Cyto-genetic studies in elite genotype of *Pongamia pinnata* (L.), a versatile legume

A THESIS

submitted by

AADI MOOLAM RAMESH

for the award of the degree

of

DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOTECHNOLOGY

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI-781039, ASSAM, INDIA.

January, 2014

Cyto-Genetic Studies in Elite Genotype of Pongamia pinnata (L.), a Versatile Legume

AADI MOOLAM RAMESH
Roll no: 09610610

Department of Biotechnology

A thesis submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Indian Institute of Technology Guwahati,
Guwahati – 781039, Assam, India
January, 2014

TH-1255_09610610

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
Department of Biotechnology
Guwahati – 781039

STATEMENT

I do hereby declare that the matter embodied in this thesis is the result of investigations carried out by me in the Department of Biotechnology, Indian Institute of Technology Guwahati, India, under the guidance of Dr. Latha Rangan and Dr. Bithiah G Jaganathan.

In keeping with the general practice of reporting scientific observations, due acknowledgements have been made wherever the work described is based on the findings of other investigators.

January, 2014 **Aadi Moolam Ramesh**

TH-1255_09610610

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
Department of Biotechnology
Guwahati – 781039

January, 2014

CERTIFICATE

It is certified that the work described in this thesis, entitled "**Cyto-genetic studies in elite genotype of *Pongamia pinnata* (L.), a versatile legume**", done by Mr. Aadi Moolam Ramesh for the award of degree of Doctor of Philosophy is an authentic record of the results obtained from the research work carried out under our supervision in the Department of Biotechnology, Indian Institute of Technology Guwahati, India, and this work has not been submitted elsewhere for a degree.

Dr. Latha Rangan	Dr. Bithiah G Jaganathan
Associate Professor	Assistant Professor
(Thesis Supervisor)	(Thesis Co-Supervisor)

TH-1255_09610610

Dedicated to my beloved parents

TH-1255_09610610

ACKNOWLEDGEMENT

It is with my deepest sense of appreciation that I express my heartiest acknowledgement to my research supervisor, Dr. Latha Rangan, Department of Biotechnology, for her guidance, encouragement, consistent support, discussions and providing me with so many learning opportunities, which helped me in the successful culmination of the present endeavor. I must acknowledge the unconditional freedom to think, plan, execute and express, that I was given in every step of my research work, while keeping faith and confidence on my capabilities.

I would like to profound regards my co-supervisor Dr. B.G. Jaganathan and doctoral committee, chairperson, Dr. A. Ramesh; members, Dr. A. Limaye and Prof. P. Mahanta for their constructive criticism and precious suggestions.

I am thankful to HODs of Department of Biotechnology IIT Guwahati for providing instrument and academic facilities during my PhD tenure. I acknowledge with gratitude for the kind corporation and facilities provided by CIF and Sila forest Office, North Guwahati to accomplish my PhD project.

I gratefully acknowledge all the faculty, technical staff members (Mr. Nurul, Mr. Raghu, Mr. Niranjan, Mrs. Anitha, Mrs. Prarthana and Mrs. Rashmi) and non teaching staffs of my department for their affection and support.

The financial support provided by Indian Institute of Technology, Guwahati, DIT, DST and DBT is gratefully acknowledged.

I would like to thank to Dr. A. Parida, Director of Biotechnology, MSSRF for allowing me to work in his lab and use the necessary tools. I am extremely grateful to Dr. Sankar, Ms. Divya, Mr. Ashok and other members of MSSRF for their support which accelerated my enthusiasm and in turn lighted my spirit to complete such an elaborated task pertaining to PhD thesis.

I appreciate and thank my research team members Tushar, Supriyo, Sudipta, Sarath, Rahul, Anuma, Reshmi and Rimjhim for their suggestions, cooperation, time, kindness and their help to me in various ways throughout my project. I also thank my seniors Dr. Vigya and Dr. Archana for sharing their knowledge and providing useful counsel and help.

I owe special gratitude to my teachers Dr. Rajan Sriraman & Dr. Mohan Subramanian, Indian Immunological Ltd., Hyderabad. I am also indebted to my sincere friends Reddy, Santhosh, Surya, Nani, Dr. Ganesh at IITG for their nice company and all my friends for their love and affection.

Special thanks to my parents (Amma, Bapu, Pinni), brothers (Santhosh, Raja Shekar) and sister (Vasanta) for being near to my heart and providing the blessings, moral support and encouragement. This work would not have been possible without my family's love and support.

Thanks to the Almighty "GOD" whose blessings have always enabled me to perceive and pursue higher studies in my life and proceeding strength to take up and complete the task successfully.

Date: **Aadi Moolam Ramesh**

TH-1255_09610610

CONTENTS

Abstract	i
Graphical abstract	ii
Synopsis	iii-v
Abbreviations	vi-viii
Units	ix
List of Figures	x-xii
List of Tables	xiii
1. INTRODUCTION	1-8
1.1 Background	1
1.2 *Pongamia pinnata* and distribution	2
1.3 Botanic description	3
1.4 Applications	5
1.5 Current research in *Pongamia*	5
1.6 Specific objectives	7
2. LITERATURE REVIEW	9-24
2.1 Background	9
2.2 cDNA library construction	12
2.3 Fatty acid desaturases	14
2.4 Nuclear genome content and chromosome number determination	18
2.4.1 Buffer compositions	22
2.4.2 Chromosome study	24
3. cDNA LIBRARY CONSTRUCTION FROM SEEDS OF *Pongamia pinnata* (ELITE GENOTYPE) AND SCREENING FOR FULL LENGTH GENES	25-49
3.1 Introduction	25
3.2 Material and methods	27
3.2.1 Plant tissue collection	27
3.2.2 RNA extraction using CTAB-GITC	27
3.2.3 RNA quantification	30

TH-1255_09610610

3.2.4 Isolation of mRNA from total RNA 30
3.2.5 Formaldehyde agarose RNA gel electrophoresis 30
3.2.6 cDNA synthesis and cDNA library construction 30
3.2.7 First strand cDNA synthesis 31
3.2.8 Second strand cDNA synthesis 31
3.2.9 Sfi digestion of ds-cDNA 33
3.2.10 Size fractionation of Sfi digested ds cDNA 33
3.2.11 Ligation of ds cDNA with the plasmid vector 34
3.2.12 Transformation of ligated cDNA and its efficiency 34
3.2.13 Transfer of bacterial colonies onto nitrocellulose filter 35
3.2.14 Replication of nitrocellulose filter from master filter 35
3.2.15 Lysis of bacterial colonies on nitrocellulose replica 36
3.2.16 Amplification of fatty acid biosynthetic genes (partial) for probe preparation 36
3.2.17 32P radio-labeling of probe using random oligo-nucleotides as primers 37
3.2.18 Hybridization 37
3.2.19 Autoradiography and identification of positive Clones 38
3.2.20 Processing and analysis of cDNA sequences 38
3.3 Results and discussion 39
3.3.1 General precautions for RNA handling 39
3.3.2 Total RNA isolation 39
3.3.3 cDNA library construction 42
3.3.4 Size distribution, evaluation of clone inserts and screening for full length genes 44
3.4 Conclusion 49
4. STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF Δ9-STEROYL-ACP- DESATURASE GENE 50-78
4.1 Introduction 50
4.2 Material and methods 52
4.2.1 Plant material and growth 52
4.2.2 Nucleic acid (DNA & RNA) isolation and quantification

TH-1255_09610610

4.2.3 Complementary DNA cloning and sequencing 53
4.2.4 Sequence analysis 54
4.2.5 RNA quantification, primer design and transcript analysis by RT-PCR 54
4.2.6 Real-time RT-PCR 55
4.2.7 Southern hybridization 56
4.2.8. Structural characterization of PpSAD by homology modeling 57
4.3 Results and discussion 57
4.3.1 Characterization of the P. pinnata SAD cDNA 57
4.3.2 RNA expression studies 65
4.3.3 Southern analysis 70
4.3.4. Structural characterization of PpSAD by homology modeling 70
4.3.5. Structural assessment of the 3D model 76
4.4 Conclusion 78
5. FUNCTIONAL CHARACTERIZATION OF FATTY ACID DESATURASE (FAD2) GENES 79-100
5.1 Introduction 79
5.2 Material and methods 80
5.2.1 Plant materials 80
5.2.2 RNA isolation and quantification 81
5.2.3 RNA transcript analysis and expression studies 81
5.2.4 DNA extraction and quantification 82
5.2.5 Blot hybridization 82
5.3 Results and discussion 83
5.3.1 Isolation and sequence analysis of two microsomal fatty acid desaturase genes from *P. pinnata* 83
5.3.2 Genomic organization and tissue specific expression of PpFAD2 93
5.4 Conclusion 100
6. NUCLEAR GENOME ESTIMATION AND CHROMOSOME NUMBER DETERMINATION IN *P. pinnata* 101-143
6.1 Introduction 101
6.2 Material and methods 103

TH-1255_09610610

6.2.1 Plant materials and culture 103
6.2.2 Sample preparation and isolation of nuclei 104
6.2.3 Flow cytometric analyses 107
6.2.4 Test for inhibitors 108
6.2.5 Test for concentration effect of inhibitors 110
6.2.6 Effect of antioxidant on flow cytometric estimation of nuclear DNA content 110
6.2.7 Statistical analysis 110
6.2.8 Cytological analysis 111
6.2.8.1 Collection and storage of root tips 111
6.2.8.2 Pre-treatment and fixation 111
6.2.8.3 Hydrolysis and staining 111
6.2.8.4 Slide preparation, squash and observation 111
6.3 Results and discussion 113
6.3.1 Test for robustness 113
6.3.2 Reference standards for flow cytometry analysis 115
6.3.3 Test for inhibitors 116
6.3.4 Pseudo-internal standardization 120
6.3.5 Internal standardization 122
6.3.6 Effect of antioxidants 124
6.3.7 Test for homogeneity 129
6.3.8 Cytological analysis 138
6.3.8.1 Collection of root tips 138
6.3.8.2 Pretreatment 139
6.3.8.3 Hydrolysis and staining 139
6.3.8.4 Slide preparation, squash, and observation 139
6.4 Conclusion 142
7. SUMMERY 144-146
8. REFERENCES 147-161
9. PUBLICATIONS AND PARTICIPATIONS 162-163
10. BIOGRAPHY 164

TH-1255_09610610

ABSTRACT

The leguminous tree, *Pongamia pinnata* (L.) Pierre is one of the major biofuel crop which produces oilseed suitable for biodiesel production. The potentiality of *Pongamia,* as a sustainable source of feedstock for the biodiesel industry is dependent on an extensive knowledge of the genome structure of the plant.

In the present study, earlier characterized (based on vegetative and reproductive characters) candidate plus tree (CPT-NGPP46) of *Pongamia* germplasm was collected from Sila forest range, North Guwahati, Assam for further experimental study. cDNA library was constructed from early immature seeds (90-DAF) and full length genes that are involved in desaturation of fatty acid biosynthesis were fished out.

Full length cDNA clone encoding for steroyl-ACP desaturase (SAD) gene was isolated from the cDNA library characterized based on the sequence similarities in NCBI database. The clone (Pp*SAD*) contains an open reading frame of 1182 bp and shares similarity with SAD from other plants. Characteristic of the deduced protein were predicted and when analyzed using molecular modeling, its 3-Dimensional structure strongly resembled the crystal structure of *Ricinus communis* (RSAD). Southern blot and expression analysis using quantitative real time PCR indicated that Pp*SAD* is a multiple copy gene having marked distinct expression during different stages of seed development. The PpSAD expression analysis, combined with existing research, suggests that SAD may be involved in the regulation of plant seed growth and development.

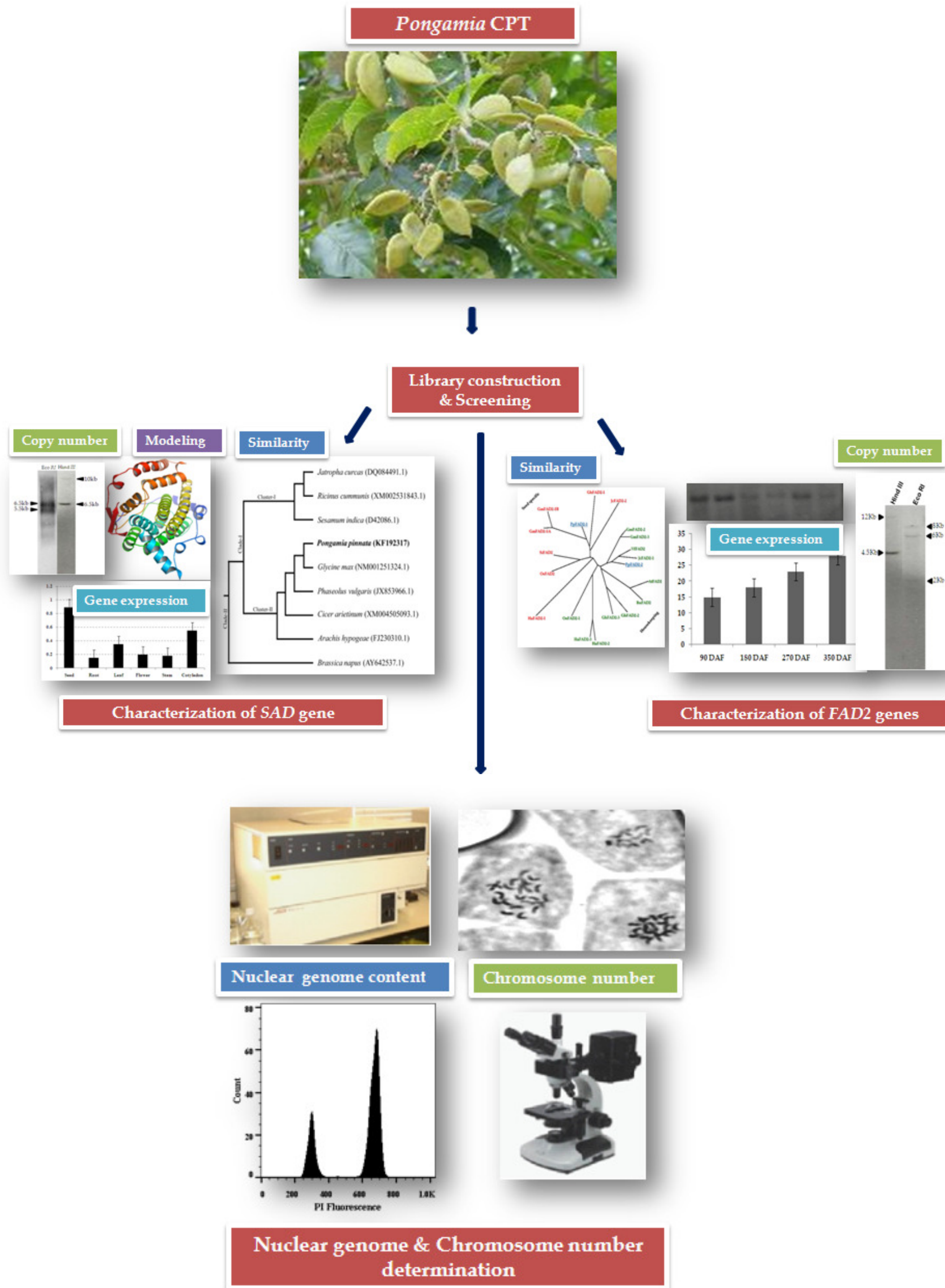
Further two full length cDNA clones were also isolated from the library coding for fatty acid desaturase enzymes (FAD2-1 and FAD2-2). The deduced amino acid sequence revealed that both the sequences have 83 % homology among them, contain eight histidines essential for membrane-bound desaturases activity and motif in the C-terminal for endoplasmic reticulum retention which are conserved in other oil yielding plant species too. Semi-quantitative RT-PCR and northern blot analysis showed that *PpFAD2-1* gene is restricted to the various seed developmental stages and *PpFAD2-2* gene is constitutively expressed in both vegetative tissues and developing seeds. Southern hybridization revealed that *PpFAD2* is a multi copy gene.

Flow cytometry, with propidium iodide (PI) as the DNA stain, was used to estimate the nuclear DNA content of *P. pinnata,* with respect to *Zea mays* `CE-777' as standard. The internal and pseudo-internal standardization was followed on account of the inhibitory effect of secondary compounds on PI intercalation. The antioxidants (PVP-40 and β-mercaptoethanol) were added to the nuclear isolation buffer for the reduction of inhibitory effect of *P. pinnata* cytosol. Nuclear DNA content estimation was done for *Pongamia* leaves from different altitudes (37-117m height from sea level) of Assam. Flow cytometry analysis indicated that the nuclear DNA content of *P. pinnata* is 2.66 pg with predicted 1C value of 1300 Mbp using *Z. mays* as standard. Coefficient of variation in flow cytometric analysis was within the limit of 5 % indicating that the results were reliable. The similar results were obtained with little variation in nuclear genome content when analyzed with different tissue types except flowers. The nuclear genome content was also estimated in individuals collected from various parts of Assam and India showing slight variation which could be due to difference in repetitive sequence elements and chromatin condensation. Somatic chromosome numbers were counted from root-tip cells and was found to be 2n=22 corresponding to the diploid level (x=11) true-to-type legumes. A decreasing trend in the nuclear DNA content was observed for the species of different altitudes. The study also focused on the relationship between average chromosome length and the DNA contents of five individuals collected from Assam. The positive correlation between nuclear DNA content and average chromosome length in all five individuals were observed.

Isolation and characterization of three fatty acid biosynthetic genes from cDNA library of *Pongamia* along with estimation of genome content and chromosome study will promote advance research and development towards understanding the germplasm to improve the key traits in *Pongamia*.

TH-1255_09610610

GRAPHICAL ABSTRACT



SYNOPSIS

Nowadays, biofuel crops are getting much attention due to its multiple applications in replacing fossil fuels and alleviating the environmental pollution leading to global warming. Exploration of genome based knowledge towards understanding cyto-genetic studies may improve the possibility of obtaining key traits in biofuel crops.

Pongamia pinnata is a commercially important tree species used to produce biofuels that can grow on waste or unproductive land. Extensive characterization of *Pongamia* plant has been studied from North Guwahati, including candidate plus tree (CPT) identification, germplasm utilization and mass multiplication, genetic diversity study, developmental expression of storage protein profile and physico-chemical characterization of crude oil (Kesari et al. 2008, 2009, 2010, 2012). Though it is characterized well and has superior characters, there is a lack of information at molecular and cyto-genetic level. There is a greater need to find out expression profiles of the genes involved in desaturation of fatty acid biosynthesis, nuclear genome content and chromosome number.

The present investigation was focused on cyto-genetic studies in elite genotype of *P. pinnata* that promotes advance research and development towards understanding the germplasm to improve the key qualities in *Pongamia* with the following objectives.

- cDNA library construction from early immature seeds (90-DAF) of elite genotype of *Pongamia* (earlier characterized based on vegetative and reproductive characteres) to screen for full length genes that are involved in fatty acid biosynthetic desaturation pathway.
- Structural and functional characterization of steroyl-ACP-desaturase (SAD) gene obtained from cDNA library constructed from seeds of elite genotype of *P. pinnata*.
- Functional characterization of fatty acid desaturase (FAD2) genes.
- Total nuclear genome content estimation of elite genotype of *Pongamia* by using flow cytometer. Flow cytometry was also carried out to test the homogeneity of nuclear genome content between the various tissues of elite genotype, samples collected from various parts of Assam, India and *in vitro* raised tissue culture plants and the somatic chromosome number determination.

Overall the thesis has been divided into seven chapters as described below. Results of the current investigation are presented in four chapters (3-6). These chapters are preceded by **Chapter One** which gives a brief introduction of the present research.

TH-1255_09610610

Chapter Two which includes the detailed review of literature emphasizing the biofuel crops and its role in agroforestry system, distribution of *Pongamia*, fatty acid biosynthesis in plants, nuclear genome estimation towards development of high quality genotypes.

Chapter Three describes the standardization of RNA extraction protocol that yields good quality of RNA from poly phenolics containing *Pongamia* tissues. cDNA library was constructed from early immature seeds of *Pongamia* and screening of fatty acid biosynthetic genes were performed by using partial gene sequences as probes. Sequence analysis reveals the isolated genes are of full length, involved in fatty acid biosynthesis desaturation process.

Chapter Four explains the characterization of steroyl-ACP-desaturase (SAD) gene. The results revealed that the isolated gene is having complete coding region with 3' and 5' UTR regions. The full length gene sequence showed sequence similarity with SAD gene of other oil yielding plants. Comparative homology modeling, validation tools showed PpSAD is a good model. Further, transcriptional analysis of PpSAD studied using real-time PCR revealed that the gene shows various expression levels in different tissues and increased as seed grows. Southern blot hybridization revealed that the *Pongamia* genome has more than two copies of the gene.

Chapter Five unveils the functional characterization of two fatty acid desaturase genes that are screened from the cDNA library. Transcript analysis revealed that both genes (PpFAD2-1 & PpFAD2-2) shows different expression pattern. PpFAD2-1 shows tissue specific expression which is restricted to seed development stages. Whereas PpFAD2-2 shows constitutive expression pattern and the same was confirmed by performing real-time PCR and northern blot analysis. Furthermore the gene copy number was also studied and was found to be containing more than two copies of FAD2 in *Pongamia* genome.

Chapter Six describes the nuclear genome estimation in CPT, various tissues of CPT, individuals collected from Assam and different location from India. Results revealed that the genome content is 2.66 pg and there is a little variation observed between the tissues and individuals collected. Decrease in genome in flowers may be due to the presence of floral components and secondary metabolites which may reduce the binding properties of propidium iodide (PI) to nucleus. Slight variation in nuclear genome content between the individuals may be due to the arrangement of repetitive sequence elements and chromatin condensation. Furthermore somatic chromosome number in five individuals revealed consistent chromosome number and there exist positive correlation between chromosome length and nuclear genome content for five individuals collected from Assam.

Chapter Seven unveils the summary and future prospective of the present study.

TH-1255_09610610

The work described in the thesis and related research carried out during the period of doctoral research has been peer reviewed and resulted in following international journal publications.

Published articles:

- Aadi Moolam Ramesh, Vigya Kesari, Latha Rangan (2013). Characterization of a stearoyl-acyl carrier protein desaturase gene from potential biofuel plant, *Pongamia pinnata*. L. *Gene*. 542:113-121
- Aadi Moolam Ramesh, Supriyo Basak, Rimjhim Roy Chowdhury, Latha Rangan (2013). Development of Flow Cytometric Protocol for Nuclear DNA Content Estimation and Determination of Chromosome Number in *Pongamia pinnata* L., a Valuable Biodiesel Plant. Applied Biochemistry and Biotechnology. 172(1):533-548.
- Rimjhim Roy Choudhury, Supriyo Basak, Aadi Moolam Ramesh and Latha Rangan (2013). Nuclear DNA content of *Pongamia pinnata* L. and genome size stability of in vitro regenerated plantlets. *Protoplasma*. 251:703-709.

Manuscripts under review:

- Ramesh et al., (2014). Molecular characterization of microsomal oleoate desaturase genes from *Pongamia pinnata*. *Gene* (Under review).

Book chapter:

- Vigya Kesari, Aadi Moolam Ramesh and Latha Rangan. "*Pongamia pinnata*: The biodiesel plant biology, tissue culture and genetic enhancement" *Biotechnological Applications for Environmental Protection*. 2013 (Springer), Accepted.

GenBank submissions:

- Ramesh, A.M., Kesari, V. and Rangan, L. NCBI GenBank Accession number KF192317- *Millettia pinnata* steroyl-ACP desaturase mRNA, complete cds.
- Ramesh, A.M., Kesari, V. and Rangan, L. NCBI Geneank Accession number KF651985- *Millettia pinnata* microsomal omega-6-desaturase (FAD2-1) mRNA, complete cds; nuclear gene for microsomal product.
- Ramesh, A.M., Kesari, V. and Rangan, L. NCBI Geneank Accession number KF651986- *Millettia pinnata* microsomal omega-6-desaturase (FAD2-2) mRNA, complete cds; nuclear gene for microsomal product.

