarized diagram shows *Jatropha* oleosin gene isolation using degenerated PCR mediated method…….148
(B) Agarose gel electrophoresis obtained from genomic (D) and cDNA (R) using 5' and 3' primers of each oleosin genes……...148
Appendix 2 Nucleic acid alignment of *J. curcas* oleosins against other plant species……..149
Appendix 3A Expression of *J. curcas oleosin* genes in different tissues using semi-quantitative PCR……..153
3B Expression of *J. curcas oleosin* genes during fruit development 7, 21, 35, 56, and 70 days after fertilization (DAF)….153
3C Expression of *J. curcas oleosin* genes during seedling using semi-quantitative PCR…….153
Appendix 4 Deduced amino acid alignment *of J. curcas* oleosins against other plant species…….....154
Appendix 5 Electrophoresis separation of protein purified with organic solvents on SDS-PAGE (12.5 %)……...…….155
Figure 6A Mass spectrometry data obtained by conventional LC-MS/MS of in-gel digestion of SDS-PAGE protein bands (see figure 5) from purified oil bodies of *J. curcas* accession India…….....156
6B Mass spectrometry data obtained by conventional LC-MS/MS of in-gel digestion of SDS-PAGE protein bands (see figure 5) from purified oil bodies of *J. curcas* accession Indonesia……....157
Appendix 7 Immunoassay for *JcOleosin3* in *J. curcas* protein extracted from seed and leaves using organic solvent……..……..158
Appendix 8 Thin Layer Chromatography of oil extracted from *J. curcas* oil bodies……..159
Appendix 9 (A) PCR products amplified using JcFLO3F and JcFLO3R primers of *J. curcas* (Jc), *J. podagrica* (Jp), *J. gysipifolia* (Jg), *J. mutifida* (Jm), *J. maheswarii* (Jm), *and J. interegima* (Ji)……160

(B) Nucleic acid and (C) amino acids alignments (D) Phylogenetic tree generated from the multiple oleosin3 sequence alignment obtained from *Jatropha integerima* (JinterO3), *J. gysipifolia* (JgossO3), *J. maheswarii* (JmahesO3) compare with olesoin3 of *J. curcas*…….161
Appendix 10 Immunological detection of chloroform/methanol 11/7 seed extracted protein of *J. curcas* using antiserum against *JcOleosin*3…..162

Chapter 4

Figure 1 Structure of the 623-*JcOleosin3* promoter with glucuronidase reporter gene (*gus*) in the binary vector pGreen0229…….……...170
Figure 2 Histological analysis of GUS activity in T1 transgenic *Arabidopsis* during fruit development…….……....171
Figure 3 Expression of GUS in transgenic *Arabidopsis* seedlings……..172
Supplemental figure 1 Schematic representation locations of putative cis-acting motifs for seed storage proteins of *JcOlesoin3* regulatory element (EU543568) predicted by the PLACE data base……180
Supplemental figure 2 (A) Histological analysis of GUS activity in 2 DAI of T1 transgenic *Arabidopsis* seedling, no expression observe in seed-coat
(B) GUS activity seedling of 13 DAI of transgenic *Arabidopsis*
(C) GUS activity leaves and stem of transgenic *Arabidopsis*
(D) GUS activity in root of transgenic *Arabidopsis*…….181

Chapter 5

Figure 1 The TAG synthesis and breakdown cycle……..……..183
Figures 2 (A) Two-Dimensional gel images of protein profiles of resting and germinating stages resolved on 18 cm strips……...……..189
(B) Close-up sections of 2D gels, showing unique protein spots number 1 to 7 present in germinating but not in resting seeds…….….190
(C) A distribution of protein spots in the two stages of seed development.190
Figure 3 Peptide fragments obtained from peptide mass fingerprinting of unique spots 1-7 from the 2D gel. of germinating seeds…….191
Figure 4 Scheme of TAGs lipid cycling of *J. curcas*…….203

Chapter 6

Figure 1 Summary flow chart of the present study……..215

LIST OF TABLES

Chapter 1

Table 1 Total allowable emissions for the entire 21st century and number of years that the allowed CO_2 emissions of each region could continue at constant 2007 level, provided that it is reduced to zero after that time span. (adapted from Sorensen 2008)……4

Table 2 Table showing absolute biodiesel production (Sharma, Y.C. & Singh, B., 2009)…….....9

Table 3 Comparison of properties between biodiesel and petro-diesel *Adapted from*: Mixon *et al*., (2003)……11

Table 4 Comparison of properties of Diesel and *Jatropha* biodiesel…….17

Table 5 Table showing the list of various uses of the *J. curcas* plant……19

Chapter 2

Table 1a Global origin of *Jatropha* accessions and species used……..67

1b List of primers used in this study……67

Table 2a Phorbol Ester content in *Jatropha* kernels……..82

2b Phorbol Ester content in *Jatropha* hybrid kernels……...82

Table 3 Comparative range of critical seed properties for interspecific d erivatives and *J.* curcas……...83

Table 4a *J. curcas* seed properties relevant to biodiesel……84

4b *J. curcas* oil properties relevant to biodiesel……84

4c Seed oil properties of *Jatropha* species……85

Table 5a Differential Water Use Efficiency (WUE) of *J. curcas* plants in the field……86

5b Differential Water Use Efficiency (WUE) of *J. curcas* plants in the CER……86

Chapter 3

Table I. Salient features of *J. curcas* oleosins……118

Table II. Identified proteins in oil bodies purified from *J. curcas* accession from Indonesia and India……119

Table III. Relationship between fatty acid profile, oil yield and presence (1) or absence (0) of *oleosins* and their expression……...120

Supplemental Table I. List of primers used in the study……142

Supplemental Table II Information on *oleosins* from *Arabidopsis*, Ficus, Coffee, *Ricinus*, Olive……143

Chapter 4

Table I Seed specific and transcription initiation motifs of the *Jcolesoin3* regulatory element (EU543568) predicted by using PLACE programme……..169

Chapter 5

Table 1 Protein spots identified from germinating seeds of *J. curcas* using MALDI-TOF MS……..193

Table 2 Sequence analysis of triacylglycerides (TAGs) lipid cycling genes in *J. curcas*……..195

Chapter 1: *Jatropha curcas* as a novel, non-edible oilseed plant for biodiesel

1.1 Summary:

The negative environmental impacts, the limited sources and rising prices of fossil fuels pose significant environmental and socio-economic challenges. Globally, major national and international initiatives are underway to identify, revive, research and recommend renewable sources of energy. One such renewable source is esterified vegetable oil i.e. biodiesel. Crop plants yielding edible oilseeds can be diverted to the biodiesel market only to a limited extent due to their value in the food sector. One route to meeting the gap between the demand for food-oils and the need for alternative fuel-oils is the use of non-edible oil seeds from *Jatropha curcas* plants. *J. curcas* or physic nut is a member of the *Euphorbiaceae* family and has been the subject of much interest as a source of biodiesel due to a number of perceived advantages. For example, the by-products of *J. curcas* based biodiesel production have potential as a nutritious seed cake for fodder, as a soil amendment or as a biogas feedstock. Glycerol, a by product of the biodiesel industry can be used in a variety of industrial applications. *J. curcas* leaf, stem and bark extracts have uses in the medicinal, cosmetics, plastics and insecticide/pesticide industry. *J. curcas* can grow and yield on marginal and wastelands promoting effective land use, gender empowerment and soil rehabilitation thus contributing to sustainable rural development. However, neither *J. curcas* nor any other potentially useful non-edible oilseed plant is currently grown commercially. In fact, such plants are generally undomesticated and have yet to be subject to any genetic improvement with respect to yield quality or quantity. Also, many *J. curcas* accessions can be toxic to humans and animals due to the presence of toxic compounds such as curcins and phorbol esters. Thus, despite the enthusiasm in some countries for widespread plantation cropping, *J. curcas* is currently not commercially viable as a biodiesel feedstock without genetic improvement either through conventional breeding or molecular engineering. The unpredictable yield patterns for seed quality and quantity, varying and often low oil content, the presence of toxic and carcinogenic compounds, high male to female flower ratio, asynchronous and multiple flowering flushes, low seed germination frequency, plant height and its susceptibility to biotic and abiotic stresses are some of the common limiting factors to the success of *J. curcas*. This chapter presents a review of the potential of *Jatropha* as a model, non-edible, oilseed plant and the research needed to realize its potential as a bioenergy crop.

1.2 The need for renewable sources of energy

1.2.1 Depletion of non-renewable resources

Non-renewable fossil fuels are a limited resource that supplies nearly 90% of the world's energy demand. Sustained economic growth in some countries and a desire for similar trends in others, especially developing countries like India and China has led to an exponential increase in global energy consumption (Figure 1). Although it is much debated as to whether substantial oil reserves lie undiscovered, inaccessible or environmentally hazardous to recover, it is widely accepted that the rate of consumption will continue to increase.

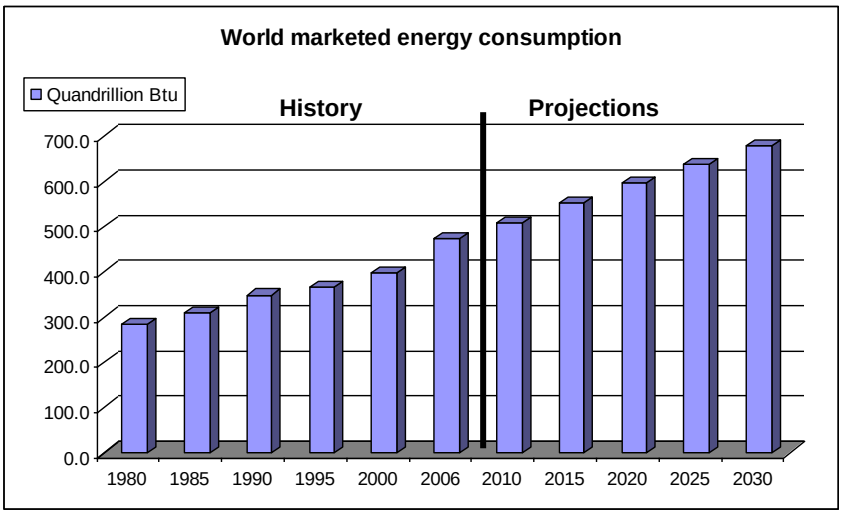


Figure 1: World marketed energy consumption 1980-2030 (Adapted from www.eia.doe.gov/iea)

Considering anticipated energy demands it is expected that oil reserves will last many years less than projections based on the current rate of consumption and there is a general consensus that non-renewable energy resources will become limited sooner rather than later. Predictions for world population of up to 8 billion by 2025 foresee most of this increased energy demand coming from developing and transition economies including China, India and Brazil (Matthews, 2007; Focacci, 2005). China and India, regarded as the fastest growing economies at present, accounted for 10% of the world energy consumption in 1990. This increased to 19% in 2006 and is projected to be 28% by 2030, compared to 17% for USA.(IEO, 2009)

1.2.2 Environmental damage

Fossil fuel combustion contributes to global warming and acid rain. Sun and Hansen (2003) have used surface air temperature change data from 1951 to 2000 to compute anticipated changes for 2050 and suggest that surface air temperatures will increase by an average of 0.3 to 0.4°C. This is close to the average surface air temperature rise of 0.5 C in the last century (Jones *et al.,* 2003; 2010). By the end of the 21 century, temperatures are expected to rise by 6.4°C (Friedlingstein, 2008). Recent suggestions put forward to stabilize the rising temperatures by 1.5°C by controlling CO_2 emissions is still very hypothetical and is a process that will take many years to achieve. The following table illustrates the concept (Sorensen, B 2008).

The IEO in its report on CO_2 emissions, revealed that the annual emissions from non-OECD countries (developing economic countries) increased by 30% in 2010 compared that by OECD countries (developed and stable economic countries). However, this difference is projected to be 77 % by 2030, thus making the issue of reducing CO_2 emissions a daunting task. The increases in surface air temperature are expected to continue even if carbon dioxide emissions were stopped today. Barnett *et al.,* 2005 concluded that the most obvious reasons for temperature rise in the past 50 years were the increasing release of greenhouse gases. Solar variability and volcanic activites were small and partial reasons for the same. Affects such as rise in sea levels, unpredictable drought-flood water cycles, quantitative and habit change in flora and fauna have all been related to climate change. The two most important components of greenhouse gases are carbon dioxide (CO_2) and methane (CH_4). Fossil fuel combustion in power plants, transport, and industry is the major source of CO_2 emissions and accounts for nearly 80% of worldwide anthropogenic CO_2 emissions. Carbon dioxide levels are also affected by deforestation where the loss of plants that trap CO_2 leads to higher atmospheric levels and thus to a rise in temperature. Similarly, the release of CH_4 from agricultural enterprises, including animal rearing and rice farming, along with the consumption of natural gas also contributes to global warming.

Countries	Accumulated emission allowance 2000–2100 (Gt CO_2)	Years to zero-CO_2 allowance at constant 2007 emissions
United States	255	42
Canada Australia New Zealand	44	30
Japan	65	49
West Europe	179	50
East Europe B	34	49
Russia Ukraine Belarus	78	39
Middle East	161	79
China	250	64
Other Asia	155	76
India	183	144
Latin America	86	41
Africa	213	115
Int. Ship/Air	80	95
World	**1783**	**61**

Table 1: Total allowable emissions for the entire 21st century and number of years that the allowed CO_2 emissions of each region could continue at constant 2007 level, provided that it is reduced to zero after that time span. (adapted from Sorensen 2008)

1.2.3 Rising prices

Fossil fuels encompass a range of commodities such as crude oil, natural gas and coal. Although prices for each of these commodities are driven by different factors, as a global average they follow similar trends. Since crude oil is the source of a number of petroleum products feeding the transportation and other industries, the economic impact of oil prices are easily visible through rising prices of other commodities. Oil prices have risen 6-fold over the last century but over the period 2002-2008 oil prices have increased 4-fold from $30 to more than $130 per barrel and are projected to rise to $200 per barrel by 2030. Higher oil prices are becoming restrictive to economic growth; hence energy efficiency and energy diversification have become a top priority in terms of both energy generation and energy use in several countries.

1.3 The case for renewables (biomass)

Current political and economic scenarios surrounding energy reserves and their utilization encourage investment in the discovery and mining of new oil and gas reserves whilst simultaneously developing highly efficient conversion technologies and considering alternative, renewable energy sources. At a regional scale renewable fuels can be very cost effective if considering ancillary benefits such as rural development, land and soil reclamation and environmental amelioration through the use of biomass-based fuels. Wind, water, solar and geothermal-based energy can also be highly suitable and cost-effective for selective areas. However, the highly centralized and largely efficient fossil fuels-based supply and distribution system currently puts renewables at a disadvantage. This is because the localized, domestic nature of renewables threatens those countries and corporations that control the source, production and distribution of fossil fuels. However, over the next 20 years developing countries will account for a large part of the projected energy demand and as such are major sources and markets for renewable energy plants. Renewables must therefore become an integral part of a diversified energy portfolio. Global awareness of the limited supply and environmental damage caused by fossil fuels combined with their increasing prices has raised government, industry and researcher interests in renewable energy. Currently, because of its importance to the transport sector, biodiesel is regarded as one of the main solutions.

1.4 The need for replacement of transport fuels

1.4.1 Why the transport sector

The total energy consumed in any country is primarily divided into the following types: electricity, heat, power and fuel. These are in turn used mainly by the residential, commercial, industrial and transport sectors. Although the transport sector globally uses 22% of the total energy and accounts for 27% of the global CO_2 emmissions (deLa Rue du Can and Price, 2008), it has become the immediate target for renewables. This is because the transport sector is expected to see the greatest and most immediate surge in energy demand, with a projected 2.7% increase in demand per year until 2030. Although this is likely to be the case in developing and transition economies such Brazil, China and India, where vehicle fitness and emission regulations are not strictly imposed, it may also impact on cities of more developed countries where transportation remains the largest emitter of CO_2 (deLa Rue du Can and Price, 2008) since energy

conversion efficiency and emission control in transportation is generally lower than in other energy sectors.

1.4.2 The options

Almost 97% of the world's transportation related energy demand whether road, rail, air or sea, runs either on petroleum-based products or natural gas (deLa Rue du Can and Price, 2008). The major non-fossil alternative fuels are biofuels that include biodiesel, bioethanol, biogas and bioelectricity produced using hydrogen fuel cells. With the exception of bioethanol produced directly from plant sugars or starches (e.g. sugarcane, sugar beet, cassava or cereal grains), both biogas and bioalcohols (including biopropanol and biobutanol being investigated by some companies) remain restrictively expensive for use as vehicle fuels. Brazil has set an example for the use of bioethanol where nearly 40% cars run on bioethanol, but for most other countries biodiesel is the biofuel of choice.

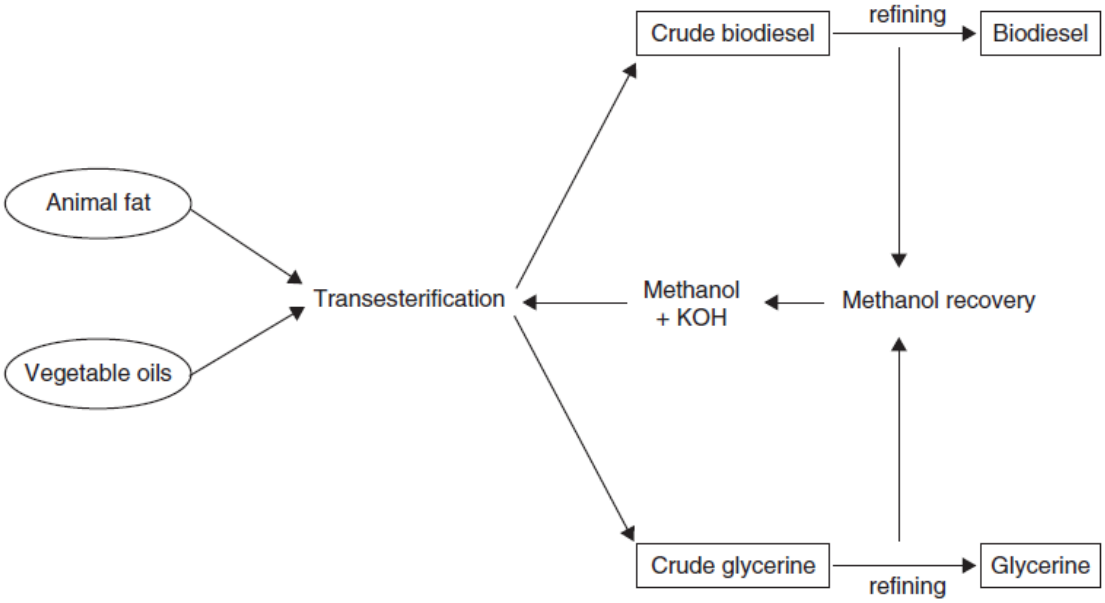


Figure 2: Basic steps in biodiesel production.

1.5 Why biodiesel?

A mix of fatty acid methyl esters (FAMEs) made from a renewable biological source such as vegetable oil or unused or used animal fat and algae is called biodiesel. Biodiesel is simple to use as it requires no major engine modifications. It is also non-toxic biodegradable and free of sulfur and aromatics. The following points illustrate why biodiesel may be the fuel for the future.

- It is one of the most tested and analysed of the alternate sources of fuels. Research by various departments of energy and renewable energy labs, have shown that biodiesel performance is comparable to that of petrodiesel and has much more benefits in terms of environment and human health.
- Since 15 years the biodiesel industry has set up the quality standards for its usage. Biodiesel fuel blends and various other performance based specifications are analysed and updated. (www.astm.org). Biodiesel fuel blend was shown to have a higher cetane number than the petroleum product. It also has a similar fuel efficiency rates, torque and horsepower. It was also recorded to have a highest BTU (British thermal unit: 1 BTU is equal 1055.056 joules) of any of the alternate fuel energy sources.
- The launch of new engine emission control system and new vehicles in 2010, can also accept 5% biofuels blend. The major concerns being the oil dilutions level which however have now been proved by the VOLKSWAGEN and National renewable energy labs to be of an uncritical level. Using high end biodiesel up to ASTM specifications, will allow users to be assured of high engine performance. All US automobile industries and also some of the companies in Europe have accepted lower level of biofuels blends (B5) to be used. (NBB, 2010, www.biodiesel.org/resources/oems.)
- National Renewable Energy Laboratory (NREL) studies on various fronts like storage, sampling, blending, shipping and distribution showed that biodiesel meets all the quality standards (NREL, 2008). Biodiesel can also work in cold temperatures. Just like normal petrodiesel, the higher freezing point components in the biodiesel crystallize at cold temperatures. However pure biodiesel is known to have a higher cloud point, hence blending it with petroleum products (B5 or B20) will increase its cold weather compatibility (www.biodiesel.org/cold).
- Biodiesel reduces the carbon emission by 60 to 80 percent and emits negligible amounts of sulfur or other harmful gases. It gives 4.5 units of energy for every

unit of energy used to produce it, thus making it the highest energy balance option (NBB, 2009). Also it is produced from a variety of natural resources and from plants already cultivated as crops. Biodiesel is thus one of the most diverse fuel on the planet (NBB, 2009). It also provides opportunity to address country-specific energy security issues because almost every country has a choice of local oilseed plant as a source of biodiesel.

- Biodiesel has become a global industry and in many cases adds substantially to the Gross Domestic Product (GDP). The biodiesel industry can potentially create job opportunities around the world. A study carried out by NBB, estimated that for every 100 million gallons of biodiesel from algae, 16,455 jobs will be created adding upto $1.461 billion to the GDP.

FAMEs can be produced by different methods of esterification, though most follow a similar basic approach (Figure 2) in which the oil is first filtered to remove water and contaminants and then pretreated to remove, or esterified to transform, free fatty acids into biodiesel. Generally, only recycled animal fat containing more than 4% w/v free fatty acids needs pretreatment. Following pre-treatment the oils are mixed with methanol and a catalyst (usually sodium or potassium hydroxide) to trans-esterify the triglycerides into FAMEs and glycerol. The glycerol can then be separated and the FAMEs used as biodiesel. The trans-esterification of vegetable oils is not new and has been practiced since the 1800s when it was used to produce glycerin for soaps. In soap production, however, the FAMEs are the by-products. The use of vegetable oils in transport is also not new. More than a century ago Rudolph Diesel, the inventor of the diesel engine first powered his engine, with peanut oil. It was not till the 1920s when petro-diesel became readily available that biodiesel went out of favour.

Of the different biofuel options currently available, biodiesel is the only one that is economically and technologically feasible and is expected, along with the Brazilian model of bioethanol, to remain competitive at least in the short term (Johnston and Holloway, 2007). Within the European Union, biodiesel production and consumption over the last decade has continued to increase. The total EU27 biodiesel production for 2006 was over 4.8 million metric tonnes, an increase of 54% over the 2005 (EBB, 2008). Whilst the USA has been producing biodiesel from soy oil to a limited extent for many years, it has in recent years dramatically increased production from 20 million gallons in 2003 to over 450 million gallons by 2007 leading to 2.2 billion gallons in 2008 (NBB, 2009). Table 2 shows that Malaysia is at the top of the list in terms of absolute biodiesel production. The raw material used for biodiesel in the nations listed above is 28% soyabean oil, 22% palm oil, 20% animal fat, 11% coconut oil and 5% from other plants which include rapeseed, sunflower and olive oil.

Top 10 countries in terms of absolute biodiesel potential

Rank	Country	Volume potential (L)	Production ($/L)[a]
1	Malaysia	14,540,000,000	$0.53
2	Indonesia	7,595,000,000	$0.49
3	Argentina	5,255,000,000	$0.62
4	USA	3,212,000,000	$0.70
5	Brazil	2,567,000,000	$0.62
6	Netherlands	2,496,000,000	$0.75
7	Germany	2,024,000,000	$0.79
8	Philippines	1,234,000,000	$0.53
9	Belgium	1,213,000,000	$0.78
10	Spain	1,073,000,000	$1.71

[a] Average production cost per liter is calculated from all available lipid feedstock prices, increased by a $0.12 refining cost and decreased by $0.04 for the sale of by-products.

Table 2: Table showing absolute biodiesel production (Sharma, Y.C. & Singh, B., 2009)

Biodiesel is the only biofuel for which there is a strong market demand at present. The temporal and spatial flexibility in biodiesel production due to the variety of feedstocks available is a major strength. Most OECD countries in the Americas and Europe are self sufficient in their production of edible oil which is generally surplus to their requirements. This is why most of Germany and France's biodiesel is made from rapeseed oil while most of that from the USA is derived from soybean. Malaysia and Indonesia use coconut oil for biodiesel synthesis. India on the other hand does not

produce enough edible oil and hence proposes the use of non-edible oilseed crops such as *J. curcas* and *Pongamia pinnata*. Other vegetable oils have also been used for producing biodiesel, for example oil from peanut, sunflower, safflower, canola, linseed and palm. However, these generally cater to local small-scale consumption. Heterotrophic Algae (Chisti, 2007) and bacteria (Kalscheuer *et al.,* 2006) have also been used following genetic engineering to produce oil for conversion to biodiesel. In presence of sunlight algae was shown to convert CO_2 into sugars and proteins. However when starved of nitrogen, it produced oil. *Chlorella protothecoides* was grown both autotrophically and hetrotrophically and produced 14.57% and 55.20% amount of lipid respectively, which can act as a raw material for biodiesel production (Sharma, Y.C & Singh, B. 2009).

Blending biodiesel with petro-diesel is more popular than the use of B100 (100% biodiesel). It can be blended with petrodiesel at any levels to create a biodiesel blend. The advantages of biodiesel over petro-diesel depend on several factors but mostly on the composition of the blend used, for example, a 20% blend of biodiesel (B20) with petro-diesel works well in terms of economies of scale and allied benefits. Based on the average diesel consumption of the US and on soya-based biodiesel production, B100 yields 3.2 units of fuel energy compared to the 0.83 unit yield of petro-diesel per unit of fossil energy consumed (Mixon *et al.,* 2003). In countries with lower transportation and feedstock production costs, this yield comparison is higher. Table 3 compares some critical properties of biodiesel and petro-diesel. Based on these properties there are obvious environmental benefits in using biodiesel. For example, in B100 there is zero sulphur, so sulphate and sulphur dioxide (contributors to acid rain) are eliminated. It also has 67% less unburned hydrocarbons reducing the carcinogen concentrations and smog forming capacity in air. There is 48% less CO and 47% less particulate matter (PM) than is found in low-sulphur petro-diesel. Biodiesel does, however, increase nitrous oxide (NOx) by 10% but the total smog forming potential is much lower and the absence of SO_2 should allow the introduction of NOx control technologies such as exhaust gas recirculation (EGR) in diesel engines to be employed (Agarwal *et al.,* 2006). The high cetane number of biodiesel leads to similar power, torque and fuel economy in terms of performance while higher lubricity provides yet another advantage (Agarwal *et al*., 2001). Biodiesel is inferior to petro-diesel in engine performance in cold temperatures and in its corrosive effects on rubber components. However, both these issues were addressed using antifreeze and alternative tubing

materials respectively. Nonetheless, more research is needed to address the issue of cold temperature performance if biodiesel is to be used more extensively in colder regions.

Property	Biodiesel	Petroleum diesel (CARB low-sulfur diesel)
Cetane number	51–62	44–49
Lubricity	More than diesel, comparable to oil lubricants	Low-sulfur fuel has very low lubricity factor
Biodegradability	Readily biodegradable	Poorly biodegradable
Toxicity	Essentially non-toxic	Highly toxic
Oxygen	Up to 11% free oxygen	Very low
Aromatics	No aromatic compounds	18–22%
Sulfur	None	0.05%
Cloud point	Slightly worse than diesel	NA
Flash point	300–400°F	125°F
Spill hazard	None	High
Material compatibility	Degrades natural rubber	No effect on natural rubber
Shipping	Non-hazardous and non-flammable	Hazardous
Heating value	2–3% higher than diesel	1
Renewable supply	Renewable fuel	Non-renewable
Supply	USA estimated 2 billion gal/year	Limited
Energy security	Domestic raw material	Mix of domestic and imports
Alternative fuel	Yes	No
Production process	Chemical reaction	Reaction + fractionation

Table 3: Comparison of properties between biodiesel and petro-diesel. (Adapted from: Mixon *et al.*, 2003)

Despite these minor technical drawbacks and policy frameworks that are still not so supportive of biodiesel compared to petro-diesel, the demand for biodiesel continues to rise (Figure 4). EU production and consumption is rising and there is a clear road map for EU countries to replace 5.75% of all its transport fuel with biofuels by 2010 with a planned increase to 10% by 2020 (BRAC, 2006). Total energy supplied by renewables is planned to be 20% by 2020. Similarly, countries such as China and India are also moving towards renewable fuels. India has an ambitious target of replacing 20% of its transport fuel with biodiesel by 2010 (Biodiesel 2020, 2008). Therefore, biofuels in general and specifically biodiesel is expected to be the fuel of choice over the next 5 to 10 years (Biodiesel 2020, 2008). It is argued, however, that a coherent vision on the global role of biofuels in energy, economic and environmental policy is still missing, but that for long-term efficiency biofuels industry should develop on a "level trade policy playing field" and it is important that the sector does not receive inordinate ‘shelter’ (Motaal, 2008). However, at present biofuel operations are heavily subsidized globally and in 2007 alone tax credits in the developed world totaled 10bn (Anonymous, 2008), prompting a UK legislation requiring biofuels to show

significantly smaller carbon footprints than their petroleum-based equivalents in order to keep their government subsidies.

1.6 Oils for food or fuel?

1.6.1 Non-edible oilseed plants for biodiesel

An agricultural revolution is brewing to feed a projected world population of 10 billion or more. Any population growth will be accompanied by desires for adequate energy. Petroleum based energy sources have been sufficient until now, however, that will most certainly not be the case in the long-term future. Current and future demands for biodiesel are encouraging the diversion of food and feed crops into fuel. Under present tax and subsidy regimes the price increases in US corn are predicted to increase the world corn price by 20% by 2010 (Fairless, 2007; Nature editorial, 2007). According to a World Bank report a 1% increase in the price for staple food leads to a 0.5% drop in calorific consumption (Fairless, 2007; Nature editorial, 2007) which is likely to impact most significantly on some of the world's poorest people. Using edible crops for fuel will also drastically alter land use pattern because the need for food supplies will have to be maintained. Whilst there are growing concerns over the use of edible crops such as rape, soy and palm for fuel oil, the use of non-edible seed oils or the use of direct bioconversions from waste (Demirbas, 2007; 2008a; 2008b;) is now considered a major alternative. In principle any oil rich plant seed can be used as a source of biodiesel. Azam *et al.,* (2005) conducted an extensive study that compared 75 non-traditional oilseed plant species containing at least 30% w/w seed oil. The FAME compositions, iodine value (IV) and cetane number (CN) were compared to assess the suitability as a feedstock for biodiesel production. Their analysis revealed that nearly one third of the plants analyzed contained suitable seed oils and made a strong case against using edible oils for biodiesel. In fact the seed oil of 26 potentially useful species actually meets the diesel standards of the USA, Germany and the European Standards Organization (ESO; Azam *et al.,* 2005). The list of alternative oilseed plants can be narrowed further to a few useful plants according to their oil productivity per hectare; potential economically usef ul by-products; growth habit (tree or shrub); habitat (arid, semi-arid, and tropical) and cultivation requirements such as fertilizers, water and the need for plant protection. Taking all of these criteria into consideration *J. curcas* was recommended as one of the most suitable sources of non-edible oilseeds for biodiesel feedstock (Azam *et al.,* 2005; Chettri *et al.,* 2008) without comparisons to the

edible oilseed-based biodiesel sector for yield of oil per unit area. However, the conclusions of an international conference on *J. curcas* in 2007 stated that 'the positive claims on *J. curcas* are numerous, but that only a few of them can be scientifically substantiated (Jongschaap *et al.,* 2007) . The claims that have led to the popularity of the crop are based on the incorrect combination of positive characteristics which are not necessarily present in all *J. curcas* accessions and have certainly not been proven beyond doubt in combination with its oil production'(Jongschaap *et al.,* 2007). Nevertheless the popularity of *J. curcas* has been pushing research on claimed benefits and a number of these are being increasingly documented through peer-reviewed publications as is partly obvious with regard to the co-product utilization (see Table 5) and *J. curcas* life-cycle sustainability assessment for rural development potential (Achten *et al.,* 2007). If most 'traditional' claims can be validated then the improved elite varieties of *J. curcas* may live up to projected promises. Francis (2008) argued that the success of *J. curcas* as a biofuel crop would largely depend on a 'refined set of agronomic plantation management practices' and realization of the value of principal co-products, especially the seed cake. Recently, methods for detoxifying the seed cake have been proposed and also a number of uses for toxic seed cake have been proposed. These are discussed below,

1.7 The case for *Jatropha curcas*

From the data presented by Azam *et al.,* (2005), it is clear that the two plants most suitable for biodiesel are *J. curcas* and *Pongamia pinnata*. The presence of toxins like phorbol esters, curcins, trypsin inhibitors, saponins, phytate, lecithin in *J. curcas* and fluranoflavones, fluranoflavanols, chromenoflavones, flavones and furanodiketones as toxins in *P. pinnata* make these crops non edible. Of the two, *P. pinnata* is a bigger tree and not readily amenable to pruning whereas *J. curcas* is a shrub that can be checked from growing tall and a pruning regime standardized to balance the vegetative and the reproductive growth for maximum yields (Jongschaap *et al*., 2007). Therefore, the ease of harvesting the fruit from *J. curcas* in comparison to *P. pinnata* makes the former a more acceptable source of seed oil when grown as a large-scale, plantation crop. Recently, the strengths of *Jatropha* have been summarized (Figure 3) by Brittaine and Lutaladio (2010).

In terms of oil quantity, seeds of soyabean contain 20% oil, seeds of rapeseed and *Jatropha* contain 40% oil and that or *Pongamia* contain 33% oil. Furthermore,

Shuit *et al.,* (2010) demonstrated a one step reactive extraction procedure for *J. curcas* which led to maximum conversion of 99.8% oil to biodiesel.

Biodiesel production from non-food feedstocks is thus gaining interest. China recently announced an area the size of England for production of *Jatropha* and other important feed stocks. India also decided to use 60 million hectares of non-arable land for *Jatropha* (Biodiesel 2020, 2009). In India, the reseacher estimates 15 billion litres of *Jatropha* biodiesel could be produced by cultivating the crop on 11 million hectares of wasteland (Mandal, 2005). Brazil and Africa are also investing in *Jatropha* plantations in order to secure the biofuel program.

The origin of *J. curcas* is still unclear but there is some evidence that it originated from Central America (Heller, 1996). *J. curcas*, also known as 'physic nut', 'pignut', 'vomit nut' or 'fig nut' is a perennial, monoecious shrub and Heller (1996) has provided a detailed morphological description of the plant. The inflorescence is a complex cyme formed terminally and the unisexual male flowers contain 10 stamens arranged in 2 distinct whorls of 5 each while the female flowers contain the gynoecium with 3 slender styles dilating to a massive bifurcate stigma. The fruit is an ellipsoid capsule 2.5-3 cm long, 2-3 cm in diameter, yellow, turning black that contains 3 black seeds per fruit. Seeds are ellipsoid, triangular-convex in shape and measure 1.5-2 x 1-1.1 cm. Recently, a number of physical and mechanical properties of the fruits, nuts and the kernels have been described in terms of their importance in harvesting, handling and processing the oil for biodiesel production (Sirisomboon *et al*., 2007).

- *Jatropha* has the potential, through varietal improvement and good farming practices, for a high level of oil production per unit area in the subhumid tropical and subtropical environments.
- *Jatropha* grows and is potentially productive in semi-arid on degraded and saline soils.
- *Jatropha* can be used for halting and reversing land degradation.
- *Jatropha* grows fast, as compared to many tree-borne oilseed.
- *Jatropha* trees remain small, enabling ease of management.
- *Jatropha* has periodic leaf shedding which facilitates nutrient recycling and dry season irrigated intercropping which short-term crops.
- *Jatropha* leaves are unpalatable to grazing livestock, making it a good barrer hedge to protect crops.
- *Jatropha* oil has physical and chemical properties that make it highlly suitable for processing into biodiesel.
- *Jatropha* oil can be used directly in suitable diesel engines, lamps and cooking stoves.
- *Jatropha* by-products have potential value, such as using seed cake as fertilizer, animal feed (non-toxic varieties) or biogas, and using fruit shells and seed husks for biogas and combussion.
- *Jatropha* oil has markets other than for fuel, such as the production of soap, medicines and pesticides
- *Jatropha* seed are storable and processing can be delayed, which makes production suited to remote areas.
- *Jatropha* has attracted investment, mainly from the private sector, into plant breeding, which increases the likelihood of developing Jatropha varieties with improved and stable oil yields.

Figure 3: The summation strengths of *J. curcas* (adapted from Brittaine and Lutaladio (2010)

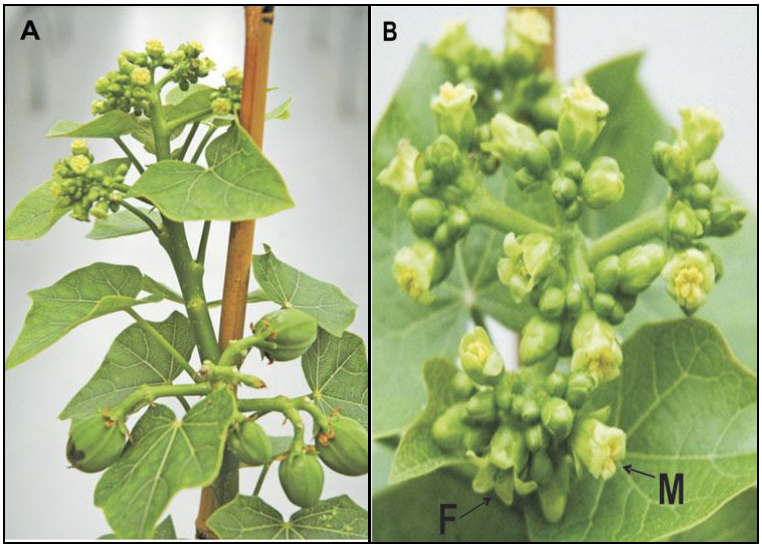


Figure 4: (A) Young *J. curcas* plant with flowers and developing seed pods. (B) J. curcas inflorescence containing both male staminate flowers (M) and female pistillate flowers (F). (King *et al.*, 2009)

1.7.1 Seed oil and biodiesel characteristics of J. curcas

The physical characteristics of *J. curcas* seeds vary depending on their geographical origin. Generally, seed weight varies from 0.4 to 0.7 g and seed dimensions vary with length and width from 15-17 mm and 7- 10 mm, respectively (Martinez-Herrera *et al.*, 2006).

Figure 5: (C) Cross-section of a *J. curcas* seed pod containing three developing seeds. (D) Mature seeds of *J. curcas.* (King *et al.*, 2009)

The kernel to shell ratio is usually around 60:40 and the seed oil content is in the range of 25 to 40% (w/w). Investigating the physico-chemical properties of *J. curcas* from Benin, Kpoviessi *et al.* (2004) reported a higher range with an upper limit close to 50%. Kaushik *et al*. (2007) observed significant differences in seed size, 100-seed weight and oil content between 24 accessions from different agro-climatic zones of the northern state of Haryana in India. They concluded that habitat and prevailing environmental conditions were more important than genotype in determining

phenotypic variation (higher co-efficient of variation) in *J. curcas* and its suitability as a fuel crop.

Oil samples of different accessions mainly contain unsaturated fatty acids in the form of oleic acid (18:1; 43 to 53%) and linoleic acid (18:2; 20 to 32%). Akintayo (2004) characterized *J. curcas* seed oil from Nigeria and found oleic acid as the most abundant fatty acid. Martnez-Herrera *et al.* (2006) have characterized four provenances of *J. curcas* from different agro-climatic regions of Mexico (Castillo de Teayo, Pueblillo, Coatzacoalcos and Yautepec) and showed that seed kernels were rich in crude protein (31–34.5%) and lipid (55–58%). Here also the major fatty acids found in the oil samples were oleic (41.5–48.8%) and linoleic acid (34.6–44.4%) though palmitic and stearic acids were also reported in lower amounts (10- 13% and 2.3-2.8% respectively) . Work in our own laboratories (Popluechai *et al.,* 2008) shows that the oil yield of *J. curcas* seeds from 6 provenances in Thailand varies from 20 to 35% (w/w). The major fatty acids found in the oil samples were oleic (36–44%), linoleic (29–35%), and palmitic (12–14%) and stearic (8–10%); these values are in general agreement with the *J. curcas* oil compositions reported elsewhere. The elevated levels of oleic and linoleic acids make the respective FAMEs suitable for biodiesel. Reksowardojo *et al*., (2006) compared five diesel types–petrodiesel, *J. curcas* B10, B100 and palm oil B10 and B100 and found the biodiesels to be more efficient in direct injection (DI) engines. A comparison of properties of petroleum diesel with *J. curcas* biodiesel (Table 4) shows that is does provide a suitable replacement for diesel.

Table 4: Comparison of properties of Diesel and *Jatropha* biodiesel

Parameter	**Diesel**	***Jatropha* Oil**
Energy content (MJ/kg)	42.6 - 45.0	39.6 - 41.8
Spec. weight (15/40 °C)	0.84 - 0.85	0.91 - 0.92
Solidifying point (°C)	-14.0	2.0
Flash point (°C)	80	110 - 240
Cetane value	47.8	51.0
Sulphur (%)	1.0 - 1.2	0.13

(Taken from http://www.jatropha.net/use-of-oil.htm)

Furthermore, Foidl *et al.,* (1996) showed that both methyl and ethyl esters of *J. curcas* fatty acids could be used without engine modification. However, the use of pure *J. curcas* biodiesel (B100) is contested (Wood, 2008) due to high NOx emissions (Reksowardojo *et al*., 2006) and it has been shown that engine performance can be improved with petro- and biodiesel blends (Pramanik, 2003). Blending with mixes of biodiesel e.g. from *J. curcas* and palm oil showed better stability at low temperature

and also improves oxidation stability compared to the use of either *J. curcas* or palm oil biodiesel alone (Sarin *et al.,* 2007). Corrosion tests on engine parts and emissions analysis showed that both *J. curcas* and palm oil biodiesel make acceptable substitutes for petro-diesel (Kaul *et al.,* 2007; Reksowardojo *et al*., 2006). The effects of using biofuels in internal combustion engines have recently been reviewed by Agarwal (2006), while Rao *et al*., (2007) specifically compared diesel to *J. curcas* biodiesel and its blends with diesel. Their results indicated that ignition delay, maximum heat release rate and combustion duration were lower for *J. curcas* biodiesel and its blends compared to diesel. *J. curcas* had lower tail pipe emissions than diesel except for nitrogen oxides (Reksowardojo *et al*., 2006). Pradeep and Sharma (2007) recently addressed the latter with *J. curcas* biodiesel and concluded that hot exhaust gas recirculation (EGR) was an effective solution to reducing nitrous oxides. Research has also demonstrated that effective routes can be found to obtaining *J. curcas* biodiesel with minimal free fatty acids (FFA; Berchmans and Hirata, 2008; Tiwari et al., 2007) so that seed oil from varieties with relatively large FFAs can also be used as long as the C16:C18 ratio is acceptable.

1.7.2 Potential J. curcas co-products

Co-products produced during and after the extraction and conversion of the oil, provide added value to the use of *J. curcas* as a bioenergy crop. Table 5 lists some of the products that can be obtained from different parts of the plant and are identified as useful in the energy, chemical, medical, cosmetics or other industries. An array of low- to high-value products obtained from the plant would help to realize the concept of a 'biorefinery' whereby the entire plant is used to ensure maximum commercial, social and environmental impact. The potential of *J. curcas* seed oil as a source of biodiesel and the expected availability of residue materials for exploitation as co-products, is encouraging research on their processing and economic validation. For example, biodiesel production generates fruit husk that can be used as feedstock for open core down draft gasifiers (Vyas *et al*., 2007) for wood gas. Singh *et al*., (2008) reported on using combustion, gasification and oil and biodiesel extraction from *J. curcas* husk, shells and seeds for a holistic approach to optimal utilization of all parts of the fruits towards obtaining maximum possible energy. Recently, Sricharoenchaikul *et al*., (2008) demonstrated that *J. curcas* waste pyrolyzed at 800^{o}C followed by KOH activation could generate a low-cost activated carbon as adsorbent with desirable surface properties for application in various industries. The glycerin generated as a by-product

of transesterification of the oil can be used to convert it to a number of value added products (Pagliaro *et a*l, 2007). The seed cake provides a rich source of protein when detoxified of its toxins, co-carcinogens and antinutrients factors providing an excellent animal feed (Devappa and Swamylingappa, 2008). Also, Mahanta *et al*., (2008) have demonstrated a more high-value use of the seed cake, without the need for detoxification, using solid-state fermentation with *Pseudomonas aeruginosa* to produce enzymes such as proteases and lipases.

Plant Part	Uses	References
Seeds	Wastewater copper biosorption by seedcoat	Jain *et al*. (2008)
	Seed diterpenes and toxins as antimicrobial	Goel *et al*. (2007)
	Antitumour, molluscicidal and insecticidal	Luo *et al*. (2007); Liu *et al*. (1997)
	Jatropholone used to prevent gastric lesions	Pertino *et al*. (2007)
Seed oil	Biodiesel	Kaushik *et al*. (2007)
	Resins and varnishes	Patel *et al*. (2008)
	Soap, lubricants and illuminants	Roy (1998)
	Antimicrobial	Eshilokun *et al*. (2007)
	Biopesticide	Goel *et al*. (2007)
Seed cake	Protein-rich feed	Devappa and Swamylingappa (2008)
	Organic fertilizer	Mendoza *et al*. (2007)
	Lipase and protease production	Mahanta *et al*. (2008)
	Activated carbon	Sricharoenchaikul *et al*. (2008)
Fruit husk	Used as fuel through combustion	Singh *et al*. (2008)
	Pyrolysed for biogas	Vyas and Singh (2007)
Leaf/extract	Feed for silkworms in silviculture	Grimm *et al*. (1997)
Oil	Larvicidal	Rahuman *et al*. (2008)
	Bactericidal	Eshilokun *et al*. (2007)
	Fungicidal	Onuh *et al*. (2008)
Stem/bark	Dark blue dye	Srivastava *et al*. (2008)
	Waxes and tannins	Burkill (1985)
	Jatrophone is antineoplastic	Biehl and Hecker (1986)
Subterranean stems	Jatrophone used against snake bite	Brum *et al*. (2006)
Root	Antibacterial diterpenoids	Alyelaagbe *et al*. (2007)

Table 5: Table showing the list of various uses of the *J. curcas* plant

The seed cake has also been used as a nutrient rich substrate to grow *Fusarium monoliforme* for gibberellin production (Udhaya Kannan – personal communication). In the seed cake the co-carcinogenic compounds, diterpene phorbol esters, are a major cause of concern limiting the wide-spread commercialization of *J. curcas.* However, some diterpenes are known for their antimicrobial and antitumor activities. For example, *J. curcas* was shown to contain jatropholone A and B (Ravindranath *et al*.,

2004) which have been recently shown to have gastro-protective and cytotoxic effects (Pertino *et al* ., 2007). Other terpenoids such as Jatrophatrione and acetylaleuritolic acid obtained from other species of *Jatropha* have also been shown to have antitumor activities (Torrance *et al*., 1976; 1977). SJ23B, a jatrophane diterpene activates classical PKCs and displays strong activity against HIV in vitro (Bedoya *et al.,* 2009). The terpenoid compounds can be used as biopesticides and bioinsecticides. Goel *et al*., (2007) have recently reviewed the documented beneficial effects of phorbol esters from *Jatropha spp* and concluded that any such uses must also consider the fate of the terpenoids in the soil, water, plants and human health following application. Burkill (1985) has provided an extensively referenced list of uses of various parts of *J. curcas* in the Kew Royal Botanic Gardens entry for the plant. Table 5 provides a list of some of those and additional new uses that have recently been documented. Min and Yao (2007) have strongly recommended the use of *J. curcas* for high-value compounds along with the biodiesel industry. These high-value co-products paint a rather rosy scenario for the economic feasibility of *J. curcas*-based biodiesel enterprise, but it must be noted that a number of high-value uses of the seed cake studied (Table 5) were on unmodified, toxic seed-cake. It remains to be seen if seed cake that has been detoxified or modified for no/minimal toxins/PE can still be a source of such compounds. If the answer is negative, then does the economics of 'toxin-less' seed cake as a source of animal feed make *J. curcas* -based biodiesel enterprise economically feasible? Use of *J. curcas* for biodiesel or high-value products will require specialized varieties for each of the two sectors, but in either case dealing with plant toxins and irritants is an issue and must be addressed.

1.7.3 Rural, socio-economic and ecological sustainability

Traditionally, *J. curcas* is used as a biofence in India and other countries in Africa and Asia. Although its primary purpose is to keep cattle away from crop plants, planting *J. curcas* also serves as a means of reducing soil erosion. It is claimed that as a drought tolerant perennial in which the root system helps to hold the soil structure together in poor soil types, *J. curcas* helps maintain valuable top-soil, which might otherwise be lost due to erosion. Recently, Achten *et al*. (2007a) have reported on studies on *J. curcas* root architecture and its relation to erosion control. Their preliminary results exhibited 'promising erosion control potential' but they stressed that further research was needed to optimize the agro-forestry and plantation systems for *J. curcas*. In another study the same group (Achten *et al*., 2007b) described the rural,

socio-economic development potential and soil rehabilitation potential of *J.curcas* plantations and highlighted the need to consider the local land, water and infrastructure conditions before undertaking *J. curcas* plantations. Studies by Maes *et al.,* (2009) on 83 *Jatropha* plantations worldwide, showed that *Jatropha* isn't common in regions with arid, semi-arid and frosty climate. They require much more humid climate than is generally specified (average rainfall of 944 mm/year). Also in frosty conditions, the plant immediately starts shedding leaves and will eventually die if the conditions don't improve quickly (Maes *et al.,* 2009a). Adequate irrigation and frost risk free facilities are needed to get optimum yields. The qualitative sustainability assessment of Maes *et al.* (2009b), focusing on environmental impacts, socio-economic issues and labour intensive cultivation methodologies involved, suggested that *J. curcas* based biodiesel enterprises could help support rural development provided wastelands or degraded lands were planted. Francis *et al.*, (2005) described the potential of *J. curcas* in simultaneous wasteland reclamation, fuel production, and socio-economic development in areas with degraded land in India. Similar but also preliminary studies with examples from different countries in Asia and Africa have been described in the proceedings of the FACT Foundation symposium on *J. curcas* (Jongschaap *et al*., 2007). These studies supported the potential of *J. curcas* for land/soil rehabilitation and for contributing to rural socio-economic development in a sustainable manner. However, the same report warned on extrapolating results from small-scale episodic planting of *J. curcas* to plantation-scale because as a crop the plant would indeed have its requirements of water, nutrients, and fertilizers for optimal yield even from unimproved varieties.

An economic feasibility study for *J. curcas* undertaken by the business school of the University of California Berkeley, for opportunities in India considered both the rural development business model and the large-scale industrial model (Khan *et al.,* 2006). Identifying the cost drivers to crop production, the study showed that the cost required during the first three years was substantially higher compared to the ongoing marginal costs of maintaining the crop once established with these low marginal cost contributing to the high profit margins. Investments in the first three years could be recovered by the sixth year and profits thereafter were seen to be high. Forecasting cash flows over 50 years–the productive lifetime of a shrub–positive net present value and high internal rate of return were reported. It was concluded that due to immense interest in its potentialities for holistic and sustainable development on different fronts, *Jatropha* -based business was viable provided appropriate tax, subsidy and biofuels policies created enough demand for the biodiesel. In both models the critical point was

minimising the risk and uncertainty of feedstock supply through research and development.

1.8 Research and development issues for *Jatropha*

The projected socio-economic, environmental and political advantages of *J. curcas* as an alternate energy crop have attracted interest from both businesses and governments. Countries in Asia, Africa and Latin America with vast tracts of non-agrarian marginal land are hastily putting policies in place to promote biofuels. Starting with India and Thailand (Bhasabutra and Sutiponpeibun, 1982; Takeda, 1982; Banerji *et al* ., 1985), *J. curcas* is becoming increasingly popular not only with other developing and transition economies in the region but even with Brazil where *J. curcas* biodiesel potential is being explored for trucks and lorries (URL 1), since their bioethanol markets cater mainly for cars. It is even claimed that in Parana state of Brazil there are some *J. curcas* varieties that are frost tolerant/resistant (URL 2). If true, this is expected to increase interest in *J. curcas* as the more temperate countries of Europe and the North America might also adopt it as a potential alternative to edible oils. Molecular responses of *J. curcas* to cold stress have been recently reported to better understand routes to generating cold tolerant varieties, having shown how early acclimation of photosystem II and the late stage H_2O_2 scavenging play an important role in cold tolerance (Liang *et al*., 2007). Although, Thailand is now increasingly exploring the bioethanol from cassava route for biofuels, *J. curcas* remains highly popular as a source of biodiesel in India. Worryingly, however, the increasing popularity of *J. curcas* is fast outpacing research efforts to improve it for delivering on its potential at a large scale-whether through small farm cooperatives or through extensive plantations. There are major issues surrounding the large-scale adoption of *J. curcas* for biodiesel and all its ancillary benefits. Despite the proposed benefits of *J. curcas* for biodiesel, biogas, biofertilizer, biopesticide, medicine, cosmetics, rural development and soil and environmental amelioration, *J. curcas* remains an undomesticated plant with variable, often low and usually unpredictable yield patterns. To obtain optimal benefits from a potentially useful plant, *J. curcas* needs to be improved with respect to the following key research areas.

1.8.1 Levels of toxins and co-carcinogens

Oil from *J. curcas* seeds is more valuable than the seed meal. However, the seeds contain elevated levels of toxic and anti-nutritive factors (ANFs) such as saponins, lectins, curcins, phytate, protease inhibitors and phorbol-esters (Martinez-

Herrera *et al.,* 2006) are generally considered toxic to humans and animals when ingested. Other studies also report on the toxicity of seeds, oil and press cake (Gubitz *et al.,* 1999) and these were recently reviewed by Gressel (2008). Methods to detoxify the seed cake have been proposed (Devappa and Swamylingappa, 2008) however the co-carcinogenic phorbol esters (PEs) and curcins are not easily and cost-effectively removed. The difficulty of removing PEs not only from the seed cake but also from the oil and even from the biodiesel itself, currently limits the commercial exploitation of *J. curcas* and its co-products, despite their potential added value. Recently, Devappa *et al*. (2010) developed phorbol esters (PEs) extraction method without affecting PEs bioactivity. PEs can be used as an agricultural bio-control and pharmaceutical applications. The ability to use the seed meal as an animal feed will increase the economic feasibility of *J. curcas* due to its use both as fuel and feed (King *et al.,* 2009). Gressel (2008) proposed that transgenic strategies would be the most effective route to obtaining toxin and PE free varieties of *J. curcas*. Use of transgenic approaches to achieve reduction/elimination of PE would require the identification of the relevant PE synthesis genes in *Jatropha* and the availability of appropriate plant transformation protocols. Although transformation of *J. curcas* using *Agrobacterium* and particle bombardment were reported (Li *et al*., 2008 and Purkayastha *et al*, 2010), transforming *J. curcas* to remove or minimise PE contents might not be straightforward because the diterpenoid biosyntheis pathway used for PE synthesis is the same as that for key plant hormones such as gibberellins and abscisic acid. Recently, *Jatropha* 3-Hydroxy-3-methylglutaryl coenzyme A reductase (JcHMGR), a key regulatory enzyme for phorbol esters synthesis in mevalonic acid (MVA) pathway, has been isolated and characterized (Lin *et al.*, 2009). Seed specific gene silencing of JcHMGR might be applied to reduce/eliminate PEs only in *Jatropha* seed without effectted key plant hormones. It is important to note that PEs found in high concentrations in the *J. curcas* kernels of different accessions were not detected in the samples from Castillode Teayo, Pueblillo and Yautepec in Mexico (Makkar *et al.*, 1998) where the seeds are traditionally eaten as roasted nuts. This offers the possibility of using conventional plant breeding to generate cultivars lacking PEs. However, as we discuss below, the limited genetic variability amongst *J. curcas* accessions from across the world did not result in reduced PE cultivars when toxic and non-toxic accessions were crossed (Sujatha *et al.,* 2005). However, more extensive breeding efforts may offer the prospect of reducing PE content in the plant by conventional or marker assisted breeding. We have successfully obtained hybrids with markedly lower PE levels using inter-specific breeding

programmes with *J. curcas* and *J. integerrima* (Popluechai *et al*., 2009). This approach has been corroborated by recent work in Dr Hong Yan's group (personal communication) who also found lower PE levels in some hybrids of *J. curcas* and *J. integerrima.* Recently, a cost-effective method for the reduction of PE to less than 15 ppm in the seed cake has been reported (H.P.S. Makkar–personal communication) but the method is currently in the process of being patented. The method uses the carp fish as a model, which is sensitive to more than 15 ppm PE, however, for other animal and human exposures much lower ppm of PE is desirable. It is not clear to us if the method at this stage is suitable for generating seed cake with the desired reduced levels of PE. It was shown previously that deacidification and bleaching could reduce the content of phorbol esters up to 55% in the seed oil of *J. curcas* (Hass *et al.,* 2000). Recent studies by Makkar *et al.,* 2009 revealed that a stripping/deodorization of oil at 260°C at 3 mbar pressure with 1% steam injection can lead to either removal or degradation of phorbol esters from *J. curcas* oil.

1.8.2 Limited genetic variability

Genetic diversity amongst *J. curcas* accessions was compared by Reddy *et al.,* (2006) using AFLP/RAPD on 20 Indian accessions for identification of drought and salinity tolerance. An 8-10% AFLP- and 14-16% RAPD-mediated polymorphism was found amongst accessions suggesting that intra-specific genetic variation was limited in *J. curcas*. Recently, Basha and Sujatha (2007) used 42 accessions of *J. curcas* from different regions in India to identify genetically divergent materials for breeding programmes. They also included a non-toxic genotype from Mexico and reported inter-accession molecular polymorphism of 42.0% with 400 RAPD primers and 33.5% with 100 ISSR primers indicating only modest levels of genetic variation amongst Indian cultivars. However, the Mexican variety could easily be differentiated from the 42 Indian accessions suggesting that almost all Indian accessions had a similar ancestry and that using accessions from different parts of the world may reveal genetic variation suitable for breeding programmes. Ranade *et al*., (2008) used single-primer amplification reaction (SPAR) to compare 21 *J. curcas* accessions from different parts of India using wild unknown accessions and classified accessions held in research institutes. Three accessions from the North East states were clearly different among themselves and from other accessions analysed, while most other accessions were highly similar. Shen *et al* (2010) used AFLP technique to explore genetic diversity of 38 *Jatropha* accessions grown in Hainan province, China. Low genetic diversity has been

observed as previous report above. Sudheer *et al*., (2008) studied differences among seven species of *Jatropha* from India and found *J. integerrima* to be most closely related to *J. curcas*.

We have compared 17 *J. curcas* and one *J. podagrica* accession using RAPD (Figure 6). The 17 accessions were obtained from Thailand (14; 6 provenances), India (2) and Nigeria (1), while the *J. podagrica* was also obtained from Thailand. UPGMA-mediated cluster analysis revealed two major clusters one containing all of the 17 *J. curcas* accessions and the second containing *J. podagrica*, which showed a genetic similarity coefficient of 52% with *J. curcas*. This was the same as recently reported using 26 RAPD primers by Ram *et al*., (2007). We also showed that the non-toxic accession from Mexico clustered separately from other *J. curcas* accessions and that the 6 Thai provenances could not be separated. However, the genetic similarity coefficient between the Thai and the Mexican accessions was high (0.76) as was also noted by Basha and Sujatha (2007). These results show the importance of testing accessions from wider eco-geographic regions and including assessments of other *Jatropha* species in addition to *J. curcas.* We have recently extended these studies to 38 *J. curcas* accessions from 13 countries on 3 continents, along with 6 different species of *Jatropha* from India using a novel approach of combinatorial tubulin-based polymorphism (cTBP; Breviario *et al.,* 2007). Once again, except for the accessions from Mexico and Costa Rica, which separated into independent groups, all other *J. curcas* accessions from different countries clustered together reiterating a narrow genetic base for *J. curcas* (Popluechai *et al*., 2009). Similarly, Ram *et al*., (2007) obtained 3 clusters; one cluster contained all five of the *J. curcas* accessions from India; the second contained 6 of the 7 different species investigated while the third contained a single species *J. glandulifera*. *J. curcas* accessions exhibit monomorphism even with microsatellite markers (personal communication, Gen Hua Yue). Recent studies by Basha *et al.,* 2009, demonstrated a narrow genetic base between 72 different *J. curcas* accessions from 13 different countries. These studies indicate the limitations of using intra-specific breeding for *J. curcas* improvement; as exemplified by the failure to obtain low PE hybrids in crosses between *J. curcas* toxic accessions from India and the non-toxic accession from Mexico (Sujatha *et al.*, 2005). In our laboratories in Thailand, India and the UK, we have employed inter-specific breeding and obtained hybrids with low PE and improved agronomic traits (Popluechai *et al.*, 2009). For most future efforts targeting *J. curcas* improvement, we would recommend using inter-specific hybridization. Therefore, characterization of different species of *Jatropha* with a view

to selecting the hybridizing parent carrying the desirable trait becomes even more important. Recently, Sunil *et al*., (2008) recorded the phenotypic traits of plants *in-situ* for plants from 4 different eco-geographical regions of India, to develop a method for identification of superior lines of *J. curcas*. Rich diversity found in Mexican genotypes in terms of phorbol ester content and distinct molecular profiles indicates the need to exploit and further research of germplasm from Mexico in *J. curcas* breeding programmes (Basha *et al.*, 2009). Also development of molecular markers will help differentiate the non-toxic varieties from the toxic ones in the mixed population (Pamidimarri *et al.*, 2009). This well help to improve the *Jatropha* species through marker assisted breeding programs and thus produce agronomically improved varieties of *Jatropha*.

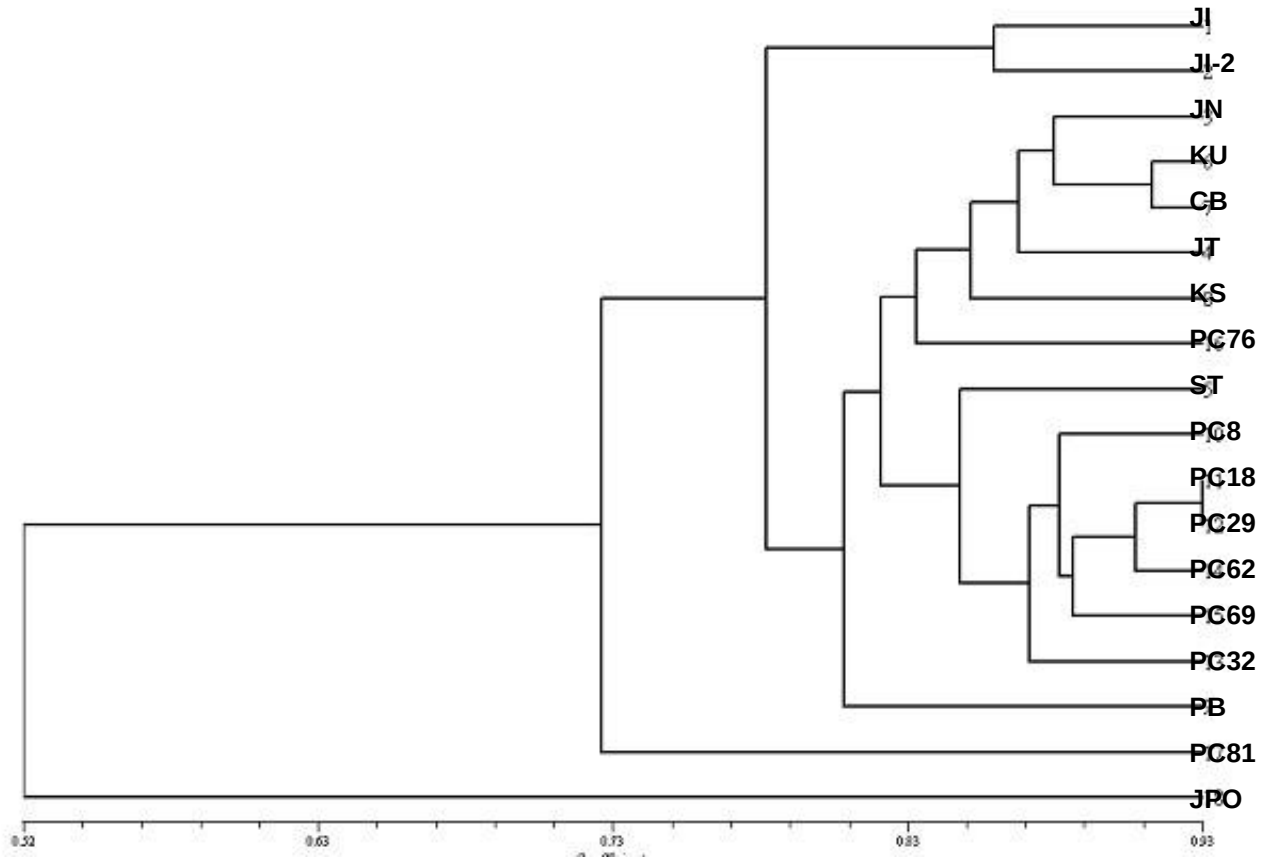


Figure 6: UPGMA analysis of RAPD conducted on 17 *J. curcas* and one *J. podagrica* accession. Origin of accessions identified as follows: JI & JI-2 – India; JN – Nigeria; JT – Thailand. All other accessions were from different provenances of Thailand represented here as TP. ST – TP1; KU – TP2; CB – TP3; KS – TP4; PB – TP5; PC8, 18, 29, 32, 62, 69, 76, 81 – TP6 and JPO – Thailand *J. podagrica.*

1.8.3 Seed germination

J. curcas is propagated through cuttings and seeds. Cuttings are useful because they maintain the clonal nature and hence the plants original characteristics. However, the vigor, vitality and yield of the plant suffer as a consequence of successive propagation (Goleniowski *et al.,* 2003; Clark and Hoy 2006); thus seed germination of identified stock is preferred. We have observed that the germination efficiency of fresh *J. curcas* seeds is good for up to six months but that after this period germination rate is markedly reduced and can reach as low as 30% in seeds stored for more than a year. Due to its thick seed coat *J. curcas* germination needs good moisture conditions while the temperature, light and oxygen conditions are, as for other seeds, also important. *J. curcas* seed germination is similar to the epigeal germination of the castor seed where the *hypocotyl* elongates and forms a hook, pulling rather than pushing the cotyledons and apical meristem through the soil. Germination rates in raised beds were found to be higher than in sunken beds (Sharma, 2007). Recent studies demonstrated that pre-sowing seed treatments (seeds kept under stone sand and moistened once with the water) led to 95.85% success rate in seed germination (Islam *et al.,* 2009). Thus, the development of standardized plantation-scale seed germination practices, founded on detailed analyses of seed metabolites in relation to varying germination rates, are needed. Studies on seed germination of *Ricinus*–a related *Euphorbiaceous* plant, have provided a valuable insight into seed oil content, composition, and hydrolysis (Kermode *et al*., 1985a, b, c; 1986; 1988; 1989a, b). Recently, *Jatropha* phospholipase D (JcPLD), a key enzyme involved in phospholipid catabolism, has been isolated and characterized (Liu *et al*, 2009). We used the 2D-PAGE analysis approach and identified 6 patatin-like lipases that were up-regulated in germinating seeds. Given that during castor seed development lipases have also been shown to be up-regulated (Eastmond, 2004), we have cloned these genes to investigate their role in determining seed oil content, composition and susceptibility to hydrolysis.

1.8.4 Reproduction

The plant produces unisexual flowers in cymose inflorescences with the male and female flowers produced in the same inflorescence. Female flowers are different to the male flowers in shape, are relatively larger and are borne on central axes. The life span of male flower is 2 days and its pollen viability is relatively higher in 9 hours after blooming until approximately 40 hours. The life span of female flowers is 5-12 days and its stigmatic

receptivity is stronger in the first 4 days and it falls or looses receptivity completely by the 9th day (Kun *et al.,* 2007). In *Jatropha*, various issues of flower development directly affect yield. For example, a high ratio of male to female flowers in the inflorescence leads to few female flowers setting seeds. Solomon-Raju *et al.,* (2002) reported that one inflorescence can produce 1–5 female and 25–93 male flowers. The average male to female flower ratio is 29:1 (Solomom-Raju and Ezradanam 2002). However, this ratio might be expected to vary amongst the different provenances. Prakash *et al.,* (2007) have reported that the male to female flower ratio increased in the second year crop and that flowering flushes numbered from one to three per year. Aker *et al*., (1997) monitored the episodic flowering dependent on precipitation in Cabo Verde in 1994. First flowering started in early May within 2-5 days of the first rains at the end of April and lasted nearly one month, being the largest of the four flowering episodes. The second flush resulted from rain on 17 May and started on 9 June with few flowers. Wet season started with rains again on 3 August lasting till mid-November during which time two flowering flushes were observed, one in mid-August and another in the third week of September. Solomon-Raju and Ezradanam (2002) working in South-East India reported a single flowering spread along the wet season from late July to late October, while Sukarin *et al.,* (1987), working in Thailand, observed two flowering peaks, one in May and the other in November. The flowering depends on the location and agro-climatic conditions. Normally in north India flowering occurs once a year, however in South-East state of Tamil Nadu fruiting occurs almost throughout the year. A few lines are bimodal, and flower twice a year. The flowering is mostly continuous in such types in the presence of optimum moisture (Gour, 2006). Although synchronous development of the male and female flowers has been reported by Solomon-Raju and Ezradanam (2002), the asynchronous maturity of fruits makes the harvesting of *J. curcas* fruits difficult because a certain stage of fruit development is needed for picking; manual harvesting becomes highly time consuming and problematic due to the number and variability of cycles during the season. At each cycle, picking at the correct stage also requires the retention of skilled pickers. Pollination from male flowers of other inflorescences effectively reduces the chances of success. Studies on self and cross pollination revealed that the success of fertilization in *J. curcas*, is irrespective of the type of pollination (Abdelgadir *et al.,* 2009). Hence studies on genetic and environmental factors affecting flower ratio, timing and cycles are an important area of research and development. For mechanical harvesting to be feasible and in a commercial operation this will be

essential, it will be necessary to synchronize flowering/fruit set as well as reduce plant height. There is substantial scope here for plant breeding initiatives to address and resolve these critical issues for genetic improvement. Also plants with desirable male to female ratio (lower ratio) as reported by Prakash *et al.,* (2007) will act as a marker to use these plants as parents for the future breeding programs. Interestingly, *J. curcas* accessions that exhibit synchronous flowering and low male to female flower ratio have now been recorded (Personal communication-Dr. Sujatha Mulpuri).

1.8.5 Propagation

There are two major methods used to propagate *Jatropha*, generative and vegetative propagation. Generative propagation is normally through direct seedling to plantation or transplanting from the nursery. The seeding method, seeding depth, seasonal timing, quality of the seed, and spacing are all important factors effecting growth and yield in *Jatropha*. Chikara *et al.,* (2007) have reported on the effect of plantation spacing on yield and showed that seed yield per plant was increased significantly by increasing the spacing between plants; obviously this reduced the seed yield per hectare. The optimum spacing could vary in different climatic and soil types. Feike *et al.,* (2007) observed from his studies on the *Jatropha* plants, that propogation by distal cuttings is more reliable than by seeds as it showed the highest biomass production. For vegetative propagation, rooting of stem cuttings is a crucial step in the propagation of woody plants and there is great variability in the rooting ability of different species. Kochhar *et al.,* (2005) reported on the effect of auxins on the rooting and sprouting behaviour of *J. curcas* and *J. glandulifera* and on the physiological and biochemical changes accompanying germination. The results showed that the sprouting of buds took place much earlier than rooting in both species. Application of indole butyric acid (IBA) and naphthalene acetic acid (NAA) increased survival percentages with IBA more effective in *J. curcas* and NAA in *J. glandulifera*. The number and type of roots per cutting is also an important factor. Roots of *J. curcas* seedlings are of the tap root system and as such are characteristically drought tolerant and play an important role in erosion control. However, roots of stem cuttings lack the main taproot and are more fibrous in nature affecting their water harvesting potential. Propagation using *in vitro* manipulations to induce strong and deep roots and the development of suitable de-differentiation and regeneration protocols is therefore important for micro-propagation and prospective genetic engineering.

1.8.6 Biotic stress assessment and management

Although *J. curcas* was not widely cropped until about 5 years ago, its potential as an energy crop was identified earlier through a project in Mali (Henning, 1996). Consequently, the International Plant Genetic Resources Institute (IPGRI) commissioned a monograph by Joachim Heller that was published in 1996. With regard to the various pests and pathogens that affect *J. curcas*, the monograph lists the following diseases and their causal organisms: Damping off caused by *Phytophthora spp*., *Phythium spp.*, and *Fusarium spp*.; leaf spot caused by *Helminthos porium tetramera*, *Pestalotiopsis spp*. *Cercospora Jatrophae-curces*; die-back due to infection with *Ferrisia virgata* and *Pinnapsis strachani*; fruit sucking damage caused by insects such as *Calidea dregei* and *Nezara viridula*; leaf damage by *Spodoptera litura* and *Oedaleua senegalensis*, seedling loss due to *Julus spp* (Heller, 1996), and root rot incited by *Rhizoctonia bataticola* in India (Sharma, S. & Kumar, K., 2009) Despite the plant's richness in pest and pathogen-repellant toxins this is a limited list and one that is likely to grow as more plantations are developed in different parts of the world. For example, following the recent establishment of a number of *J. curcas* plantations in India, Narayana *et al*., (2006) reported an increased incidence of the *Jatropha* mosaic virus disease. Cucumber mosaic virus was reported to infect *J. curcas* in India (Raj *et al.*, 2008). *Bemisia tabaci*, the carrier of mosaic virus disease and a known pest for cassava in Uganda, was reported to infect *J. curcas* (Sseruwagi *et al*., 2006). In Nicaragua, Heteroptera an insect order has been reported to have 15 species that affect the *J. curcas* nut. Even in Australia, 60 species of 21 families belonging to 4 orders of phytophagous insects affect *J. curcas* (Shankar *et al.*, 2006). Figure 7 shows few the other insects affecting different parts of *Jatropha* plant. New sightings for the leaf pathogens on *J. curcas* leaves in the Brazilian Cerrado region was also recently noted (Dianese *et al.*, 2009). Thus, a careful assessment under different agroclimatic conditions is necessary to determine the potential biotic stressors and their management in *J. curcas* plantations, a challenge that may be confounded by manipulating through breeding and genetic engineering of toxin production in *Jatropha*.

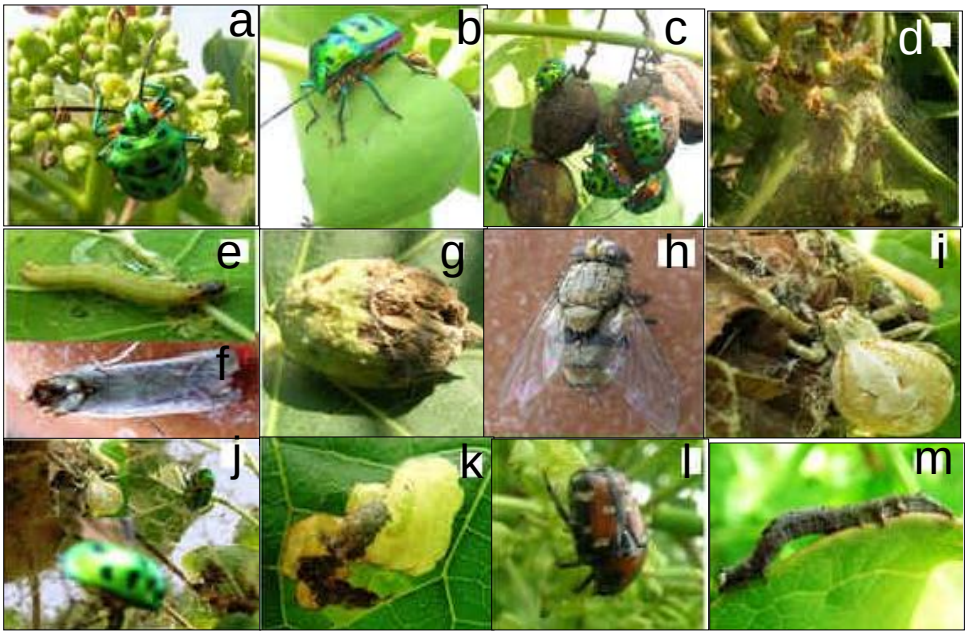


Figure 7: Pests and natural enemies of *J. curcas*. ***a***, *Scutellera nobilis* inserting stylet into flowers; ***b***, Fruits; ***c***, Bunch. ***d***, *Pempelia morosalis* webbing inflorescence. ***e***, Larvae; ***f***, Adult; ***g***, Capsule damage. ***h***, Parasite of *Pempelia*. ***i***, Spider, *Stegodyphus* sp. ***j***, Bugs ensnared in web. ***k***, *Stomphastis thraustica*. ***l***, *Oxycetonia versicolor*; ***m***, *Archaea janata* (adapted from Shankar *et al* ., 2006)

1.8.7 Agronomic requirements

Work on agronomic issues such as water requirement, the timing of its application, and the rate and timing of fertilizers, spacing etc., are largely unknown and poorly characterised. Jongschaap *et al*., (2007) have provided one-step information on different experiments conducted in different countries with regard to understanding and optimising agronomic practices for optimal *J. curcas* cultivation. The differences demonstrate that it is likely that these practices will depend not only on prevailing soil and climatic conditions, but also on any socio-economic and environmental impact expected from *Jatropha* cultivation in different countries. This will in turn be guided by wasteland availability along with other economic factors. Depending on the range of soils capable of supporting *J. curcas* plantations, soil science and agronomy have a major role to play in delivering on biodiesel *per se* and on the soil conservation and rehabilitation potential of *Jatropha*. Maes *et al.,* (2009) argued that even though there is a general consensus that *Jatropha* can grow on marginal lands and extreme weather conditions, the plant didn't function efficiently in arid, semi-arid and frosty areas. Investigating *Jatropha* seedling response to drought, they

also reported that the inability of the plants to cope up with drought is because of leaves with a higher adaxial stomatal density although the succulent stem plays a critical role in maintaining water economy. Achten *et al* (2010) also reported plant seedling biomass allocation and root/shoot were significantly influenced by drought. Fundamental studies on plant physiology of *J. curcas* will also help to define the correct agronomic practices for maximal yields. Recently, Ghosh *et al* (2010) reported application of paclobutrazol significantly reduces unwanted vegetative growth and increase seed yield as high as 1127 % relative to control. In our own work we have noted a clear difference in the water use efficiency (WUE) between 4 accessions tested in the field and under glasshouse conditions (Popluechai *et al*., 2009) indicating one accession to be much superior compared to others. An overall view of the agronomic requirements for this plant seems to rest on the fact that, natural evolutionary forces from different environmental conditions around the world, have acted upon this plant leading to create an uncertain diversity and variability within and across the *Jatropha* species.

1.8.8 Gene mining and expression analysis

In a recent article Gressel (2008) stressed the need to identify and understand the expression patterns of different genes that needed to be up-regulated or silenced as part of a transgenic approach to *Jatropha* improvement. It is only since its popularity as a biodiesel feedstock over the last 5 years that basic information on properties such as genome, genes and genotypes of *J. curcas* has become available. We have obtained pictures for chromosome spreads of the pro-metaphase of the root-tip cells (Populechai *et al.*, 2008b). This work has also confirmed the presence of 11 pairs of chromosomes, as reported earlier (Soontornchainaksaeng *et al.*, 2003). However, Carvalho *et al*., (2008) have reported the genome size to be 416 Mb. It is not clear why there is such variability in genome size or why there seems to be no consensus in plant materials that all contained 2n = 22 chromosomes. Given the importance of the seed oil, the fatty acid biosynthesis (FAB) genes in *J. curcas* were among the first to be investigated and identified. DNA sequences of some FAB genes and others of *J. curcas* have only recently become available. Important amongst these are those for steroyl acyl desaturase (Tong *et al*., 2006), acyl-acyl carrier protein thioesterase (Wu *et al.*, 2009), putative beta-ketoacyl-acyl carrier protein (ACP) synthase III (Li *et al* ., 2008), beta-ketoacyl-ACP synthase (Li *et al*., 2008), UDP-glycosyltransfersase (EMBL AAL40272.1), phosphoenolpyruvate carboxylase (EMBL

AAS99555.1), acetyl-CoA carboxylase (EMBL ABJ90471.1), omega-3 and omega-6 fatty acid desaturase (EMBL ABG49414.1), diacylglycerol acyltransferase (EMBL ABB84383.1), acyl-ACP thioesterase (EU106891), ribulose-1,5-bisphosphate carboxylase/oxygenase (EU395776). Some toxin biosynthesis genes like curcin gene (Luo *et al*., 2007) and 3-hydroxy-3-methylglutaryl coenzyme A reductase gene (Lin *et al.,* 2009) are available. Additional gene sequences now available from EMBL and NCBI are DRE-binding ERF3 genes (Tang *et al*., 2007), aquaporin (Zhang *et al.,* 2007), betaine-aldehyde dehydrogenase (Zhang *et al*., 2008), DNA-binding with one finger (Dof) (FJ605173, Yang *et al*., 2010), allene oxide cyclase (AOC) (Liu *et al*, 2010 and phenylalanine ammonia lyase (DQ883805). Recently, Ye *et al* (2009) developed high throughput screening of *Jatropha* genes function assay in leaves using virus induced gene silencing (VIGS) technique. Thirteen *Jatropha* genes involving in FA biosynthesis, developmental regulation and toxin biosynthesis have been funtionally characterized. Eswaran *et a*l (2010) used yeast funcrional screening technique to identify genes involved in abiotic stress (salt stress) in *J. curcas* seedling. These techniques can be applied to accelerate *Jatropha* gene characterization in future.

Recently the Synthetic Genomics Inc. and Asiatic Centre for Genome Technology sequenced the *J. curcas* genome. They reported a genome size of 400 Mb which is similar to the size of rice genome. The complete plastid genome and 13,201 embryo EST are also available at NCBI (access 23/03/10).

We have identified 3 oleosin genes and the upstream promoter sequence for one of them along with the gene for grain softness protein (Popluechai *et al.*, 2008) and have used these, together with those of published sequences, to study the expression patterns in different accessions and in different tissues of the plant. We have also isolated a seed specific promoter for one of the oleosin genes which will be helpful in engineering any genes targeting either the oil improvement or any other metabolic activity targeted to the seeds. We have also isolated complete genomic sequence for Phosphoenolpyruvate carboxylase gene in order to understand the basic biology of the plant. These studies are now being extended using EST, cDNA and BAC libraries to identify other genes of potential for the improvement of *Jatropha* as a fuel crop.

1.9 Conclusions

A number of options are now available as potential sources of renewable bioenergy including a variety of routes to the production of biodiesel, bioethanol and biogas. Apart from the Brazilian model of bioethanol production, only biodiesel is applicable and cost-effective over the short to medium term. For biodiesel the non-edible oilseed plants are favoured due to questions of whether good agricultural land should be used for food or fuel. As a non-edible oil crop, *J. curcas* has attracted much attention in this respect. However, the unsubstantiated claims and known facts must be balanced to have an objective view on the chances of this plant delivering its perceived potential (Jongschaap *et al*., 2007). Continued evaluation of *J. curcas* accessions in terms of their agronomic performance suggests that the optimization of simple agronomic practices can help deliver quantifiable increases in yields. However, to realize the full potential of *J. curcas* as a bioenergy crop requires the development of elite lines and elimination/reduction of the toxins and co-carcinogens such as curcin and PE. Currently the transgenic approaches needed to achieve this in *J. curcas* have been developed and optimized (He *et al.,* 2009, Purkayastha *et al*, 2010) Transformation of *J. curcas* by relevant and important genes to produce a transgenic variety with desirable traits can be seriously considered. Additionally, conventional plant breeding approaches are hampered by lack of knowledge on the genetic variability in global accessions of *J. curcas*. Vegetative and apomictic reproduction (Datta *et al*., 2005) may have contributed to maintaining and propagating a limited stock of *J. curcas* over time thereby limiting genetic variability. The capacity for apomixis in *J. curcas* can be used to stabilise hybrid vigor and to help maintain high yielding stocks. Recent descriptions of fixing apomixis in *Arabidopsis* (Ravi *et al*., 2008) may be transferrable to *Jatropha*. The importance of genetic variability is evident from the work on using inter-specific hybrid lines to breed for desirable traits despite the fact that *J. integerrima* itself lacks the desirable range of targeted traits. However, although genetically polymorphic, it is sufficiently closely related (Sudheer *et al*., 2008) to produce viable hybrids. Much has yet to be done if *Jatropha* is to become established as a commercially viable source of biodiesel and little should be expected from the widespread planting of existing varieties. There is an immediate need to include resistance to a variety of insects, pests and pathogens in the variety improvement programs. Sujatha *et al*., (2008) recently reviewed the role of biotechnology in the improvement of *J. curcas.* A combination of plant breeding and transgenic approaches should be used as a platform to develop elite varieties capable of

delivering the necessary yields and which are free of toxins. Whether or not this happens depends very much on global commodity markets and the competition from other energy technologies and energy crops.

1.10 Thesis rationale and overview

Conventional breeding could provide a useful starting point for *Jatropha* improvement because appropriate accessions such as the low PE non-toxic Mexican (NTM) accession, are available. Molecular tools to assess genetic variability within *J. curcas* would facilitate this approach. Chapter 2 outlines a route and provides supporting data for the fast-track identification of useful breeding lines by using several different screens and methods of assesment. Expression analysis of genes essential for the quantity and quality of seed/oil were targeted as molecular screens. Known sequences of four Fatty Acid Biosynthesis genes in *Jatropha* (*acc*, *sad*, *fad* and *dgat*) were used to investigate the link between their transcript levels and the FA profile. This analysis suggested that comparatively high *fad* and/or low *dgat* transcript levels was an effective screen for relatively inferior accessions. Despite low genetic variation among global accessions, a variation of phenotypes has been observed, suggesting the effect of biotic/abiotic stress via epigenetic mediated mechanisms.

An understanding of triglyceride biosynthesis-storage-utilization in *Jatropha* seed will provide important information to improve *Jatropha* in the near future. Recently, a number of fatty acids and triglyceride biosynthesis genes have been isolated and characterized (see above). These gene sequences are available at the National Centre for Biotechnology Information (NCBI). However, no information is available on the oilbodies or oleosins of *J. curcas.* Considering the increasing popularity of this plant as a source of non-edible seed oil in preference to other plants for biodiesel, it is highly pertinent that its seed oil storage organelles and their components are characterized. The nature and size of oilbodies is largely dependent on the protein component of these organelles. Chapter 3 reports the characterization of the major component protein of oilbodies – the oleosins, at the gene, transcript and protein levels. Proteomic tools were also used to identify the minor protein components of *J. curcas* oilbodies. In addition, biochemical properties of *Jatropha* oilbodies have also been characterized.

The highly seed specific oleosin promoter is a very useful tool for seed specific gene manipulation. Oleosin promoters from several plants, for example *Brassica*,

Arabidopsis, corn, *Perilla*, and soybean have been characterized. Genetic modification for the desired phenotype in *Jatropha* seed may depend on being able to limit expression of the gene of interest to the seed. Under there circumstances, seed-specific promoters will be essential to limit gene expression to the *Jatropha* seed. In chapter 4, 623 nucleotides upstream of the *JcOleosin3* gene isolated from *J. curcas* have been characterized in transgenic *Arabidopsis* plants via the *β-glucuronidase* (GUS) reporter gene.

Triacylglycerols (TAGs) form major food reserves in oilseed plants and are accumulated in the oilbodies. When the seeds germinate, the TAGs are degraded in order to produce a carbon source that supplies the embryo's postgerminative growth until the photosynthetic system of the juvenile plant becomes active. Lipase plays an important role in the conversion of TAGs in germinating seeds. In this study (Chapter 5), patatin-like lipases were identified from *J. curcas* seedlings using systematic proteome approaches.

Genetic manipulation of oil quality and quantity in oilseed crops such as *Arabidopsis thailiana* and *Brassic napus* has been carried out successfully because the triacylglycerols biosynthesis pathway of these oilseed plant are well understood. However, this information is still not available for *J. curcas*. Only some TAGs biosynthesis genes are available at NCBI. In this study, a scheme of triacylglycerols biosynthesis pathway was proposed based on the bioinformatic analysis of existing TAGs genes and data obtained from this present study (Chapter 5).

1.11 References:

Abdelgadir, H.A., Johnson, S.D. and Van Staden, J. (2009) Pollen viability, pollen germination and pollen tube growth in *Jatropha curcas*–A potential oil seed crop for biodiesel. *South African Journal of Botany* 75: 391.

Achten, W.M.J, Aerts, R., Mathijs, E., Verchot, L., Singh, V.P & Muys B (2007a) *Jatropha* biodiesel fueling sustainability? *Biofuels Bioproducts and Biorefining* . 1: 283-291.

Achten, W.M., Reubens, B., Maes, W., Mathijs, E., Verchot, L., Singh, V.P., Poesens, J. and Muys, B. (2007b) Root architecture of the promising bio-diesel plant *Jatropha*. *Communications in Agricultural and Applied Biological Sciences* 72: 81-85.

Achten, W.M.J., Maes, W.H., Reubens, B., Mathijs, E., Singh, V.P., Verchot, L. and Muys, B. (2010) Biomass production and allocation in *Jatropha curcas* L. seedlings under different levels of drought stress. *Biomass and Bioenergy* 34: 667-676.

Agarwal D., Sinha, S., Kumar, A. and Agarwal, A (2006) Experimental investigation of control of NOx emissions in biodiesel-fueled compression ignition engine. *Renewable Energy* 31: 2356–2369.

Agarwal, A.K. (2006) Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines. *Progress in Energy and Combustion Science* 33: 233–271.

Agarwal, A.K. and Das, L.M. (2001) Biodiesel development and characterization for use as a fuel in compression ignition engine. *Journal of Engineering for Gas Turbines Power* 123, 440–447.

Aker, C.L. (1997) Growth and reproduction of *Jatropha curcas*. Biofuels and industrial products from *Jatropha curcas*. Developed from the Symposium "*Jatropha* 97"Managua, Nicaragua. February 23 – 27. Dbv-Verlag für die Technische Universität Graz, 2-18.

Akintayo, E.T. (2004) Characteristics and composition of Parkia biglobbossa and *Jatropha curcas* oils and cakes. *Bioresource Technolog* 92: 307-310.

Alyelaagbe, O.O., Adesogan, K., Ekundayo, O. and Gloer, J.B. (2007) Antibacterial diterpenoids from *Jatropha podagrica* hook. *Phytochemistry* 68: 2420-2425.

Anonymous (2008) Subsidy shake up to roil biofuel market. *Chemistry & Industry*. 6: 15.

Azam, M.M., Waris, A. and Nahar, N.M (2005) Prospects and potential of fatty acid methyl esters of some non-traditional seed oils for use as biodiesel in India. *Biomass and Bioenergy*. 29: 293–302.

Bedoya, L.M., Márquez, N., Martínez, N., Gutiérrez-Eisman, S., Álvarez, A., Calzado, M.A., Rojas, J.M., Appendino, G., Muñoz, E. and Alcamí, J. (2009). SJ23B, a jatrophane diterpene activates classical PKCs and displays strong activity against HIV in vitro. *Biochemical Pharmacology* 77:965-978.

Banerji, R., Chowdhury, A.R., Misra, G., Sudarsanam, G., Verma, S.C. and Srivastava, G.S. (1985) *Jatropha curcas* seed oils for energy. *Biomass* 8: 277–282.

Barnett, T., Zwiers, F., Hegerl, G., Allen, M., Crowley, T., Gillett, N., Hasselmann, K., Jones P., Santer, B., Schnur, R., Scott, P., Taylor, K. and Tett, S. (2005) Detecting and attributing external influences on the climate system: A review of recent advances. *Journal of climate* 18: 1291-1314.

Basha, S.D. and 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, S.D., Francis, G., Makkar, H.P.S., Becker, K. and Sujatha, M. (2009). A comparative study of biochemical traits and molecular markers for assessment of genetic relationships between *Jatropha curcas* L. germplasm from different countries. *Plant Science*. 176: 812-823.

Berchmans, H.J. and Hirata, S. (2008) Biodiesel production from crude *Jatropha curcas* L. seed oil with a high content of free fatty acids. *Bioresource Technology* 99: 1716-1721.

Bhasabutra, R. and Sutiponpeibun, S. (1982) *Jatropha curcas* oil as a substitute for diesel engine oil. *Renewable Energy Review Journal* 4: 56-60.

Biehl, J. and Hecker, E. (1986) Irritant and antineoplastic principles of some species of the genus *Jatropha* (*Euphorbiaceae*). *Planta Medica* 52: 430-430.

BIODIESEL 2020: Global Market Survey, Feedstock Trends and Forecasts (2008 release), Multi-Client Study, 2nd Edition 685 pages Published by Emerging Markets Online, pp. 1-8.

BRAC: Biofuels Research Advisory Council Report (2006) A vision for 2030 and Beyond, pp. 1-39.

Breviario, D., Baird, Wm.V., Sangoi, S., Hilu, K., Blumetti, P. and Giani, S. (2007) High polymorphism and resolution in targeted fingerprinting with combined beta-tubulin introns. *Molecular Breeding* 20: 249-259.

Brittaine, R. and Lutaladio, N. (2010) *Jatropha*: A small holder bioenergy crop the potential for Pro-Poor development. *Integrated Crop Management* Vol. 8–2010. Food and agriculture organization of the united nations, pp. 1-17.

Brum, R.L., Cardoso, C.A.L., Honda, N.K., Andreu, M.P. and Viana, L.L.D.S. (2006) Quantitative determination of Jatrophone in "Cachaqa" prepared with *Jatropha* elliptica. *Chemical and Pharmaceutical Bulletin* 54: 754-757.

Burkill, H.M. (1985) The useful plants of west tropical Africa, Vol 2. Royal Botanical Gardens, Kew, pp. 90-94.

Carvalho, C.R., Clarindo, W.R., Praca, M.M., Araujo, F.S. and Carels, N. (2008) Genome size, base composition and karyotype of *Jatropha curcas* L., an important biofuel plant. *Plant Science* 174: 613–617.

Chhetri, A.B., Tango, M.S., Budge, S.M., Watts, K.C. and Islam, M.R. (2008) Non-edible plant oils as new sources for biodiesel production. *International Journal of Molecular Sciences* 9: 169-180.

Chikara, J., Ghosh, A., Patolia, J.S., Chaudhary, D.R. and Zala, A. 'Productivity of *Jatropha curcas* L. under different spacing', FACT SEMINAR on *Jatropha* curcas L.: agronomy and genetics, Wageningen, The Netherland, 26-28 March 2007, FACT Foundation; p. article number 9.

Chisti, Y. (2007) Biodiesel from microalgae. *Biotechnology Advances* 25: 294-306.

Clark, C.A. and Hoy, M.W. (2006) Effects of common viruses on yield and quality of Beauregard sweetpotato in Louisiana. *Plant Disease* 90: 83-88.

Datta, S.K., Bhattacharya, A. and Datta, K. (2005) Floral biology, floral resource constraints and pollination limitation in *Jatropha* curcas L. *Pakistan-Journal-of-Biological-Sciences* 8: 456-460.

de La Rue du Can, S. and Price, L. (2008) Sectoral trends in global energy use and greenhouse gas emissions. *Energy Policy* 36: 1386–1403.

Dianese, A.C. and Dianese, J.C. (2009). New records for the Brazilian Cerrado of leaf pathogens on *Jatropha curcas*. *Phytopathology* 99: S28-S28.

Demirbas, A. (2007) Progress and recent trends in biofuels. *Progress in Energy and Combustion Science* 33: 1-18.

Demirbas, A. (2008a) Biomethanol production from organic waste materials. *Energy sources part A–recovery utilization and environmental effects* 30: 565-572.

Demirbas, A. (2008b) The importance of bioethanol and biodiesel from biomass. *Energy sources Part B- Economics planning and policy* 3: 177-185.

Demirbas, A., Pehivan, E. and Altun, T. (2006) Potential evolution of Turkish agricultural residues as bio-gas, bio-char and bio-oil sources. *International Journal of Hydrogen Energy* 31, 613-620.

Devappa, R.K. and Swamylingappa, B. (2008) Biochemical and nutritional evaluation of *Jatropha* protein isolate prepared by steam injection heating for reduction of toxic and antinutritional factors. *Journal of the Science of Food and Agriculture* 88: 911-919.

Devappa, R.K., Maes, J., Makkar, H.P.S., Greyt W.D. and Becker K (2010) Quality of biodiesel prepared from phorbol ester extracted *Jatropha curcas* oil. *J Am Oil Chem Soc*. 87: 697-704.

Eastmond, P.J. (2004) Cloning and characterization of the acid lipase from castor beans. *Journal of Biological Chemistry* 279: 45540-45545.

EBB: European Biodiesel Board (2008) Statistics. Available at: http://www.ebb-eu.org/stats.php (accessed 25/10/10)

Eshilokun, A.O., Kasali, A.A., Ogunwande, I.A., Walker, T.M. and Setzer, W.N. (2007) Chemical composition and antimicrobial studies of the essential oils of *Jatropha* integerrima Jacq (leaf and seeds). *Natural Producs Communications* 2, 853-855.

Eswaran, N., Parameswaran, S., Sathram, B., Anatharaman, B., Kumar G R.K. and Tangrirala, S.J. (2010) Yeast functional screen to identifygenetic determinants capable of conferring abiotic stress tolerance in *Jatropha curcas*. *BMC Biotechnology* 10:23.

Fairless, D. (2007) Biofuel: The little shrub that could-maybe. *Nature* 449: 652-655.

Feike, T., Weis, K., Claupein, W. and Mueller, J. (2007). Propogation of physic nut (*Jatropha curcas* L.) on Leyte island, Philippines. *Conferenceon International Agricultural Research for development.* Tropentag 2007 Witzenhausen, Germany, pp. 1-6.

Focacci, A. (2005) Empirical analysis of the environmental and energy policies in some developing countries using widely employed macroeconomic indicators: the cases of Brazil, China and India. *Energy Policy* 33: 543–554.

Foidl, N., Foidl, G., Sanchez, M., Mittelbach, M. and Hackel, S. (1996) *Jatropha curcas* L. as a source for the production of biofuel in Nicaragua. *Bioresource Technology* 58: 77-82.

Francis, G. (2008) *Jatropha* as a biofuel crop: potential and issues. Proceedings International Technical Workshop on the Feasibility of Non-edible Oil Seed Crops for Biofuel Production May 25-27, 2007. Chaing Rai, Thailand. Eds. Syers KJ, Wood D and Thongbai P. pp. 38-48.

Francis, G., Edinger, R. and Becker, K. (2005) A concept for simultaneous wasteland reclamation, fuel production, and socio-economic development in degraded areas in India: Need, potential and perspectives of *Jatropha* plantations. *Natural Resources Forum* 29: 12–24.

Friedlingstein, P. (2008) A steep road to climate stabilization. *Nature* 451: 297-298.

Goel, G., Makkar, H.P.S., Francis, G., Becker, K. (2007) Phorbol esters: Structure, biological activity and toxicity in animals. *International journal of Toxicology* 26: 279-288.

Goleniowski, M.E., Flamarique, C., Bima, P. (2003) Micro propagation of oregano (*Origanum vulgare* xapplii) from meristem tips. *In vitro Cell and Developmental Biology* 39: 125-128.

Ghosh, A., Chikara, J., Chaudhary, D.R., Prakash, A.R., Boricha, G. and Zala, A. (2010) Paclobutrazol arrests vegetative growth and unveils unexpressed yield potential of *Jatropha curcas Journal of Plant Growth Regulation*, 29: 307-315.

Gour, V.K. (2006) Production practices including post-harvest management of *Jatropha curcas*. *Biodiesel Conference Towards Energy Independence – Focus on Jatropha on 9 – 10 June 2006* Eds. Singh, B., Swaminathan, R. and Ponraj, V., pp. 223-251.

Gressel, J. (2008) Transgenics are imperative for biofuel crops. *Plant Science* 174: 246-263.

Grimm, C., Somarriba, A. and Foidl, N. (1997) Development of eri silkworm *Samiacynthia ricini* (Boisd.) (Lepidoptera: *Saturniidae*) on different provenances of *Jatropha curcas* leaves. Biofuels and Industrial Products from *Jatropha* curcas. Symposium "*Jatropha* 97" Managua, Nicaragua. February 23 – 27. Available at: http://www.jatropha.de/conferences/ abstracts-Jatropha 97.htm (accessed 25/10/10)

Gubitz, G.M., Mittelbach, M. and Trabi, M. (1999) Exploitation of the tropical oil seed plant *Jatropha curcas* L. *Bioresource Technology* 67: 73-82.

Hass, W and Mittlebach, M. (2000). Detoxification experiments with the seed oil from *Jatropha curcas* L. *Industrial crops and products* 12: 111-118.

He, Y., Pasapula, V., Li, X., Lu, R., Niu, B., Hou, P., Wang, Y., Xu, Y. and Chen, F. (2009) *Agrobacterium* tumefaciens-mediated Transformation of *Jatropha curcas*: Factors Affecting Transient Transformation Efficiency and Morphology Analysis of Transgenic Calli. *Silvae Genetica* 58: 123-128.

Heller, J. (1996) Physic nut *Jatropha curcas* L.: Promoting the conservation and use of under utilized and neglected crops 1. Monograph Published by International Plant Genetic Resources Institute, Rome, pp. 1-62.

Henning, R. (1996) Combating Desertification: The *Jatropha* Project of Mali, West Africa. Aridlands Newsletter 40. Available at : http://ag.arizona.edu/OALS/ALN/aln40/*Jatropha*. Html (accessed 25/10/10)

Islam, A.K.M.A., Anuar, N. and Yakoob, Z. (2009) Effect of genotypes and pre-sowing treatments on seed germination behavior of *Jatropha*. *Asian journal of Plant Sciences* 8: 433-439.

IEA outlook 2009. Available at: http://www.eia.doe.gov/oiaf/ieo/emissions.html (accessed 25/10/10)

Jain, N., Joshi, H.C., Dutta, S.C., Kumar, S. and Pathak, H. (2008) Biosorption of copper from wastewater using *Jatropha* seed coat. *Journal of Scientific and Industrial Research* 67: 154-160.

Johnston, M. and Holloway, T. (2007) A global comparison of national biodiesel production potentials. *Environmental Science and Technology* 41: 7967-7973.

Jones, P.D., and A. Moberg (2003) Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. *Journal of Climate* 16: 206-223.

Jones, P.D., Parker D.E., Osborn T.J. and Briffa K.R. (2010) Global and hemispheric temperature anomalies-land and marine instrumental records. In Trends: A compendium of data on global change. Carbon dioxide information analysis center. Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A. doi: 10.3334/CDIAC/cli.002. Available at: http://cdiac.ornl.gov/trends/temp/jonescru/jones. html (accessed 25/10/10)

Jongschaap, R.E.E., Corre, W.J., Bindraban, P.S. and Brandenburg, W.A. (2007) Claims and facts on *Jatropha* curcas L: global *Jatropha curcas* evaluation breeding and propagation programme. Plant Research International Wageningen Report, pp. 1-66.

Kalscheuer, R., Stolting, T. and Alexander, S.A. (2006) Microdiesel: *Escherichia coli* engineered for fuel production. *Microbiology* 152: 2529–2536.

Kaul, S., Saxena, R.C., Kumar, A, Negi, M.S., Bhatnagar, A.K., Goyal, H.B. and Gupta, A.K. (2007) Corrosion behavior of biodiesel from seed oils of Indian origin on diesel engine parts. *Fuel Processing Technology* 88: 303 – 307.

Kaushik, N., Kumar, Kumar, K., Kumar, S., Kaushik, N. and Roy, S. (2007). Genetic variability and divergence studies in seed traits and oil content of *Jatropha* (*Jatropha curcas* L.) accessions. *Biomass and Bioenergy* 31: 497-502.

Kermode, A.R. and Bewley, J.D. (1985a) The role of maturation drying in the transition from seed development to germination. *Journal of Experimental Botany* 36: 1906-1915.

Kermode, A.R. and Bewley, J.D. (1985b) The role of maturation drying in the transition from seed development to germination .2. postgerminative enzyme-production and soluble-protein synthetic pattern changes within the endosperm of *Ricinus communis* L seeds. *Journal of Experimental Botany* 36: 1916-1927.

Kermode, A.R., Gifford, D.J, Bewley, J.D. (1985c) The role of maturation drying in the transition from seed development to germination 3 insoluble protein synthetic pattern changes within the endosperm of *Ricinus communis* L seeds. *Journal of Experimental Botany* 36: 1928-1936.

Kermode, A.R. and Bewley, J.D. (1986) The role of maturation drying in the transition from seed development to germination 4 protein-synthesis and enzyme-activity changes within the cotyledons of *Ricinus communis* L seeds. *Journal of Experimental Botany* 37: 1887-1898.

Kermode, A.R., Pramanik, S.K. and Bewley, J.D. (1988) The role of maturation drying in the transition from seed development to germination 5 responses of the immature castor bean embryo to isolation from the whole seed-a comparison with premature desiccation *Journal of Experimental Botany* 39: 487-497.

Kermode, A.R., Pramanik, S.K. and Bewley, J.D. (1989a) The role of maturation drying in the transition from seed development to germination 6 desiccation-induced changes in messenger-RNA populations within the endosperm of *Ricinus communis* L seeds. *Journal of Experimental Botany* 40: 33-41.

Kermode, A.R., Pramanik, S.K. and Bewley, J.D. (1989b) The role of maturation drying in the transition from seed development to germination 7 effects of partial and complete desiccation on abscisic-acid levels and sensitivity in *Ricinus communis* L seeds. *Journal of Experimental Botany* 40: 303-313.

Khan, A., Rajagopal, D. and Yoo, K.J. (2006) Plant oils-entrepreneurial opportunities in alternative fuel production and rural development in India. Report-MOT-UNIDO Bridging the Divide 2005-2006. University of California, Berkeley. pp. 1-18. Available at: http://www.bio2power.org/dmdocuments/India_Source_of_fuel_deepak_oct05.pdf (accessed 25/10/10)

King, A.J., He, W., Cuevas, J.A., Freudenberger, M., Ramiaramanana, D and Graham, I. (2009). Potential of *Jatropha curcas* as a source of renewable oil and animal feed. *Journal of experimental botany* 60: 2897-2905.

Kochhar, S., Kochhar, V.K., Singh, S.P., Katiyar, R.S. and Pushpangadan, P. (2005) Differential rooting and sprouting behaviour of two *Jatropha* species and associated physiological and biochemical changes. *Current Science* 89: 936-939.

Kpoviessi, D.S.S, Accrombessi, G.C., Kossouoh, C., Soumanou, M.M. and Moudachirou, M. (2004) Physicochemical proper ties and composition of pignut (*Jatropha curcas*) non-conventional oil from different regions of Benin. C. R. *Chimie* 7: 1007–1012.

Kun, L., Wei-Lun, Y. and Chang-Wei, L. (2007). Breeding system and pollination ecology in *Jatropha curcas. Forest Research* 20: 775-781.

Li, J., Li, M.R., Wu, P.Z., Tian, C.E., Jiang, H.W. and Wu, G.J. (2008) Molecular cloning and expression analysis of a gene encoding a putative -ketoacyl-acyl carrier protein (ACP) synthase III (KAS III) from *Jatropha* curcas. *Tree Physiology* 28: 921-927.

Liang, Y., Chen, H., Tang, M.J., Yang, P.F. and Shen, S.H. (2007) Responses of *Jatropha curcas* seedlings to cold stress: photosynthesis-related proteins and chlorophyll fluorescence characteristics. *Phyiologia Plantarum* 131: 508-517.

Lin , J., Jin, Y.J., Zhou, M.Q., Zhou, X. and Wang, J.Y. (2009) Molecular cloning, characterization and functional analysis of a 3-hydroxy-3-methylglutaryl coenzyme A reductase gene from *Jatropha curcas*. *African Journal of Biotechnology* 8: 3455-3462.

Liu, S.Y., Sporer, F., Wink, M., Jourdane, J., Henning, R., Li, Y.L. and Ruppel, A.(1997) Anthraquinones in *Rheum palmatum* and *Rumex dentatus* (Polygonaceae) and phorbolesters from *Jatropha* curcas (Euphorbiaceae) with molluscicidal activity against the schistosomias vector snails *Oncomelania*, *Biomphalaria* and *Bulinus*. *Tropical Medicine and International Health* 2: 179–188.

Liu, B., Yao, L., Wang, W., Gao, J., Chen, F., Wang, S., Xu, Y., Tang, L. and Jia, Y. (2009) Molecular cloning and characterization of phospholipase D from *Jatropha curcas. Molecular Biology Reports* 37: 939-946.

Luo, M.J., Liu, W.X., Yang, X.Y., Xu, Y., Yan, F., Huang, P. and Chen, F. (2007) Cloning, expression, and antitumor activity of recombinant protein of curcin. *Russian Journal of Plant Physiology* 54: 202-206.

Maes, W.H., Trabucco, A, Achten, W.M.J. and Muys, B.(2009a). Climatic growing conditions of *Jatropha curcas* L. *Biomass and bioenergy* 33: 1481-1485.

Maes, W.H., Achten, W.M.J., Reubens, B., Raes, D., Samson, R. and Muys, B. (2009b) Plant–water relationships and growth strategies of Jatropha curcas L. seedlings under different levels of drought stress. *Journal of Arid Environments* 73: 877-884.

Mahanta, N., Gupta, A. and Khare, S.K. (2008) Production of protease and lipase by solvent tolerant *Pseudomonas aeruginosa* PseA in solid state fermentation using *Jatropha curcas* seed cake as substrate. *Bioresource technology* 99: 1729-1735.

Makkar, H.P.S., Becker, K. and Schmook, B. (1998) Edible provenances of *Jatropha curcas* from Quintana Roo state of Mexico and eff ect of roasting on antinutrient and toxic f actors in seeds. *Plant Foods for Human Nutrition* 52: 31-36.

Makkar, H., Maes, J., De Greyt, W. and Becker, K. (2009) Removal and degradation of phorbol esters during pre-treatment and transesterification of *Jatropha curcas* Oil. *Journal of the American Oil Chemists' Society* 86: 173-181.

Mandal, R. (2005) 'The Indian Biofuels Programme–national mission on bio-diesel', Presentation to the International Conference and Expo, Biofuels 2012: Vision to Reality, New Delhi, 17–18 October 2005. Available at: http://www.teiin.org/events/docs/5biofuel. htm (accessed 25/10/10)

Martnez-Herrera, J., Siddhuraju, P., Francis, G., Dvila-Ortz, G. and Becker, K. (2006) Chemical composition, toxic/antimetabolic constituents, and eff ects of different treatments on their levels, in four provenances of *Jatropha* curcas L. from Mexico. *Food Chemistry* 96: 80-89.

Mathews, J. (2007) Seven steps to curb global warming. *Energy Policy* 35: 4247–4259.

Mendoza, T.C., Castillo, E.T. and Aquino, A.L. (2007) Towards making *Jatropha curcas* (tubing bakod) a viable source of biodiesel oil in the Philippines. *Philippine Journal of Crop Science* 32: 29-43.

Min, E. and Yao, Z.L. (2007) The development of biodiesel industry in recent years-Peculiarity, predicament and countermeasures. *Progress in Chemistry* 19: 1050-1059.

Mixon, J., Dack, J., Kraukunas, I. and Feng, J. (2003) A case for biodiesel. pp. 1-37. Available at: http://depts.washington.edu/poeweb/gradprograms/envmgt/2003symposium/biodiesel_finalpaper.pdf (accessed 25/10/10)

Modi, M.K., Reddy, J.R.C., Rao, B.V.S.K. and Prasad, R.B.N. (2006) Lipase-mediated transformation of vegetable oils into biodiesel using propan-2-ol as acyl acceptor. *Biotechnology Letters* 28: 637-640.

Motaal, D.A. (2008) The biofuels Landscape: Is there a role for WTO. *Journal of World Trade* 42: 61-86.

Narayana, D.S.A., Shankarappa, K.S., Govindappa, M.R., Prameela, H.A., Rao M.R.G. and Rangaswamy K.T. (2006) Natural occurrence of *Jatropha* mosaic virus disease in India. *Current Science* 91: 584-586.

Nature Editorial (2007) Kill king corn. *Nature*. 449, 637.

NBB: National Biodiesel Board – USA (2008) Statistics. pp. 1-31. Available at: http://www.biodiesel.org/resources/reportsdatabase/reports/gen/20090914_gen-396.pdf (accessed 25/10/10)

Onuh, M.O., Ohazurike, N.C. and Emeribe, E.O. (2008) Efficacy of *Jatropha curcas* leaf extract in the control of brown blotch disease of cowpea (*Vigna unguiculata*). *Biological Agriculture and Horticultur* 25: 201-207.

Pamidimarri, D.V.N. Mastan, S., Rahman, H. and Reddy, M. (2009) Molecular characterization and identification of markers for toxic and non-toxic varieties of *Jatropha curcas* L. using RAPD, AFLP and SSR markers. *Molecular biology reports* 36: 1357-1364.

Pagliaro, M., Ciriminna, R., Kimura, H., Rossi, M. and Della Pina, C. (2007) From Glycerol to Value-Added Products. *Angwandte Chemie-International edition* 46: 4434-4440.

Patel, V.C., Varughese, J., Krishnamoorthy, P.A., Jain, R.C., Singh, A.K. and Ramamoorty, M (2008) Synthesis of alkyd resin from *Jatropha* and Rapeseed oils and their applications in electrical insulation. *Journal of Applied Polymer Science* 107: 1724-1729.

Pertino, M., Schmeda-Hirschmann, G., Rodriguez, J.A. and Theoduloz, C. (2007) Gastroprotective effect and cytotoxicity of semisynthetic jatrpholone derivatives.*Planta medica* 73: 1095-1100.

Popluechai, S., Breviario, D., Mulpuri, S., Makkar, H., Raorane, M., Reddy, A., Palchetti, E., Gatehouse, A., Syers, K., O'Donnell, A. and Kohli, A. (2009) Narrow genetic and apparent phenetic diversity in *Jatropha curcas*: initial success with generating low phorbol ester interspecific hybrids. pp. 1-44. Available at: http://hdl.handle.net/10101/npre.2009 .2782.1 (accessed 25/10/10)

Popluechai, S., Raorane, M., O'Donnell, AG and Kohli, A (2008) Research progress towards *Jatropha* as alternate oilseed for biodiesel. Proc. International Technical Workshop on the Feasibility of Non-edible Oil Seed Crops f or Biofuel Production May 25-27, 2007. Chaing Rai, Thailand. Eds. Syers KJ, Wood D and Thongbai P., pp. 136-145.

Pradeep, V and Sharma, R.P. (2007) Use of Hot EGR for NOx control in a compression ignition engine fuelled with biodiesel from *Jatropha* oil. *Renewable Energy* 32: 1136-1154.

Prakash, A.R., Patolia, J.S., Chikara, J. and Boricha, GN. 'Floral biology and flowering behaviour of *Jatropha curcas*', FACT SEMINAR on *Jatropha curcas* L.: agronomy and genetics, Wageningen, The Netherland, 26-28 March 2007, FACT Foundation; p. article number 2.

Pramanik, K. (2003) Properties and use of *Jatropha* curcas oil and diesel fuel blends in compression ignition engine. *Renewable Energy* 28: 239-248.

Purkayastha, J., Sugla, T., Paul, A., Solleti, S.K., P. Mazumdar, P., Basu, A., Mohommad, A., Ahmed, Z. and Sahoo, L. (2010) Efficient *in vitro* plant regeneration from shoot apices and gene transfer by particle bombardment in *Jatropha curcas. Biologia plantarum* 54: 13-20.

Rahuman, A.A., Gopalakrishnan, G., Venkatesan, P. and Geetha, K. (2008) Larvicidal activity of some Euphorbiaceae plant extracts against *Aedes aegypti* and *Culex quinquefasciatus* (Diptera: Culicidae). *Parasitology Research* 102: 867-873.

Raj, S.K., Kumar, S., Snehi, S.K. and Pathre, U. (2008) First report of Cucumber mosaic virus on *Jatropha curcas* in India. *Plant Disease* 92: 171-171.

Ram, S.G., Parthiban, K.T., Kumar, R.S., Thiruvengadam, V. and Paramathma, M (2008) Genetic diversity among *Jatropha* species as revealed by RAPD Markers. *Genetic Resources and Crop Evolution* 55: 803-809.

Ranade, S.A., Srivastava, A.P., Rana, T.S., Srivastava, J. and Tuli, R. (2008) Easy assessment of diversity in *Jatropha curcas* L. plants using two single-primer amplification reaction (SPAR). *Biomass and Bioenergy* 32: 533-540.

Rao, G.L.N, Pr asad, B.D, Sampath, S. and Rajagopal, K. (2007) Combustion analysis of diesel engine fueled with *Jatropha* oil methyl-ester diesel blends. *International Journal for Green Energy* 4: 645-658.

Ravi, M., Marimuthu M.P.A. and Siddiqi, I. (2008) Gamete for mation without meiosis in Arabidopsis, *Nature* 451: 1121-1124.

Ravindranath, N., Ravireddy, M., Ramesh, C., Ramu, R., Prabhakar, A., Jagdeesh, B. and Das, B (2004) New lathyrane and podocarpane diterpenoids from *Jatropha* curcas *Chemical and Pharmaceutical Bulletin* 52: 608—611.

Reddy, J.N. and Ramesh, A. (2006) Parametric studies for improving the performance of a *Jatropha* oil-fuelled compression ignition engine. *Renewable Energy* 31: 1994–2016.

Reddy, M.P., Chikara, J., Patolia, J.S. and Ghosh, A. 'Genetic improvement of *Jatropha curcas* adaptability and oil yield', FACT SEMINAR on *Jatropha curcas* L.: agronomy and genetics, Wageningen, The Netherland, 26-28 March 2007, FACT Foundation. Availablae at: http://www.jatropha-alliance.org/fileadmin/documents/knowledgepool/ReddyChikara _Genetic_Improvement.pdf (accessed 25/10/10)

Reksowardojo, I.K., Dung, N.N., Tuyen, T.Q., Sopheak, R., Brodjonegoro, T.P., Soerwidjaja, T.H., Ogawa, I.H. and Arismunandar, W. (2006) The comparison of effect of biodiesel fuel from palm oil and physic nut oil (*Jatropha curcas*) on an direct injection (DI) diesel engine. In: Proceedinds of FISITA 2006 conference in Yokohama, Japan. Available at: http://www.fisita.com/publications/papers?id=4944 (accessed 25/10/10)

Roy, S. (1998) Biofuel production from *Jatropha curcas*. 10th European Conference and Technology Exhibition on Biomass for Energy and Industry, Jun 08-11, 1998 WURZBURG GERMANY. *Biomass for energy and industry*. 613-615.

Sarin, R., Sharma, M., Sinharay, S. and Malhotra, R.K. (2007) *Jatropha*-Palm biodiesel blends: An optimum mix for Asia. *Fuel* 86: 1365-1371.

Shanker, C. and Dhyani, S.K. (2006) Insect pests of *Jatropha curcas* L. and the potential for their management. *Current Science* 91: 162-163.

Sharma, N. 'Effect on germination on raised bed and sunken bed nursery in different provenances of *Jatropha curcas* L.', FACT SEMINAR on *Jatropha curcas* L.: agronomy and genetics, Wageningen, The Netherland, 26-28 March 2007, FACT Foundation; p. article number 26.

Sharma, S. and Kumar, K. (2009) Root rot of *Jatropha curcas* incited by rhizoctonia bataticola in India. *Indian Forester* 135: 433-434.

Sharma, Y.C and B. Singh. (2009) Development of biodiesel: Current scenario. *Renewable and Sustainable Energy Reviews* 13:1646-1651.

Shen, J.-l., Jia, X.-n., Ni, H.-q., Sun, P.-g., Niu, S.-h. and Chen, X.-y. (2010) AFLP analysis of genetic diversity of *Jatropha curcas* grown in Hainan, China. *Trees-Structure and Function* 24: 455-462.

Shuit SH, Lee KT, Kamaruddin AH and Yusup S. (2010) Reactive extraction and *in situ* esterification of *Jatropha curcas* L. seeds for the production of biodiesel. *Fuel* 89: 527-530.

Singh, R.N., Vyas, D.K., Srivastava, N.S.L and Narra, M. (2008) SPRERI experience on holistic approach to utilize all parts of *Jatropha curcas* fruit for energy. *Renewable Energy* 33: 1868-1873.

Sirisomboon, P., Kitchaiya, P., Pholpho, T. and Mahuttanyavanitch, W. (2007) Physical and mechanical properties of *Jatropha curcas* L. fruits, nuts and kernels. *Biosystems Engineering* 97: 201-207.

Solomon Raju, A.J. and Ezradanam, V. (2002) Pollination ecology and fruiting behaviour in a monoecious species, *Jatropha curcas* L. (*Euphorbiaceae*). *Current Science* 83: 1395-1398.

Soontornchainaksaeng, P. and J enjittikul, T. (2003) Karyology of *Jatropha* (*Euphorbiaceae*) in Thailand. *Thai Forest Bulletin* (Bot) 31: 105-112.

Sorensen, B. (2008) Pathways to climate stabilisation. *Energy Policy* 36: 3505-3509.

Sricharoenchaikul, V., Pechyen, C., Aht-ong, D. and Atong, D. (2008) Preparation and characterization of activated carbon from the pyrolysis of physic nut (*Jatropha curcas* L.) waste. *Energy & fuels* 22: 31-37.

Srivastava, S.K., Tewari, J.P. and Shukla, D.S. (2008) A folk dye from leaves and stem of *Jatropha curcas* L. used by Tharu tribes of Devipatan division. *Indian Journal of Traditional Knowledge* 7: 77-78.

Sseruwagi P, Mar uthi M.N., Colvin J., Rey M.E.C., Brown J. K. and Legg J.P. (2006) Colonization of non-cassava plant species by cassava whiteflies (*Bemisia tabaci*) in Uganda. *Entomologia Experimentalis et Applicata* 119: 145-153.

Sudheer Pamidiamarri, D.V.N., Pandya, N., Reddy, M.P. and Radhakrishnan, T. (2008) Comparative study of inter-specific genetic divergence and phylogenic analysis of genus *Jatropha* by RAPD and AFLP: Genetic divergence and phylogenic analysis of genus *Jatropha Molecular Biology Reporter* 36: 901-907.

Sujatha, M., Makkar, H.P.S and Becker, K. (2005) Shoot bud proliferation from axillary nodes and leaf sections of non-toxic *Jatropha curcas* L. *Plant Growth Regulation* 47: 83-90.

Sujatha, M., Reddy, T.P. and Mahasi, M.J. (2008) Role of biotechnological interventions in the improvement of castor (*Ricinus communis* L.) and *Jatropha curcas* L. *Biotechnology Advances* 26: 424-435.

Sukarin, W., Yamada, Y. and Sakaguchi, S. (1987) Character istics of physic nut, *Jatropha curcas* L. as a new biomass crop in the Tropics. *Japanese Agricultural Research Quarterly, Japan* 20: 302-303.

Sun, S. and Hansen, J.E. (2003) Climate simulations for 1951-2050 with a coupled atmosphere-ocean model. *Journal of Climate* 16: 2807-2826.

Sunil, N., Varaprasad, K.S., Sivar aj, N., Kumar, T.S., Abraham, B. and Prasad, R.B.N (2008) Assessing *Jatropha curcas* L. germplasm in-situ-A case study. *Biomass and Bioenergy* 32: 198-202.

Takeda, Y. (1982) Development study on *Jatropha curcas* (sabu dum) oil as a substitute for diesel engine oil in Thailand. *Journal of Agriculrural Association, China*. 120: 1-8.

Tang, M., Sun, J., Liu, Y., Chen, F. and Shen, S. (2007) Isolation and functional characterization of the *JcERF* gene, a putative AP2/EREBP domain-containing transcription factor, in the woody oil plant *Jatropha curcas*. *Plant Molcular Biology* 63: 419-428.

Tiwari, A.K., Kumar, A. and Raheman, H. (2007) Biodiesel production from *Jatopha* (*Jatropha curcas*) with high free fatty acids: An optimized process. *Biomass and Bioenergy* 31: 569-575.

Tong, L., Shu-Ming, P., Wu-Yuan, D., Dan-Wei, M., Ying, X., Meng, X. and Fang, C. (2006) Characterization of a new stearoyl-acyl carrier protein desaturase gene from *Jatropha curcas*. *Biotechnology Letters* 28: 657–662.

Torrance, S.J., Wiedhopf, R.M. and Cole, J.R. (1977) Antitumor agents from *Jatropha* macrorhiza (Euphorbiaceae) III: Acetylaleuritolic acid. *Journal of Pharmaceutical Sciences* 66: 1348-1349.

Torrance, S.J., Wieldhopf, R.M., Cole, J.R., Arora, S.K., Bates, R.B., Beavers, W.A. and Cutler, R.S. (1976) Antitumor agents from *Jatropha* macrorhiza (Euphorbiaceae). II. Isolation and Characterization of *Jatropha* trione. *Journal of Organic Chemistry* 41: 1855-1857.

URL 1: http://www.abelard.org/news/archive-oil1-2.htm#oi151202 (accessed 25/10/10)

URL 2: http://www.abelard.org/briefings/energy-economics.asp (accessed 25/10/10)

Vyas, D.K. and Singh, R.N. (2007) Feasibility study of *Jatropha* seed husk as an open core gasifier feedstock. *Renewable energy* 32: 512-517.

Wood, D. (2008) Target properties for biofuels in Thailand. Proceedings International Technical Workshop on the Feasibility of Non-edible Oil Seed Crops for Biofuel Production May 25-27, 2007. Chaing Rai, Thailand. Eds. Syers, K.J., Wood, D. and Thongbai, P. pp. 50-61.

Wu, P.-Z., Li, J., Wei, Q., Zeng, L., Chen, Y.-P., Li, M.-R., Jiang, H.-W. and Wu, G.-J. (2009) Cloning and functional characterization of an acyl-acyl carrier protein thioesterase (JcFATB1) from *Jatropha curcas*. *Tree Physiology* 29:1299-1305.

Ye, J., Qu, J., Bui, H.T.N. and Chua, N.-H. (2009) Rapid analysis of *Jatropha curcas* gene functions by virus-induced gene silencing *Plant Biotechnology Journal* 7: 964-976.

Yang, J., Ming-Feng, Y., Dan, W., Fan, C. and Shi-Hua, S. (2010) JcDof1, a Dof transcription factor gene, is associated with the light-mediated circadian clock in *Jatropha curcas*. *Physiologia plantarum* 139: 324-334.

Young, A.L. (2009). Finding the balance between food and biofuels. *Environmental Science and Pollution Research* 16:117-119.

Zhang, F.L., Niu, B., Wang, W.C., Chen, F., Wang, S.H., Xu, Y., Jiang, L.D., Gao, S., Wu, J., Tang, L. and Jia, Y.R. (2008) A novel betaine aldehyde dehydrogenase gene from *Jatropha curcas*, encoding an enzyme implicated in adaptation to environmental stress. *Plant Science* 174: 510-518.

Zhang, Y., Wang, Y.X., Jiang, L.D., Xu, Y., Wang, Y.C., Lu, D.H. and Chen, F. (2007) Aquaporin *JcPIP2* is involved in drought responses in *Jatropha curcas*. *Acta Biochimicaet Biophysica Sinica* 39: 787-794.

Chapter 2: Narrow genetic and apparent phenetic diversity in *Jatropha curcas*: initial success with generating low phorbol ester interspecific hybrids

2.1 Abstract

Due to the increasing popularity of *Jatropha curcas* as a feedstock for biodiesel, generating, nontoxic and high yielding varieties of the plant requires genotypic characterization towards identifying breeding lines. There is little information on the phylogenetic relationships between its global accessions and species. Assessing genetic variation by RAPD, AFLP and combinatorial tubulin based polymorphism (cTBP) in 38 *J. curcas* accessions from 13 countries on 3 continents revealed narrow genetic diversity while the 6 *Jatropha* species from India exhibited pronounced genetic diversity indicating higher possibilities of improving *J. curcas* by interspecific breeding. The relatively unexplored cTBP approach was an efficient molecular tool. Presence of the cocarcinogenic phorbol esters (PE) in the seeds, seed-cake and biodiesel is undesirable. We report initial success in obtaining interspecific F1 and back cross (BC1) plants with low PE and improved agronomic traits. Further BC populations will lead to varieties with targeted traits. Despite the limited genetic diversity within *J. curcas* accessions, appreciable variability exists in important phenotypic, physiological and biochemical traits such as seed size, water use efficiency and seed oil content respectively. This implicates fundamental epigenetic regulatory mechanisms and posits *J. curcas* as a unique system to study them.