TH-1255_09610610

ABBREVIATIONS

ACP	Acyl Carrier Protein
ASTM	American Society for Testing and Materials
ATP	Adenosine Triphosphate
AVP	Acclimatized in vitro plants
cDNA	Complementary DNA
β-ME	Beta mercaptoethanol
Cl	Chloride
Cam	Chloramphenicol
CPTs	Candidate Plus Trees
CTAB	Cetyl Trimethyl Ammonium Bromide
DAF	Days After Flowering
DBT	Department of Biotechnology
DNA	Deoxyribo Nucleic Acid
DEPC	Diethylpyrocarbonate
dNTPs	Deoxy Nucleotide Triphosphates
E. coli	*Escherichia coli*
EDTA	Ethylene Diamine Tetra-acetic Acid
EtBr	Ethidium Bromide
EtOH	Ethyl Alcohol
FAD	Fatty Acid Desaturase
GC	Gas Chromatography
GC-MS	Gas Chromatography Mass spectrophotometry
GITC	Guanidinium thiocyanate
GOI	Government of India
IIT	Indian Institute of Technology
IRP	In vitro rooted plantlets
Ka	Kanamycin
KCl	Potassium Chloride
LD PCR	Long distance polymerase reaction
MFI	Mean/Median Fluorescence intensity

TH-1255_09610610

MEGA	Molecular Evolutionary Genetics Analysis
$MgCl_2$	Magnesium Chloride
Mn	Manganese
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	Messenger RNA
NaCl	Sodium Chloride
NaOH	Sodium Hydroxide
NCBI	National Centre for Biotechnology Information
NG	North Guwahati
NGO	Non Governmental Organization
NGPP	North Guwahati *Pongamia pinnata*
NP-40	Nonidet P-40
NS	Nodal segments
NZY	NZ amine (casein hydrolysate) Yeast extract Medium
PCR	Polymerase Chain Reaction
Ph	Potentiometric Hydrogen Ion Concentration
PI	Principal Investigators
PVP	Polyvinylpyrrolidone
R&D	Research & Development
RNA	Ribose Nucleic Acid
rRNA	Ribosomal RNA
RT-PCR	Real Time Polymerase Chain Reaction
SAD	Stearoyl-ACP desaturase
SD	Standard Deviation
SDS	Sodium Dodecyl Sulfate
SE	Standard Error
SMART	Switching Mechanism at 5' End of RNA Template
TAE	Tris acetic acid Ethylenediaminetetraacetic acid
Tet	Tetracycline
TAGs	Triacylglycerols
TE	Tris Ethylenediaminetetraacetic acid
Tris HCl	Tris Hydrochloride

TH-1255_09610610

TYE	Trypton yeast extract
UV	Ultra Violate
VI	Volume Index
WP	Woody plant
YEM	Yeast Extract Monitol
min	Minutes
mg/ml	Milligram per millilitre
mL	Microliters
mM	Millimolar
mM	Millimolar
ppm	Parts per million
g	Grams
w/v	Weight/Volume
wt	Weight
µg	Micrograms
µM	Micro Meter
µg/µl	Microgram per microlitre
µg/ml	Microgram per milliliter
µl	Microlitre

TH-1255_09610610

UNITS

Å	Angstrom
cm	Centi meters
^oC	Degree Centigrade
h	Hours
ha	Hectare
kDa	Kilo Dalton
Kb	Kilo base
m	Meter
M	Molar
Mbp	Mega base pairs
mg	Milligram
min	Minutes
mg/ml	Milligram per millilitre
mL	Microliters
mM	Millimolar
mM	Millimolar
ppm	Parts per million
pg	Picogram
g	Grams
w/v	Weight/Volume
wt	Weight
µg	Micrograms
µM	Micro Meter
µg/µl	Microgram per microlitre
µg/ml	Microgram per milliliter
µl	Microlitre

TH-1255_09610610

LIST OF FIGURES

Figure no.	Figure description	Page no.
1.1	*P. pinnata* habit and habitat.	4
2.1	Various applications of *P. pinnata*	10
2.2	Preparation and Screening of a cDNA library	13
2.3	Fatty acid biosynthesis in plants	17
2.4	Sequential desaturation of fatty acids	17
2.5	A cell biology and flow cytometry interface for nuclear genome estimation	20
3.1	Flow diagram of the cDNA library construction	28
3.2	Flow chart of the Creator™ SMART™ cDNA Library construction kit protocol	32
3.3	Electrophoresis profile of a *Pongamia* total RNA samples extracted from different methods i.e., Trizol, GITC and CTAB-GITC	40
3.4	Graph depicts the UV-Spectrophotometric readings of total RNA	40
3.5	Purified mRNA as smear	43
3.6	Amplified cDNA profile after 22 cycles of PCR	43
3.7	*E. coli* growth on 30 µg/ml chloramphenicol selective medium after platings with various transformed cell suspensions	45
3.8	A 1.2 % agarose gel electrophoresis of PCR products of cDNA inserts selected randomly from the amplified cDNA library	46
3.9	High density filters of the early immature seed *P. pinnata* cDNA lilbrary with FAD2 and SAD probes	48
4.1	A complete nucleotide sequence and deduced polypeptide of SAD cDNA from *P. pinnata*	59
4.2	Amino acid sequence alignment of PpSAD from *P. pinnata* with SAD from various species	61
4.3	Protein structural analysis	63
4.4	Hydropathy plot of the Pp*SAD* gene product	64
4.5	Phylogenetic analysis of PpSAD proteins	66

TH-1255_09610610

Figure no.	Figure description	Page no.
4.6	RNA expression studies. (A) Total RNA extracted from various tissues of *P. pinnata*. (B) Semi quantitative analysis of steroyl-acp-desaturase gene showing expression in all tissues	68
4.7	Quantitative analysis of Pp*SAD* expression in *P. pinnata*	69
4.8	Southern blotting analysis of gene Pp*SAD* from *P. pinnata*	71
4.9	Sequence alignment between PpSAD (KF192317) and template 1AFR (*R. communis*).	72
4.10	Secondary structure of the PpSAD predicted by homology modeling	73
4.11	(A) Three-dimensional model of PpSAD deduced by Swiss-Model and Swiss-Pdb-viewer. (B) Superimposition of target-PpSAD (cyan) and templat-1AFR (Green) structure	74
4.12	Ribbon diagram of the PpSAD (cyan) predicted by homology modeling	75
4.13	Ramachandran plot for the PpSAD model	77
4.14	3D Profile of PpSAD using ERRAT server.	77
5.1	A complete nucleotide sequence and deduced polypeptide of (A) FAD2-1 and (B) FAD2-2 cDNA from *P. pinnata*.	85
5.2	Comparison of the deduced amino acid sequences of two microsomal FA desaturases	88
5.3	Comparison of the deduced amino acid sequences of *Pongamia FAD2* genes	89
5.4	Phylogenetic analysis of microsomal FAD2 desaturases	91
5.5	Hydropathy plot of *Pongamia* (A) PpFAD2-1; (B) PpFAD2-2	92
5.6	Southern blot hybridization	94
5.7	Semi quantitative PCR analysis of Pp*FAD2*-1 and Pp*FAD2*-2 genes	96
5.8	Northern hybridization analysis of PpFAD2-1 expression of *Pongamia* tissues	96
5.9	Expression of PpFAD2-1	98
6.1	Collection and maintenance of standard germplasm	105

TH-1255_09610610

Figure no.	Figure description	Page no.
6.2	Flow diagram shows the process of sample in nuclear extraction buffer (NEB) and its analysis with Flow cytometer.	106
6.3	Step-by-step gating regime during standardization	109
6.4	Maintenance of *Pongamia* plant in hydroponics culture	112
6.5	Test for robustness (Internal standard)	114
6.6	Appearance of G2 peak of *P. pinnata* in internal standardization.	117
6.7	Histogram of PI stained nuclei from simultaneously processed *P. pinnata*	117
6.8	Effect of β-mercaptoethanol and PVP on PI binding	119
6.9	Pseudo internal standardization process	121
6.10	Internal standardization with *Z. mays*	123
6.11	Effect of antioxidants on PI fluorescence and light scattering properties	126
6.12	Comparison of the standard nuclei peak location and the *Pongamia* peak location	127
6.13	Effect of antioxidant on *Z. mays*	127
6.14	Nuclear genome content estimation in different tissue of *Pongamia*	130
6.15	Histogram of mean fluorescence intensity in different individuals of *P. pinnata* collected from various parts of India	132
6.16	Histogram of mean fluorescence intensity in different individuals of *P. pinnata*	134
6.17	Correlation between genome size with geographical distribution of *P. pinnata* in Assam	135
6.18	Histograms of relative DNA content of nuclei	137
6.19	Somatic metaphase chromosome in *P. pinnata*	141
6.20	Correlation between nuclear genome size (pg) and length (μM) of *Pongamia*	141

TH-1255_09610610

LIST OF TABLES

Table no.	Table description	Page no.
2.1	Various nuclear isolation buffers used base on the tissue type of the sample	23
3.1	Elite genotype identification characters in *P. pinnata*	27
3.2	CTAB-GITC buffer composition	29
4.1	Reagents and chemicals	53
4.2	Primer sequences used in the experiment	55
4.3	Similarity matrix indicating percentage of PpSAD amino acid and Nucleic acid sequences (open reading frame and untranslated sequences) conserved between species	60
4.4	Quality scores of modeled PpSAD protein using different quality control programs.	76
5.1	Primer sequences used in the experiment	82
5.2	Sequence comparison of deduced amino acids and the conserved histidine boxes	86
5.3	Semi-quantitative PCR analysis	95
6.1	Study site characteristics-samples collected from various parts of Assam and India	104
6.2	Time course of the ratio of the mean (±SD) fluorescence of nuclei (standard 1) of standard leaves simultaneously processed and stained with *P. pinnata* leaves to the mean fluorescence of added PI stained nuclei (standard 2) from independently processed standard leaves	116
6.3	PI fluorescence of nuclei from mixtures of independently chopped *P. pinnata* and standard leaf samples	120
6.4	Mean PI fluorescence of nuclei from standard leaf nuclei and *P. pinnata* leaves simultaneously processed in different volumes of buffer	122
6.5	DNA content of *P. pinnata* estimated by internal and pseudo-internal standard	122
6.6	Effect of antioxidants on flow cytometric estimation of 2C DNA content in *P. pinnata*	125
6.7	Nuclear genome content of different tissues in *Pongamia*	129

TH-1255_09610610

6.8	Genome content of *P. pinnata* saplings collected from various parts of India	131
6.9	Correlation between genome size, temperature and height of *Pongamia* individuals from Assam	133
6.10	2C DNA content of *P. pinnata* estimated by using *Z. mays* as an internal reference standard	136
6.11	Chromosome number and its length of *Pongamia* individuals collected from Assam	140

TH-1255_09610610

Chapter 1

Introduction

The chapter describes brief back ground on the biofuel plant *Pongamia pinnata* and its applications. Based on the background information specific objectives were designed to work on current research programme.

TH-1255_09610610

INTRODUCTION

1.1 Background

During the past several years the cultivated agricultural lands are becoming waste/uncultivated lands (approximately 55 million ha) which is mainly due to the improper land use and population pressure in India (Srivastava & Prasad 2000). On the other hand, with the economic development, increase in population, greater urbanization and land reforms have resulted in the depletion of coal and natural resources (Kesari et al 2008, 2010). Also the continuous use of petroleum intensifies problems of air pollutants, carcinogens etc. Increasing price of petroleum products and raising concerns about oil production are likely to have serious implications for the automobile industry in the near future.

The answer to the above requirement is to search for an alternative to the fast depleting reserves of fossil fuel from renewable natural resources (Martini & Shell 1998; Srivastava & Prasad 2000). This has resulted in strengthening the research and development in the areas of biofuel. In recent years there has been lot of research on alternative fuel sources considering the increased demand for fossil fuel and its limited availability. Since biofuel is renewable, eco-friendly, safe to use with broad applications, as well as biodegradable, it has become a major focus on intensive global research.

Leguminosae is one of the largest families of angiosperms with more than 18,000 species belonging to about 650 genera (Polhill & Raven 1981). The family is about a 12th of all known flowering plants with extremely diverse characters. In terms of economic importance, Leguminosae is one of the most important families in the Dicotyledonae (Harborne 1994). The economic importance of the family is likely to increase as human pressure places greater demand on the marginal land. Many legume species are characteristic of open and disturbed places and are thus well adapted to grow under poor conditions. Numerous legumes also show symbiotic relationship with *Rhizobium* bacteria which helps to convert atmospheric nitrogen into nitrogenous compounds useful to plants.

TH-1255_09610610

Most biodiesel is produced from soya, oilseed rape, maize (corn) and palm oils, although more or less any vegetable oil can be used for biodiesel production. But the problem arises if edible oil is employed in biodiesel industry; the world food market is affected soon. In order to address the shortages of feedstock for the production of biodiesel, new non-edible oil sources need to be exploited (Lin et al 2011; Mustafa 2011). Some of the plant species i.e., *Jatropha curcas*, *Ricinus communis* and *Pongamia pinnata* that produce oil seeds and can grow on marginal lands are falls into this category. Among the other biofuel crops *Pongamia* has got special attention as biofuel crop because of its survival at different climatic conditions (Daniel 2001). *Pongamia pinnata* (L.) Pierre (Synonyms: *Derris indica* Lam., Bennet, *Pongamia glabra* Vent., *Cytisuspinnaus* L.) tree belonging to family Leguminosae/Fabaceae, sub family, Papilonaceae, popularly known as 'Karanj' or 'Karanja' in Hindi is known for its multipurpose benefits and as a potential source of biodiesel (Kesari et al 2010b; Meera et al 2003; Sharma & Singh 2008).

1.2 *Pongamia pinnata* and distribution

P. pinnata belonging to fabaceae/papilionoideae family and sub family papilionaceae is known by various scientific names i.e., *Millettia pinnata*, *Pongamia glabra*, *Derris indica*, and has a common names of Indian beech (English), Karanj (Hindi), karchaw trade poonga (Assam), dalkaramcha (Tamil), and kanuga (Telugu). *P. pinnata* is a commercially important tree species used to produce alternative fuels and can grow on waste land or unproductive lands. The tree is suitable for afforestation especially in watershed areas and in drier part of the country. The tree is adaptable to wide agro-climatic conditions. The natural distribution of this species is along coasts and river banks in India and Burma. *Pongamia* is native to humid and subtropical environments including Asian subcontinent and lowlands in the Philippines, Malaysia, Australia, the Seychelles, the United States and Indonesia. *Pongamia* is common along waterways or seashores, with its roots in fresh or salt water and grows well in full sun or partial shade. However maximum growth rates are observed on well drained soils with assured moisture. According to Duke (1983) *Pongamia* reproduction takes place naturally by seeds and common by root suckers. Since last two decades large number of *Pongamia* trees has been cultivated on road side, railway tracks, canal banks and open farm lands throughout India except temperate regions. The tree is extremely tolerant of saline

TH-1255_09610610

conditions and alkalinity, and occurs naturally in lowland forest on limestone and rocky coral outcrops on the coast, along the edges of mangrove forest and along tidal streams and rivers. *Pongamia* withstand in areas having an annual rainfall ranging from 500 to 2500 mm. In its natural habitat, the maximum temperature is from 27-38 ^{o}C and the minimum can be as low as 1 ^{o}C. *Pongamia* full grown trees can withstand water logging, are resistant to high wind, drought and slight hoarfrost/cold temperatures (Gilman & Watson 1994). *Pongamia* can also grow on most soil types ranging from stony to sandy to clayey, including dry sands, saline soils and grows elevations of 1200 m but in the Himalayan foothills it is not found above 600 m (Daniel 2001). In dry areas, it is one of the few species remaining green during the summer season.

1.3 Botanic description

P. pinnata is a medium sized briefly deciduous, tree (grows 15-25 m height), with straight or crooked trunk (up to 80 cm) or more in diameter and broad crown of spreading branches. Bark is grey-brown, smooth or faintly vertically fissured (Daniel 2001). Leaves are alternate, imparipinnate with long slender leafstalk, pinkish-red in color when tender stage. During the mature stage it will turn to glossy dark green above and dull green with prominent veins beneath. Inflorescence is raceme-like, axillary, 6-27 cm long, bearing pairs of strongly fragrant flowers; calyx campanulate about 4-5 mm long, truncate, delicately pubescent. Flowers are white tinged with pink or violet in color and approximately 2-4 together, short-stalked, pea-shaped, 15-18 mm long. Generally in *Pongamia* pods borne in quantities, smooth, slanted oblong to ellipsoid, flattened but slightly swollen and curved with short, curved point (beaked), brown in color, thick-walled, thick leathery to subwoody, rigid, indehiscent, 1-2 seeded, short stalked. Seeds are compressed ovoid or elliptical, bean like structure with a brittle coat long, flattened, dark brown and oily (Figure 1.1).

TH-1255_09610610

Figure 1.1: *P. pinnata* habit and habitat. A) Tree occurring in North Guwahati, Assam, India; B) Inflorescence; C) individual flowers; D) early immature pods; E) late mature pods; F) mature (dry) seeds.

TH-1255_09610610

1.4 Applications

P. pinnata a versatile oleaginous tree has many attributes and multiple applications and has attracted the world's attention as a sustainable substitute for petroleum products and its ability to grow and reclaim marginal lands (Kesari et al 2010a; Scott et al 2008). All parts of the plant have been used as a crude drug for the treatment of tumors, piles, skin diseases, itches, abscess, painful rheumatic joints, wounds, ulcers, diarrhea etc (Meera et al 2003). More recently, the effectiveness of *Pongamia* as a source of biomedicines has been reported, specifically as antimicrobial and therapeutic agents targeting host pathways and processes (Brijesh et al 2006; Kesari et al 2010b). Besides, it is well known for its application as animal fodder, timber and fish poison (Bottoms 2000). It has also been recognized to possess applications in agriculture and environmental management, with insecticidal and nematicidal activity (Chopade et al 2008). The tree itself is nitrogen fixing and so can grow on waste land or unproductive land and is adaptable to wide agro-climatic conditions (Kesari et al 2010a; Scott et al 2008). The seeds or nuts of the tree are its most useful product. The non-edible oil extracted from seeds used as fuel, lubricant and in soap making. The cake is non-edible; it can be used as an organic fertilizer and is rich in nitrogen and micronutrients. Rural communities in India are familiar with this species because its oil has been used traditionally for lighting lamps in households. Its ability to fix atmospheric nitrogen is another advantage of this species (Kesari et al 2013).

1.5 Current research in *Pongamia*

Extensive characterization of *Pongamia* plant has been studied from north Guwahati, Assam, including candidate plus tree (CPT) identification, germplasm utilization, mass multiplication, genetic diversity study, developmental expression of storage protein profile and physico-chemical characterization of crude oil (Kesari et al 2008, 2009a&b, 2010, 2012). Though the plant has been characterized for various aspects earlier by our group, there exists a lack of information at molecular and cyto-genetic level. Genomic information provides the starting point for understanding the instructions for molecular machines and the systems needed to control and operate them. Through a network of pathways, chemistry and mechanics, this machinery makes the cell and the organisms come alive. Understanding the operation, function and coordination of genomic

TH-1255_09610610

information is a necessary step for the emerging field of systems biology. There is a critical need for detailed study in *Pongamia* using molecular biology tools and technologies. There is a greater need to find out functional and structural profiles of the genes involved in desaturation of fatty acid biosynthesis, DNA content and chromosome number.

Fatty acid composition is a vital for biodiesel quality and storage as it influences viscosity, melting point, cold flow, cetane number and oxidation stability (Knothe 2005). These qualities are prejudiced by fatty acid chemical structure, in particular the carbon number and degree of saturation process. Since fatty acid composition greatly influences biodiesel properties; characterization of fatty acid biosynthetic genes which decide the ratio of saturation to unsaturation fatty acids is an important step.

To understand the molecular mechanism and the regulation of expression of fatty acid biosynthetic genes in *Pongamia,* isolation of full length desaturases genes is a prerequisite and this will be only possible if we construct cDNA library that helps to screen the desired full length functional genes. Constructing a cDNA library lays basic foundation for finding relevant genes and investigating their functions. cDNA library construction from immature seeds of *J. curcas* (Jatinder et al 2010) and *R. communis* (Lu et al 2011) a potential biofuel crops has been reported earlier. Fatty acid (FA) biosynthesis in oilseeds is catalyzed by a set of condensing and desaturation enzymes located in plastids and endoplasmic reticulum (ER). Desaturation of 18:0 FA into 18:1cis Δ^{9}, 18:2cis $\Delta^{9,12}$ is performed by group of enzymes namely steroyl ACP desaturase (SAD) and fatty acid desaturase (FAD2) which have been cloned in plants like flax, olive, soybean and others (Fofana et al 2004; Vrinten et al 2005). Especially in plants the fatty acid biosynthetic gene expression analysis is very important to gain insight into the function of those (desaturase) genes. Functional characterization of fatty acid desaturase genes like *SAD*, *FAD2* has been well studied in oil yielding plants like *J. curcas* (Pingzhi et al 2013), *Glycine max* (Lingyong et al 2007), *Gossypium hirsutum* (Irma et al 2001) and *Ginkgo biloba* (Huanli et al 2013).

Genome size and chromosome numbers are important cytological characters that significantly influence various organismal traits (Jana et al 2007). According to the Pedro (2006) the analysis of chromosome numbers represents an important step in studies of genetic variation, phylogeny, taxonomy and evolution, as well as in studies

TH-1255_09610610

on the structure and diversity of the genomes (e.g., genome sizes, ploidy levels, nuclear architecture). Chromosome numbers among tropical woody species and in some economic flowering plants has been determined (Atchison 1951). Sarbhoy (1977) has revealed cyto-genetic studies (mitotic & meiotic studies) in *P. pinnata*. Furthermore nuclear DNA amount and genome size (C-value) are important biodiversity characters, whose study provides a strong unifying element in biology with practical and predictive uses. The genome content was thoroughly studied in angiosperms (Bennett & Leitch 2005a) compared with other taxonomic groups. No efforts have been made till date to estimate and characterize the genome of *Pongamia* which is important for all future research. Flow cytometry (FCM) using DNA selective fluorochromes is a fast and accurate method for the measurement of nuclear DNA content and thereby estimation of genome size. Accurate estimations of nuclear genome sizes are important for mapping of plant genomes and development of strategies for isolation of important plant genes. Despite *Pongamia* immense usefulness not much work has been done. The full potentiality of *Pongamia* is yet to be realized, from its cytogenetics background; particularly genome size of *P. pinnata* has not been documented.

Therefore, based on the back ground information available, the current study was focused on the cyto-genetic studies i.e., construction of cDNA library from the seeds of elite genotype of *P. pinnata*, characterization of fatty acid biosynthetic desaturase genes (*SAD* and *FAD2*), chromosome number determination and nuclear genome content estimation towards development of high quality *Pongamia* genotypes.

1.6 Specific objectives

By considering the comprehensive literature survey on availability and versatile applicability of *P. pinnata* in India, the present study was focused on the cyto-genetic aspects in elite genotype of *P. pinnata* from North Guwahati, Assam, India. In order to establish its potential applications in molecular and genetic level for developing high quality *Pongamia* genotypes following detailed investigations were performed in the current research:

1. Germplasm collection of earlier characterized (based on vegetative and reproductive characters) elite genotype of *Pongamia* from Sila forest range, North Guwahati, Assam, India.

TH-1255_09610610

2. cDNA library construction from early immature seeds (90-DAF) of elite genotype of *Pongamia* to screen for full length genes that are involved in fatty acid biosynthetic desaturation pathway.

3. Structural and functional characterization of steroyl-ACP-desaturase (*SAD*) gene obtained from cDNA library constructed from seeds of elite genotype of *P. pinnata*.

4. Functional characterization of fatty acid desaturase (*FAD2*) genes.

5. Total nuclear genome content estimation of elite genotype of *Pongamia* by using flow cytometer. It also encompassed to test the homogeneity of nuclear genome content between the various tissues of elite genotype, samples collected from various parts of Assam, India and *in vitro* raised tissue culture plants and the somatic chromosome number determination.

Indian Institute of Technology Guwahati

TH-1255_09610610

Literature Review

The chapter includes the detailed review of literature emphasizing the biofuel crops and its role in agroforestry system, distribution of *Pongamia*, fatty acid biosynthesis in plants and nuclear genome estimation towards development of high quality genotypes.

TH-1255_09610610

LITERATURE REVIEW

The main focus of this dissertation is the implementation of plant biology methods to dissect multipurpose legume tree *Pongamia pinnata*. Employing the techniques of molecular biology related (cDNA library construction, full length gene identification) functional and structural characterization using tools of bioinformatics, nuclear genome estimation and chromosome number determination in *Pongamia* will be examined in detail. The literature review will thus encompass the aforementioned topics, with a brief background on biofuel crop *Pongamia*, a versatile legume.

2.1 Background

Global demand for petroleum is predicted to increase 40 % by 2025. Continued use of petroleum sourced fuels is now widely recognized as unsustainable because of depleting supplies and the contribution of these fuels to the accumulation of carbon dioxide in the environment. Renewable, carbon neutral, transport fuels are necessary for environmental and economic sustainability (Chisti 2007). India meets nearly 75-80 % of its total petroleum requirements through imports. To answer the above problems there is need to search for an alternative to the fast-depleting reserves of fossil fuel from renewable natural resources (Bandana et al 2011, 2013; Martini & Shell 1998; Srivastava & Prasad 2000). This fuel can be termed as Biodiesel, which is biodegradable, non-toxic and has low emissions profiles as compared to petroleum diesel. Application of biodiesel will allow a balance to be sought between agriculture economic development and the environment. Several oil yielding plants are available, among them *Pongamia* has got greater attention towards biofuel research. *Pongamia* oil is non-edible oil extracted from its seeds. The tree belongs to the family Fabaceae, sub-family, Papillionaceae and is known for its varied benefits and as a potential source of biodiesel (Kazakoff et al 2012; Kesari et al 2008, 2009a; Meera et al 2003; Sharma & Singh 2008; Sharon et al 2013) (Figure 2.1).

TH-1255_09610610

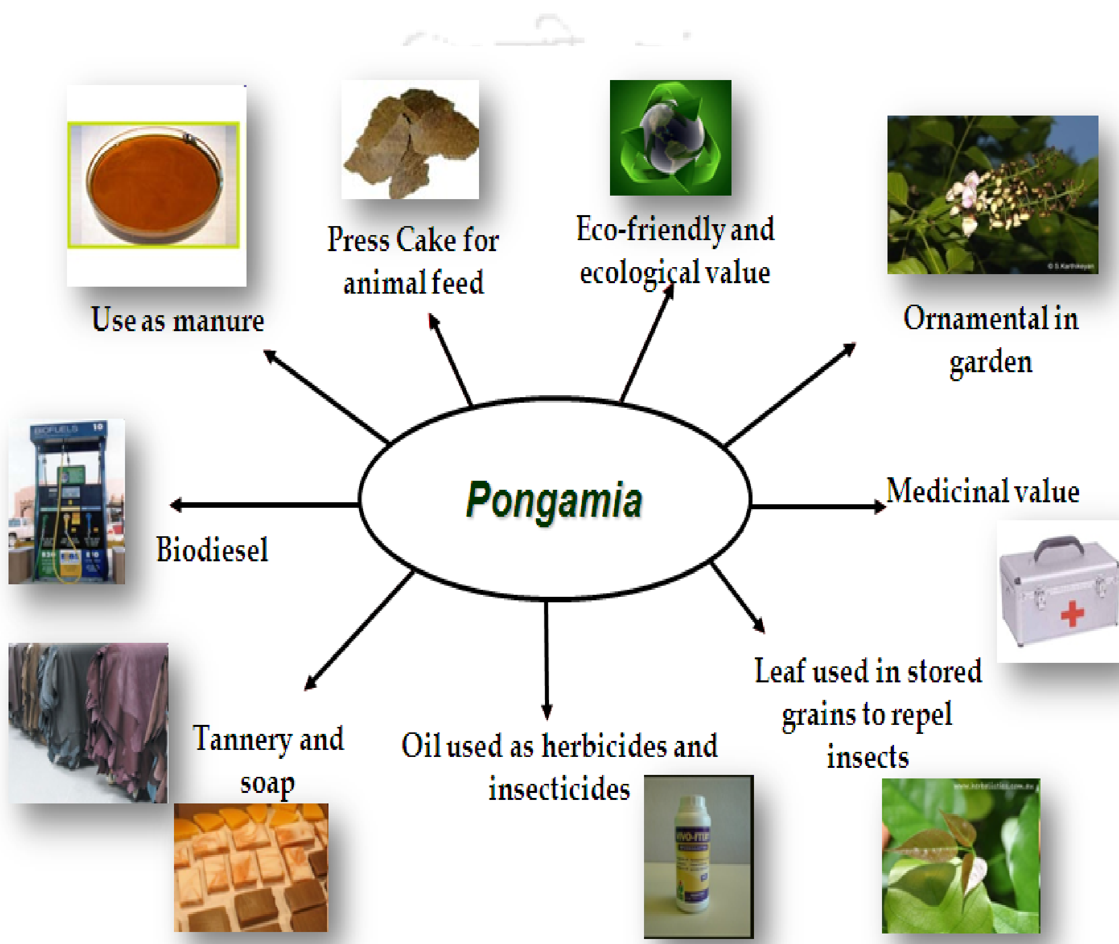


Figure 2.1: Various applications of *P. pinnata*

TH-1255_09610610

The potential of *P. pinnata* oil as a source of fuel for the biodiesel industry is well recognized (Azam et al 2005; Karmee & Chadha 2005). Moreover, the use of oils from plants such as *P. pinnata* has the potential to provide an environmentally acceptable fuel with reduction in engine emissions (Raheman & Phadatare 2004). Importantly, the successful adoption of biodiesel is reliant on the supply of feedstock from non-food crops with the capacity to grow on marginal land and not destined to be used for the cultivation of food crops (Hill et al 2006). *P. pinnata* is a strong candidate to contribute significant amounts of fuel feedstock, meeting both of these criteria. The seeds of *P. pinnata* contain 30 to 40 % oil (Natanam et al 1989; Nagaraj & Mukta 2004) which can be converted to biodiesel (fatty acid methyl esters; FAMEs) by trans-esterification with methanol in the presence of KOH. The predominant fatty acid is oleic acid (C18:1; 40 % to 55 %) with palmitic acid (C16:0; 5 % to 15 %), stearic acid (C18:0; 5 % to 10 %) and linoleic acid (C18:2; 15 % to 20 %), and to a much lesser extent arachidic acid (C20:0), eicosanoic acid (C20:1), behenic acid (C22:0) and lignoceric acid (C24:0). Number of reports are available on oil aspects from biodiesel crops (Kesari et al 2008, 2009; Kpikpi 2002; Akintayo 2004; Kaushik et al 2007). Reports were also available on fuel related properties and fatty acid profile of *Pongamia* seed oil (Kesari et al 2008, 2009; Ahmad et al 2003; Sarma et al 2005; Sharmin et al 2006). The composition of the seed oil and the properties of the FAMEs meet American and European standards (Azam et al 2005; Karmee & Chadha 2005).

P. pinnata has been documented to include variable forms with a wide range of pod as well as seed size and shape (WOI 1969; Kesari et al 2008). Characterization and selection of candidate plus trees (CPTs) is essential for the improvement of this species in addition to experiments on controlled crossing among selected genotypes (Kesari et al 2008, 2009a,b). Tree breeding consists of selection of superior germplasm followed by large-scale propagation of true-to-type plus trees in the seed orchard for the production of improved seeds, which can be used for high quality plantations (Kesari et al 2010). A plus tree is an individual tree of a species possessing superior morphological and reproductive characters than other individuals of the same species. Despite the importance of this versatile plant and the availability of appropriate molecular genetics tools, the full potential of *Pongamia* is far from being realized (Kesari et al 2010).

TH-1255_09610610

2.2 cDNA library construction

cDNA library is a random collection of cDNA fragments, classically representing the entire mRNA of a target tissue, that have been inserted into a cloning vector. It could be easier to find out if a particular gene sequence is required from a cDNA library rather than a gene library, which is larger. However, the cDNA library screened would have to be from a selective tissue where the desired genes were being expressed (Brown 2001). A flowchart illustrating processes involved in the preparing and screening of a cDNA library is given in figure 2.2.

cDNA libraries are simpler to construct, because cDNA fragments, like their parental mRNAs, are already fairly short, so an entire cDNA can be spliced into a single vector. Other advantages over genomic libraries include absence of introns, so there is no danger of pieces of the target gene being chopped onto separate clones (Onken 1997; Old & Primrose 1994). A wide range of technical and theoretical advances over the last couple of decades has enabled cDNA libraries to be constructed from small quantities of mRNA. A variety of reliable methods have been developed to identify cDNA clones corresponding to extremely rare species of mRNA (Sambrook & Russell 2001). Methods exist to generate and amplify cDNA libraries, and amplify genes and single RNA transcripts from a single cell without cloning (Jena et al 1996). When isolating RNA or mRNA from a small number, or even individual cells, certain techniques and precautions can be taken to establish cDNA libraries of useful size. These include extracting RNA immediately from freshly harvested material and using a scaled down version of standard RNA isolation protocols, using total RNA rather than poly(A)+ mRNA as template for first strand cDNA synthesis (Lambert & Williamson 1993), using a single buffer for synthesis of single strand and double strand cDNA, and addition of linkers and amplifying either single strand or double strand cDNA by PCR (McCarrey & Williams 1994).

TH-1255_09610610

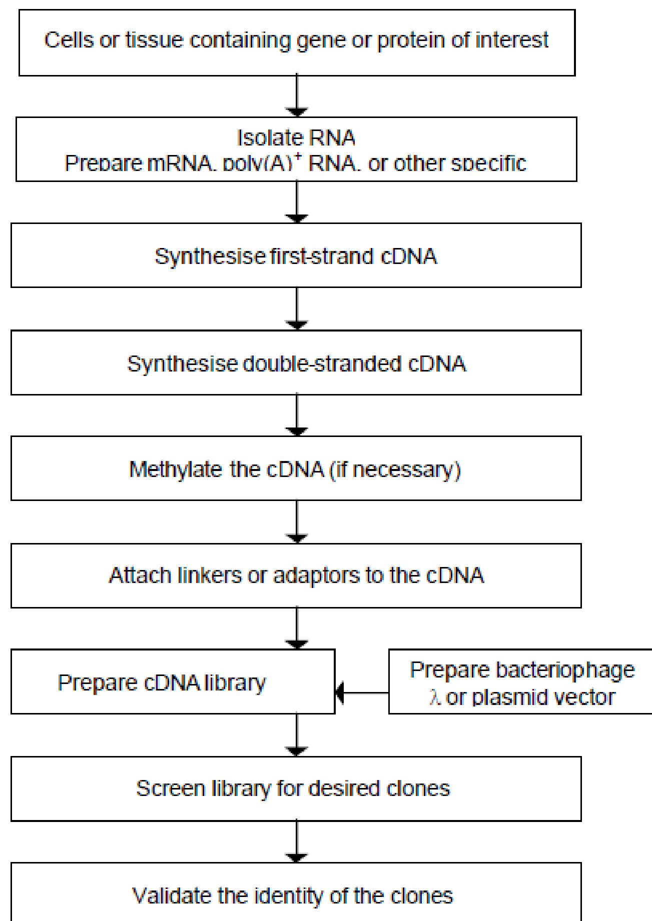


Figure 2.2: Preparation and screening of a cDNA library (Sambrook & Russell 2001).

Indian Institute of Technology Guwahati

TH-1255_09610610

Many reports are available on cDNA library that are constructed in various plants, including biofuel crops. Jatinder et al (2010) were succeeded in construction of cDNA library from immature seeds of *Jatropha curcas*. The library contained 6 x 10^6 clones and the titer of amplified library was 1 x 10^{11} pfu/ml; Kohchi et al (1995) constructed cDNA library from *A. thaliana* and Kim (1999) also constructed cDNA library from rice in which the primary library titer was 1.8 x 10^7 pfu /ml and the titer of amplified library was 2.5 x 10^{10} pfu/ml. The average size of inserts was 1-0.75 kb and the recombinant efficiency was 94.4 %. Recently Mochida et al (2013) constructed a mixed full-length cDNA library from 21 various tissues of *Brachypodium distachyon* Bd21, and were able to isolate more than 75,000 high quality expressed seqeucnce tags (ESTs). Reports are also available on cDNA library constructed using seed endosperms in biofuel plants such as *Jatropha* and *Ricinus*. Jatinder (2010) has constructed cDNA library with early stage seeds and achieved more than 1 x 10^6 clones which harbor full length genes along with ESTs. Lu et al (2011) analyzed the ESTs of developing caster endosperm using a full-length cDNA library. Leif Schauser et al (1995) used *Lotus japonicus* and Soybean Root nodule to construct cDNA libraries and some functional genes identified from the library by ESTs. The titer of unamplified library was 1 x 10^7 and the size of insert was about 0.75-3 kb. Seven out of 11 ESTs obtained from the library have a significant homology with certain functional genes.

Since full-length cDNAs carry complete protein coding sequences and UTRs, they are indispensable for the identification of genes and for the determination of primary protein structure. In particular, full-length cDNAs represent a valuable resource for functional gene studies (Draper et al 2001, Mochida et al 2013). These full-length clones not only serve as a resource for functional analysis but also give valuable information such as transcriptional start sites (Suzuki et al 2003, Jatinder et al 2010).

2.3 Fatty acid desaturases

Plants synthesize a variety of fatty acids which shows diverse role. These are the products of fatty acid biosynthetic metabolism, which takes sucrose derived from photosynthesis and converts it into three major storage components, namely protein, starch and fatty acids. Fatty acids are synthesized by a well defined pathway involving two carbon elongation and bond desaturation (Figure 2.3). The poly unsaturated fatty acid (linoleic and α-linolenic) contributes major portions of the fatty acid profile in oil yielding plants.

TH-1255_09610610

The desaturation of fatty acids determined by the ratio of saturated fatty acids to unsaturated fatty acid is an important aspect in oil biochemistry (Knutzon et al 1992; Mikkilineni & Rocheford 2003). During lipid biosynthesis, the formation of the first double bound in stearic acid (18:0) to produce the monounsaturated oleic acid ($18{:}1\Delta^{9}$) is catalyzed by the soluble plastidial stearoyl-ACP Δ^{9} desaturase. Subsequently, oleic acid is incorporated into the glycerolipids either in plastids or in the endoplasmic reticulum (ER) membranes before being further desaturated by the membrane-bound ω-6 (Δ^{12}) desaturases (Figure 2.4).

The first double bond is introduced by the Δ9 desaturase into stearic acid. Once the mono unsaturated fatty acid is converted then the other two desaturases (Δ12 and Δ6) introduce a double bond into fatty acids which has a double bond at the Δ9 position (Murata & Wada 1995; Higashi & Murata 1993). According to Higashi & Murata (1993) the ω3 desaturase (fatty acid desaturase) introduces a double bond into fatty acids that have a double bond at the Δ12 position. Importantly all the desaturase enzymes are characterized by the presence of eight histidine amino acids in three histidine boxes, which are localized at robustly conserved positions in the amino acid sequence of each protein (Murata & Wada 1995). It has been suggested that these histidine boxes could be involved in the formation of the active site of each desaturase, which is extensively demonstrated in other di-iron enzymes by Shanklin et al 1994; Fox et al 1993, 1994.

The comparative study of the acyl-lipid desaturases of cyanobacteria with the corresponding desaturases of higher plants reveals the occurrence of similarly positioned histidine boxes in the enzymes from all types of organisms (animal, yeast and fungal cells) (Murata & Wada 1995; Shanklin et al 1994). Furthermore, the details of the histidine boxes are specific to the type of desaturase, when the desaturases are classified with respect to the position in the acyl-chain at which each introduces an unsaturated bond. Site-directed mutagenesis of the stearoyl-CoA desaturase of the rat (Fox et al 1993; Shanklin et al 1994) and of the Δ12 acyl-lipid desaturase of *Synechocystis* sp. (Avelange et al 1995) revealed that substitution of amino acid in one of the conserved histidine residues leads to the loss of enzymatic activity. The loss of enzyme activity could be due to the inability of each mutant enzyme to bind ferric iron at the necessary sites (Avelange et al 1995; Schneider et al 1992).

TH-1255_09610610

According to Shanklin et al (1994) and Fox et al (1994) each acyl-ACP desaturase binds two atoms of iron to form a reactive complex with oxygen (Fe-O-Fe). This complex converts the C-C single bond to a C=C double bond in the carbon chain of the fatty acid substrate.

For the first time Lindqvist et al (1996) studied the crystallographic analysis of stearoyl-ACP desaturase from caster seeds and the results revealed that the desaturase forms a di-iron active center. The two iron atoms are bound in a highly symmetric environment, with one of them interacting with the side chains of E196 and H232 and the other interacting with the side chains of E105 and H146. The deep channel were also observed that extends from the surface into the interior of the enzyme and the channel might be the site to which the fatty acyl chain binds. These findings provide clues to the mechanism of desaturation, as well as to the mechanisms whereby specific sites of desaturation are determined.

The expression of genes that are involved in desaturation is most important since it provides the molecular basis for the acclimation of organisms to changing environmental temperatures, in particular in the case of plants and other eukaryotes (; Murata & Wada 1995; Nishida & Murata 1996). Cloning of genes for desaturases from a variety of organisms over the past decade has allowed an examination of the expression of these genes and the roles of their protein products in acclimation (Dmitry & Murata, 1998). The fatty acid desaturase (*FAD2*) gene of *Arabidopsis thaliana*, which encodes the microsomal Δ12 desaturase, does not respond to a shift in temperature, at least, at the level of its transcript (Okuley 1994). By contrast to *A. thaliana*, soybean has two *FAD2* genes (Heppard et al 1996). One of the genes is expressed in developing seeds exclusively, whereas the other gene is expressed constitutively in both vegetative and reproductive tissues. The former gene might be responsible for the synthesis of polyunsaturated fatty acids in storage lipids (Heppard et al 1996). The similar kind of expression was also reported in olive by Luisa et al (2005).

TH-1255_09610610

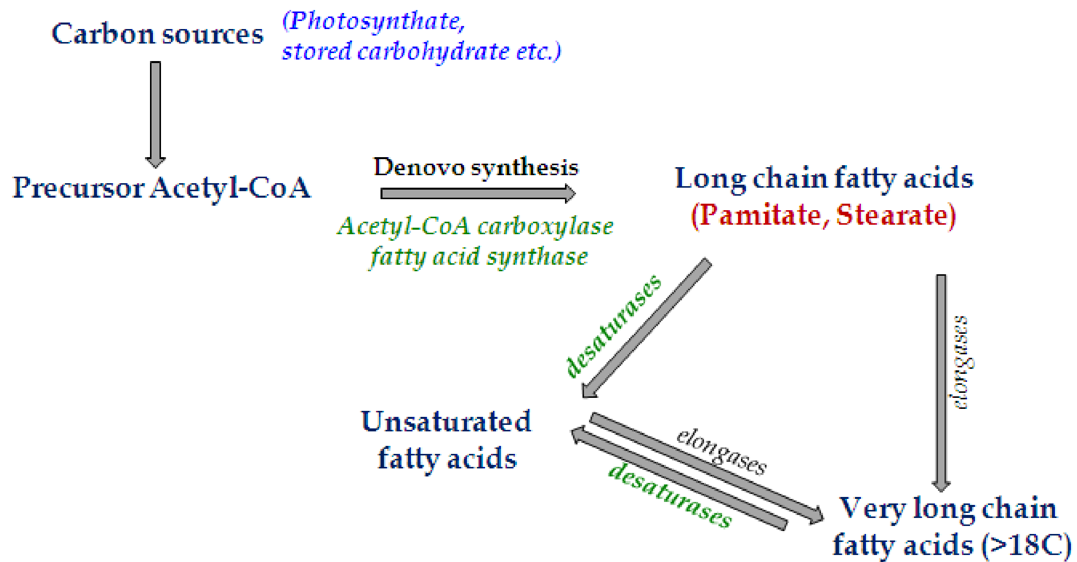


Figure 2.3: Fatty acid biosynthesis in plants.

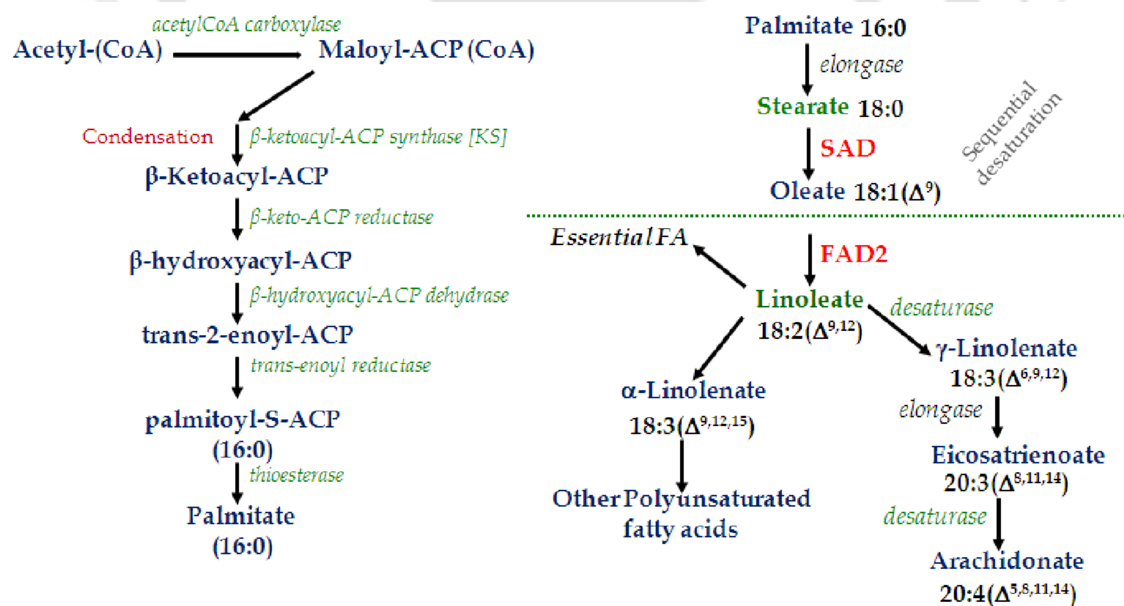


Figure 2.4: Sequential desaturation of fatty acids.

TH-1255_09610610

In *A. thaliana* two genes have been isolated, namely *FAD2* and *FAD6*, encoding the microsomal and the plastidial ω-6 desaturases, respectively (Falcone et al 1994; Okuley et al 1994). cDNAs encoding for microsomal ω6 desaturases have been isolated from several plant species including soybean (Heppard et al 1996), parsley (Kirsch & Hahlbrock 1997), peanut (Jung et al 2000), sesame (Jin et al 2001), sunflower (Martinez-Rivas et al 2001), cotton (Pirtle et al 2001 & Dyer et al 2002). Similarly, cDNA that encodes for SAD have been reported in various plants (*Carthamus tinctorius, Thunbergia alata, Pricinus communis, Solanum tubersoum*). Structural and functional characterizations of the aforementioned genes were determined by Davydove et al (2005) and Lindquvist et al (1996).

2.4 Nuclear genome content and chromosome number determination

Genome size is a fundamental parameter in many genetic and molecular biological studies (Baird et al 1994). DNA C-value remains a key character in biology and biodiversity. Knowledge of the haploid nuclear DNA content (C-value) is important for basic and applied studies involving genome organization, species relationships, gene expression analysis, and germplasm improvement (Bennett 1995). For example, genome size estimates are important when constructing and screening genomic or cDNA libraries (Clarke & Carbon 1976). It is also necessary for developing linkage maps for genetic analysis and breeding purposes, and in efforts to estimate the recombinational length of nuclear genomes and correlates this genetic distance with physical distance (Meagher et al 1988). Finally, this information can be useful in evaluating reproductive and somatic compatibility, an important parameter in scion breeding and rootstock selection programs, especially for those using inter-specific crosses (Baired et al 1994).

Furthermore genome size (nuclear DNA content) data also provide information useful in various fields of plant biology, including systematics, evolution and conservation (Bennett & Leitch 2005b). Genome size has many important practical implications at many different levels. *Arabidopsis thaliana* was the first plant chosen for genome sequencing, partly because it had one of the smallest C-values known for an angiosperm (NSF 1990; Anderson 1991). A grass in the genus *Brachypodium* (e.g. diploid *B. distachyon*-1C=0.25±0.3 pg) was proposed as a first monocot for genome sequencing on similar grounds (Bablak et al 1995; Catalan et al 1995), but rice (*Oryza sativa*,

TH-1255_09610610

1C=~0.5pg) was chosen because it has the smallest C-value among the world's major cereal crops (Sasaki 2002). Accurate estimations of nuclear genome sizes are important for mapping of plant genomes and development of strategies for isolation of important plant genes. Bennett & Smith (1976), Bennett et al (1982) compiled estimates for the nuclear DNA amounts of about 1000 angiosperm species which were either scattered in scientific journals or unpublished. For some species several different values were reported. Despite its usefulness in understanding plant evolution and diversification, genome size of *P. pinnata* is not well documented, estimates for only a few species have been published (Kew Royal botanical garden). Ohri D & Kumar (1986) determined by flow cytometry 2C=3.60 pg in *Pongamia glabra* and 2C=4.0 in *Millettia ovalifolia*, who observed 2C-value is ranging from 1.68 to 4.60 pg in 10 different legume tree plants. In addition, flow-cytometric nuclear DNA amounts for 16 perennial shrubs which belongs to leguminosae family has been reported in Kew Royal botanical garden. In plants holoploid genome sizes (1C-values) vary strikingly, ranging from about 0.065 to 127.4 pg. Genome size variation has significant consequences at cellular, tissue and organismal levels and also influences phenological and ecological behavior. Olszewska & Osiecka (1984) used cytophotometry to study genome size in *Laburnum alpinum* (2C=3.80 pg).

There are many approaches towards determination of 2C DNA content (a nucleus in G0/G1 phase of the cell cycle, with two copies of un-replicated genome). One of them is measurement of individual nuclei which offers mush higher precision, but is technically more demanding (Dolezel and Bartos 2005). Early measurements of individual nuclei relied on the absorption of UV light by the DNA molecule. Later the nuclei were stained by Feulgen method, which is considered specific for DNA, and the absorption of visible monochromatic light was quantified (Swift 1950). Due to inherent limitations and deficiencies (e.g., expense, time and sample size) of above methods, use of flow cytometry (FCM) for measurement of individual nuclei became the method of choice (Galbraith 1989; Michaelson et al 1991a). Because the particles are analyzed individually and at high speed, large populations can be measured in a short time and the presence of subpopulations may be detected (Shapiro 2003).

TH-1255_09610610

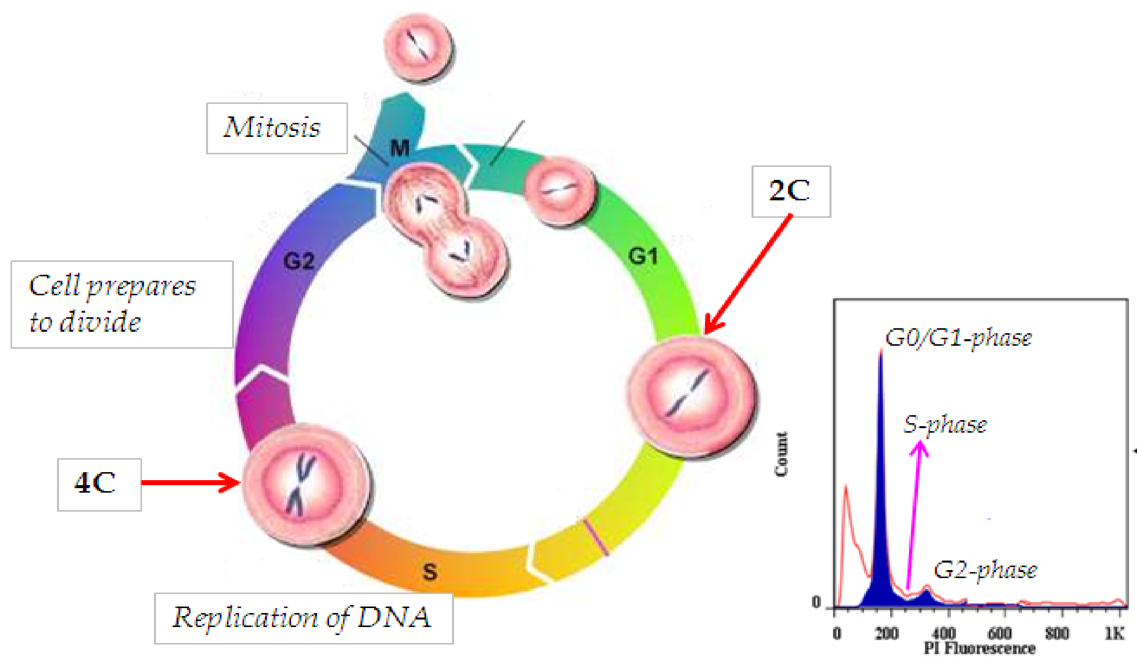


Figure 2.5: A cell biology and flow cytometry interface for nuclear genome estimation.

Indian Institute of Technology Guwahati

TH-1255_09610610

The unreplicated haploid chromosome complement (n) is termed as C-value (Sweift 1950). Hence a nucleus in G1 phase of the cell cycle, which two copies of unreplicated genome has a 2c DNA amount. The cell cycle is the series of events that take place in a cell leading to its division and duplication (replication). Interphase proceeds in three stages, G_1, S and G_2 preceded by the previous cycle of mitosis and cytokinesis. The most significant event is the replication of genetic material (DNA) in S phase. The interpretation of cell cycle events and the flow cytometry has been shown in figure 2.5. The G0-G1 peak is characterised by the 2C DNA amounts and the G2 phase has been characterised by 4C DNA amounts. So the fluorescence area ratio of G2 and G0/G1 peak is always 2.0. So the researchers could get the estimation of nuclear DNA amounts from the G0-G1 peak of the above histogram. C-value has been characterized by the DNA amounts in the unreplicated nuclear genome of an organism. In Go/G1 stage, two copy of the chromosome are present. So in G0/G1 stage the estimation is said to be 2C-value.

FCM provides an accurate and simple means to determine nuclear DNA contents (C-value) within plant homogenates (Galbraith 2009). In another study a cotton column was designed and polyvinylpyrrolidone-40 was added to the buffer to remove phenolic impurities and cytoplasmic compounds from plant nuclei, making the suspension suitable for FCM. This simple and highly efficient protocol enabled isolation of intact nuclei from plant tissues containing high levels of polysaccharides, calcium oxalates, and other metabolites (Lee & Lin 2005).

The results of the study of flow cytometric estimation of nuclear DNA amount in diploid bananas showed that the size of nuclear genome of Musa is smaller than previously estimated (Dolezel et al 1994). A comparative flow cytometric estimation of nuclear DNA content in oil palm (*Elaeis guineensis* Jacq) tissue cultures and seed-derived plants was studied by Rival et al (1997). Embryogenic calli and plants showed the same ploidy level, but the measured 2C DNA values differed significantly (Rival et al 1997). Microfluorometric analysis of the nuclear DNA contents of the somatic tissues of *Arabidopsis thaliana* has revealed extensive endoreduplication, resulting in tissues that comprise mixtures of polyploid cells (Galbraith et al 1991). FCM was used to determine DNA content levels of Brome grass germplasm (Tuna et al 2001). Usually young, rapidly

TH-1255_09610610

growing leaves are used for analysis. The usefulness of seeds or their specific tissues for FCM genome size estimation was investigated by (Sliwinska et al 2005).

2.4.1 Buffer compositions

The buffer should protect the nucleus from degradation and provide an appropriate environment for specific and stoichiometric staining of the nucleus, including the minimization of negative effects of some cytosolic compounds on DNA staining (Loureiro et al 2007). Although various popular nuclear isolation buffer formulas are available, quantitative data on performance of the most buffers has showed that none of them worked well with all species that represented different types of leaf tissues and different nuclear DNA content estimates, and researchers are encouraged to optimize protocols for their specific applications (Loureiro et al 2006, 2007) (Table 2.1). Buffer choice, staining period and PI concentration has been shown to have statistically significant effect on nuclear DNA content estimates, altering the 1C-values (Dolezel et al 2007; Greilhuber et al 2007; Loureiro et al 2006). Previous researchers speculated that inhibitors that decrease dye fluorescence of nuclei were common in plants (Price et al 2000) and showed that cytosolic compounds could bias nuclear DNA content estimates by up to 20 % (Noirot et al 2000). The effect of cytosolic compounds present in the leaves could not be avoided entirely despite the fact that antioxidants (β-mercaptoethanol and PVP) were used in nuclei-isolation buffer (Noirot et al 2000, 2003).

Various lysis buffers i.e., Galbraith's buffer (Galbraith et al 1983), LB01 (Dolezel et al 1989), Otto's buffer (Dolezel & Gohde 1995), Tris.$MgCl_2$ (Pfosser et al 1995) and Marie & Brown (1993) were used to prepare the sample from leaf tissues of seven plant species (*Sedum burrito, Oxalis pes-caprae, Lycopersicon esculentum, Celtis australis, Pisum sativum, Festuca rothmaleri* and *Vicia faba*). The species were selected to cover a wide range of genome sizes (1.30-26.90 pg per 2C DNA) and a variety of leaf tissue types (Loureiro et al 2006).

TH-1255_09610610

Table 2.1: Nuclear isolation buffers used for flow cytometric analysis.

Buffer	Composition	References
Galbraith's buffer	45mM $MgCl_2$; 30mM siduyn cutrate; 20mM MOPS; 0.1% (w/v) TritonX-100; pH 7.0.	Galbraith et al (1983)
LB01	15mM TRIS; 2mM Na2EDTA; 0.5mM spermine.4HCl; 80mM KCl; 20mM NaCl; 15mM β-mercaptoethanol; 0.1% (v/v) Triton X-100; pH 7.5.	Dolezel et al (1989)
Arumuganathan and Earle	9.53 mM $MgSO4.7H_2O$; 47.67 mM KCl; 4.77 mM HEPES; 6.48 mM DTT; 0.25% (w/v) Triton X-100; pH 8.0.	Arumuganathan & Earle (1991)
Marie's nuclear isolation buffer	50mM glucose; 15mM KCl; 15mM NaCl; 5mM Na_2EDTA; 50mM sodium citrate; 0.5% (v/v) Tween 20; 50mM HEPES; 0.5% (v/v) β-mecaptoethanol; pH 7.2.	Marie & Brown (1993)
Otto buffers	Otto-I buffer: 100mM citric acid; 0.5 % (v/v) Tween 20 (pH approx. 2.3).	Otto (1990)
	Otto-II buffer: 400mM $Na_2HPO4.12H_2O$ (pH approx. 8.9).	Dolezel and Gohde (1995)
Tri-$MgCl_2$	200mM TRIS; 4mM $MgCl_2.6H_2O$; 0.5% (v/v) Triton X-100; pH 7.5.	Pfosser et al (1995)

Nuclear DNA content (C-value) varies approximately 1000-fold across the angiosperms and this variation has been reported to have an effect on the quality of AFLP fingerprints (Fay et al 2005). Standard DAPI protocol was employed to evaluate the performance of 60 air-dried species, spanning more than 100 fold-ranges of nuclear DNA amounts (Suda & Travnicek 2005). FCM was used to measure fluorescence of isolated *Gossypium* nuclei stained with propidium iodide (Hendrix & Stewart 2005). The current procedures for estimation of absolute DNA amounts in plants using FCM, with special emphasis on preparation of nuclei suspensions, stoichiometric DNA staining and the use of DNA reference standards were employed to classify the nuclei according to their relative fluorescence intensity or DNA content (Dolezel & Bartos 2005). In Flow cytometric analysis SYBRGreen-I was used for genome size estimation in coffee and tested as an alternative cytometric protocol using the SYBRGreen-I as a fluorochrome (Clarindo and Carvalho 2009). Monocot genomic diversity includes striking variation at many levels. Various genomic characters (e.g., range of chromosome packaging and organization, genome size) between monocots and the remaining angiosperms are compared to discern just how distinctive monocot genomes are (Leitch et al 2010). Estimation of plant nuclear DNA content is usually performed using young leaves; however, seeds can be used as an alternative material because they are easier to transport

TH-1255_09610610

and can be stored for prolonged periods (Sliwinska 2006). Sliwinska (2009) showed that cell cycle activity in seeds corresponds with their physiological state and therefore it can be used by seed growers to monitor seed development, maturation and germination, as well as to control pre-sowing priming treatments. Coordinated changes in cell cycle/endo-reduplication activity in developing seeds with respect to storage protein and carbohydrate/oil synthesis was used to determine a more suitable marker of seed developmental stages for assessing optimal harvest times (Rewers & Sliwinska 2012).