Key words: *Jatropha curcas*, biodiesel, phorbol esters, genotyping, breeding, phenotypic plasticity, epigenetics

2.2 Introduction

The *J. curcas* is becoming an increasingly popular plant (Fearless, 2007; Nature editorial 2007) for its proposed value in the biodiesel, biopharmaceuticals, cosmetics and biopesticides industry (Gubitz *et al*., 1999; Kumar & Sharma, 2008). Meagre genotypic characterization, sub-optimal agronomic practices, limited information on the genome of *J. curcas* and a scarce number of its isolated genes require major research initiatives in agronomy, breeding and molecular biotechnology of *J. curcas* for it to live up to its potential (Popluechai et al., 2008). Gressel (2008) recommended the transgenic approach for improvement of biofuel crops including *J. curcas*. *Agrobacterium* mediated transformation of *J. curcas* was recently reported (Li *et al*., 2008), allowing transgenic *J. curcas* to be generated for the desired traits. The 'omics' based technologies can allow fast-track genome sequencing and relevant gene identification. The major constraint in the widespread acceptance of *Jatropha* is the presence in seed/oil, of toxins and antinutrient factors (ANFs) such as curcins, saponins, protease inhibitors, phytate and the cocarcinogenic phorbol esters (PE; Makkar *et al*., 1997). The seeds are inedible and harmful, leading to *J. curcas* seeds also being called black vomit nut and purge nut and its oil being called hell oil or oilinfernale. *J. curcas* cannot be an economically viable feedstock without reduction/elimination of the ANFs and varietal improvement for oil yield (Gressel, 2008; Kohli *et al*., 2009). Elimination of toxins such as curcins was recommended as one of the primary targets (Gressel, 2008) because curcin being similar to ricin has an immediate toxic, often fatal effect (Gressel, 2008). Lin *et al*., (2003) achieved the cloning, expression and characterisation of the *Jatropha* curcin as a ribosome inactivating protein, and Luo *et al*., (2006) described its useful antitumor effects. Makkar & Becker, (1999) proposed a detoxification method for curcin but whether such processing is cost effective or not would depend on the scale of the operation and the prevalent market forces. Importantly however, exposure to *J. curcas* PEs is of equal, if not higher concern due to their co-carcinogenic effect that becomes evident much later after exposure. Currently up to 55% of the original PE content is retained in the biodiesel after the degumming and deodorization processes and even after additional deacidification and bleaching (Haas & Mittelbach, 2000). Seed PEs may additionally limit the commercial value of *J. curcas* by making the otherwise protein-rich seed-cake unsuitable as an animal feed. Recently Devappa & Swamylingappa (2008) reported detoxification of the seed-cake. Another method reducing

seed-cake PE to undetectable levels (sensitivity of the method: 5 ppm) is being considered for patent protection. The performance of carp fish fed the detoxified seed meal at a level of 75% replacement of fishmeal protein is at par with soyabean meal (Dr Harinder P.S. Makkar – personal communication). However, human exposure to PEs up to the stage of generating the seed-cake still remains in both methods. The reduction of PEs through the RNAi approach would require the yet unidentified *J. curcas*-specific PE synthesis genes because the diterpene biosynthesis pathway involved in PE production comprises divergent genes. Additionally, the pathway is similar to that of plant hormones gibberellins and abscisic acid, thus demanding precision in manipulation is required to avoid the interfering with plant hormones production. Incidentally, certain *J. curcas* accessions from specific regions in Mexico are edible and assessed to contain no or a minimal amount of PE that is not harmful to humans (Makkar *et al*., 1998). The use of such accessions, including the ones with variable yield parameters, in conventional breeding could be a beneficial parallel approach. To gain maximum benefit from breeding efforts, identification of breeding lines with target traits in genetically polymorphic background is desirable. As discussed later, existing information on genetic diversity of *J. curcas* is confined to accessions from India and reveals a narrow genetic base. No data exists for comparison of accessions from different parts of the world. We used RAPD, AFLP and/or cTBP to assess genetic polymorphism between 38 *J. curcas* accessions from 13 countries on 3 continents. Six additional *Jatropha* species from India were also assessed. Results indicated high degree of monomorphism in *J. curcas* accessions obtained from around the world. Only the accessions from Mexico and Costa Rica exhibited polymorphism but were still nearly 70% similar to other accessions in our analysis, indicating limited use of intra-specific breeding programs even if global accessions were used. Additional *Jatropha* species exhibited polymorphism indicating *J. gossypifolia* and *J. integerrima* as closer to *J. curcas*. Interspecific breeding of *J. curcas* x *J. integerrima* was successfully achieved earlier (Sujatha & Prabakaran, 2003). *J. integerrima* exhibits certain agronomically undesirable traits such as relatively thin stems and the seeds contain relatively low oil and high levels of PE. However, Sujatha & Prabakaran (2003) reported hybrid plants with intermediate and acceptable phenotype in terms of stem size and seed shattering. In the present study we report additional interspecific crosses and further analysis of back cross (BC) of the hybrids to *J. curcas* that identified plants with improved yield traits and reduced PE. However, the challenge remains to obtain and select advanced BC generation plants, not necessarily

introgressed for, but introduced with heterozygosity for the desirable traits. This may not be difficult to achieve because hybrid plants that flower within 5-6 months instead of 16-24 months were earlier identified and maintained (Sujatha & Prabakaran, 2003). Our results demonstrate the utility of cTBP as a simple and efficient genotyping tool. More importantly, the cTBP results with parental and hybrid plants indicate that increased heterozygosity in F1 plants of interspecific hybrids may indeed explain the wide range, and thus improved plants, for the desirable traits. We obtained marked phenetic variability in morphological, biochemical and physiological traits such as seed size, total oil content and composition and water use efficiency respectively in *J. curcas* accessions with high genetic similarity. The phenomenon of phenetic differences in genetically identical organisms is well known (Lewontin & Goss, 2004). Epigenetic mechanisms for such differences in an undomesticated plant have been explored in the popular model *Arabidopsis* (Blodner *et al*., 2007). However, *J. curcas* as a novel and undomesticated plant system uniquely exhibits naturally widespread genetic monomorphism at a global level. An understanding and the importance of the epigenetic/stochastic events within such a unique genome (allelome) have not been explored. With the construction of gene libraries and its genome sequencing underway (ISAAA, 2008), *J. curcas* may be a useful system to further define the molecular mechanisms underlying pronounced phenotypic variability.

2.3 Materials and Methods

Plant and seed material

The number of accessions of *J. curcas* and additional *Jatropha* species obtained from around the world are listed in Table 1. Plants were either grown in the fields in Hyderabad, India (Latitude 17n20 and longitude 78e30, average temperature 27ºC, average relative humidity 57% and soil type red sandy loam) and Chiang Rai, Thailand (Latitude 19n54 and longitude 99e50, average temperature 26ºC, average relative humidity 76% and soil type sandy/clayey loam) or in controlled environment chambers (CER). In the latter case plants were grown at 28ºC (daytime) and 18ºC (night time) on a 16 h photoperiod of light intensity (PAR) of 300 µmol $m^{-2}s^{-1}$ and a relative humidity of 40%. The *Jatropha* interspecific hybrids and backcrosses were generated as described earlier (Sujatha & Prabakaran, 2003)

DNA extraction and RT-PCR

Jatropha DNA was extracted using modified CTAB method as described earlier (Basha & Sujatha, 2007). For RT-PCR, the Invitrogen RT-PCR kit was used as per the manufacturer's instructions using 1 µg of total RNA with oligo dT and Superscript II reverse transcriptase for first strand cDNA synthesis. Semi-quantitative PCR was carried out using gene specific primers (see table 1b, page 67) on first strand cDNA. Two microlitre of first stand cDNA reaction was used to perform PCR as described above. The following conditions were set up: 94^{o}C – 5min, 94^{o}C - 45sec, 55^{o}C-30 sec, 72^{o}C -1 min for 28 cycles, 72^{o}C for 5 min and store at -20^{o}C for further analysis. The products were run in agarose gel (1.5 % containing 0.5µg/ml of ethidium bromide) and viewed under UV illuminator. Densitometry analysis was performed using BioRad Quantity One Fluor-S Multi-Imager software.

RAPD, AFLP and cTBP analysis

Each RAPD reaction mixture of 20 µl contained 50 ng DNA template, 10 mM Tris-HCl pH 8.3, 50 mM KCl, 0.1% Triton-X100, 0.2 µM 10-mer primer (Operon

Technologies), 100 µM of each of 4 dNTPs, 2 mM MgCl2, and 0.75 unit *Taq* DNA polymerase (Bioline). Amplifications were performed in a Thermal Cycler (Perkin-Elmer) programmed for an initial incubation at 94°C for 3 min, followed by 35 cycles of 30 sec at 94°C, 45 sec at 36°C and 1 min 30 sec at 72°C. The samples were incubated at 72°C for 4 min and held at 4°C prior to analysis. Amplified products were resolved on 0.8 % (w/v) agarose gel electrophoresis. Gels were stained with 0.5 µg/ml ethidium bromide solution and visualized on a UV transilluminator (Bio-Rad). GenRulerTM DNA Ladder (Fermentas) was used as a standard DNA marker. The AFLP procedure was carried out as per the manufacturer's instructions (Invitrogen). Thirty-two primers [E1-E4 combine with M1-M8; Invitrogen) were used for amplification. *Taq* DNA polymerase from Invitrogen was used in this study. PCR products were analyzed as described earlier (Lin *et al*., 1996). The silver staining process included fixing the gel in 10% (v/v) acetic acid for 20 minutes, rinsing in de-ionized water (3x5 minutes), staining for 30 minutes in a solution containing 0.2% (w/v) silver nitrate and 0.015% (v/v) formaldehyde. The stained plate was rinsed with de-ionized water for 10 seconds and developed in a cold (4–10°C) developer solution containing 3% (w/v) sodium carbonate (BDH Analar grade), 0.015% (w/v) formaldehyde, and 0.002% (w/v) sodium thiosulphate until the DNA bands became visible. The gel was rinsed with distilled water and airdried. RAPD or AFLP bands were recorded and cluster analysis performed using the SIMQUAL analysis method. The results were presented as a dendrograms using UPGMA cluster analysis in NTSYSpc-2.01 program (Rohlf, 1997). The cTBP was performed and analyzed as described earlier (Breviario *et al*., 2007)

Total lipid and FAMES and phorbol ester analysis

Up to 25 g of *J. curcas* seed kernels were ground using a homogenizer. Total lipid was obtained from powdered seed kernels by soxhlet extraction using hexane as a solvent. Oil recovery was calculated from ratio of oil weight to total seed weight. To a 100 µl oil sample, 1 ml methanol and 1 ml 30 % KOH were added. The samples were then incubated at 75°C overnight. After cooling to room temperature samples were extracted with 1 ml diethyl ether three times. The combined ether extract was dried under a stream of nitrogen. Dried samples were mixed with 1 ml dichloromethane, 1 ml phase transfer catalyst solution (0.1 M tetrabutylammonium hydrogen sulphate in 0.2 M aqueous NaOH) and 25 µl iodomethane using horizontal shaker at 100 rpm for 30 min and then left at room

temperature until two phases of the mixtures were separated. The lower phase was transferred to a new tube and then dried under a stream of nitrogen at 40°C. Dried samples were dissolved in hexane and stored at -20°C.

Oil properties

Saponification number (SN), iodine value (IV), cetane number (CN), and gross energy (GE) were calculated from fatty acid methyl ester compositions of oil using the following equations 1 – 4 (Krisnangkura, 1986; Freeman & Bagby, 1989; Kalayasiri *et al*., 1996)

$SN = \Sigma (560 \times Ai)/MWi$: (1)

$IV = \Sigma (254 \times D \times Ai)/MWi$: (2)

$CN = 46.3 + 5458/SN - 0{:}225 \times IV$: (3)

$GE = 618000/SN – 0.08 \times IV – 430$: (4)

Ai is the percentage, D is the number of double bonds and MWi is the molecular mass of the respective fatty acid (mainly palmitic, stearic, oleic and linoleic acid). The phorbol ester content in the kernels of *J. curcas* and other species of *Jatropha* was determined using HPLC as described earlier (Makkar *et al*., 1997).

Gas Chromatography

The Fatty Acid Methyl Esters (FAMEs) were separated by capillary gas chromatography (Perkin Elmer) through a 25 mm methyl silicone (OV-1) capillary column and a flame ionization detector. The temperature profile was run from 100°C to 210°C at a rate of 4°C/min, then held at 210°C for 15 minutes, then increased to 240°C at the rate of 20°C/min and held for 11 minutes to complete the program. Retention time and peak area of each FAME was recorded using a TRIO computing integrator (Trivector System). Comparing the retention times with those of a standard FAMEs mixture identified individual FAMEs.

Determination of water use efficiency (WUE)

Jatropha plants were established in the field under natural photoperiod in Hyderabad, India with a planting width of 2 x 2 m. All measurements were taken on nine month old plants. CO_2 assimilation rates and transpiration were determined by using a portable infrared gas analyzer (LCpro+, Analytical Development Company, Hoddesdon, England). The average light intensity (PAR) was 1500 μmol $m^{-2}s^{-1}$ and the CO_2 partial pressure was 37.5 Pa. Air flow through the analyzer was adjusted to maintain cuvette relative humidity near ambient levels of 55-65% during measurements. The rates of CO_2 exchange were determined on five plants at regular intervals. The measurements were taken on clear days during three periods: early summer (May, 2007), midsummer (June) and late summer (August). Measurements were taken on fully expanded mature leaf number three counted from the shoot apex. The condition for controlled environment chambers (CER) was described above in Plant and seed material. Water use efficiency was calculated as Photosynthetic rate (A mmol m-2s-1)/ Transpiration rate (E mmol m-2s-1) (See Table 5).

Statistical analysis

Statistical analysis was performed using SPSS analysis by Dr Clare V. Lanyon. Data of Table 4 in this chapter was analyzed using One-way Multivariate Analysis of Variance (MANOVA). Dendrogram of Figure 8b was analyzed using Cluster analysis

Table 1a Global origin of *Jatropha* accessions and species used in this chapter

Continent	**Country**	***J. curcas* accessions**	***J. spp* accessions**
Asia	India	10	6*
	Thailand	7	1§
Africa	Egypt	1	
	Mali	1	
	Nigeria	2	
	Tanzania	4	
Americas	Cape Verde	1	
	Costa Rica	4	
	Ecuador	1	
	Mexico	1	
	Paraguay	1	
	Surinam	4	
	Uganda	1	
Total		**38+**	**7 = 45**

**J. glandulifera, J. gossypifolia, J. integerrima, J. multifida and J. podagrica, J. maheshwarii, §J. Podagrica*

Table 1b List of primers used for gene expression expereiment in this chapter

Primer	Primer sequence	Gene target	Tm (^{o}C)
Jc-acc-F	5’ GAGAGAGTCATGCCCGTGTC 3’	acetyl-CoA carboxylase (acc; DQ987702)	65
Jc-acc-R	5’ CCATAAGCCGCAATATCAGC 3’		64
Jc-sad-F	5’-TGGAAACACTGCCAGACTTG-3’	steroyl-ACP desaturase (sad; DQ084491)	57
Jc-sad-R	5’-TACAGATGCCCCCTATCAGC-3’		60
Jc-fad-F	5’-CTCCTTCCTCAACCCCTCTC-3’	delta-12 fatty acid desaturase (fad; DQ157776)	61
Jc-fad-R	5’-CCTCTCTCCACATTGCCTTG-3’		60
Jc-dgat-F	5’ GTGATGGTGTTGCTGAGTCG 3’	diacylglyceol acyltransferase (dgat; DQ278448)	64
Jc-dgat-R	5’ TTAATATCCGACCTCCCACG 3’		63
Jc-18S-F	5’ AAACGGCTACCACATCCAAG 3’	18S ribosomal RNA	63
Jc-18S-R	5’ TCATTACTCCGATCCCGAAG 3’		63

2.4 Results

Genetic diversity in Jatropha

Four accessions (JI, JI-2, JN and JT) from India, Nigeria and Thailand respectively, were supplemented with 13 additional *J. curcas* accessions from six provenances of Thailand and one accession of *J. podagrica,* as an outgroup*,* also from Thailand. These 18 samples were initially examined using 10 RAPD primers. The resultant profiles subjected to UPGMA-mediated cluster analysis revealed two major clusters (Figure 1). One cluster contained all of the 17 *J. curcas* accessions and the second contained the out-group *J. podagrica*, which showed an overall similarity of 52% with *J. curcas*. Among the *J. curcas* accessions the similarity coefficient was high (0.78) indicating a narrow genetic base. The two Indian accessions clustered separately while the Nigerian accession clustered with the remaining 14 Thai accessions.

The 6 provenances of Thai accessions could not be clearly differentiated, reinforcing the narrow genetic base between provenances. The *J. curcas* accession from Mexico known to bear seeds containing minimal amounts of antinutrient factors (ANFs) and phorbol esters (PEs) was expected to be genetically diverse. This non-toxic Mexican accession (NTMA) was then investigated together with 6 Thai accessions from 4 provenances using AFLP with 32 primer pairs. Results showed the former to be genetically distant from the Thai accessions (Figure 2), yet 76% similar, whereas the Thai accessions again turned out to be highly similar between themselves (90%). These results indicated the importance of a) testing accessions from wider eco-geographic regions; b) assessing other *Jatropha* species in addition to *J. curcas,* and c) assessing natural and artificial interspecific hybrids for genetic variability. We thus investigated the following material: a set of *J. curcas* accessions obtained from various parts of the world, including NTMA; different species of *Jatropha* obtained from India; a natural hybrid occurring in India (Prabakaran & Sujatha, 1999); synthetic hybrids and a somaclonal variant generated earlier in India (Sujatha & Prabakaran, 2003) and additional hybrids generated for this study between *J. curcas* and *J. maheshwarii*/*J. multifida* (Please refer to Table 1 for details on different accessions and species used). A novel, relatively unexploited technique defined of cTBP [combinatorial Tubulin Based Polymorphism; Breviario *et al*., 2007] was used. The method uses variation in the length of the first and second intron of members of the plant β-tubulin

gene family. The approach was successfully used earlier to detect intra and inter-species polymorphism in diverse plants including oilseed plants-rapeseed and peanut (Breviario *et al*., 2007) and palm (Breviario *et al.*, 2008). Two sets of cTBP were performed on *J. curcas* accessions from different parts of the world. In the first set cTBP analysis on 10 *J. curcas* accessions from 9 different countries in Asia, Africa and the Americas revealed no polymorphism for either intron (Figure 3 and appendix 1). Although this was surprising in that nearly 10 different plant genera and their respective species and accessions put through cTBP analysis never exhibited monomorphism earlier (Dr Deigo Breviario - unpublished results), it was not surprising in view of the RAPD and AFLP results obtained with *J. curcas* earlier. In the second set 16 accessions from 6 countries were analyzed. Whereas some accessions in the second set were identical to the first set, those from Tanzania, Surinam and Costa Rica were new. Results showed that the four accessions from Costa Rica were clearly different to those from other parts of the world and they also exhibited intra-specific polymorphism in both intron I and II in particular accession number 10 (Figure 4; see *). The ability of cTBP to differentiate between accessions was then used in the following cTBP analysis to differentiate between species. The NTMA which revealed polymorphism by AFLP analysis earlier was tested together with some other *J. curcas* accessions; different *Jatropha* species; natural and synthetic hybrids of *Jatropha* and a somaclonal variant of *J. integerrima*. These investigations showed that the NTMA clustered separately from the other *J. curcas* accessions and indeed the different species of *Jatropha* exhibited marked genetic variability (Figure 5). These results demonstrate the value of cTBP analysis in fingerprinting the *Jatropha* genus. UPGMA analysis of the cTBP bands gave genetic similarity coefficients over the range of 0.07 to 1.0 (Figure 5), with three of the four *J. curcas* accessions typically showing a high similarity amongst them and with NTMA. Other species were highly dissimilar. As expected the 2 accessions of *J. integerrima* that differed in flower colour and the somaclonal variant were markedly similar to each other but also to *J. rosea*. The cTBP analysis was again used to assess the genetic variability in the F1 generation compared to the parental species for hybrids of *J. curcas* with *J. integerrima, J. maheshwarii* or *J. multifida* (Figure 6). Clear polymorphism was indicated between the parental types and the resulting F1 progeny plants with facile identification of parental bands in the hybrids. The 1:1 inherited band pattern of the F1 hybrid with respect to parental contribution indicates the homozygous status of the parental lines that may not be restricted just to the alleles of the beta-tubulin gene family. Our

results reiterate the value of cTBP as a genotyping tool for plant accessions, species and hybrids. The method is simple and rapid and can be applied to any plant genome of unknown sequence.

Seed phorbol ester content

Table 2a shows that *J. curcas* accessions and species have a wide range of PE content with the NTMA showing undetectable to very low amounts of PE in the seed. The material for a potentially successful conventional breeding approach is thus available in nature. An earlier attempt at trying to obtain low PE plants after crossing the NTMA with standard *J. curcas* was not successful (Sujatha *et al*., 2005), most likely due to a limited genetic variability. To circumvent this limitation interspecific hybrids generated earlier (Sujatha & Prabakaran, 2003) and for this study were analysed. We compared the F1 population of interspecific reciprocal hybrids between *J. curcas* as one parent and *J. integerrima, J. maheshwarii or J. multifida* as the other parent for yield and desirable agronomic traits including the reduction in PE content in glasshouse experiments. We obtained one line that exhibited comparatively better agronomic traits and low PE. This line was backcrossed to *J. curcas* NTMA accession in order to further reduce the PE content. Indeed, we obtained two BC lines (ID 222 and 223) with PE levels comparable to those in NTMA (Table 2b). Along with reduced PE these lines exhibited quantitatively and qualitatively improved seed as seen through the number of fruits per inflorescence (23-25), seed weight, oil content, oleic acid to linoleic acid (O/L) ratio and PE level (Table 3). This indicated the potential for breeding elite *Jatropha* cultivars. These data suggest that the chances of improvements in commercially desirable traits of *J. curcas* are higher through interspecific breeding rather than the intraspecific breeding approach, which is limited by the low genetic variability in globally distributed accessions of *J. curcas*.

Phenotypic variability: seed size, oil content and composition and water use efficiency

High genetic similarity in *J. curcas* seeds of accessions collected from different countries and provenances suggested that seed phenotype parameters such as seed size and percent kernel weight might not be significantly different. However, our data confirmed high variability in seed phenotypic characteristics (Table 4a). For example, standard

deviation (SD) and sample variance (SV) in mean seed weight (n=10) of seeds from 7 different eco-geographic regions in Table 1 was 0.1019 and 0.0102 respectively. However, SD and SV within the 10 seeds of each region varied from 0.017 to 0.022 and 0.00061 to 0.00073 respectively. This indicated higher variability between the regions than within the regions, in turn indicating that despite high genetic similarity there was high phenotypic variability in seed weight, most likely due to differences in the respective environmental and eco-geographical conditions. Boschin *et al*. (2008) recently demonstrated that in oilseed lupin when a total of six cultivars from Spain, France and Italy were assessed in three different climates, the variation in total oil content and fatty acid composition was more dependent on the genotype than on genotype environment (GE) interaction This suggested limited variability in these traits among the *J. curcas* accessions exhibiting high genetic similarity. However, the variation in total oil content of *J. curcas* seeds from different regions (Table 4b) exhibited a high SD and SV of 7.5 and 56.6 respectively suggesting environment-mediated changes in oil content. In parallel, since the biodiesel quality depends on high O/L ratio, we analysed the O/L ratio of different accessions coming from different eco-geographic regions The O/L ratio (Table 4b) also exhibited high SD and SV (0.25 and 0.06 respectively) compared to a low SD and SV (0.094 and 0.008 respectively) within 10 different samples from the same field in the case of accession JI, once again implicating environmental variables for the noticed differences. Saleem *et al*., (2008) demonstrated that in sunflower, oil composition was affected by environmental variables of temperature and humidity arising due to different sowing times.

The Δ-12 fatty acid desaturase (FAD; DQ157776) is responsible for conversion of oleic acid to linoleic acid contributing to the final O/L ratio in the seed oil. We used semi-quantitative RT-PCR for the *fad* gene on a small set of three *J. curcas* accessions (JI, JN and JT; see table 1b for primers used), to see if a relationship existed between the O/L ratio and the *fad* transcript level. As evident from Table 4b, JT was chosen due to its lowest O/L ratio and JI and JN were chosen for being at the higher end of the O/L ratio values amongst the ten accessions. The RT-PCR analysis showed 1.6 times more *fad* transcript in JT than in JI or JN (Figure 7), suggesting a higher amount of FAD function leading to increased conversion of oleic acid to linoleic acid and in turn leading to the low O/L ratio observed. A higher amount of *fad* transcript accumulation in JT seeds is more likely due to epigenetic rather than genetic reasons as discussed below.

Water use efficiency (WUE) in a drought tolerant plant such as *J. curcas* is an important physiological trait. We assessed four different Indian accessions of *J. curcas* for WUE under identical environmental conditions in the field. One of these; TCR193 clearly exhibited a relatively high WUE (Table 5). However, Table 5 also shows that differences in WUE between the four accessions were mainly because of differences in CO2 fixation efficiency rather than to differences in transpiration rate, which were largely similar. Such relatively large differences in the photosynthetic rate under identical environmental conditions implicate strong genetic determinants. Since these accessions were not compared for genetic variability, it was not clear if differences in the photosynthetic rate had a genetic base. Therefore, we compared the WUE of JI – one of the accessions used in the field study – to two other accessions JN and JT, in a controlled environment chamber. JI, JN and JT were genetically more than 80% similar (Figures 1 and 5). The dissimilarities are not necessarily expected in essential (hence largely conserved) genes such as the ones involved in photosynthesis. Hence, photosynthetic rate, transpiration rate and the WUE were expected to be largely similar. However, once again as in the field study, we noticed similar transpiration rates and WUE values but relatively large differences in photosynthetic rates under identical growth conditions. These results implicate epigenetic imprint-mediated expression differences in otherwise similar genomes. Our results for variability in seed phenotype and FA content and composition suggest epigenetic mediated adaptation and those with WUE suggest such adaptations being passed to the next generation. Recent studies do indicate that ecologically adaptive epigenetic variation in natural populations can be independent from genetic variation, and that such environmentally induced epigenetic changes are inherited by future generations in some cases (Bossdorf *et al*., 2007).

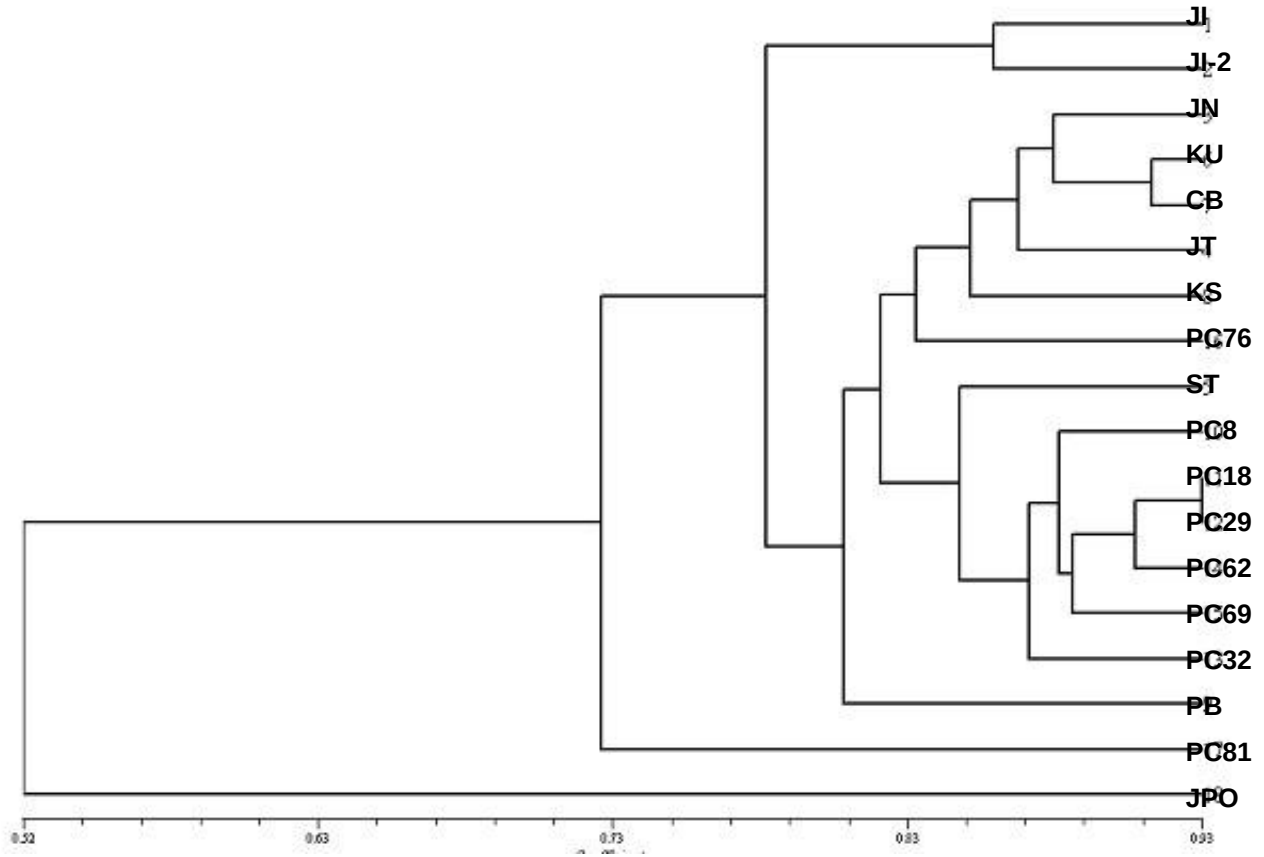


Figure 1 UPGMA analysis of RAPD conducted on 17 *J. curcas* and one *J. podagrica* accession. Origin of accessions identified as follows: JI (1) & JI-2(2) – India; JN(3) – Nigeria; JT(4) – Thailand. All other accessions were from different provenances of Thailand represented here as ST (5); KU – (6); CB – (7); KS – (8); PB – (9); PC8 (10), PC18 (11), PC29 (12), PC32 (13), PC62 (14), PC69 (15), PC76 (16), PC81 (17) and JPO (18) – Thailand *J. podagrica.*

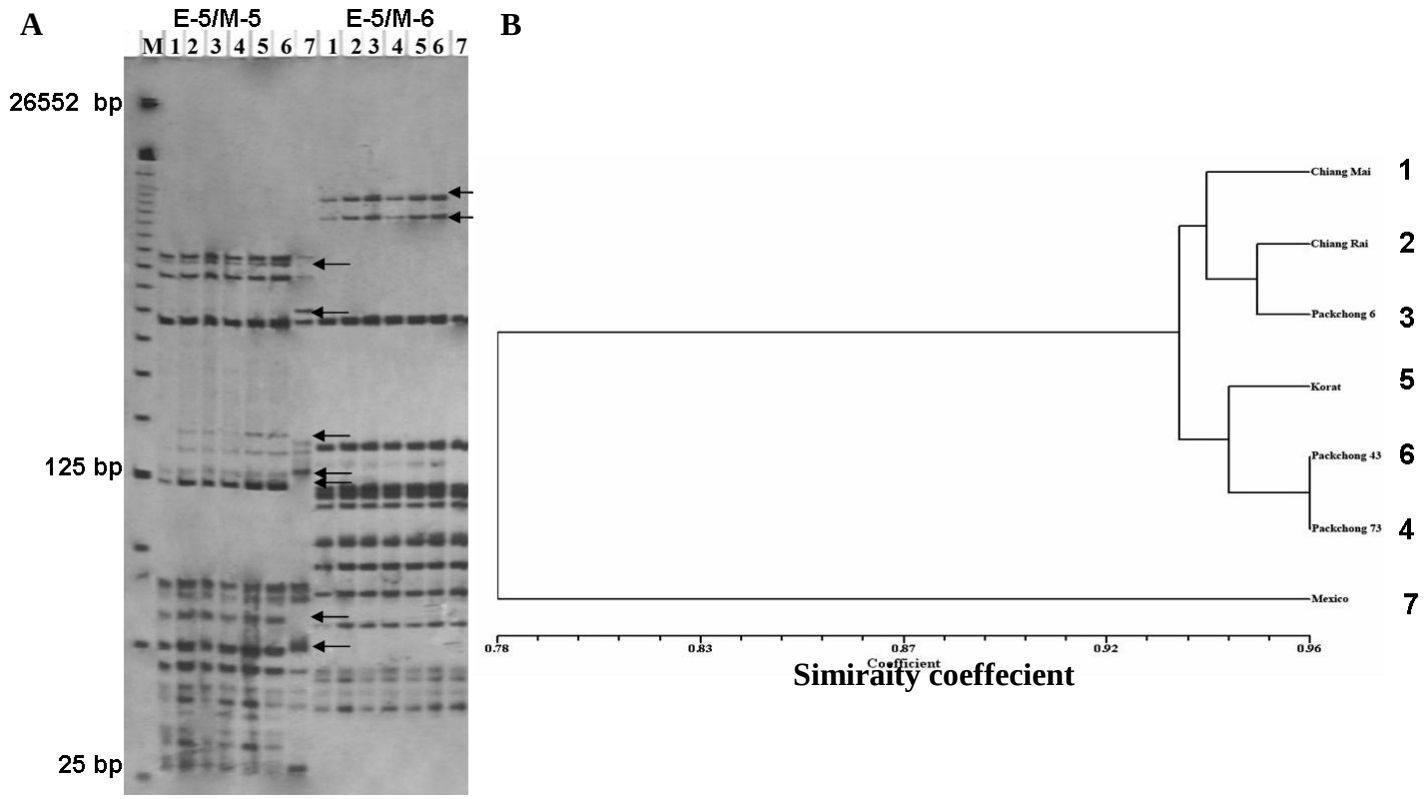
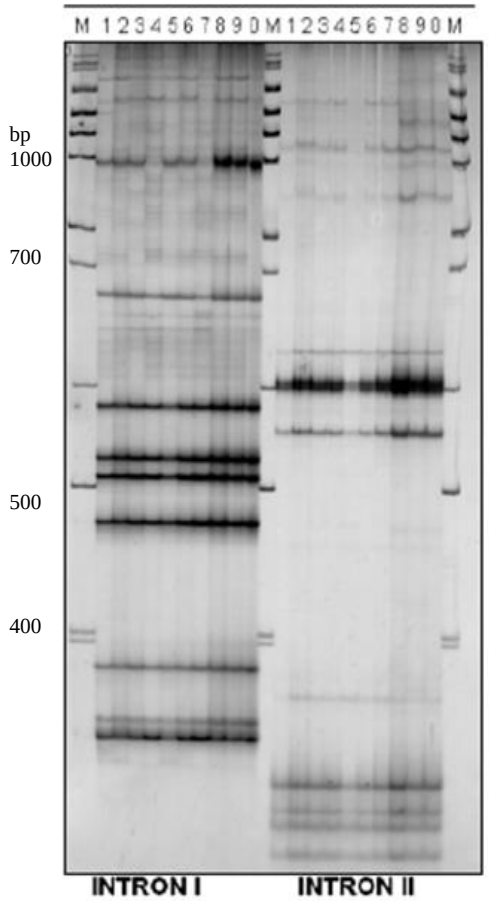


Figure 2A Representative gel showing DNA amplicons obtained using two separate AFLP primer sets (I and II) on six *J. curcas* accessions (lane 1 to 6) from Thailand and one (lane 7) from Mexico (Non-toxic; NTMA). M: DNA size marker. Lack of polymorphism in the Thai accessions was representative of similar results obtained with additional primer sets and other *J. curcas* accessions from India andNigeria. The polymorphic bands (present: blue circles; absent: yellow circles) in NTMA indicate its genotypic diversity in comparison to the Thai (and Indian and Nigerian) accessions analysed.

Figure 2B UPGMA analysis of the AFLP results with the 32 primer pairs. Numbers represent the following accessions: 1:Chaing Mai 1, 2: Chiang Rai 1, 3: Packchong 6, 4: Packchong 73, 5: Korat 1, 6: Packchong 43, 7: NTMA. The analysis revealed high similarity between the Thai accessions.

Figure 3 cTBP-mediated detection of monomorphism in *J. curcas*. Combinatorial Tubulin Based Polymorphism (cTBP) for β-tubulin gene family introns. Representative gel for cTBP on 10 accessions from 9 countries (1: Cape Verde; 2: Egypt; 3: Ecuador; 4: India (JI); 5: Indonesia; 6: Nigeria (JN); 7: Thailand (JT); 8: Mali; 9: Nigeria; and 10: Paraguay. M: DNA size marker). Gel shows a distinct lack of polymorphism in both intron I and II.



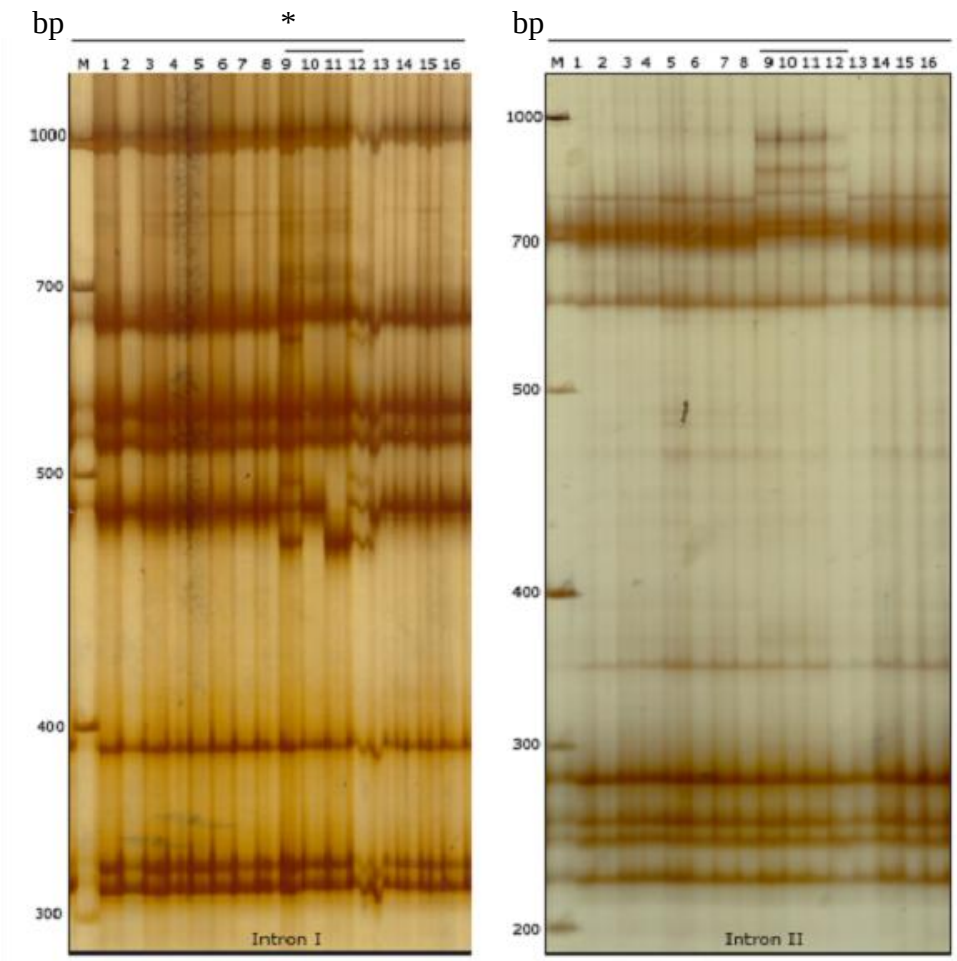


Figure 4 cTBP-mediated genotyping of *J. curcas*. cTBP analysis for intron I (a) and II (b) for 16 *J. curcas* accessions from 6 countries - Tanzania (1 to 4 and 14); Surinam (5 to 8); Costa Rica (9 to 12); India (13); Egypt (15) and Thailand (16). M: DNA size marker. Accessions from Costa Rica were clearly different within and between the countries (Intron I; see *).

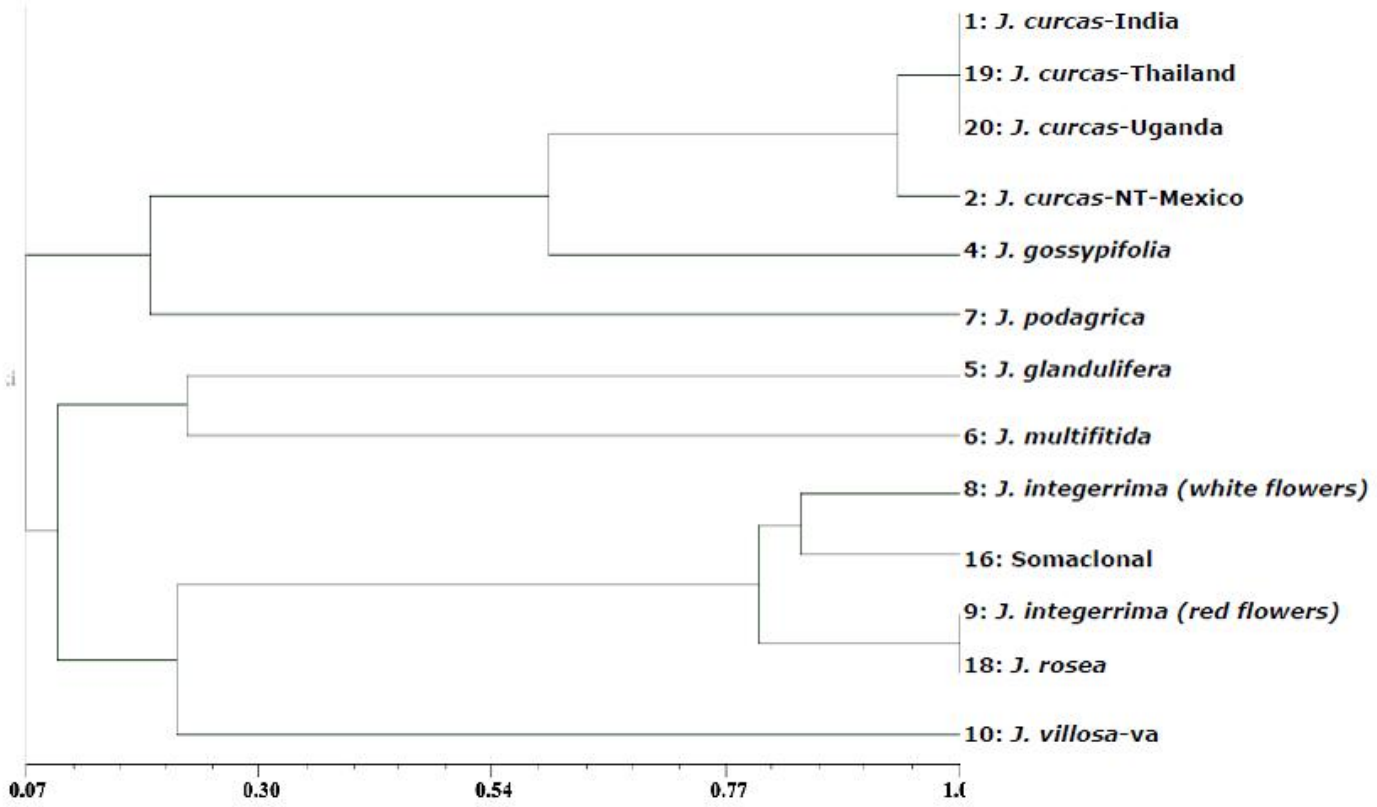
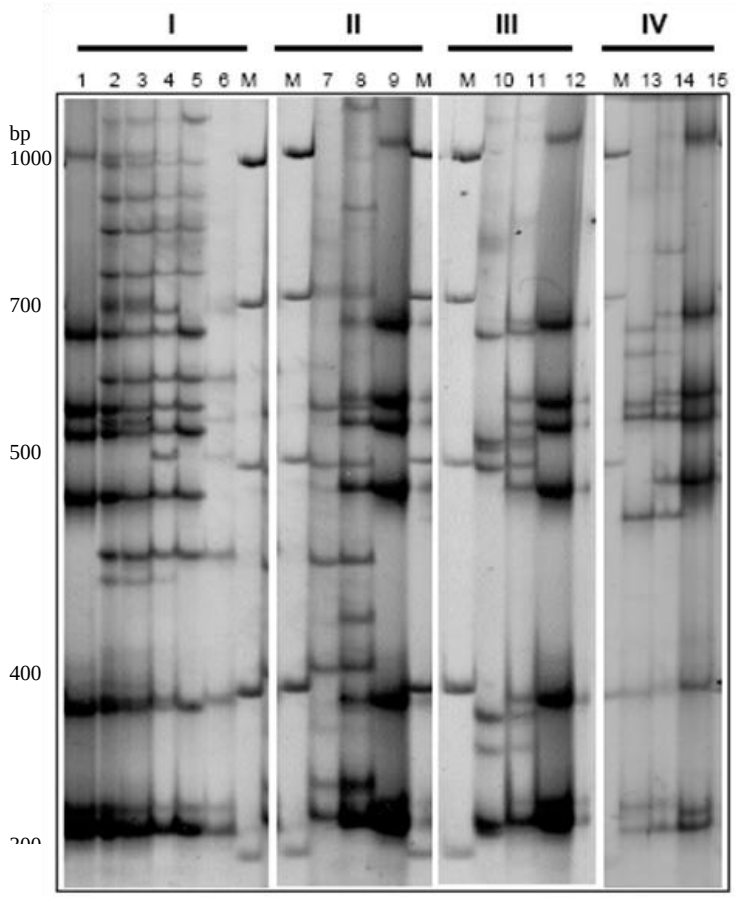


Figure 5 UPGMA of cTBP analysis on different species of *Jatropha*. *Jatropha* species were clearly different from each other and from *J. curcas* accessions which were often closely related.

Figure 6 cTBP-mediated genotyping of the inter-specific hybrids and parents. cTBP analysis showing polymorphism in intron I between the parental and hybrid plants for I: *J. curcas* X *J. integerrima;* lanes 1, 2, 3, 4, 5 and 6: J.c, hybrid 1, hybrid 2, hybrid 3, hybrid 4 and J.i respectively; II: *J. curcas* X *J. gossypifolia = J. tanjorensis;* lanes 7, 8 and 9: J.g, J.t and J.c respectively; III: *J. curcas* X *J. multifida*; lanes 10, 11 and 12: J.mu, hybrid and J.c respectively; IV: *J. curcas* X J. *maheshwarii* lanes 13, 14 and 15: J.ma , hybrid, J.c respectively; M: DNA size marker)



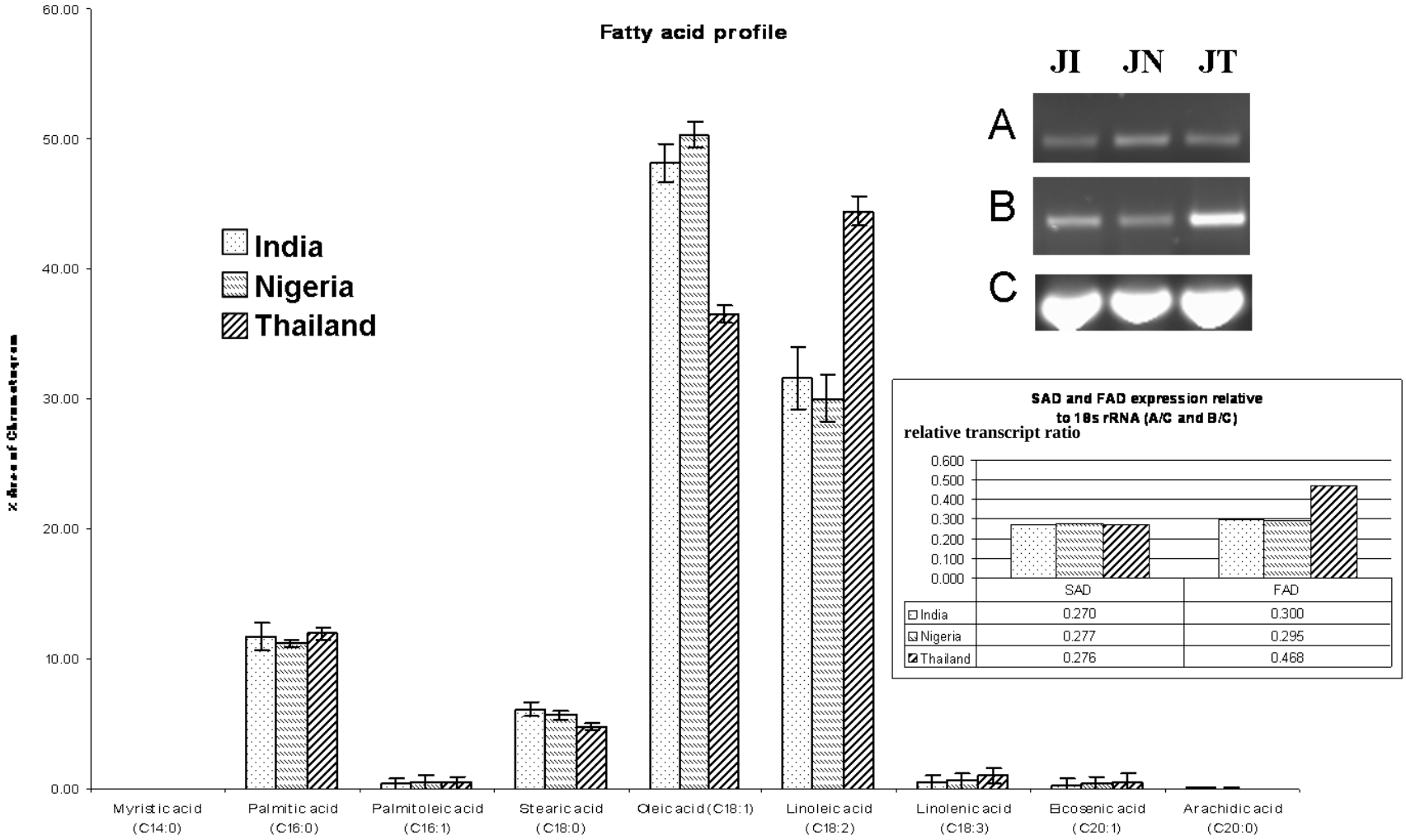
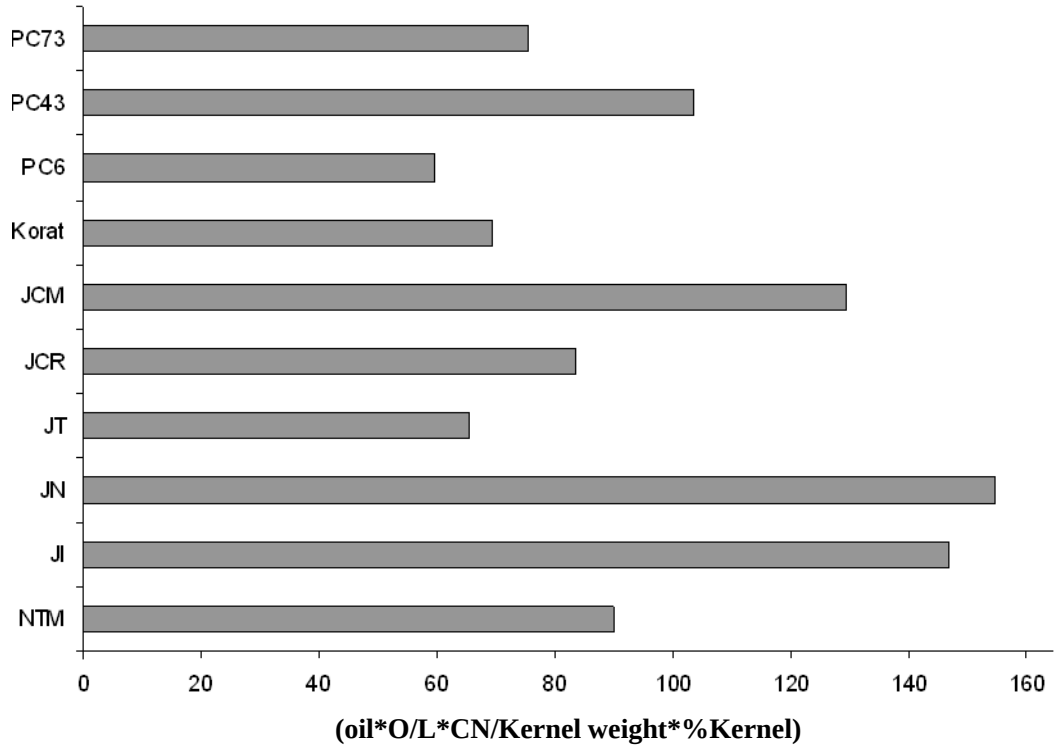


Figure 7 Relationship between FA biosynthesis gene transcripts and FA profile in the seeds of *J. curcas*. Gas chromatographic analysis of seed fatty acid profiles of accessions JI, JN and JT showing higher amounts of linoleic acid than oleic acid in JT, also reflected in the larger amount of the *fad* transcript (which converts oleic to linoleic acid; see inset) in JT. A is transcript level for *sad*, B is transcript level for *fad*. C is transcript level for *18S*.



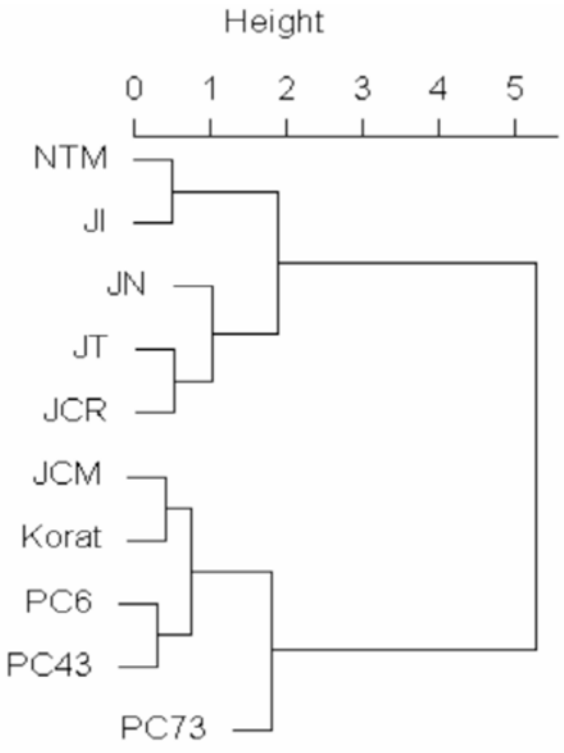


Figure 8a **Figure 8b**

Figure 8 Classification of *J. curcas* accessions based on seed/oil properties listed in Table 4. Measured values for shaded properties in Table 4 were computed to assess the accession with an optimal trait combination (oil*O/L*CN/Kernel weight*%Kernel). 4a shows that accession JN and JI were superior. However, low PE content in JI makes it an even better candidate. 4b shows the relationship between the accessions based on the standardized data set [(x-meanx)/Sdx]. Relatedness reflected the eco-geographical influence on seed properties on a set of accessions highly similar at genotypic level. The 6 provenances of the Thai accessions were clearly separated which was not the case through genotypic clustering (see main text).

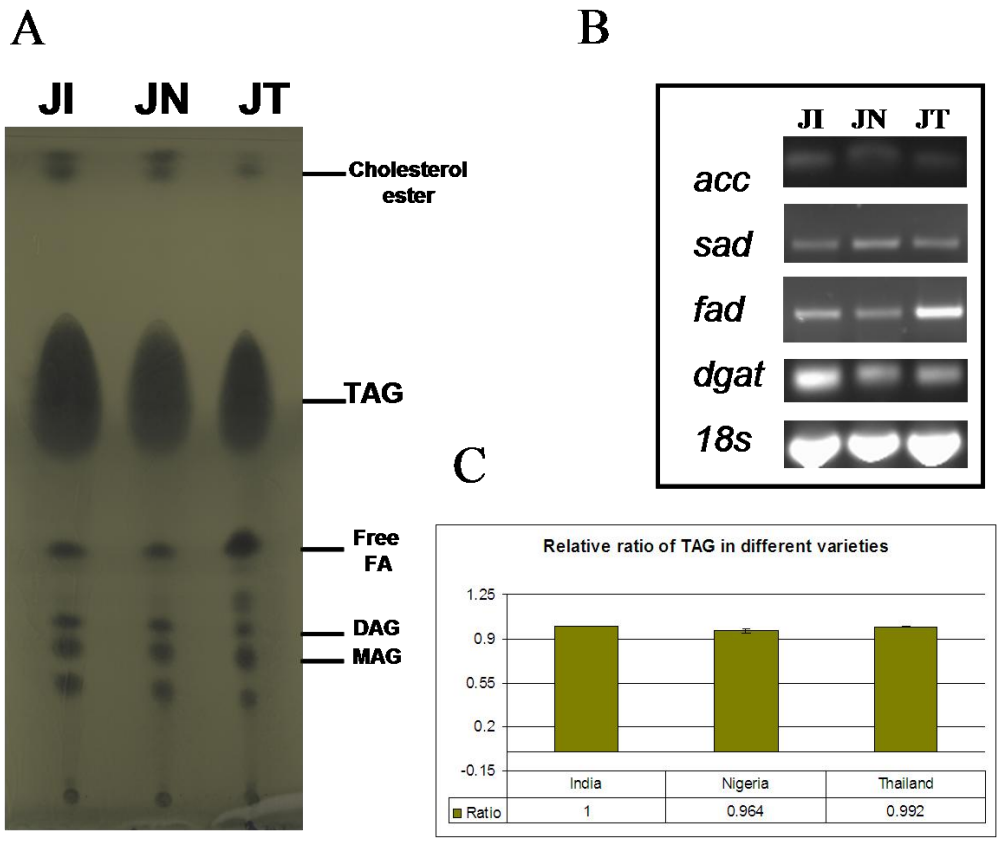


Figure 9 Relationship between FA biosynthesis gene transcripts and TAG accumulation in the seeds of *J. curcas***.** A) Thin layer chromatography-mediated lipid profile visualisation (TAG: Triacyl glycerides; DAG: Diacyl glyceride; MAG: Monoacyl glycerides B) RT-PCR-mediated analysis showing the difference in transcript level for the genes *acc, sad, fad* and *dgat* in the seeds of JI, JN and JT. The obvious difference in *fad* transcript levels and its manifestation in oleic:linoleic acid content in JT is shown in Figure 7 of the manuscript. C) Estimation of the difference between the TAG content in JI, JN and JT. The intensity and the area of TAG spot on the TLC was measured with a densitometer to obtain relative amounts of TAG. These differences were not as pronounced as at the transcript level. The y-axis on the graph (C) is the relative ratio of TAG of each variety compared to India variety.

Table 2a Phorbol Ester content in *Jatropha* kernels

Species	Sample	PE (mg gm-1 kernel)
J. curcas	NTMA*	ND** - 0.11
J. curcas	JI	1.62
J. curcas	IJC-4	2.22
J. curcas	IJC-1	2.62
J. curcas	IJC-2	2.65
J. curcas	IJC-11	3.81
J. curcas	Kerala	3.90
J. curcas	JI-1	4.12
J. curcas	JI-2	4.30
J. curcas	Chintapally	5.04
J. integerrima	IJI	7.92
J. glandulifera	IJGL	2.78
J. podagrica	IJP	4.50
J. multifitida	IJM	9.09

Table 2b Phorbol Ester content in *Jatropha* hybrid kernels

Species	Sample	PE (mg gm-1 kernel)
Hybrid (Jc.Jint)	ID 222	0.08
Hybrid (Jc.Jint)	ID 223	0.12
Hybrid (Jc.Jint)	ID 50	1.71
Hybrid (Jc.Jint)	ID 112	2.20
Hybrid (Jc.Jint)	ID 111	2.52
Hybrid (Jc.Jint)	ID 93	2.80
Hybrid (Jc.Jint)	ID 225	2.82
Hybrid (Jc.Jint)	ID 226	3.31
Hybrid (Jc.Jint)	IH 13	4.10
Hybrid (Jc.Jint)	IH 300	4.43

Hybrid, analysis of seeds on *J. curcas* obtained after pollinationby pollen of F1 flowers of *J. curcas* (f) x *J. integerrima* (m) cross. Each sample represents an independent back cross (BC1) *NTMA: Non-Toxic Mexican accession. **ND = Not Detected

Table 3 Comparative range of critical seed properties for interspecific derivatives and *J. curcas*

Trait	**Interspecific**		***J. curcas***	
	n	Range	n	Range
Seed weight (g)	19	0.23-0.78	5	0.22-0.77
Seed oil content (%)	14	16.39-34.47	4	29.80-33.58
Oleic (%)	22	24.92-52.95	6	36.28-51.18
Linoleic (%)	22	25.66-44.89	6	27.27-60.92
Phorbol esters (mg/g)	12	0.04-9.04	8	0.01-5.25

Table 4a *J. curcas* seed properties relevant to biodiesel

Name	Origin	Seed			Kernel			Shell	Kernel	Shell	Kernel
		Weight (gm)	Diameter (mm)	Length (mm)	Weight (gm)	Diameter (mm)	Length (mm)	Weight (gm)	% weight	% weight	Density gm/cc
NTM	Mexico	0.692	10.00	16.99	0.444	8.56	13.91	0.248	64.10	35.90	0.100
JI	India	0.633	11.33	17.60	0.422	9.19	13.70	0.211	66.66	33.34	0.089
JN	Nigeria	0.653	11.10	17.80	0.435	8.92	14.00	0.217	66.65	33.35	0.093
JT	Thailand	0.682	11.45	17.50	0.454	9.29	14.33	0.227	66.67	33.33	0.089
JCR	Thailand	0.820	11.46	19.23	0.519	9.68	15.24	0.301	63.30	36.70	0.088
JCM	Thailand	0.745	11.58	19.31	0.434	8.84	15.27	0.311	58.28	41.72	0.089
Korat	Thailand	0.843	11.53	18.98	0.529	9.52	14.66	0.313	62.83	37.17	0.096
PC6	Thailand	0.918	11.59	19.19	0.537	9.25	14.55	0.381	58.45	41.55	0.100
PC43	Thailand	0.844	11.30	19.63	0.532	8.92	14.79	0.312	63.00	37.00	0.098
PC73	Thailand	0.877	11.25	19.24	0.511	9.07	15.25	0.365	58.31	41.69	0.100

Table 4b *J. curcas* oil properties relevant to biodiesel

Name	Oil	FAMEs (%)						SN	IV	CN	GE
	% (w/w)	P	S	O	L	M	O/L				kCal/mol
NTM	37.40	14.9	7.60	37.50	30.60	9.40	1.22	177.01	88.65	56.13	3054.04
JI	52.20	11.60	6.60	48.10	32.10	1.60	1.49	189.31	98.08	53.06	2826.72
JN	51.10	11.10	6.50	49.90	30.10	2.40	1.65	189.03	97.82	53.17	2831.56
JT	49.50	11.90	6.10	36.10	44.60	1.30	0.81	191.34	112.15	49.49	2790.88
JCR	37.40	15.90	10.80	40.30	30.30	2.70	1.33	188.54	89.01	55.22	2830.78
JCM	39.40	15.10	10.30	43.60	28.80	2.20	1.51	189.91	89.01	55.01	2817.04
Korat	37.80	15.50	8.20	38.80	35.00	2.50	1.11	189.42	93.33	55.09	2825.45
PC6	30.70	16.10	8.10	38.80	34.90	2.10	1.11	190.64	93.33	54.91	2804.55
PC43	49.60	14.60	8.10	42.00	32.90	2.40	1.27	189.75	93.33	55.04	2819.87
PC73	38.10	19.40	9.50	35.40	32.90	2.80	1.07	189.01	89.11	55.15	2832.52

Table 4c Seed oil properties of *Jatropha* species

Name	**Oil**	**FAMEs (%)**						**SN**	**IV**	**CN**	**GE**
	% (w/w)	P	S	O	L	M	O/L				kCal/mol
J. integerrima	40.20	8.64	4.39	11.98	74.75	0.24	0.16	190.90	139.21	43.57	2796.12
J. gossypifolia	47.90	11.33	5.16	23.05	59.95	0.51	0.38	190.68	123.16	47.21	2801.16
J. multifitida	50.10	17.86	5.32	26.32	49.15	1.35	0.53	190.14	107.33	50.85	2811.64
J. podagrica	42.70	8.92	6.13	24.47	59.43	1.05	0.41	189.20	123.48	47.36	2826.44
J. glandulifera	53.30	7.68	7.92	20.33	64.05	0.02	0.31	190.96	127.91	46.10	1796.04

Note FAMEs = Fatty Acid Methyl Esters**,** P = Palmitic acid, S = Stearic acid, O = Oleoic acid, L = Linoleic acid, M= micellanues (other fatty acids), O/L = Oleic to Linoleic acids ratio, SN = Saponification number, IV = Idoine value, CN = Cetane number, GE =Gross Energy

Table 5a Differential Water Use Efficiency (WUE) of *J. curcas* plants in the field

Species	**Sample**	**Photosynthetic rate** (A mmol m-2s-1)	**Transpiration rate** (E mmol m-2s-1)	**Water use efficiency (WUE)** (A/E mmol/mol)
J. curcas	JI	30.12	6.03	4.99
J. curcas	TCR31	28.26	5.87	4.81
J. curcas	TRC193	35.38	5.05	7.01
J. curcas	TNC-10	22.46	5.85	3.84

Table 5b Differential Water Use Efficiency (WUE) of *J. curcas* plants in the CER

Species	**Sample**	**Photosynthetic rate** (A mmol m-2s-1)	**Transpiration rate** (E mmol m-2s-1)	**Water use efficiency (WUE)** (A/E mmol/mol)
J. curcas	JI	31.12	7.03	4.42
J. curcas	JN	37.31	7.85	4.75
J. curcas	JT	28.21	7.91	3.57

2.5 Discussion

The *Jatropha* plant is a new system and only recently exposed to molecular investigations, mainly due to its increasing popularity as a biodiesel feedstock and valuable co-products (Kohli *et al*., 2009). The few studies on genotyping of *J. curcas* are based on assessing Indian accessions. Reddy *et al.* (2007) used AFLP/RAPD on 20 Indian accessions while Basha & Sujatha (2007) used 400 RAPD and 100 ISSR primers on 42 accessions from different regions in India. They also included the NTMA in their analysis and although it could be easily differentiated from the Indian accessions, there was 70% similarity between the two. Ranade *et al*., (2008) used single-primer amplification reaction (SPAR) to compare 21 accessions from different parts of India and demonstrated that 3 North East accessions were different among them and from other accessions analysed. Ram *et al*., (2008) and Pamidiamarri *et al*., (2009) compared *J. curcas* with additional *Jatropha* species from India and demonstrated clear divergence. In all these studies *J. curcas* accessions from different ecogeographic regions of India were 60 to 80% similar. It is important to test *J. curcas* accessions from wider eco-geographic regions to obtain genetic information of *J. curcas* at global scale. The outcome information will be used in breeding promgramme. Our analysis of 38 accessions from 13 countries around the world, along with 6 different species of *Jatropha* from India, again indicated 75% similarity among the global *J. curcas* accessions. We also showed that the NTMA clustered separately from other *J. curcas* accessions and that the genetic similarity coefficient between the Thai and the NTMA was high (0.76) as similarly noted by Basha & Sujatha (2007). Our assessment of nearly 52% similarity between the Thai *J. curcas* and *J. podagrica*, was also the same as reported by Ram *et al*., (2008) in the case of Indian *J. curcas* and *J. podagrica*. Our comparison of global accessions of *J. curcas* to other *Jatropha* species resulted in *J. integerrima* and *J. gossypifolia* being closer to *J. curcas* than other species as noted by Pamidiamarri *et al*., (2009) in comparing the Indian *J. curcas* to other *Jatropha* species from India. Such an overlap between local and global results indicates a narrow genetic diversity in *Jatropha*.

Incidentally, preliminary data suggests that *J. curcas* accessions exhibit monomorphism even with microsatellite markers (Drs. Gen Hua Yue and Hong Yan, Temasek Lifescience Laboratories, Singapore-personal communication), indicating very low genetic divergence among accessions spread around the globe. The reasons for the globally low genetic variability seen in *J. curcas* are not clear. Most likely, the anthropogenic and environmental influences in generating genetic variability are

missing because a) it is not a crop, b) as a well-surviving, undomesticated plant, it is highly stress tolerant due to adaptive genomic characters probably acquired before its global distribution and c) a limited stock has been vegetatively and apomictically propagated, since *J. curcas* is known to exhibit apomixis (Bhattacharya *et al*., 2005). The studies mentioned above indicate the limitations of using intra-specific breeding for *J. curcas* improvement; as exemplified by the failure to obtain low PE hybrids in crosses between *J. curcas* toxic accessions from India and the NTMA (Sujatha *et al.*, 2005). Hence we tried the interspecific hybridization approach. Of the three out-species we used (*J. integerrima, J. multifida* and *J. maheshwarii*) the hybrids that survived the best were those of *J. curcas* x *J. integerrima*, reiterating the relatedness between the two species. *J. integerrima* does not exhibit most agronomic or commercially useful traits to desirable degrees, for example it contains the highest amount of seed PE, second only to *J. multifida* (Table 2a). It is also a plant with very thin stem, an intermediate level of oil content and a low O/L ratio (Table 4c) in contrast to the high value desirable. However, analysis of the F1 plants revealed advantages gained through hybridisation, in agronomic and commercial traits (Table 3) including low PE levels (Table 1b). A back cross with *J. curcas* NTMA further reduced the PE levels comparable to NTMA – a highly desirable target (Table 2). Our analysis revealed higher genetic distance (GD) between *J. curcas* and *J. integerrima* than between *J. curcas* accessions. Higher GD favourably affects F1 performance (Kiula *et al*., 2008). Interspecific hybridization with compatible species leads to generation of valuable material due to heterozygosity at several loci. The effect could be more pronounced if most loci in the parents were homozygous to start with. Our cTBP analysis of the hybrids (Figure 6) indeed suggests parental lines homozygous for alleles of the beta-tubulin gene family – one that is known to be variable (Breviario *et al*., 2007; 2008), in turn suggesting the possibility of homozygosity at the other loci. Additionally, a larger numbers of genes express differential alleles in a hybrid (Guo *et al*., 2008), which if resulting from genetically diverse parents permits larger numbers of permutations and combinations, in turn resulting in pronounced range of differences in the progeny in comparison to the parents. A combination of such possibilities may underlie our observations of improved hybrids despite *J. integerrima* apparently not being an ideal material. Our preliminary data with selection of plants from additional back cross progeny (BC3) exhibits stable inheritance of the improved characters illustrated in Table 3. In any case, our results suggest that excluding the Indian accessions due to their narrow genetic base and trying to find sufficient diversity within *J. curcas* around the world for selecting breeding lines

may have very limited value due to the globally narrow genetic base of *J. curcas* that we illustrate. There are a few reports on genetic characterisation of *J. curcas* and even fewer on systematic assessment of phenotypic characteristics of different accessions. We noticed a large variation in seed size and oil content in seeds from various parts of the world, while they exhibited high genetic similarity. Kaushik *et al*., (2007) reported the predominant role of the environment in the higher phenotypic coefficient of variation in some characters in 24 Indian accessions *J. curcas* originating from different agroclimatic zones. Although they did not conduct genetic similarity tests, our results and other genotyping studies mentioned above suggest a high likelihood of a narrow genetic base in the accessions tested. Our results of large differences in morphological and biochemical phenotype for seed size and seed oil content and composition respectively, in accessions with a narrow genetic base, implicated environmental factors-mediated changes. Recently, Sunil *et al*., (2008) recorded the phenotypic traits of *J. curcas* plants *in-situ* at 4 different eco-geographical regions of India. They noticed pronounced differences in the 9 characters they assessed for a total of 162 accessions in the 4 zones. For example, the plant height of 80% accessions in one zone was less than 1.5 meters while in another zone 60% of the accessions were larger than 1.5 meters. Similar differences were noticed in number of fruits and seed oil content and composition. Environment related developmental differences noticed in *J. curcas* by Sunil *et al*., (2008) are reminiscent of such differences noticed in *Achillea* in the classical study by Clausen *et al*., (1958) wherein an identical genotype (clonal cutting) grew to be the tallest at low elevations but was the shortest plant at medium elevations and the second tallest at high elevations. Since Sunil *et al*., (2008) also did not undertake genetic characterisation of the accessions, the reason for the variability was not clear. However, once again, considering our results and the genetic studies conducted in India and mentioned above, it is not unreasonable to expect a narrow genetic base in the 162 accessions tested. If highly related Indian accessions demonstrated large phenotypic variability then in our study, similarly related accessions from totally different countries are prone to showing even larger phenetic variability. Our analysis of WUE in *J. curcas* is the first such report to the best of our knowledge. Although our WUE results provide only a general idea under a single set of conditions, both in the field and in the CER, we obtained reproducible WUE values in the field in India and in the CER in the UK, which compared well with WUE studies on other hardy plants. For example, the range of WUE of olive trees, a hardy plant like the *J. curcas*, in Tunisia on marginal rain-fed sandy soil and under similar but irrigated land varied

between 4.7 to 7.6 mmol/mol (Ben-Rouina *et al*., 2007). Similar transpiration rates in different accessions of *J. curcas* may indicate similar stomatal density. Since the three accessions tested in CER originated from three countries, differences in photosynthetic rates and hence in WUE in genetically highly similar accessions under similar environmental conditions implicate allopatric, epigenetic and stochastic factors.

One character that has been well studied for environment dependent changes in plants is flowering in plants such as *Arabidopsis and Antirrhinum*. In the case of *Jatropha*, flowering phenology was studied for an accession from Cape Verde on a field in Nicaragua (Aker, 1997). Flowering time, number and male:female flower ratio all varied substantially depending on soil fertility, soil moisture, precipitation, Piche evaporation and temperature. Our unpublished observations are similar to environment-dependent differences noticed by Aker (1997) in that a *J. curcas* accession exhibits one major flowering flush in one eco-geographical zone (around Hyderabad in India) whereas in another zone (the more humid parts of Kerala in India) it exhibits episodic flowering depending on rain and moisture conditions. A pronounced phenotypic plasticity is in itself a genotypic trait that allows the plant to respond to different environments through morphological and physiological changes for its survival (Richards *et al*., 2006). In our study, although genetic analysis clustered accessions from a single region separately (Figures 1 and 2), interestingly, clustering the accessions based on the data in Table 4 indicated similarity between accessions depending on the origin of the seeds, highlighting the influence of the local conditions. For example, accessions from particular provenances of Thailand clustered together through such an analysis (Figure 8). We propose that high genetic similarity in global accessions of *J. curcas* can exhibit large phenotypic variability largely through epigenetic mechanisms. The potential for epigenetic marks to contribute to heritable differences depending on environmental differences has been recently examined (Bossdorf *et al*., 2008; Grant-Downton & Dickinson, 2005) including in similar seed traits of *Arabidopsis* (Blodner *et al*., 2007) as studied for *J. curcas* in this report. The molecular basis of phenetic variability in genetically similar backgrounds may likely reside in stochastic and/or epigenetic factors regulating gene transcription. We noticed a relatively higher amount of the *fad* gene transcript, most likely responsible for the low O/L ratio in JT seeds, as compared to that in JI and JN seeds. Similarly, JI exhibited larger amounts of seed triacylglyceride (TAG) and revealed larger amount of the diacylglyceride acyl transferase (*dgat*; DQ278448*)* transcript (Figure 9). Given the general genetic similarity, it is unlikely that the difference we noticed in the transcript level of essential genes such

as *fad* and *dgat* in JT and JI respectively were due to differences in the respective promoters of the genes. Blodner *et al.*, (2007) recently discussed the temperature-based differential transcription of fatty acid desaturases in other oilseed plants. We have cloned and compared the coding region and promoter of the phosphoenol pyruvate carboxylase (*pepc*) gene and three oleosin genes from JI and JN and found them to be identical. Transcript-based expression analysis of these genes exhibit differences similar to the *fad* gene. Moreover, the lack of polymorphism resulting after AFLP analysis of *J. curcas* has been resolved in favour of an epigenetics-based hypothesis due to a high degree of polymorphism noticed with methylation-sensitive AFLP, implicating differential methylation of genes in various accessions (Dr. Hong Yan, Temasek Lifescience Laboratory Singapore – personal communication). Finally, phenetic variability in the face of widespread homozygous loci as preliminarily hypothesised through our cTBP results on parental and F1 hybrids (Figure 6) adds to the proposition of epigenetics-mediated reasons for the changes noticed and suggests the strength of proteomic-based approach.

For *J. curcas* to be a commercially viable source of biodiesel much research and development is needed and little should be expected from the present widespread planting of wild varieties. Sujatha *et al.*, (2008) recently reviewed the role of biotechnology in the improvement of *J. curcas.* A combination of breeding and transgenic approaches could deliver elite varieties. Whether or not this happens depends very much on global commodity markets and the competition from other energy technologies and energy crops. However, the simultaneous value of *J. curcas* as a novel research system must be appreciated. Its difference in habit as a shrub compared to model herb *Arabidopsis* and model tree poplar may be important in terms of the extent of DNA and histone modifications; just as such difference exists between *Arabidopsis* and poplar.

2.6 References:

Aker CL. 1997. Growth and reproduction of *Jatropha curcas*. Biofuels and Industrial Products from *Jatropha curcas*. In: Gubitz GM, Mittelbach M, Trabi M, eds. *Proceedings of 'Jatropha 97'.* Dbv-Verlag für die Technische Universität Graz, 2-18.

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.

Ben Rouina B, Trigui A, d'Andria R, Boukhris M, Chaieb M. 2007. Effects of water stress and soil type on photosynthesis, leaf water potential and yield of olive trees (*Olea europaea* L. cv. Chemlali Sfax) *Australian Journal of Experimental Agriculture.* 47: 1484-1490.

Bhattacharya A, Datta K, Datta SK. 2005. Floral biology, floral resource constraints and pollination limitation in *Jatropha curcas* L. *Pakistan Journal of Biological Sciences*. 8: 456-460.

Blodner C, Goebel C, Feussner I, Gatz C, Polle A. 2007. Warm and cold parental reproductive environments affect seed properties, fitness, and cold responsiveness in *Arabidopsis thaliana* progenies. *Plant, Cell & Environment*. 30:165-175.

Boschin G, D'Agostina A, Annicchiarico P, Arnoldi A. 2008. Effect of genotype and environment on fatty acid composition of *Lupinus albus* L. seed. *Food Chemistry* 108: 600-606.

Bossdorf O, Richards CL, Pigliucci M. 2008. Epigenetics for ecologists. *Ecological Letters* 11: 106-115.

Breviario D, Baird WV, Sangoi S, Hilu K, Blumetti P, Gianì S. 2007. High polymorphism and resolution in targeted fingerprinting with combined beta-tubulin introns. *Molecular Breeding* 20: 249-259.

Breviario D, Morello L, Manca A, Gianì S. 2008. The importance of being an intron by wild type tubulin genes. In: Baird WV, Blume Y, Breviario D, eds. *The Plant Cytoskeleton: genomic and bioinformatics tools for biotechnology and Agriculture*. Springer AK/NATO Publishing Unit, pp. 199-218.

Clausen K, Keck D, Hiesey W. 1958. Experimental studies in the nature of species – 3. Environmental responses of climatic races of *Achillea*. *Carnegie Institution of Washington Publication* 581: 1-129.

Devappa RK, Swamylingappa B. 2008. Biochemical and nutritional evaluation of *Jatropha* protein isolate prepared by steam injection heating for reduction of toxic and antinutritionalfactors. *Journal of the Science of Food and Agriculture* 88: 911-919.

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

Freedman B, Bagby MO. 1989. Heats of combustion of fatty esters and triglycerides. *Journal of American Oil Chemical Society* 66: 1601.

Grant-Downton RT, Dickinson HG. 2005. Epigenetics and its implications for plant biology-1. The epigenetic network in plants. *Annals of Botany* 96: 1143-1164.

Gressel J. 2008. Transgenics are imperative for biofuel crops. *Plant Science* 174: 246-263.

Gubitz GM, Mittelbach M, Trabi M. 1999. Exploitation of the tropical oil seed plant *Jatropha curcas* L. *Bioresource technology* 67: 73-82.

Guo M, Yang S, Rupe M, Hu B, Bickel DR, Arthur L, Smith O. 2008. Genome-wide allelespecific expression analysis using Massively Parallel Signature Sequencing (MPSS (TM)) Reveals cis- and trans-effects on gene expression in maize hybrid meristem tissue. *Plant Molecular Biology* 66: 551-563.

Haas W, Mittelbach M. 2000. Detoxification experiments with the seed oil from *Jatropha curcas* L. *Industrial crops and Products* 12: 111-118.

ISAAA. 2008. First draft of Palm oil genome and progress on *Jatropha* genome sequencing. June 13 issue. Available at: http://www.isaaa.org/kc/cropbiotechupdate/ biofuels/default.asp?Date=6/13/2008.

Kalayasiri P, Jeyashoke N, Krisnangkura K. 1996. Survey of seed oils for use as diesel fuels. *Journal of American Oil Chemical Society* 73: 471–474.

Kaushik N, Kumar K, Kumar S, Kaushik N, Roy S. 2007. Genetic variability and divergence studies in seed traits and oil content of *Jatropha* (*Jatropha curcas* L.) accessions. *Biomass and Bioenergy* 31: 497–502.

Kiula BA, Lyimo NG, Botha AM. 2008. Association between AFLP-based genetic distance and hybrid performance in tropical maize. *Plant Breeding* 127: 140-144.

Kohli A, Raorane M, Popluechai S, Kannan U Syers KJ, O'Donnell AG. 2009. *Jatropha curcas* as a novel, non-edible oilseed plant for biodiesel. In: Ferry N, Gatehouse AMR, eds. *Environmental impact of genetically modified novel crops*. Chapter 14. CAB International, London, UK, pp. 294 -322.

Krisnangkura K. 1986. A simple method for estimation of cetane index of vegetable oil methylesters. *Journal of American Oil Chemical Society* 63: 552–553.

Kumar A, Sharma S. 2008. An evaluation of multipurpose oil seed crop for industrial uses (*Jatropha curcas* L.): A review. *Industrial crops and products* 28: 1-10.

Lewontin R, Goss P. 2004. Developmental canalization, stochasticity and robustness. In: Jen E, ed. *Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies*. Oxford University Press, New York, pp. 21–44.

Li MR; Li HQ; Jiang HW, Pan XP, Wu JG. 2008. Establishment of an *Agrobacterium*-mediated cotyledon disc transformation method for *Jatropha curcas*. *Plant Cell Tissue and Organ Culture* 92: 173-181.

Lin JJ, Kuo J, Ma J. 1996. A PCR-based DNA fingerprinting technique: AFLP for molecular typing of bacteria. *Nucleic Acids Research* 28: 3649-3650.

Lin J, Chen Y, Xu Y, Yan F, Tang L, Chen F. 2003. Cloning and expression of curcin, a ribosome-inactivating protein from the seeds of *Jatropha curcas*. *Acta Botanica Sinica* 45: 858-863.

Luo MJ, Yang XY, Liu WX, Xu Y, Huang P, Yan F, Chen F. 2006. Expression, purification and anti-tumor activity of curcins. *Acta Biochimica Et Biophysica Sinica* 38: 663-668.

Makkar HPS, Becker K, Sporer F, Wink M. 1997. Studies on nutritive potential and toxic constituents of different provenances of *Jatropha curcas*. *Journal of Agricultural and Food Chemistry* 45: 3152-3157.

Makkar HPS, Aderibigbe AO, Becker K. 1998. Comparative evaluation of non-toxic and toxic varieties of *Jatropha curcas* for chemical composition, digestibility, protein degradability and toxic factors. *Food Chemistry* 62: 207-215.

Makkar HPS, Becker K. 1999. Plant toxins and detoxification methods to improve feed quality of tropical seeds – Review. *Asian-Australasian Journal of Animal Sciences* 12: 467-480.

Nature Editorial. 2007. Kill king corn. *Nature* 449: 637.

Pamidiamarri S, Nirali Pandya N, Reddy MP, Radhakrishnan T. 2009. Comparative study of interspecific genetic divergence and phylogenic analysis of genus *Jatropha* by RAPD and AFLP: Genetic divergence and phylogenic analysis of genus *Jatropha*. *Molecular Biology Reports* 35: 901-907.

Popluechai S, Raorane M, Syers KJ, O'Donnell AG. Kohli A. 2008. Research needs to make Jatropha a viable alternate oilseed crop. *Proceedings of the International Workshop on Feasibility of Non-Edible Oil Seed Crops for Biofuel Production,* May 25 – 27, 2007, Mae Fah Luang University, Chiang Rai, Thailand. Eds. Syers KJ and Wood D. MFLU publication, pp. 56 - 60.

Prabakaran AJ, Sujatha M. 1999. *Jatropha tanjorensis* Ellis & Saroja, a natural interspecific hybrid occurring in Tamil Nadu, India. *Genetic Resources and Crop Evolution* 46: 213-218.

Ram SG, Parthiban KT, Kumar RS, Thiruvengadam V, Paramathma M. 2008. Genetic diversity among *Jatropha* species as revealed by RAPD Markers. *Genetic Resources and Crop Evolution* 55: 803-809.

Ranade SA, Srivastava AP, Rana TS, Srivastava J, Tuli R. 2008. Easy assessment of diversity in *Jatropha curcas* L. plants using two single-primer amplification reaction (SPAR). *Biomass and Bioenergy* 32: 533-540.

Reddy MP, Chikara J, Patolia JS, Ghosh A. 2007. 'Genetic improvement of *Jatropha curcas* adaptability and oil yield', In: FACT SEMINAR on *Jatropha curcas* L.: agronomy and genetics. Wageningen, The Netherland, 26-28 March 2007, FACT Foundation. Availablae at: http://www.jatropha-alliance.org/fileadmin/documents/knowledgepool/ReddyChikara_Genetic_Improvement.pdf (accessed 25/10/10)

Richards CL, Bossdorf O, Muth NZ, Gurevitch J, Pigliucci M. 2006. Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters* 9: 981-993.

Rohlf FJ. 1997. NTSYSpc: version 2.01. Exeter Publishing, Setauket, NY.

Saleem MF, Ma BL, Malik MA, Cheema MA, Wahid MA. 2008. Yield and quality response of autumn-planted sunflower (*Helianthus annuus* L.) to sowing dates and planting patterns. *Canadian Journal of Plant Science* 88: 101-109.

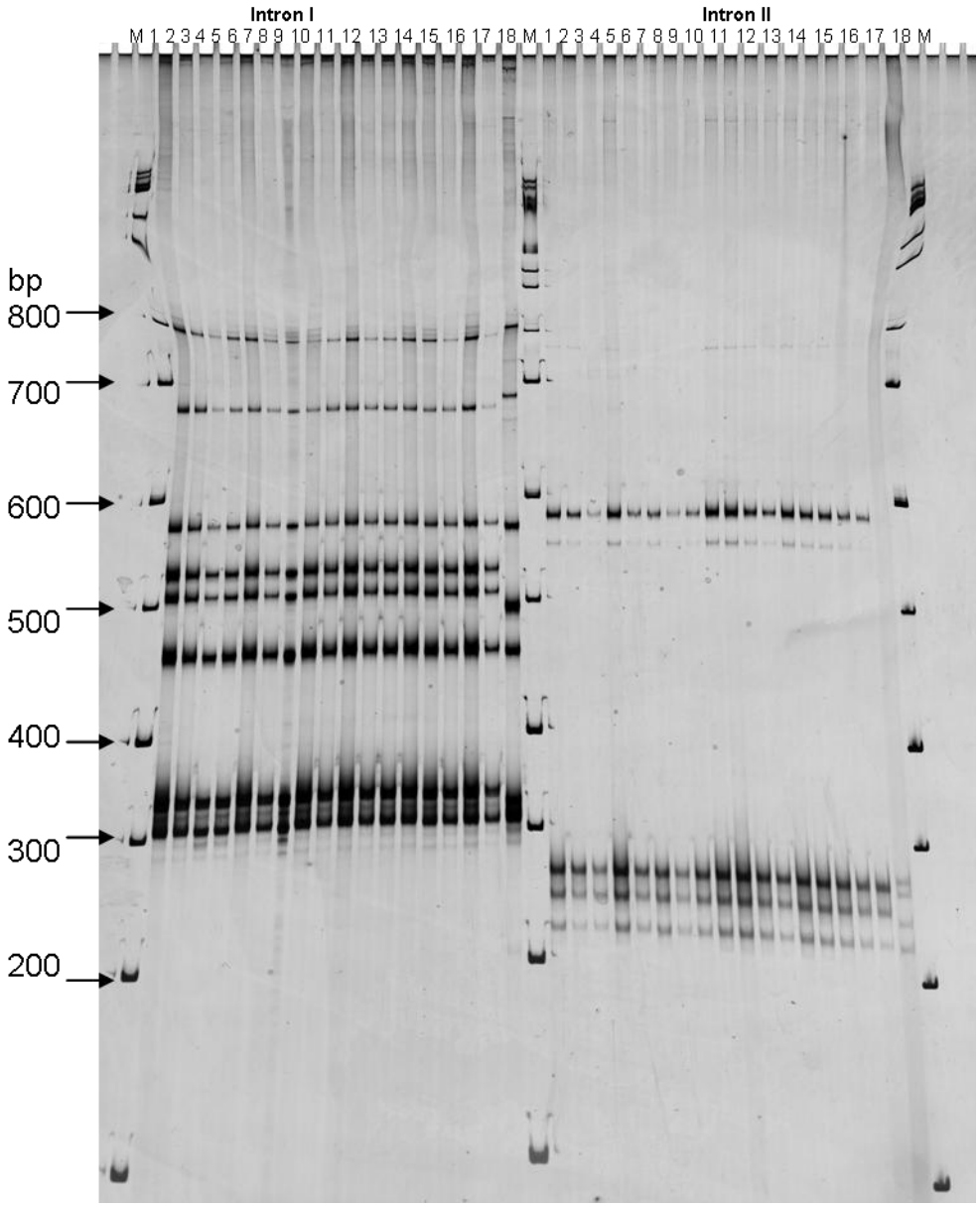
Sujatha M, Reddy TP, Mahasi MJ. 2008. Role of biotechnological interventions in the improvement of castor (*Ricinus communis* L.) and *Jatropha curcas* L. *Biotechnology Advances* 26: 424-435.

Sujatha M, Makkar HPS, Becker K. 2005. Shoot bud proliferation from axillary nodes and leaf sections of non-toxic *Jatropha curcas* L. *Plant Growth Regulation* 47: 83-90.

Sujatha M, Prabakaran AJ. 2003. New ornamental *Jatropha* hybrids through interspecific hybridization. *Genetic Resources and Crop Evolution* 50: 75-82.

Sunil N, Varaprasad KS, Sivaraj N, Kumar TS, Abraham B, Prasad RBN. 2008. Assessing *Jatropha curcas* L. germplasm in-situ-A case study. *Biomass and Bioenergy* 32: 198-202.

2.7 Appendix



Appendix 1 Combinatorial Tubulin Based Polymorphism (cTBP) for β-tubulin gene family introns analysis conducted on 17 *J. curcas* and one *J. podagrica* accession. Accessions identified as follows: 1-JI; 2-JI-2; 3-JN; All other accessions were from different provenances of Thailand represented here as 4- JT; 5-ST; 6-KU; 7-CB; 8-KS; 9-PB; 10-PC8; 11-PC18; 12-PC29; 13-PC32;14-PC62; 15-PC69; 16-PC76; 17-PC81 and 18 – Thailand *J. podagrica.*

Chapter 3: Understanding *Jatropha curcas* seed oil storage components: characterization of seed oilbody oleosins

3.1 Abstract

The *Euphorbiaceous* plant *Jatropha curcas* is becoming popular as a source of biodiesel. As in other oilseed plants such as rapeseed and soybean, the seed oil in *J. curcas* is stored in sub-cellular organelles called oilbodies. No information exists on *J. curcas* oilbodies. We report proteome composition of *J. curcas* oilbodies using liquid chromatography-mass spectrometry (LC-MS/MS) which revealed oleosins, caleosin, steroleosin, aquaporin and curcin. The oleosins were characterized at the gene, transcript and protein level. Using degenerate primers, RACE and gene walking, three *Jatropha* oleosins were isolated and sequenced. All three genes contained the proline knot domain characteristic of oleosins. Quantitative real-time PCR analysis revealed a single copy of each of the three genes. An intron identified in one of the *oleosins*, *Jcolesoin3,* was reminiscent of an intron containing *Arabidopsis oleosin*. Isolation of the 623 bp upstream regulatory sequence of *Jcolesoin3* confirmed the presence of a number of seed specific *cis*-regulatory motifs. The sequence was an addition to a limited number of promoters known from *J. curcas*. Transcript expression analysis confirmed preferential expression of the *J. curcas* oleosins during seed development. Immunoassays with antibodies raised against *Arabidopsis* oleosins or against the *JcOleosin3* were used to confirm the relationship of *J. curcas* oleosins with those from other plants. The antibodies and the intron sequence of *Jcoelosin3* were used to assess the diversity of oleosins in *J. curcas* accessions, species and hybrids. This analysis identified SNPs associated with *JcOleosin3* alleloforms. These data are important for varietal improvement of *J. curcas*, targeting qualitative and/or quantitative modification of the seed oil because oleosins largely dictate the size and number of oilbodies and hence the oil storage capacity in oilseed plants.

Keywords: Biodiesel, intron, *Jatropha curcas*, oilbodies, oleosin, proteome, seed-specific, SNP.

3.2 Introduction

Jatropha curcas belongs to the family *Euphorbiacea* and inhabits different areas of Central and South America, South East Asia, India and Africa. It is vegetatively propagated and can be established in a wide variety of soils. *J. curcas* seeds contain 27-40% non-edible oil, which can be used for the production of bio-diesel (Achten et al., 2007). The non-edible nature of the oil is due to the presence of toxins like toxalbumin or curcin and co-carcinogenic phorbol esters. Although recent information on its unsuitable comparative water footprint as a bioenergy crop (Gerbens-Leenes et al., 2009) has been strongly argued (Maes et al., 2009), support for it as a suitable feedstock for biodiesel under specific conditions remains high (Jongschaap et al., *20*09). In order to make it a successful bioenergy crop, substantial research is required to improve the oil yield. One obvious priority for such research is fatty acid (FA) biosynthesis in *J. curcas* seeds. Any manipulation of it requires an understanding of FA biosynthesis and storage in the *J. curcas* seeds.

In plants, lipids are stored in the form of triacylglycerols (TAGs) in oilbodies called lipid droplets or lipid bodies (Huang 1992). Seeds are the most common site for accumulation of lipids as storage reserves for germination and post-germinative growth (Murphy et al., 2001). An oilbody (0.5-2μm in diameter) consists of a core of TAGs matrix surrounded by a monolayer of phospholipids. The phospholipid layer is stabilized by the embedded proteins (Tzen and Huang, 1992). Oilbodies routinely consist of 94-98% lipids, 0.5-2% phospholipids and 0.5-3.5% proteins (Chen et al., *20*04) and 80-90% of these proteins belong to the class of proteins known as oleosins (Qu and Huang, 1990) which cover the entire surface of the oilbody and dictate its size (Capuano et al., *20*07). Oleosins were first identified in seed cotyledons of sunflower and safflower by Slack et al., (1980) and are ubiquitous in oil storing seeds of monocots or dicots including oilseed rape, maize, cotton, tea, coffee, soybean, sunflower, safflower and rice (Capuano et al., *20*07). They are low-molecular weight amphipathic proteins of about 15-26 kDa classified into low and high molecular weight isoforms based on their molecular weights (Tzen et al., 1990). Evolution of oleosins can be traced by the presence of high or low molecular weight isoforms. Angiosperms contain both low and high molecular weight isoforms, whereas the oleosins identified in gymnosperms such as *Pinus koraiensis*, *Ginkgo biloba*, and *Cycas revoluta* contain only the low molecular weight isoform (Wu et al., 1999).

Oleosins are characterized by the presence of three domains, the N-terminal and C-terminal domains which are hydrophilic and a central long hydrophobic domain. The N and C-terminal domains vary both in their amino acid residue length and composition. The central hydrophobic domain is composed of approximately 75 uninterrupted and uncharged amino acid residues (Huang, 1996). The C-terminal portion is present approximately 30 residues from the hydrophobic domain and forms an amphipathic α-helical structure horizontally with charged phosphate and choline groups of the phospholipid layer on the oil body surface (Tzen et al., 1992; Hsieh and Huang, 2004). The N-terminal region contains many repeats of short peptides that are not conserved. These peptides have several glycine residues. Hence oleosins are sometimes referred to as Gly-rich proteins. Amino acid studies from 40 different seed oleosins revealed that the central hydrophobic domain is highly conserved and rich in aliphatic amino acids, alanine, glycine, leucine, isoleucine and valine (Li et al., 2002; Capuano et al., 2007). The hydrophobic domain is the longest hydrophobic stretch found in any protein in any organism and is the hallmark of oleosin protein. The domain forms a hairpin structure which penetrates the surface phospholipid monolayer of an oilbody into the matrix. The centre of the hydrophobic stretch has three proline residues and one serine residue. These residues interact and result in the formation of a "Proline knot". Proline residues in polypeptides contribute to the breaking or turning of the α-helix and β-sheet structures (Hsieh and Huang, 2004).

Oleosins play an important role in stabilization of oilbodies by preventing coalescence of oilbodies during desiccation. Correlation between oilbody size and oleosin levels suggests that accumulation of oleosins modulate oilbodies *in vivo* (Siloto et al., 2006; Ting et al., 1996). Species containing higher amounts of oleosins have smaller oilbodies compared with those of lower oleosin content (Hu et al., 2009). Leprince et al., (1998) proposed that oleosins do not play a major role in lipid-body biogenesis but a key role in post-germinative rehydration of seeds. Oleosins and oleosin like proteins have also been identified in plant tissues other than the seeds. Tapetal cells of anthers of *Arabidopsis* and *Brassica* contain oleosin-like proteins where the oilbody-like structures are called tapetosomes (Kim et al., 2002; Dunbar, 1973). Additional oleosin-like proteins such as the caleosins and steroleosins are also known (Naested et al., 2000; Lin et al., 2002). The latter are unique in the location of the proline knot in the N-terminal region (D'Andrea et al., 2007a)

To date, no information is available on oilbodies or oleosins from *J. curcas.* Considering its increasing popularity as a source of biodiesel, it is highly pertinent that

its seed oil storage organelles and their components are characterized. Since the nature and size of oilbodies is largely dependent on the protein component of these organelles, we used proteomic tools to identify the different protein components of the oilbodies. We characterized the major component protein of oilbodies, the oleosins, at the gene, transcript and protein levels. Three *oleosins* of *J. curcas* were successfully isolated and characterized. The upstream regulatory promoter region of *JcOleosin3*, the highly seed-specific oleosin, was isolated and characterized *in silico* for *cis*-active regulatory elements (CAREs), while an intron sequence specific to this gene was discovered in its coding region. The intron was used as a marker along with anti-oleosin antibodies to investigate the diversity of *oleosins* in various *J. curcas* accessions and other *Jatropha* species. Transcript and protein analyses revealed that all three isoforms express predominantly in the seed, although two of these also express at low levels in other tissues. Our results form the basis for an understanding on *J. curcas* oilbodies/oleosins. The comparative differences and similarities between *J. curcas* oilbodies/oleosins with those from other plants are discussed. We also discuss the importance of these results in appreciating the role of lipid storage organelles in the face of popular proposals to modify the *J. curcas* seed FA composition for improved quality and quantity.

3.3 Materials and Methods

Seed material used

The *J. curcas* samples used in this study were *J. curcas* accessions from India (JI); Indonesia, (JIN); Mexico, (Non-toxic; NTM); India-Mettupalem (JIM); Uganda (JU); Nigeria, (JN); Thailand, (18 accessions; JT1 – JT18); Backcross (BC_1F_1) hybrid clones of *J. curcas* and *J. interegima* (Hybrid clone 91-1, 92, 96, 111, 201, 202, 205, 213, 222, 223, 225, and 226), and *Jatropha* sp. (*J. glandulifera, J. multifida, J. podagrica, J. gossypifolia, J. maheswarii, and J. interegima*). Most seed material was obtained from the collection Dr. Sujatha Mulpuri at the Directorate of Oilseed Research, Hyderabad, India. These included the hybrid clones; while others were acquired from *Jatropha* seed collections of the institutes participating in this research.

DNA extraction, PCR, cloning, and sequencing

DNA was extracted from 18 seed samples based on the procedure of Sujatha et al. (2003). PCR was carried out according to Sambrook et al. (2001). PCR reaction was composed of 1X buffer: 5 U of Taq DNA polymerase (Invitrogen); 1.5 mM of $MgCl_2$; 0.2 mM of dNTP Mix; 0.2 pmol of forward and reverse primers; 100 ng of template DNA, and DEPC-treated water up to 50 μl. The following conditions were set up: 94 ^oC – 5min, 94^oC - 45sec, 50-65 ^oC -30sec, 72^oC -1 min for 35 cycles, 72^oC for 5 min and store at -20^oC for further analysis. The products were run in agarose gel and viewed under UV illuminator. DNA cloning was carried out using Strataclone Cloning Kit (Stratagene). DNA sequencing was carried out using BigDye^{TM} version 3.1 Terminator Cycle Sequencing Kit (PE Applied Biosystems, MA). The sequences were obtained as electropherogram files and were viewed through Finch TV (version 2.0), blasted using nblast to check the sequences similar to the insert. The sequences were then aligned using Clustal X version 2.0 alignment tool.

PCR cloning of cDNA fragments encoding oleosins

J. curcas oleosins were cloned using a degenerate primer (named JODEGF, Supplemental Table 1), designed after sequence alignment of multiple oleosins from *Ricinus communis, Helianthus annuus,* and *Arabidopsis thaliana,* in combination with oligo $(dT)_{20}$ (Invitrogen) for PCR amplification of seed derived cDNA as template. The 5' upstream sequences were obtained with 5' RACE (Invitrogen) using three specific primers (JcO1R, JcO2R, and JcO3R; see Supplemental Table 1). Complete coding sequences for three oleosin isoforms were obtained after aligning overlapping sequences of 3 fragments each obtained from the 2 PCRs, the one with the degenerate primer and Oligo dT and the other after respective 5'-RACE reactions. Complete coding sequences were finally amplified using terminal gene specific primers shown in Supplemental Table 1. The analysed sequences were submitted to NCBI.

Isolation of upstream promoter sequence from J. curcas genomic DNA

Upstream promoter sequences of *J. curcas* oleosin were determined using DNA Walking SpeedUp™ kit (Biogene, United Kingdom) following the procedures described by the manufacturers. Briefly, the DW-ACP primer set from Seegene kit was used with a series of gene specific primers (O3DWGSP1–3, Supplemental Table 1) to extend the sequence as obtained by using the above procedure. The DNA walking reaction was performed using the *Taq* polymerase PCR kit (Fermentas) and conditions as described above. Sequences obtained from various amplification steps were assembled in a single contig.

RNA extraction and cDNA synthesis

RNA extraction of *Jatropha* seeds were performed using TRIzol reagent (Helena) according to the manufacturer's instructions. The RNA was treated with RNase-Free DNaseI (Invitrogen) to remove contaminating DNA. RNA concentration was determined using a NanoDrop® ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies). First strand cDNA from various tissues or different physiological stages was synthesized according to the manufacturer's instructions (Superscript II Reverse

Transcriptase (RT), Invitrogen). The cDNA was stored at -20^{0}C until further analysis by semi-quantitative RT PCR or quantitative PCR.

Quantitative and semi-quantitative PCR

Quantitative PCR was carried out with cDNA prepared above as recommended by the manufacturer (SYBR Green qPCR Supermix Universal, Invitrogen); primers used in this study are shown in Supplemental Table 1**.** The cDNA synthesized from seed samples was used as template. Twenty five microlitre of reaction was set with 12.5µl of 2X Master Mix, 0.5 µl [10 pmol] of forward primer, 0.5 µl [10 pmol] of reverse primer, 9 µl of DEPC-treated water and 2.5µl of cDNA was added. The following thermal procedure was used: 50^{o}C for 2 min, 95^{o}C for 10 minutes, following by 40 cycles of 95^{o}C for 15 seconds, then 60^{o}C for 1 minute. Melting curve was established between 60^{o}C to 95^{o}C at 0.2 seconds intervals. The samples were analysed and compared to reference genes, *J. curcas* 18S or *J. curcas* beta-ketoacyl-ACP synthase III gene (KAS III; DQ987701), which was previously reported to exist as a single copy in the *J. curcas* genome and used for similar comparisons (Li et al., *20*08). A graph was plotted using the ratio values obtained and standard deviation of these values were calculated and shown as error bars in the graph.

Semi-quantitative PCR was carried out using qPCR gene specific primers on first strand cDNA obtained from above. Two microlitre of first stand cDNA reaction was used to perform PCR as described above. The following conditions were set up: 94^{o}C – 5min, 94^{o}C - 45sec, 55^{o}C-30 sec, 72^{o}C -1 min for 28 cycles, 72^{o}C for 5 min and store at -20^{o}C for further analysis. The products were run in agarose gel (1.5 % containing 0.5µg/ml of ethidium bromide) and viewed under UV illuminator. Densitometry analysis was performed using BioRad Quantity One Fluor-S Multi-Imager software.

Total protein extraction

Plant tissue was ground in liquid N_2 and 100 mg samples extracted in 1 ml of 50 mM Tris-Cl buffer containing 8M Urea, 4% CHAPS, 40 mM DTT, and 1:100 dilution of plant protease inhibitor cocktail (Sigma) by agitation for 2 h at 4^{o}C. Then the mixture was centrifuged at 15,000 g for 15 min at 4^{o}C and the supernatant retained for further analysis.

Oleosin and lipid bodies extraction

Oleosin extraction was carried out according to D'Andrea et al (2007b). In brief, half gram (0.5 g) of *J. curcas* seed kernel was ground (~1 seed) in 5 ml of 50 mM Na_2CO_3 pH 9. Aliquots of 200 µl of this mixture were dispensed in 2 ml Eppendorf tubes and add 9 vols. (1800 ul) of chloroform/methanol (11/7) mix was added to it. The mixture was then vortexed and sonicated for 30 seconds in water bath sonicator and spun at 20,000 g for 15 min at 20^{o}C. The organic solvent was transfered to a new tube and evaporated under nitrogen for 1 hr. The protein pellet was analyzed (or kept at -20^{o}C until further use). Protein concentration was estimated using 2D quant kit (Amersham). Lipid bodies extraction was carried out using flotation by centrifugation method as previously described in Jolivet et al (2004).

Antibody production

An presumably antigenic peptide of 15 amino acids was selected from the alignment of *J. curcas* oleosin3 against *Arabidopsis thaliana* oleosins and synthesized coupled to Keyhole Limpet Hemocyanin (KLH) at Genscript Corporation, USA. The antigenic peptide (200 mg) was subcutaneously inoculated to two rabbits according to manufacturer procedure (Genscript Corporation, USA). Antiserum titre was tested by dot blot using synthetic peptide samples as standards.

SDS-PAGE and Immuno assay by western blot

Protein electrophoresis was carried out in ready-to-use NuPAGE Novex 12% Bis-Tris gels (Invitrogen, France) and NuPAGE MES SDS running buffer from Invitrogen (France) were used in Figure 6a. Whereas the NuPAGE Novex 12% Bis-Tris gels (Invitrogen) and Tricine running buffer were used in 6b. [D'Andrea et al. (2007b). Alternatively, Tris-glycine gels containing 12 % acrylamide and Tris-glycine buffer were used in Figure 5 and 6c, according to Laemmli (1970). Molecular mass markers were SeeBlue® Plus2 Pre-Stained Standard (for Figure 5, 6a and 6b; Invitrogen) and Spectra™ Multicolor Broad Range Protein Ladder (for Figure 6c, Feremntas). Protein fractions (oil body proteins and solvent extracted proteins) were re-suspended in 2X Laemli buffer and separated on 12% polyacrylamide gel by SDS-PAGE. Gels were stained with Coomassie blue or proteins were transferred to PVDF for immunoblot

analysis with antibodies directed against the N terminal region of *A. thaliana* oleosin S2 (anti-rS2N) at 1/1000 dilution, or N terminal region of *J. curcas* oleosin3 (anti-rJO3N) antibody at 1/5000 dilution. Primary antibodies were revealed with Peroxidase-conjugated goat anti-rabbit IgG antibody (1:5000) from Sigma (UK). Peroxidase activity was revealed using ECL Plus™ Western Blot Detection Reagents (Amersham, GE healthcare, UK) according to manufacturer protocol

Protein Identification by LC-MS/MS

After staining with Coomassie blue, protein bands from SDS PAGE were excised, washed, reduced (DTT), alkylated (iodoacetamide) and trypsin digested. The resulting peptides were extracted and analysed by LC-MS/MS as previously described by Jolivet et al. (2009). Protein identification was performed with Bioworks 3.1 software using the non redundant database downloaded from the NCBI FTP site. The algorithm Sequest uses cross-correlation (Xcorr) and delta correlation (dCn) functions to assess the quality of the match between tandem mass spectrum and amino acid sequence information in the database. The data were evaluated in terms of Xcorr magnitude (up to 1.7, 2.2, and 3.3 for, respectively, mono-, di-, and tricharged peptides) and dCn (higher than 0.1)

Alkaline-phosphatase treatment of oleosin

Oleosin (100 µg) isolated from *J. curcas* seed was resuspended in 100 µl of 0.05 M Tris-HCl, 5 mM $MgCl_2$, pH 8.5 buffer (Roche, UK) containing five units of Shrimp Alkaline phosphatase. The reaction was incubated at 37^{o}C for 30, 60, 120, and 90 min and inactivated by increasing the temperature to 65^{o}C for 10 min. Samples were then analysed by SDS-PAGE and immunoassay using anti-rJO3N antibody as described above.

Lipid analysis

Total lipid extraction, trans-esterification, and Gas Chromatography were performed as previously described by Popluechai et al., (2009). Thin Layer Chromatography (TLC) was performed using Merck silica gel 60 F 254 plastic sheets. The sheets were developed with hexane/diethyl ether/acetic acids (80:20:1) and then sprayed with a 5% solution of molybdophosphoric acid in ethanol. The sheets were visualised and quantified using Fluoro spectrophotometer (BioRad). Surface properties of *Jatropha* seed oil bodies were performed according to Wu et al., (1999). The investigations were included experiment of low pH aggregation and trypsin digestion (Wu et al., 1999).

Bioinformatics analysis

Nucleic acid and amino acid sequence analysis was performed by using CLUSTAL (Higgins et al, 1996) or MEGA4 (Tamura et al, 2007) programs analysis.

3.4 Results

Isolation and characterization of J. curcas oleosins at the gene level

Full-length cDNA fragments encoding three putative oleosins were isolated by a combination of PCR, RACE and gene walking strategies, initially using a degenerate forward primer (5'-AGYCCIRTIHTISTICCIGC-3') with Oligo dT as the reverse primer. The degenerate primer was designed after sequence alignment of multiple oleosins from *Ricinus communis*, *Helianthus annuus*, and *Arabidopsis thaliana* (see appendix 1A for detail). Gene sequences of the three oleosins, designated *JcOleosin1*, *JcOleosin2* and *JcOleosin3*, are available under accession numbers EU234462, EU234463 and EU234464 respectively at the NCBI site (see appendix 2 for nucleic acid alignment). Table I shows the salient features of the three *J. curcas* oleosin cDNAs and their translated products in terms of total nucleotide numbers per oleosin; number of nucleotides in the 5'- and 3'-untranslated region (UTR); number of nucleotides comprising the open reading frame (ORF); deduced amino acids (AA), their molecular weight (MW), isoelectric point (pI) and percent sequence identity to known oleosins from other plants.

Copy number of the three *J. curcas oleosins* was determined through Q-rtPCR using the *J. curcas* KAS III, a single copy gene in the *J. curcas* genome used for similar comparisons previously (Li et al., *20*08). The results shown in Table I as the Ct value for the ratio between oleosin copy number to KAS copy number (O/KAS) revealed that the three identified oleosins of *J. curcas* were each present as a single copy gene in the *J. curcas* genome.

JcOleosin3 shows similarity to the *Arabidopsis* oleosin which contains an intron. To determine whether any of the three *J. curcas* oleosins contain introns, PCR amplification was performed for the three genes using genomic DNA and primers terminally placed on the newly determined cDNA sequence. Genomic DNA amplification products revealed fragments of expected sizes for *JcOleosin1* and *JcOleosin2* but a larger fragment for *JcOleosin3* (see appendix 1B for detail). Sequencing of this fragment revealed an 87 bp intron. The location and sequence of this intron in relation to the cDNA sequence is shown in Figure 1. The intron did not exhibit the classical GU-AG concensus splice site (Brown 1986; Brown and Simpson, 1998). However, it was AU rich (AT 69.3 %) which is a characteristic reported to be an important factor in pre-mRNA splicing in plants (Goodall and Filipowicz, 1989)

Since oleosins are known to express preferentially in the seeds of oilseed plants, a fact that was confirmed for *J. curcas* oleosins as described in the next section, the upstream regulatory sequence of *J. curcas oleosins* was characterized. A promoter sequence of 623 nucleotides upstream of the *JcOleosin3* gene was isolated (EU543568) by PCR assisted gene walking and the sequence analyzed for *cis*-acting regulatory elements (CAREs) using the web-based PLACE program (http://www.dna.affrc.go.jp/PLACE/). Typical promoter motifs were identified for example, a TATA box (TATAAAT) and a CAAT box (CCAAT) was located at -27 and -46 bp upstream of the *JcOleosin3* transcription start site (TSS) respectively. Other motifs known for seed-specific gene expression were also noted. Transgenic *Arabidopsis* plants containing *gus* driven by the *JcOleosin3* promoter show seed-specific GUS expression. Confirmation of this result in homozygous plants and testing GUS expression during germination is still awaited. Details of promoter analysis as well as the validation of the isolated sequence as a seed-specific promoter in *Arabidopsis*, will be reported separately (Popluechai et al., see chapter 4).

Characterization of J. curcas oleosins at the transcript level

Expression of the three oleosins was observed and quantified in different tissues of *J. curcas* at different developmental stages using reverse transcriptase PCR (RT-PCR) and quantitative real-time PCR (qPCR) with the 18S ribosomal RNA gene as a comparative and technical control. Seed, leaf, stem, and root tissues were analyzed and the RT-PCR (see appendix 3A for detail) and qPCR results confirmed predominant expression of all three *Jatropha* oleosins in the seed. Expression of all three isoforms was also detected in the leaf at a very low level in comparison to expression in the seed, but no expression was detected in either the root or stem (Figure 2). In seed, *JcOleosin3* was expressed 4.9 and 4.6 fold higher compared to *JcOleosin2* and *JcOleosin1* respectively.

Previous studies on sesame (Tai et al., 2002), olive (Giannoulia et al., 2007) and coffee (Simkin et al., 2006) indicated oleosin expression as a function of fruit development. To determine the expression pattern of oleosins in relation to fruit development, seeds were obtained from fruits at 7, 21, 35, 56 and 70 days after fertilization (DAF) (see appendix 3B for gel electrophoresis of RT-PCR). Pollen from male flowers (10 days after first appearance of flower bud) was also analyzed since some oleosins are known to express in the pollen (Kim et al., 2002). Additionally,

diacylglycerol acyl transferase (DGAT; EU477378), a key enzyme for the synthesis of triglycerides that are encapsulated in the oilbodies formed by oleosins in the seeds, is expected to be active during seed development. This was included for transcript expression analysis. Semi-quantitative PCR results for a time course expression analysis revealed that *JcOleosin1* and *JcOleosin3* transcript level was high by 7-DAF and was maintained at that level, with minor variations, until fruit maturity i.e. 70-DAF (Figure 3a). However, the *JcOleosin2* transcript was not detected during the 7-DAF and 21-DAF stages while the later stages exhibited transcript levels comparable to the other two oleosins. Transcript accumulation of all three oleosins gradually increased during fruit development up to 56-DAF but decreased near to the stage of fruit maturity (70-DAF). The amount of transcript for *JcOleosin3* remained more stable over time compared to the other two oleosins. No transcript was detected for any of the three oleosin isoforms in the pollen. Since the pollen was not sampled from different developmental stages, it is not certain whether any expression of oleosin occurred in the pollen at any developmental stage. However, the *dgat* transcript was detected in pollen, although at a relatively low level, while in seeds its accumulation was high by 7-DAF and steadily decreased there after, yet it remained comparable to the level of oleosin transcript accumulation. Such a correspondence between the oleosin and the *dgat* transcript accumulation may be physiologically expected for the seed to have lipids for storage while simultaneously generating storage organelles.

Earlier studies on gene expression of oleosins in coffee beans during germination suggested a role for oleosins in controlling the coalescence of oilbodies, facilitating controlled TAGs utilization (Simkin et al., 2006). Therefore, the transcript level of the three oleosins during *J. curcas* seed germination was investigated using semi-quantitative PCR. The three oleosins briefly exhibited a slight increase in transcript levels at an early stage of germination (1 to 7 days after imbibition; DAI; Figure 3b). Later, the level of the transcripts significantly decreased until they disappeared at 14 DAI. However, *JcOleosin3* transcript was detectable at moderate levels until 21 DAI (see appendix 3C for gel electrophoresis of RT-PCR).

Characterization of J. curcas oleosins at the protein level

In silico analysis of the deduced amino acid sequences of the three *J. curcas* oleosins using protein family (Pfam) database tools for protein motifs/domains revealed all three to contain a hydrophobic region of 82, 83, and 79 amino acids for *JcOleosin 1, 2* and *3* respectively. The oleosin signature motif of the proline knot containing a serine residue in the centre (PXXXXXSPXXXP) was present in all three deduced proteins (Figure 4a). Hydrophobicity plots for each *J. curcas* oleosin clearly indicated the presence of a large central region with a negative value, corresponding to the hydrophobic region. Tzen et al., (1990) classified oleosins into H- and L-forms, depending on the presence of an 18 amino acid insertion in the C-terminal region. The protein sequence alignment of *J. curcas* oleosins with other oleosins previously identified as either H- or L-form revealed the 18 residue insertion in the C-terminal in *JcOleosin1* and *JcOleosin2* but not in *JcOleosin3* (Figure 4b). The *JcOleosin1* and *JcOleosin2* therefore represent H-form oleosins while *JcOleosin3* is an L-form oleosin. Phylogenetic tree analysis also supported these alignments (Figure 4c) (see appendix 4 for complete sequences alignment).

The *J. curcas* oleosins were purified as hydrophobic proteins recovered from oilbodies obtained after sucrose gradient and organic solvent extraction as described above. A chloroform/methanol ratio (11/7, v/v) yielded the best results in terms of extraction of a major protein band of approx 17 kDa visualized after SDS-PAGE and coomasie blue staining (see appendix 5 for detail). These conditions were therefore used for subsequent analysis of 2 separate accessions, one each from India (JI) and Indonesia (JIO). Bands visualized on SDS PAGE (Figure 5) were identified by LC-MS/MS which revealed that most of the proteins identified belonged to the sub-set of proteins closely associated with seed oilbodies, including caleosin and steroleosin (Table II, see appendix 6 for Mass spectrometry data obtained from LC-MS/MS). Oleosins were the major proteins in the two accessions and had apparent M_r of between 17 to 20 kDa. Densitometric scanning and image analysis of gels provided an approximate quantification indicating that oleosins represented up to 87% of oil body proteins. The toxic protein curcin, for which *J. curcas* is notorious in terms of non-edible nature of its seeds, was also identified in this set of proteins. Whether it is actually associated with the oilbodies is not certain, but it co-purified with most oilbody associated proteins. Potentially, the *J. curcas* homologue of *Arabidopsis* microspore oleosin (SM2) was also identified while a further protein similar to an *Arabidopsis* aquaporin (TIP3) was also

detected. Caleosin was predicted by homology with *Arabidopsis* and *Vitis vinifera* protein and a protein similar to *Coffea canephora* or *Ricinus communis* steroleosin was also present. Finally, a storage protein (60 kDa), identified from its homology with proteins from *R. communis* or *Quercus serrata* was also shown to be present.

Oleosins were not finely resolved under SDS-PAGE due to their close molecular weights. *J. curcas* oleosins were identified from multiple bands of varying sizes. For example, with reference to Figure 5 and Table II, in the Indonesian accession, *JcOleosin2* was identified in bands 1 and 2, while in the Indian accession *Joleosin2* was present in bands 1, 2, and 3. These results most likely indicate multimer breakdown products. The evidence for the existence of multimers is also provided by immunoassay as described below. Additionally, differential movement in the gel due to protein modification (*JcOleosin2* and *JcOleosin3* can be potentially phosphorylated; please see below) may have caused either the same oleosin to occur in different bands or different oleosins to occur in the same band.

The predicted AA sequence of the three *J. curcas oleosins* exhibited 45-69% identity to *Arabidopsis oleosins*. To test for the relationship between *Arabidopsis* oleosins and *J. curcas* oleosins, immunoassays were conducted using *Arabidopsis* oleosin S2 antibodies (anti-rS2N) on total seed protein extracted from 5 different accessions of *J. curcas* (2 from India and 1 each from Indonesia, Nigeria and Thailand). We observed cross reactivity for all accessions with the anti-rS2N antibody, with common bands at an apparent size around 13 kDa and 26 kDa, potentially corresponding to monomeric and dimeric forms of oleosins (Figure 6a). Bands possibly specific for some of the analyzed *J. curcas* accessions were also detected (Figure 6a) while the high molecular weight bands (> 30kDa) shown in Figure 6a represent nonspecific signals, detected with pre-immune sera (data not shown). To confirm the cross reactivity observed with anti-rS2N antibody on oleosins and to discriminate from non specific signals, immunoassays were also carried out on purified fractions, chloroform/methanol (11/7, v/v) and oilbodies protein extracts from JI and JIO accessions. We observed a cross reactivity with this *A. thaliana* antibody for all purified protein extracts with the presence of one strong band with an apparent size around 16 kDa (Figure 6b). However, when solvent extraction was used, a larger band at nearly 30 kDa was observed. This might be possibly as a result of dimerization of oleosin since oleosins has strong tendency to closely interact together *in vitro* and *in vivo* (Pons et al., 1998). Results suggesting that different extraction procedures effect oligomerization of oleosin *in vitro*. However, these experiments confirmed the presence of proteins,

certainly oleosins, with cross reactivity with *A. thaliana* antibodies in the oilbodies of *J. curcas* seeds.

To obtain conclusive evidence as to the identity of the larger-sized bands and to acquire specific information on *J. curcas* oleosins, a 15 AA peptide (EHPQSQHVGQQPR) unique to the amino-terminal of *JcOleosin3* was used as antigen to generate anti-*JcOleosin3* antibodies. Subsequent immunoassays revealed the expected specific band of approximately 16 kDa for the organic solvent extracted hydrophobic proteins (Figure 6c).

Since qPCR revealed *JcOleosin3* transcripts in *J. curcas* leaves (Figure 2) immunoassays were also performed on total leaf protein using anti-*JcOleosin3* antibody. However, no signal was detected indicating that the low level of *JcOleosin3* transcripts in leaves either gave rise to an undetectable amount of protein or no protein at all (see appendix 7 for detail). This result was important in terms of verifying that the *JcOleosin3* promoter sequence was seed-specific. Perhaps it is leaky at the transcript level but not at the protein level.

Oleosins are responsible for surface properties of J. curcas oil bodies

The levels and integrity of oleosins in the oilbodies are responsible for their size and distribution. Oilbodies isolated from the *J. curcas* seed were maintained as stable individual particles in a 0.1 M potassium phosphate buffer of pH 7.5 at 23^{o}C (Supplemental Figure 1a). Aggregation of *J. curcas* oilbodies could be induced by lowering the pH from 7.5 to 6.5 since this reduced the electronegative charge (Supplemental Figure 1b). Exposing the oilbodies to limited trypsin digestion breaks up the surface oleosins, again reducing electronegative charge and steric hindrance afforded by the oleosins, thus leading to coalescence of oilbodies (Supplemental Figure 1c). These results were similar to stability of oilbodies of other oil seed plants.

Lipid content in J. curcas oil bodies

Oilseed plants store Triacylglycerides (TAGs) as a food reserve for germination and to support growth of the seedling (Huang, 1992). To determine lipid composition in *Jatropha* oil bodies, TLC was performed. These results showed that *Jatropha* oil bodies, in common with other oilseed plants, predominantly composed of TAGs (see appendix 8 for detail).

Diversity of oleosins in Jatropha sp

The diversity of *oleosins* in *Jatropha* was investigated by analyzing 19 samples (Table III**)**, which represented three categories; different accessions of *J. curcas*, different species of *Jatropha* and different F1 lines from an inter-specific hybridization between *J. curcas* and *J. integerrima* (generated previously; Popluechai et al., *20*09). Diversity in the *oleosins* across the 19 samples was investigated by PCR and RT-PCR using primers either in the 5'- and 3'UTR so that the amplified product spanned most of the full length (FL) gene, or gene specific primers (GSP) that amplified the conserved central hydrophobic region. Amplification products of the expected size were obtained from genomic DNA template for the hydrophobic domain of all three oleosins in the 19 samples investigated. The *JcOleosin3* in all but one sample amplified a band diagnostic of the presence of the intron. This result confirmed the presence of an intron in the hydrophobic region of *JcOleosin3* across the samples and reiterated the conserved nature of this region. The result also confirmed that all three oleosins were present in all 19 samples. However, genomic DNA amplification using primers on 5'- and 3'UTR did not amplify a product in some samples for at least one of the three oleosins (Table III; value '0' under 'FL' for each oleosin). This was indicative of variation in the 5- and/or 3'UTR of the three oleosins across the samples

For *JcOleosin3* intron, an amplification product of 197 bp was expected from the genomic DNA template. However, all samples except hybrid clone 111 and the *J. curcas* accession from India exhibited two products, a 197 bp fragment and a 110 bp fragment. The hybrid clone 111 contained only the 110 bp product while the *J. curcas* accession from India contained only the 197 bp product. Sequencing of these two products from five different samples revealed that the 197 bp product indeed comprised the expected *JcOleosin3* sequence (exon1:intron:exon2), identical in all 5 samples. The 110 bp product was also identical in all 5 samples and comprised the expected

JcOleosin3 exon1:exon2 sequence without the intron sequence. Interestingly, the sample in which the *JcOleosin3* intron was initially detected and characterized was the same *J. curcas* accession from India that showed the single 197 bp band. Furthermore, the inter-specific hybrids were generated in India from local *J. curcas* (most likely one band: 197 bp) and *J. integerrima* (most likely two bands: 197 and 110 bp), hence the identification of one band (hybrid clone 111) or two bands (other samples) in the progeny of the cross. These results indicated the presence of two alleles of *JcOleosin3* in the genus *Jatropha*, distinguishable for the presence or the absence of an intervening intron, the position of which is strongly conserved.

To check for any variations in the intron, the 87 bp fragment was sequenced from *J. curcas* India, Non-Toxic Mexican (NTM), *Jatropha* hybrid clone numbers 222 and 223, *J. integerrima*, *J. maheshwarii, J. gossypifolia* and *J. podragrica*. DNA sequences were then aligned for comparison. The results showed no variation in the intron sequence occurring within the *J. curcas* species while single nucleotide polymorphisms, including 2 cases of a single base insertion at 12 different positions, were disclosed in the four other *Jatropha* species, *J. maheshwarii, J. gossypifolia* and *J. podragrica* (Figure 7).

The present study was also extended to investigate the presence of *JcOleosin3* gene in *J. podagrica* (Jp), *J. gysipifolia* (Jg), *J. mutifida* (Jm), *J. maheswarii* (Jm), *and J. interegima* (Ji) by PCR mediated amplification using JcFLO3F and JcFLO3R primers. The result showed 750 bp of *JcOleosin3* gene was present in *J. gysipifolia* (hereafter *JgossO3*), *J. maheswarii* (hereafter *JmahesO3*), *and J. interegima* (hereafter *JinterO3*). Cloning and sequencing analysis was performed. Deduced amino acids alignment of the *JgossO3, JmahesO3, and JinterO3* genes to the *JcOleosin3* showed more than 97 % identical. (see appendix 9 for detail). The *JcOleosin3* diversity was determined at the protein level using anti-*JcOleosin3* antibody. Organic solvent extracted proteins obtained from 18 samples were used in this study and compared with these present in the India accession. The results showed the presence of JcOleosin3 protein in all samples, indicating that *JcOleosin3* was conserved among *Jatropha spp* (see appendix 10 for detail).

Relationship of oleosin expression to oil yield and fatty acid profile

Results obtained above revealed variation in the oleosin genes present in different species investigated within the genus *Jatropha*. Such a variability encompassed samples with only one alleloform of *JcOleosin3,* with or without intron (JI and clone111 respectively); samples with both alleloforms in different genetic backgrounds of various accessions, species and hybrids; a sample containing *JcOleosin3* with 10 SNPs (*J. podagrica*) compared to the rest of the samples and those with differences in the 5'- and 3'UTRs. To determine whether the presence of such different organization and sequence of the oleosin genes could relate to differences in seed oil yield and FA profile, these two parameters were measured in all 19 *Jatropha* samples. Results of oil yield and FA profile did not correlate with the type of oleosin present. For example, Table III revealed that a similar level of oil yield (30% w/w) was obtained irrespective of whether both 110 bp and 197 bp alleloforms were present (S4) or only one alleloform of 110 bp was present (S10). Also, similar level of oil yield (39% w/w) was obtained in S3 and S19 that differed by 10 SNPs in *JcOleosin3*. Furthermore, species, accessions and hybrids similar in terms of the presence of both of the *JcOleosin3* alleloforms and similar with respect to 5'- and 3'- UTRs in *JcOleosin1* and *JcOleosin2*, exhibited a range of oil yield from 18 to 44%. Similarly the oleic acid to linoleic acid ratio was not related to the kind of oleosin present. These results indicate that there is no direct correlation between the type of oleosin present, the oil yield or the oil profile.

Table I. Salient features of *J. curcas* oleosins. [a]Total number of nucleotides. [b]Number of amino acids comprising the N-terminal + Hairpin + C-terminal part of the protein. UTR=Untranslated region. ORF = Open Reading Frame. AA = Amino acid. MW = Molecular weight. [1]Simkin et al, 2006; [2]Chua et al, 2008; [3]Salanoubat et al, 2000; [4]Eastmond 2004; [5]Not published; [6]van Rooijen et al, 1992. O/KAS = Ct value for the ratio between oleosin copy number to beta-ketoacyl synthase III (KAS) copy number. [c]Ratio calculated from real-time PCR efficiencies of each oleosin to KAS imply copy number of *Jatropha* oleosin gene. [d]Standard deviation calculated from triplicate data. [*]post-hairpin i.e. the intron occurs in the gene after the nucleotides coding for the proline knot mediated protein hair pin.