2.4.2 Chromosome study

Chromosomal survey which primarily involve in the determination of chromosome number and meiotic behavior is an important step in understanding the cytogenetical constitution of species, relationship among taxa and to provide a base for future improvement programmes (Gill & Singhal 1998b). The chromosomal studies in flowering plants started at the end of 19th century by a German cytologist Strasburger in 1882 who for the first time counted chromosomes in orchids (Fedorov 1969). The cytological investigations of monocot zinger family have been well documented (Mukherjee 1970). According to Bandel (1974) the cytological investigations are not thoroughly done in leguminosae and only 18 % of the information on chromosome is available. The chromosome number of the genus *Pongamia*, as far as the records reveal, indicates the presence of both X=11 (Patel & Narayana 1937; Raghavan & Arora 1958) and X=10 (Atchison 1951) chromosomes in the basic set. According to Sarbhoy (1977) investigations in *P. pinnata* confirms 11 and 22 as haploid and diploid number of chromosomes. The two basic numbers X=10 and X=11 in the genus *Pongamia* are to be considered. It seems that there exists two races of *P. pinnata* with basic numbers one with X=10 and other with X=11 in nature. It is clear that basic numbers 10 and 11 are derived one from the other. The chromosome number and its length were determined in *Allium cepa* by Bandaru et al (2011). In his experiments he has also drawn a comparison of chromosome length with nuclear genome size. Till date there are no reports available on characterization of fatty acid desaturases, nuclear genome estimation and chromosome number determination in biofuel tree *Pongamia*. ***This dissertation is an attempt towards cDNA library construction to fish out useful genes, chromosome number and its length measurement and correlation with nuclear genome content of Pongamia.***

TH-1255_09610610

Chapter 3

cDNA Library Construction from Seeds of *Pongamia pinnata* (Elite Genotype) and Screening for Full Length Genes

This chapter describes the standardization of RNA extraction protocol, cDNA library construction and screening of fatty acid biosynthetic genes using gene probes.

TH-1255_09610610

cDNA LIBRARY CONSTRUCTION FROM SEEDS OF *PONGAMIA PINNATA* (ELITE GENOTYPE) AND SCREENING FOR FULL LENGTH GENES

3.1 INTRODUCTION

In recent years, we have witnessed an outbreak of DNA sequence data from various model systems from prokaryotes to eukaryotes (including plants) genome projects. However, the quick increased in nucleotide sequence data have not been matched by a correspondingly large increase in the identification of gene expression patterns or understanding of gene functions. As complete genome sequences become available, it will be possible to study the regulation and expression of genes on a full genome basis, rather than gene by gene. Complementary DNA (cDNA) is the reverse transcriptase product of messenger RNA (mRNA) and represents the coding sequence of all transcribed genes at the time of mRNA isolation. In any attempt to study the gene function of an organism, the incorporation of a cDNA-based approach is unavoidable. Indeed, one of the fundamental tools of molecular biology is the enzymatic conversion of poly $(A)^+$ mRNA to double stranded (ds) cDNA, and the insertion of this DNA into both prokaryotic and eukaryotic vectors (Kimmel & Berger 1987).

cDNA libraries are physical archives of the genes that are transcribed in living organisms (Forrell 2005). In spite of the high rate of information generation which is the trademark of high-throughput sequencing, obtaining expressible sequence information by means of cDNA libraries has its own virtues. The most important advantage is that, when a transcript of interest is discovered in a cDNA library, the cloned sequence is physically available and analysis and/or manipulation thereof can directly commence. Another advantage is that nucleic acid probes can be used to quickly identify cDNA clones which contain inserts of interest if genetic sequence information is already available for a similar transcript in a related organism. Furthermore, full-length cDNA sequences can be obtained from cDNA libraries and if a library was constructed using expression vectors, the proteins for which they encode can be directly expressed and isolated. Unlike genomic DNA that has introns in it,

TH-1255_09610610

cDNA contains an open reading frame ready for expression. Therefore, a cDNA library can be used not only to screen the target genes required, but also to express them (Sambrook & Russell 2001). A disadvantage of the method is that low abundance transcripts are often lost during the many steps of the library construction protocol. Library construction is also a more labour intensive approach than high-throughput cDNA end sequencing.

Efficient isolation of high quality and quantity of total RNA from plant tissues is prerequisite for molecular biology experiments including construction of a good-quality cDNA library. The isolation of RNA can be quite difficult particularly in plant tissues especially storage organs such as seeds that are rich in polyphenolics, polysaccharides, proteins and other unidentified compounds (Bugos et al 1992; Sharma et al 2003;). These compounds tend to co-precipitate with RNA and contaminate RNA extracts because they display physical and chemical properties similar to those of nucleic acids (Azevedo et al 2003; Salzman et al 1999; Sharma et al 2003; Singh et al 2003). Moreover phenolic compounds readily oxidized to form covalently linked quinones (Loomis 1974; Salzman et al 1999) and avidly bind nucleic acids. In addition to this RNA degradation by ribonucleases (RNases) drastically reduces the recovery of intact RNA (Birtic & Kranner 2006; Murillo et al 1995). Existing RNA isolation methods, commercial kits (including Trizol, Gibco-BRL Life Technologies; RNeasy plant kit, QIAGEN) specifically alleviate one or two of these challenges, but are usually designed for a single species rather than a broad biodiversity (Ding et al 2008). Even plant species belonging to the same genus or related genera can exhibit enormous variability in the complexity of pathways of dispensable functions (Sharma et al 2002). Thus the biochemical compositions in plant tissues of different species are expected to vary considerably. This renders RNA unusable for molecular biology methods such as semi-quantitative RT-PCR, cDNA library construction and investigating gene expression profiles.

Genetic intervention in *P. pinnata* requires an understanding of the biosynthetic pathways, cloning of the genes that code for the enzymes involved in metabolic pathways (fatty acid biosynthetic pathway) in seed development stages. Molecular studies in *Pongamia* are very limited and many researchers are focusing on its biodiesel aspects. In this study the main objective was to extract good quality and quantity of

TH-1255_09610610

total RNA and to construct a full-length cDNAs from the multipurpose legume crop, *P. pinnata*. This chapter describes the main approach to making the full length library and its potential uses in gene functional analysis.

3.2 MATERIAL AND METHODS

3.2.1 Plant tissue collection

Seeds from early immature stage [90-Days after flower (DAF)] of earlier characterized elite genotype of Candidate Plus Tree-North Guwahati Pongamia Pinnata [(CPT-NGPP46) (based on vegetative and reproductive characters, Table 3.1)] of *Pongamia* samples were collected from sila forest range, North Guwahati, Assam, India. The collected seeds were first washed with distilled water, followed by washing with 70 % ethyl alcohol twice and subsequently stored at -80 ^{o}C until total RNA extraction for cDNA library construction. The overall process of cDNA library construction starting from sample collection is depicted in Figure 3.1.

Table 3.1: Elite genotype identification characters in *P. pinnata*

Vegetative characters	Reproductive characters
• Girth of the main stem at breast level (m) • Plant height (m) • Number of leaves /g (wt) • Canopy size (m)	• Number of buds/inflorescence • Number of flowers/inflorescence • Number of seeds/inflorescence • Pod traits viz; Length (cm), Breadth (cm) and 100 Pod wt(g) • Seed traits viz; Length (cm), Breadth (cm) and 100 seed wt (g) Pod to seed ratio

3.2.2 RNA extraction using CTAB-GITC

Standard RNA isolation procedures of Chang et al (1993) and using Trizol reagent, failed to work when applied to *P. pinnata* tissues especially to seeds which are known to contain polysaccharides, proteins, lipids and other secondary products. The extraction method used in the current study is quite simple, inexpensive, rapid and reproducible cetyl trimethyl ammonium bromide (CTAB) extraction method of Chang et al (1993) having slight modification. The described protocol allowed RNA extraction from tissues of elite genotype of *P. pinnata* and were reverse transcribed to construct cDNA library in which other procedures were unsuccessful.

TH-1255_09610610

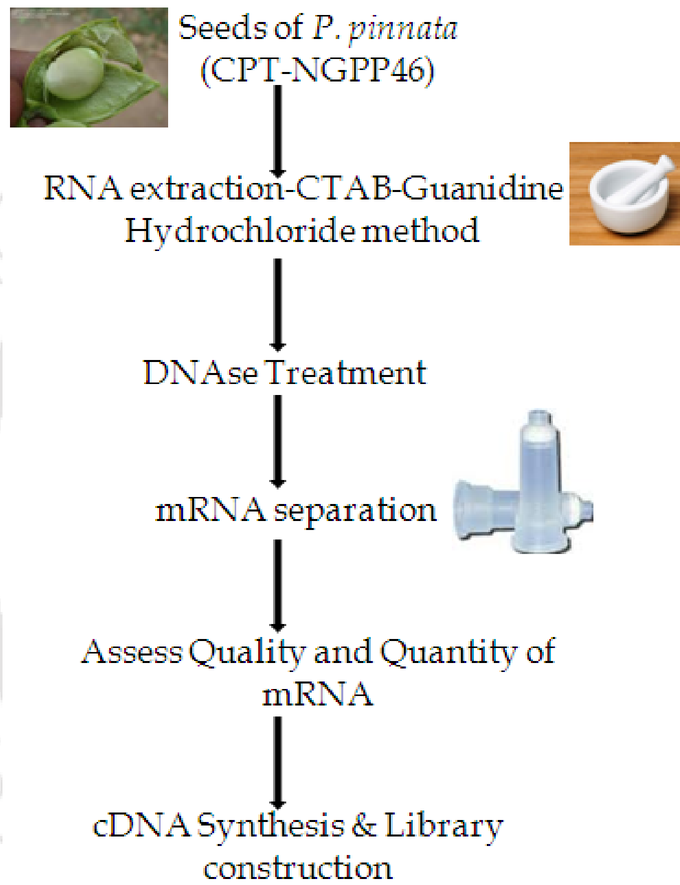


Figure 3.1: Flow diagram of the cDNA library construction.

Indian Institute of Technology Guwahati

TH-1255_09610610

The modification of CTAB extraction procedure published by Chang et al (1993) was used to extract total RNA from seed tissues (Table 3.2). It incorporates the addition of guanidinium iso thio cyanate (GITC) to facilitate the removal of excess miscellaneous and acidic phenol to bring down DNA contamination. Approximately 3 g of seed cotyledon was taken and ground in liquid nitrogen and 5ml of the CTAB extraction buffer (heated to 65 ^{o}C) was added and kept for 30 min in a water bath set at 65 ^{o}C. The contents were centrifuged at 12,000 rpm for 15 min. To the supernatant 2.5 ml of guanidium HCl buffer was added, shaken well and centrifuged at 12,000 rpm for 15 min. Equal volume of chloroform was added to the supernatant and shaken vigorously and centrifuged at 12,000 rpm for 15 min. The supernatant was again extracted with equal volume of phenol: chloroform (1:1) of pH-4.8 and centrifuged at 12,000 rpm for 15 min. RNA from the aqueous phase was precipitated at room temperature for 15 min by adding 1/10th volume of 3 M Sodium acetate (pH-5.2) and equal volume of isopropanol. Total RNA was treated with 200 units of RNase-free DNase I (Invitrogen) at 37 °C for 15 min. DNase I was added to eliminate nucleic acid and the timing of the step was optimized. DNase was removed by phenol/chloroform extraction. Total RNA pellet was recovered by sodium acetate precipitation; re-suspended in 50-70 µl DEPC-treated water, vortexed and incubated the tubes at 65 °C for 10 min to suspend the pellet thoroughly. Tubes containing ribonucleic acids were briefly centrifuged and were quantified by using spectrophotometer by monitoring the absorbance at 260 nm.

Table 3.2: CTAB-GITC buffer composition [Modified buffere (Cheng et al 1993)].

Component	**Concentration**	**Significance**
CTAB & GITC	2 & 1%	Disrupts plant cell wall
PVP	2 %	Precipitate polyphenolics
Tris-HCl	100 mM	Membrane permeability
EDTA	25 mM	Chelating agent
NaCl	5 M	Nucleic acid-neutralizer
Spermidine	0.05 %	Nucleic acid-precipitator
B-Mercaptoethanol	2 %	Reducing agent
Chloroform:isoamylalcohol	24:1	

TH-1255_09610610

3.2.3 RNA quantification

RNA was quantified through spectro-photometric recording of the absorbance at 260 nm. Recording the absorbance at 280 nm was used to verify the purity of RNA. Purity of RNA was also verified by running 2 µl of RNA on a 1.5 % formaldehyde-agarose gel containing 0.5 µg/ml of EtBr. Equal intensity of rRNA bands was taken as the criteria for equal quantity and uniform quality of the sample.

3.2.4 Isolation of mRNA from total RNA

About 1 mg of total RNA was used for mRNA purification using Qiagen mRNA separation kit, following the manufacturer's procedure with minor changes. The concentration of the mRNA was measured with spectrophotometer. Quantity of mRNA obtained was verified using 1.5 % formaldehyde agarose gel electrophoresis by comparing with known concentration of mammalian mRNA supplied with Clontech's library construction kit. Prime RNase inhibitor was added to the mRNA to protect it from RNase contamination and aliquoted samples were stored at -80 ^{o}C for library construction.

3.2.5 Formaldehyde agarose RNA gel electrophoresis

Agarose gel (1.2 %) was melted in a solution made with 10 ml 10 x MOPS buffer (200 mM 3-[*N*–morpholino] propane-sulfonic acid, 50 mM sodium acetate, 10 mM EDTA, pH 6.5-7.0) and 85ml sterile distilled water (121 °C). The solution was cooled to 50 °C and 4.5 ml 37 % formaldehyde added and poured into the gel casting tray. Samples of RNA were re-suspended in 10 µl formaldehyde gel loading buffer, heated at 65 °C for 5-10 min, chilled on ice and then loaded. The gel was run in 1 x MOPS buffer at 5 V/cm.

3.2.6 cDNA synthesis and cDNA library construction

cDNA synthesis was performed using Long distance PCR of Creator^{TM} SMART^{TM} (Switching Mechanism At the 5'-end of RNA Transcript) cDNA library Kit (cat. no. K1053-1). The Long Distance Polymerase Chain Reaction (LD PCR) was employed for the generation of full-length cDNA. The mRNA obtained from total RNA extracted

TH-1255_09610610

from seeds of *P. pinnata* was in sufficient quantities, and directly used for cDNA synthesis through LD PCR method (Figure 3.2).

3.2.7 First strand cDNA synthesis

The following reagents were combined in a flat caped PCR tube.

Component	Quantity
Total RNA	3 µl (~200 ng)
SMART IV oligonucleotide	1 µl
CS III/3' PCR primer	1 µl

The contents were mixed, spun briefly and incubated at 72 °C for 2 min. The tube was cooled on ice for 2 min, spun briefly to collect the contents at the bottom of the tube and the following reagents were added to the reaction tube.

Component	Quantity
5 x First strand buffer	2 µl
DTT (20 mM)	1 µl
dNTP mix (10 mM)	1 µl
Powerscript Reverse Transcriptase	1 µl

The contents were mixed by pipetting and incubated at 42 °C for 1hour in a PCR machine (Applied Biosystems, India). The first strand reaction was terminated by keeping the tube on ice and stored at -20 °C until further use.

3.2.8 Second strand cDNA synthesis

Component	Quantity
First strand cDNA	2 µl
Deionised water	80 µl
10 x advantage 2 PCR buffer	10 µl
50 x dNTP mix	2 µl
5' PCR primer	2 µl
CDS III/3' PCR primer	2 µl
50 x Advantage 2 polymerase mix	2 µl

TH-1255_09610610

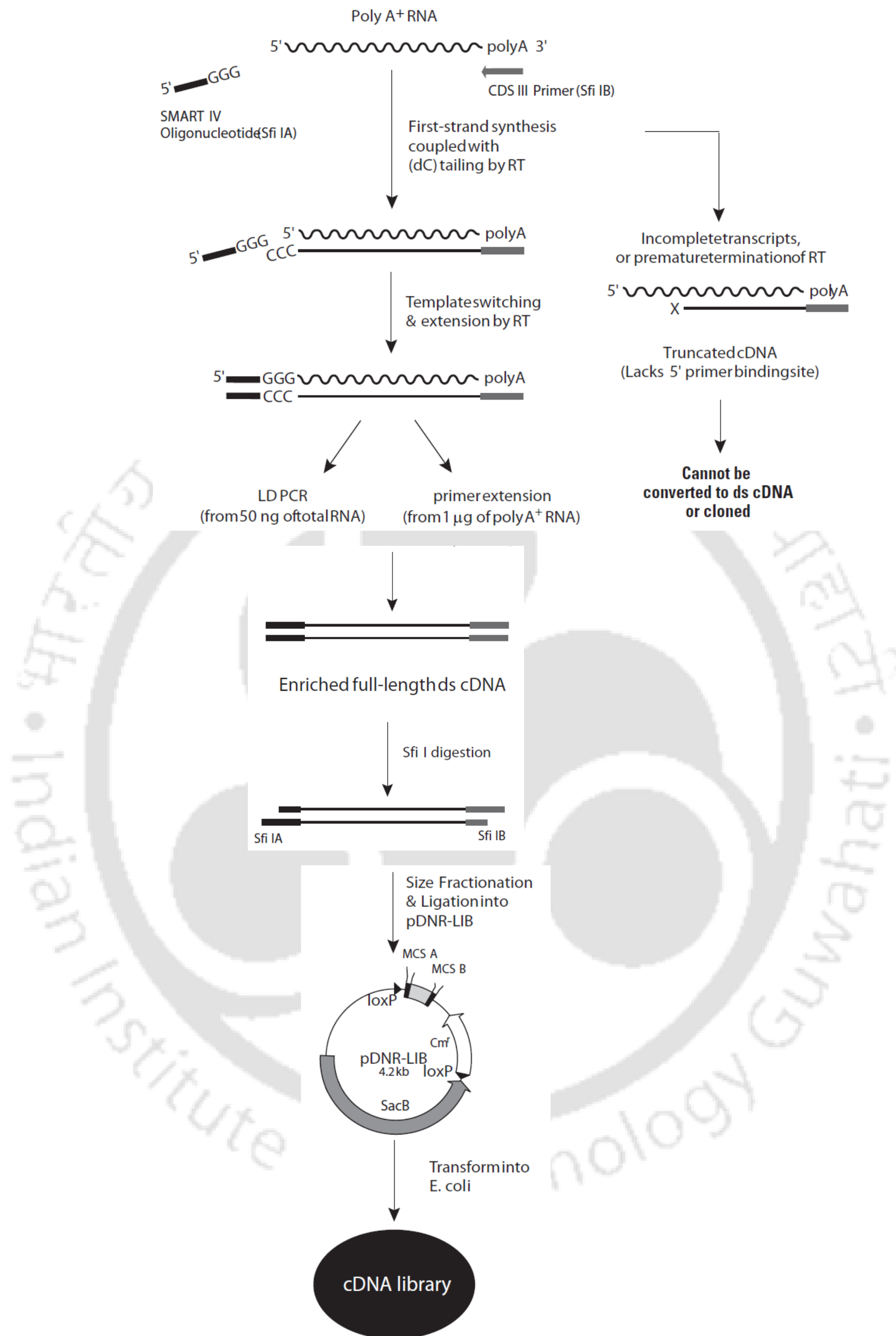


Figure 3.2: Flow chart of the Creator^{TM} SMART^{TM} cDNA Library construction kit protocol. The right side of the flow chart shows the fate of incomplete transcripts caused by RNA degradation or premature termination of reverse transcription (Source: Creator^{TM} SMART^{TM} cDNA Library construction kit Manual).

TH-1255_09610610

Above contents were mixed and centrifuged briefly before starting the second strand cDNA synthesis using PCR. The reaction conditions were as follows: 95 ^{o}C for 20 sec; 20 cycles of: 95 ^{o}C for 5 sec and 68 ^{o}C for 8 min. The ds cDNA was analyzed on 1 % TAE agarose gel to verify the size distribution. The cDNA product was treated with proteinase K to remove DNA polymerase activity. About 50 µl of ds cDNA that correspond to approximately to 2-3 µg of total RNA was used. Two microlitre of proteinase K (20 µg/µl) was added to the ds cDNA in a PCR tube and incubated at 45 ^{o}C for 20 min. Phenol: Chloroform extraction of the proteinase K treated ds cDNA was done following the manufacturer's instructions and pellet was re-suspended in 79 µl of deionized water.

3.2.9 *Sfi* digestion of ds-cDNA

The following components were added, mixed and incubated in a PCR machine (Applied Biosystems, India) for 2 hours and 15min at 50 ^{o}C. Similarly the plasmid vector pDNR-LIB was also digested with same enzyme, followed by dephosphorylation.

Component	**Quantity**
ds cDNA in H_2O	79 µl
10X *Sfi* Restriction buffer	10 µl
Sfi Restriction enzyme	10 µl
100x BSA	1 µl

3.2.10 Size fractionation of *Sfi* digested ds cDNA

CHROMA SPIN-400 columns provided by the manufacturer (Clontech) were used for size fractionation. The flow rate of the column was adjusted to 1-drop/40 sec. After the column storage buffer stopped dripping and the column bed became visible, 100 µl of *Sfi* I digested ds cDNA dyed with 2 µl of 1 % xylene cyanol was added exactly at the center of the bed. After the cDNA containing solution got absorbed into the column, 100 µl of fresh column buffer was added and allowed to drip out of the column. After the dripping ceased, 600 µl of column buffer was again added and single-drop fractions were collected immediately using 1.5 ml micro centrifuge tubes. Twenty such fractions

TH-1255_09610610

of 40 µl volume were collected. Size distribution of collected cDNA fractions were verified by loading 3 µl aliquots through TAE-agarose gel electrophoresis run at 150 V for 10 min. Double stranded *Sfi* I digested cDNA fractions showing higher size distribution on agarose gels were grouped together and precipitated following the manufacturer's instructions and the cDNA pellet was suspended in 7 µl of deionised water.

3.2.11 Ligation of ds cDNA with the plasmid vector

The cDNA was ligated to *Sfi* I digested, dephosphorylated to pDNR-LIB vector provided in the library construction kit following the manufacturer's instructions. Three ligation reactions with 0.5 µl, 1.0 µl and 1.5 µl of cDNA and 100 ng (1 µl) of pDNR-LIB vector were done (insert to vector ratio: 0.5:1, 1:1 and 1.5:1). The ligation reactions were carried out at 16 ^{o}C over night and the purified cDNA pellets were suspended in 5 µl of deionised H_2O.

3.2.12 Transformation of ligated cDNA and its efficiency

Escherichia coli-DH5α cells were used for transformation of ligated products. The *E. coli* cells were suspended in 0.1 M $CaCl_2$ and stored in 200 µl aliquots in 0.5 ml eppendorf tubes at -80 ^{o}C. After the cells were taken out from the freezer, they were thawed in hand and then immediately left on ice for 10 min. The competent cells were mixed with DNA sample (DNA-plasmid chimeras of three different combinations) and were left on ice for another 45 min, followed by heat shock treatment at 42 ^{o}C for 90 sec. Around 800 µl of NZY medium (NZ amine Yeast extract) was added and then incubated at 37 ^{o}C for 1hour to allow expression of antibiotic resistance gene. Eventually the cells were plated out by glass spreader on selective agar (NZY Tet^+ Kan^+ Cam^+ agar plate) at certain different concentration to achieve the best transformation results. To know the efficiency of the transformation, positive control (vector without insert) and negative control (bacterial cell suspension without plasmid/insert) were kept as control reactions. The obtained colonies were counted and the number of total colonies was considered to check the efficiency of the transformation. Percentage of recombinant clones were determined by screening the colonies using M13 primers [M13F (5'-TGTAAAACGACGGCCAGT-3') and M13R (5'-CAGGAAACAGCTATGACC-3')] provided in the library construction kit along

TH-1255_09610610

with advantage 2 PCR kit (BD Biosciences, cat. no. K1910-y). Colonies were randomly picked using sterile toothpicks and inoculated into 50 µl of TE buffer in separate 1.5 ml micro centrifuge tubes to perform colony PCR. The tubes were boiled for 5 min and the reaction was setup as follows:

Component	**Quantity**
Boiled colony lysates	1 µl
10 x Advantage 2 PCR buffer	2.5 µl
dNTP mix (10 mM each)	0.5 µl
M13 forward primer (20 µM)	0.5 µl
M13 reverse primer (20 µM)	0.5 µl
Deionised water	19.5 µl
50 x Advantage 2 Polymerase mix	0.5 µl

The following program was used to amplify the cDNA inserts: 94 ^{o}C for 30 sec; 25 cycles of: 94 ^{o}C for 30 sec followed by 68 ^{o}C for 2 min and; 1 cycle of: 68 ^{o}C for 5 min. The PCR products were run on a 1.2 % TAE-agarose gel with DNA size markers to check for the presence or absence and the size distribution of cDNA inserts.

3.2.13 Transfer of bacterial colonies onto nitrocellulose filter

After the selection of possible transformants was completed using selective agar, the transformed colonies were picked out from the agar plate with sterile toothpicks carefully and transformed onto nitrocellulose filters on selective agar by streaking diagonally on the gridded filters (Hanahan & Meselson 1980, 1983) and the plates were then incubated at 37 ^{o}C for overnight.

3.2.14 Replication of nitrocellulose filter from master filter

This method of production of replica filter for different hybridization has been described by Maniatis et al (1982) and Davis et al (1980). Three pieces of sterilized 3 MM paper were laid down on a clean, alcohol wiped glass plate with the top piece being wet with sterile water. A master filter from the selective agar was removed and placed with colonies up on the wet 3 MM paper. A nitrocellulose filter was then put accurately on the master filter. A wet piece of 3 MM paper was then placed on the

TH-1255_09610610

filters followed by two other dry 3 MM papers. A glass plate was then placed on the whole stack and was pressed gently and evenly. The filter copy was removed and grown on selective agar overnight at 37 ^{o}C whilst the master filter was stored at -20 ^{o}C for further use.

3.2.15 Lysis of bacterial colonies on nitrocellulose replica

Replica filter was removed from the selective agar plate and was placed on 3 MM papers soaked with 10 % SDS for 3 min, then denaturing solution (0.5 M NaOH, 1.5 M NaCl, 1 mM EDTA) for 5 min, neutralizing solution (3 M NaCl; 0.5 M Tris.Cl, pH 7.0; 1 mM EDTA) for 5 min, and followed by 2 x SSC wash (0.15 M NaCl; 15 mM sodium citrate, pH 8.0) for 5 min. The replica filter was dried on 3 MM paper in between each step. The filter with released plasmid DNA was then air dried, baked at 80 °C for 2 hours in a vacuum oven and was ready for hybridization.

3.2.16 Amplification of fatty acid biosynthetic genes (partial) for probe preparation

The primers for fatty acid biosynthetic genes [steroyl-ACP-desaturase (SAD), fatty acid desaturase (FAD2)] were designed based on sequence information of the SAD and FAD2 cDNA clones (AAB00859, NM_001251324.1) of *Glycine max.* Approximately 1 µg of total RNA extracted from seeds of *Pongamia,* was taken as template and One-step RT PCR reaction mix (Qiagen, India) along with gene specific primers were used to amplify the partial sequences of fatty acid biosynthetic genes (*SAD* and *FAD2*) separately. The PCR cycle conditions were as follows: Initially the reaction was set up at 50 ^{o}C for 30 min to synthesize single stranded cDNA and then followed by 95 ^{o}C for 15 min, 30 cycles of initial denaturation at 95 ^{o}C for 30 sec, annealing (Tm varies according to primers) for 30 sec, and extension at 72 ^{o}C for 1 min. Final extension was kept at 72 ^{o}C for 10 min and the PCR product was run on 1.2 % agarose gel to check the amplification of the genes followed by purification with gel extraction kit (Qiagen, India) following the manufacturers protocol. Eluted DNA was diluted to a concentration of 25 ng in 45 µl of Tris-EDTA buffer (pH 8.0), denatured at 95 ^{o}C for 10 min in a thermal cycler.

TH-1255_09610610

Primer	Sequence (5'-3')
SAD-F1	GAAGCCATTCACTCCTCC
SAD-R1	TCAACTCGACCACTCAAG
FAD2-F	GCGTTTAGCGATTATCAG
FAD2-R	GCTTTTCGGTTTCGGCAC

3.2.17 ^{32}P radio-labeling of probe using random oligo-nucleotides as primers

The denatured probe DNA was added to the tube of Amersham Rediprime random primer labeling kit (cat. no. RPN 1633) and the tube was flicked to mix the contents. About 5 µl of α-^{32}P-dCTP was added, mixed and incubated at 37 ^{o}C for one hour. The reaction was terminated by an addition of 2 µl of 0.2 M EDTA, pH 8.0. Finally ^{32}P labeled probe was purified using Sephadex G-50 column, denatured by boiling in a water bath for 5 min and immediately transferred to ice.

3.2.18 Hybridization

To hybridize probes (SAD/FAD2) with blotted plasmid-nucleic acid sequences, hybridization method given by Amersham (1985) was used. Firstly, pre-hybridization solution was prepared as follow:

Solution	Volume	Final concentration
20 x SSC	7.5 ml	6 x SSC
100 x Denhardt's solution	1.25 ml	5 x
10 % SSC	1.25 ml	0.5 %

The solution was prepared up to 25 ml with sterile distilled water. Then 0.5 ml of a 1mg/ml solution of Salmon Sperm DNA was denatured by heating in a boiling water bath for 5 min; chilled on ice and added to the pre-hybridization solution. Secondly, the filters were incubated with rotating for at least 1 h at 65 °C in pre-hybridization buffer in hybridization bottle. Hybridization solution (25 ml) was prepared just like the prehybridization solution. About 0.5 ml of 1 mg/ml solution of Salmon Sperm DNA and probe were denatured, added to prehybridzation buffer and

TH-1255_09610610

incubated for 12 h at 65 °C with gentle shaking. After hybridization, the filters were incubated in 50 ml 2 x SSC at 65 °C for 15 min. Finally, the filters were transferred to 50 ml of 2 x SSC containing 0.1 % SDS, and incubated at 65 °C for 30 min. The filters were air dried and wrapped in Saran wrap, followed by autoradiography.

3.2.19 Autoradiography and identification of positive clones

X-ray film (X-OMATTM AR Scientific Imaging Film, Kodak, Rochester, NY, USA) was exposed after pre-flashing once with γ-ray onto the filters. After keeping 1-2 days at -80 °C the film was taken out in the dark room and placed in the developer solution for 8min. It was then rinsed with tap water, drained and immersed into fixer for 3 min. The fixer was washed under tap water for about 20 min, and then dried.

After the development of the autoradiograph, the blot was aligned with the master filter in accurate and correct orientation. The dark dots on the film represented the positive recombinants and the corresponding colonies on the master filter were picked out and were streaked on selective agar plate to obtain single colony. The selected clones were sequenced by Macrogen sequencing service, South Korea.

3.2.20 Processing and analysis of cDNA sequences

After sequencing, the contigs were assembled from redundant reads using DNA baser software, which also removes the poor quality sequences generated by standard sequencer. The sequences of fatty acid biosynthetic genes have been sequenced in other oil yielding and legume plants (i.e., *Jatropha*, Olive, Sesame, Soybean and Peanut), thus the fatty acid biosynthetic gene sequences are available to identify the cloned cDNA sequences at the NCBI genomic BLAST web site (www.ncbi.nlm.nih.gov). The processed cDNA sequence was used to query the fatty acid biosynthetic genes with a Blastn algorithm (default parameters). Sequences with an E-value of e-0 or up to e-100 were classified as significantly matching when compared with other fatty acid biosynthetic desaturase genes. Those sequences with an E-value of e-0 were given special consideration because this could represent an mRNA matching to various exons of the same gene (Scheetz et al 2003).

TH-1255_09610610

3.3 RESULTS AND DISCUSSION

The results obtained for RNA isolation, mRNA purification, the various steps of cDNA library construction, and the screening of libraries using gene probes are presented and discussed.

3.3.1 General precautions for RNA handling

Care was taken to avoid contamination with RNase enzymes which cause destruction of RNA. All glasswares were baked at 80 °C in an oven for overnight. Plastic wares were incubated with 0.1 % DEPC overnight and then autoclaved for 40 min. General lab equipments, pipettes and working surfaces were cleaned with the solution containing 1 % NaOH and 10 % SDS, which is effective in eliminating RNase contamination. Gloves were frequently changed.

3.3.2 Total RNA isolation

Finding a suitable protocol for RNA extraction was a high priority, as RNA integrity is of vital importance for the construction of full-length cDNA libraries.

Effort has made to construct cDNA library from the early immature seeds (90-DAF) of *P. pinnata,* failed to yield high quality intact RNA by using standard CTAB method of Chang et al (1993) and Trizol reagent. This was possibly due to either no RNA recovery or an inability to reverse-transcribe the RNA produced. To circumvent this problem, total RNA was isolated from seeds (90-DAF) using CTAB method of Chang et al (1993) with slight modifications (addition of GITC and increase in NaCl concentration from 2 M to 5 M). Large quantities of high quality RNA could be extracted before downstream implementation by CTAB-GITC method. Fluorescence profile was observed when the RNA samples were electrophoresed on a formaldehyde (denaturing) 1.2 % agarose gel. The distinct two banded profile (Figure 3.3) was in bright contrast to the conventional two banded profile which is characteristic of most plant total RNA preparations, and the faint fluorescence signal that could be discerned below the two rRNA bands, was an indication of excellent RNA integrity. The bands which corresponds to CTAB-GITC extraction process is showing two intact bright bands; 28S rRNA and 18S rRNA where the former was equal to or more abundant than the latter indicating that little or no RNA degradation or contamination occurred during

TH-1255_09610610

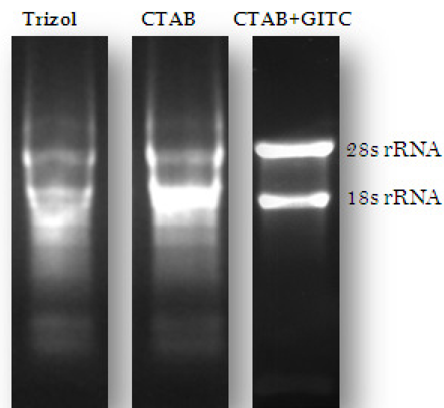


Figure 3.3: Electrophoresis profile of a *Pongamia* total RNA samples extracted from different methods i.e., Trizol, GITC and CTAB-GITC. Approximately 100 ng aliquot loaded in 6 x Formaldehyde loading dye and electrophoresed on a 1.2 % w/v formaldehyde agarose gel run in RNase-free 1 x TAE at 4 V/cm for 45 min.

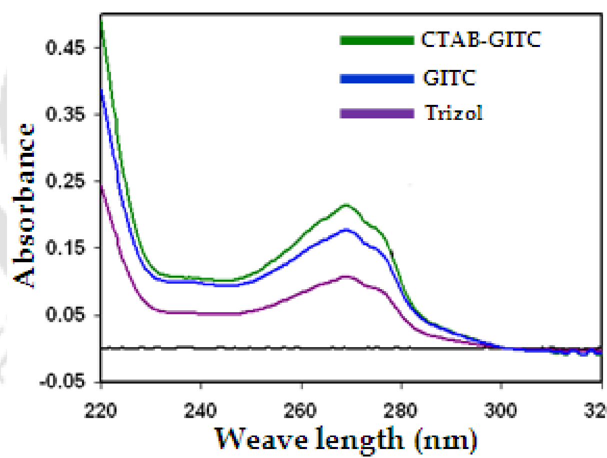


Figure 3.4: Graph depicts the UV-Spectrophotometric readings of total RNA extracted by various methods (Trizol, GITC and CTAB-GITC).

TH-1255_09610610

isolation. But in case of total RNA which is extracted from other two methods (Trizol and CTAB) is not clear, though it gives two intense bands which corresponds to 28s and 18s rRNA, lot of degradation was observed. Due to this reason total RNA was not considered for further experimental analysis.

Spectrophotometer analysis revealed that A_{260}/A_{280} ratio of RNA sample is 1.97, elucidating that there was no contamination of proteins, DNA, polyphenolics, and polysaccharides in isolated RNA by using modified CTAB method described. The A_{260}/A_{280} ratio of the RNA samples isolated by CTAB method of Chang et al (1993) and Trizol were very low (1.50 and 1.45) (Figure 3.4). By using the modified protocol of Chang et al (1993) high-quality RNA from early immature seeds of *P. pinnata* was successfully isolated. The RNA yields ranged from 350-500 µg/g, sufficient enough to carry out cDNA library construction. RNA yield obtained were 70-400 µg/g from CTAB method of Chang *et al.*, (1993) and 55-200 µg/g from Trizol method. Using the protocol described in the current study, the yields of RNA were higher or within the range of previously published data. For example, RNA yields of 200-800 µg/g (Ruuska & Ohlrogge 2001) and 150-200 µg/g (Vicient & Delseny 1999) have been reported from seeds of *Arabidopsis thaliana*. Although no explanation could be found for the elevated ratio of the purified mRNA samples (~50 ng/µl), a survey of the literature indicated that this is a general phenomenon, and the samples could be regarded as adequately purified for cDNA library construction.

The original method of Chang et al (1993) and Trizol reagent was ineffective for the isolation of good quality RNA from *Pongamia* samples. The main problem was that the homogenate of *Pongamia* seeds formed clumps like structure that could not be properly mixed in CTAB extraction buffer as well as in Trizol reagent. This interferes in initial phase separation (aqueous and organic) by centrifugation. In addition another problem was faced in the last step i.e., RNA precipitation using standard protocols of CTAB and Trizol reagent. Pellet was found to contain lot of mucilage which co-precipitate with RNA and formed a glue-like gel structure that cannot be dissolved or suspended in DEPC-treated water. The main reason is that tree species often contain large amounts of proteins, DNA, polyphenolics, polysaccharides and phenolic compounds that are difficult to separate from RNA but are easily identified, because

TH-1255_09610610

they make the RNA pellet sticky and gelatinous (polysaccharides) (Gao et al 2001). In this modified protocol combined use of CTAB with GITC and high concentration of 5 M NaCl, pre-chilled isopropanol at the precipitation step significantly increased the efficiency of proteins, secondary metabolites and polysaccharide removal. The method of CTAB-GITC provided solutions for all of the difficulties encountered with other extraction protocols. DNase I was applied to eliminate genomic DNA, when DNase is added at the end of the RNA isolation procedure, RNA migration on the agarose gel is abarrent. Therefore, DNaseI was applied immediately after precipitation with isopropanol.

3.3.3 cDNA library construction

mRNA purified from the total RNA was sufficient and was observed as complete smear between 1.0 kb to 2.6 kb which reveals most of the transcripts are full length. The experiment was performed in duplicates and the results were consistent (Figure 3.5).

Construction of the cDNA library was accomplished without much difficulty. The only alteration to the Creator™ SMART™ Library construction kit protocol was in the number of cycles used for cDNA amplification. These had to be optimized, as the proposed number of cycle repeats resulted in a cDNA smear that continued up to the high molecular weight region of the gel. Since this is usually an indication of over cycling (Chenchik et al 1998), the number of cycles was lowered until good results were obtained (Figure 3.6). The size distribution of the ds cDNA product ranged from ~0.3 kb to ~2.6 kb correlated well with the size ranges reported for other cDNA libraries constructed from early immature seed tissues. The high intensity of the smear observed between 0.7 kb to 1.8 kb represented highly expressed sequences in the transcriptome of the *Pongamia*. The smear did not contain cDNA less than 0.3 kb in size or impurities. Since the results indicated that reverse transcription and cDNA amplification had been successful, the amplified product was digested with *Sfi*I-enzyme and preceded for ligation with vector.

TH-1255_09610610

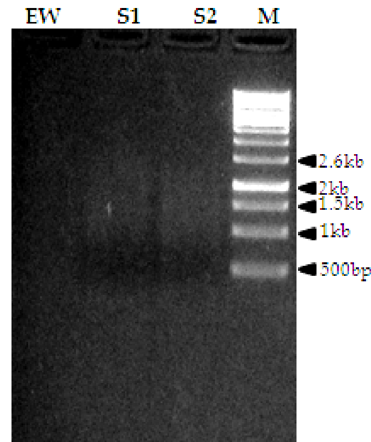


Figure 3.5: Purified mRNA as smear; M-1Kb DNA marker. EW-Empty well; S1, S2-mRNA as duplicates from 90-DAF seeds of *P. pinnata*.

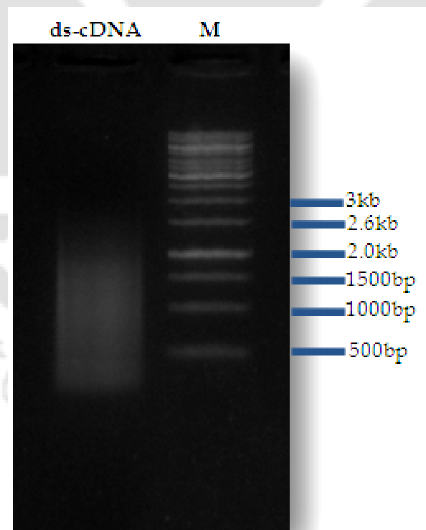


Figure 3.6: Amplified cDNA profile after 22 cycles of PCR (5 µl aliquots electrophoresed on a 1.2 % agarose gel. M-1 kb DNA marker (Neubiogene).

TH-1255_09610610

All transformation reactions were successful, with nearly confluent plates obtained from all but the negative control reaction (Figure 3.7). Lack of growth on the negative control plate indicated that there had been no contamination of the ligation reaction (with insert), or the transformation reaction (with uncut vector). It also suggested that the antibiotic selection was effective, and that growth on the other plates was a result of antibiotic resistance acquired through plasmid transformation. This was substantiated by a control plate which contained no chloramphenicol, and onto which an aliquot of the negative control reaction was plated. Confluent growth on this plate demonstrated the efficiency of antibiotic selection on all other plates. Based on the bacterial colonies obtained according to vector to insert ratio, the clones were calculated and the cDNA library consisted of approximately 1.2×10^4 (0.5:1), 1.3×10^4 (1:1) and 1.5×10^4 (1.5:1) respectively. Clones in all transformation reactions that give confidence to screen the desired clones by using gene specific probes were selected to screen the library.

3.3.4 Size distribution, evaluation of clone inserts and screening for full length genes

The size-distribution of the library was estimated by PCR amplification (using M13 forward and reverse primers) of the inserts of randomly picked colonies from ~1.5×10^4 clones of 1.5:1 vector to insert ratio. Clones that produced multiple products (identified by multiple banding in the electrophoresis profile), were discarded from the analysis, as these results were unclear. Nearly all of the remaining amplification products migrated in the 2.3 kb to 0.7 kb range, with most of them migrating between 2.0 kb to 0.8 kb of the ladder range (Figure 3.8). The largest insert was estimated to be ~1.6 kb (clone no.137); the bands which are less than 0.8 kb were eliminated from library to consider full length genes. These data suggested the synthesis of *Pongamia* early seed (90-DAF) cDNA and ligation with pDNR, transformation of *E. coli* with recombinants and subsequent inserts size determination were all successful.

TH-1255_09610610

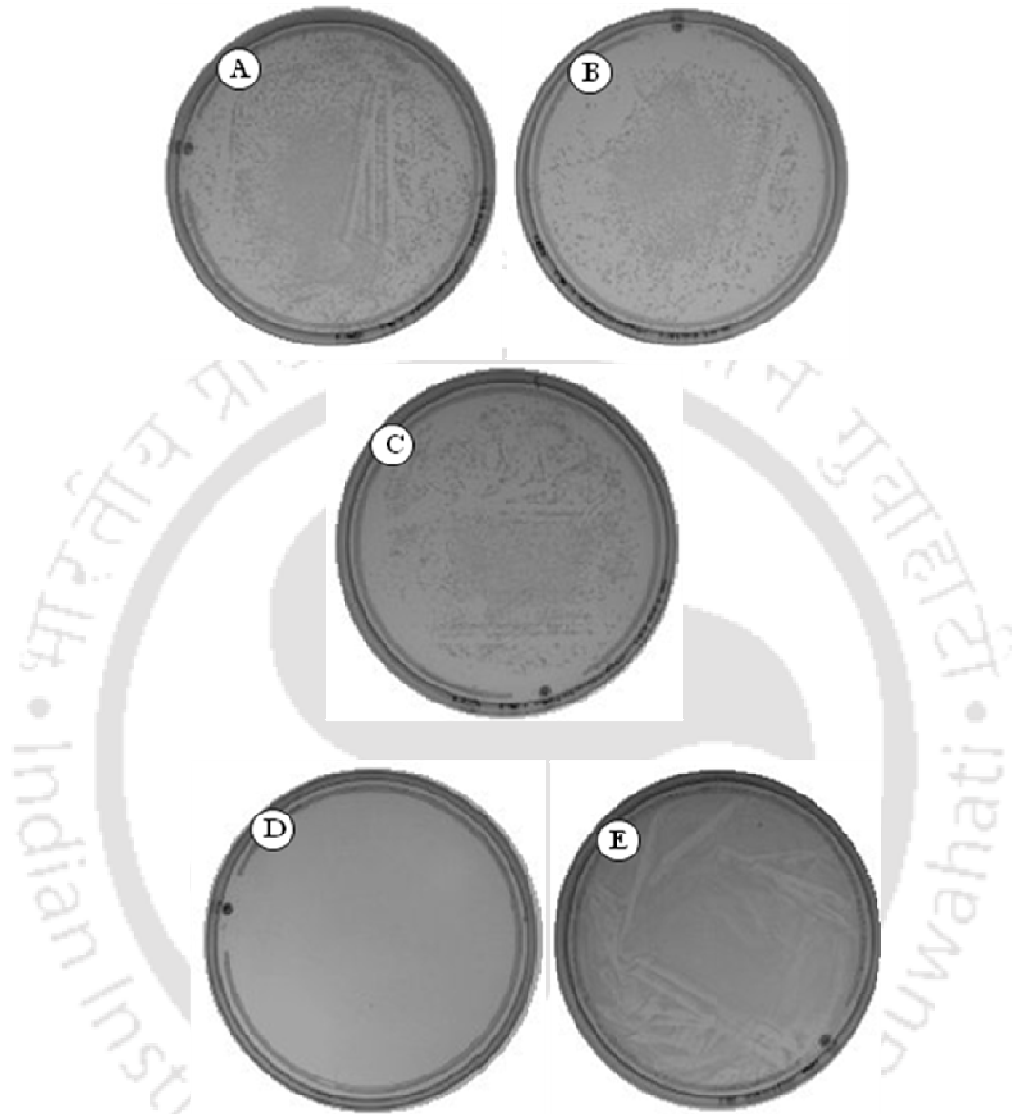


Figure 3.7: *E. coli* growth on 30 µg/ml chloramphenicol selective medium after platings with various transformed cell suspensions. Plates (A) Ligation-A, (B) Ligation-B and (C) Ligation-C represents ligation reactions with volumetric vector:insert ratios of 0.5:1, 1:1 and 1.5:1. Controls were provided by platings of the negative control reaction onto selective (D), and non-selective medium (E).

TH-1255_09610610

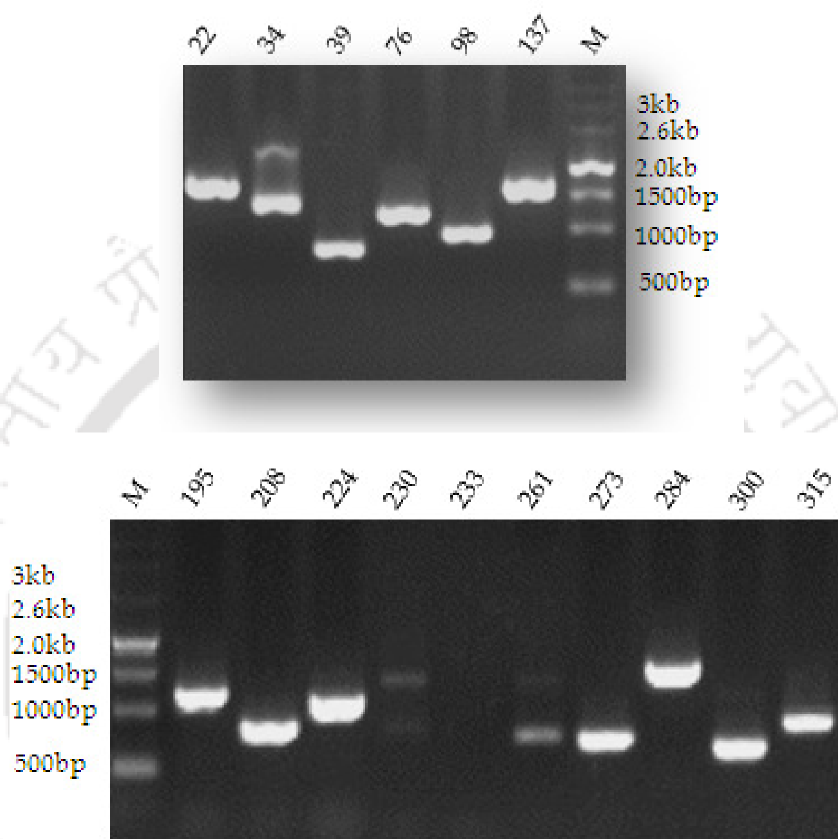


Figure 3.8: A 1.2 % agarose gel electrophoresis of PCR products of cDNA inserts selected randomly from the amplified cDNA library. A part of colony selected randomly and insert amplified by PCR; M-1kb DNA marker.

TH-1255_09610610

Two separate replica filters were produced from the master filters, the bacterial colonies were then lysed and fixed on nitrocellulose filters by baking at 80 °C for 2hrs so that the probes could later hybridized with them. The two different types of probes (SAD/FAD2) were radiolabelled by random primed labeling method separately. Both the probes were purified along Sephadex G-50 column and by liquid scintillation counting. The labeled DNA probes of SAD and FAD2 were hybridized with two separate sets of replica filters and autoradiographed. According to Geiger-Muller counter checked, the radioactive strength of the filters hybridized with probe SAD was strong [about 50 counts per second (cps)] while the other set that hybridized with probe FAD2 was moderate (about 33cps). Therefore, autoradiograph of SAD was exposed for only four hours while the other was exposed for one day. Hybridization between the probes and the filters were revealed on the autoradiographs after being developed (Figure 3.9). Both the autoradiographs were superimposed onto the master filters. The bacterial colonies corresponded to the dark spots on the autoradiograph were picked out (1 and 2 colonies from SAD and FAD2 respectively), the positive clones were streaked on selective agar plates (containing chloramphenicol as selective marker) to obtain single colony and the colony PCR with M13 primers resulted in several amplified products ranging between approximately 1.2-1.6kb. The PCR products were eluted, purified/cleaned and were further sequenced by standard sequencing method.

The raw sequences obtained from sequencing service (Macrogen, Korea) were processed by generating contig assembly. The software DNA baser was able to completely remove all ambiguous/poor quality sequences from the forward and reverse sequences obtained from standard DNA sequencer. Final contigs were used for BLAST analysis to find out suitable genes (*SAD/FAD2*) that are involved in fatty acid biosynthetic pathway. The results showed that all the selected positive clones matched with high similarity scores their respective counterparts in the GenBank database namely *FAD2-1, FAD2-2* and *SAD* of fatty acid biosynthetic genes. The positive clones were stored as glycerol stocks (20 %) at -80 ^{o}C for further experimental analysis.

TH-1255_09610610

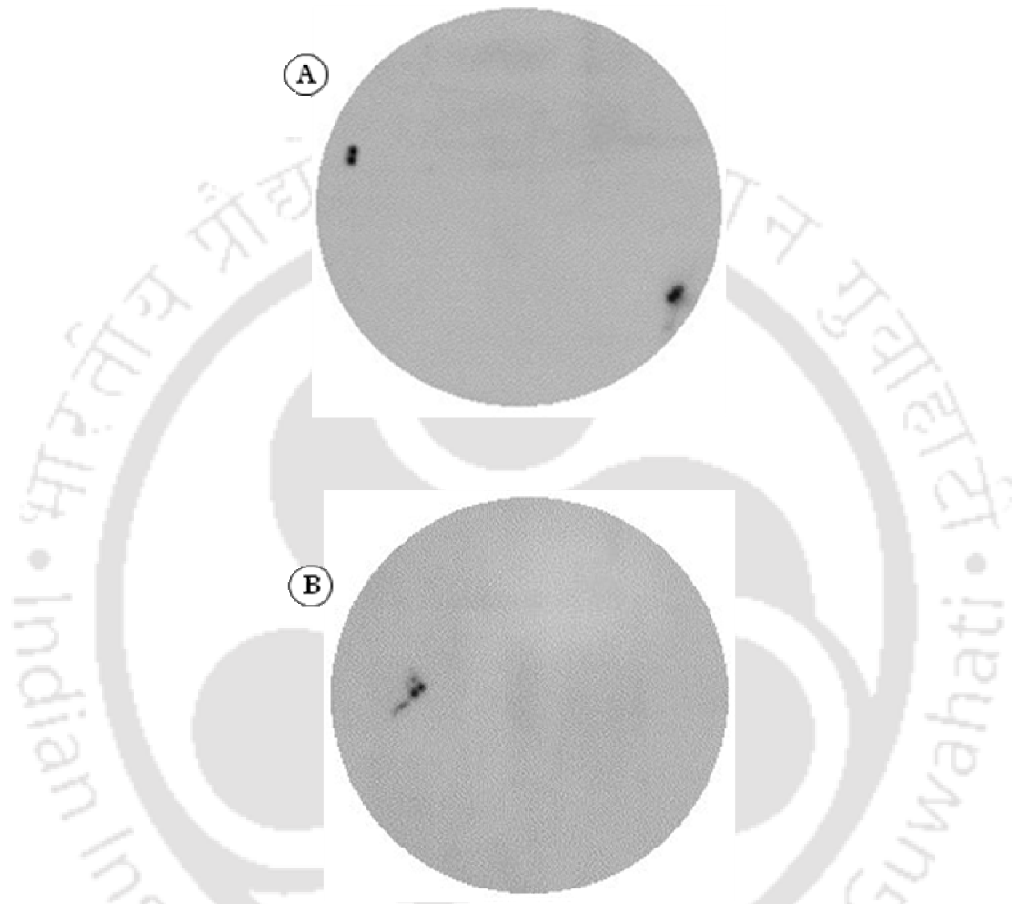
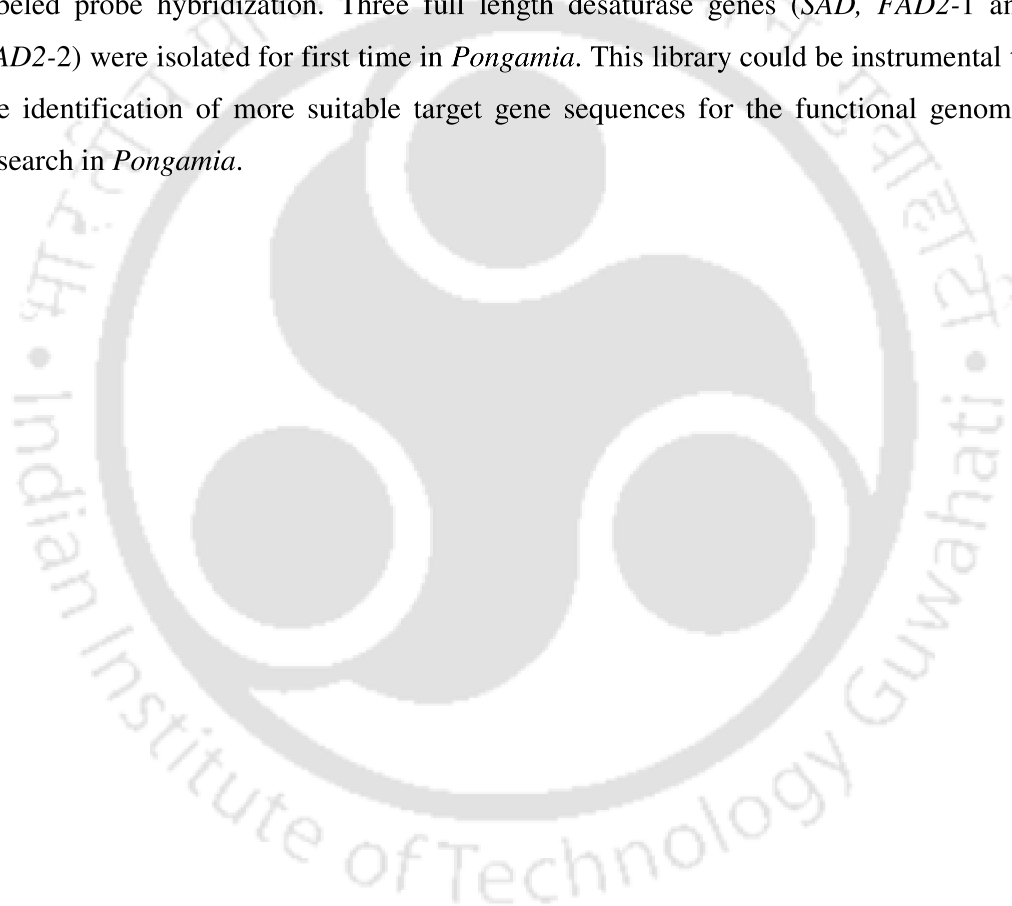


Figure 3.9: High density filters of the early immature seed *P. pinnata* cDNA lilbrary with (A) FAD2 and (B) SAD probes. Each clone was duplicated in a predetermined pattern represented by the close doublet signals. Two hybridization signals were observed for FAD2, but only one for SAD.