Gene	Total[a]	5'-UTR	ORF	3'-UTR	Intron	Total AA	AA NT+HP+CT[b]	MW (Da)	pI	AA similarity	% Identity	O/KAS Mean[c]	O/KAS SD[d]
JcOleosin 1	652	67	443	141	No	147	23+82+42	15,581	10.28	*C. arabica* (AY14574)	60[1]	1.02	0.19
										F. pumila (ABQ57396)	60[2]		
										A. thaliana At3g01570	45[3]		
JcOleosin 2	685	50	467	155	No	155	30+83+42	16,608	9.99	*R. communis* (AAR15171)	69[4]	1.03	0.13
										F. pumila (ABQ57396)	63[2]		
										A. thaliana At3g01570	55[3]		
JcOleosin 3	876	72	413	304	Yes[*]	137	17+79+40	14,499	9.65	*C. sinensis* (CAA88360)	78[5]	1.05	0.08
										R. communis (AAR15172)	75[4]		
										A. thaliana At4g25140	69[6]		

Table II. Identified proteins in oil bodies purified from *J. curcas* accession from Indonesia and India

Band	Identified protein	Apparent molecular mass (kDa)	Theoretical molecular mass (kDa)	NCBI accession number/A. thaliana hit	Peptide number	Coverage (%)
Indonesia						
1	*Jatropha* oleosin 3	18	14.40	ABW90150	3	10.9
	Jatropha oleosin 1		15.58	ABW90148	11	23.8
	Jatropha oleosin 2		16.60	ABW90149	14	31.6
2	*Jatropha* oleosin 2	19.5	16.60	ABW90149	16	40.0
3	*Arabidopsis* oleosin SM2	21	18.13	NP_188487/At3g18570	1	7.2
4	*Arabidopsis* α-TIP3	29	28.30	NP_177462/At1g73190	1	3.7
5	*Arabidopsis* ATS1	31	28.01	AAL36241/At4g26740	1	3.3
	Vitis vinifera caleosin		26.48	CAO71462	1	5.0
6	*Jatropha* curcin precursor	32	32.77	AAL86778	7	25.6
7	*R. communis* steroleosin	46	39.92	EEF34360	1	2.5
	Coffee steroleosin		40.05	AAX49394	1	3.1
8	*Quercus serrata* legumin	60	13.52	BAB12450	1	12.4
	R. communis protein storage		40.10	EEF35875	1	3.6
India						
1	*Jatropha* oleosin 3	18	14.40	ABW90150	2	10.9
	Jatropha oleosin 1		15.58	ABW90148	4	16.3
	Jatropha oleosin 2		16.60	ABW90149	3	18.7
2	*Jatropha* oleosin 2	19.5	16.60	ABW90149	3	13.5
3	*Jatropha* oleosin 2	21	16.60	ABW90149	4	18.1
4	*Jatropha* curcin precursor	32	32.77	AAL86778	6	24.6
5	*R. communis* storage protein	60	40.10	EEF35875	1	3.6

Table III. Relationship between fatty acid profile, oil yield and presence (1) or absence (0) of *oleosins* and their expression. Presence/absence was judged by PCR on genomic DNA (gDNA) and expression was judged by RT-PCR on cDNA. Primers amplified either a full length (FL) product wherein the 5'- and/or 3'- regions were included or a shorter gene-specific product (GSP) wherein the coding region could be amplified from the cDNA. Oil yield was estimated as mg oil per 100 mg of seed. The ratio O/L defines suitability as a source of biodiesel, with higher value being more suitable thus indicating high oleic and less linoleic acid content in the seed oil.

	Fatty acid profiles					**Oil yield (%w/w)**	**Oleosin1**		**Oleosin2**		**Oleosin3**			
							FL gDNA	**GSP cDNA**	**FL gDNA**	**GSP cDNA**	**FL gDNA**	**GSP primers**		
Sample	**Palmitic acid**	**Stearic acid**	**Oleic Acid (O)**	**Linoelic Acid (L)**	**O/L**							gDNA 110 bp	gDNA 197 bp	**cDNA**
Jg (S1)	8.93	10.01	25.02	56.04	0.45	33.5	1	1	1	1	1	1	1	1
Jm (S2)	16.20	8.06	30.71	45.03	0.68	20.6	0	1	0	1	1	1	1	1
Jp (S3)	8.87	6.44	16.78	67.91	0.25	39.69	1	1	1	1	1	1	1	1
Jc NTC (S4)	12.75	8.24	39.66	39.34	1.01	30.62	1	1	1	1	0	1	1	1
Jc Met (S5)	13.66	7.92	46.79	31.63	1.48	29.42	0	1	0	1	1	1	1	1
Jc Ugan (S6)	15.21	6.06	36.23	42.50	0.85	20.7	1	1	1	1	1	1	1	1
Hy 91-1 (S7)	10.22	9.79	48.96	31.03	1.58	27.89	1	1	0	1	0	1	1	1
Hy 92 (S8)	14.57	7.03	27.06	51.34	0.53	7.2	1	1	0	1	0	1	1	1
Hy 96 (S9)	11.55	9.89	20.07	58.50	0.34	4.62	1	1	1	1	1	1	1	1
Hy 111 (S10)	8.70	9.11	44.71	37.48	1.19	31.55	0	1	0	1	0	1	0	1
Hy 201 (S11)	10.57	6.50	41.12	41.80	0.98	28.3	1	0	1	1	1	1	1	1
Hy 202 (S12)	8.65	15.27	50.27	25.81	1.95	33.4	1	1	0	1	0	1	1	1
Hy 205 (S13)	15.02	8.62	47.76	28.60	1.67	24.29	1	1	1	1	1	1	1	1
Hy 213 (S14)	12.78	7.88	41.99	37.35	1.12	33.2	1	1	1	1	1	1	1	1
Hy 222 (S15)	11.76	7.86	38.24	42.14	0.91	44.62	0	1	1	1	0	1	1	1
Hy 223 (S16)	12.58	8.95	39.38	39.09	1.01	34.6	1	1	0	1	0	1	1	1
Hy 225 (S17)	12.78	10.12	48.09	29.01	1.66	18.29	0	1	0	1	0	1	1	1
Hy 226 (S18)	11.81	7.54	40.33	40.33	1.00	28.52	1	1	0	1	0	1	1	1
Jc India	11.6	6.6	48.1	32.1	1.49	38	1	1	1	1	1	0	1	1

```
DNA-JcOleosin3      GCCTCAGGCGGTTTTGGGGTCGCTGCCATTTTTGTCTTGTTTTGGATTTATAGGTTTAAA
mRNA-JcOleosin3     GCCTCAGGCGGTTTTGGGGTCGCTGCCATTTTTGTCTTGTTTTGGATTTATAGGT-----

DNA-JcOleosin3      GAAATTTGCCTACTTCCTTTCATTTCTTTCAACGCATAATATTAAAAAGACAATGTTTGA
mRNA-JcOleosin3     ------------------------------------------------------------

DNA-JcOleosin3      CATTGGAGACATGAATTGCAGGTATGTTACTGGAAAGCATCCCCCAGGAGCAGAAAATCT
mRNA-JcOleosin3     ----------------------ATGTTACTGGAAAGCATCCCCCAGGAGCAGAAAATCT
```

Figure 1 Partial sequence alignment of gDNA and cDNA of *JcOleosin 3* showing the position and sequence of the intron in bold letters.

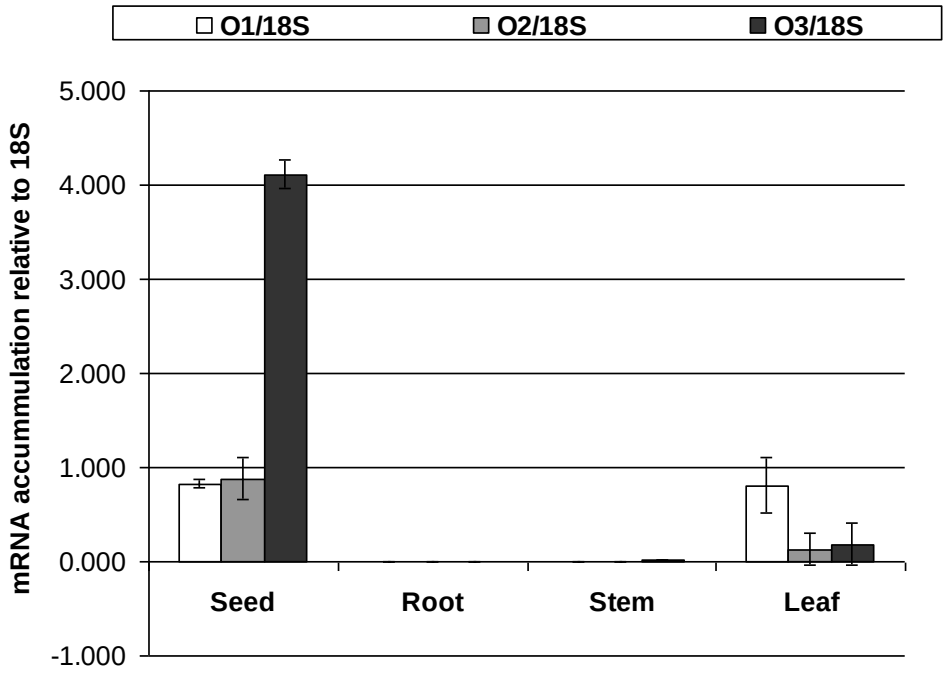
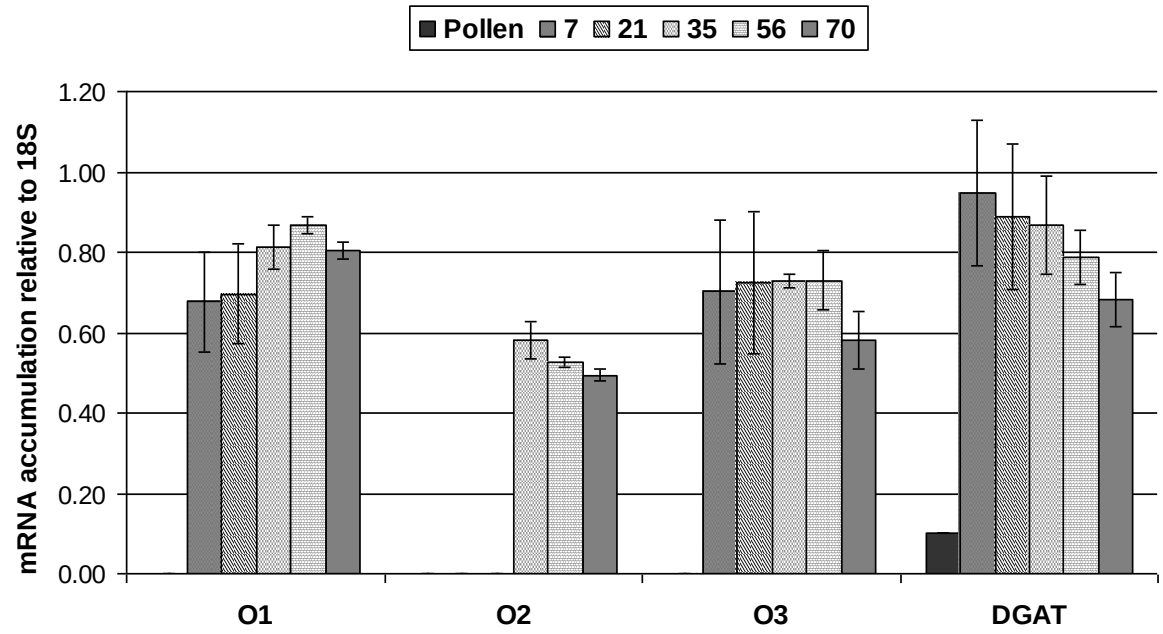


Figure 2 Expression of *J. curcas oleosins* in different tissues using quantitative real-time PCR (qPCR). O1, O2 and O3 represent *JcOleosin1*, *JcOleosin2* and *JcOleosin3* respectively. The 18S ribosomal RNA expression was used as the internal positive and comparative control for quantification.

(a)



(b)

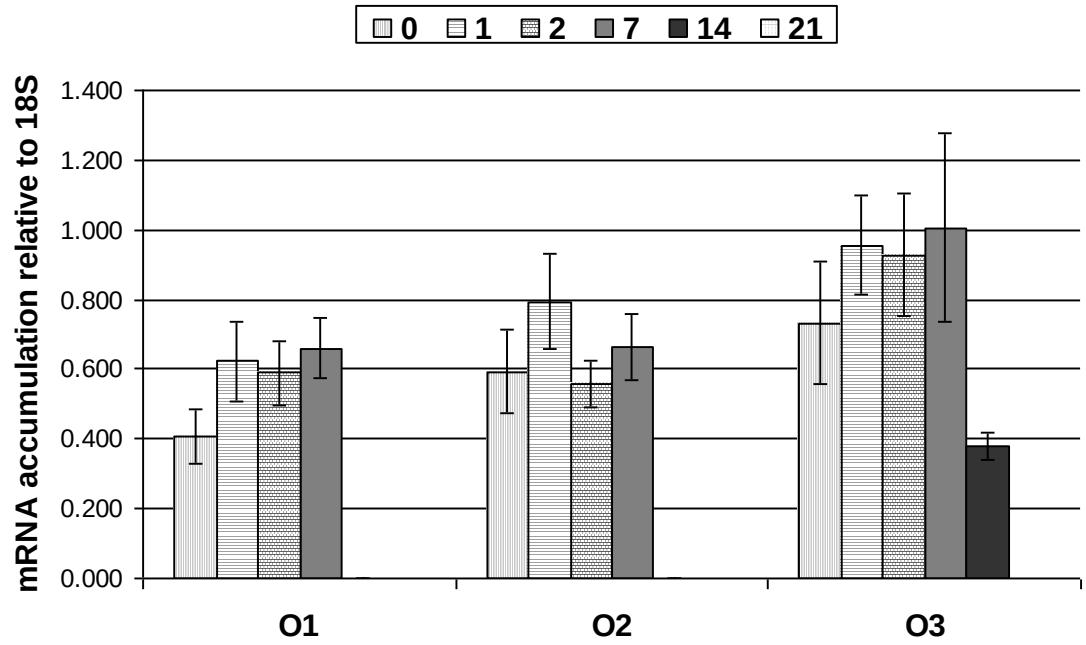


Figure 3 (a) Accumulation of *oleosin* and *dgat* mRNA during *J. curcas* fruit development. Numbers indicate days after fertilization (DAF). (b) Accumulation of mRNA of *oleosins* during *J. curcas* seedling development. Numbers indicate days after initiation (DAI) implying the time when the seeds were first soaked in water. O1, O2 and O3 represent *JcOleosin1*, *JcOleosin2* and *JcOleosin3* respectively.

(a)

```
                                          Proline knot
RicinusH        LAVLTFLPVGGGLLSLSGITLTNTLIGMAIATPLFILFGPIILPAAVVIGLAMMAFMVAG 93
JatrophaH2      LAVITLLPVGGGLLALAGITLVGTLIGLAITTPLFVIFSPVLVPAALVIGLSVMAFLASG 94
OeOLEH          LAVVTLLPVGGTLLALAGLTLVESLIGLAVTTPLFIIFSPVLVPATILVGLAVTAFLTSG 100
FicusH          LAVLTLLPVGGTLLFLAGLTFIGTLIGLALSTPVFILFSPVLVPAAITIGLAITGFLTSG 94
AtS4            LSLLIGVPVVGSLLALAGLLLAGSVIGLMVAIPLFLLFSPVIVPAALTIGLAMTGFLASG 114
AtS2            LALIAGVPIGGTLLTLAGLTLAGSVIGLLVSIPLFLLFSPVIVPAALTIGLAVTGILASG 112
CofOLE1H        LAVVTLLPVAGVLLGLSGLILVGTVIGLAVTTPLFVIFSPILVPAVFTLGLTLAGFLTSG 89
JatrophaH1      LAVVTLLPVSGTLLFLAGITLTGTLIALAVATPLFVICSPVRGPAALVIGLSVLGFLTSG 88
RicinusL        VKAGTAATAGSSLLFLSGLTLTGTVIALALATPLMVLFSPVLLPAVIIISLIGAGFLTSG 79
JatrophaL       VKAATAVTAGGSLLVLSGLTLAGTVIALTVPTPLLVIFSPVMVPTVITVCLIITGFLASG 79
FicusL          VKAATAVTAGGSLLVLSALILAGTVIALTIATPLFVIFSPVLVPAVITLGLIIIGFLASG 91
CofOLE2L        VKAATAVTAGGSLLVLSGLILAGTVIALALATPLLVIFSPVLVPAGITVFLLVTGFLSSG 74
AtS3            AKAATAVTAGGSLLVLSSLTLVGTVIALTVATPLLVIFSPILVPALITVALLITGFLSSG 103
CofOLE4L        VKTTTAVAVGGSLMLLSGLTLAGTIIGLVLATPLLVIFSPVIVPAAVTFFLILAGFFISG 77
                       .  . *: *:.: :  ::*.: :. *:::: .*:  *: . . *   .:: :*
```

(b)

```
                                H-form insertion
RicinusH        AKKRMQDMAGYVGMKTKEVGQDIQRKAQEGK
JatrophaH2      AKKRMQDMAGFVGQKTKEVGQEIQRKAHEGK
FicusH          AKRRVQDMAGYTGQKAKEVGQEVQSKGQEGK
AtS2            AKRRMADAVGYAGMKGKEMGQYVQDKAHEAR
RiceH           ARRRMAEAAAHAGHKTAQAAHAIQGRADQAG
JatrophaH1      AKRRAQETTGQLGQKAREVGQKAQEVAKT--
RicinusL        ARLKLAG------------------KAREMK
JatrophaL       ARLKLAG------------------KAREMK
FicusL          ARHKLAS------------------KAREMK
AtS5            ARMKVGS------------------RVQDTK
RiceL           AKARLAS------------------KARDIK
```

(c)

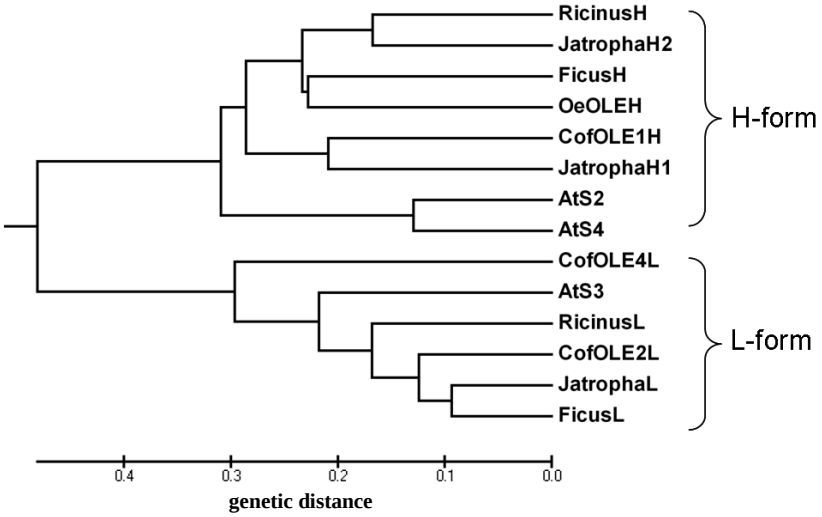


Figure 4 (a) Amino acid sequence alignment of *J. curcas* oleosins with those known from other plants to show the conserved sequence from the proline knot. (b) Optimal alignment of the protein region containing the 18 residue insertion at the C-terminal distinguishing the H- and L-form oleosins. (c) Phylogenetic tree generated from the multiple oleosin sequence alignment clearly distinguishes the L- and the H-form oleosins. JatrophaH1 is JcOleosin1, JatrophaH2 is JcOleosin2, and JatrophaL is JcOleosin3.

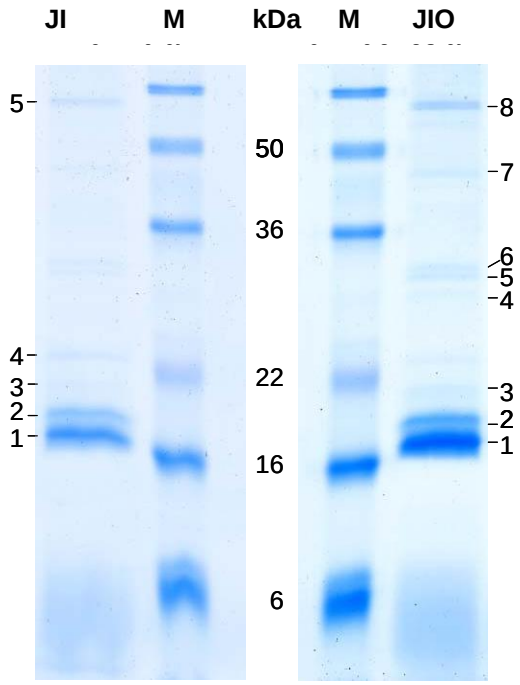


Figure 5 Protein profiles of chloroform/methanol (11/7) seed extract of *J. curcas* accessions from India (JI) and Indonesia (JIO). Numbers indicate the bands that were analyzed with LC-MS/MS and described in Table II.

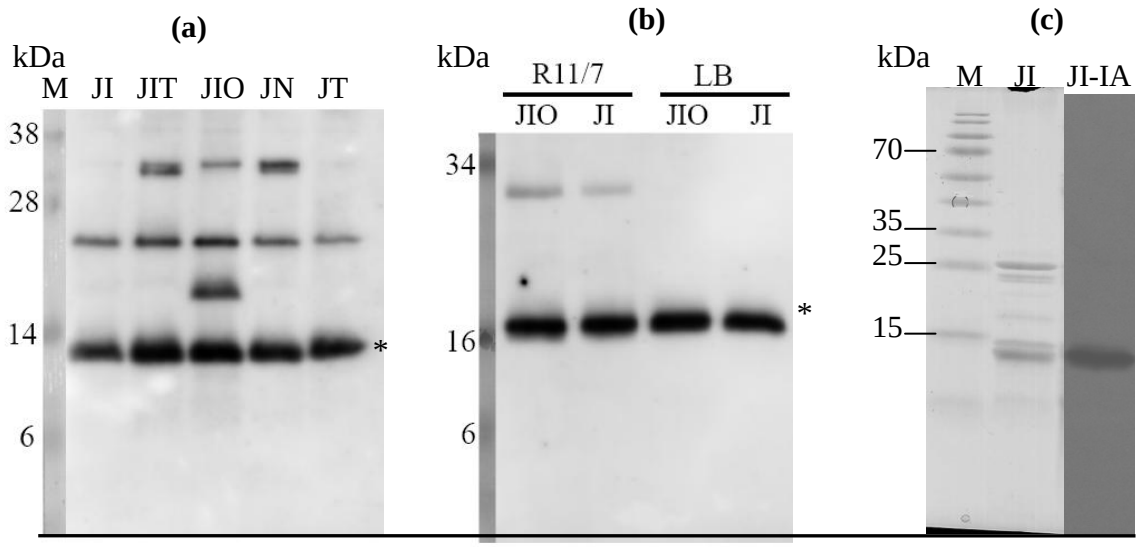


Figure 6 Immunological detection of oleosins using antiserum raised against *A. thaliana* oleosin S2 (anti rS2N). (a) Total seed protein of *J. curcas* extracted from accessions belonging to India (JI and JIT), Indonesia (JIO), Nigeria (JN) and Thailand (JT). (b) Seed protein of JIO and JI extracted with methanol/choloroform (11/7) or from lipid bodies (LB). (c) Immunological detection of *JcOleosin*3 using antiserum against *JcOleosin*3 on seed protein extract of JI using chloroform/methanol (11/7). The panel on the left with multiple bands is the coomasie-blue stained protein gel while the one on the left shows the immunoassay (JI-IA) result. M represents molecular weight (kDa) marker. Please see detail of gels and protein marker used in this chapter in Material and Methods

```
JcO3InIndia      -TTAAAGAAATTTGCCTACTTCCTTTCATTTCTTTCAACGCATAATATTAAAAAGACAAT 59
JcO3InNTMA       -TTAAAGAAATTTGCCTACTTCCTTTCATTTCTTTCAACGCATAATATTAAAAAGACAAT 59
JcO3InHy222      -TTAAAGAAATTTGCCTACTTCCTTTCATTTCTTTCAACGCATAATATTAAAAAGACAAT 59
Jc3InHy223       -TTAAAGAAATTTGCCTACTTCCTTTCATTTCTTTCAACGCATAATATTAAAAAGACAAT 59
Jint-O3In        -TAAAAGAAATTTGCCTACTTCCTTTAATTTCTTTCAACACATAATATTAAAAAGACAAT 59
Jmh-O3In         -TCAAAGAAATTTGCCTACTTCCTTTCATTTCTTTCAACACATAATATTAAAAAGACAAT 59
Jpod-O3In        TAAAAAGAAATTTGCCTACTTTCTTTCATTTCTTTTAACGCATAGTATTAAAAAGACAAT 60
Jgos-O3In        -AAAAAGAAATTTGCCTACTTTCTTTCATTTCTTTTAACGCATAATATTAAAAAGACAAT 59
                    ****************** **** ******** *** **** ***************

JcO3InIndia      GTTTGACATT-GGAGACATGAATTGCAGGT 88
JcO3InNTMA       GTTTGACATT-GGAGACATGAATTGCAGGT 88
JcO3InHy222      GTTTGACATT-GGAGACATGAATTGCAGGT 88
Jc3InHy223       GTTTGACATT-GGAGACATGAATTGCAGGT 88
Jint-O3In        GTTTGACATG-GGAGACATGAATTGCAGGT 88
Jmh-O3In         GTTTGACATT-TGAGACATGAATTGCAGGT 88
Jpod-O3In        GTTTGACGTT-GGAGACACGAATTGCAGGT 89
Jgos-O3In        GTTTGACATTTGGAGACACGAATTGCAGGT 89
                 ******* *   ****** ***********
```

Figure 7 Alignment of the *JcOleosin3* intron sequences obtained from various *Jatropha* accessions. The shaded bases indicate the SNPs

3.5 Discussion

J. curcas oleosin genes

The three *oleosin*s, *JcOlesoin1, JcOlesoin2* and *JcOlesoin3,* isolated from *J. curcas,* irrespective of accession or origin, were found to be comparable in size with known *oleosins* from other plants. The ORF for each of the three genes (Table I) varied in size, as previously reported for oleosins from *Arabidopsis thaliana* (Kim et al., 2002), *Coffea canephora* (Simkin et al., 2006), *Ricinus communis* (Eastmond, 2004), *Ficus pumila* (Chua et al., 2008), and *Olea europaea* L. (Giannoulia et al., 2007) which all show ORF variation from 558 to 908 nucleotides. In *J. curcas*, ORF of *Jcolesoin3* (L-form) is longer than *Jcolesoin1* and *Jcolesoin2* (H-form). The ORF of *Jcolesoin2* was 72 % identical to H-form-*Ricinus* oleosin1 and 73 % identical to L-form-*Ricinus* oleosin2 (Supplemental Table II) suggesting a close relationship between *Jatropha* and *Ricinus* both of which are members of the *Euphorbiaceae*. Supplemental Table II also shows the relationship between the 5'and 3' untranslated regions (5'UTR; 3'UTR) of oleosins from different plants. The UTRs play an important role in post-transcriptional control of mRNA such as translational control, subcellular localization and mRNA stability (Pesole et al., 2001). 5'UTRs are generally shorter than 3'UTRs with the length of the former ranging between 100-200 nucleotides compared to between 200 to more than 1000 nucleotides for 3'UTRs, depending on the organism. The 3'UTRs exhibit higher variation due to greater length (Pesole et al., 2001). The three *J. curcas* oleosins also showed longer 3'UTRs (Table I). Similar findings were reported for *A. thaliana* (Kim et al., 2002), *C. canephora* (Simkin et al., 2006), *Ricinus communis* (Eastmond, 2004), *Ficus pumila* (Chua et al., 2008) and *Olea europaea* (Giannoulia et al., 2007; Supplemental Table II). The conserved hydrophobic region of *oleosins* ranges from between 200-250 nucleotides and all three *J. curcas oleosins* contain nucleotides within this range similar to *A. thaliana* (Kim et al., *20*02) and *C. canephora* (Simkin et al., *20*06).

Results from quantitative PCR indicated that the three *J. curcas oleosin*s existed as single copy genes. Previous studies also reported that specific *oleosin*s in sesame (Tai et al., 2002), and coffee (Simkin et al., 2006) exist as single copy genes. Repeated efforts to obtain this information through Southern hybridizations, by addressing various parameters, were not successful except with *JcOleosin2* which did reveal a single copy character on a poor quality autoradiogram. This may be due to the presence

of phenolics and complex carbohydrates co-extracted with the DNA, which interfere with efficient restriction digestion.

Using PCR amplification on genomic DNA of seeds of *J. curcas* accession from India, gene structure of each isoform was characterised. The results showed that *JcOleosin3* contained one intron of 87 nucleotides immediately following the sequence encoding the protein hairpin stretch. In *Arabidopsis*, 13 out of 16 *oleosins* have one intron preceding or immediately following the sequences encoding the hairpin stretch (Kim et al., 2002). The importance of having no intron or an intron before or after the sequence encoding the hairpin stretch remains to be seen. The role of introns in gene expression, mRNA stability and/or translation processes has been well documented (Rose and Last, 1997; Morello and Breviario, 2008).

A sequence of 623 nucleotides upstream of the *JcOleosin3* gene was isolated (EU543568) from the genome of *J. curcas* accession from India (JI). *In silico* analysis using PLACE programme indicated seed specific motifs such as DPBFCORE, SEF4, GAT-box, E-box, Martbox, and RY repeat. These motifs have previously been reported in oleosin promoters from other plants such as *Perilla frutescens* (Chung et al., 2008), *Coffee canephora* (Simkin et al., 2006) and *Brassica napus* (Keddie et al., 1994). Recently the *J. curcas* curcin gene promoter was reported to contain similar seed-specific motifs such as the DPBFCORE and E-box. Seed-specific expression of this promoter was validated in transgenic tobacco plants (Qin et al., 2008). Our analysis of the oilbody proteome also reveals the presence of curcin (Table II) in the seeds. Our results on testing the *JcOleosin3* promoter in transgenic *Arabidopsis* will be reported separately.

J. curcas oleosin transcripts

Expression analysis through transcript detection for the three oleosin isoforms in seeds, leaves, roots and stem revealed high expression of all three isoforms in the seeds, with very low expression in leaves (Figure 2). Roots and stems showed no expression of these isoforms. Similarly, in *Coffea arabica* the expression of all isoforms was found to be very high in seeds when compared to expression in leaves. In leaves, where low expression was detected, one of the high molecular weight isoforms, *JcOleosin1* was found to be relatively more abundant. This expression was found to be similar in *Coffea arabica* which showed expression of the high molecular weight isoform CcOLE-4 (AY841274) in leaves (Simkin et al., 2006) and no expression of low molecular weight

isoforms. The expression of the low molecular weight isoform is found to be high in seeds of *J. curcas* in agreement with expression of the low molecular weight isoform Oleosin S3 in *Arabidopsis thaliana*, which accounts for approximately 65% of oil-body associated proteins (Jolivet et al., 2004; Siloto et al., 2006).

The pattern of transcript accumulation for *JcOleosin1, 2,* and *3* during fruit development (Figure 3a) was similar to that of coffee oleosin (*CcOLE-2* to *CcOLE-4*; Simkin et al., 2006). It was also similar to the expression patterns of seed *oleosins* (S1, S2, and S3) during silique maturity in *Arabidopsis* (Kim et al., 2002). The pattern of *oleosin* expression in *J. curcas* was also similar to the H-form *oleosin* in olive (*OeOLE*) which exhibits a bell shaped pattern of expression during fruit development (Giannoulia et al., *20*07). Recent studies by Annaroa et al. (2008) demonstrated that lipid accumulation in *Jatropha* seed during fruit development was found after approximately 3 weeks after fertilization reaching maximal levels approximately 5 weeks after fertilization. This evidence strongly correlated with the expression profiles of *oleosins* in *J. curcas* which also reached its maximum at the same period (Figure 3a). Similar expression patterns were found for diacylglycerol-acetyl-CoA acetyltransferase **(***dgat)* in olive mesocarp tissue by Giannoulia et al. (2000) where *dgat* gene expression was high in the early and mid stages of mesocarp growth, a period of triacylglycerols (TAGs) accumulation. As growth proceeded, transcript level dropped. In *Jatropha* seed, active TAGs synthesis around the 4^{th} week of fruit development reported by Annarao et al. (2008) indicated a correlation between the *oleosins*' transcript and TAGs accumulation.

Expression analysis revealed that none of the *J. curcas oleosin* transcripts were detected in the pollen in contrast to *Arabidopsis* (Kim et al., 2002). Evidently, the three isoforms of *J. curcas* oleosin are seed-specific. Although transcripts for oleosins were detected in leaves, no protein could be detected, as discussed below. Additionally, during *J. curcas* seed germination, *JcOleosin1*, *2*, and *3* briefly exhibited a slight increase in transcript level at the early stage (1-7 days after initiation; DAI; Figure 3b) of germination and disappeared after 14 DAI. These results were similar to *oleosin* expression profiles in coffee during germination (Simkin et al., 2006). However the transcript for *JcOleosin3* (L-form) remained detectable until 21 DAI. Interestingly, the coffee L-form oleosin also exhibited the same pattern during seed germination. Froese et al. (2003) and Simkin et al. (2006) suggested that such a pattern of expression may be required to prevent coalescence of oilbodies during TAGs utilization. In this respect *JcOleosin3* may have a similar role, although this is yet to be confirmed.

The deduced amino acid sequences of the three *J. curcas* oleosins contained the oleosin-specific central hydrophobic region with conserved proline knot and a conserved serine at its centre. The hydrophobicity profiles of the three oleosins were similar to previously published profiles for several seed specific oleosins (Huang, 1992). Oleosins belong to either the H-form (high molecular weight) or L-form (low molecular weight) depending on their molecular weight. Previous studies by Tai et al (2002) reported that the difference in molecular mass between H- and L-oleosin in several plant species was 2 kDa, accounted for by the addition of 18 amino acid residues (~2 kDa) at the C-terminal domain of the H-oleosins. The difference in the number of amino acids between either *JcOleosin1* or *JcOleosin2* with *JcOleosin3* was 18 residues (~2 kDa), confirming that *JcOleosin1* and *JcOleosin2* must be considered H-oleosins while *JcOleosin3* belongs to the class of L-oleosin.

In silico analysis of the deduced amino acid sequence of the three *J. curcas* oleosins using Motif Scan (SwissProt) revealed 4 potential phosphorylation sites at the carboxyl terminal (97-99:TGK) of *JcOleosin3*. Holbrook et al. (1991) demonstrated that a 32 kDa oilbody protein of *Brassica napus* was phosphorylated. In the present study immunoassays with alkaline phosphatase-treated protein extracts revealed a single band indicating a lack of phosphorylation in *JcOleosin3,* in contrast to the *in silico* prediction (Supplemental Figure 2). Interestingly the *JcOleosin3* immunoassay (lane JI-IA) shown in Figure 6c does not show any multimers of the oleosins. This may be a cultivar specific trait because Supplemental Figure 3 for a similar immunoassay with the anti-*JcOleosin3* antibody shows two bands in some cultivars, one of which could most likely be the dimer, while a single band is seen in other cultivars. Oligomers of oleosins have been previously reported for peanut oleosin (Pons et al., 1998) and *Arabidopsis* oleosins (Roux et al., 2004). For unclear reason, the different in size of oleosins detected in Figure 5, 6a and 6b can not be explained properly [see Figure 5 band number1, Figure 6 indicated as *]. It may be due to different migration rate of oleosins in the different buffer systems [Tris-Glycine buffer for gel of Figure 5, MES buffer for gel of Figure 6a, and Tricine buffer for gel of Figure 6b, respectively]. However, the different in size of oleosins might be resulted in molecular weight estimation base on the different apparent molecular weight values of prestained protein marker which was run on the different gel buffer systems (see above). The different apparent molecular weight values of

prestained protein marker have been observed that when run on different gel buffer systems (URL 1).

Mass-spectrometry based protein identification of samples separated by SDS-PAGE revealed the presence of several proteins in addition to oleosins (Table II). According to the normalized spectral abundance factor (NSAF; Zybailov et al., 2006) calculations which consider the number of spectral counts for each protein divided by the protein length and divided by the sum of all SAFs for proteins in the sample, it appeared that *JcOleosin2* was the most abundant oleosin. The *J. curcas* accession from Indonesia consisted of 66% *JcOleosin2*, 25% *JcOleosin1*, 7% *JcOleosin3* and 2% SM2 oleosin ortholog, and the accession from India consisted of 61%, 25% 14% of *JcOleosin2, -1, -3* respectively). These results contradict the comparatively higher amounts of *JcOleosin3* observed during transcript analysis (Figure 2 and 3). However, it is known that transcript and protein abundance are not always directly correlated and that translational efficiencies vary widely within related genes. It is possible that the intron and/or the UTRs of *JcOleosin3* have a role in higher transcription and/or lower transcript turn over rate of the gene, but a lower translation efficiency of its mRNA. On the contrary, the transcript analysis results for *JcOleosin3* in the present study may be valid if its protein turn over rate is comparatively high. These aspects and their impact on oilbody quality and quantity need to be addressed in future research.

Although, oleosins were the major proteins present in *J. curcas* oilbodies, curcin was also recovered in oilbodies of both the Indonesian and Indian accession with a total coverage of 36% for the precursor sequence. This finding is in agreement with those of Mourgue et al. (1961) who suggested that curcin was predominantly expressed in the seeds. In addition to oleosins and curcin, other minor proteins were identified by similarity with plant proteins such as the *A. thaliana* SM2 oleosin; an aquaporin TIP3:1 or TIP3:2; and caleosin and steroleosin-like proteins. SM2 oleosin was expressed in both pollen and maturing seed (Kim et al., 2002). Caleosin and steroleosin are known to be associated with the oilbodies of several oil seed plant species. Caleosins posses a single conserved EF-hand calcium-binding domain and an oleosin like association with lipid bodies. It is postulated that the function of caleosin is modulated by Ca^{2+}- binding and phosphorylation in the cell, as seen in signal transduction proteins, and may play an important role in lipid trafficking process (Murphy et al., 2001). In this study caleosins were found to persist through seed desiccation and dormancy up to 6 days post-germinative seedling development. Steroleosins are present in most plants and derive their name from the presence of a steroid binding site (Lin et al., 2002). Those in

sesame seeds have been classified as hydroxysteroids dehydrogenases belonging to a very large group of NAD(P)-dependent oxidoreductases (Lin and Tzen, 2004; D'Andrea 2007a). They also possess a hydrophobic anchor (proline knot) located at the N-terminal end of the protein, as well as a sterol and a nucleotide binding sites. The role of stabilization of oilbodies by the presence of oleosins and caleosin was demonstrated by construction of artificial oilbodies (Chen et al., 2004). Furthermore, it is known that in *Arabidopsis* the amount of oleosins present on the surface of the oil body determines the size of the oilbody (Siloto et al 2006) and that the absence of these proteins in the cell lines resulted in irregular and disrupted oilbodies. Aquaporin was also found as minor protein in *J. curcas* oilbodies. The role of aquaporin in oilbodies is still unclear; however it is believed to play a role in glycerol and/or water transportation (Weig and Jacob, 2000). Evidently, the protein component of *J. curcas* oilbodies is rather similar to that of *A. thaliana* (Jolivet et al., 2004) and *Brassica napus* (Katavic et al., 2006; Jolivet et al., 2009) which, like *J. curcas*, all contain oleosins, caleosin, steroleosin and aquaporin along with some contaminant storage proteins.

The question as to whether an increase or decrease in oleosin levels of one type or another in *J. curcas* can affect oil yield or oil profile remains unclear. However, in *Arabidopsis* it has been shown that down-regulation of the 18 kDa oleosin resulted in larger oilbodies, decreased seed oil and minor changes in FA profile (Siloto et al., 2006).

J. curcas oleosin diversity

Investigation of the diversity of *J. curcas oleosins* revealed that the size of the PCR amplification product that spans the hydrophobic region was conserved between different oleosins. However, variations among the accessions and species in the 5' and 3'UTRs were noticed. Li et al. (2004) proposed that variation in 5'UTRs could regulate gene expression by affecting transcription and translation whereas variation in 3'UTRs may primarily affect RNA accumulation. PCR-based investigations for *JcOleosin3* intron diversity across *Jatropha* accessions, species and hybrids revealed the expected band of 197 bp, inclusive of an 87 bp long intron, as in the accession from India originally used to isolate and sequence the intron. However, all seed samples except those of hybrid clone 111 showed 2 fragments, a 197 bp fragment and a 110 bp fragment. Sequence analysis revealed similarity to *JcOleosin3* for both fragments with the only difference being in the absence of 87 nucleotides of the intron sequence in the

110 bp fragment. Alignment of the intron sequence isolated from different *Jatropha sp.* showed 12 SNPs but otherwise the intron sequence was largely conserved across *Jatropha* accessions and species (Figure 7). The presence of an intron at a conserved position of the coding sequence might suggest a role in gene expression that is currently under investigation. Additional diversity analysis identified various accessions and species that contain the two alleles (with and without intron) of *JcOleosin3*. This may be interpreted as the result of a differential rate of intron evolution, loss or gain, occurring at the level of some allele-specific locus and can also represent a useful tool for the characterization of the products resulting from cross-genetic hybridizations.

Acknowledgements

SP gratefully acknowledges the Royal Thai government for a graduate scholarship. We thank Prapassorn D. Eungwanichayapant, MaeFah Lunag University, Thailand and Manish Raorane and Udhaya Kannan, Newcastle University, UK for technical assistance in some of the experiments.

3.6 References:

Achten W.M.J., Mathijs E., Verchot L., Singh V.P., Aerts R., Muys B., *Jatropha* Biodiesel fueling sustainability?, Biofuels Bioprod. Bioref. 1 (2007) 283-291.

Annarao S., Sidhu O.P., Roy R., Tuli R., Khetrapal C.L., Lipid profiling of developing *Jatropha curcas* L. seeds using H-1 NMR spectroscopy, Bioresource Technology 99 (2008) 9032-9035.

Brown J.W.S., A catalogue of splice junction and putative branchpoint sequences from plant introns. Nucleic Acids Res. 14 (1986) 9549–5959

Brown J.W.S., Simpson C.G., Spice site selection in plant pre-mRNA splicing, Annu. Rev. Plant Physiol. Plant Mol. Biol. 49 (1998) 77–95.

Capuano F., Beaudoin F., Napier J.A., Shewry P.R., Properties and exploitation of oelosins, Biotechnology Advances 25 (2007) 203-206.

Chen M.C.M., Chyan C.L., Lee T.T.T., Huang S.H., Tzen J.T.C., Constitution of stable oil artificial oilbodies with triacylglycerol, phospholipids and caleosin, J. Agric. Food Chem. 52 (2004) 3982-3987.

Chua A.C., Jiang P., Shi L., Chou W., Tzen J.T.C., Characterization of oil bodies in jelly fig achenes, Plant physiology and biochemistry 46 (2008) 525-532.

Chung K., Hwang S., Hahn B., Kim K., Kim J., Kim Y., Yang J., Ha S., Authentic seed-specific activity of the *Perilla* oleosin 19 gene promoter in transgenic *Arabidopsis*, Plant Cell Reports 27 (2008) 29-37.

D'Andréa S., Canonge M., Beopoulos A., Jolivet P., Hartmann M.A., Miquel M., Lepiniec L., Chardot T., At5g50600 encodes a member of the short-chain dehydrogenase reductase superfamily with 11 beta- and 17 beta-hydroxy steroid dehydrogenase activities associated with Arabidopsis thaliana seed oil bodies, Biochimie 89 (2007a) 222-229.

D'Andrea S., Jolivet P., Boulard C., Larré C., Froissard M., Chardot T., Selective one-step extraction of *Arabidopsis thaliana* seed oleosins using organic solvents, Journal of Agricultural and Food Chemistry 55 (2007b) 10008-10015.

Dunbar A., Pollen development in the *Eleocharis palustris* group (Cyperaceae). I. Ultrastructure and ontogeny, Bot Not 126 (1973) 197–254

Eastmond P.J., Cloning and characterization of the acid lipase from castor beans, J. Biol. Chem. 279 (2004) 45540-45545.

Froese C.D, Nowack L., Cholewa E., Thompson J.E., Molecular composition and surface properties of storage lipid particles in wax bean (*Phaseolus vulgaris*), Journal of Plant Physiology 160 (2003) 215–225.

Gerbens-Leenes W., Hoekstra A.Y., van der Meer T.H., The water footprint of bioenergy, Proc Natl Acad Sci USA 10 (2009) 10219–10223.

Giannoulia K., Banilas K., Hatzopoulos P., Oleosin gene expression in olive, Journal of Plant Physiology 164 (2007) 104—107.

Giannoulia K., Haralampidis K., Poghosyan Z., Murphy D.J., Hatzopoulos P., Differential expression of diacylglycerol acyltransferase (DGAT) genes in olive tissues, Biochemical Society Transactions, 28 (2000) 695-697.

Goodall G.J., FilipowiczW., The AU rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell 58 (1989) 473–483.

Holbrook H.A., van Rooijen G.J.H., Wilen R.W., Moloney M.M., Oilbody proteins in microspore-derived embryos of *Brassica napus*: hormonal, osmotic, and developmental regulation of synthesis, Plant Physiol. 97 (1991) 1051-1058.

Higgins DG, Thompson JD, Gibson TJ., Using CLUSTAL for multiple sequence alignments, Methods Enzymol 266 (1996) 383–402.

Hsieh K., Huang A.H.C., Endoplasmic reticulum, oleosins, and oil in seeds and Tapetum cells, Plant Physiology 136 (2004) 3427-3434.

Hu Z., Wang X., Zhan G., Liu G., Hua W., Wang H., Unusually large oilbodies are highly correlated with lower oil content in *Brassica napus*, Plant Cell Reports 28 (2009) 541-549.

Huang A.H.C., Oilbodies and oleosins in seeds, Annu. Rev. Plant Physiol. Plant Mol. Biol. 43 (1992) 177-200.

Huang A.H.C**.,** Oleosins and oil bodies in seeds and other organs, Plant Physiol**,** 110 (1996**)** 1055-1061**.**

Jolivet P., Roux E., D'Andrea S., Davanture M., Negroni L., Zivy M., Chardot T., Protein composition of oilbodies in *Arabidopsis thaliana* ecotype WS, Plant Physiology and Biochemistry 42 (2004) 501-509.

Jolivet P., Boulard C., Bellamy A., Larre C., Barre M., Rogniaux H., d'Andrea S., Chardot T., Nesi N., Protein composition of oil bodies from mature *Brassica napus* seeds, Proteomics 9 (2009) 3268-3284.

Jongschaap R.E.E., Blesgraaf R.A.R., Bogaard T.A., van Loo E.N., Savenije H.H.G., The water footprint of bioenergy from *Jatropha curcas* L., Proc. Natl. Acad. Sci. USA. 106 (2009) 35.

Katavic V., Agrawal G.K., Hajduch M., Harris S.L., Thelen J.J., Protein and lipid composition analysis of oilbodies from two *Brassica napus* cultivars, Proteomics 6 (2006) 4586-4598.

Keddie J.S., Tsiantis M., Piffanelli P., Cella R., Hatzopoulos P., Murphy D.J., A Seed-Specific *Brassica napus* oleosin promoter interacts with a G-box-specific protein and may be bidirectional, Plant Molecular Biology 24 (1994) 327-340.

Kim H.U., Hsieh K., Ratnayake C. Huang A.H.C., A novel group of oleosins is present inside the pollen of *Arabidopsis*, J Biol. Chem. 277 (2002) 22677-22684.

Laemmli U.K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680-685.

Leprince O., van Aelst A.C., Pritchard H.W, Murphy D.J., Oleosins prevent oil-body coalescence during seed imbibition as suggested by a low-temperature scanning electron microscope study of desiccation-tolerant and -sensitive oilseeds, Planta 204 (1998) 109-119.

Li M., Murphy D.J., Lee K.K., Wilson R., Smith L.J., Clark D.C. and Sung J.Y., Purification and structural characterization of the central hydrophobic domain of oleosin,. The Journal of Biological Chemistry 277 (2002) 37888-37895.

Li Y.C., Korol A.B., Fahima T., Nevo E., Microsatellites within genes: structure, function, and evolution, Molecular Biology and Evolution 21 (2004) 991-1007.

Li J., Li M.R., Wu P.Z., Tian C.E., Jiang H.W., Wu G.J., 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 Physiology 28 (2008) 921-927.

Lin L.J., Tai S.S.K., Peng C.C., Tzen J.T.C., Steroleosin, a sterol-binding dehydrogenase in seed oilbodies, Plant Physiol. 128 (2002) 1200-1211.

Lin L.J., Tzen J.T.C., Two distinct steroleosins are present in seed oilbodies, Plant Physiology and Biochemistry 42 (2004) 601–608.

Maes W.H., Achten W.M.J., Muys B**.,** Use of inadequate data and methodological errors lead to an overestimation of the water footprint of *Jatropha curcas*, Proc. Natl. Acad. Sci. USA 106 (2009) 34.

Morello L., Breviario D., Plant spliceosomal introns: not only cut and paste, Current Genomics 6 (2008) 227-238.

Murphy D.J., The biogenesis and functions of lipid bodies in animals, plants and microorganisms, Progress in lipid research 40 (2001) 325-438.

Naested H., Frandsen G.I., Jauh G-Y., Heranandez Pinzon I., Nielsen H.B., Murphy D.J., Rogers J.C., Mundy J., Caleosins: calcium binding proteins associated with lipid bodies, Plant Molecular Biology 44 (2000) 463-476.

Pesole G., Mignone F., Gissi C., Grillo G., Licciulli F., Liuni S., Structural and functional features of eukaryotic mRNA untranslated regions, Gene. 276 (2001) 73–81.

Pons L., Olszewski A., Guéant J.L., Characterization of the oligomeric behavior of a 16.5 kDa peanut oleosin by chromatography and electrophoresis of the iodinated form, J Chromatogr B Biomed Sci Appl. 706 (1998) 131-140

Popluechai S., Breviario D., Mulpuri S., Makkar H.P.S., Raorane M., Reddy A., Palchetti E., Gatehouse A.M.R., Syers K.J., O'Donnell A., Kohli A., Narrow genetic and apparent phenetic diversity in *Jatropha curcas*: initial success with generating low phorbol ester interspecific hybrids. Available from Nature Precedings (2009) 1-44. http://hdl.handle.net/10101/npre.2009.2782.1 (accessed 25/10/10)

Qin X., Zhang J., Shao C., Lin S., Jiang L., Zhang S., Xu Y., Chen F., Isolation and characterization of a curcin promoter from *Jatropha curcas* L. and its regulation of gene expression in transgenic tobacco plants, Plant Mol Biol Rep. (2008) 275-281.

Qu R., Huang A.H.C., Oleosin kDa 18 on the surface of oil bodies in maize genomics and cDNA sequence and the deduced protein structure, Journal Biol. Chem. (1990) 2238-2243.

Rose A.B., Last R.L., Introns act post-transcriptionally to increase expression of the *Arabidopsis thaliana* tryptophan pathway gene *PAT1*. The Plant Journal 11 (1997) 455-464.

Roux E., Baumberger S., Axelos M.A., Chardot T., Oleosins of *Arabidopsis thaliana*: expression in *Escherichia coli*, purification, and functional properties, J Agric Food Chem. 11 (2004) 5245-5249.

Salanoubat M., et al. Sequence and analysis of chromosome 3 of the plant *Arabidopsis thaliana*. Nature, 408 (2000) 820-822.

Sambrook J., Fritsch E.F., Maniatis T., Molecular Cloning, A Laboratory Manual. 2001, CSHL Press, New York.

Siloto R.M.P., Findlay K.F., Lopez-Villalabos A., Yeung E.C., Nykiforuk C.L., The accumulation of oleosins determines the size of seed oilbodies in *Arabidopsis*, The plant cell 18 (2006) 1961-1974.

Simkin A.J., Qian T.; Victoria C.; Franck M.; Mohamed B.R.; Lin C.; Steve T., James M., Oleosin gene family of *Coffea canephora*: Quantitative expression analysis of five *oleosins* in developing and germinating coffee grain, Journal of Plant Physiology 163 (2006) 691-708.

Slack C.R., Bertaud W.S., Shaw B.D., Holland R., Browse J., Wright H., Some studies on the composition and surface properties of oilbodies from the seed cotyledons of safflower (*Carthamus tinctorius*) and linseed (*Linum ustatissimum*), Biochem J. 190 (1980) 551–561.

Sujatha M., Prabakaran A.J., New ornamental *Jatropha* hybrids through interspecific hybridization, Genetic Resources and Crop Evolution 50 (2003) 75-82.

Tai S.S.K., Chen M.C.M., Peng C-C., Tzen J.T.C., Gene family of oleosin isoforms and their structural stabilization in sesame seed oilbodies, Biosci Biotechnol Biochem 66 (2002) 2146-2153.

Tamura K., Dudley J., Nei M., Kumar S., MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0., Mol Biol Evol 24 (2007) 1596–1599.

Ting J.T., Lee K., Ratnayake C., Platt K.A., Balsamo R.A., Huang A.H.C, Oleosin genes in maize kernels having diverse oil contents are constitutively expressed independent of oil contents - Size and shape of intracellular oilbodies are determined by the oleosins oils ratio, Planta. 199 (1996) 158-165.

Tzen, J.T.C., Lai, Y.K., Chan, K.L., Huang A.H.C., Oleosin isoforms of high and low-molecular weights are present in the oilbodies of diverse seed species, Plant Physiology 94 (1990) 1282-1289.

Tzen J.T.C., Huang A.H.C., Surface structure and properties of plant seed oilbodies. J. Cell Biol. 117 (1992) 327-335.

URL 1: http://www.neb.com/nebecomm/products/faqproductP7709.asp (accessed 25/10/10)

van Rooijen, G.J.H., Terning, L.I., Moloney, M.M., Nucleotide sequence of an *Arabidopsis thaliana* oleosin gene, Plant Molecular Biology 18 (1992) 1177-1179.

Weig A.R., Jakob C., Functional identification of the glycerol permease activity of *Arabidopsis thaliana* NLM1 and NLM2 proteins by heterologous expression in *Saccharomyces cerevisiae*, FEBS Lett. 481 (2000) 293-298.

Wu L.S.H., Hong G.H.H., Hou R.F., Tzen J.T.C., Classification of the single oleosin isoform and characterization of seed oilbodies in gymnosperms, Plant and Cell Physiology 40 (1999) 326-334.

Zybailov B., Mosley A.L., Sardiu M.E, Coleman M.K., Florens L., Washburn M.P. Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*, Journal of Proteome Research *5* (2005) 2339-2347.

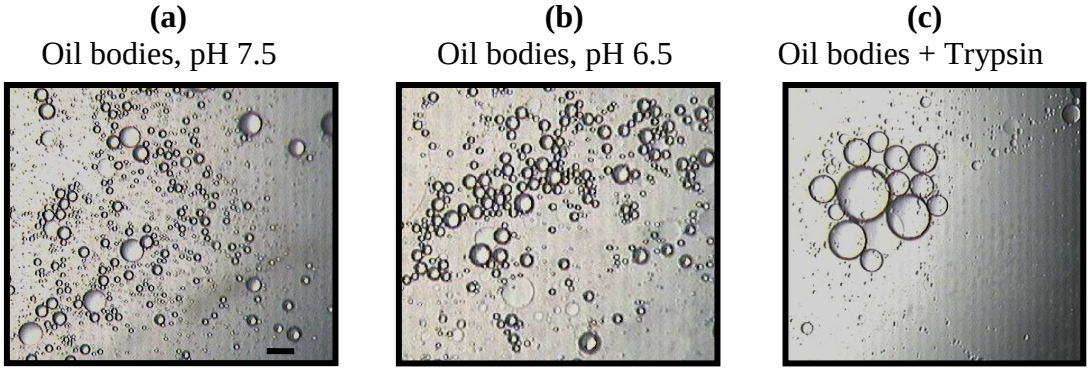
3.7 Supplemental data

Supplemental Table I. List of primers used in chapter 3.

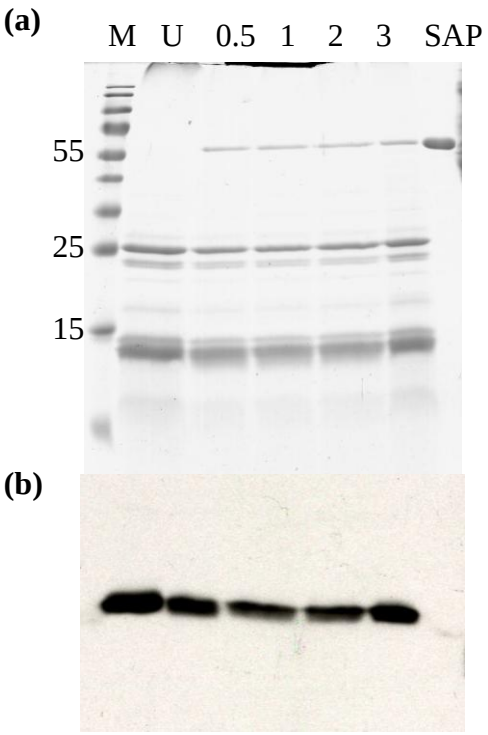
Primer	Primer sequence	Target	Tm°C
JODEGF	5' AGYCCIRTIHTISTICCIGC 3'	Oleosin	58.84
JcO1R	5′ AACCCAATGACCAACGCCGC 3′	Oleosin1	64
JcO2R	5′ AATGCCATGACCGAAAGG 3′	Oleosin2	54
JcO3R	5′ ATGGCAGCGACCCCAAAACCG 3′	Oleosin3	75.63
JcFLO1F	5′ CTTCCATGGCCGAGTATGAG 3′	Full length Oleosin1	62
JcFLO1R	5′ ACCAAAGCTATGGCAAACA 3′		58
JcFLO2F	5′ TTATCATCAACAAATGGCTGAA 3′	Full length Oleosin2	58
JcFLO2R	5′ CCAAAACCCAGCTCTCTGAC 3′		62
JcFLO3F	5′ TGCTTCGATCGCGTCAATGG 3′	Full length Oleosin3	62
JcFLO3R	5′ CAGAGACCTTCAACACATGACAA 3′		66
O3DWGSP1	5′ CTGAGGCTAAAAACCCTGTGAT 3′	DNA walking upstream olesoin3	64
O3DWGSP2	5′ GGTCAAACCGGAGAGAACTAGA 3′		66
O3DWGSP3	5′ CACTACTTGGTGGGATCGTGGC 3′		70
O3inF	5′CCGCTCGAGTTAAACAAATTTGCCTACTTC 3′	Intron of olesoin3	86
O3inR	5′GGCCTCGAGACCTGCAATTCATGTCTC 3′		84
Rt-O1	F -5′ GGAATCACTGCGCTTTCTTC 3′	qPCR oleosin1	60
	R -5′ AATTGACCCGTCGTTTCTTG 3′		58
Rt-O2	F- 5′ GGAAAGCACATGAAGGGAAA 3′	qPCR oleosin2	58
	R- 5′ TCCCCCATAAACCAACTCAA 3′		58
Rt-O3	F- 5′ GCCGACGGTTATCACAGTTT 3′	qPCR oleosin3	60
	F- 5′ GGGGATGCTTTCCAGTAACA 3′		60
Rt-KAS	F- 5′ GCAGTTGCGACAGCATTAGA 3′	qPCR KAS	60
	R- 5′ TCACTTTCCCACTTCGAACC 3′		60
Rt-dgat	F 5′ TGGGTCTGCTGTTCTGTCTG 3′	qPCR dgat	62
	R 5′ GAGTTGGCAATGGCTCTCAT 3′		60
Rt-18S	F 5′ AAACGGCTACCACATCCAAG 3′	qPCR 18S	60
	R 5′TCATTACTCCGATCCCGAAG 3′		60

Supplemental Table II Information on *oleosins* from *Arabidopsis*, Ficus, Coffee, *Ricinus*, Olive. NR = Not reported. ND = Not Determined.

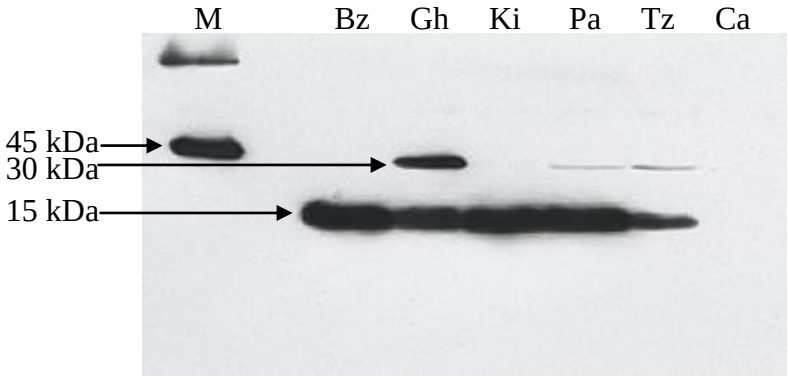
Name	**Accession no**	**mRNA size (nt)**	5'UTR **(nt)**	**ORF (nt)**	3'UTR **(nt)**	**Intron before or after HP**	**Nucleic acids NT+HP+CT**	**Oleosin isoform**	**DNA seq Identity to *Jatropha* oleosin/coverage**
AtS3	At4g25140	908	68	521	318	After	135+216+168	L-form	71 % to JcO3/ 28%
AtS2	At3g01570	872	43	599	229	Before	153+225+189	H-form	65 % to JcO2/ 62%
AtS4	At3g27670	874	96	575	202	Before	108+225+189	H-form	ND
Ficus OLEH	EF583923	669	28	468	173	NR	NR	H-form	75 % to JcO2/ 68 %
Ficus OLEL	EF583922	586	10	462	114	NR	NR	L-form	71 % to JcO3/ 45 %
Coffee OLE1H	AY841271	629	78	551	104	NR	NR+243+NR	H-form	66 % to JcO2/ 37%
Coffee OLE2L	AY841272	558	87	471	64	NR	NR+219+NR	L-form	72 % to JcO3/ 45 %
Coffee OLE4L	AY841274	720	74	428	218	NR	NR+216+NR	L-form	ND
Ricinus OLE1H	AY360218	720	68	462	190	NR	NR	H-form	72 % to JcO2/ 68 %
Ricinus OLE2L	AY360219	711	86	417	208	NR	NR	L-form	73 % to JcO3/ 59 %
Olive H	AY083161	792	46	498	248	NR	NR	H-form	ND



Supplemental Figure 1 Surface properties of *Jatropha* seed oil bodies. Stable oil bodies were maintained as individual particles in a medium of pH 7.5 at 23°C. Aggregations of *Jatropha* oil bodies occurred by inducing the pH of medium to pH 6.5. Coalescence of *Jatropha* oil bodies were induced by Trypsin digestion. The results suggested that the stability of *Jatropha* oil bodies could possibly be maintained by proteins on the surface of *Jatropha* oil bodies (mainly oleosins) using electrostatic repulsion and steric hindrance. The bar represents 5 um.

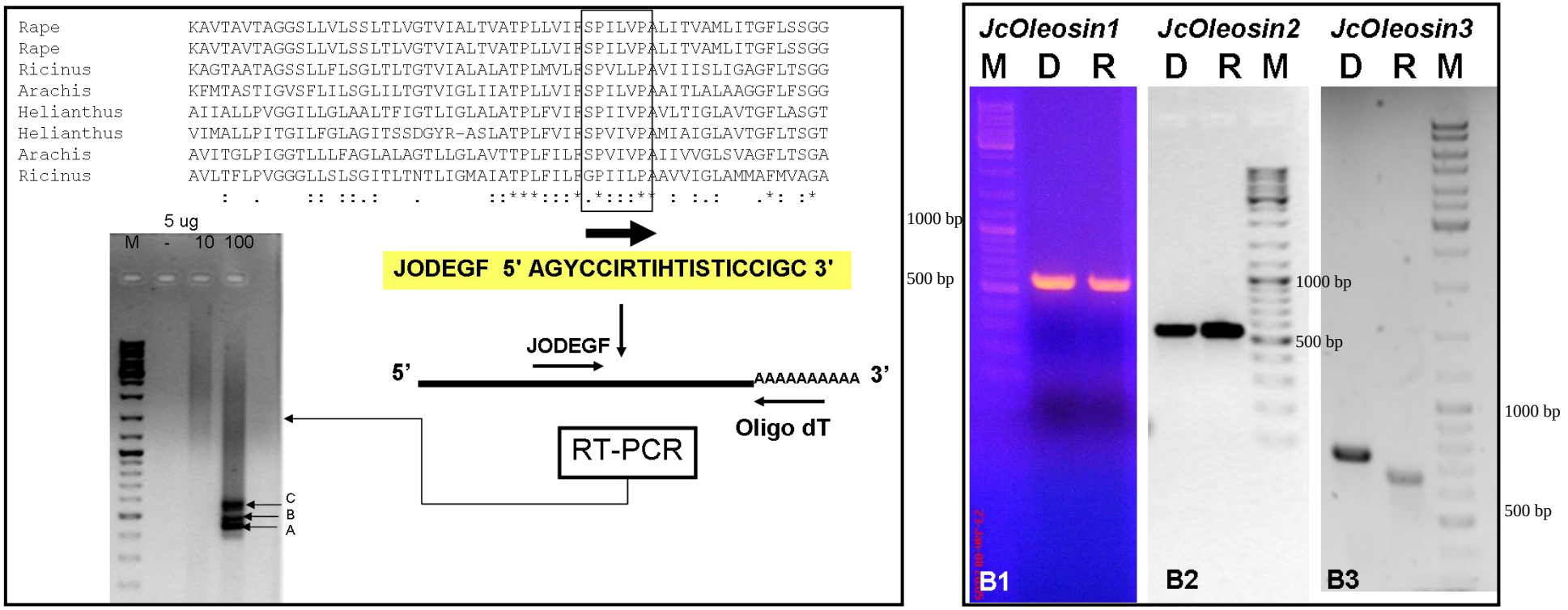


Supplemental Figure 2 (a) SDS-PAGE and (b) immunoassay of alkaline phosphatase treatment on *JcOloesin3* using Shrimp Alkaline Phosphatase (SAP) at different time intervals (0.5, 1, 2, and 3 hour). M represents molecular weight (kDa) marker. U represents untreated sample.