TH-1255_09610610

3.4 CONCLUSION

In conclusion, the proposed RNA extraction protocol in this study is an efficient, inexpensive and highly reproducible technique to isolate total RNA from *P. pinnata* tissues, recalcitrant to extraction by standard protocols of Cheng et al (1987) and Trizol reagent. *P. pinnata* early immature seeds cDNA library was successfully constructed using LD PCR approach. cDNA library consisted of 1.5 x 10^4 clones that were sufficient enough to screen for the desired clones. Screening for full length genes which are involved in the desaturation of fatty acid biosynthesis were performed by radio labeled probe hybridization. Three full length desaturase genes (*SAD, FAD2*-1 and *FAD2*-2) were isolated for first time in *Pongamia*. This library could be instrumental to the identification of more suitable target gene sequences for the functional genomic research in *Pongamia*.


Indian Institute of Technology Guwahati

TH-1255_09610610

Chapter 4

Structural and Functional Characterization of Δ^9-Steroyl-ACP- Desaturase Gene

The chapter explains the characterization of steroyl-ACP-desaturase gene including sequence identification, transcriptional studies, copy number determination, structural comparison by homology modeling and its validation.

TH-1255_09610610

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF Δ^9-STEROYL-ACP- DESATURASE GENE

4.1 INTRODUCTION

Fatty acids (FA) have many and diverse roles in plants. Broadly speaking long chain fatty acids are synthesized *de novo* from small precursors (sucrose) ultimately derived from photosynthesis and convert it into three major storage components, namely protein, starch and fatty acids. Fatty acids are synthesized by a well defined pathway involving two carbon elongation and bond desaturation. They are the major structural component of membrane lipids, provide a substantial reserve of free energy and serve as key precursors for the biosynthesis of messengers in signal transduction mechanisms that influence plant growth, development and responses to environmental cues (Browse & Somerville 1991; Munnik et al 1998; Nandi et al 2003; Somerville et al 2000; Weber 2002). In oil seed crops, seeds store a large proportion of their reserves of energy and reduced carbon as lipids. In plants the *de novo* synthesis of fatty acids occurs primarily in the plastid. These fatty acids are used for the synthesis of plastidial and other cellular membranes in all cells. In certain plant tissues, most notably in seeds, they are also used for the synthesis of storage oils such as triacylglycerols (TAGs).

A major fraction of the FA in plants are the polyunsaturated fatty acids, linoleic ($18:2\Delta^{9,12}$) and a-linolenic ($18:3\Delta^{9,12,15}$). Consequently, the desaturation of fatty acids is an important aspect in oil biochemistry since it determines the level of unsaturation (Knutzon et al 1992; Mikkilineni & Rocheford 2003). Fatty acid desaturase enzymes catalyze the conversion of saturated fatty acids to unsaturated fatty acids by introducing the first double bond into saturated fatty acids. Group of enzymes involved in desaturation reaction have been identified in all eukaryotes, cyanobacteria and in some *Bacillus* bacteria as well (Bloomfield & Bloch 1960; Fulco 1974). Three types of desaturases are noticeable depending on the kind of compounds esterified to fatty acids (Murata & Wada 1995). Desaturases identified in plants, animals and yeast are membrane-bound proteins with acyl chain attached to CoA or lipids. The only known soluble desaturase is the plant stearoyl-ACP desaturase (SAD) specific to stearic acid

TH-1255_09610610

localized in plastids. SAD catalyzes the desaturation of stearoyl-ACP to oleoyl-ACP and plays a key role in determining the ratio of saturated fatty acids to unsaturated fatty acids in plants (Lindqvist et al 1996) and this ratio are closely related to many functions of plants, especially with regard to acclimatization to low-temperature (Kodama et al 1995).

Although there are other fatty acid desaturases in plants, the SAD catalyzed introduction of the first double bond at position nine in the FA carbon chain is the most important desaturation for the gel state to liquid crystal state transformation of membranes (Los & Murata 1998). SAD also play a key role in environmental changes like cold resistance in plants (Kodama et al 1995; Tasseva 2004), senescence regulation, resistance to fungal infection and mechanical damage (Kachroo 2003; Lea et al 2003). Many genes coding for SAD have been cloned from different plants (*Carthamus tinctorius*, *Thunbergia alata*, *Ricinus communis, Solanum tuberosum*) and the structure and functions of several SAD have been studied (Davydov et al 2005; Lindqvist et al 1996). Antisense expression of *Brassica rapa SAD* gene in *Brassica napus* led to dramatically increased stearate levels (up to 40%) in the seeds of transgenic *B. napus* (Knutzon et al 1992). In the reverse, when the *SAD* gene from *Lupinus luteus* was over expressed in tobacco, the transgenic tobacco contained very high level of oleic acid (up to 60 %) in comparison with control plants (Zaborowska et al 2002). These imply that SAD promises to modify the composition of plant fatty acids by introducing/manipulating *SAD* gene.

Pongamia pinnata (L.) Pierre, an arboreal legume, is a member of the subfamily Papilionoideae and family Leguminasae, native to tropics and temperate Asia including part of India, China, Japan, Malaysia, and Australia. Commonly it is referred as karanj, pongam, dalkaramch etc. *Pongamia* is drought resistant, semi-deciduous, nitrogen fixing leguminous tree. The tree is well suited to intense heat and sunlight and its dense network of lateral roots and thick long tap roots make it drought tolerant (Bobade & Khyade 2012). Historically, this plant has been used in India and neighboring regions as a source of traditional medicines, animal fodder, green manure, timber, fish poison and fuel. More importantly, *P. pinnata* has recently been recognized as a viable source of oil for the burgeoning biofuel industry (Kesari et al 2008, 2009; Scott et al 2008). *Pongamia* has received much attention because of its high content of seed oil (~28–39

TH-1255_09610610

%, Kesari et al 2012), which contains high amounts of C18 fatty acids, of which oleic acid and linoleic acid represents about 46 and 27.1 %, respectively.

Most research on *P. pinnata* has focused on techniques for extracting oil from seed (Bobade & Khyade 2012; Shweta et al 2004). No reports till date on cloning of fatty acid genes in *P. pinnata*. As a first step towards understanding the reaction mechanism and the regulation of expression of fatty acid biosynthetic genes in *Pongamia*, we isolated and characterized a cDNA containing the complete coding region of the fatty acid gene (*PpSAD*), and analyzed its expression in different tissue types. As is well known, the function of protein is determined essentially by its corresponding 3D structure. Therefore, in the present investigation, based on the crystal structure of *Ricinus communis* (PDB code 1AFR), which was reported in 1996 by Lindqvist, the 3D structure of *Pongamia* steroyl ACP desaturase (PpSAD) is constructed theoretically and then the binding site is also evaluated by homology modeling.

4.2 MATERIAL AND METHODS

4.2.1 Plant material and growth

Seed sampled at different time intervals were collected from sila forest range, North Guwahati, 26 °14’6” N, 91 °41’28” E, Assam, India. Collected samples were rinsed twice with 70 % ethanol and immediately stored at -80 ^{o}C for further experimental purposes. Seedlings were grown under green house conditions [16 h photoperiod at 28±2 ºC with a relative humidity of 75 %; light intensity of 125 $\mu mol/m^2/S^1$ was provided by fluorescent light (Philips India Ltd., Thane, India)] and the different tissues were collected from grown saplings for expression studies.

4.2.2 Nucleic acid (DNA & RNA) isolation and quantification

For DNA extraction healthy seeds from elite genotype of *Pongamia* were used (~5 g fresh weight). Genomic DNA isolation from the seeds of *P. pinnata* was performed according to the protocol described by Kesari et al (2009). The list of chemical and reagents were used in DNA extraction process was mentioned in Table 4.1. DNA concentration was measured by obtaining the $A_{260/280}$ ratio with a spectrophotometer (Cary 100Bio; Varian). DNA quality and concentration was also checked by running 2

TH-1255_09610610

µl of DNA from each sample on a 1 % agarose gel in 1 x Tris-acetate-EDTA buffer containing 0.5 µg/ml of EtBr.

Table 4.1: Reagents and chemicals

Chemical	Concentration
Tris Base (pH 8.0)	1 M
EDTA (pH 8.0)	0.5 M
Sodium Dodecyl Sulphate (SDS)	10 %
NaCl	5 M
β-mecaptoethanol	As per required
Potassium acetate	5 M
Sodium acetate (pH 5.2)	3 M
Phenol:Chloroform	1:1 v/v
Isopropanol	-20 ^{o}C
70% ethanol, Absolute alcohol	As per required
TE buffer	10 mM Tris-HCl, pH 8; 1 mM EDTA, pH 8.0
RNase A	10 mg/ml
CTAB	1 %
DNA extraction buffer	100 mM Tris-HCl (pH 8.0), 25 mM EDTA (pH 8.0), 1.5 M NaCl, 1.25 % SDS and 5 % β-mercaptoethanol (v/v) added immediately before use.

Total RNA from different tissue types [root, leaf, stem, flower, cotyledon and different developmental stages of seeds (90-, 180-, 270-, 350-DAF (Day after flowering)] were isolated according to the protocol described in chapter-3, section 3.2.2 and recovered by ethanol and sodium acetate precipitation. The RNA quality and relative quantities were estimated by resolving 1 µg RNA on 1.2 % agarose/formaldehyde gels with a final concentration of 0.22 M HCHO (formaldehyde), followed by staining with ethidium bromide (0.5 mg/ml). Equal intensity of rRNA bands were taken as the criteria for equal quantity and uniform quality of the samples.

4.2.3 Complementary DNA cloning and sequencing

cDNA clone was screened from full length cDNA library constructed using 0.5µg of poly $(A)^{+}$ RNA extracted from early immature seeds of *Pongamia.* The partial sequence of *SAD* gene probed with ^{32}P radio isotope was used to elute full length SAD cDNA from the lawn of clones (Complete details were described in chapter-3, section 3.2.16-

TH-1255_09610610

17). The screened cDNA insert (approximately 1.6 kb in size) was gel eluted using gel extraction kit (cat. no. 28704), purified and finally precipitated with ethanol and sodium acetate method. The purified product was sequenced by the dideoxy chain termination method using an automatic sequencer (ABI 777, Macrogen sequencing service, Korea).

4.2.4 Sequence analysis

The DNA sequences (duplicates) obtained from sequencing service was used to generate contigs and compared with sequences deposited in the GenBank database using the BLAST program (Altschul et al 1997). Sequence data of corresponding cDNA from different species was analyzed using the ClustalX2.1 package (Thompson et al 1997). The physicochemical properties of the deduced protein were predicted by Protparam (http://www.expasy.ch/tools/protparam). Active sites of the protein sequence were analyzed with the PROSITE database. Protein domains were analyzed by SMART (http://smart.embl-heidelberg.de/). The sub-cellular location of the protein was predicted by the TargetP 1.1 Server (http://www.cbs.dtu.dk/services/TargetP/). Hydrophobicity analysis was performed on ProtScale (http://www.expasy.org/cgi-bin/protscale.pl) by using Kyte & Doolittle (1982), and transmembrane topology prediction was performed using TMHMM Server version 2.0 (http://www.cbs.dtu.dk/services/TMHMM). The phylogenetic tree was constructed with ClustalX2.1 based on neighbor-joining method.

4.2.5 RNA quantification, primer design and transcript analysis by RT-PCR

RNA quality and quantity was estimated by measuring the A_{260}/A_{280} ratio and absorbance at 260 nm with a spectrophotometer respectively. The integrity of total RNA extracted was characterized by resolving 2 µl RNA on 1.2 % agarose gel under denaturing conditions with (0.4 M) formaldehyde followed by staining with EtBr (0.5 µg/ml).

The oligonucleotide primers for *Pongamia SAD* gene were designed based on sequence information obtained from *Pongamia* full length *SAD* gene sequence. Table 4.2 shows the sequence of the gene specific, housekeeping primers and their approximate amplified product sizes. Glyceroldehyde-3-phospate dehydrogenase (GAPDH) gene was used as an internal control (housekeeping gene).

TH-1255_09610610

For RT-PCR reaction 1µg of total RNA isolated from root, leaf, stem, flower, cotyledon and 4 different developmental stages of seeds (90-, 180-, 270- & 350-DAF) were used as a template in first strand cDNA synthesis using an oligo dT primer (0.5 µg/ reaction) and ImProm-II^{TM} Reverse Transcription system (Promega), according to the manufacturers protocol. The products were quantified by running 2 µl cDNA on 1.2 % agarose gel and the samples were stored at 4 °C for short term storage and -20 °C for long term storage. For second strand synthesis corresponding to fatty acid gene namely *SAD*, gene specific primer was used and GAPDH primer as housekeeping gene was kept as standard to confirm the efficiency of the reaction. PCR amplification was performed in 25 µl reaction volume containing cDNA amount corresponding to 200 ng total RNA as a template, together with 200 µM of each dNTP, 2 mM $MgCl_2$, 15 µM of each oligo, 5U *Taq* DNA polymerase, using a DNA Thermal cycler (Applied Biosystems, USA). The PCR thermal-cycling parameters were as follows: an initial denaturizing step of 95 °C for 5 min, followed by 35 repeats of 95 °C for 1 min, annealing temperature for 1 min 30sec annealing time was 1 min with varied temperatures as per the melting temperature of the gene specific primer and housekeeping primer, extension at 72 °C for 2 min and final extension at 72 °C for 7 min. Amplified products were loaded in 1.2 % agarose gel containing 0.5 µg/ml of EtBr and documented by a gel documentation system (Bio Rad, USA).

Table 4.2: Primer sequences used in the experiment.

Primer	**Sequence (5'-3')**
SAD-F2	TGGACAAGGGCATGGACTGC
SAD-R2	TGTTCTTTGGCAAGTCTGGC
GAPDH-F	GCAGGAACCCTGAGGAGATC
GAPDH-R	TTCCCCCTCCAGTCCTTGCT

4.2.6 Real-time RT-PCR

For Real-time RT-PCR, First-strand cDNA was synthesized from 3µg of total RNA using Superscript II (Invitrogen, USA). The reverse transcribed cDNA samples were used for real time PCR, which was performed on an ABI Step One sequence detection system (Applied Biosystems, USA). A PpSAD cDNA fragment (136 bp) was amplified

TH-1255_09610610

with gene-specific SAD-F2 and SAD-R2 primers. *P. pinnata GAPDH* gene, amplified with the primers *GAPDH*-F and *GAPDH*-R, giving a product of 120 bp, was used as a reference for normalizing the PpSAD cDNA amounts (Table 4.2). Each PCR was performed in a 25 µl reaction mix containing 1 µl of template cDNA or the standard, 1 x SYBR Premix Ex Taq (TaKaRa, Japan) and 0.3 µM of each primer. Thermal cycling conditions were: 95 °C for 10 s; 40 cycles of 95 °C for 5 s, 60 °C for 31 s; then 95 °C for 15 s, 60 °C for 20 s and 95 °C for 15 s for the dissociation stage. After the real time PCR, the absence of unwanted by products was confirmed by automated melting curve analysis and agarose gel electrophoresis of the PCR product. The relative expression ratio of *PpSAD* gene was calculated based on real-time PCR efficiencies and the crossing point differences of each sample versus a control sample. *PpSAD* gene cDNA levels were normalized with those of *GAPDH* as a control in the same samples qualified in the same manner, and the final relative cDNA amounts of *PpSAD* gene were the means of three replicates. Statistical differences in expression between the mean values of various samples were analyzed by one-way analysis of variance (ANOVA).

4.2.7 Southern hybridization

Approximately 10 µg of total DNA was digested separately with *Eco*RI and *Hind*III (Fermentas), restriction endonuclease according to supplier's instructions. The completely digested product was separated on 0.8 % agarose gel and this was confirmed by visualizing by post staining with 0.5 mg/ml EtBr. The gel was soaked in 0.5 M NaOH, 1.5 M NaCl for 30 min and blotted on to a Hybond N^+ membrane (Amersham Biosciences, Buckinghamshire, U.K). Probes for DNA blot was generated through PCR amplification of the cDNA inserts using the primers SAD-1F and 1R (described in chapter-3 and section 3.2.16). The probes were gel eluted using Qiagen gel extraction kit (cat. no. 28704), following the manufacturer's protocol. ^{32}P-labeled 437 bp PpSAD probe was prepared using Amersham Rediprime random primer labeling kit (cat. no. RPN1633) and was purified using Probequant G-50 column (Pharmacia Biotech, India) according to the manufacturer's instructions. The purified probes were probed to the filters and the hybridization was performed in 6 x SSPE, 10 % Denhardt's solution, 0.5 % SDS, 100 µg/ml denatured Salmon Sperm DNA

TH-1255_09610610

fragments overnight at 65 ^{o}C, and the filters were washed three times in 2 x SSC, 0.5 % SDS for 30 min at the same temperature.

4.2.8. Structural characterization of PpSAD by homology modeling

Homology modeling combines computational chemistry and bioinformatics tools to calculate atomic-resolution model of the target protein from its amino acid sequence and one or more experimental three-dimensional structures of related homologous proteins, i.e. the templates. In current study homology modeling studies have been conducted for deriving the structure of PpSAD protein; basically it consists of the four following steps: (i) template selection by database search, (ii) sequence alignment between the target and template proteins, (iii) model building and refinement, and (iv) a final model validation (Liu *et al* 2011). Due to the sequential completion of steps i–iv any error propagates to the later steps decreasing the model accuracy. Alignment errors are unrecoverable, therefore both finding the proper template and making a good alignment need special attention. The secondary and 3D structures of the deduced protein were predicted by PredictProtein (http://www.predictprotein.org/) and Swiss Model respectively.

Structural analysis was done based on the homology model constructed using MODELER ModWeb server Ver. SVN.r1368M (Sali lab). Crystal structure of *R. communis* (PDB ID: 1AFR) was selected as a template for construction of the model. The initial models were refined by energy minimization using DeepView/Swiss-Pdv viewer software package (Guex & Peitsch 1997). Accuracy of the predicted model was examined by Ramchandran Plot obtained from the PROCHECK (Laskowski et al 1993) and structural features were validated using different online server programs (PDBsum, VADAR, ModFold server, QMEAN server, SAVER andPDB2PQR Server). For 3D modeling PyMOL Version 1.5 software package was used.

4.3 RESULTS AND DISCUSSION

4.3.1 Characterization of the *P. pinnata* SAD cDNA

cDNA containing the complete coding region of the *SAD* gene was screened from the full length cDNA library constructed from 90-DAF seeds of *P. pinnata*. Cloning and sequencing of a cDNA *SAD* transcript revealed that the cDNA is 1509 bp long with a

TH-1255_09610610

5'-untranslated region of 63 bp, an open reading frame (ORF) of 1182 bp and a 3'-untranslated region of 264 bp (Figure 4.1). The completed nucleotide sequence indicates that PpSAD cDNA encoded a polypeptide 393 amino acids with molecular weight of 45.04 kDa. The gene sequence was deposited in GenBank public domain and the accession number for the *PpSAD* gene was allotted to be KF192317. The alignment of deduced amino acid sequence of *Pongamia* Steroyl-ACP-desaturase to deduced amino acid sequences of other SAD also indicates the presence of two sequence motifs, E-E-N-R-H and D-E-K-R-H which are normally found in stearoyl-ACP desaturases and other oxygen activating enzymes. These amino acid sequences represent a biological motif used for the creation of reactive catalytic intermediates. Thus, the process of fatty acid desaturation may proceed via enzymatic generation of a high-valent iron-oxo species derived from the diiron cluster (compounds and intermediates in which iron is found in a formal oxidation state) (Fox et al 1993).

In addition, PpSAD shows high identity to conserved domain of ferritin-like structures, which suggests PpSAD has a ferritin with a Fe-O-Fe center (Fox et al 1994). According to Lindqvist (1996) the six amino acids of E138, E176, H179, E229, E262 and H265 in SAD peptide are the ligands of ferritin's di-iron. They are all located in the conserved domain of SAD. Sequence alignment of PpSAD with SAD from other plants revealed that PpSAD contains two conserved domains which plays crucial role in introducing the double bonds. These conserved regions suggest that PpSAD likely to have the same function as other reported SAD proteins (Tong et al 2007). Moreover the high identity of PpSAD to other plant SADs indicate that SADs have been highly conserved during evolution and further demonstrates their critical enzymatic roles in fatty acid synthesis in plants.

TH-1255_09610610

```
GGATTGTGGACCAATAACTAACATCACAAACCATGTATTGTTTTTCGGAGAGGCCAATCT    60

GAAATGGCTCTCAGACTAAACCCTTTCCCAACTCAAACATCTACCTTCTCTCTTCCTCAA   120
   M  A  L  R  L  N  P  F  P  T  Q  T  S  T  F  S  L  P  Q      19
ATGGCCAGTCTCAGATCTCCCAGGTTCCGCATGGCCTCTACCCTCCGCTCCGGTTCCAAA   180
M  A  S  L  R  S  P  R  F  R  M  A  S  T  L  R  S  G  S  K      39
GAGGTTGAAAATATTAAGAAGCCATTCACTCCTCCCAGAGAAGTGCATGTTCAAGTAACC   240
E  V  E  N  I  K  K  P  F  T  P  P  R  E  V  H  V  Q  V  T      59
CACTCTATGCCTCCTCAGAAGATTGAGATCTTTAAATCTTTAGAGGGTTGGGCTGAACAA   300
H  S  M  P  P  Q  K  I  E  I  F  K  S  L  E  G  W  A  E  Q      79
AACATCTTGACTCATCTTAAGCCTGTTGAAAAATGTTGGCAACCACAAGATTTTTTACCG   360
N  I  L  T  H  L  K  P  V  E  K  C  W  Q  P  Q  D  F  L  P      99
GATCCTTCCTCTGATGGATTTGAAGAGCAAGTGAAGGAACTGAGAGAGAGGGCAAAGGAG   420
D  P  S  S  D  G  F  E  E  Q  V  K  E  L  R  E  R  A  K  E     119
CTTCCAGATGATTACTTTGTTGTTCTGGTCGGAGACATGATCACAGAGGAAGCCCTGCCT   480
L  P  D  D  Y  F  V  V  L  V  G  D  M  I  T  E  E  A  L  P     139
ACTTACCAAACAATGTTAAATACTTTGGATGGAGTTCGTGATGAAACAGGTGCCAGCCTT   540
T  Y  Q  T  M  L  N  T  L  D  G  V  R  D  E  T  G  A  S  L     159
ACTTCCTGGGCAATTTGGACAAGGGCATGGACTGCTGAAGAAAACAGACACGGTGATCTT   600
T  S  W  A  I  W  T  R  A  W  T  A  E  E  N  R  H  G  D  L     179
CTTAACAAATATCTTTACTTGAGTGGTCGAGTCGACATGAAACAAATTGAAAAGACAATT   660
L  N  K  Y  L  Y  L  S  G  R  V  D  M  K  Q  I  E  K  T  I     199
CAGTACCTTATTGGGTCTGGAATGGATCCTCGAACCGAGAACAGCCCCTACCTTGGTTTC   720
Q  Y  L  I  G  S  G  M  D  P  R  T  E  N  S  P  Y  L  G  F     219
ATTTACACTTCATTTCAAGAGAGGGCAACCTTCATATCCCACGGAAACACAGCCAGGCTT   780
I  Y  T  S  F  Q  E  R  A  T  F  I  S  H  G  N  T  A  R  L     239
GCTAAGGAGCACGGCGATATAAAGTTGGCCCAGATCTGCGGCATGATTGCCTCAGATGAG   840
A  K  E  H  G  D  I  K  L  A  Q  I  C  G  M  I  A  S  D  E     259
AAACGCCATGAGACTGCATACACGAAGATAGTGGAAAAGCTATTTGAGATTGATCCTGAC   900
K  R  H  E  T  A  Y  T  K  I  V  E  K  L  F  E  I  D  P  D     279
GGTACAGTTATGGCGTTTGCCGACATGATGAGGAAGAAGATTTCTATGCCAGCACACCTC   960
G  T  V  M  A  F  A  D  M  M  R  K  K  I  S  M  P  A  H  L     299
ATGTATGACGGCCGTGATGACAACCTTTTTGATCATTACTCTGCCGTCGCGCAGCGTATT  1020
M  Y  D  G  R  D  D  N  L  F  D  H  Y  S  A  V  A  Q  R  I     319
GGGGTCTACACCGCCAAGGACTATGCTGACATACTTGAGTTTCTGGTGGGGAGGTGGAAG  1080
G  V  Y  T  A  K  D  Y  A  D  I  L  E  F  L  V  G  R  W  K     339
GTGGAGGACCTAACTGGACTTTCAGGCGAGGGAAGAAAGGCTCAAGAATACATTTGCGGG  1140
V  E  D  L  T  G  L  S  G  E  G  R  K  A  Q  E  Y  I  C  G     359
CTGCCGCCAAGAATCAGAAGGTTGGAGGAGAGAGCTCAAGCAAGAGCAAAGGAATCATCA  1200
L  P  P  R  I  R  R  L  E  E  R  A  Q  A  R  A  K  E  S  S     379
ACAATACCATTCAGTTGGATTCATGACAGGGAAGTACTACTCTGACCATGCTAAGAATCA  1260
T  I  P  F  S  W  I  H  D  R  E  V  L  L  *                   393
GGTTGAATAGATGAAAATTCAATTTGTCCTCTTGTTTGGGGTTTTTTGTTATATGTCACT  1320
GTAAGGTGAAGCTTCAGTTAAATGCATAGTCTGCAAGCTTAAGCATTAGGGGGCAGCTGT  1380
AGTAATAGAAATGCTATTTTTTGTTTCCCTTTTCTTTGTGGTAGTGATGTTTCGAAGTAT  1440
AAGTAAACGTTTTTCTCTGGCAATTTTGATGATAAAGAAAATTTAGTTCTATCCAAAAAA  1500
AAAAAAAAA                                                     1509
```

Figure 4.1: A complete nucleotide sequence and deduced polypeptide of SAD cDNA from *P. pinnata*. The start and stop codons are marked by arrows and also the stop codon denoted by asterisk (*). The two amino acid sequence motifs are boldly under lined. The 5' untranslated region (UTR) is 63 bp and the 3' UTR is 264 bp.

TH-1255_09610610

The A+T content and G+C content in the sequences were 56.40 % and 43.60 % respectively. Blastn and Blastp analyses, combined with a multiple sequence alignment analysis showed that *Pongamia* SAD's nucleotide and amino acid sequences are highly homologous to those of SAD from other plants (Figure 4.2). The closest nucleotide and amino acid sequences were from *G. max* (NM_001251324.1) with 98 % identity to PpSAD deduced amino acid sequence and also shows high similarity to other dicot SAD sequences e.g. *P. vulgaris* (JX853966.1), *A. hypogeae* (FJ230310.1), *S. indicum* (D42086.1), *J. curcas* (DQ084491.1), *R. cummunis* (XM_002531843.1), *C. arietinum* (XM_004505093.1) and *B. napus* (AY642537.1). Despite high degree of homology in the coding region, the nucleotide sequence of *PpSAD* gene showed only 81, 80, 58, 49, 46, 48 % similarity at the 3'-UTRs to GmSAD, PvSAD, SiSAD, AhSAD, JcSAD and RcSAD respectively (Table 4.3).

Table 4.3: Similarity matrix indicating percentage of PpSAD amino acid and Nucleic acid sequences (open reading frame and untranslated sequences) conserved between species: *P. pinnata* PpSAD (KF192317), *G. max* (NM_001251324.1), *P. vulgaris* (JX853966.1), *A. hypogeae* (FJ230310.1), *S. indicum* (D42086.1), *J. curcas* (DQ084491.1), *R. cummunis* (XM_002531843.1), *C. arietinum* (XM_004505093.1) and *B. napus* (AY642537.1). Sequences were aligned using ClustalX2.1 (Thompson *et al* 1997) with default parameters. The untranslated region is composed of 63 nucleotides upstream of the start codon and 264 nucleotides of downstream of the stop codon.

Species	Protein	Nucleic acid ORF	3' UTR
GmSAD	97	97	81
PvSAD	95	93	80
SiSAD	86	78	58
AhSAD	90	86	49
JcSAD	85	81	46
RcSAD	84	80	48

TH-1255_09610610

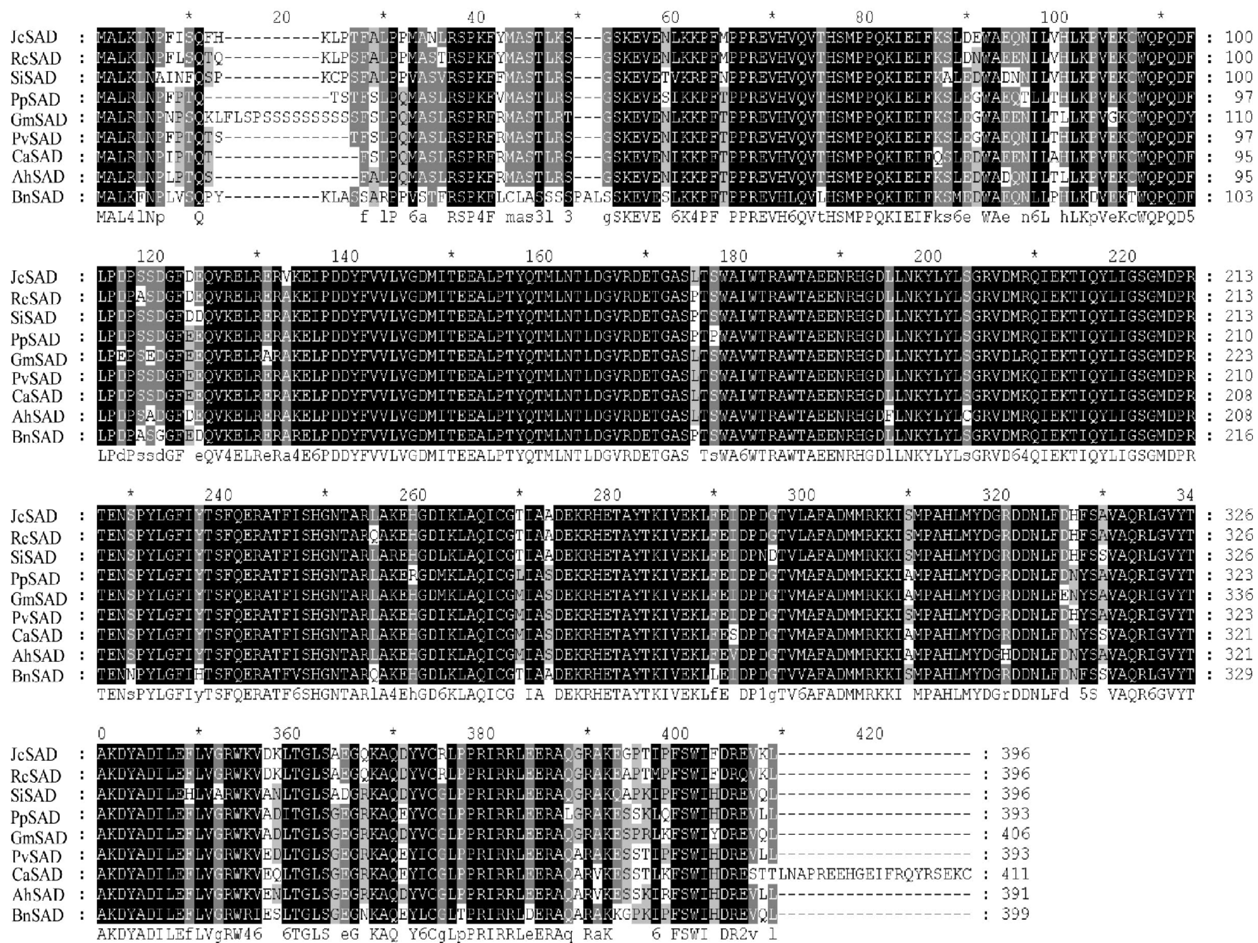


Figure 4.2: Amino acid sequence alignment of PpSAD from *P. pinnata* with SAD from various species. The alignments were performed using ClustalX2.1. Shading was performed by Multiple Sequence Alignment Editor and Shading Utility (Genedoc, version 2.6.002) in conservation mode. Black shading indicates residues that are 100 % conserved and are identical in all sequences; grey shading indicates residues that are identical in most of sequences or functionally identical in all of them. The accession numbers and sources of the above proteins: *P. pinnata* PpSAD (KF192317), *J. curcas* JcSAD (DQ084491.1), *R. cummunis* RcSAD (XM_002531843.1), *S. indicum* SiSAD (D42086.1), *G. max* GmSAD (NM_001251324.1), *P. vulgaris* PvSAD (JX853966.1), *C. arietinum* CaSAD (XM_004505093.1), *A. hypogeae* AhSAD (FJ230310.1) and *B. napus* BnSAD (AY642537.1).

TH-1255_09610610

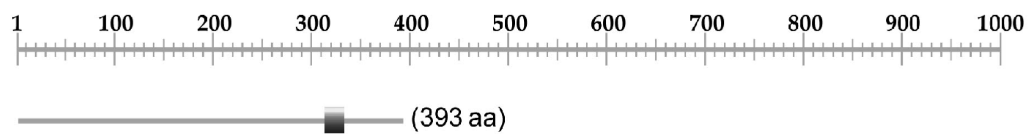
Active site analysis of PpSAD's 393 amino acids with the PROSITE database showed that the sequence contains a fatty acid desaturase family-2 signature (PS00574, SAvaqRIgvytakDYadILE) situated at amino acid 313 to 332 (Figure 4.3A).

Two kinds of protein conserved domains are found in the deduced amino acid sequence of PpSAD and the roles are as follows: Domain-1 (65-371 amino acids) is Acyl ACP desaturase, ferritin-like diiron-binding domain; is a mu-oxo-bridged diiron-carboxylate enzyme, which belongs to a broad superfamily of ferritin-like proteins and catalyzes the NADPH and O_2 dependent formation of a cis-double bond in acyl-ACPs. Acyl-ACP desaturases are found in higher plants and a few bacterial species (*Mycobacterium tuberculosis*, *M. leprae, M. avium and Streptomyces avermitilis, S. coelicolor*). In plants, Acyl-ACP desaturase is a plastid localized, covalently ACP linked, soluble desaturase that introduces the first double bond into saturated fatty acids, resulting in the corresponding monounsaturated fatty acid. Members of this class of soluble desaturases are specific for a particular substrate chain length and introduce the double bond between specific carbon atoms. Domain-2 (63-391 amino acids) which is fatty acid desaturase involved at the end of the desaturation process, and converts saturated fatty acid to unsaturated fatty acids (Figure 4.3B). Hydrophobicity plot analysis of deduced amino acid sequences of PpSAD and other SADs showed that these enzymes are likely to be water soluble proteins (Figure 4.4). Sub-cellular localization of the deduced PpSAD amino acid sequence using the PSORT tool revealed that it was localized in chloroplast.

Physico chemical property predictions of the deduced amino acid sequence indicated that the chemical formula of PpSAD is $C_{2006}H_{3139}N_{549}O_{597}S_{17,}$ its molecular weight is 45.04 kDa, and its theoretical PI is 5.88. PpSAD's half-life at *in vitro* conditions is predicted to be approximately 30 h and its instability index (II) was computed to be 42.74 which are above 40, thus classifying *Pongamia* SAD as an unstable protein. Amino acid composition analysis of this protein showed that Leu (8.9 %), Glu (8.4 %), Ala (6.9 %) Asp (6.1 %), Arg (6.9 %), Thr (6.6 %), Lys (5.9 %) and Ser (6.4 %) are the most abundant residues. PpSAD contains 57 negatively charged residues (Asp + Glu) and 50 positively charged residues (Arg + Lys). Its grand average of hydropathicity (GRAVY) was estimated to be -0.499, suggesting that this protein is hydrophilic.

TH-1255_09610610

(A)



(B)

(i)

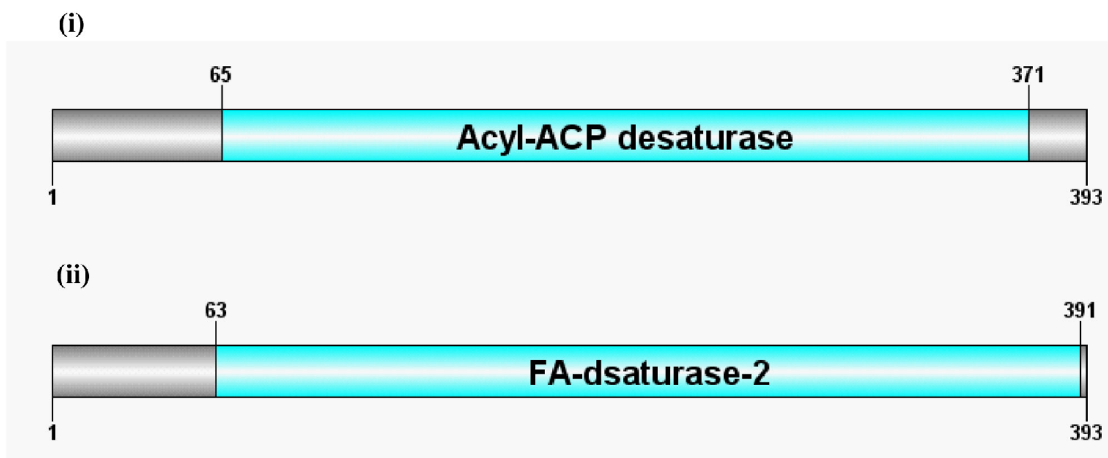


Figure 4.3: Protein structural analysis: (**A**) Active site map of PpSAD; (**B**) Protein conserved domain analysis of PpSAD. (**i**) Acyl-ACP-Desaturase (65-371 amino acids). (**ii**) Fatty acid desaturase (63-391).

TH-1255_09610610

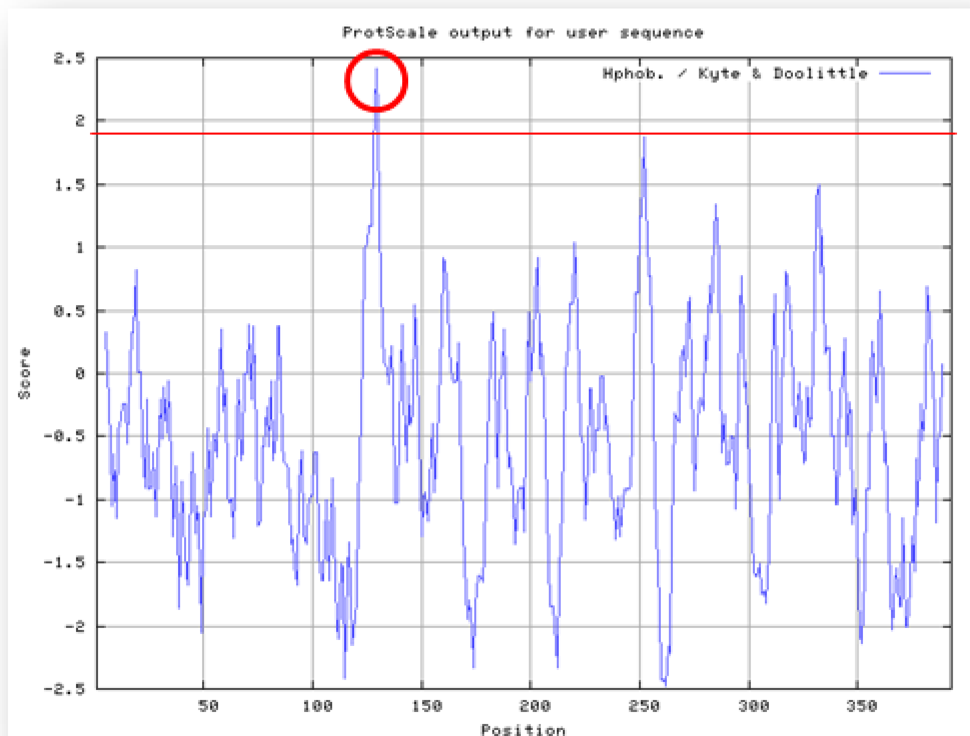


Figure 4.4: Hydropathy plot of the *PpSAD* gene product. Average hydropathicity values were determined for spans of 20 residues using the method of Kyte and Doolittle (1982). Numbers on the x-axis are amino acid residues and the numbers on the y-axis are hydrophobic means.

TH-1255_09610610

A phylogenetic tree based on deduced amino acid sequences of PpSAD and other SADs from different flora was constructed to investigate the evolutionary relationships among various SADs. We selected total 8 steroyl-ACP-desaturase amino acid sequences from NCBI gene bank database and included PpSAD for a phylogenetic analysis. The results revealed that the phylogenetic tree grouped the SAD into two different clades. Analysis of SAD shows that the evolutionary relationship among SAD from higher plants is close (Clade-I), but SAD from *B. napus* (AY642537.1) was distinct (Clade-II) from other legume plants. The clade-I was again divided into two clusters in which one cluster shows SAD from other higher plants [*J. curcas* (DQ084491.1), *R. communis* (XM_002531843.1), and *S. indicum* (D42086.1)] which does not belong to legume family but are potential biofuel plants. The second cluster from clade-I shows only plants with fabaceae/leguminosae family [*G. max* (NM_001251324.1), *P. vulgaris* (JX853966.1), *C. arietinum* (XM_004505093.1), *A. hypgeae* (FJ230310.1), *P. pinnata* (KF192317)] in which PpSAD is also part of this cluster (Figure 4.5).

4.3.2 RNA expression studies

The expression pattern of the *PpSAD* gene in different tissues may provide vital information regarding the key role in fatty acid biosynthesis. In the current study based on sequence information available from full length *SAD* gene sequence, oligonucleotide primers were designed from the coding region. To determine the steady state level of *PpSAD* gene expression in different tissues of *Pongamia*, RT-PCR analysis were carried out using RNA samples extracted from various tissues including root, leaf, stem, flower, cotyledon, and four different stages of seeds (90-350DAF). SAD is a key enzyme in fatty acid biosynthetic metabolism of higher plants which is present in plastid stroma, and the main function of the SAD is it catalyzes desaturation of stearoyl-ACP, which introduces a double bond into the fatty acid chain between C9 and C10 to form oleoyl-ACP (Tong et al 2007).

TH-1255_09610610

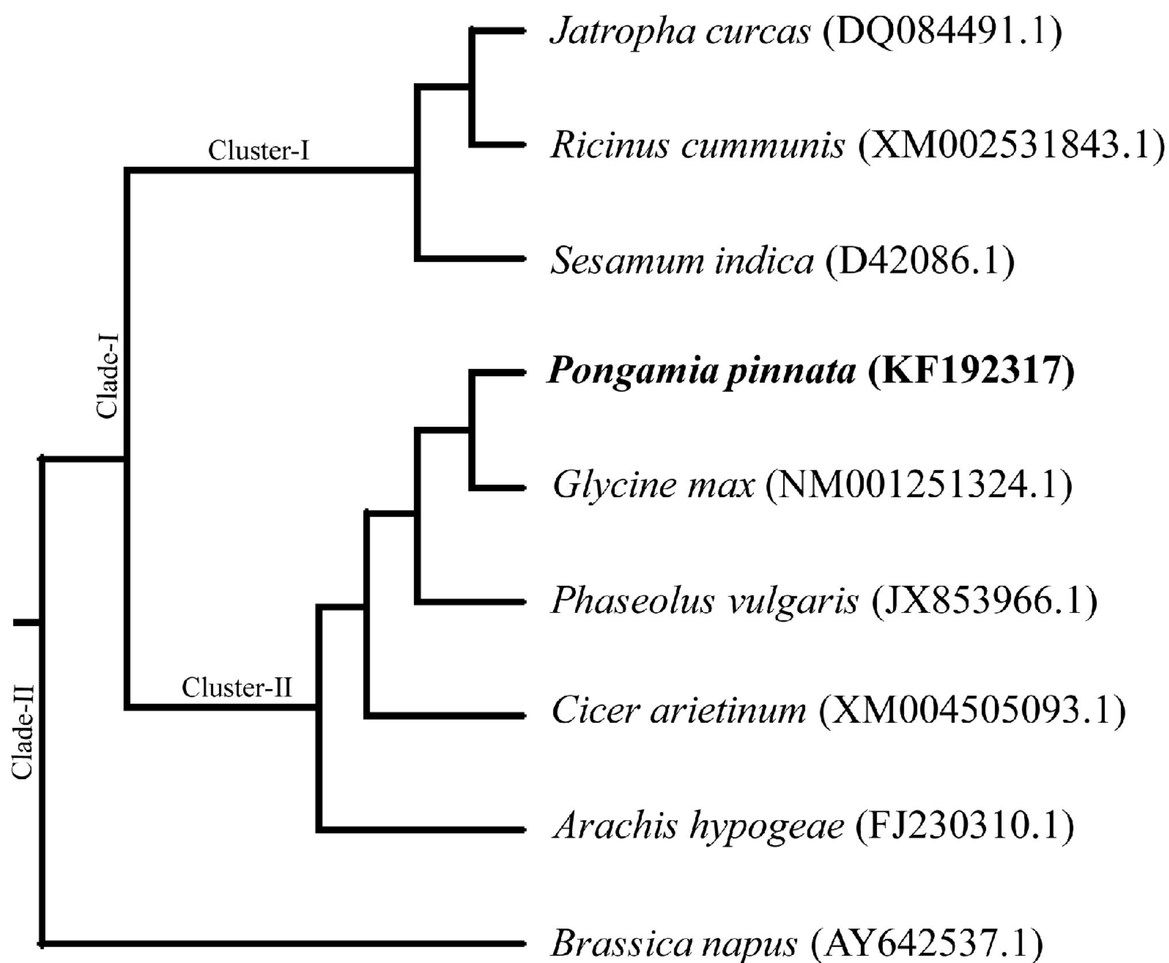


Figure 4.5: Phylogenetic analysis of PpSAD proteins. The phylogenetic tree was constructed with ClustalX2.1 based on neighbor-joining method.

TH-1255_09610610

The desaturation of stearoyl-ACP is the last step of de novo synthesis of fatty acids in plants, but its desaturated product (oleoyl-ACP) is the substrate forming many poly-unsaturated fatty acids, such as linoleic acid and linolenic acid, because double bonds are further introduced into the monounsaturated acyl-lipids and saturated acyl groups do not generally serve as substrates for desaturation at the C6, C12, or C15 position in the carbon chain. Therefore, SAD plays a key role in determining the ratio of saturated to unsaturated fatty acids in plants (Thompson et al 1991). SAD is closely related with many life activities of plants based on catalyzing desaturation of stearoyl-ACP. RT PCR analysis revealed that the *SAD* gene displays tissue-specific expression patterns in *P. pinnata* (Figure 4.6A&B). The expression level in seeds and in leaf was significantly higher than in other tissues. Low level of expression was detected in root, stem and flower (Figure 4.7A). Similar to *Jatropha* SAD, *Elaies* SAD, *PpSAD* is constitutively expressed in every tissue examined, including non-photosynthetic tissues, such as seed and root, suggesting that it may be active in all cells and tissues to produce poly unsaturated fatty acids for normal membrane functions (Okule et al 1994, Wallis & Browse 2002), besides its presumed role in synthesizing PUFA in seed oil. In addition, comparative analysis among seeds in different development stages showed that the expression levels of *SAD* gene differed markedly among different stages of seeds. The expression levels in seeds increased along with seed development, it was lowest in 90-DAF and was highest in 350-DAF (Figure 4.7B). Thus, quantitative analysis showed that the expression levels of *PpSAD* gene differed markedly among tissues but shows constitutive expression levels. It was reported that rape *SAD* is mainly expressed in developing seeds, but that its expression level in flowers is also relatively high (Slocombe et al 1994). This comparison of *PpSAD* gene expression strategy in the different tissue levels revealed that the pattern of *PpSAD* gene expression is consistent in all tissues and increases with seed developmental stages. The higher level in seeds indicates that *SAD* gene may be involved with growth and development, while the expression in leaves may be related to SAD's chloroplast location. These results serve as a foundation for further studies of the mechanisms regulating *SAD* gene expression and may eventually lead to the development of higher quality *Pongamia* genotypes.

TH-1255_09610610

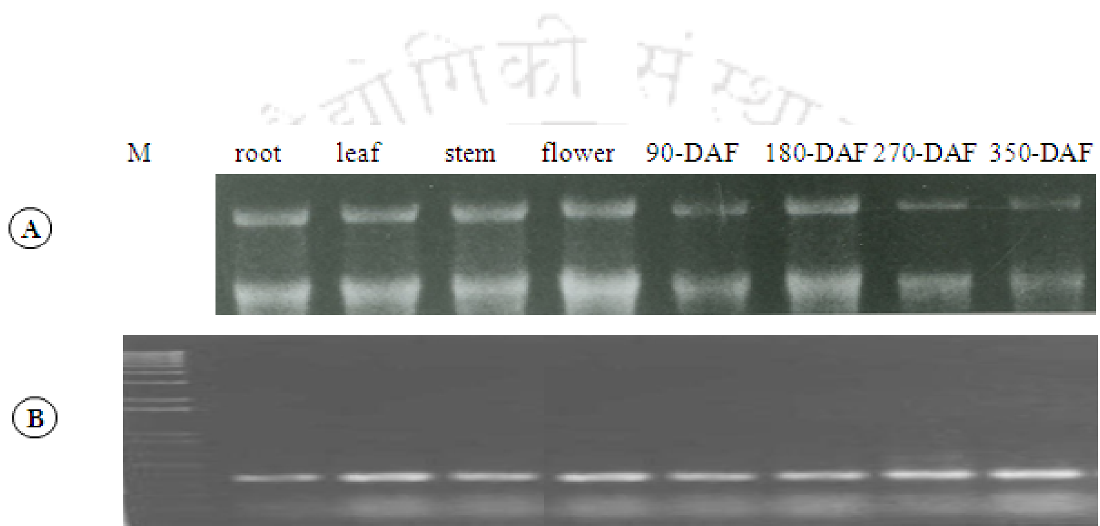


Figure 4.6: RNA expression studies. (**A**) Total RNA extracted from various tissues of *P. pinnata* (root, leaf, stem, flower, 90DAF, 180DAF, 270DAF, 350DAF) (**B**) Semi quantitative analysis of steroyl-acp-desaturase gene showing expression in all tissues.

TH-1255_09610610

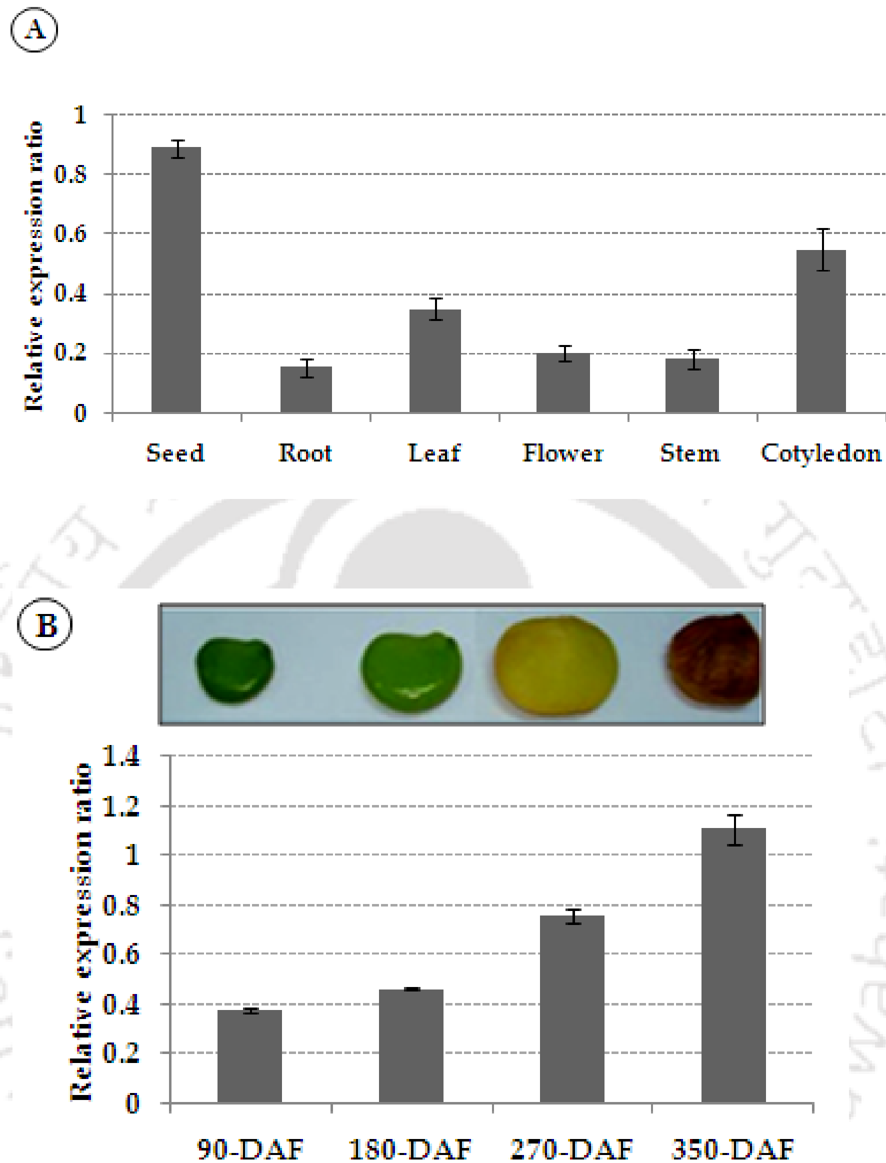


Figure 4.7: Quantitative analysis of *PpSAD* gene expression in *P. pinnata*. (**A**) Expression of Pp*SAD* in different tissues analyzed by real-time PCR. Relative expression ratio of each sample is compared with the seed sample. (**B**) Temporal expression of *PpSAD* gene. Real time PCR was performed with cDNA isolated from seeds of early immature (90-DAF), early mature (180-DAF), late immature (270-DAF) and late mature (350-DAF). Relative expression ratio of each sample is compared with the sample at 350-DAF. GAPDH from *Pongamia* was used as internal control. The final relative cDNA amounts of PpSAD are means of three replicates. The relative expression ratios of *PpSAD* gene were significantly different at $p<0.05$.