Supplemental Figure 3 Immunoassay for *JcOleosin3* in *J. curcas* accessions from Brazil (Bz), Ghana (Gh), Kilimanjaro (Ki), Paraguay (Pa), Tanzania (Tz) and an accession of *Camelina* (Ca) a different oilseed plant. *J. curcas* accession from Bz, Pa and Tz show a *JcOleosin3* dimer at around 30 kDa, while accessions from Gh and Ki exhibit a ca. 15 kDa monomer. There is no cross hybridization with *Camelina* oleosins.

3.8
Appendixes



Appendix 1 (A) Summarized diagram showing *Jatropha* oleosin gene isolation using degenerated PCR mediated methods. Sequence alignment of oleosin oil seed plants used for degenerated primer design, indicated in the rectangular box and the arrow shows direction of primer from 5' to 3' (A1). The degenerated primer sequnce, designated JODEGF, is shown in box (A2). *Jatropha* oleosin genes were isolated using degenerated and oligo dT primers. Agarose gel electrophoresis showed that three PCR products (indicated by A, B, and C) were obtained at 100 uM primers (A3). (B) Agarose gel electrophoresis obtained from genomic (D) and cDNA (R) using 5' and 3' primers foreach oleosin gene (*JcOleosin1* for B1, *JcOleosin2* for B2, *JcOleosin3* for B3).

```
JcOLE1H         ------------------------------------------------------------
OeOLH           ------------------------------------------------------------
JcOLE2H         ------------------------------------------------------------
RcOLE1H         ------------------------------------------------------------
FpoleH          ------------------------------------------------------------
CofOLE-1        ------------------------------------------------------------
AtOLEO2H        -------------------------ATTACAAAGAAAATAGGTAAAAACAATTTCTCATT 35
AtOLEO4H        TTTATATGTACGTTCTCGTAGACTTATCTCTATATACCCCCTTTAATTTGTTTGCTCTTA 60
JcOLE3          ------------------------------------------------------------
RcOLE2          ------------------------------------------------------------
CofOLE-2        ------------------------------------------------------------
FpoleL          ------------------------------------------------------------
AtOLEO1L        ------------------------------------------------------------AT 2
CofOLE-4        ------------------------------------------------------------


JcOLE1H         ---------------------------------------CTGCAAAAATCTACT--CTTT 19
OeOLH           ---------------ATTCAATTTCCTCTTTCTTTCACACCAAAAAAAATTGAA--TTTT 43
JcOLE2H         ----------------------------ATCTGTTTTTCTCTACATTAACATATC---TTT 30
RcOLE1H         ---------------GTTCGACTGATAATCTCTTTCTTTGTACTCTTTTTTGGCGACTTC 45
FpoleH          ------------------------------------------------------CTCCT 5
CofOLE-1        ---------------GGACCATCTTTAGCACTCAAAAAGGGGAAAGCAAAAAACTCCGAT 45
AtOLEO2H        AGCTTACAATGGCGGATACACACCGTGTCGACCGTACTGATAGACAC---TTTCAATTTC 92
AtOLEO4H        GCCTTTACTTTATAGTTTTATATCATATCAATCGACATGGCGAATGTGGATCGTGATCGG 120
JcOLE3          ----------------------------------------------ACCCACAA-C 9
RcOLE2          ---------------------------------------------CATCGCTCCCTTACTC 16
CofOLE-2        ------------------------------------------------TGCAGAATTC 10
FpoleL          ------------------------------------------------------------
AtOLEO1L        ATACACATCTTTTTGATCAATCTCTCATTCAAAATCTCATTCTCTCTAGTAAACAAGAAC 62
CofOLE-4        ------------------------------------------------------------


JcOLE1H         TTCTCCCTTCCAAGAAAAAGCTTATTACGATATCTATTTTAAGCTTCCATG--GCCGAGT 77
OeOLH           CTCATGGCGGAACGTGATCGTCCACAGCCACATCAGGTTCAAGTTCACA-----CTCAGT 98
JcOLE2H         CTCACACTTATCATCAACA---AATGGCTG-AACGTTCACAGCCACACCAA--GTTCAAG 84
RcOLE1H         TTTGAACTTGGAAGCTACTGACAATGGCTG-ATCGTCCGCAGCCGCACCAG--GTGCAAG 102
FpoleH          TGAAAAAAAAAAAAAAAACGAAAAATGGCCG-ACCGTCCTCAAGCTCACCAG--ATCCAAG 62
CofOLE-1        TTTTTTCCCATAAGCTACCGTCCACCGCCCTTCATGGCTGAGCACTACCAG--CTGCAGC 103
AtOLEO2H        AGTCGCCCTATGAA-------GGCGGCCGAGGTCAA-GGTCAG--TATGAA--GGTGACC 140
AtOLEO4H        CGTGTGCATGTAGACCGTACTGACAAACGTGTTCATCAGCCAAACTACGAA--GATGATG 178
JcOLE3          CGTATATCAAATATCACTGATCGTAAACCTCCAT-TGAATTTACTCTGTG---CTTCGAT 65
RcOLE2          CGCAAAAAAGCAATCATCT-CCATATATTTTATTATATATTAACCAGAAG---CTCTGAT 72
CofOLE-2        GGTACCACAACAAAAACTT----CAATTCCCACCATC----CGCTTGTCC---CTTCGCC 59
FpoleL          -----------------TC----CGAATCCGATGGCGGAG-CCCCAGTCG---CTGCAGC 35
AtOLEO1L        AAAAAAATGGCGGATACAGCTAGAGGAACCCATCACGATATCATCGGCAGAGACCAGTAC 122
CofOLE-4        ------ATCCCAATTCCCAACTCCATTTCCATCCATTG---TGCTGAAAATCCCAACAAA 51


JcOLE1H         ATGAGCGACGACCA-GGCGGAGAAGGGCACAAAGGAGCCCTAAAGGAGAAGGGCGGCTTC 136
OeOLH           -CACGCTACGATCAAGGCGGCGGAATGAAGAGTGTCCTCCCAAAAAAG------GGTCCC 151
JcOLE2H         TGCACCCACAAC-ACCGCTATGAAGCTGCCTTCAAAGGTCAACAAAAA------GGTCCA 137
RcOLE1H         TGCACC-GCTAT-GATCCAACTA--CTGGCTACAAGGGCCAACAGAAG------GGTCCA 152
FpoleH          TCCATCCACAAT-ACGGTAAGGGGTTTCAGCAGAGAGGAGAACAGCAA------GGCCCG 115
CofOLE-1        AACGCC--CCACAGAGGCCGTCA-----AAAGCTTCCTTCCTCAGAAG------GGTCCA 150
AtOLEO2H        GTGGTT-ACGGTGG---TGGCGGTTACAA----GAGCATGATGCCTGAAA--GTGGCCCA 190
AtOLEO4H        TCGGTT-TTGGTGGCTATGGCGGTTATGGTGCTGGTTCTGATTATAAGAGTCGCGGCCCC 237
JcOLE3          CGCGTCA-------ATGGCGGAGCACCCAC--AGTCGCAGCATGTTGGCC--AACAGCCA 114
RcOLE2          TCTGCAAGCCTATTATGGCTGAGCATCAAC--AATCACCGGTTGTTAGTC--ACCGCCCA 128
CofOLE-2        TTAGGAATCCTCCTCCTCTTCCTCCTC-AT--GGCGGACATTCGTCAG-C--AGCAGCCG 113
FpoleL          GCGGGGAGCGGGGGGAGCAGCTGCAGCTAC--AGCTGCAG--CAGCAG-C--AGCACCCG 88
AtOLEO1L        CCGATGATGGGCCGAGACCGAGACCAGTACCAGATGTCCGGACGAGGATCTGACTACTCC 182
CofOLE-4        AAATTGAAGGAAAAAAAAAGAAATGGCAAC--CCTCCCTGACCAGCCGCAGACACAGCAC 109


JcOLE1H         TCGACCTCACAGATTTTAGCGGTTGTCACTCTTTTGCCGGTGAGTGGGACGCTTCTTTTC 196
OeOLH           TCAACTAGCCAAGTTTTGGCAGTCGTCACTCTCCTCCCTGTTGGCGGAACCCTACTGGCA 211
JcOLE2H         TCAGCTCAGAAAGTGTTAGCTGTGATCACCCTCCTGCCGGTAGGCGGTGGGCTTCTTGCA 197
RcOLE1H         TCAGCCTCCAAAGTGTTAGCTGTGTTAACCTTTCTGCCGGTTGGTGGTGGTCTTCTATCT 212
FpoleH          TCGGCGGGCAAAATCCTGGCCGTTCTGACACTCCTTCCGGTGGGCGGGACTCTGCTGTTC 175
CofOLE-1        TCAACTTCACATGTGTTAGCAGTTGTCACGCTCCTCCCAGTTGCGGGAGTCCTGCTAGGC 210
AtOLEO2H        TCTAGTACCCAAGTATTGTCCCTGTTGATTGGAGTCCCTGTCGTCGGTTCGCTACTTGCC 250
AtOLEO4H        TCCACTAACCAAATCTTGGCACTTATAGCAGGAGTTCCCATTGGTGGCACACTGCTAACC 297
JcOLE3          CGATCCCACCAAGTAGTGAAGGCAGCCACTGCAGTCACTGCCGGTGGATCTCTTCTAGTT 174
RcOLE2          CGGGTGAACCAACTGGTGAAGGCTGGAACCGCCGCCACAGCTGGTAGCAGTCTCCTATTT 188
```

Appendix 2 (continue next page)

```
CofOLE-2        CTATCCCACCAGGTGGTCAAGGCGGCCACCGCGGTGACGGCTGGTGGGTCGCTTCTGGTT 173
FpoleL          CGGTCCCACCAGGTGGTGAAGGCAGCGACAGCCGTGACCGCCGGCGGGTCCTTGCTGGTC 148
AtOLEO1L        AAGTCTAGGCAGATTGCTAAAGCTGCAACTGCTGTCACAGCTGGTGGTTCCCTCCTTGTT 242
CofOLE-4        TCAACTAGTCAGGTGGTAAAAACCACCACAGCAGTCGCCGTGGGGGGTTCACTCATGCTT 169
                         *  *                        *       *     *  *

JcOLE1H         CTTGCTGGAATAACGCTTACTGGAACACTCATTGCGCTAGCTGTTGCGACTCCGCTGTTC 256
OeOLH           CTCGCCGGATTGACACTTGTCGAAAGCCTTATTGGCCTTGCTGTCACTACTCCTCTTTTC 271
JcOLE2H         CTAGCTGGTATAACCTTAGTAGGTACCCTAATCGGCCTAGCTATTACTACCCCTCTTTTT 257
RcOLE1H         CTCTCTGGCATAACCTTAACGAATACGCTCATCGGGATGGCCATTGCCACCCCACTTTTT 272
FpoleH          CTGGCGGGTCTGACTTTCATCGGGACCCTTATTGGGCTGGCCCTGAGCACGCCGGTGTTC 235
CofOLE-1        CTTTCTGGGCTGATTCTCGTCGGAACGGTCATCGGTCTGGCGGTGACAACCCCGCTTTTC 270
AtOLEO2H        TTGGCTGGATTACTTCTAGCTGGTTCGGTGATCGGCTTAATGGTTGCTTTACCACTATTT 310
AtOLEO4H        CTAGCTGGACTCACTCTAGCCGGTTCGGTGATCGGCTTGCTAGTCTCCATACCCCTCTTC 357
JcOLE3          CTCTCCGGTTTGACCCTTGCTGGTACGGTCATTGCACTGACCGTACCCACTCCATTGCTG 234
RcOLE2          CTTTCTGGTTTGACCCTCACTGGTACGGTTATTGCACTCGCACTGGCCACTCCGTTGATG 248
CofOLE-2        CTTTCGGGCCTGATCCTAGCAGGGACTGTGATTGCACTAGCTCTGGCCACGCCTTTGCTG 233
FpoleL          CTGTCGGCGCTGATCCTCGCCGGAACGGTGATCGCGCTGACGATCGCCACGCCGCTGTTC 208
AtOLEO1L        CTCTCCAGCCTTACCCTTGTTGGAACTGTCATAGCTTTGACTGTTGCAACACCTCTGCTC 302
CofOLE-4        CTGTCGGGGCTGACACTGGCTGGAACAATAATAGGCCTAGTTTTGGCAACGCCCCTTTTG 229
                 *  *     *     *           * ** *   *     *       **  *  *

JcOLE1H         GTGATTTGTAGTCCGGTGAGGGGGCCGGCGGCGTTGGTCATTGGGTTGTCGGTGTTAGGG 316
OeOLH           ATCATCTTCAGCCCAGTTCTCGTCCCGGCCACCATTCTTGTTGGCCTTGCAGTCACTGCT 331
JcOLE2H         GTCATCTTTAGCCCTGTTCTTGTCCCAGCTGCATTAGTTATTGGCCTTTCGGTCATGGCA 317
RcOLE1H         ATTCTCTTCGGCCCTATAATTCTTCCTGCTGCCGTTGTTATTGGCCTTGCTATGATGGCA 332
FpoleH          ATTCTCTTCAGCCCGGTTCTGGTCCCGGCGGCCATAACTATAGGGCTAGCCATTACTGGG 295
CofOLE-1        GTTATCTTTAGCCCCATTTTGGTCCCAGCTGTATTTACCCTAGGGCTGACCCTGGCCGGG 330
AtOLEO2H        CTCCTCTTCAGCCCGGTTATAGTCCCAGCGGCTCTAACTATCGGGCTTGCAATGACAGGC 370
AtOLEO4H        CTCCTCTTCAGTCCGGTGATAGTCCCGGCGGCTCTCACTATTGGGCTTGCTGTGACGGGA 417
JcOLE3          GTCATATTCAGTCCGGTGATGGTGCCGACGGTTATCACAGTTTGTTTGATAATCACAGGG 294
RcOLE2          GTCCTTTTCAGCCCAGTGCTTCTTCCTGCTGTGATCATAATTAGTCTTATTGGCGCTGGG 308
CofOLE-2        GTGATTTTCAGCCCGGTGCTGGTCCCCGCGGGGATCACGGTGTTTCTGCTGGTCACCGGG 293
FpoleL          GTGATATTCAGCCCGGTGCTGGTGCCGGCTGTGATCACGCTGGGGCTGATCATCATCGGG 268
AtOLEO1L        GTTATCTTCAGCCCAATCCTTGTCCCGGCTCTCATCACAGTTGCACTCCTCATCACCGGT 362
CofOLE-4        GTGATTTTCAGTCCTGTGATAGTCCCAGCTGCTGTAACCTTCTTTCTGATTCTTGCAGGG 289
                 *  * *   * **  *       **  *     *     *     *          *

JcOLE1H         TTTTTGACGTCCGGTGCTTTTGGAATCACTGCGCTTTCTTCACTGTCGTGGATGGTTAAT 376
OeOLH           TTTCTTACTTCCGGTGCCTTCGGGCTAACCGGGCTGTCATCGCTTTCTTGGGTTGTGAAT 391
JcOLE2H         TTTTTGGCTTCCGGGGCTATGGGGCTCACAGGGTTGTCATCACTCTCGTGGGTTCTGAAA 377
RcOLE1H         TTTATGGTTGCTGGAGCTCTCGGGCTGAGCGGGCTGACGTCGCAGTCGTGGGCGTTGAAG 392
FpoleH          TTTCTGACCTCGGGGGCTTTCGGAGTAACGGGGCTGTCGTCTCTGTCTTGGGTGTTCAAC 355
CofOLE-1        TTCTTGACCTCCGGTGCTTTCGGGATCACTGCACTTGCTTCATTGTCGTGGATGCTGAAC 390
AtOLEO2H        TTTTTAGCCTCGGGGATGTTCGGTCTAACCGGGCTTAGCTCAATCTCATGGGTCATGAAC 430
AtOLEO4H        ATCTTGGCTTCTGGTTTGTTTGGGTTGACGGGTCTGAGCTCGGTCTCGTGGGTCCTCAAC 477
JcOLE3          TTTTTAGCCTCAGGCGGTTTTGGGGTCGCTGCCATTTTTGTCTTGTTTTGGATTTATAGG 354
RcOLE2          TTCTTGACTTCTGGCGGTTTCGGCTTCGGAGCAATTCTCGTCCTGTCATGGATCTATAGG 368
CofOLE-2        TTCCTGTCTTCGGGAGGGTTTGGGGTGGCAGCGTTGAGCGTGCTGTCCTGGATTTACGGG 353
FpoleL          TTCCTGGCGTCAGGAGGCTTCGGCGTGGCGGCGCTCACTGTTCTGTCGTGGATCTACCGG 328
AtOLEO1L        TTTCTTTCCTCTGGAGGGTTTGGCATTGCCGCTATAACCGTTTTCTCTTGGATTTACAAG 422
CofOLE-4        TTCTTTATCTCTGGAGGTCTTGGCGTAACTGCCACCTTTATTTTCTACTGGATGTTCAGG 349
                 *  *     * **     * **  *    *              *  ***

JcOLE1H         TATATTCGCAGGATGAGAGGATCTTGGACGATGCAGATGGAAATGGCAAAGCGGCGCGCG 436
OeOLH           TTCTTGAGGCAGGTTAGTGGGTCC------ATGC---TGGATCTAGCGAAGAGCCGGATG 442
JcOLE2H         TACCTCCAGGAGGTCACTCGGCGAATGCCGGAGCAGCTGGATATAGCTAAGAAGCGTATG 437
RcOLE1H         TATTTCAGGGAAGGTACT-GCC--ATGCCAGAATCACTGGACCAGGCGAAGAAGCGCATG 449
FpoleH          TACTTCAAGCGGACCTCG------GTGCCGGACCAGCTGGACTACGCCAAGAGGCGCGTG 409
CofOLE-1        TACATCCGACTCATGAAGGCGTCTTCCCAGGAGCAAATGGACCTCGCAAAGTGGCGCGTG 450
AtOLEO2H        TATCTTCGTGGGACAAGGAGAACTGTGCCTGAGCAATTGGAGTATGCTAAGAGGAGAATG 490
AtOLEO4H        TACCTCCGTGGGACGAGTGATACAGTGCCAGAGCAATTGGACTACGCTAAACGGCGTATG 537
JcOLE3          TATGTTACTGGAAAGCATCCCCCAGGAGCAGAAAATCTGGACCAGGCACGCTTAAAGTTG 414
RcOLE2          TATGTTACAGGAAAGCAGCCACCAGGAGCGGAGAGCTTGGACCAGGCTCGCTTGAAGCTA 428
CofOLE-2        TACGTGACGGGGAAGAACCCGCCGGGAGCGGACCAGCTGGACCGGGCAAGGCAAAAGCTG 413
FpoleL          TACGTGACCGGGAAGCACCCGCCGGGTGCCGACCAGCTGGACCAGGCCCGCCACAAGCTG 388
AtOLEO1L        TACGCAACGGGAGAGCACCCACAGGGATCAGACAAGTTGGACAGTGCAAGGATGAAGTTG 482
CofOLE-4        TATGCGACTGGGAAGCACCCCATCGGGGCAGATCAGCTGGATCGTGCGAGAGAAAAGATT 409
                *                                     ****    **

JcOLE1H         CAAGAAACGACGGGTCAATTGGGCCAGAAGGCTAGGGAGGTGGGCCAGAAG---GCCCAG 493
OeOLH           GGGGATGCGGCCATACAAGTTGGGCAGAAAACAAAGGAGACTGGGCAGACG---ATACAG 499
JcOLE2H         CAAGACATGGCTGGTTTTGTGGGACAGAAGACCAAGGAAGTAGGCCAAGAA---ATCCAG 494
RcOLE1H         CAGGACATGGCCGGTTATGTTGGGATGAAGACCAAGGAAGTGGGACAAGAC---ATCCAG 506
FpoleH          CAGGACATGGCTGGCTACACCGGCCAGAAGGCCAAGGAAGTTGGCCAGGAG---GTGCAG 466
CofOLE-1        CAGGACACTGCCGGCCAAGTTGGTCAGAAAGCGAGAGACGTGGGCCAGAGA---ACTCAA 507
AtOLEO2H        GCTGATGCGGTTGGCTACGCAGGACAAAAGGGCAAAGAAATGGGCCAGCAT---GTGCAG 547
```

Appendix 2 (continue next page)

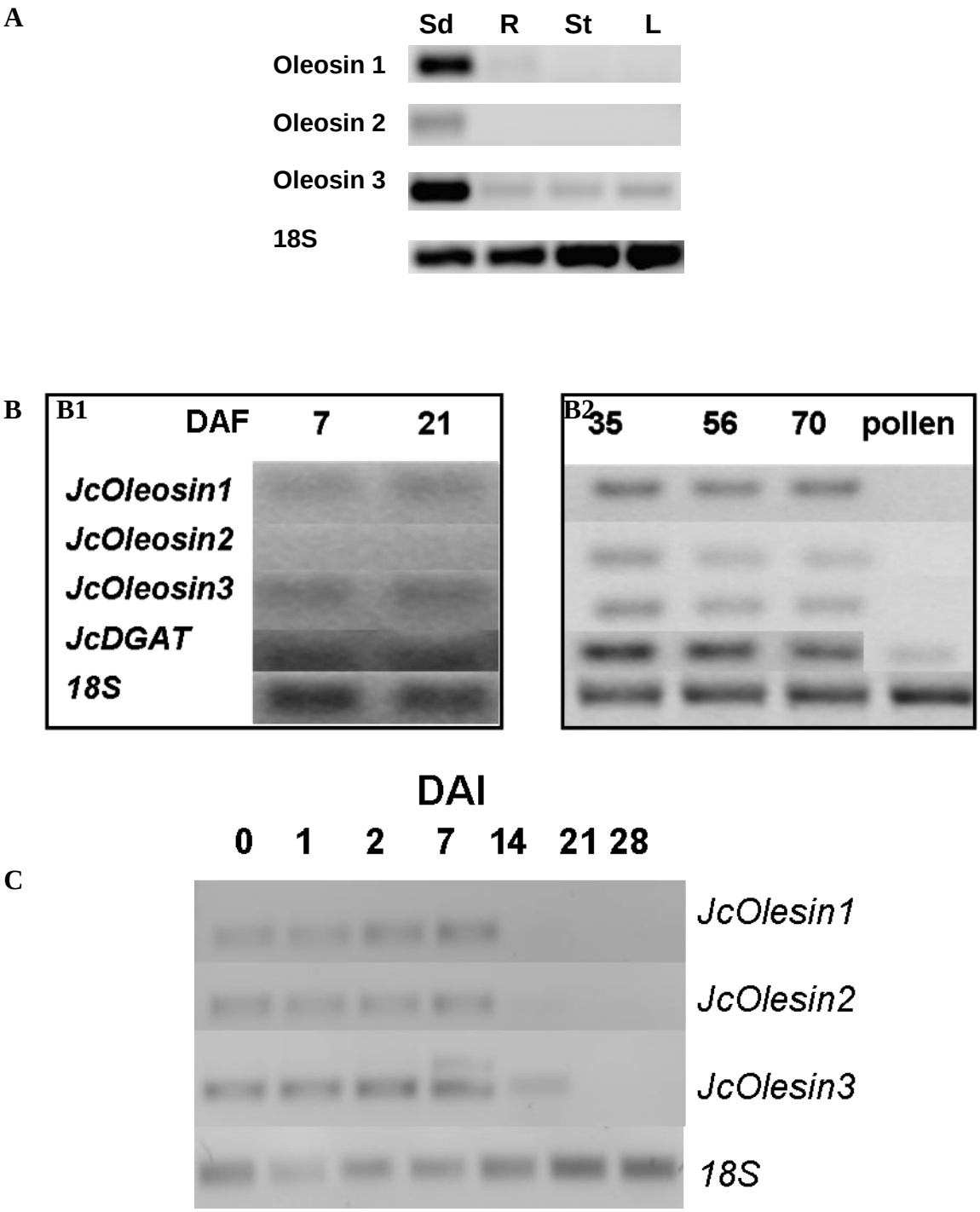
```
AtOLEO4H        GCTGATGCGGTAGGCTATGCTGGTATGAAGGGAAAAGAGATGGGTCAGTAT---GTGCAA 594
JcOLE3          GCGGGAAAGGCAAGAGAGATGAAGGACAGAGCTGAGCAGTTTGGACAGCATGTAACAGGT 474
RcOLE2          GCTGGGAAGGCAAGGGAGATGAAGGACAGGGCTGAGCAGTTTGGACAGCATGTAACCGGT 488
CofOLE-2        GCTCTTAAGGCTAGGGAGATGAAGGACAGGGCTGAGCAGTATGGGCAGCAG--AACATAC 471
FpoleL          GCCAGCAAGGCCAGGGAGATGAAGGACAAGGCGGAGCAGTTTGGCCAGCAG--CACCTCA 446
AtOLEO1L        GGAAGCAAAGCTCAGGATCTGAAAGACAGAGCTCAGTACTACGGACAGCAACATACTGGT 542
CofOLE-4        GCTCATGCAGCTAAGGAGATGAGGGACAAAGCTGAACATTTTGGGCAGCAG---GCTCAG 466
                                          *         *    ** **

JcOLE1H         GAAGTTGCAAAGACATGATCA-------GGTTGT---CTAGCAGTAGTTGTT-------- 535
OeOLH           CAGAAAGCACCAGAAGGGAAA-------GAAAGTA--CCGGTGGTAGAACTT-------- 542
JcOLE2H         AGGAAAGCACATGAAGGGAAA-------TAAGTA------GAGAGAAATGCT-------- 533
RcOLE1H         AGGAAGGCACAAGAAGGGAAA-------TGATAATAATGTAAGAGGAATAAT-------- 551
FpoleH          AGCAAGGGCCAGGAAGGGAAG-------AGGACCTAGCGAGAGAAAGATAGA-------- 511
CofOLE-1        GATGTAGCCAGAGCATGAGCT----------------CCTGAAGTGGGCTTT-------- 543
AtOLEO2H        AACAAGGCCCAAGATGTTAAACAA----TATGATATTTCTAAGCCACATGACACTACCAC 603
AtOLEO4H        GATAAGGCTCATGAGGCT----------CGTGAGA---CTGAGTT-CATGAC-------- 632
JcOLE3          CAACAGACTTAAAAGAGAAGTGGGTTTCGGTGGTTGATGCAGAGAACATAAAATGGCTTT 534
RcOLE2          CAGCAGACTTCCTAGACA---GTGCTTCCATTTCATGCTTGATGGATATAAAATGGCTTT 545
CofOLE-2        CCGCCGGTTCCCAGCAGCACT-------------AGATTTGAAATTTAAGC--CAATGAT 516
FpoleL          CCTCCGGCCAACAGCAGTCTTCTTAATTTACGGCGGTTCTGACGATGACGTGACGATGAT 506
AtOLEO1L        GGGGAACATGACCGTGACCGTACTCGT-GGTGGCCAGCACACTACTTAAGTTACCCCACT 601
CofOLE-4        CAGCAGATTAAAGGTTCCCCAGATCAA-GATACTTGATGGAATGTTGTGCATGCAATTAA 525


JcOLE1H         ---GTTTGGTA--------AATAATAT-AAGCTAATAT-----AGTT-------TGAGTT 571
OeOLH           ---GAA-GATG--------GATGATAT-AGTACAGAATTT--AAGAT-------CCAGTT 580
JcOLE2H         ---GTTGGTTATT------AAGGGTTG-AAGTCAGAGAGC-TGGGTT-------TTGGTA 575
RcOLE1H         ---GTTGGGTTTTTATATGAAGGGTTT-ATGTTTGTTTGTATGGGGT-------CTGCTT 600
FpoleH          ---GAGAGATAGAGA----GAGAGAGA-GAGAGAGAGAGAAGGGAGT-------GTATCT 556
CofOLE-1        ---GTGTCACGAG------CTCGGCCC-AATTTTGTAT---TGGGTT-------TTGATG 583
AtOLEO2H        TAAGGGTCATGAGACTCAGGGGAGGAC-GACGGCTGCATGATGAGTTTTCAGTATGAACG 662
AtOLEO4H        TGAGACCCATGAG---CCGGGTAAGGC-CAGGAGAGGCTCATAAGCT---AATATAAATT 685
JcOLE3          GATATATATATAAAAATATATGAGTCCTTTAATTAACATGGAAGGCCATTGCTTGTCTGG 594
RcOLE2          GATATAGATTT------------GTC------TTGGTTCAAAAGGCCATCGCTTGTCTAG 587
CofOLE-2        CAGCTCCAATT---------CCAGT---------GGTGTGGATAGGTGGAATTGTCATGT 558
FpoleL          CATTTAGGGTTG--------TTTGTC-------AGGAATAAGAAAGAGATATTTGCATGC 551
AtOLEO1L        GATGTCATCGTCATAGTCCAATAACTCCAATGTCGGGGAGTTAGTTTATGAGGAATAAAG 661
CofOLE-4        TGTATGATGATG------CATGAATAAACTCTGTGCTTTAATTAGCCGGGATTGTGACTT 579


JcOLE1H         T--GATGATGGATTTCATAAATATGTAATAAGACTTGTTTGTTTGCCCATAGCTTTGGTT 629
OeOLH           TCCATTGCTGAAATAATGCAGTTTCTGCTGTTATAAGTTTCTTGTTTGTTTGTGTATGTG 640
JcOLE2H         CGAGACTACTTTT--GTTGTATGGGTGTTGTTTTGAGGTAGTTGAGTTGGTTTA--TGGG 631
RcOLE1H         CGGGGACTCTTCTTCGTTATGTTTGTGTTTCTTTACTTTTCTTAATTTTATTCTGTCTTT 660
FpoleH          AGTGATGGC------AAAGGAAAGGGGAAATAATGTGGGAATTTGCATGGTTTTTTAGAT 610
CofOLE-1        TTCGTTTTGTTTCTGATCCTTGTTAATCTTGCCTATGGCTTTTATA-------------- 629
AtOLEO2H        GTAGATATGTGTTTTCACTATTATGTCGTTTTTTCTGCATTTTCAATATGATGTTATGTG 722
AtOLEO4H        GCGGGAGTCAGTTGGAAACGCGATAAATGTAGTTTTACTTTTATGTCCCAGTTTCTTTCC 745
JcOLE3          TGTAT-TTCTTTTTTCGCTTAAATCTTTTTGTTTCTTCTATGTAATTTGTTAACGGGTCT 653
RcOLE2          TGT---TTCTATTTTCGGTTCTTTGTTGGTTGATCTTCTTGTTAATGAGTCTTTTGTGAG 644
CofOLE-2        ------------------------------------------------------------ 
FpoleL          TTTGC-TGTGTGGATACGAGGAAAATTAAGAATTAA------------------------ 586
AtOLEO1L        TGTTTAGAATTTGATCAGGGGGAGATAATAAAAGCCGAGTTTGAATCTTTTTGTTATAAG 721
CofOLE-4        TGTATGTTTCCTAGTTAATCGGCAATTCAATTGTTGTTGTGACTGTTGTAGAGAGTGGAG 639


JcOLE1H         GGCTGTATGTTTCTG-TTTACTGA------------------------------------ 652
OeOLH           AACTGTGTGTTTTTGATTTGGTGATTGCTTTCTGTGGAACGTAGTAGAGAAACATATGAA 700
JcOLE2H         GGATCTTCTTTGTTGCATTTGTATTTCTCTTTTTGGTCTTTTGGGTTCCATTGT------ 685
RcOLE1H         GGAGTTCAATTGTTAAATTTCAATGTGCTTATCTTGTTAGTTAAAAAAAAAAAAAAAAAA 720
FpoleH          TTACATAAAATGTTA-ATGTACGGCTTTTTTTATTTTATTTTAAAAAAAAAAAAAAAAAA 669
CofOLE-1        ------------------------------------------------------------ 
AtOLEO2H        TTTTTTTTGTTTGGCTTTTTGTTGAACCGTGTATGTGTTTTATGTTTTTGTAAGCATGAA 782
AtOLEO4H        TCTTTTAAGAATATCTTTGTCTATATATGTGTTCGTTCGTTTTGTCTT-GTCCAAATAAA 804
JcOLE3          TCCCTCTTTTGAGAGCTGTCTAAAATAATTTTCTTGTCATGTGTTGAAGGTCTCTGTAAT 713
RcOLE2          TGCTTTATAGTAATAATAATAATAATAATTTCTTTTTACCAAAAAAAAAAAAAAAAAAAA 704
CofOLE-2        ------------------------------------------------------------ 
FpoleL          ------------------------------------------------------------ 
AtOLEO1L        TAATGTTTATGTGTGTTTCTATATGTTGTCAAATGGTACCATGTTTTTTTTCCTCTCTTT 781
CofOLE-4        ATGAGTTGCTGTTTCTCTCAACTCAGCTTGTTCTGCGAGTTTGTCGTTGGATACGTCGTT 699
```

Appendix 2 (continue next page)

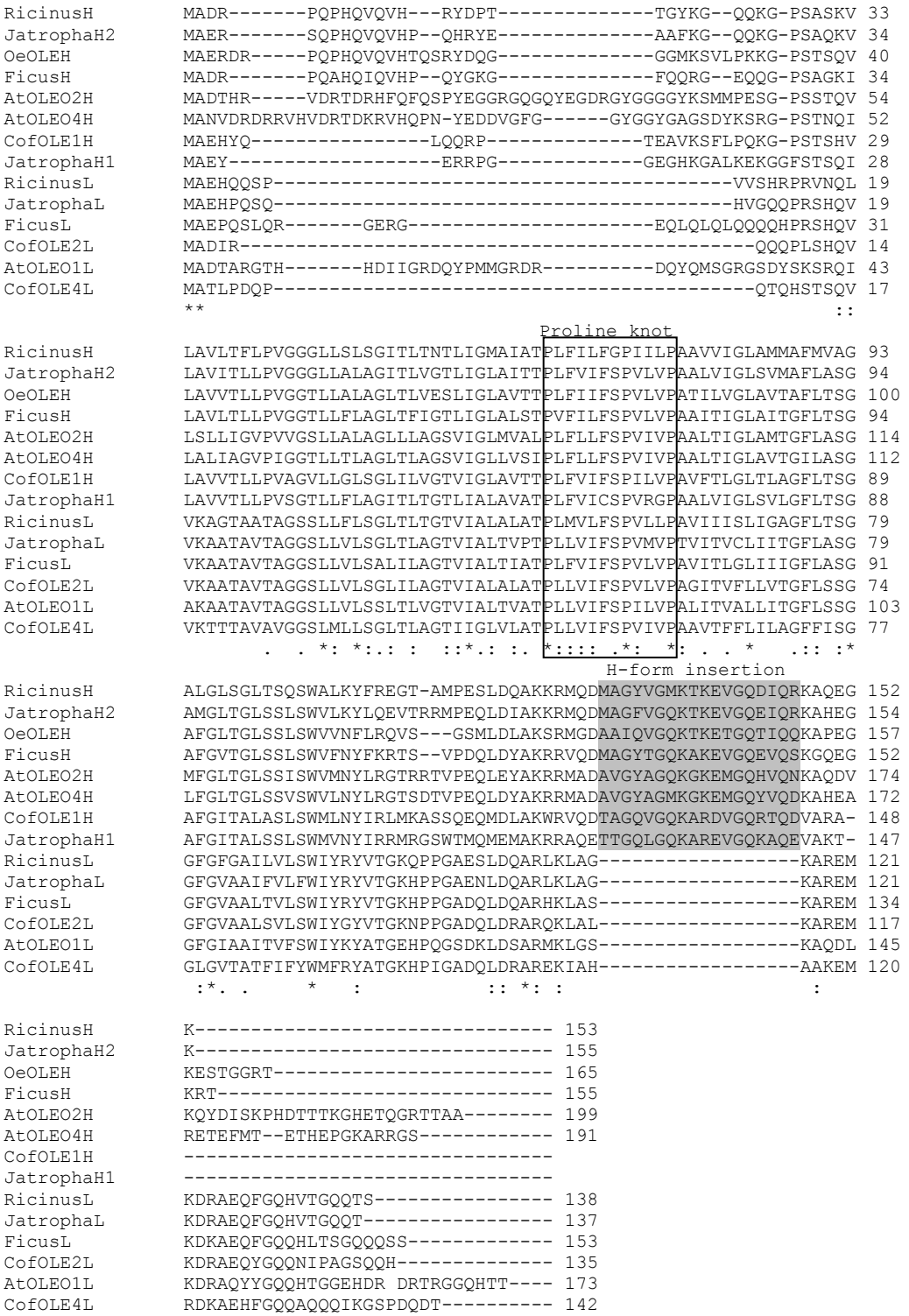
```
JcOLE1H         ------------------------------------------------------------
OeOLH           ACATGCAAGTTGTGGTTTAAGCTCATTTGTGCAATTACTTGATGTTTTTCTTCCTAATTA 760
JcOLE2H         ------------------------------------------------------------
RcOLE1H         ------------------------------------------------------------
FpoleH          ------------------------------------------------------------
CofOLE-1        ------------------------------------------------------------
AtOLEO2H        AGATCGCAAGTGTTGTGGTAATATTTGAATGTAATAATATGATAAGTTGATAAATCATGG 842
AtOLEO4H        A-ATCCTTGTTAGTGAAATAAGAAATGAA-ATAAATATGTTTTCTTTTTTGAGATAACCA 862
JcOLE3          CAAGTCTTACTTTGTTTGTTGATGAATTTGATGTTTCCATACTTTTTGGTTTTAAAAAAA 773
RcOLE2          AAAAAAA---------------------------------------------------- 711
CofOLE-2        ------------------------------------------------------------
FpoleL          ------------------------------------------------------------
AtOLEO1L        TTGTAACTTGCAAGTGTTGTGTTGTACTTTATTTGGCTTCTTTGTAAGTTGGTAACGGTG 841
CofOLE-4        TCATGTGTGCTAATACATTTG-------------------------------------- 720


JcOLE1H         ------------------------------------------------------------
OeOLH           CATTTGCAAATTCAAGTAAAAAAAAAAAAAAA---------------------------- 792
JcOLE2H         ------------------------------------------------------------
RcOLE1H         ------------------------------------------------------------
FpoleH          ------------------------------------------------------------
CofOLE-1        ------------------------------------------------------------
AtOLEO2H        GAACATTTAAATTAGGTGGACATGTTTAGC------------------------------ 872
AtOLEO4H        GAAATCTCATAC------------------------------------------------ 874
JcOLE3          AAATAACAAAAAAAAAA------------------------------------------ 789
RcOLE2          ------------------------------------------------------------
CofOLE-2        ------------------------------------------------------------
FpoleL          ------------------------------------------------------------
AtOLEO1L        GTCTATATATGGAAAAGGTCTTGTTTTGTTAAACTTATGTTAGTTAACTGGATTCGTCTT 901
CofOLE-4        ------------------------------------------------------------


JcOLE1H         -------
OeOLH           -------
JcOLE2H         -------
RcOLE1H         -------
FpoleH          -------
CofOLE-1        -------
AtOLEO2H        -------
AtOLEO4H        -------
JcOLE3          -------
RcOLE2          -------
CofOLE-2        -------
FpoleL          -------
AtOLEO1L        TAACCAC 908
CofOLE-4        -------
```

Appendix 2 Nucleic acid alignment of *J. curcas* oleosins against other plant species, Accession numbers of the aligned oleosin sequences are: NP_194244 (AtOLEO1L), XP_002511014 (RicinusL), ABQ57397 (FicusL), AAX49390 (CofOLE2L), AAX49392 (CofOLE4L), AF302807 (RiceL), AAX49389 (CofOLE1H), NP_198858 (AtOLEO2H), NP_189403 (AtOLEO4H), XP_002521458 (RicinusH), ABQ57396 (FicusH), AAL92479 (OeOLEH), U97700 (RiceH)

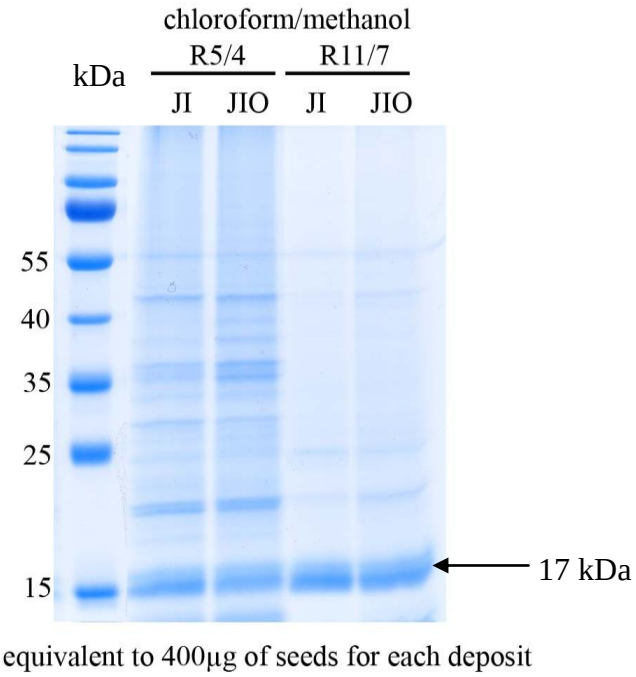
.



Appendix (3A) Expression of *J. curca*s *oleosin* genes in different tissues using semi-quantitative PCR (**3B, B1)** Expression of *J. curca*s *oleosin* genes during fruit development 7, 21 days after fertilization (DAF). **(3B, B2)**Expression of *J. curca*s *oleosin* genes during fruit development 35, 56, and 70 days after fertilization (DAF) and pollen. Please note that gel electrophoresis of B1 and B2 was performed separately. **(3C)** Expression of *J. curca*s *oleosin* genes during seed germination using semi-quantitative PCR. Numbers indicate days after initiation (DAI) implying the time when the seeds were first soaked in water.



Appendix 4 Deduced amino acid alignment *of J. curcas* oleosins against other plant species. Accession numbers of the aligned oleosin sequences are: NP_194244 (AtOLEO1L), XP_002511014 (RicinusL), ABQ57397 (FicusL), AAX49390 (CofOLE2L), AAX49392 (CofOLE4L), AF302807 (RiceL), AAX49389 (CofOLE1H), NP_198858 (AtOLEO2H), NP_189403 (AtOLEO4H), XP_002521458 (RicinusH), ABQ57396 (FicusH), AAL92479 (OeOLEH), U97700 (RiceH).



equivalent to 400µg of seeds for each deposit

Appendix 5 Electrophoresis separation of protein purified with organic solvents on SDS-PAGE (12.5 %). Proteins were stained with Coomassie Brilliant Blue G250.

M*: oxidation of methionine residue (+ 16)
cysteine residues were carbamidomethylated (+ 57)

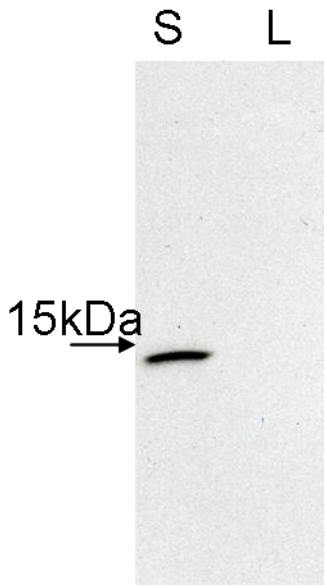
band number	Protein name	Peptide sequence	MH^+	Charge state	Xcorr	Delta Cn	experimental ions/theoretical ions	identified peptides number	identified amino acids number
1	oleosin 1	R.RAQETTGQLGQK.A	1316.69	2	2.796	0.540	15/22	5	24
		R.AQETTGQLGQK.A	1160.59	2	3.489	0.681	16/20		
		R.EVGQKAQEVAKT.-	1287.69	2	2.606	0.447	14/22		
		K.AQEVAKT.-	746.40	2	2.603	0.614	11/12		
		K.AQEVAKT.-	746.40	1	1.544	0.502	8/12		
	oleosin 2	R.RM*PEQLDIAK.K	1216.64	3	2.570	0.307	15/36	4	30
		R.M*PEQLDIAK.K	1060.53	2	2.617	0.514	14/16		
		R.M*QDM*AGFVGQK.T	1243.55	2	3.439	0.594	17/20		
		K.EVGQEIQRK.A	1086.59	2	2.264	0.479	10/16		
	oleosin 3	K.DRAEQFGQHVTGQQT.-	1701.79	2	3.576	0.703	17/28	2	15
		R.AEQFGQHVTGQQT.-	1430.67	2	3.023	0.576	17/24		
2	oleosin 2	R.RM*PEQLDIAK.K	1216.64	2	2.211	0.354	11/18	3	21
		R.M*PEQLDIAK.K	1060.53	2	2.378	0.531	14/16		
		R.M*QDM*AGFVGQK.T	1243.55	2	3.005	0.647	15/20		
3	oleosin 2	K.YLQEVTR.R	908.48	2	2.579	0.285	10/12	4	28
		R.RM*PEQLDIAK.K	1216.64	2	2.289	0.554	12/18		
		R.M*PEQLDIAK.K	1060.53	2	2.871	0.476	13/16		
		R.M*QDM*AGFVGQK.T	1243.55	2	2.947	0.589	16/20		
6	curcin precursor	K.VGGTSYFFNDPESLADAK.T	1917.89	2	3.505	0.590	18/34	6	72
		R.ANVHREDVDLGVQALDNYIYTLEK.S	2775.40	3	5.012	0.615	34/92		
		K.SSKPADIAKPLVGFIEM*VPEAAR.F	2442.31	3	3.978	0.603	25/88		
		K.SSKPADIAKPLVGFIEM*VPEAAR.F	2442.31	2	2.357	0.585	17/44		
		K.PADIAKPLVGFIEM*VPEAAR.F	2140.15	2	2.459	0.583	15/38		
		K.VLSQISK.T	774.47	1	1.636	0.442	7/12		

Figure 6A Mass spectrometry data obtained by conventional LC-MS/MS of in-gel digestion of SDS-PAGE protein bands (see figure 5) from purified oil bodies of *J. curcas* accession India (continue next page)

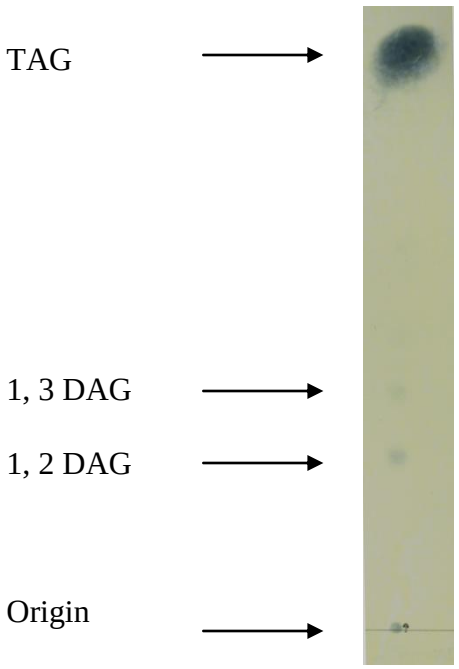
M*: oxidation of methionine residue (+ 16)
cysteine residues were carbamidomethylated (+ 57)

band number	Protein name	Peptide sequence	MH^+	Charge state	Xcorr	Delta Cn	experimental ions/theoretical ions	identified peptides number	identified amino acids number
1	oleosin 2	K.YLQEVTR.R	908.48	2	2.452	0.382	11/12	17	39
		K.YLQEVTR.R	908.48	1	2.273	0.340	10/12		
		R.RMPEQLDIAK.K	1200.64	2	2.841	0.665	13/18		
		R.RMPEQLDIAK.K	1200.64	1	1.621	0.435	8/18		
		R.M*PEQLDIAK.K	1060.53	2	2.729	0.539	14/16		
		R.MPEQLDIAK.K	1044.54	1	2.094	0.413	10/16		
		R.M*QDM*AGFVGQK.T	1243.55	2	2.766	0.545	16/20		
		R.M*QDM*AGFVGQK.T	1243.55	1	1.677	0.359	10/20		
		R.M*QDMAGFVGQK.T	1227.55	2	2.777	0.295	17/20		
		R.MQDM*AGFVGQK.T	1227.55	2	2.775	0.363	16/20		
		R.MQDMAGFVGQK.T	1211.56	2	2.676	0.518	17/20		
		R.MQDMAGFVGQK.T	1211.56	1	2.180	0.577	15/20		
		K.TKEVGQEIQR.K	1187.64	3	2.705	0.481	21/36		
		K.TKEVGQEIQR.K	1187.64	2	3.200	0.506	15/18		
		K.EVGQEIQR.K	958.50	2	2.017	0.332	10/14		
		K.EVGQEIQRK.A	1086.59	2	2.492	0.381	12/16		
		K.EVGQEIQRK.A	1086.59	1	1.624	0.319	10/16		
	oleosin 1	R.GSWTM*QM*EM*AK.R	1347.54	2	3.160	0.483	11/20	11	35
		R.GSWTM*QMEM*AK.R	1331.54	2	2.364	0.247	12/20		
		R.GSWTMQM*EM*AK.R	1331.54	2	2.525	0.337	12/20		
		R.GSWTM*QMEMAK.R	1315.55	2	3.026	0.650	13/20		
		R.GSWTMQM*EMAK.R	1315.55	2	3.558	0.068	14/20		
		R.GSWTMQMEM*AK.R	1315.55	2	3.314	0.150	14/20		
		R.GSWTMQMEMAK.R	1299.55	2	2.622	0.499	16/20		
		R.RAQETTGQLGQK.A	1316.69	2	3.568	0.542	16/22		
		R.AQETTGQLGQK.A	1160.59	2	3.606	0.714	16/20		
		R.EVGQKAQEVAKT.-	1287.69	2	2.267	0.351	14/22		
		R.EVGQKAQEVAKT.-	1287.69	1	1.847	0.617	13/22		
	oleosin 3	R.LKLAGK.A	629.43	1	1.532	0.164	8/10	5	24
		R.EMKDRAEQFGQHVTGQQT.-	2089.97	2	2.011	0.416	12/34		
		K.DRAEQFGQHVTGQQT.-	1701.79	2	2.910	0.704	20/28		
		R.AEQFGQHVTGQQT.-	1430.67	2	3.179	0.634	18/24		
		R.AEQFGQHVTGQQT.-	1430.67	1	3.355	0.594	14/24		
2	oleosin 2	K.YLQEVTR.R	908.48	2	2.247	0.320	11/12	17	39
		K.YLQEVTR.R	908.48	1	1.686	0.224	10/12		
		R.RM*PEQLDIAK.K	1216.64	2	2.210	0.507	12/18		
		R.RMPEQLDIAK.K	1200.64	2	3.308	0.474	15/18		
		R.M*PEQLDIAK.K	1060.53	2	2.732	0.555	13/16		
		R.M*PEQLDIAK.K	1060.53	1	1.537	0.413	10/16		
		R.MPEQLDIAK.K	1044.54	2	2.530	0.423	13/16		
		R.M*QDM*AGFVGQKTK.E	1472.69	3	2.582	0.605	24/48		
		R.M*QDM*AGFVGQK.T	1243.55	2	3.226	0.594	17/20		
		R.M*QDMAGFVGQK.T	1227.55	2	3.117	0.288	18/20		
		R.MQDM*AGFVGQK.T	1227.55	2	2.218	0.651	16/20		
		R.MQDMAGFVGQK.T	1211.56	2	3.005	0.595	17/20		
		K.TKEVGQEIQRK.A	1315.73	2	2.972	0.288	13/20		
		K.TKEVGQEIQR.K	1187.64	3	3.110	0.560	20/36		
		K.TKEVGQEIQR.K	1187.64	2	3.033	0.468	16/18		
		K.EVGQEIQR.K	958.50	1	1.986	0.243	11/14		
		K.EVGQEIQRK.A	1086.59	2	2.330	0.404	11/16		
3	orthologue NP 188487 oleosine SM2 Ath	R.IYDTASHVDYAR.E	1539.67	2	2.612	0.439	15/24	1	12
5	alpha-TIP3;1 Ath	R.AFGPALVGWR.W	1073.59	2	3.158	0.484	15/18	1	10
6	AtCLO1	R.DEEGFLSK.E	924.43	2	2.576	0.253	13/14	1	8
7	curcin precursor	K.NYAQFIK.D	883.47	2	2.174	0.250	10/12	7	75
		K.NYAQFIK.D	883.47	1	1.953	0.326	10/12		
		K.NYAQFIKDLR.E	1267.68	2	2.751	0.085	13/18		
		R.EAFGFSYSSHEIPVLR.A	1838.91	2	3.499	0.541	17/30		
		K.YGGTSYFFNDPESLADAK.T	1917.89	2	2.539	0.464	13/34		
		K.TYLFTDTK.Q	988.50	2	2.226	0.565	12/14		
		K.SSKPADIAKPLVGFIEM*VPEAAR.F	2442.31	3	3.560	0.468	31/88		
9	orthologue coffee SLO	K.AALINFYETLR.F	1311.51	2	2.005	0.035	15/20	1	11

Appendix 6B Mass spectrometry data obtained by conventional LC-MS/MS of in-gel digestion of SDS-PAGE protein bands (see figure 5) from purified oil bodies of *J. curcas* accession Indonesia.

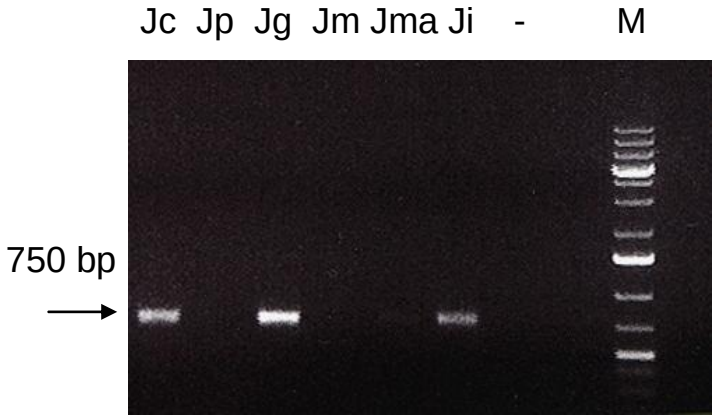


Appendix 7 Immunoassay with Western blots of *JcOleosin3* in *J. curcas* protein extracted from seed and leaves using organic solvent (as described in Materials and Methods).



Appendix 8 Thin Layer Chromatography of oil extracted from *J. curcas* oil bodies. TLC was performed with Merck silica gel 60 F 254 plastic sheets. The sheets were developed with hexane/diethyl ether/acetic acids (80:20:1) and then sprayed with a 5% solution of molybdophosphoric acid in ethanol. The sheets were visualised and quantified using Fluoro spectrophotometer (BioRad). The results showed that the lipid extract from oil bodies are predominantly composed of Triglycerides (Rf = 0.88), 1, 3 Diglycerides (Rf = 0.38), and 1, 2 Diglycerides (Rf = 0.25)

(A)



(B)

```
JinOle3G        --------------------------CGTTGTCCTGTGTTGTTGGTCTCTTCTGTACAGC 34
JmahOle3G       --------------------------CCTTGCCCTGTGTCGCAGGTCGCTGGCCACCAGC 34
Jgosole3G       --------------------------CGGTAGCCTCTGTCGCAGGACGTTGGCCAACAGC 34
JcInOle3G       CCTTTGCTTCGATCGCGTCAATGGCGGAGCACCCACAGTCGCAGCATGTTGGCCAACAGC 60
JcOle3CDS       --------------------ATGGCGGAGCACCCACAGTCGCAGCATGTTGGCCAACAGC 40
                                     **   ** *  *     *      ****

JinOle3G        CACGATCCCACCAAGTAGTGAAGGCAGCAACTGCAGTCACTGCCGGTGGATCTCTTCTAG 94
JmahOle3G       CACGATCCCACCAAGTAGTGAAGGCAGCCACTGCAGTCACTGCCGGTGGATCTCTTCTTG 94
Jgosole3G       CACGATCCCACCAAGTAGTGAAGGCAGCCACTGCAGTCACTGCCGGTGGATCTCTTCTAG 94
JcInOle3G       CACGATCCCACCAAGTAGTGAAGGCAGCCACTGCAGTCACTGCCGGTGGATCTCTTCTAG 120
JcOle3CDS       CACGATCCCACCAAGTAGTGAAGGCAGCCACTGCAGTCACTGCCGGTGGATCTCTTCTAG 100
                ************************** ************************** *

JinOle3G        TTCTCTCCGGTTTGACCCTTGCTGGTACGGTCATTGCACTGACCGTAGCCACTCCATTGC 154
JmahOle3G       TTCTCTCCGGTTTGACCCTTGCTGGTACGGTCATTGCACTGACCGTAGCAACTCCATTGC 154
Jgosole3G       TTCTCTCCGGTTTGACCCTTGCTGGTACGGTCATTGCACTGACCGTAGCCACTCCTTTGC 154
JcInOle3G       TTCTCTCCGGTTTGACCCTTGCTGGTACGGTCATTGCACTGACCGTAGCCACTCCATTGC 180
JcOle3CDS       TTCTCTCCGGTTTGACCCTTGCTGGTACGGTCATTGCACTGACCGTACCCACTCCATTGC 160
                ********************************************* * ***** ****

JinOle3G        TGGTCATATTCAGTCCGGTTCTTGTTCCTGCGGTTATCACTGTTTGTTTGATAATCACAG 214
JmahOle3G       TGGTCATATTCAGTCCGGTTCTTGTTCCAGCGGTTATCACAGTTTGTTTGATAATCACAG 214
Jgosole3G       TGGTCATATTCAGTCCGGTTCTCGTTCCTGCGGTTATCACAGTTTGTTTGATAATCACAG 214
JcInOle3G       TGGTCATATTCAGTCCGGTTCTTGTTCCTGCGGTTATCACAGTTTGTTTGATAATCACAG 240
JcOle3CDS       TGGTCATATTCAGTCCGGTGATGGTGCCGACGGTTATCACAGTTTGTTTGATAATCACAG 220
                ******************  * ** **  ********** ******************

JinOle3G        GGTTTTTAGCCTCAGGCGGTTTTGGCGTCGCTGCCATTTCTGTCTTGTCATGGATTTATA 274
JmahOle3G       GGTTTTTAGCCTCAGGCGGTTTTGGCGTCGCTGCCATTTCTGTCTTGTCATGGATTTATA 274
Jgosole3G       GGTTTTTAGCCTCAGGCGGTTTCGGCGTCGCTGCCATTTCTGTCTTGTCATGGATTTATA 274
JcInOle3G       GGTTTTTAGCCTCAGGCGGTTTTGGCGTCGCTGCCATTTCTGTCTTGTCCTGGATTTATA 300
JcOle3CDS       GGTTTTTAGCCTCAGGCGGTTTTGGGGTCGCTGCCATTTTTGTCTTGTTTTGGATTTATA 280
                ********************** ** ************ ********  **********

JinOle3G        GGTTCAAA-GAAATTTGCCTACTTCCTTTCATTTCTTTCAACACACACTATTAAAAAGAC 333
JmahOle3G       GGTTCAAA-GAAATTTGCCTACTTCCTTTCATTTCTTTCAACACATAATATTAAAAAGAC 333
Jgosole3G       GGTTAAAAAGAAATTTGCCTACTTTCTTTCATTTCTTTTAACGCATAATATTAAAAAGAC 334
JcInOle3G       GGTTTTAA-GAAATTTGCCTACTTCCTTTCATTTCTTTCAACGCATAATATTAAAAAGAC 359
JcOle3CDS       GGT------------------------------------------------------- 283
                ***

JinOle3G        AGTGTTTGACATTGGAGACATGAATTGCAGGTATGTTACTGGAAAGCATCCACCAGGAGC 393
JmahOle3G       AATGTTTGACATTTGAGACATGAATTGCAGGTATGTCACTGGAAAGCATCCACCAGGAGC 393
Jgosole3G       AATGTTTGACATTGGAGACACGAATTGCAGGTATGTCACTGGAAAGCATCCACCAGGAGC 394
JcInOle3G       AATGTTTGACATTGGAGACATGAATTGCAGGTATGTTACTGGAAAGCATCCACCAGGAGC 419
JcOle3CDS       -------------------------------ATGTTACTGGAAAGCATCCCCCAGGAGC 311
                                              **** ************** ********
```

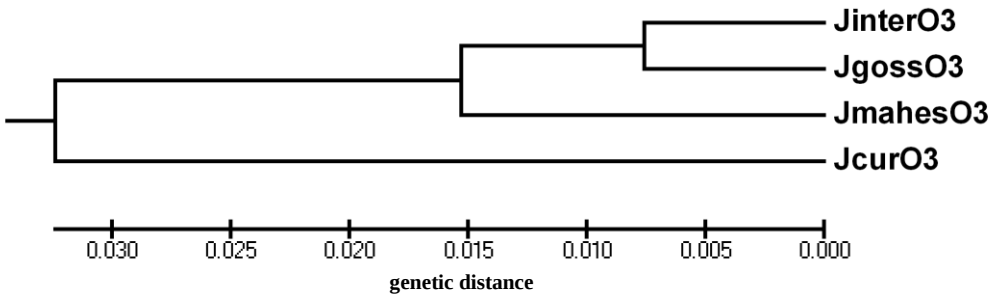
Appendix 9 (continue next page)

```
JinOle3G        AGAAAGTCTGGACCAGGCACGCTTGAAGTTGGCGGGAAAGGCAAGAGAGATGAAGGACAG 453
JmahOle3G       AGAAAGTCTGGACCAGGCACGCTTGAAGTTGGCGGGAAAGGCAAGAGAGATGAAGGACAG 453
Jgosole3G       AGAAAGTCTGGACCAGGCACGCTTGAAGTTGGCGGGAAAGGCAAGAGAGATGAAGGACAG 454
JcInOle3G       AGAAAGTCTGGACCAGGCACGCTTAAAGTTGGCGGGAAAGGCAAGAGAGATGAAGGACAG 479
JcOle3CDS       AGAAAATCTGGACCAGGCACGCTTAAAGTTGGCGGGAAAGGCAAGAGAGATGAAGGACAG 371
                ***** ****************** **********************************

JinOle3G        AGCTGAGCAGTTTGGACAGCATGTAACAGGTCAACAGACTTAAAAGAGAAGTGGGTTTTG 513
JmahOle3G       AGCTGAGCAGTTTGGACAGCATGTAACAGGTCAACAGACTTAAAAGAGAAGTGGGTTTTG 513
Jgosole3G       AGCTGAGCAGTTTGGACAGCATGTAACAGGTCAACAGACTTAGAAGAGAAGTGGGTTTTG 514
JcInOle3G       AGCTGAGCAGTTTGGACAGCATGTAACAGGTCAACAGACTTAAAAGAGAAGTGGGTTTTG 539
JcOle3CDS       AGCTGAGCAGTTTGGACAGCATGTAACAGGTCAACAGACTTAA----------------- 414
                ****************************************

JinOle3G        GTGG-TTGATGGAGAGAACATAAAATGGCTTC----CATATATATATATAAACATATATA 568
JmahOle3G       GAGG-TTGATGGAGAGAACATAAAATGGCTTTGATATATATATATATATAAAAATATATA 572
Jgosole3G       GTGGGTTGATGGAGAGAACACAAAATGGCTTTG------ATATATATATAAACA----TA 564
JcInOle3G       GTGG-TTGATGCAGAGAACATAAAATGGCTTTG------ATATATATATAAAAATATATG 592
JcOle3CDS       ------------------------------------------------------------

JinOle3G        AGTCCCTTAATGAACATGAAAGGCCATTGCTTGTCTGGTGTATTTCTTTTTTTG-CTTAA 627
JmahOle3G       AGTCCTTTAATTAACATGGAAGGCCATTGCTTATCTGGTGTATTTCTTTTTTTG-CATAA 631
Jgosole3G       AGTCCTTTAATTAACATGGAAGGCCATTGCTTGTCTGGTGTATTTCTTTTTTTGGCTTAA 624
JcInOle3G       AGTCCTTTAATTAACATGGAAGGCCATTGCTTGTCTGGTGTATTTCTTTTTTCGCTTAAA 652
```

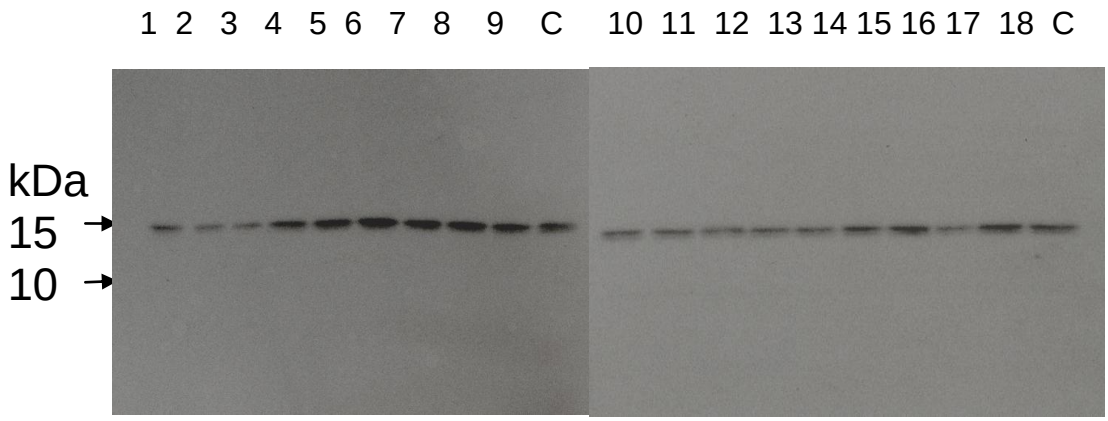
(C)

```
JmahesO3AA      ----PVSQVAGHQPRSHQVVKAATAVTAGGSLLVLSGLTLAGTVIALTVATPLLVIFSPV 56
JgossO3AA       ----PLSQDVGQQPRSHQVVKAATAVTAGGSLLVLSGLTLAGTVIALTVATPLLVIFSPV 56
JinterO3AA      ----PQSQQVGQQPRSHQVVKAATAVTAGGSLLVLSGLTLAGTVIALTVATPLLVIFSPV 59
JcurO3AA        ----PQSQHVGQQPRSHQVVKAATAVTAGGSLLVLSGLTLAGTVIALTVPTPLLVIFSPV 60
                    * **  .*:************************************.**********

JmahesO3AA      LVPAVITVCLIITGFLASGGFGVAAISVLSWIYRYVTGKHPPGAESLDQARLKLAGKARE 116
JgossO3AA       LVPAVITVCLIITGFLASGGFGVAAISVLSWIYRYVTGKHPPGAESLDQARLKLAGKARE 116
JinterO3AA      LVPAVITVCLIITGFLASGGFGVAAISVLSWIYRYVTGKHPPGAESLDQARLKLAGKARE 119
JcurO3AA        MVPTVITVCLIITGFLASGGFGVAAIFVLFWIYRYVTGKHPPGAENLDQARLKLAGKARE 120
                :**:********************** ** ***************.**************

JmahesO3AA      MKDRAEQFGQHVTGQQT 133
JgossO3AA       MKDRAEQFGQHVTGQQT 133
JinterO3AA      MKDRAEQFGQHVTGQQT 136
JcurO3AA        MKDRAEQFGQHVTGQQT 137
                *****************
```

(D)



Appendix 9 (A) PCR products amplified using JcFLO3F and JcFLO3R primers of *J. curcas* (Jc), *J. podagrica* (Jp), *J. gysipifolia* (Jg), *J. mutifida* (Jm), *J. maheswarii* (Jm), *and J. interegima* (Ji). (B) Nucleic acid and (C) amino acids alignments (D) Phylogenetic tree constructed from the multiple alignment of deduced amino acid of *Oleosin3* genes obtained from *J. integerima* (JinterO3), *J. gysipifolia* (JgossO3), *J. maheswarii* (JmahesO3) *J. curcas* (JcurO3).



Appendix 10 Immunological detection of chloroform/methanol 11/7 seed extracted protein of *J. curcas* using antiserum against *JcOleosin* 3. Samples used in this studied are *J. glandulifera* (S1), *J. multifida* (S2), *J. podagrica* (S3), *J. curcas* non-toxic Mexican accession (S4), *J. curcas* Mettupalem accession (S5), and *J. curcas* Uganda accession (S6), Hybrids clone of *J. curcas* and *J. interegima* (Hybrid clone 91-1 (S7), 92 (S8), 96 (S9), 111 (S10), 201 (S11), 202 (S12), 205 (S13), 213 (S14), 222 (S15), 223 (S16), 225 (S17), and 226 (S18), and *J. curcas* (control).

Chapter 4: Characterization of the *Jatropha, JcOleosin3* promoter in transgenic *Arabidopsis*

4.1 Abstract

The *Jatropha* (*Jatropha curcas*) oleosin gene, *JcOleosin3*, produces a 14-kDa protein that is highly expressed in seeds. In this study, the activity of the 623 bp 5'-upstream regulatory region of this gene was determined in transgenic *Arabidopsis* plants using *β-glucuronidase* (GUS) as a reporter gene. Results showed the *JcOleosin3* promoter strongly directs expression of the *β-glucuronidase* gene in seed but not in leaves, root, stem, and silique wall. The *JcOleosin3* promoter also drives GUS activity during late pollen development and during early seedling development.

Keywords: *Jatropha curcas, JcOleosin3* promoter, *Arabidopsis*, *β*-glucuronidase

4.2 Introduction

Oilbodies consist of 94-98% lipids, 0.5-2% phospholipids and 0.5-3.5% proteins (Chen et al., 2004). Between 80-90% of the protein contents belong to the class of proteins known as oleosins (Roberts et al., 2008), which cover the entire surface of the oilbody and dictate its size (Capuano et al., 2007). Oleosins were first identified in cotyledons of sunflower and safflower by Slack et al. (1980). They are low-molecular weight amphipathic proteins of about 15-26 kDa, predominantly found in the seeds of oilseed plants. Oleosins are now classified into low and high molecular weight isoforms based on their molecular weights (Tzen et al., 1990). Oleosins are ubiquitous in oil storing seeds of monocots or dicots including oilseed rape, maize, cotton, tea, coffee, soybean, sunflower, safflower and rice (Capuano et al., 2007).

The *J. curcas* is becoming an increasingly popular plant (Fearless, 2007; Nature editorial 2007) for its proposed value in the biodiesel, biopharmaceuticals, cosmetics and biopesticides industries (Gubitz et al., 1999; Kumar & Sharma, 2008; Kohli et al., 2009). Given meager genotypic characterization, sub-optimal agronomic practices, limited information on the genome of *J. curcas* and a scarce number of its isolated genes is required major research initiatives in agronomy, breeding and molecular biotechnology of *J. curcas* for it to live up to its potential (Popluechai et al., 2008). Gressel (2008) recommended a transgenic approach for improvement of biofuel crops including *J. curcas*. *Agrobacterium* mediated transformation of *J. curcas* was recently reported (Li et al., 2008), allowing transgenic *J. curcas* to be generated for the desired traits.

Preferentially seed targeted oleosin promoter is a very useful tool for seed specific gene manipulation/expression. Oleosin promoters from several plants, for example *Perilla* (Chung et al., 2008), coffee (Simkin et al., 2006), and *Brassica napus* (Keddie et al., 1994) have been characterized. Genetic modification of desired phenotypes in *J. curcas* seed may depend on being able to limit or increase the expression of the genes of interest specifically in the seed. Towards this goal, seed targeted promoters play an important role. In *J. curcas*, only ribosome-inactivating protein (RIPs) promoters have been studied (Qin et al., 2009a; Qin et al., 2009b). In the present study, 623 nucleotides upstream of the *JcOleosin3* gene isolated from *J. curcas* has been isolated and characterized in transgenic *Arabidopsis* plant via *β-glucuronidase* (GUS) as a reporter gene.

4.3 Materials and Methods

Plant material and growth conditions

Genomic DNA of *J. curcas* was extracted from accession India (Hyderabad, India). *Arabidopsis thaliana* wild type (Col-0) and transgenic lines were grown in pots containing a mixture of vermiculite and soil (1/4, v/v) or on plates with MS medium (Murashige and Skoog, 1962) and were kept in a growth room at 22℃ and 70% humidity with a 16-h light/8-h dark cycles. The transgenic line (JcOleosin3::GUS) was in the Columbia (Col-0) background; Columbia was also used as the wild-type control. The transgenic *Arabidopsis* was selected using BASTA (phosphinothricin) selection. The selection is herbicide (phosphinotricin) resistance selection. Only transgenic plant harbouring *bar* gene (phosphinotricin resistance gene) will be survived under the selection.

Isolation of upstream regulatory element from the J. curcas genomic DNA

An upstream sequence of *J. curcas* oleosin was determined using DNA Walking SpeedUp™ kit (Biogene, United Kingdom) following the procedures described by the manufacturers. Briefly, the DW-ACP primer set from Seegene kit was used with a series of gene specific primers (JcLMW1-TSP1: 5' GAAGAAAGCGCAGTGATTCC 3'; JcLMW1-TSP2: 5' CTAGCGCAATGAGTGTTCCA 3'; JcLMW1-TSP3: 5'ACCGGCAAAAGAGTGACAAC 3') to extend the sequence as obtained by using the above procedure. The DNA walking reaction was performed using the *Taq* polymerase PCR kit (Fermentas) and conditions as previously described (Chapter 3, Materials and Methods). The sequences obtained from various amplification steps were assembled in a single contig. The sequence was analyzed as previously described (Chapter 3, Materials and Methods).

Recombinant vector construction

An upstream regulatory region consisting of 623 nucleotides, of the *J. curcas oleosin3* gene was amplified using PCR mediated with the JcO3PKpn-F 5' ACGGGGTACCCAGCAGCAGAGCTTTCAAGATG 3' and JcO3PXho-R 5' CCGCTCGAGCGTGATCGAGTTGTGAAATTACG 3'. The PCR product was digested with *Kpn*I (Fermentas) and *Xho*I (Fermentas) before being cloned into pGreen0229-GUS (Gift from Dr. Maria Manuela R. Costa), designated pJcO3P623::GUS (see Figure 1 for the physical map of the T-DNA cassette). The recombinant vector was confirmed by restriction enzymes analysis and DNA sequencing.

Plant transformation

The pJcO3P623::GUS vector was used to transform *Agrobacterium tumefaciens* strain C58C1 pGV101 pMP90 by electroporation. *Arabidopsis* plants were then transformed by floral dipping (Clough and Bent, 1998). BASTA resistant transformants were isolated after spraying T1 plants with 100 ug/L of BASTA and confirmed by PCR.

Histochemical GUS assay

Histochemical GUS assays using X-Gluc (5-Bromo-4-chloro-3-indoxyl-beta-D-glucuronide) as a substrate (Jefferson et al., 1987) were performed on seeds and seedlings at different times after imbibition/germination (0 h, 1 h, 2 h, 4 h, 8 h, 16 h, 1 d, 2 d, 5 d, 10 d) as well as siliques in different maturation stages and other several organs from 30-day old *Arabidopsis* plants.

4.4 Results

Isolation and analysis of the 5' regulatory regions of the JcOleosin3 gene

A promoter sequence of 623 nucleotides upstream of the *JcOleosin3* gene was isolated (GenBank accession no. EU543568) by PCR assisted gene walking as described in the Materials and Methods. This sequence was then analyzed for cis-acting regulatory elements (CAREs) using the web-based PLACE programme (http://www.dna.affrc.go.jp/PLACE/; Higo et al., 1999). The analysis identified motifs typical of promoters (Table 1). For example, a TATA box (TATAAAT) and a CAAT box (CCAAT) was located at -26 and -51 bp upstream of the *JcOleosin3* transcription start site (TSS) respectively. Other motifs known for seed-specific gene expression were also noted and these are described in Table 1. These include: the sequence ACACNNG (-68 to -73) which known as the seed-specific DPBFCOREDCDC3 motif, is recognised by bZIP transcription factors induced by abscissic acid (Kim et al., 1997). The sequence AAAG (-276 to -279, -420 to -424, -570 to -573, -575 to -578) known as the seed-specific DOFCORE motif, is recognised by Dof transcription factors involved in endosperm specific expression (Yanagisawa S, 2000). The sequence CANNTG (-68 to -73, -393 to -398,-458 to -463,-502 to -507, -537 to -542) known as the E-box is involved in seed–specific expression of the napin protein of *Brassica napus* (Stalberg et al, 1996). The sequence RTTTTTR (-185 to -191) known as the SEFMOTIFGM75 motif is a soybean embryo factor 4 (SEF4) binding site (Lessard et al, 1994). The sequence CATGCATG (-119 to -125, -535 to -540) known as the RY repeat box lends seed-specific expression to the legumin gene in *Vicia faba* (Baumelin et al., 1992). A late pollen-specific motif [AGAAA; (-572 to -576; Bate and Twell, 1998)] was also identified.

Characterization of transgenic Arabidopsis

To investigate the *623-JcOleosin3* gene promoter, the upstream regulatory element was fused to the GUS reporter gene as described in Material and Methods (Figure 1). Twelve positive clones of *Arabidopsis* transgenic (T0) were obtained after BASTA selection. The T0 *Arabidopsis* transgenic plants were grown to obtain homozygous transgenic plants (T1). Segregation analysis of BASTA resistance identified transgenic

homozygous T1 lines and PCR analysis was used to confirm the presence of the transgene and homozygosity (Dr. Maria Manuela R. Costa; personal communication).

Tissue-specific accumulation of GUS driven by the JcOleosin3 promoter

β-glucuronidase histochemical analysis was performed in transgenic *Arabidopsis* leaves, stem, root, silique, flower, and mature seed. GUS expression was detected during late pollen development including in the pollen tube (Figure 2A-2B). Strong GUS expression was detected only in seed but not in the silique wall (Figure 2C-2E). No GUS signal could be detected in other tissues (Supplemental Figure 2). These results suggest that the 623 bp upstream regulatory element of the *JcOleosin3* gene is functionally active during late pollen development and in the seed.

Accumulation of GUS driven by the JcOleosin3 promoter in transgenic seedlings

Studies carries out during the course of this thesis showed that during germination the transcript for *JcOleosin3* (L-form) remained detectable until 21 DAI (Chapter 3, Figure 3b). The presence of the *JcOleosin3* transcript in *J. curcas* seedlings suggested that a transgenic approach could be used to study the role of the *JcOleosin3* promoter. To this end GUS activity under the control of *JcOleosin3* promoter was investigated during transgenic *Arabidopsis* seedling development. Results showed that *Arabidopsis* seedlings showed strong GUS expression 2 days after imbibition (DAI), but not before (Figure 3A). The strong expression was detected only in the cotyledons but not in the true leaves (Figure 3B and 3C). Furthermore, expression studies also demonstrated GUS activity in the root tip (Figure 3E), but that this disappeared 10 days post imbibition. These results indicate that whilst the 623-*JcOleosin3* promoter drives gene activity, this is limited to the early seedling stage.

Table 1 Seed specific and transcription initiation motifs of the 623-JcOleosin3 regulatory element (EU543568) predicted by using PLACE programme

Motif	Function	Publication	Location
TATA (TATAAAT)	critical for accurate initiation of transcription	Grace et al, 2004	-26 to -32
CCAAT	critical for accurate initiation of transcription	Rieping et al, 1992	-47 to -51
DPBFCOREDCDC3 (ACACNNG)	bZIP transcription factors inducing by ABA	Kim et al, 1997	-68 to -73
ABRE motif (MACGYGB)	ABA responsive element	Nakashima et al, 2006	-149 to -156
SEFMOTIFGM75 (RTTTTTR)	SEF4 binding site	Lessard et al , 1994	-185 to -191
DOFCORE (AAAG)	Dof proteins binding site	Yanagisawa *S*, 2000	-276 to -279 -420 to -424 -570 to -573 -575 to -578
GATABOX (GATA)	Binding with ASF-2	Reyes et al, 2004	-308 to -311 -314 to -317 -367 to -370 -387 to -390 -464 to -468
EBOXBNNAPA (CANNTG)	R response element (E box)	Stalberg et al, 1996	-68 to -73 -393 to -398 -458 to -463 -502 to -507 -537 to -542
MARTBOX (TTWTWTTWTT)	T box scaffold attachment region	Gasser et al, 1989	- 470 to -479
RY repeat box (CATGCATG)	seed specific expression response to ABA	Reidt et al., 2000	-119 to -125 - 535 to - 540
Late pollen-specific motif (AGAAA)	Require for late pollen specific expression	Bate et al, 1998	-572 to -576

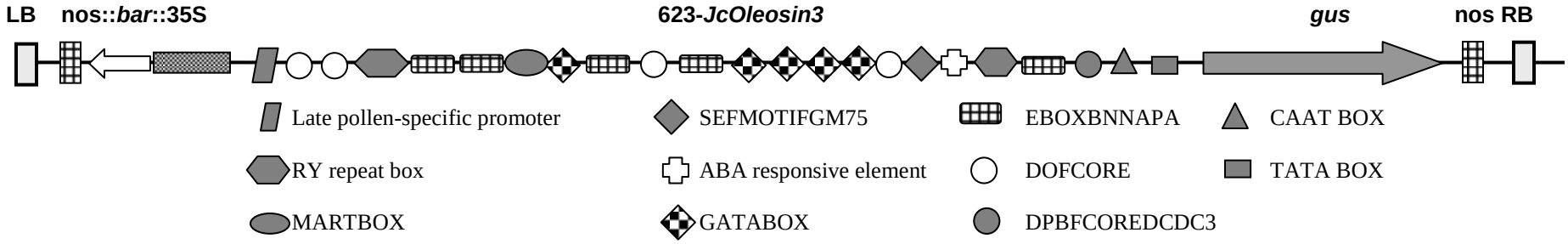


Figure 1 Structure of the 623-*JcOleosin3* promoter with glucuronidase reporter gene (*gus*) in the binary vector pGreen0229. LB, left border of the transfer DNA; 35S::*bar*::nos cassette (phosphinothricin resistance gene under 35S promoter); 623-JcOleosin3::*gus*::nos (glucuronidase reporter gene under 623-*JcOleosin3* promoter); nos, nopaline synthase terminator; RB, right border of the transfer DNA.

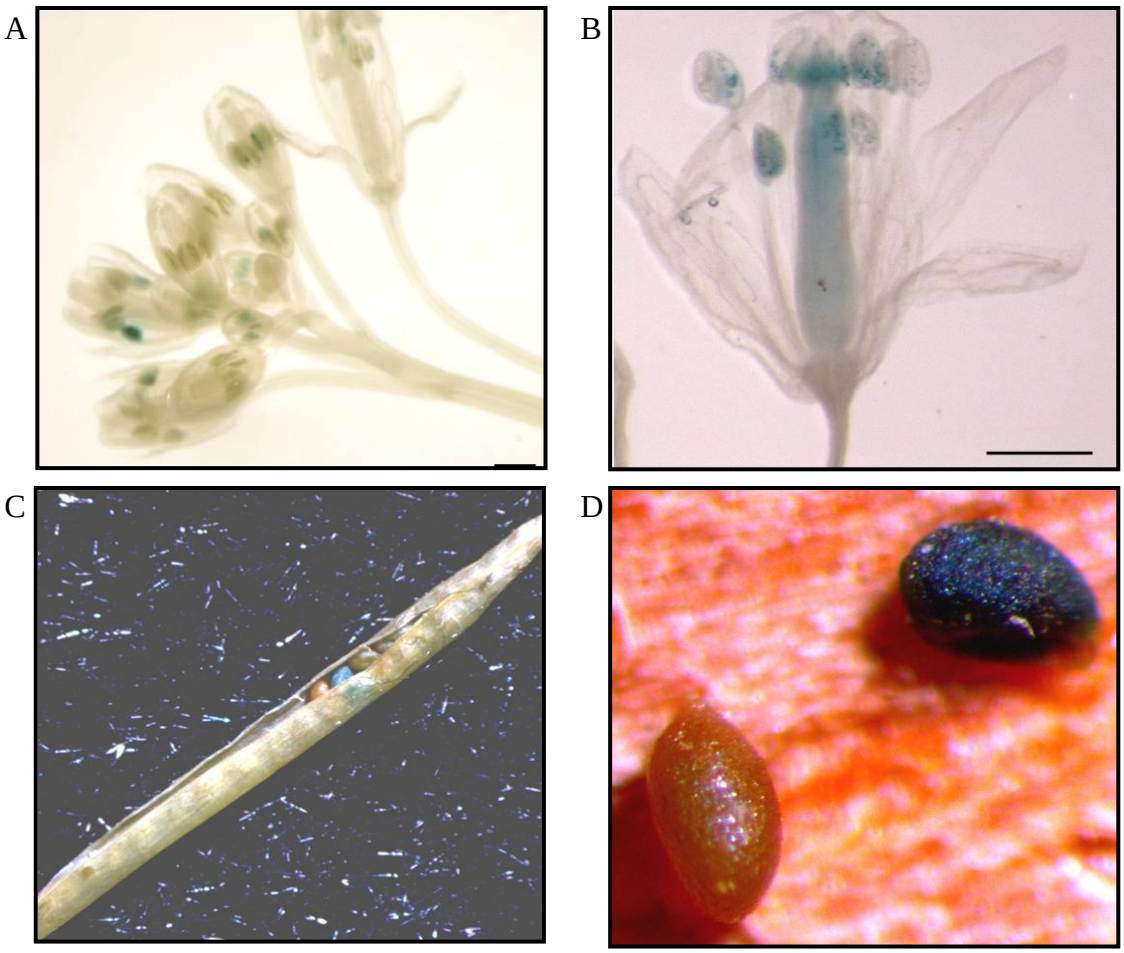


Figure 2 Histological analysis of GUS activity in T1 transgenic *Arabidopsis* during fruit development. (A) GUS activity in the inflorescence; (B) GUS activity in stamens of a fully expanded flower; (C) GUS expression was specific to the mature seeds; (D) GUS activity in segregated mature seeds of transgenic *Arabidopsis*.

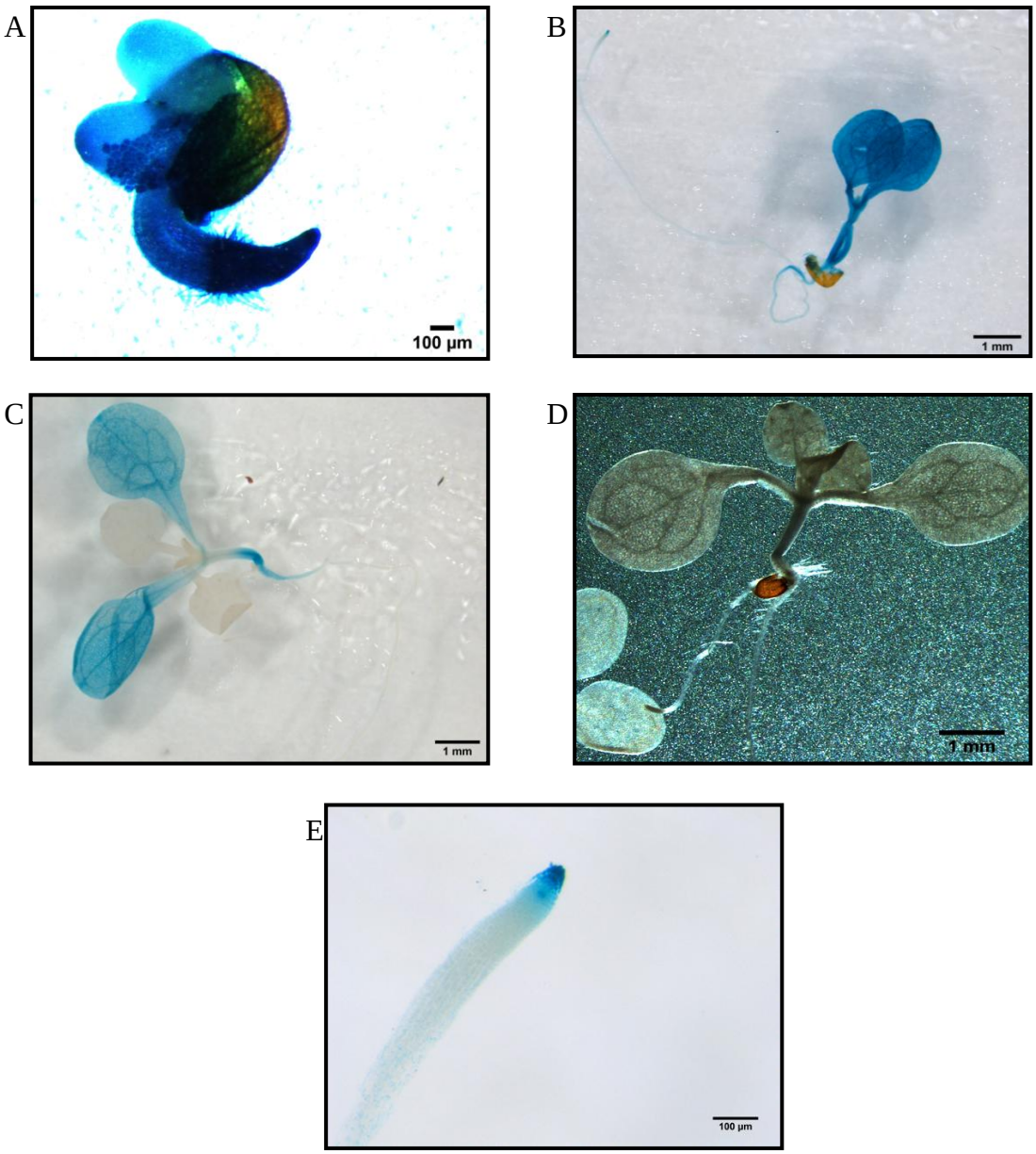


Figure 3 Expression of GUS in transgenic *Arabidopsis* seedlings. Histological analysis of GUS activity during seedling development of transgenic *Arabidopsis* at 2 days (A); 5 days (B); 10 days after imbibitions (DAI) (C); control *Arabidopsis* at 10 DAI (D); Gus activity in the root cap of 10 DAI of transgenic *Arabidopsis* (E).

4.5 Discussion

Oleosins are known to express preferentially in the seeds of oilseed plants. Our studies confirmed this with results that the *JcOleosin3* transcript and protein was predominantly expressed in the *J. curcas* seed (Chapter 3, Figure 3). The upstream regulatory sequence of *J. curcas oleosins3* was characterized. A sequence of 623 nucleotides upstream of the *JcOleosin3* gene was isolated (EU543568) from the genome of *J. curcas* (India accession). *In silico* analysis using PLACE programme indicated several seed specific motifs such as the DOFCORE motif, the DPBFCORE motif, the SEF4 motif, the GAT-box motif, the E-box motif, the MARTbox motif, and the RY repeat motif (Table 1). Similar motifs have also been found in oleosin promoters from several plants such as *Perilla frutescens* L. cv. *Okdong* (Chung et al., 2008), *Coffee canephora* (Simkin et al., 2006), and *Brassica napus* (Keddie et al., 1994). NanYing et al. (2009) demonstrated the crucial role of two neighboring RY elements as LEC2 binding domain for the expression of oleosin. Recently, the *J. curcas* curcin gene promoter was reported to contain similar seed-specific motifs such as the DOFCORE and E-box. Seed-specific expression of this promoter was validated in transgenic tobacco plants (Qin et al., 2009a). The authors found that the 0.62-kb *J. curcas* curcin promoter is specifically active in endosperm tissue and initiated at the heart-shaped embryonic stage during tobacco seed development. The authors also found that DOFCORE and E box motifs are important for the regulation of expression in the endosperm. Results from the present study demonstrated that the 623-*JcOleosin3* promoter exhibits DOFCORE and E box motifs similar to the *J. curcas* curcin promoter. However, whether these two motifs in the 623-*JcOleosin3* promoter play a crucial role like the *J. curcas* curcin promoter remains to be investigated. Additionally, the stress-inducible-curcin-L promoter has also been isolated and characterized in transgenic tobacco (Qin et al., 2009b)

Results from this study show that the 623-*JcOleosin3* promoter drives GUS expression during late pollen development in transgenic *Arabidopsis* (Figure 2A-2B). These results were supported by the presence of the late pollen-specific motif (pollenlate52 motif; Bate and Twell, 1998). This motif plays an important role by controlling transcriptional activation in the mature pollen of tomato.

The finding that the 623-*JcOleosin3* promoter strongly drives GUS expression only in seed, but not in the silique wall (Figure 2C-2E) is supported by the presence of several seed specific motifs in this promoter. Interestingly, the *Perilla* oleosin 19 promoter also exhibited the same pattern of expression (Chung et al., 2008). These results suggest that the 623-*JcOleosin3* promoter is preferentially seed targeted promoter and might be a good candidate for controlling seed preferentially gene expression in the studies with transgenic *J. curcas*.

The 623-*JcOleosin3* promoter also drives GUS expression in the cotyledon but not in true leave (Figure 3C). Interestingly, the *Perilla* oleosin 19 promoter again also exhibited the same pattern of expression (Chung et al., 2008). This result might explain the expression of *JcOelsoin3* gene in *Jatropha* leave (Chapter 3, Figure 2).

Moreover, in this study the 623-*JcOleosin3* promoter also drives GUS expression in the root tip of transgenic *Arabidopsis* seedlings. Murphy et al. (2001) and Karimi et al. (2002) observed the transient expression of GUS in the root tip of transgenic *Arabidopsis* seedlings when using oleosin and pollenin promoters, respectively. Murphy et al. (2001) speculated that the root tips may accumulate lipid oilbodies for use as gravity sensors.

Regulation of gene expression of 623-*JcOleosin3* promoter in transgenic *Arabidopsis* might be different to *Jatropha*. Whether regulation of gene expression of 623-*JcOleosin3* promoter in *Jatropha* similar to that in transgenic *Arabidopsis* remains to be confirmed.

In summary, we have isolated and *in silico*/*in vivo* characterized the *J. curcas* 623-*JcOleosin3* promoter. Although deletion analysis needs to be investigated in order to confirm the role of each motif in the *JcOleosin3* promoter, our study demonstrated that the 623 nucleotide upstream of the regulatory element of the *JcOleosin3* gene is sufficient to act as a preferentially seed targeted promoter for subsequent gene manipulation studies, for example silencing of the genes involved in the biosynthesis of various toxins, the presence of which makes *J. curcas* seed oil less than ideal for biodiesel. Furthermore, the fatty acids profiles (quality/quantity) could also be manipulated in the near future using this promoter.

4.6 References:

Bate, N. and Twell, D. (1998) 'Functional architecture of a late pollen promoter: pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements', *Plant Molecular Biology*, 37, (5), pp. 859-869.

Bäumlein, H., Nagy, I., Villarroel, R., Inzé, D. and Wobus, U. (1992) '*Cis*-analysis of a seed protein gene promoter: the conservative RY repeat CATGCATG within the legumin box is essential for tissue-specific expression of a legumin gene', *The Plant Journal*, 2, pp. 233-239.

Capuano, F., Beaudoin, F., Napier, J.A. and Shewry, P.R. (2007) 'Properties and exploitation of oelosins', *Biotechnology Advances*, 25, pp. 203-206.

Chen, M.C.M., Chyan, C.L., Lee, T.T.T., Huang, S.H. and Tzen, J.T.C. (2004) 'Constitution of stable oil artificial oilbodies with triacylglycerol, phospholipids and caleosin', *J .Agric. Food Chem.*, 52, pp. 3982-3987.

Chung, K., Hwang, S., Hahn, B., Kim, K., Kim, J., Kim, Y., Yang, J. and Ha, S. (2008) 'Authentic seed-specific activity of the *Perilla* oleosin 19 gene promoter in transgenic *Arabidopsis*', *Plant Cell Reports*, 27, pp. 29-37.

Clough, S.J. and Bent, A.F. (1998) 'Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*', *The Plant journal*, 16, pp. 735-743.

Fairless, D. (2007) 'Biofuel: The little shrub that could maybe', *Nature*, 449, pp. 652-655.

Grace, M.L., Chandrasekharan, M.B., Hall, T.C. and Crowe, A.J. (2004) 'Sequence and spacing of TATA box elements are critical for accurate initiation from the beta-phaseolin promoter', *Journal of Biological Chemistry*, 279, pp. 8102-8110.

Grasser, K.D., Maier, U.G. and Felix, G. (1989) 'A nuclear casein type-II kinase from maize endosperm phosphorylating HMG proteins', *Biochemical and Biophysical Research Communications*, 162, pp. 456-463.

Gressel, J. (2008) 'Transgenic are imperative for biofuel crops', *Plant Science*, 174, pp. 246-263.

Gubitz, G.M., Mittelbach, M. and Trabi, M. (1999) 'Exploitation of the tropical oil seed plant *Jatropha curcas* L', *Bioresource technology*, 67, pp. 73-82.

Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. (1999) 'Plant cis-acting regulatory DNA elements (PLACE) database:1999', *Nucleic Acids Research*, 27, pp. 297-300.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) 'GUS fusions: b-glucuronidase as a sensitive and versatile gene fusion marker in higher plants' *EMBO J.*, 6, pp. 3901-3907.

Karimi, M., de Oliveira Manes, C.L., Van Montagu, M. and Gheysen, G. (2002) 'Activation of a Pollenin Promoter upon Nematode Infection', *J Nematol*, 34, pp. 75-79.

Keddie, J.S, Tsiantis, M., Piffanelli, P., Cella, R., Hatzopoulos, P. and Murphy, D.J. (1994) 'A seed-specific *Brassica napus* oleosin promoter interacts with a G-box-specific protein and may be bidirectional', *Plant Molecular Biology*, 24, pp. 831-831.

Kim, S.Y., Chung, H.J. and Thomas, T.L. (1997) 'Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the Dc3 promoter using a modified yeast one-hybrid system', *Plant Journal* 11, pp.1237-1251.

Kohli, A., Raorane, M., Popluechai, S., Kannan, U., Syers, K.J., O'Donnell A.G. (2009) '*Jatropha curcas* as a novel, non-edible oilseed plant for biodiesel', In: Ferry N, Gatehouse AMR, eds. *Environmental impact of genetically modified novel crops*. Chapter 14. CAB International, London, UK. pp. 294 -322.

Kumar, A. and Sharma, S. (2008) 'An evaluation of multipurpose oil seed crop for industrial uses (*Jatropha curcas* L.): A review', *Industrial crops and products,* 28, pp. 1-10.

Lessard, P.A., Allen, R.D., Fujiwara, T. and Beachy, R.N. (1993) 'Upstream regulatory sequences from 2 beta-conglycinin genes', *Plant Molecular Biology*, 24, pp. 873-885.

Li, M.R, Li, H.Q., Jiang, H.W., Pan, X.P., Wu, J.G., 2008, Establishment of an *Agrobacterium*-mediated cotyledon disc transformation method for *Jatropha curcas*. *Plant Cell Tissue and Organ Culture* 92: 173-181

Murashige, T. and Skoog, F. (1962) 'A revised medium for rapid growth and bioassays with tobacco cultures', *Physiol. Plant*, 15, pp. 473-497.

Murphy, D.J., Hernández-Pinzón, I. and Patel, K. (20010 'Role of lipid bodies and lipid-body proteins in seeds and other tissues', *Journal of Plant Physiology*, 158, (4), pp. 471-478.

Nakashima, K., Fujita, Y., Katsura, K., Maruyama, K., Narusaka, Y., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2006) 'Transcriptional regulation of ABI3-and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of *Arabidopsis*', *Plant Molecular Biology*, 60, pp. 51-68.

NanYing, C., Yang, Y., YanDong, L., Lili, W., Ping, H., Yin, G. and ChengCai, A. (2009) 'Efficient LEC2 activation of OLEOSIN expression requires two neighboring RY elements on its promoter', *Sci China Ser C – Life Sci*, 52, pp. 854–63.

Nature editorial. (2007) ‘Kill king corn’, *Nature*, pp 449 637

Popluechai, S., Raorane, M., Syers, K.J., O’Donnell, A.G. and Kohli, A. (2008) ‘Research needs to make *Jatropha* a viable alternate oilseed crop’, Proceedings of the International Workshop on Feasibility of Non-Edible Oil Seed Crops for Biofuel Production, May 25–27, 2007, Mae Fah Luang University, Chiang Rai, Thailand. Eds. Syers K.J. and Wood D. MFLU publication pp. 56-60.

Qin, X., Zhang, J., Shao, C., Lin, S., Jiang, L., Zhang, S., Xu, Y. and Chen, F. (2009a) ‚Isolation and characterization of a curcin promoter from *Jatropha curcas* L and its regulation of gene expression in transgenic tobacco plants’, *Plant Mol Biol Rep*, 27, pp. 275-281.

Qin, X., Zhang, X., Shao, C., Gao, J., Jiang, L., Zhu, X., Yan, F., Tang, L., Xu, Y. and Chen, F. (2009b) ‘Stress-induced curcin-L promoter in leaves of *Jatropha curcas* L. and characterization in transgenic tobacco’, *Planta*, 230, pp. 387-395.

Reidt, W., Wohlfarth, T., Ellerström, M., Czihal, A., Tewes, A., Ezcurra, I., Rask, L. and Bäumlein, H. (2000) ‚Gene regulation during late embryogenesis: the RY motif of maturation-specific gene promoters is a direct target of the FUS3 gene product’ *Plant Journal*, 21, pp. 401-408.

Reyes, J.C., Muro-Pastor, M.I. and Florencio, F.J. (2004) ‘The GATA family of transcription factors in *Arabidopsis* and rice’, *Plant Physiology*, 134, pp.1718-1732.

Rieping, M. and Schoffl, F. (1992) ‘Synergistic effect of upstream sequences, CCAAT box elements, and HSE sequences for enhanced expression of chimeric heat shock genes in transgenic tobacco’, *Molecular & General Genetics*, 231, pp. 226-232.

Roberts, N.J., Scott, R.W. and Tzen, J.T.C. (2008) ‘Recent biotechnological application of using oleosins’, *The Open Biotechnology Journal*, 2, pp. 12-21.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory. Cold Spring Harbor, New York, USA

Slack, C.R., Bertaud, W.S., Shaw, B.D., Holland, R., Browse, J. and Wright, H. (1980) 'Some studies on the composition and surface properties of oilbodies from the seed cotyledons of safflower (*Carthamus tinctorius*) and linseed (*Linum ustatissimum*)', *Biochem J,* 190, pp. 551–561.

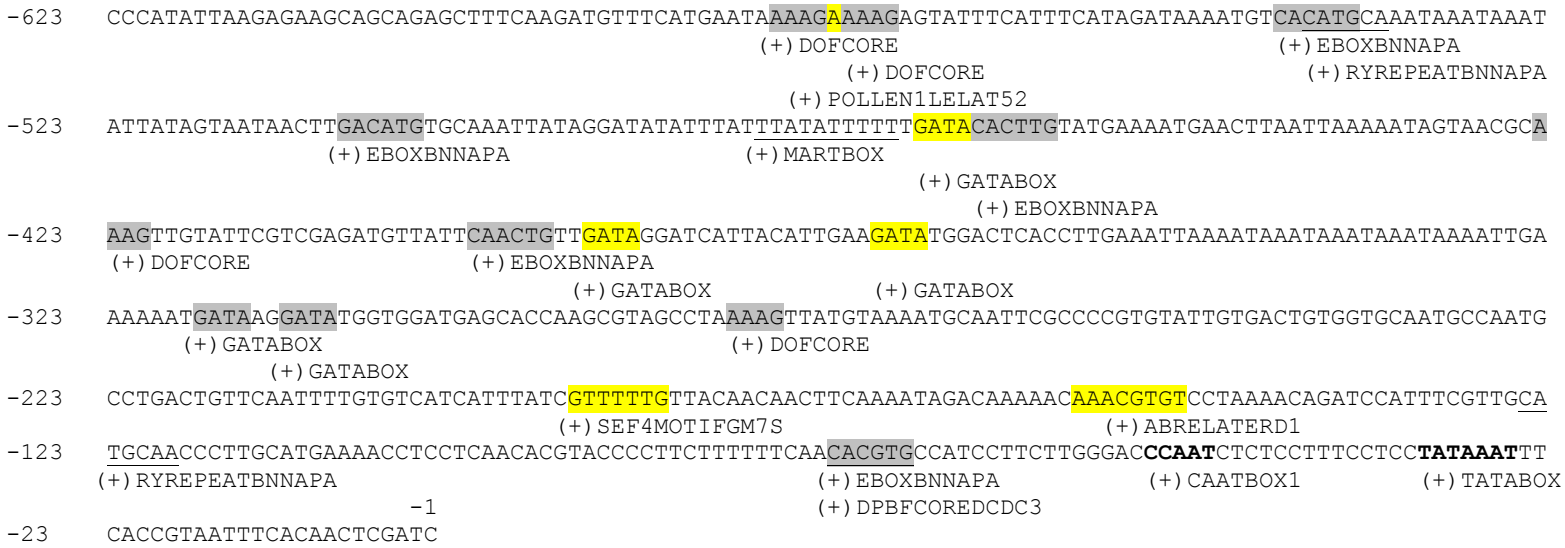
Simkin, A.J., Qian, T.; Victoria, C.; Franck, M.; Mohamed, B.R.; Lin, C.; Steve, T. and James, M. (2006) 'Oleosin gene family of *Coffea canephora*: Quantitative expression analysis of five *oleosins* in developing and germinating coffee grain', *Journal of Plant Physiology,* 163, pp. 691-708.

Stålberg, K., Ellerstöm, M., Ezcurra, I., Ablov, S. and Rask, L. (1996) ,Disruption of an overlapping E-box/ABRE motif abolished high transcription of the napA storage-protein promoter in transgenic *Brassica napus* seeds', *Planta,* 199, pp. 515-519.

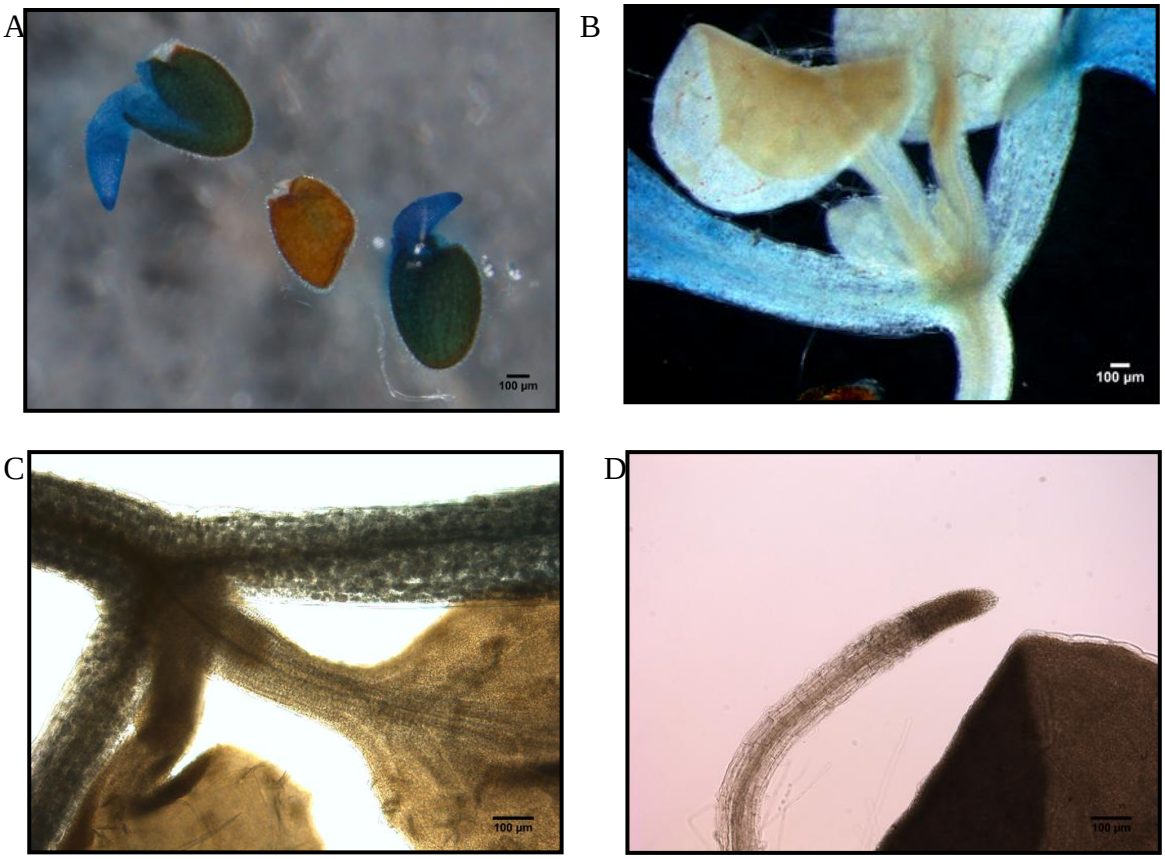
Tzen, J.T.C., Lai, Y.K., Chan, K.L. and Huang, A.H.C. (1990) 'Oleosin isoforms of high and low-molecular weights are present in the oilbodies of diverse seed species', *Plant Physiology,* 94, pp. 1282-1289.

Yanagisawa, S. (2000) 'Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize', *Plant Journal,* 21, pp. 281-288.

4.7 Supplental data



Supplemental figure 1 Schematic representation locations of putative cis-acting motifs for seed storage proteins of *JcOlesoin3* regulatory element (EU543568) predicted by the PLACE data base



Supplemental figure 2 (A) Histological analysis of GUS activity in 2 DAI of T1 transgenic *Arabidopsis* seedling, no expression observe in seed-coat **(B)** GUS activity seedling of 13 DAI of transgenic *Arabidopsis* **(C)** GUS activity leaves and stem of transgenic *Arabidopsis* **(D)** GUS activity in root of transgenic *Arabidopsis*

Chapter 5: Identification of patatin-like lipase proteins and presentation for a Triacylglycerides cycling scheme in *Jatropha curcas*

5.1 Abstract

Jatropha curcas has been identified to have a major potential as a biofuel crop. To be able to fully exploit its potential, an understanding of the molecular basis of its seed oil production, accumulation and utilization is necessary. However, little is know about the oil mobilization during seed germination in *J .curcas*. To understand the molecular basis of seed germination, proteomic studies were performed on resting and germinating seeds of *J. curcas*. Studies using 2-DE and MALDI-TOF MS revealed the presence of patatin-like lipases in the germinating seeds but not in the resting seeds. These proteins showed sequence similarity and conserved domains, as found in the patatin lipases from *Solanum tuberosum* and *Arabidopsis thaliana*. Patatin lipases play a significant role in the mobilization and lipolysis of TAGs during the germination of oil seeds. Further, there is little knowledge on the triacylglycerides (TAGs) lipid cycling in *J. curcas*. Bioinformatic data-mining confirmed TAGs cycling in *J. curcas* to include the expected stages of synthesis, storage and lipolysis of fatty acids. Based on the data-mining results for the presence of various genes for TAGs storage, synthesis and utilization, a scheme is proposed for TAGs cycling in *J. curcas*.

5.2 Introduction

Lipid biosynthesis

The biosynthesis of triacylglycerides (TAGs) in oilseed plants involves three stages.The first is synthesis in the plastids, followed by modification of the fats catalyzed by enzymes in the endoplasmic reticulum (as shown in Figure 1). In the final stage the newly syntheszed TAGs are stored in oil bodies (oleosomes) derived from the endoplasmic reticulum (Chen et al., 2008; Quettier and Eastmond, 2009).

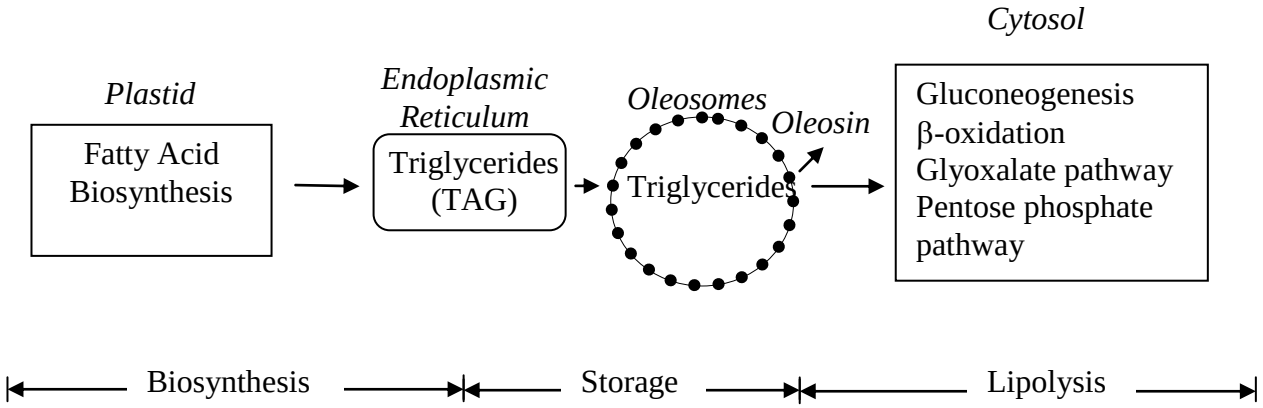


Figure 1 The TAG synthesis and breakdown cycle

Biosynthesis of TAGs in developing seeds follows the Kennedy pathway in a reaction catalyzed by acylCoA: diacylglycerol acyltransferase (DGAT; Chen et al., 2008; Baud and Lepiniec, 2009). A second pathway mediated by phospholipid:diacylglycerol acyltransferase (PDAT) has been described. This pathway can also be catalyzed by DGAT using two molecules of DAG to produce TAG and monoglycerol. Various factors mediate fatty acid synthesis at the transcriptional, translational and enzyme activity level (Chen et al., 2008; Baud and Lepiniec, 2009).

Lipid profiles in developing *J. curcas* seed have been studied. Lipid biosynthesis starts during the early stages of development, immediately after fertilization, and proceeds until maturation of the plant. It involves early development (till 22 days after fertilization) and maturation stages (Annarao et al., 2008; Baud and Lepiniec, 2009; Rao et al., 2009). High levels of primary sterols important in the hormonal regulation of water permeability of the phospholipids layer for cell division and embryogenesis are detected in the early stages and drop significantly at maturity (Annarao et al., 2008; Baud and Lepiniec, 2009; Rao et al., 2009). Campesterols, sitosterol, stigmasterols and isofucosterols, all sterols involved in early fruit development, were detected in the

developing seeds of *J. curcas*. There is a direct correlation between high levels of free fatty acids and sterols at the early stages of seed development. Environmental factors such as biotic and abiotic stress have significant effects on TAGs synthesis leading to higher free fatty acid (FFA) concentration (Annarao et al., 2008).

Three weeks after fertilization, young *Jatropha* seeds form membrane specific polar lipids. The presence of key enzymes for fatty acid synthesis showed that malonyl-CoA synthesis is implicated in the high level of fatty acid synthesis. The next stage of seed development is critical due to a tremendous fall in sterol concentration, little fatty ester formation, increased TAGs synthesis and drastic drop in FFA concentration (Annarao et al., 2008). TAGs synthesis occurs three weeks after fertilization alongwith an increase in saturated fatty acids and decrease in unsaturated fatty acids. Significant amounts of fats are synthesized at these intermediate stages of seed development (Annarao et al., 2008).

Some of the *J. curcas* genes for fatty acid biosynthesis have been identified such as, acetyl-CoA carboxylase, beta-keto acyl ACP synthase, and steroyl acyl desaturase (NCBI database)

Lipid storage

In plants, lipids are stored in sub-cellular organelles called oleosomes or oil bodies (Figure 1) which have a characteristic diameter of 0.6–2μm. The mechanism of oleosome formation in plants is not fully understood. The oleosome comprises of a protein (oleosin) coated spherical phospholipid monolayer (Chen et al., 2008; Quettier and Eastmond, 2009), which contains TAGs in its lumen. The size of the oil body determines the storage capacity for TAGs. Recently, the three oleosin genes whose products stabilize the oleosomes of *J. curcas* were studied at the gene, transcript and proteomic levels. These genes were largely conserved and the transcript expression was tissue-specific. All three oleosin proteins were shown to contain a conserved proline knot domain. Using a proteomic approach, caleosin, steroleosin, aquaporin and curcin were identified as minor proteins present in *J. curcas* oil bodies (Popluechai et al., 2008).

Lipolysis

Lipolysis is catalyzed by lipases which are esterases that hydrolyse triacylglycerol into glycerol and free fatty acids. This catalysis is important for seedling establishment, and thus plant fitness; acidic, basic and patatin-like lipases have been reported in plants (Eastmond, 2004; Eastmond, 2006; Quettier and Eastmond, 2009). A study on *J. curcas* reported activity of lipase during germination of seeds (Staubmann et al., 1999); in another study lipase activity was detected in both resting and germinating seeds with preference for hydrolysis of long-chain triacylglycerides rather than medium chain triacylglycerides (Abigor et al., 2002). Proteomic studies carried out by Yang and colleagues described oil mobilization at the germination stage of *J. curcas* involving several pathways including β-oxidation, the glyoxylate pathway, the glycolysis, the citric acid cycle, gluconeogenesis, and the pentose phosphate pathway (Yang et al, 2009). It would be expected that lipases will be up-regulated in germinating seeds in comparison to the resting seeds. However, no lipases were identified by Yang et al (2009). To investigate proteins up-regulated during germination, particularly the lipases, a 2D proteomic comparison was performed between germinating and resting seeds. Despite studies on individual genes in *J. curcas* for fatty acid biosynthesis, an integrative picture for TAGs cycling in *J. curcas* is absent. In this study, a bioinformatics data-mining-mediated integrative model for TAGs cycling in the seeds of *J. curcas* is proposed.

5.3 Materials and Methods

Sample Preparation

Ten *J. curcas* seeds were weighed, imbibed in sterile distilled water for 24 hours, and then covered with moist cotton wool for a further 8 h at 38ºC. Samples were incubated at 28ºC until the radicle immerged (approximately 5 mm of radicle). The germinating seeds were then harvested and stored at -80ºC until further analysis.

Protein extraction using urea Buffer

J. curcas seeds were ground in liquid nitrogen into a fine powder, and 100 mg of the powdered seed samples was homogenised in 8 M urea extraction buffer (2% w/v C7BzO/ 4% CHAPs and 40 mM DTT (dithiothreitol) and 10 µl/ml plant protease inhibitor cocktail (Sigma). The mixture was kept agitated for 2 h at 4ºC and centrifuged at 10,000 g for 15 minutes at 4ºC. The supernatant was transferred into new tubes and stored at -80ºC prior to analysis. The protein quantification was performed using 2-D quant kit (Amersham Biosciences, GE Healthcare, Bucks, UK) according to manufacturer's instructions. The assay is based on the specific binding of copper ions to protein. The absorbance of the standard (Bovine Serum Albumin, BSA) and samples were measured at 480 nm using spectrophotometer (SpectraMAX, Molecular Devices) and recorded; a standard curve was prepared using series BSA as a standard over a concentration.

Two-dimensional electrophoresis

Samples, containg 450 µg of proteins from dormant or germinated seeds, were cleaned-up prior to 2-D electrophoresis using Ettan™ Sample preparation Kits (Amersham Biosciences) according to the manufacturer's instructions. Isoelectric focusing (IEF) was carried out according to manufacturer's instructions on an Ettan IPGphor II (GE Healthcare, UK). In brief, 450 µg of protein were loaded onto 18 cm IPG strips with a non-linear gradient (pH 3-10) and rehydrated using 340 µl of rehydration buffer (DeStreak Rehydration Solution, GE Healthcare) at 30 V for 16 h. IEF was conducted at 20°C with an IPGphor II focusing system. The running conditions were as follows: 300 V for 30 min, followed by two gradients of 1000 V for 30 min and

5000 V for 80 min, and finally 5000 V for 25 min. The focussed IPG strips were equilibrated for 15 min in 5 ml equilibration solution, 75 mM Tris-HCl, 6 M urea, 30% w/v glycerol, 2% w/v SDS, 1% w/v (50 mg) DTT followed by equilibration in buffer containing 2.5% w/v IAA (125 mg) for a further 15 min. The second dimension was performed on a 10% SDS-PAGE, the running buffer contained 0.3% Tris, 1.44% glycine and 0.1% SDS. Protein spots in all gels were stained with Coomassie Brilliant Blue for about 1 h followed by de-staining with 25% methanol and 7% acetic acid in deionized water. The comparision of the presence and absence of protein spots between resting and germinating seeds were conducted. Spots of interests were excised from the gel, digested with trypsin (Sigma) according to the manufacturer's instruction and subjected to peptide mass fingerprinting by MALDI-TOF MS.

MALDI-TOF MS analysis

Mass fingerprint profiles of spots of interest were obtained by excising the spots out of the gels, digesting the proteins with trypsin, and determining the peptide masses by MALDI-TOF MS (Courchesne & Patterson, 1999). Determination of peptide masses was done in an Ultraflex III TOF/TOF MS (Bruker, Bremen, Germany). The database search using the peptide masses was performed with the Mascot Peptide Mass Fingerprint search programme [Matrix Science Ltd URL: http://www.matrixscience.com/home.html]. The searches were carried out at a mass accuracy of ± 0.3 Da.

Bioinformatics

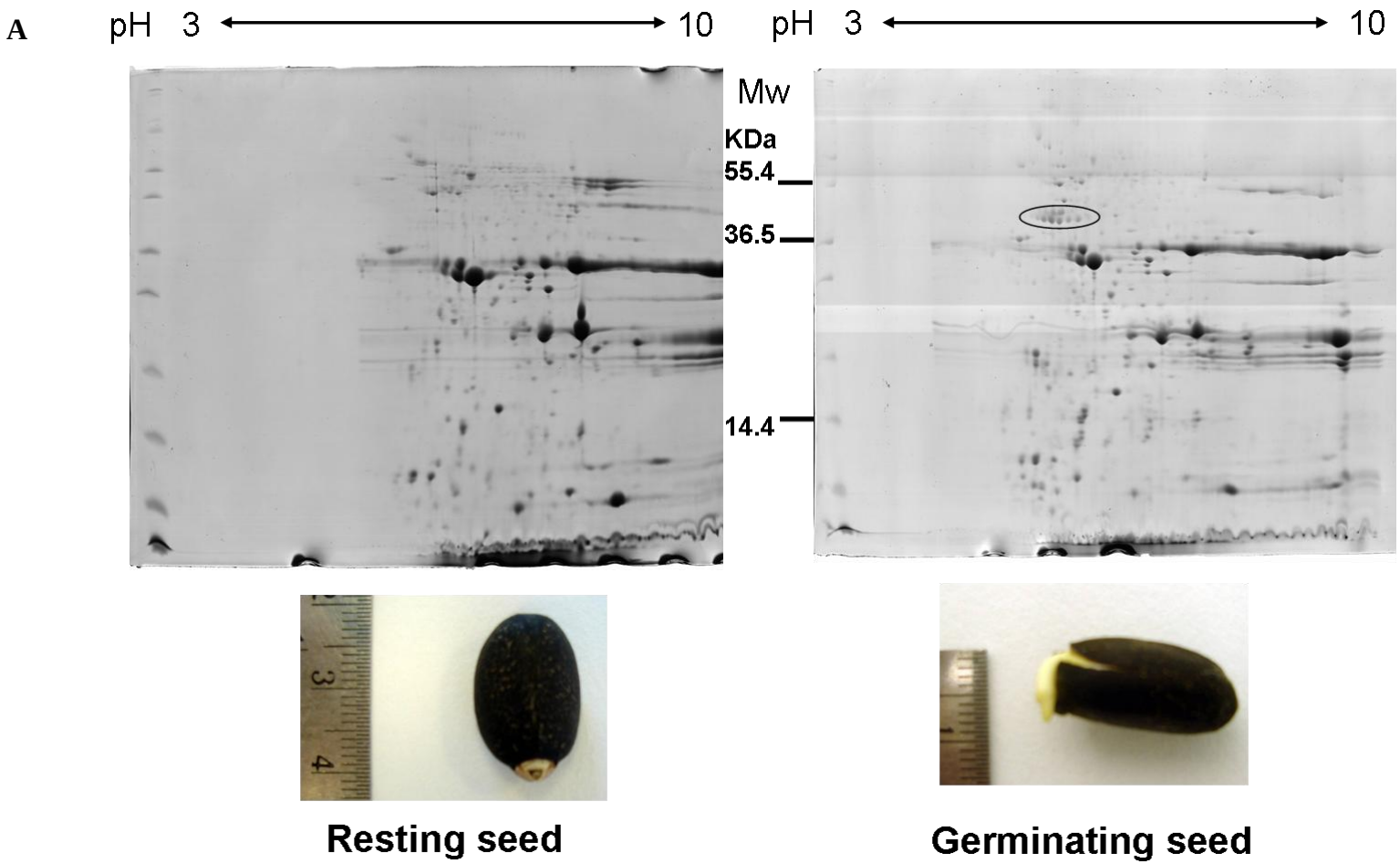
J. curcas genes involved in fatty acid biosynthesis and related genes sequences were retrieved from The National Center for Biotechnology Information (NCBI) databases and submitted to NCBI blast programme (Altschul et al., 1990), homologies were determined. Multiple sequence alignments of BLAST hits were carried out using the ClustalW at EMBL database (http://www.ebi.ac.uk/2can/tutorials/nucleotide/clust-alw.html).

5.4 Results

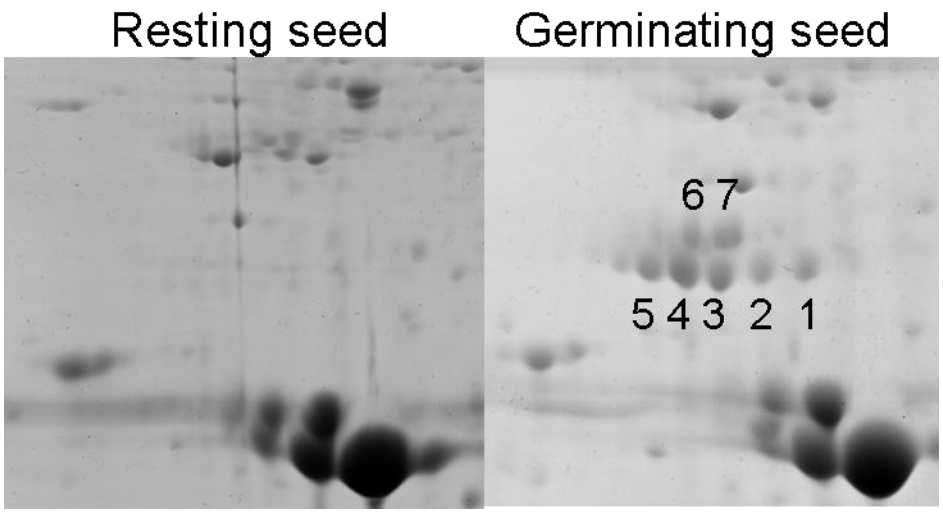
Identification of Patatin-like lipase using proteomic analysis

Protein profiles during seed germination were investigated using 18 cm 2D gel electrophoresis. In the resting seed protein profile, seventy spots were observed. Whereas eighty protein spots were observed in germinating seeds. Most spots between the two gels could be considered as homologous, albeit that their expression levels varied between these present in resting seeds or germinating seeds (Figure 2A). However, seven spots, designated 1-7, were unique to the germinating seed sample (Figure 2B), thus showing germination-specific up-regulation of these proteins. Figure 2C presents the different and common spots between the resting and germinating seeds protein profiles. The germination-specific spots 1 – 7 were analysed further using peptide mass finger-printing (PMF). MALDI-TOF MS revealed the spots to have molecular weights of approximately 41 kDa (Table 1 and Figure 3). Analyses of peptide fragments obtained from the digested protein spots based on their molecular weights, pI, sequence coverage and similarity database searches using MASCOT software and NCBI, indicated the proteins to be similar to one another. Interestingly, each showed a conserved domain similar to the patatin-like lipase protein from *Solanum tuberosum.*

Figure 2 (continue on next page)



B



C

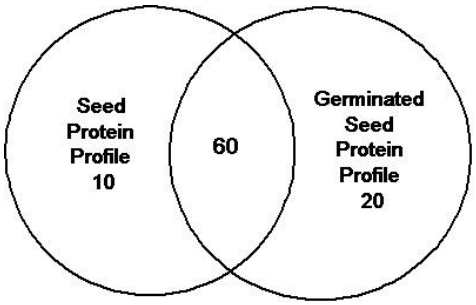


Figure 2 (A) Two-Dimensional gel images of protein profiles of resting and germinating stages resolved on 18 cm strips. A set of encircled proteins denotes proteins unique to germinating seeds; **(B)** Close-up sections of 2D gels, showing unique protein spots number 1 to 7 present in germinating but not in resting seeds; **(C)** A distribution of protein spots in the two stages of seed development.

1 MATTKSVLVL IFMILATTSS TFATLGEMVT VLSTDGGGIK **GIIPGIIPEF**
51 **LEGQLQK**MDN NADARLADYF DVIGGTSTGG LLTAMITTPN ENNRPFAAAK
101 DIVPFYFQHG PHIFNSSTGQ FFGPKYDGK**Y LMQVLQEKLG ETRVHQALTE**
151 **VAISSFDIKT NKPVIFTK**SN LAKSPELDAK MYDICYSTAA APIYFPPHHF
201 VTHTSNGATY EFNLVDGGVA TVGDPALLSL SVATRLAQED PAFSSIKSLD
251 YK**QMLLLSLG TGTNSEFDKT YTAEEAAK**WG PLRWMLAIQQ MTNAASSYMT
301 DYYISTVFQA R**HSQNNYLR**V QENALNGTTT EMDDASEANM ELLVQVGETL
351 LKKPVSK**DSP ETYEEALKR**F AKLLSDRKKL RANKASY

Spot number 1 (Q2MY40)

1 MILATTGSTC ATLGEMVTVL SIDGGGIK**GI IPATILEFLE GQLQEVDNNK**
51 **DAR**LADYFDV IGGTSTGGLL TAMITTPNEN NRPFAAAKDI VPFYFEHGPH
101 IFNSSGTIFG PMYDGK**YLLQ VLQEKLGETR VHQALTEVAI SSFDIKTNKP**
151 **VIFTK**SNLAK SPELDAKMYD ICYSTAAAPI YFPPHYFVTH TSNGDRYEFN
201 LVDGAVATVG DPALLSLSVA TR**LAQEDPAF SSIK**SLDYK**Q MLLLSLGTGT**
251 **NSEFDKTYTA EEAAK**WGPLR WLLAIQQMTN AASSYMTDYY LSTVFQAR**HS**
301 **QNNYLR**VQEN ALTGTTTEMD DASEANMELL VQVGETLLKK PVSK**DSPETY**
351 **EEALKR**FAKL LSDRKKLRAN KASY

Spot number 2 (Q3YJT3)

1 MATTKSFLIL FFMILATTSS TCATLGEMVT VLSIDGGGIK GIIPAVILEF
51 LEGQLQEVDN NKDAR**LADYF DVIGGTSTGG LLTAMITTPN ENNRPFAAAK**
101 **DIIPFYFEHG PHIFNYSGSI FGPMYDGKYL LQVLQEK**LGE TR**VHQALTEV**
151 **AISSFDIKTN KPVIFTK**SNL AKSPELDAKM YDICYSTAAA PMYFPPHYFI
201 THTSDGDIYE FNLVDGAVAT VGDPALLSLS VATR**LAQEDP AFSSIK**SLDY
251 K**QMLLLSLGT GTNSEFDKTY TAEEAAK**WGP LRWLLAIQQM TNAASSYMTD
301 YYISTVFQAH HSQNNYLR**VQ ENALTGTTTE MDDASEANME LLVQVGETLL**
351 **K**KPVSK**DSPE TYEEALKR**FA KLLSDRKKLR ANKASY

Spot number 3 (Q3YJT2)

1 MILATTGSTC ATLGEMVTVL SIDGGGIKGI IPATILEFLE GQLQEVDNNK
51 DAR**LADYFDV IGGTSTGGLL TAMITTPNEN NRPFAAAK**DI VPFYFEHGPH
101 IFNSSGTIFG PMYDGK**YLLQ VLQEKLGETR VHQALTEVAI SSFDIKTNKP**
151 **VIFTK**SNLAK SPELDAKMYD ICYSTAAAPI YFPPHYFVTH TSNGDRYEFN
201 LVDGAVATVG DPALLSLSVA TR**LAQEDPAF SSIK**SLDYK**Q MLLLSLGTGT**
251 **NSEFDKTYTA EEAAK**WGPLR WLLAIQQMTN AASSYMTDYY LSTVFQAR**HS**
301 **QNNYLRVQEN ALTGTTTEMD DASEANMELL VQVGETLLK**K PVSK**DSPETY**
351 **EEALKR**FAKL LSDRKKLRAN KASY

Spot number 4 (DQ3YJT3)

Figure 3 (continue on next page)

1 MILATTSSTF ASLEEMVTVL SIDGGGIKGI IPGTILEFLE GQLQKMDNNA
51 DARLADYFDV IGGTSTGGLL TAMITTPNEN NRPFAAANEI VPFYFEHGPH
101 IFNSSTGQFF GPKYDGK**YLM QVLQEKLGET RVHQALTEVA ISSFDIKTNK**
151 **PVIFTK**SNLA KSPELDAKMY DICYSTAAAP TYFPPHYFAT NTINGDKYKF
201 NLVDGAVATV ADPALLSVSV ATR**RAQEDPA FASIR**SLNYK K**MLLLSLGTG**
251 **TTSEFDKTHT AEETAK**WGAL QWMLVIQQMT EAASSYMTDY YLSTVFQDLH
301 SQNNYLRVQE NALTGTTTK**A DDASEANMEL LAQVGENLLK** KPVSK**DNPET**
351 **YEEALKR**FAK LLSDRKKLRA NKASY

Spot number 5 (Q3YJT4)

1 MILATTGSTC ATLGEMVTVL SIDGGGIKGI IPATILEFLE GQLQEVDNNK
51 DARLADYFDV IGGTSTGGLL TAMITTPNEN NRPFAAAKDI VPFYFEHGPH
101 IFNSSGTIFG PMYDGK**YLLQ VLQEKLGETR VHQALTEVAI SSFDIKTNKP**
151 **VIFTK**SNLAK SPELDAKMYD ICYSTAAAPI YFPPHYFVTH TSNGDRYEFN
201 LVDGAVATVG DPALLSLSVA TR**LAQEDPAF SSIK**SLDYK**Q MLLLSLGTGT**
251 **NSEFDKTYTA EEAAK**WGPLR WLLAIQQMTN AASSYMTDYY LSTVFQAR**HS**
301 **QNNYLR**VQEN ALTGTTTEMD DASEANMELL VQVGETLLKK PVSK**DSPETY**
351 **EEALKR**FAKL LSDRKKLRAN KASY

Spot number 6 (Q3YJT3)

1 MILATTGSTC ATLGEMVTVL SIDGGGIKGI IPATILEFLE GQLQEVDNNK
51 DARLADYFDV IGGTSTGGLL TAMITTPNEN NRPFAAAKDI VPFYFEHGPH
101 IFNSSGTIFG PMYDGK**YLLQ VLQEKLGETR VHQALTEVAI SSFDIKTNKP**
151 **VIFTK**SNLAK SPELDAKMYD ICYSTAAAPI YFPPHYFVTH TSNGDRYEFN
201 LVDGAVATVG DPALLSLSVA TR**LAQEDPAF SSIK**SLDYK**Q MLLLSLGTGT**
251 **NSEFDK**TYTA EEAAKWGPLR WLLAIQQMTN AASSYMTDYY LSTVFQAR**HS**
301 **QNNYLR**VQEN ALTGTTTEMD DASEANMELL VQVGETLLKK PVSK**DSPETY**
351 **EEALKR**FAKL LSDRKKLRAN KASY

Spot number 7 (Q3YJT3)

Figure 3 Peptide fragments obtained from peptide mass fingerprinting (PMF) of unique spots 1-7 from the 2D gel of germinating seeds. The bold and underlined sequences are the regions that matched to patatin-like lipases.

Table 1 Protein spots identified from germinating seeds of *J. curcas* using MALDI-TOF MS

Spot no.	Score ($p < 0.05$)	Sequence coverage (%)	Mw	pI	Protein Identification (accession number)
1	74	26	42.47	5.58	Patatin protein 11- *Solanum tuberosum* (Q2MY40)
2	85	32	41.08	5.12	Patatin - *Solanum tuberosum* (Q3YJT3)
3	69	46	42.55	5.04	Patatin - *Solanum tuberosum* (Q3YJT2)
4	96	44	41.08	5.12	Patatin - *Solanum tuberosum* (Q3YJT3)
5	86	29	41.29	5.39	Patatin - *Solanum tuberosum* (Q3YJT4)
6	79	25	41.08	5.12	Patatin- *Solanum tuberosum* (Q3YJT3)
7	71	23	41.08	5.12	Patatin - *Solanum tuberosum* (Q3YJT3)

J. curcas genes involved in TAGs biosynthesis, storage and lipolysis

Nineteen genes from *J. curcas* involves in seed oil biosynthesis, storage, and lipolysis were identified from NCBI database (accessed 01/05/10; Table 2). There were fifteen genes involves in TAGs biosynthesis, three genes involved in TAGs storage, and one gene involved in lipolysis. Analysis of translated coding sequences of these genes using BLAST alignments showed homology ranges from 60-100 % to known TAGs biosynthesis/storage/lipolysis genes in other oilseed plants (Table 2). Most genes exhibited conserved domains associated with specific classes of these genes (Table 2).

Table 2 Sequence analysis of triacylglycerides (TAGs) lipid cycling genes in *J. curcas*

J. curcas Genes	Gene accession number	Protein accession number	Maximum hit to other plants	% Homology	Conserved domain	Location in cell
rubisco	EU395776	ABZ04129	*N. tobaccum* (AAA34116)	100	No/NI*	Plastid
acc	DQ987702	ABJ90471	*R. communis* (EEF48464)	90	Yes/ NI*	Plastid
kas I	DQ987699	ABJ90468	*R. communis* (ABR12416)	92	Yes/G^{355}NTSAAS^{361}	Plastid
kas II	DQ987700	ABJ90469	*R. communis* (EEF46230)	81	Yes/G^{355}NTSAAS^{361}	Plastid
kas III	DQ987701	ABJ90470	*R. communis* (EEF32603)	88	Yes/G^{355}NTSAAS^{361}	Plastid
sad	DQ084491	AAY86086	*R. communis* (AAA74692)	96	Yes/ExxH	Plastid
omega3-fad	EU106890	ABU96743	*V. fordii* (AAD13527)	85	Yes/(HXXX(H),HXX(X)HH,HXXHH	Plastid
omega6-fad	EU106889	ABU96742	*B. napus* (AAA50157)	84	Yes/(HXXX(H),HXX(X)HH,HXXHH	Plastid
fat A	EU267122	ABX82799	*R. communis* (ABS30422)	80	Yes/hot-dog super-family	Plastid
fat B	EU106891	ABU96744	*R. communis* (EEF47013)	88	Yes/hot-dog super-family	Plastid
delta12-fad (fad2)	DQ157776	ABA41034	*V. fordii* (AAN87573)	91	Yes/(HXXX(H),HXX(X)HH,HXXHH	ER
fad3	EU267121	ABX82798	*V. fordii* (CAB45155)	83	Yes/(HXXX(H),HXX(X)HH,HXXHH	ER
gpat	FJ952147	ACR61638	*R. communis* (EEF43526)	83	Yes/NI*	ER
dgat 1	DQ278448	ABB84383	*V. fordii* (ABC94472)	87	Yes/O-acyl transferase, histidyl active site	ER
dgat 2	EU395774	NI	*R. communis* (ACB30544)	79	Yes/O-acyl transferase, histidyl active site	ER
oleosin I	EU234462	ABW90148	*C. arabica* (AY14574)	60	Yes/PXXXXXSPXXXP, Proline knot	Oilbodies
oleosin II	EU234463	ABW90149	*R. communis* (AAR15171)	69	Yes/PXXXXXSPXXXP, Proline knot	Oilbodies
oleosin III	EU234464	ABW90150	*C. sinensis* (CAA88360)	78	Yes/PXXXXXSPXXXP, Proline knot	Oilbodies
lipase	FJ233094	ACI62779	*R. communis* (EEF45491)	79	Yes/GXSXG	Cytoplasm

NI: not identified

5.5 Discussion

Identification of J. curcas patatin-like lipase using a proteomics approach

Proteomic analysis of the protein spots showed a divergence in protein profile between the resting and germinating stages of *J. curcas*. Patatin-like lipase was identified in the germinating stage of *J. curcas* seeds, this protein was up-regulated indicating its role in seed germination. Patatin-like lipase, an oil body-associated protein, has been identified in *Arabidopsis* with a conserved domain of esterase GXSXG and has been described as a storage protein found in abundance in oil bodies with the ability to degrade triacylglycerides. The identified patatin-like lipase possesses similar catalytic domains to that in *Solanum tuberosum* and SDP1 of *Arabidopsis* that has been shown to catalyse the first step in the hydrolysis of triacylglycerides, diacylglycerides and monoglycerides with a preferential activity on TAGs indicating its role in germinating oil seeds (Eastmond, 2006). This may indicate that oil mobilization starts at an early stage of seed germination (Yang et al., 2009). High activity of lipid degrading enzymes has been reported on germination of many different oilseeds which is attributed to TAGs mobilization, hence the low amount of oil detected (Eastmond, 2006).

The high lipase activity in germinating seeds leading to TAGs mobilization may be due to energy requirements as they lack an efficient source of nutrition and photosynthetic system. The energy stored in the form of lipid reserves in the oil bodies is degraded to free fatty acids that enter the β-oxidation and glyoxylate pathway, thus oil mobilization occurs during germination to provide the energy required for the necessary physiological activities of the seeds. The glyoxylate cycle, lipoxygenase-dependent oxidation of polyunsaturated fatty acids and fatty acyl-CoA-synthetase-independent pathways are described to be the routes of oil mobilization in germinating seeds. Lipoxygenase-dependent oxidation of lipids is essential in TAGs mobilization during germination of *J. curcas* (Yang et al., 2009). Failure to identify patatin-like lipase from previous work (Yang et al., 2009) may have been a consequence of different methods of protein extraction used and different sampling times. However, this hypothesis remained to be investigated.

Ribulose 1, 5-bisphosphate carboxylase is responsible for fixing CO_2 in maturing embryos of green plants. Ribulose 1, 5-bisphosphate carboxylase in *J. curcas* (NCBI EU395776) and its translated amino acids sequence is identical to that in tobacco (NCBI, AAA34116). The synthesis of fatty acids takes place in the plastid as shown in the Figure 5.

Acetyl-CoA carboxylase catalyzes the committed step that involves carboxylation of acetyl-CoA to malonyl-CoA (Baud and Lepiniec, 2009). The reaction takes place in two stages. The initial carboxylation of a biotin cofactor bound covalently to the central-biotin carboxyl carrier protein is first catalyzed by biotin carboxylase. The biotin bound CO_2 is then transferred to acetyl-CoA forming malonyl-CoA in a reaction mediated by the carboxytransferase domain of the enzyme (Dyer and Mullen, 2005; Chen et al., 2008). The translated gene encoding acetyl-CoA carboxylase (ACC) in *J. curcas* (DQ987702) shows 90 % sequence similarity to that of ACC in *Ricinus communis* (EEF48464) (Marchler-Bauer et al., 2005). The translated gene shows conserve domain of ACC-central and the carboxyl-transferase. Thus, acetyl-CoA carboxylase in *J. curcas* may play the same role of carboxylation.

Malonyl-ACP is formed from malonyl-CoA and acyl carrier protein in a reaction catalyzed by malonyl-CoA:ACP transferase (MAT); this involves the transfer of the malonyl group to the acyl carrier protein (ACP) (Dyer and Mullen, 2005; Baud and Lepiniec, 2009). However, identification of the gene encoding this enzyme has not been reported in *J. curcas*, but it is thought to be present.

The formation of fatty acids proceeds by condensation of acetyl-CoA and malonyl-ACP as shown in Figure 5. β-ketoacyl-ACP synthase III (KASIII) catalyzes the initial condensation of malonyl-ACP and acetyl-CoA to 3-ketobutyryl-ACP. β-ketoacyl-ACP synthase I (KASI) catalyzes the condensation between the growing acyl-side chain and malonyl-ACPs to produce 6-16 carbon fatty acids with palmitoyl-ACP being produced in a series of seven condensation cycles (Töpfer et al., 1995; Li et al., 2008). β-ketoacyl-ACP synthase II (KASII) catalyzes the elongation of 16:0-ACP to 18:0-ACP, thus completing the biosynthesis of 18 carbon fatty acids. The palmitate content of plants has been shown to be dependent on KASII activity. β-ketoacyl-ACP synthase III (*KASIII*) in *Jatropha* (DQ987701) encodes amino acids with 88 % identity to *KASIII* in *R. communis* (EEF32603). *JcKASI* (DQ987699) encodes translated amino acids with 92% identity to

that of KASI in *R. communis* (ABR12416). β-ketoacyl-ACP synthase II (KASII) in *J. curcas* (DQ987700) encode translated amino acids that showed a high similarity of 81% to its counterpart in *R. communis* (EEF46230). *Jc*KASI, II and III showed conserved domains in their active sites of cysteine and the condensing enzymes superfamily, as shown in Table 2 (Marchler-Bauer et al., 2005; Li et al., 2008). Hence, *Jc*KASI, II and III are thought to catalyze the condensation of fatty acids.

Modification of the 18:0-ACP includes desaturation by steroyl-ACP desaturase (SAD) which includes the introduction of a double bond into steroyl-ACPs between C9-C10 producing oleoyl-ACP (18:1-ACP). SAD determines the ratio of saturated to unsaturated fatty acids in plants (Lindqvist et al., 1996; Luo et al., 2009). *J. curcas* steroyl-ACP desaturase (*sad*) (DQ084491) and its translated amino acids show a high homology of 96% to the SAD from *R. communis* (AAA74692), Two conserved domains were detected, one region belonging to the acyl-ACP desaturase family and the second belonging to the ferritin-like family. The gene encoding the above enzyme has been isolated from *J. curcas* and shown high expression when cloned in *E. coli*. Enzymes of *de novo* fatty acid biosynthesis form a multisubunit complex referred to as fatty acid synthase (FAS) that dissociates easily (Tong et al., 2006)

In *A. thaliana* and other plants, three additional reactions have been described for the production of saturated fatty acids: the reactions catalyzed by 3-ketoacyl-ACP reductase (KAR), 3-hydroxyacyl-ACP dehydratase (HD) and enoyl-ACP reductase (ENR) (Baud and Lepiniec, 2009). However, to date, none of the genes encoding the above enzymes have yet been reported in *J. curcas*, thus more research is needed to identify whether or not they are present.

Acyl-carrier protein thioesterase is critical to chain termination during the *de novo* fatty acid synthesis; it cleaves fatty acids from acyl-ACPs. To date two thioesterases, FatA and FatB, have been described in plants. FatA removes the acyl group from ACP in 18:0-ACP and 18:1-ACP, while FatB is responsible for dissociation of palmitoyl-ACP (16:0-ACP) to palmitate, the latters thioesterase showing higher affinity for saturated acyl-ACPs (Dyer and Mullen, 2005; Venegas-Calerón et al., 2006). The fatty acid pool of 16:0-CoA and 18:0-CoA produced are transferred to the endoplasmic reticulum where further modifications in the form of desaturation and elongation take place. In *J. curcas* FatA (EU267122) has been shown to encode translated amino acids homologous to those of FatA in *R. communis* (ABS30422) with 80% identities. Similarity, the translated amino acids of FatB in *J. curcas* (EU106891) have high similarity to that from *R. communis* (EEF47013), with 88% homology. Since both FatA and FatB in *J. curcas* showed

conserved motifs in two active sites and a conserved domain to the hot-dog superfamily, it is thought that they might have similar functions to those reported in others higher plants (Marchler-Bauer et al., 2005).

In higher plants, desaturation to produce polyunsaturated fatty acids is mediated by various desaturases including omega-3, omega-6 genes of the membrane FADs-like superfamily (Dyer and Mullen, 2005). The *Jcomega-3 fad* gene (EU106890) encodes a translated amino acid sequence with high homologous (85%) to omega-3 desaturase from *Vernicia fordii* (AAD13527) whereas the *Jcomega-6 fad* gene (EU106889) encodes translated amino acids with 84 % similarity to that in *Brassica napus* (AAA50157). The *omega-3 and omega-6 fatty acid desaturase* genes in *J. curcas* possess conserved domains to di-iron ligands with catalytic histidyl residues (Table 2). Therefore, *omega-3 and omega-6* genes might have similar roles to those in higher plants of encoding enzymes (desaturases) involved in fatty acid desaturation (Dyer and Mullen, 2005).

Lipids in oilseed plants are stored as triacylglycerides (TAGs), thus the fatty acids synthesized are converted to TAGs in reactions of the Kennedy pathway. *Jc*GPAT may function in catalyzing the committed step in glycerolipid synthesis which is the initial reaction and occurs outside the plastid (Zheng et al., 2003). Glycerol-3-phosphate is acylated to TAGs in reactions catalysed by glycerol-3-phosphate acyltransferase (GPAT), lysophosphatidic acid acyltransferase (LPAT) and diacylglycerol acyltransferase (DGAT) (Dyer and Mullen, 2005; Baud and Lepiniec, 2009). DGAT 1 and DGAT 2 encoded by single genes are located in the endoplasmic reticulum and mediate the committed step in TAGs synthesis. Furthermore, DGAT 1 is expressed in various organs and has key roles in metabolism of TAGs in plants while DGAT 2 is the found in developing seeds, and functions in processing and accumulation of unusual fatty acids such as α-eleostearic acid (Shockey et al., 2006). Diacylglycerol is initially converted to phosphotidylcholine and phosphotidyl-ethanolamine before being converted to triacylglycerides (TAGs). The glycerol 3-phosphate acyltransferase (*gpat*) gene in *J. curcas* (FJ952147) encodes a protein with that shows 83% homology to that in *R. communis* (EEF43526). *JcGPAT* possesses conserved domains to the acyltransferase superfamily, with four conserved acyltransferase motifs and two transmembrane domains. The translated amino acid of *J. curcas* diacylglyceol acyltransferase (DGAT1) (DQ278448) showed 87 % identity to DGAT1 in *V. fordii* (ABC94472). *JcDGAT*2 (EU395774) amino acids showed 79% homology to DGAT2 in *R. communis* (ACB30544). *JcDGAT* showed a conserved domains belonging to the membrane bound O-acyl transferase family and conserved histidyl active site (Marchler-Bauer et al., 2005). PC and PE are desaturated by FAD2 and

FAD3 before TAGs synthesis (Dyer and Mullen, 2005; Baud and Lepiniec, 2009). Thus the desaturase enzymes might function in modification of phosphatidylcholine (PC) and Phosphatidylethanolamine (PE) in *J. curcas*. A further pathway of TAGs synthesis catalyzed by phospholipid:diacylglycerol acyltransferase (PDAT) where fatty acyl-group is transferred from phosphatidylcholine to diacylglycerol (DAG) has also been reported (Dyer and Mullen, 2005; Chen et al., 2008). The *J. curcas fad2* (*Jcdelta-12 fad2*) gene (DQ157776) encodes translated protein with 91% homology to delta-12 *fad2* in *V. fordii* (AAN87573). The *fad3* gene (EU26712) of *J. curcas* encodes protein that has over 83% similarity to those of FAD3 in *V. fordii* (CAB45155) and *G. max* (ACU17817). Both FAD2 and FAD3 possess conserved domains to di-iron ligands with catalytic histidyl residues (Table 2). Therefore, *fad2 and fad3* genes might have similar roles to those in higher plants of encoding enzymes (desaturases) involved in fatty acids desaturation (Dyer and Mullen, 2005).

Bioinformatics data-mining of Triaclycerides storage genes

TAGs is stored in small spherical droplets 'oil bodies' referred to as oleosomes that bud off from the endoplasmic reticulum (Siloto et al, 2006). The oleosomes are associated with oleosin proteins, whislt serve to maintain seed oil reserves in small droplets that have a high surface to volume ratio. The translated amino acids of the *JcOleosin1* gene (EU234462) in *J. curcas* has 60% identity to those in coffee bean (AY14574), whislt *JcOleosin2* (EU234463) translated amino acids showed 69 % homologies to *R. communis* (AAR15171). The translated amino acids of the *JcOleosin3* gene (EU234464) in *J. curcas* have 78% identity to oleosin in *Camelia sinensis* (CAA88360). All these *J. curcas* oleosins showed a conserved proline-knot domain. They function in maintaining steric hindrance and electrical repulsion between oil bodies and also preventorganelles from coalescence during seed maturation, desiccation and germination (Popluechai et al., 2008). Three oleosins with minor proteins, namely steroleosin, caleosin and aquaporin have been reported in *J. curcas*, where they play a role in TAGs degradation. High expression of *JcOleosin1*, *2* and *3* was reported in the seeds of *J. curcas*, with *JcOleosin3* being the most abundant; an increase in expression of these *JcOleosins* was found at the initial stages of germination. This expression pattern may play a role in prevention of coalescence during oil mobilization (Popluechai et al., 2008).

Bioinformatics data-mining of Triaclycerides lipolysis genes

The translated amino acid sequence of lipase in *J. curcas* (FJ233094) is 79% homologous to that in *R. communis* (EEF45491) and showed conserved domains in its catalytic active site belonging to the lipase superfamily as shown in Table 2 (Marchler-Bauer et al., 2005). Lipase activity was detected in germinating seeds of *J. curcas* (Staubmann et al., 1999), whilst a later report described lipase activity in both resting and germinating seeds of *J. curcas*, with a preference to hydrolyse long-chain triglycerides at a faster rate than medium chain triglycerides (Abigor et al., 2002). In this study, the patatin-like lipases were identified from *J. curcas* seedling indicating that patatin-like lipases might play a role during germination similar to previously report by Eastmond (2006). Storage oil mobilization has been shown to occur during seed germination, and lipolysis in seeds is thought to occur through one of the pathways catalyzed by lipoxygenase, phospholipase and triacylglycerol (TAG) lipase. TAG lipases have been isolated from oil bodies of oil seeds and are thought to play a role in lipolysis in germinating seeds by hydrolysing oxygenated fatty acids (Feussner et al., 2001). The free fatty acids are transported to the peroxisomes where they are converted to acyl-CoA and enter β-oxidation; the acetyl-CoA produced is converted to oxaloacetate and enters the glyoxylate cycle. Glycerol is converted to dihydoxyacetone phosphate; oxaloacetate and dihydroxyacetone phosphate enter the gluconeogenesis pathway to be converted to sugars. Proteomic studies carried out by Yang and colleagues described oil mobilization at the germination stage and its subsequent consumption in early seedling stage (Quettier and Eastmond, 2009; Yang et al., 2009).

Scheme of triacylglycerides cycling of J. curcas

The results obtained from bioinformatics data-mining of genes from *J. curcas* involved in biosynthesis/storage/lipolysis of TAGs showed close similarities and domain conservation to the corresponding genes of other oilseed model plants. Based on previously describes genes from other oilseed plants together with the results obtained from this study for the patatin-like lipase identification, a scheme in *J. curcas* for triacylglyceride cycling was presented, as shown in Figure 4. This scheme provides preliminary information required for oil manipulation in *J. curcas*, but still needs to be refined. However, the proposed scheme is similar to a model proposed for TAGs cycling in *Arabidopsis* (Dyer and Mullen, 2005; Quettier and Eastmond, 2009), especially the

presence of patatin-like lipases during the initial stage of seed germination (Eastmond, 2006). Thus genetic manipulation of oil yield and quality trait of *J. curcas* might be achieved using existing information from *Arabidopsis*. For example biodiesel property in *J. curcas* is determined by the ratio of oleic acid to linoleic acid. The FAD2 enzyme is responsible for conversion of oleic acid ($18:1\Delta^{9cis}$) to linoleic acid ($18:2\Delta^{9cis,12cis}$), contributing to the final O/L ratio in the seed oil. Therefore suppression of JcFAD2 activity could result in accumulation of oleic acid and a reduction in linoleic acid as similar to previous study in peanut (Yin et al., 2007). The same approach can be applied to increase *J. curcas* seed oil content. Seed overexpression of *J. curcas acc* (EU395776), *dgat2* (EU395774) genes can be resulted in increasing seed oil content as similar to previous study in *Arabidopsis* (Roesler et al., 1997; Jako et al., 2001). If time permits, additional bioinformatics data-mining such as homology modelling should be carried out to make the scheme more robust.

5.6 Conclusions and future studies

J. curcas is a potential biodiesel crop, but little is currently known regarding the molecular basis of TAGs biosynthesis, storage and mobilization. Patatin-like lipases were identified using proteomic analysis and this was shown to be similar to that already identified in *A. thaliana*. Further, it was shown to have the same conserved domain (Eastmond, 2006). This lipase has been described to play a key role in triglycerides hydrolysis and mobilization during germination of oil seed plants replenishing the energy loss due to lack of essential nutrients and photosynthetic machinery (Eastmond, 2006; Yang et al., 2009).

Study on molecular basis of seed germination is necessary to understand the limitation of seed propagation. In the present study attempts were made to isolate patatin-like lipase genes from *J. curcas* germinating seed using degenerate primers-mediated PCR amplification. The degenerate primers were designed from the peptide mass fingerprint (PMF) obtained from MALDI-TOF MS. Unfortunately, this approach was not successful. Thus, the future work gene isolation should be base on the information of PMF obtained from tandem mass spectrometry (LC-MS/MS).

The presented scheme of TAGs Biosynthesis/Storage/Lipolysis of *J. curcas* presented in this study needs to be refined and confirmed. Experimental verification such as gene silencing is required to investigate the scheme proposed in this study.

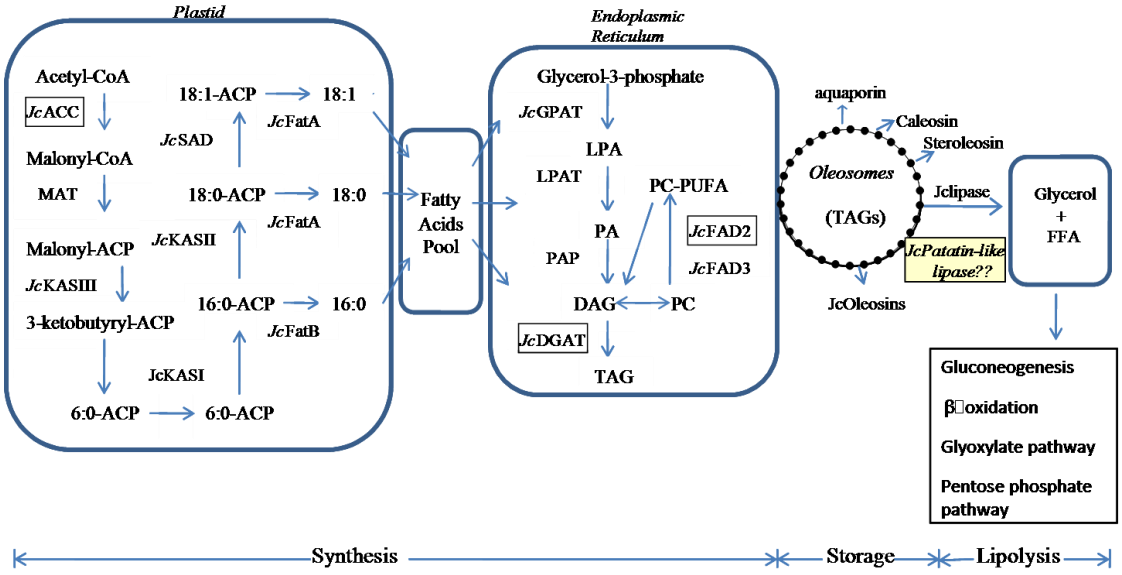


Figure 4 Scheme of TAGs lipid cycling of *J. curcas* (adapted from Dyer and Mullen, 2005; Quettier and Eastmond, 2009)

This scheme is divided into three discreet stages: synthesis, storage and lipolysis of TAGs. As shown in Figure 5 the synthesis of fatty acids take place in the plastid; acetyl-CoA carboxylase catalyses acetyl-CoA in the first step in fatty acid synthesis to malonyl-CoA. Malonyl-ACP is produced from malonyl-CoA a reaction catalyzed by MAT. In series of condensation reaction cycles catalyze by KASIII, KASI and KASII, 16 and 18-C fatty acyl-ACPs are formed; the 18-C is desaturated by SAD and the ACP is removed from the fatty acids forming the pool. Fatty acids are transferred to the endoplasmic reticulum where the triacylglycerides are produced in a reaction catalyzed by GPAT and DGAT; TAGs are stored in the oleosomes that bud from ER, the oleosins stabilise these oil bodies once dormancy is broken. Lipase/hydrolase and patatin-like lipase (in shaded-box) degrades triglycerides to glycerol and fatty acids and these are channelled into gluconeogenesis, β-oxidation, glyxoylate pathway and pentose phosphate pathway. *Jatropha* genes in rectangular boxes indicate potetial genetic manipulation to increase ol quality/quantity (see discussion)

Abbreviations; *Jc*: *Jatropha*, ACC: acetyl-CoA carboxylase, ACP: acyl carrier protein, DAG: diacylglycerol, DGAT: diacylglycerol acyltransferase, FAD: fatty acid desaturase, FFA: free fatty acids, FAT: fatty acyl-ACP thioesterase, GPAT: glycerol 3-phosphate transferase, KAS: β-ketoacyl-ACP synthase, LPA: lysophosphatidic acid, LPAT: lysophophatidic acid acyltransferase, MAT: malonyl-CoA:ACP transferase, PA: phosphatidic acid, PAP: phosphatidic acid phosphatase, PC: phosphatidylcholine, PC-PUFA: PC with polyunsaturated fatty acids, SAD: steroyl-ACP desaturase, TAG: triacylglycerol.

5.7 References:

Abigor, R.D., Uadia, P.O., Foglia, T.A., Haas, M.J., Scott, K. and Savary, B.J. (2002) 'Partial purification and properties of lipase from germinating seeds of *Jatropha curcas* L.' , *Journal of the American Oil Chemists' Society*, 79, pp. 1123-1126.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) 'Basic local alignment search tool' , *Journal of Molecular Biology*, 215, pp. 403-410.

Annarao, S., Sidhu, O.P., Roy, R., Tuli, R. and Khetrapal, C.L. (2008) 'Lipid profiling of developing *Jatropha curcas* L. seeds using 1H NMR spectroscopy' , *Bioresource Technology*, 99, pp. 9032-9035.

Baud, S. and Lepiniec, L. (2009) 'Regulation of de novo fatty acid synthesis in maturing oilseeds of *Arabidopsis*' ,*Plant Physiology and Biochemistry*, 50, (8), pp. 1463-1478.

Chen, Y., Wang, Y., Zhou, G., Li, P. and Zhang, S. (2008) 'Key mediators modulating TAG synthesis and accumulation in woody oil plants', *African Journal of Biotechnology*, 7, pp. 4743-4749.

Courchesne, P.L. and Patterson, S.D. (1999) 'Identification of proteins by matrix-assisted laser desorption/ionization mass spectrometry using peptide and fragment ion masses', *Methods in molecular biology* (Clifton, N.J.), 112, pp. 487-511.

Dyer, J.M. and Mullen, R.T. (2005) 'Development and potential of genetically engineered oilseeds', *Seed Science Research*, 15, pp. 255-267.

Eastmond, P.J. (2004) 'Cloning and Characterization of the Acid Lipase from Castor Beans', *The Journal of Biological Chemistry*, 279, pp. 45540–45545.

Eastmond, P.J. (2006) 'Sugar-dependent1 encodes a patatin domain triacylglycerol lipase that initiates storage oil breakdown in germinating *Arabidopsis* seeds', *Plant Cell* 18, pp. 665-675.

Feussner, I., Kuhn, H. and Wasternack, C. (2001) 'Lipoxygenase-dependent degradation of storage lipids', *Trends in Plant Science*, 6, pp. 268-273.

Gressel, J. (2008) 'Transgenics are imperative for biofuel crops', *Plant Science*, 174, pp. 246-263.

Jako, C., Kumar, A., Wei, Y., Zou, J., Barton, D.L., Giblin, E.M., Covello, P.S., Taylor, D.C. (2001) 'Seed-specific over-expression of an *Arabidopsis* cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight' , *Plant Physiology*, 126 (2), pp. 861-874.

Li, J., Li, M.-R., Wu, P.-Z., Tian, C.-E., Jiang, H.-W. and Wu, G.-J. (2008) 'Molecular cloning and expression analysis of a gene encoding a putative -ketoacyl-acyl carrier protein (ACP) synthase III (KAS III) from *Jatropha curcas*', *Tree Physiology*, 28, pp. 921-927.

Lindqvist, Y., Huang, W., Schneider, G. and Shanklin, J. (1996) 'Crystal structure of A9 stearoyl-acyl carrier protein desaturase from castor seed and its relationship to other di-iron proteins', *The EMBO Journal*, 15, pp. 4081-4092.

Luo, T., Deng, W.Y., Zeng, J. and Zhang, F.L. (2009) 'Cloning and Characterization of a Stearoyl–Acyl Carrier Protein Desaturase Gene from *Cinnamomum longepaniculatum*', *Plant Mol Biol Rep.*, 27, pp. 13-19.

Marchler-Bauer, A., Anderson, J.B., Cherukuri, P.F., DeWeese-Scott, C., Geer, L.Y., Gwadz, M. et al. (2005) 'CDD: A Conserved Domain Database for protein classification', *Nucleic Acids Research* 33 (DATABASE ISS.), pp. D192-D196.

Popluechai, S., Chardot, T., Gatehouse, A.M.R. and Kohli, A. (2008) 'Characterisation of Jatropha curcas Oil Bodies and Oleosins: A Systems Approach', In *Jatropha.* International Congress. Singapore.

Quettier, A.-L. and Eastmond, P.J. (2009) 'Storage oil hydrolysis during early seedling growth', *Plant Physiology and Biochemistry*, 47, (6), pp. 485-490.

Rao, K.S., Chakrabarti, P.P., Rao, B.V.S.K. and Prasad, R.B.N. (2009) 'Phospholipid composition of *Jatropha curcus* seed lipids', *JAOCS, Journal of the American Oil Chemists' Society*, 86, pp. 197-200.

Roesler, K., Shintani, D., Savage, L., Boddupalli, S., Ohlrogge, J. (1997) 'Targeting of the *Arabidopsis* homomeric acetyl-coenzyme A carboxylase to plastids of rapeseeds', *Plant Physiol*, 113, pp. 75–81.

Shockey, J.M., Gidda, S.K., Chapital, D.C., Kuan, J.-C., Dhanoa, P.K., Bland, J.M. Rothstein, S.J., Mullen, R.T. and Dyer, J.M.(2006) 'Tung Tree DGAT1 and DGAT2 Have Nonredundant Functions in Triacylglycerol Biosynthesis and Are Localized to Different Subdomains of the Endoplasmic Reticulum W.', *The Plant Cell*, 18, pp. 2294–2313.

Siloto, R.M., Findlay, K., Lopez-Villalobos, A., Yeung, E.C., Nykiforuk, C.L. and Moloneya, M.M. (2006) 'The Accumulation of Oleosins Determines the Size of Seed Oilbodies in *Arabidopsis*', *The Plant Cell*, 18, pp. 1961–1974.

Staubmann, R., Ncube, I., Gubitz, G.M., Steiner, W. and Read, J.S. (1999) 'Esterase and lipase activity in *Jatropha carcass* L. Seeds', *Journal of Biotechnology*, 75, pp. 117-126.

Tong, L., Shu-Ming, P., Wu-Yuan, D., Dan-Wei, M., Ying, X., Meng, X. and Fang, C. (2006) 'Characterization of a new stearoyl-acyl carrier protein desaturase gene from *Jatropha curcas*', *Biotechnology Letters*, 28, pp. 657–662.

Töpfer, R., Martini, N. and Schell, J. (1995) 'Modification of plant lipid synthesis', *Science*, 268, pp. 681-686.

Venegas-Calerón, M., Muro-Pastorb, A.M., Garcésa, R. and Martínez-Forcea, E. (2006) 'Functional characterization of a plastidial omega-3 desaturase from sunflower (*Helianthus annuus*) in cyanobacteria', *Plant Physiology and Biochemistry*, 44, pp. 517–525.

Yang, M.F., Liu, Y.J., Liu, Y., Chen, H., Chen, F. and Shen, S.H. (2009) 'Proteomic Analysis of Oil Mobilization in Seed Germination and Postgermination Development of *Jatropha curcas*', *Journal of Proteome Research*, 8, pp. 1441-1451.

Yin, D., Deng, S., Zhan, K. and Cui, D. (2007) 'High-oleic peanut oils produced by HpRNA-mediated gene silencing of oleate desaturase', *Plant Molecular Biology Reporter*, 25, (3), pp. 154-163.

Zheng, Z., Xia, Q., Dauk, M., Shen, W., Selvaraj, G. and Zou, J. (2003) 'Arabidopsis AtGPAT1, a Member of the Membrane-Bound Glycerol-3-Phosphate Acyltransferase Gene Family, Is Essential for Tapetum Differentiation and Male Fertility', *The Plant Cell*, 15, pp. 1872–1887.

Chapter 6: General discussion and conclusions

6.1 Genotype and phenotype analysis of *Jatropha curcas*

The present study shows that there is low genetic diversity among *J. curcas* accessions obtained from different parts of the globe, in contrast to the large variation in phenotypes (Chapter 2). This poses a series of questions:

What limits genetic variation in J. curcas?

An overlap between our results (Chapter 2) and a number of reports from different parts of the world assessing genetic diversity locally, regionally and internationally indicate that very low genetic divergence among *J. curcas* accessions spread around the globe (Basha and Sujatha 2007; Ram et al., 2008; Pamidiamarri et al., 2009). The reasons for this globally low genetic variability seen in *J. curcas* are not clear. Most likely, anthropogenic and environmental influences in generating genetic variability are missing because (a) it is not a crop, (b) as a well-surviving, undomesticated plant, it is highly stress tolerant (biotic/abiotic) due to adaptive genomic characters probably acquired before its global distribution, and (c) a limited stock has been vegetatively and/or apomictically propagated, since *J. curcas* is known to exhibit apomixis (Bhattacharya et al., 2005). The studies mentioned above indicate the limitations of using intra-specific breeding for *J. curcas* improvement; as exemplified by the failure to obtain low phorbol ester (PE) hybrids in crosses between *J. curcas* toxic accessions from India and the non-toxic Mexican variety (NTMA) (Sujatha et al., 2005).

Could epigenetics explain phenetic diversity in J. curcas?

Large differences in morphological and biochemical phenotype for seed size, and seed oil content and composition, in accessions with a narrow genetic base implicate environmental-mediated changes. Gene expression or repression under particular environmental conditions occur due to epigenetic changes in the DNA whereby the DNA sequence *per se* remains the same, While this results in low genetic diversity, gene expression shows appreciable changes manifested in a changed phenotype. At the molecular level gene silencing/expression through changes in DNA

methylation, and chromatin remodeling, inclusive of post-translational histone modifications (deacytylation/methylation), can be the cause of phenetic diversity despite limited genetic diversity (Grant-Downton and Dickinson, 2005). The underlying mechanism for phenotypic plasticity in *J. curcas* remains to be investigated. However, epigenetic phenomenon can be applied in order to increase seed and oil yield in *J. curcas* whist genetic improvement for generating elite lines progresses independently. Recently, Ghosh et al. (2010) reported that application of Paclobutrazol significantly reduced unwanted vegetative growth and increased seed yield to levels as high as 1127 % relative to the control. The authors speculated that Paclobutrazol might reduce unfavorable vegetative-growth by affecting the endogenous level of plant hormones. Paclobutrazol is a triazole type plant growth retardant which blocks gibberellin biosynthesis and is involved in reducing abscisic acid, ethylene and indole-3-acetic acid while increasing cytokinin levels. Thus, triazole promotes fruit set, but inhibits vegetative development. It is also known that cytokinin protects plants against abiotic stresses. Although detailed mechanisms of the effect of Paclobutrazol remain to be investigated, its extended effect on the level of hormones may in turn affect DNA and chromatin modifications which may be inherited. Initial studies conducted to assess differential DNA methylation using a modified RAPDs approach on plants from different eco-geographic regions of India revealed that DNA methylation did play a role in influencing the phenotype (Manish Raorane, personal communication).

Reducing/eliminating phorbol ester content using a transgenic approach versus conventional breeding

One of the major problems associated with *J. curcas* that it is contains several toxins especially phorbol esters which can be highly carcinogenic. As a consequence *J. curcas* cannot be an economically viable feedstock without reduction/elimination of the phorbol esters (Kohli et al., 2009). However, there is the potential to reduce/eliminate phorbol esters through conventional plant breeding or gene manipulation. *J. curcas* hybrid lines (*J. curcas* x *J. integerrima*) were successfully generated that exhibited low PE levels. A back cross with *J. curcas* non-toxic Mexican variety (NTMA) further reduced the PE levels comparable to NTMA variety. The mechanism of how interspecific hybrid lines exhibit low PE level remains to be investigated. However, we speculated that interspecific hybridization with *J. curcas* x *J. integerrima* leads to generation of valuable material due to heterozygosity at several loci including genes

involved in phorbol ester biosynthesis. Reduction/elimination of the phorbol ester in *J. curcas* can be done through gene manipulation approaches as discussed below.

6.2 Molecular biology of *J. curcas* oil bodies and oleosins

In the present study *J. curcas* oil bodies have been characterized. The stability of *J. curcas* oil bodies is maintained by proteins on the surface of the oil bodies (mainly oleosins) using electrostatic repulsion and steric hindrance, similar to as reported for other oilseed plants. The *J. curcas* oil bodies comprise of TAGs similar to other oilseed plants. Proteomic analyse showed that *J. curcas* oil bodies are composed of oleosins as major proteins with caleosin, stereoleosin, aquaporin, pollen oleosin as minor proteins. These results are similar to those reported for *Arabidopsis* oilbodies (Jolivet et al., 2004).

In the present study oleosin genes were characterized at the gene, transcript and protein levels. Three oleosin genes, designated *JcOleosin1, JcOleosin2* and *JcOleosin3,* were isolated and are now available under accession numbers EU234462, EU234463 and EU234464 at GenBank. Quantitative PCR analysis revealed that the isoforms exist as a single copy similar to other oilseed plants. *JcOleosin3* exhibits one intron located after hairpin structure as previously reported in *Arabidopsis* (Kim et al., 2002). These three oleosin genes are mainly expressed in the seed especially *JcOleosin3* which was expressed 4.9 and 4.6 fold higher compared to *JcOleosin2* and *JcOleosin1*, respectively. Transcript accumulation of all three oleosins gradually increased during fruit development up to 56-DAF but decreased near to the stage of fruit maturity (70-DAF). The pattern was similar to that reported for coffee oleosin (*CcOLE-2* to *CcOLE-4*; Simkin et al., 2006). It was also similar to the expression patterns of seed *oleosins* (S1, S2, and S3) during silique maturity in *Arabidopsis* (Kim et al., 2002). The three oleosins briefly exhibited a slight increase in transcript levels at an early stage of germination (1 to 7 days after imbibition). Later, the level of the transcripts significantly decreased until they disappeared at 14 DAI. However, the *JcOleosin3* transcript was detectable at moderate levels until 21 DAI. These results were similar to *oleosin* expression profiles reported for coffee during germination (Simkin et al., 2006). The presence of three olesoin genes in all samples investigated in this study (Chapter 3, table 3) suggests the importance of oleosin in *Jatropha* spp. Immunoassays and the intron sequence analysis of *Jcoelosin3* were used to assess the diversity of oleosins in *J. curcas* accessions, species and hybrids. This analysis identified SNPs associated with *JcOleosin3*

alleloforms. These results indicated *JcOleosin3* to be conserved among *Jatropha spp*. This study also attempted to determine the relationship between different organizations of the oleosin gene sequence to seed oil yield and FA profile with the assumption of the intron having an effect on expression levels. Results indicated that there was no direct correlation between the type of oleosin present, the oil yield or the oil profile.

Applications of oil bodies and oleosins

The oilbodies and oleosins have been explored for a variety of applications in biotechnology. These applications include the use of oilbobies as emulsifying agents (Deckers et al., 2001), the use of oleosins fusion proteins to produce recombinant proteins (Chaudhary et al., 1998; Moloney and Dalmia, 2004; Nykiforuk et al., 2006), and the use of oleosins for artificial oilbodies (AOBs) (Peng et al., 2004). Recently, Shimada et al. (2010) developed a rapid and non-destructive screening marker, named Fluorescence-Accumulating Seed Technology (FAST), for *Arabidopsis* using oelosin-fluorescence fusion stragety. This novel technique uses oleosin fused with green fluorescent or red flurescent protein to select for transgenic plants without destroying the plant and reducing the screening time process from 8 to 4 months. This new technology might be used to accelerate gene manipulation of *J. curcas* in near future.

6.3 *J. curcas* preferentially seed targeted promoter

In the present study a promoter sequence of 623 nucleotides upstream of the *JcOleosin3* gene was isolated (GenBank accession no. EU543568) by PCR assisted gene walking. The sequence exhibited motifs typical of promoters, for example a TATA box (TATAAAT) and a CAAT box (CCAAT) and several seed-specific motifs for example the DPBFCOREDCDC3 motif, the DOFCORE motif, the E-box motif, the SEFMOTIFGM75 motif, the RY motif, and the late pollen-specific motif. The activity of the 623 bp 5'-upstream regulatory region of this gene was determined in transgenic *Arabidopsis* plants using *β*-glucuronidase (GUS). Results showed that the *JcOleosin3* promoter directs expression of the *β-glucuronidase* gene in seed and late pollen development but not in leaves, root, stem, or silique wall.

Application of the J. curcas preferentially seed targeted promoter, JcOleosin3

The preferentially seed targeted oleosin promoter is a very useful tool for seed target gene expression. Oleosin promoters from several plants, for example *Brassica*, *Arabidopsis*, corn, *Perilla*, soybean, have been characterized. Genetic modification for the desired phenotype in *J. curcas* seed may depend on being able to limit expression of the gene of interest to the seed. Sunilkumar et al. (2006) successfully used siRNA to interrupt gossypol biosynthesis in cotton seed without affecting the gossypol profile or that of others related terpenoids in other part of the plants; this is important since these compounds are thought to play a role in plant defence. Therefore inthat study the authors silenced δ-*cadinene synthase* gene under the control of the cotton seed specific promoter, ά-*globulin B* gene.

The presence of phorbol ester in *J. curcas* seed is considered as a significant disadvantage limiting the potential of this oilseed plant. Reduced/eliminated-phorbol ester content is required. Gressel (2008) proposed transgenic approach to suppress phorbol ester levels in *J. curcas* by using an antisense/RNAi mediated approach to silence the phorbol diterpene synthase gene. In this instance the author emphasized that a seed specific promoter is needed in order to limit damage to the rest of the plant. The *J. curcas* 3-Hydroxy-3-methylglutaryl coenzyme A reductase (*JcHMGR*), a key regulatory enzyme for phorbol ester synthesis in the mevalonic acid (MVA) pathway has been isolated and characterized (Lin et al., 2009). Moreover, *J. curcas* farnesyl-diphosphate synthase (*JcFPS2*) and geranylgeranyl-diphosphate synthase (*JcGGR*) (genes related to phorbol ester biosynthesis) EST have been isolated (Costa et al., 2010). Therefore, the *JcOleosin3* preferentially seed targeted promoter might be used to silence the *JcHMGR, JcFPS2, JcGGR* genes only in *J. curcas* seed without affecting key plant hormones. This approach may also be applied to reduce/eliminate other toxins in *J. curcas* such as curcin, but only in the seed.

As previously discussed in Chapter 2, biodiesel property in *J. curcas* is determined by the ratio of oleic acid to linoleic acid. The FAD2 enzyme is responsible for conversion of oleic acid ($18{:}1\Delta^{9cis}$) to linoleic acid ($18{:}2\Delta^{9cis,12cis}$), contributing to the final O/L ratio in the seed oil. Suppression of FAD2 activity could result in accumulation of oleic acid and a reduction in linoleic acid (Dyer and Mullen, 2005). Disruption of FAD2 in transgenic peanut increased the oleic acid content up to 70% accompanied by a reduction of linoleic acid to 10% (Yin et al., 2007). Without disrupting the ratio of saturated/unsaturated fatty acids in other parts of the plant,

JcOleosin3 seed preferentially expressed promoter might be used to specifically silence the *Jcfad2* gene (DQ157776) in *J. curcas* seed resulting in appropriate O/L ratio. The same approach can be applied to increase *J. curcas* seed oil content. Seed overexpression of *J. curcas acc* (EU395776), *dgat2* (EU395774) genes can be achieved by fusion with the *JcOleosin3* preferentially seed targeted promoter.

6.4 *J. curcas* patatin-like lipases and presentation for a cycling scheme for triacylglecerides (TAGs) in *J. curcas*

Proteomic studies using 2-DE and MALDI-TOF MS revealed the presence of a patatin-like lipase in germinating seed but not in resting seed. These proteins showed sequence similarity and conserved domain to patatin lipase from *Arabidopsis thaliana* (Eastmond, 2004). Isolation of patatin-like lipase genes from *J. curcas* should be considered for future study by using degenerate primer-mediated PCR amplification. Peptide mass fingerprint obtained from LC-MS/MS can be used to design degenerate primers.

The TAGs cycling scheme for *J. curcas* has been presented, as shown in chapter 5 (Figure 5). The scheme presented in the this study provides preliminary information required for oil manipulation in *J. curcas*. This presented scheme is similar to that presented for *Arabidopsis* TAGs cycling model (Quettier and Eastmond, 2009). However, the scheme for TAGs cycling in *J. curcas* still needs to be refined and verified.

6.5 Conclusions

The following conclusions can be drawn from the present study:

1. Low genetic diversity exists in *J. curcas* from among global accessions, in contrast to a pronounced phenotypic variation in these accessions. An inter-specific breeding programme is an alternative way to broaden the genetic background of *J. curcas*.
2. Characterization of *J. curcas* oil bodies and oleosins at the gene, transcript, protein and metabolite levels illustrated similarities and differences in these structures within different *J. curcas* accessions and in comparision with other plants.
3. The *J. curcas* preferentially seed targeted promoter characterized in transgenic *Arabidopsis* supported previous studies for the role of different motifs and resulted in a

molecular tool which could be used for seed preferentially-expressed of other genes in *J. curcas*.

4. The *J. curcas* patatin-like lipase identified during seed germination was similar to other lipases known.

5. Based on bioinformatic and data-mining a TAGs lipid cycling scheme was presented.

6.6 Future prospects

Jatropha curcas has been the center of attention for several years as a potential source of biodiesel production. Several private sectors have been planting *J. curcas* on a large-scale but, unfortunately, in many regions of the world, especially in developing countries, this has been in a non-sustainable way (Achten et al., 2007). Without considerable genetic insight and good agronomic practice, the yield of *J. curcas* will remain unpredictable compared to other oilseed crops. Therefore, there are urgent research issues to be addressed to make *J. curcas* commercially sustainable.

The present study suggests that elite commercial *J. curcas* lines must be generated through either interspecific breeding or genetic manipulation due to the low genetic variation among accessions from around the globe. Whilst stable yield can be accomplished through good agronomic practices, further molecular-based research is required to produce commercially viable elite lines of *J. curcas* which exhibit low levels of phorbol esters, but high levels of oil, and oleic acids in the seeds. Genetic manipulation appears to be a promising approach to achieve this end. During the course of the project a preferentially seed targeted promoter, *JcOleosin3*, was isolated. This promoter will enable future studies to be carried out whereby toxin could be reduced/eliminated and the oil quality/quantity improved.

Based on experimental work carried out in this project together with bioinformatics, a scheme was presented for the cycling of TAGs. This presented scheme, together with established TAG cycling models for *Arabidopsis* and *Brassica* will enable future manipulation of the lipid profile in *J. curcas*.

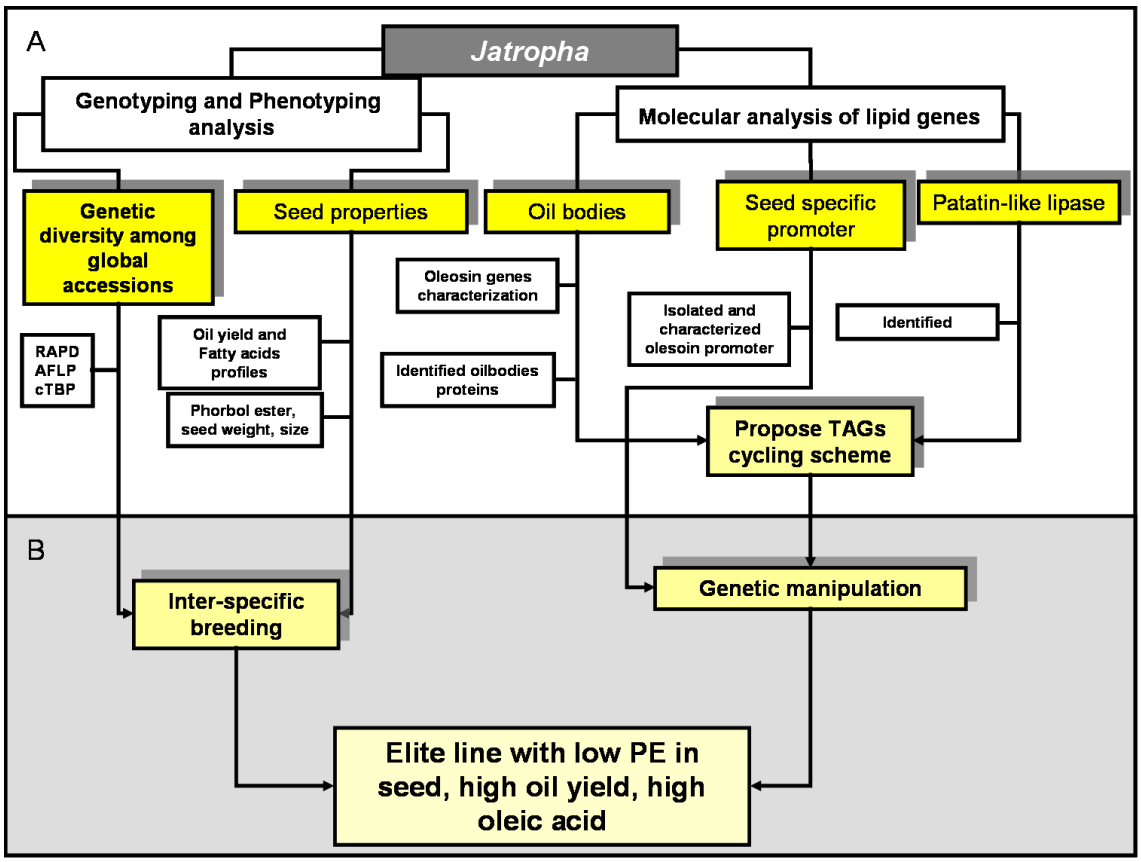


Figure 1 Summary flow chart of the present study (A). The genotype and phenotype of *J. curcas* and *Jatropha* spp from different regions of the world have been characterized. Results indicate inter-specific conventional breeding is the better choice to broaden the genetic background of *J. curcas* before producing elite lines. *Jatropha* oil bodies and the seed preferentially-expressed promoter have been isolated and characterized. Patatin-like lipase has been identified. TAGs lipid cycling scheme has been proposed based on bioinformatics and data obtained from the present study. The molecular data can be used for genetic manipulation to generate desired elite lines such as those with low phorbol ester, high oil yield, and high oleic acids in near future (B).

6.7 References:

Achten, W.M.J., Mathijs, E., Verchot, L., Singh, V.P., Aerts, R. and Muys, B. (2007) '*Jatropha* biodiesel fueling sustainability?', *Biofuels, Bioproducts and Biorefining*, 1, pp. 283–291.

Basha, S.D. and 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, pp. 375-386.

Bhattacharya, A., Datta, K. and Datta, S.K. (2005) 'Floral biology, floral resource constraints and pollination limitation in *Jatropha curcas* L', *Pakistan Journal of Biological Sciences*, 8, pp. 456-460.

Costa, G.G., Cardoso, K.C., Del Bem, L.E., Lima, A.C., Cunha, M.A., de Campos-Leite, L., Vicentini, R., Papes, F., Moreira, R.C., Yunes, J.A., Campos, F.A., Da Silva, M.J. (2010) 'Transcriptome analysis of the oil-rich seed of the bioenergy crop *Jatropha curcas* L.', *BMC Genomics*, 6, 11:462. doi:10.1186/1471-2164-11-462

Chaudhary, S., Parmenter, D.L. and Moloney, M.M. (1998) Transgenic *Brassica carinata* as a vehicle for the production of recombinant proteins in seeds *Plant Cell Rep* 17:195–200.

Deckers, H.M., Van Rooijen, G., Boothe, J., Goll, J., Mahmoud, S. and Moloney, M.M. Sembiosys Genetics Inc. Uses of oil bodies. United States patent US 6210742. 2001 April.

Dyer, J.M. and Mullen, R.T. (2005) 'Development and potential of genetically engineered oilseeds', *Seed Science Research*, 15, (4), pp. 255-267.

Eastmond, P.J. (2004) 'Cloning and characterization of the acid lipase from castor beans', *The Journal of Biological Chemistry*, 279, pp. 45540–45545.

Ghosh, A., Chikara, J., Chaudhary, D.R., Prakash, A.R., Boricha, G. and Zala, A. (2010) 'Paclobutrazol arrests vegetative growth and unveils unexpressed yield potential of *Jatropha curcas*', *Journal of Plant Growth Regulation*, pp. 1-9. Article in Press.

Grant-Downton, R.T. and Dickinson, H.G. (2005) 'Epigenetics and its implications for plant biology: 1. The epigenetic network in plants', *Ann. Bot.*, 96, pp. 1143–1164.

Gressel, J. (2008) 'Transgenics are imperative for biofuel crops', *Plant Science*, 174, pp. 246-263.

Jolivet, P., Roux, E., D'Andrea, S., Davanture, M., Negroni, L., Zivy, M., Chardot, T., (2004) 'Protein composition of oilbodies in *Arabidopsis thaliana* ecotype WS', *Plant Physiology and Biochemistry*, 42, pp. 501-509.

Kim, H.U., Hsieh, K., Ratnayake, C., Huang, A.H.C. (2002) 'A novel group of oleosins is present inside the pollen of *Arabidopsis*', *J Biol. Chem*., 277, pp. 22677-22684.

Kohli, A., Raorane, M., Popluechai, S., Kannan, U., Syers, K.J. and O'Donnell, AG. (2009) '*Jatropha curcas* as a novel, non-edible oilseed plant for biodiesel', In: Ferry N, Gatehouse AMR, eds. *Environmental impact of genetically modified novel crops*. Chapter 14. CAB International, London, UK. p 294 -322.

Lin, J., Jin, Y.J., Zhou, M.Q., Zhou, X. and Wang, J.Y. (2009) 'Molecular cloning, characterization and functional analysis of a 3-hydroxy-3-methylglutaryl coenzyme A reductase gene from *Jatropha curcas*', *African Journal of Biotechnology*, 8, pp. 3455-3462.

Moloney, M.M. and Dalmia, B.K. Sembiosys Genetics Inc. Preparation of thioredoxin and thioredoxin reductase proteins on oil bodies. United States patent US 6750046. Jun 2004.

Nykiforuk, C.L., Boothe, J.G., Murray, E.W., Keon, R.G., Goren, H.J., Markley, N.A. and Moloney, M.M. (2006) 'Transgenic expression and recovery of biologically active recombinant human insulin from *Arabidopsis thaliana* seeds', *Plant Biotechnol. J.* 4, pp. 77–85.

Pamidiamarri, S., Nirali Pandya, N., Reddy, M.P. and Radhakrishnan, T. (2009) 'Comparative study of interspecific genetic divergence and phylogenic analysis of genus *Jatropha* by RAPD and AFLP: Genetic divergence and phylogenic analysis of genus *Jatropha*', *Molecular Biology Reports*, 35, (6), pp. 901-907.

Peng, C.C., Chen, J.C.F., Shyu, D.J.H., Chen, M.J. and Tzen, J.T.C. (2004) 'A system for purification of recombinant proteins in *Escherichia coli* via artificial oil bodies constituted with their oleosin-fused polypeptide', *J Biotechnol*, 111, pp. 51–57.

Quettier, A-L. and Eastmond, P.J. (2009) 'Storage oil hydrolysis during early seedling growth', *Plant Physiology and Biochemistry*, 47, (6), pp 485-490.

Rademacher, W. (2000) 'Growth retardants: Effects on gibberellin biosynthesis and other metabolic pathways', *Annual Review of Plant Biology*, 51, pp. 501-531.

Ram, S.G., Parthiban, K.T., Kumar, R.S., Thiruvengadam, V. and Paramathma, M. (2008) 'Genetic diversity among *Jatropha* species as revealed by RAPD Markers' *Genetic Resources and Crop Evolution*, 55, (6), pp 803-809.

Shimada, T.L., Shimada, T. and Hara-Nishimura, I. (2010) 'A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*', *The Plant Journal*, 61, (3), pp. 519-528.

Simkin, A.J., Qian, T., Victoria, C., Franck, M., Mohamed, B.R., Lin, C., Steve, T. and James, M. (2006) 'Oleosin gene family of *Coffea canephora*: Quantitative expression analysis of five *oleosins* in developing and germinating coffee grain', *Journal of Plant Physiology*, 163, pp. 691-708.

Sujatha, M., Makkar, H.P.S. and Becker, K. (2005) ‘Shoot bud proliferation from axillary nodes and leaf sections of non-toxic *Jatropha curcas* L’, *Plant Growth Regulation*, 47, pp. 83-90.

Sunilkumar, G., Campbell, L.M., Puckhaber, L., Stipanovic, R.D. and Rathore, K.S. (2006) ‘Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol’, *PNAS, 103,* (48), pp. 18054-18059.

Yin, D., Deng, S., Zhan, K. and Cui, D. (2007) ‘High-oleic peanut oils produced by HpRNA-mediated gene silencing of oleate desaturase’, *Plant Molecular Biology Reporter*, 25, (3), pp. 154-163.