TH-1255_09610610

4.3.3 Southern analysis

To know the gene copy number of *PpSAD* gene within the *Pongamia* genome, Southern hybridization was performed. *P. pinnata* genomic DNA was digested with restriction enzymes *Eco*RI and *Hin*dIII separately; neither enzyme has a cutting position inside the *PpSAD* gene. About 2-4 hybridizing bands were observed in each lane which suggests that other closely related isoenzyme genes encoding SAD families may occur in the whole genome. However hybridization at high stringency revealed several hybridizing bands, demonstrating a very limited potential number of genes. The results indicated that there were at least two copies (multiple copies) of the gene, which is consistent with the diploid nature of the *Pongamia* genome and corresponds to the SAD desaturase family (Shah et al 2000; Luo et al 2006) (Figure 4.8).

4.3.4. Structural characterization of PpSAD by homology modeling

The homology model of PpSAD was obtained through Modeller online server; ModWeb (Eswar et al 2003; Sali & Blundell 1993) using 2.40 Å resolution structure of steroyl-ACP desaturase (1AFR) from *R. communis* with 91 % sequence identity. The residues from 49 to 393 positions of PpSAD were aligned with template sequence (1AFR) (Figure 4.9). The pair-wise sequence alignment between the model and template showed that most of the secondary structures were evolutionarily conserved with minimum gaps or insertions. As depicted in Figure 4.10, the core secondary structural elements of steroyl-ACP desaturase was organized into a mixed α/β structure consisting of 2 β-strands and 19 α-helices in PpSAD connected by several loops. Secondary structure prediction by PredictProtein indicated that this protein is composed of 65.2 % α-helix, 4.3 % extended strand and 28.7 % random coil. Program Swiss-Model and Swiss-Pdb Viewer (Guex and Peitsch 1997) were used to analyze PpSAD space configuration, comparative ribbon structures of target and template; super imposition of both proteins (Figure 4.11A,B & 2.12).

TH-1255_09610610

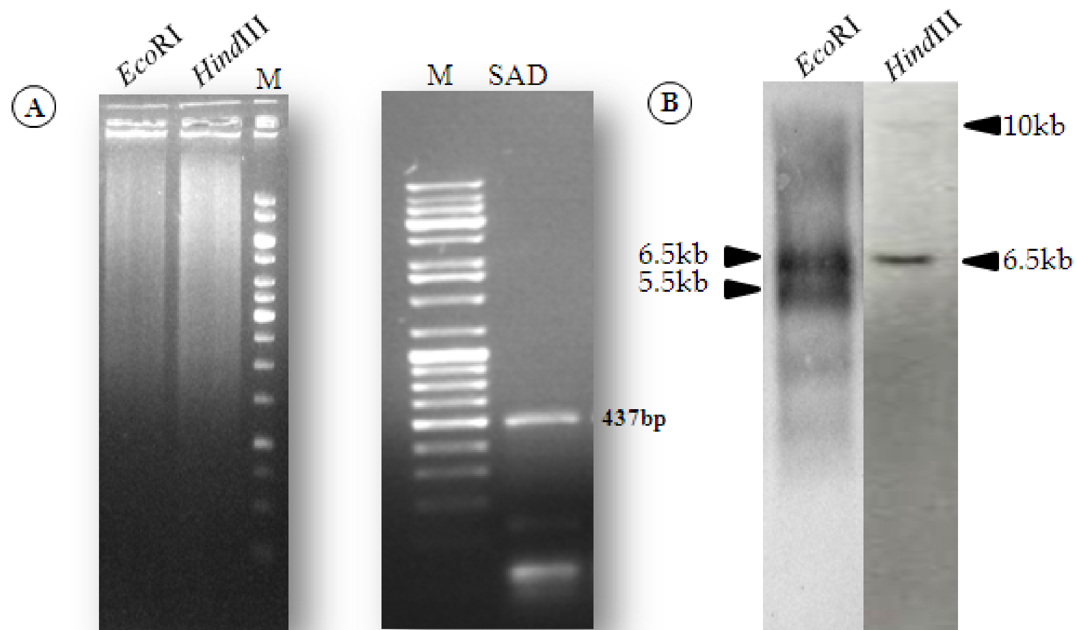


Figure 4.8: Southern blotting analysis of gene *PpSAD* from *P. pinnata*. (**A**) Approximately 10 µg of genomic DNA was digested with *Eco*RI, *Hind*III, separated in a 0.8 % agarose gel. (**B**) Transferred to a membrane followed by developed with ^{32}P-labeled 437 bp PCR fragment generated from the *PpSAD* coding sequence.

TH-1255_09610610

Figure 4.9: Sequence alignment between PpSAD (KF192317) and template 1AFR (*R. communis*).

TH-1255_09610610

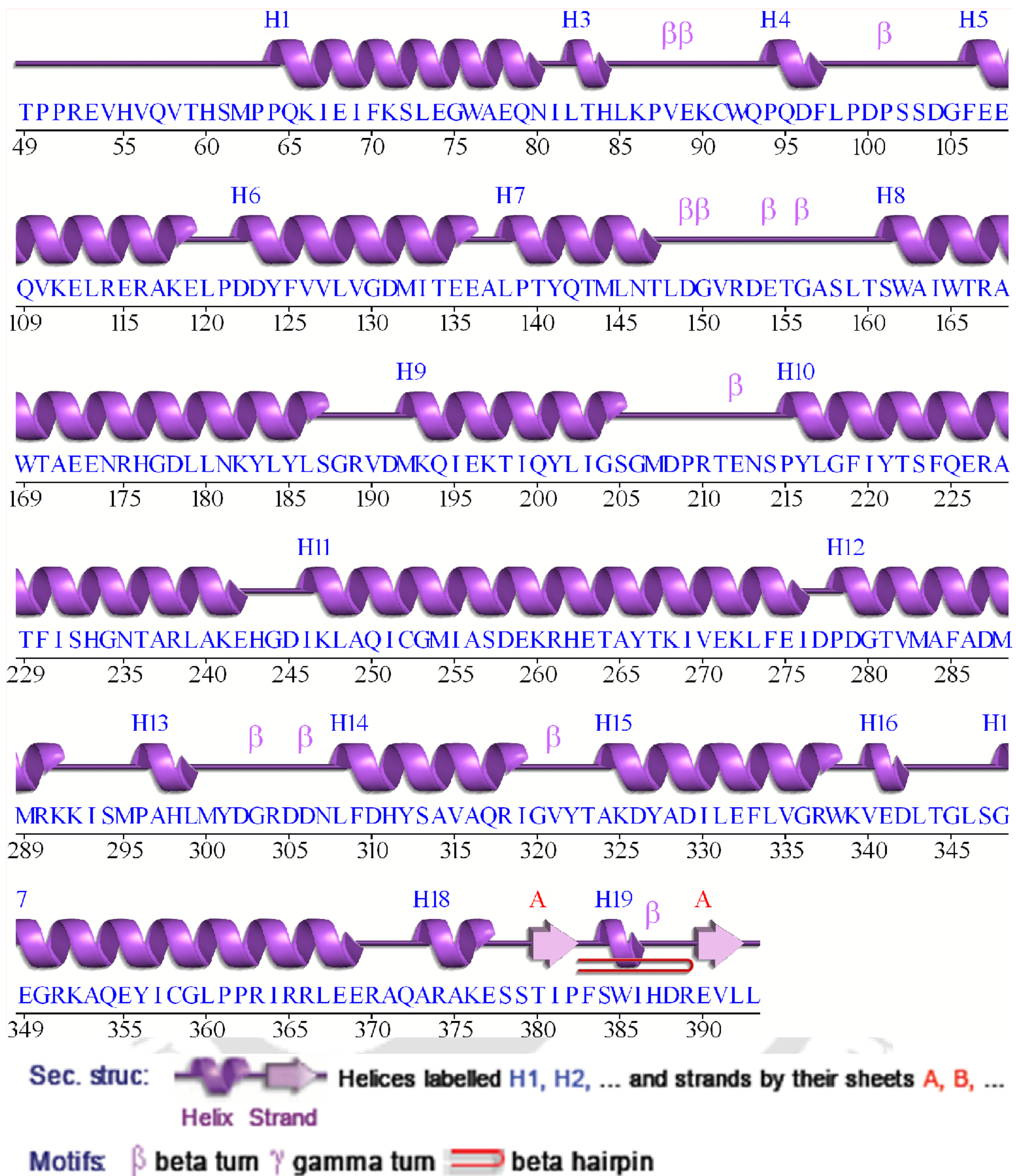


Figure 4.10: Secondary structure of the PpSAD predicted by homology modeling.

TH-1255_09610610

Figure 4.11: (**A**) Three-dimensional model of PpSAD deduced by Swiss-Model and Swiss-Pdb-viewer. (**B**) Superimposition of target-PpSAD (cyan) and templat-1AFR (Green) structure.

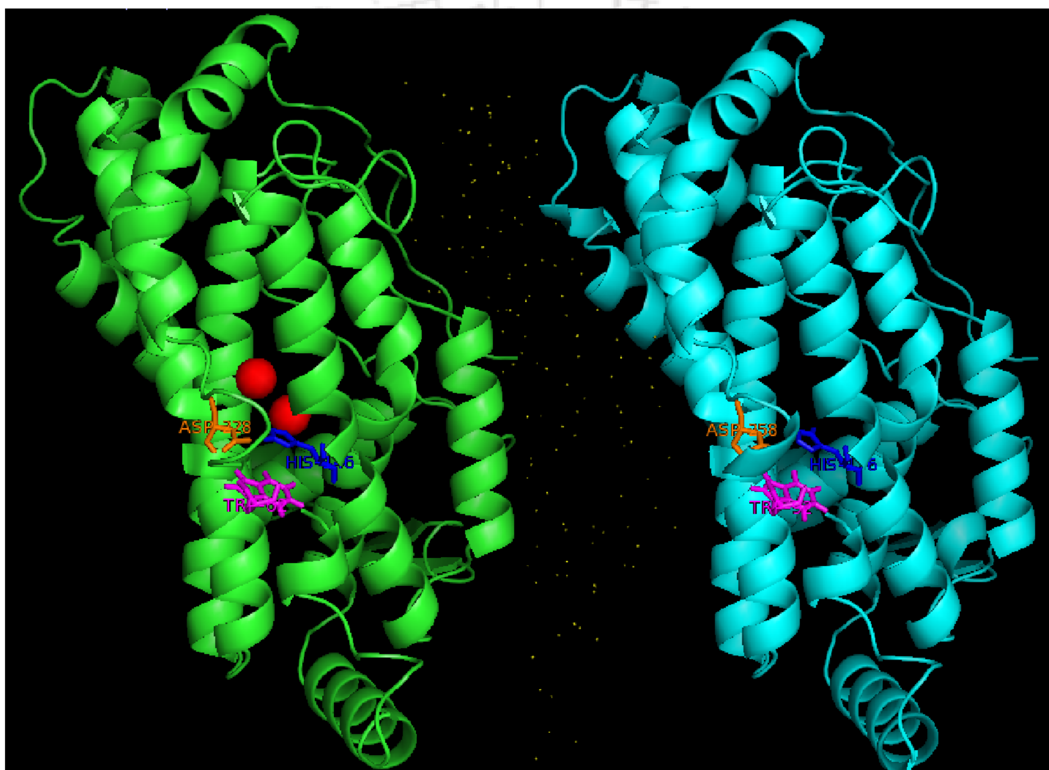


Figure 4.12: Ribbon diagram of the PpSAD (cyan) predicted by homology modeling by taking 1AFR (green) as template; three predicted iron binding sites were highlighted (Asp, Trp & His).

TH-1255_09610610

4.3.5. Structural assessment of the 3D model

The superimposition of PpSAD on the template (1AFR) structure showed a RMSD of 0.796 Å, signifying the overall tertiary structure of the model bears close resemblance to its template. Validation of the PpSAD model was carried out using different structure validation tools as mentioned in section 4.2.8. The calculated Ramachandran plot using PROCHECK compared well with that of the template (1AFR) exhibiting 94.4 % residues in core, 5.2 % in additional allowed, 0.0 % in generously allowed, and 0.3 % within disallowed regions. Altogether, the main chain and side chain parameters as evaluated by PROCHECK and Ramachandran plot (Figure 4.13) and the high scoring values obtained using various other structure validation tools (Table 4.4) confirm the overall accuracy of the model. In conclusion, it was evident from the structural data that PpSAD is structurally very similar with 1AFR (*R. communis*).

Table 4.4: Quality scores of modeled PpSAD protein using different quality control programs.

Parameter	Results	Significance	Tools
Ramachandran Plot	94.4 % core, 5.2 % allow, 0.0 % generous, 0.3 % disallowed	Above 90 %-good model	PROCHECK
G-factor	0.28	Values below -0.5* - unusual	PROCHECK
VERIFY3D	91.62	Percentage of residues with average 3D–1D score >0.2.	SAVES
ERRAT	96.726	Percentage of the protein for which the calculated error value falls below the 95% rejection limit	SAVES
Q score	0.56	QMEAN score of the whole model reflecting the predicted model reliability ranging from 0 to 1.	QMEAN
Z-score	-2.43	-(2.00-2.50)	QMEAN
Mean fractional volume	48122.6	Moderate range	VADAR

TH-1255_09610610

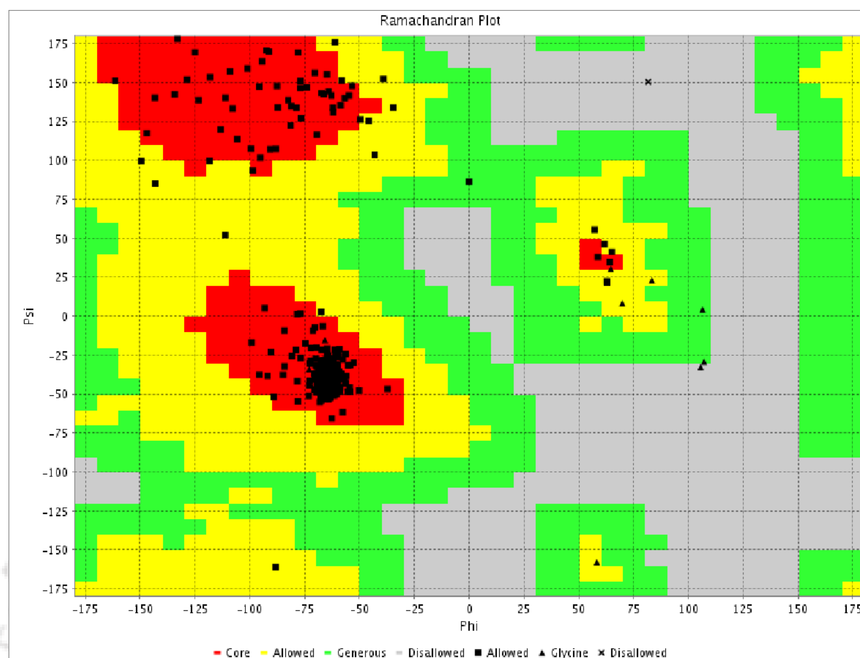


Figure 4.13: Ramachandran plot for the PpSAD model generated using PROCHEK server. Color and symbol represents Core region (Red), Allowed (Yellow), Generous (Green), Disallowed (Grey), Glycine residues (Triangle). The percentage of amino acids in each region is listed in Table 4.4.

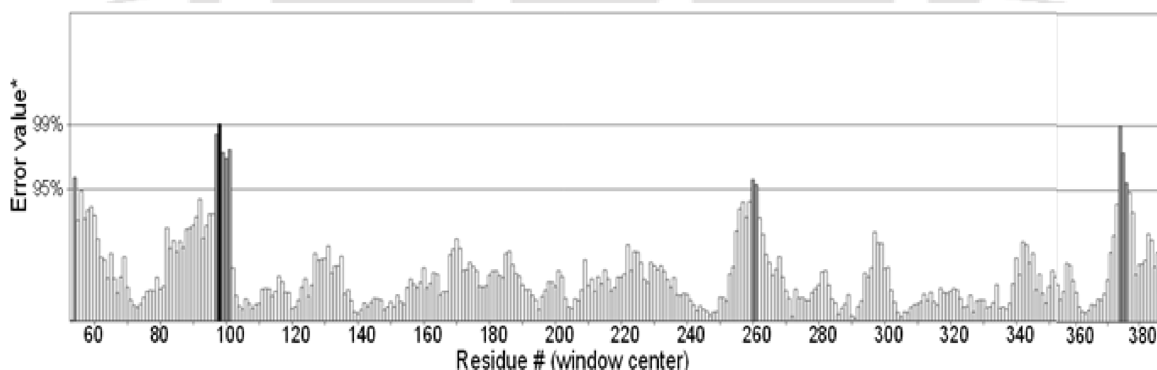


Figure 4.14: 3D Profile of PpSAD using ERRAT server. [*On the error axis, two lines are drawn to indicate the confidence with which it is possible to reject regions that exceed that error value. **Expressed as the percentage of the protein for which the calculated error value falls below the 95 % rejection limit. Good high resolution structures generally produce values around 95 % or higher. For lower resolutions (2.5 to 3 Å) the average overall quality factor is around 91 %].

TH-1255_09610610

ERRAT is a program for verifying protein structures determined by crystallography. Error values are plotted as a function of the position of a sliding 9-residue window, which works by analyzing the statistics of non bonded interactions between different atom types in the reported structure (compared to a database of reliable high-resolution structures) and a score of greater than 50 is normally acceptable. The figure (4.14) shows a plot of a final model. Regions of the structure that can be rejected at the 95 % confidence level are grey; 5 % of a good protein structure is expected to have an error value above this level. Regions that can be rejected at the 99 % level are shown in black. For our model (PpSAD), the ERRAT score is 92.441 which are well within normal range for a high quality model; the ERRAT score for the template is 96.726 (Overall quality factor**).

4.4 CONCLUSION

The well studied pathway of fatty acid biosynthesis and characterization of genes involved in the fatty acid biosynthesis have greatly facilitated the genetic modification of higher plants for oil and fatty acid production, e.g. to increase the total oil content (Lardizabal et al 2008), to alter the composition of fatty acids (Graef et al 2009), or to produce new fatty acids for nutritional improvement (Cheng et al 2010). However, it is clear that better understanding of regulatory processes of fatty acid biosynthesis and lipid metabolism and availability of more sophisticated genetic manipulation tools might be necessary for the successful engineering of potential biofuel crops for enhanced oil production. Current chapter deals with the molecular cloning and characterization of full-length cDNA of *P. pinnata* steroyl-acyl carrier protein desaturase. The deduced amino acid sequence showed a 98 % homology to the *G. max* SAD and more than 84 % to corresponding proteins from other species, including *J. curcas* and *R. communis*. The structural validation of PpSAD reveals 91 % homology with template protein (1AFR) and also validated by using various structural validation tools. The expression patterns of *PpSAD* gene in different tissues were dramatically varied and increased during seed growth and development. These results serve as a foundation for further studies of mechanisms regulating *SAD* gene expression and may eventually lead to the development of higher quality *Pongamia* genotypes.

TH-1255_09610610

Chapter 5

Functional Characterization of Fatty Acid Desaturase (FAD2) Genes

The chapter unveils the functional characterization of two fatty acid desaturase genes that are screened from the cDNA library including transcriptional analysis and gene copy number in *Pongamia* genome.

TH-1255_09610610

FUNCTIONAL CHARACTERIZATION OF FATTY ACID DESATURASE (*FAD2*) GENES

5.1 INTRODUCTION

The major structural components in the plant membrane phospholipids and tryacylglycerol (TAG) storage lipids are fatty acids. Plant oil biosynthesis, which begins in the plastid, is catalyzed by two key enzyme; acetyl-CoA caboxylase and fatty acid synthase. The final products of these two enzyme systems are 16:0 acyl carrier protein (ACP) and 18:0 ACP respectively. Once the formation of C18 fatty acid products takes place in the plastid, it is exported to the cytosol for desaturation and glycerolipid assembly (Ohlrogge 1997). Fatty acid desaturation is a biochemical process to desaturate oleic acid to linoleic acid, which catalyzes addition of single double bond to Δ12 position of oleic acid. The process of desaturation mainly takes place in both the plastidial and the endoplasmic reticulum (ER) membrane (Ohlrogge & Browse 1995). The first desaturation step for fatty acids is catalyzed by a plastidial steroyl-ACP desaturase, which introduces a double bond at 9th position of carbon and forms 18:1-ACP. Similarly the second double bond will be introduced at carbon 12 by the action of endoplasmic reticulum oleate desaturase (FAD2) to form 18:2-ACP (Mekhedov et al 2000; Okuley et al 1994). According to the Stoutjesdijk et al (2002); Hu et al (2006), an endoplasmic reticulum oleate desaturase (FAD2) is the key enzyme that controls the relative content of oleic acid, linoleic acid and linolenic acid. However, regulation of fatty acid composition is an intricate biochemical process in which many still unknown genes are involved, and exploring these unknown genes are complicated yet exciting to researchers (Xiao et al 2013).

The genes for endoplasmic reticulum (ER) and plastid derived Δ12 fatty acid desaturases have been characterized from some plant species. Several various microsomal desaturases (*FAD2, FAD6*) genes may exist, depending on the type of plant species, such as there is only one *FAD2* gene present in *Arabidopsis* (Falcone et al 1994; Okuley 1994), and two different *FAD2* genes in olive and soybean have been identified encoding both tissue (seed) specific and constitutive microsomal oleate

TH-1255_09610610

desaturases (Hernandez et al 2005; Lingyong et al 2007). More than two different microsomal oleate desaturases genes have been identified and reported in both sunflower and cotton, in which one gene shows seed specific expression and other two genes expressed in all tissue types (Liu et al 1999; Martinez-Rivas et al 2001; Pirtle et al 2001). According to the Heppard et al (1996) the gene which shows seed specific expression pattern plays a predominant role in determining the poly unsaturated fatty acid (PUFA) content of the seed storage oil.

Pongamia pinnata L. belonging to the family fabaceae/leguminoceae, is a monotypic genus and an important source of biofuel. Pongam oil, the major component of its seeds, is a complex mixture of fatty acids; palmitic (10.8 %), stearic (8.7 %), oleic (46.0 %), linoleic (27.1 %) and linolenic (6.3 %) known for its biodiesel aspects (Kesari et al 2008, 2009, 2012). Although progress has been made in the elucidation of the regulatory mechanism of endoplasmic reticulum derived Δ12 desaturase gene expression in numerous plant species, as per our knowledge there is no information on the gene cloning and characterization in tree legume.

Current chapter is therefore carried out to study the functional characterization of Δ12 desaturase; we isolated two microsomal oleate desaturases (*FAD2*-1, *FAD2*-2) genes in *Pongamia*. The expression pattern of the two genes was also investigated in different tissues and during the various seed developmental stages. The copy number determination of *FAD2* gene in *Pongamia* genome was also reported.

5.2 MATERIAL AND METHODS

5.2.1 Plant materials

P. pinnata leaves, roots and seeds collected from the sila forest, North Guwahati, 26 °14'6" N, 91 °41'28" E, Assam were used for RNA extraction. Developing seeds were harvested at different time interval corresponding to different developmental stages of the *Pongamia* (90-DAF, 180-DAF, 270-DAF and 350-DAF). All collected samples were frozen immediately in liquid nitrogen and stored at -80 °C for further experimental purpose.

TH-1255_09610610

5.2.2 RNA isolation and quantification

Total RNA was extracted from leaves, roots and various developmental stages of seeds using the CTAB-GITC method (details are described in chapter-3, section 3.2.2-3.2.3). RNA was quantified and evaluated for purity by UV spectroscopy and agarose gel electrophoresis. Prior to reverse transcription, RNA samples were treated with DNase I (TaKaRa, Japan) according to the manufacturer instructions.

5.2.3 RNA transcript analysis and expression studies

Total RNA was extracted from different tissues of *Pongamia* starting from early immature to late mature seed developmental stages (90-, 180-, 270-, 350-DAF), leaf and root. DNase I (Fermentas, USA) treatment of RNA samples were given so as to get RNA transcripts suitable for PCR amplification. The quantitative analysis of RNA was performed to evaluate the purity of sample by UV spectroscopy and agarose gel electrophoresis (described in chapter-4, section 4.2.2). Approximately 1 µg of total RNA was taken from each tissue as a template and single stranded cDNA (ss-cDNA) was synthesized by using oligo-dT primers/random hexamers of ImProm-II™ Reverse Transcription System (Promega, Madison, WI, USA). The reaction was set up at 50 °C for 30 min to synthesize ss-cDNA, followed by heating the reaction at 72 °C to deactivate reverse transcriptase enzyme. To check the semi quantitative expression pattern of fatty acid desaturases (FAD2-1 & FAD2-2) equal quantity of ss-cDNA was taken as a template for each tissue. The PCR cycle conditions were as follows: Initial denaturation at 95 °C for 5 min, 30 cycles of denaturation at 95 °C for 30 sec, annealing (Tm varies according to primers) for 30 sec, and extension at 72 °C for 1 min. Final extension was kept at 72 °C for 10 min. The primers for GAPDH (internal control) were used to check the efficiency of the PCR reactions and the PCR products were run on 1.2 % agarose gel to check the amplification of the genes.

To confirm the expression pattern of *PpFAD2*-1 & *PpFAD2*-2 genes, real time PCR was performed using Step one realtime PCR machine (Applied Biosystems). PCR reaction was performed in a 20 µl reaction mixture containing around 50 ng first strand cDNA, SYBR® Green PCR Master Mix (Applied Biosystems, Life technologies, India) and 2.5 mM of each primer (forward & reverse). Template denaturation was done for 2 min at 50 °C and 95 °C for 10 min, followed by 40 cycles of denaturation at 95 °C for

TH-1255_09610610

15 s, annealing at 60 °C for 1 min. Fluorescence signals obtained were collected at each polymerization step. The fluorescence signals obtained from the primer artifacts were eliminated by melting curve analysis. The reaction was performed in triplicates and the data are presented as means ± SE of three reactions. The *GAPDH* gene was used as an internal control and the primer sequences designed for each gene are listed in Table 5.1.

Table 5.1: Primer sequences used in the experiment.

Primer	Sequence (5'-3')
PpFAD2-1F	GCGTTTAGCGATTATCAG
PpFAD2-1R	GCTTTTCGGTTTCGGCAC
PpFAD2-2F	GGCCTGATTCTGCATAGC
PpFAD2-2R	ATACCAGCGAATGCAGCT
GAPDH-F	GCAGGAACCCTGAGGAGATC
GAPDH-R	TTCCCCCTCCAGTCCTTGCT

5.2.4 DNA extraction and quantification

To extract the large amount of DNA from healthy *Pongamia* seeds, protocol described by Kesari et al (2009) was followed and detailed methodology for DNA extraction and quantification mentioned in chapter-4, section 4.2.2.

5.2.5 Blot hybridization

To perform the Southern blot hybridization approximately 10 µg of total DNA extracted from *Pongamia* seeds were digested separately with *Eco*RI and *Hind*III restriction endonucleases according to supplier's instructions (Fermentas). The complete digested product was separated on 0.8 % agarose gel and was confirmed by visualizing by post staining with 0.5 mg/ml EtBr. The gel with DNA fragments (after digestion) was soaked in 0.5 M NaOH, 1.5 M NaCl for 30 min and blotted on to a Hybond N^+ membrane (Amersham Biosciences, Buckinghamshire, U.K). Similarly total RNA (~10 µg) extracted from leafs, roots and various seed stages of *Pongamia* tissues were electrophoresed and blotted onto Hybond N^+ membranes for northern analysis. Before loading on 1.1 % agarose/formaldehyde gels, RNA samples were heat denatured at 65 ^{o}C for 5 min.

TH-1255_09610610

The nucleic acid transferred membranes were baked for 2 h at 80 °C in hybridization oven, and the hybridization was performed at 65 °C with gene specific probes. Probes for DNA/RNA blots were generated through PCR amplification of the cDNA inserts using the primers PpFAD2-2F and 2R (Table 5.1), the amplification of PCR product with PCR reaction conditions described in Chapter 3, section1.2.16. The probe was gel eluted and purified using Qiagen gel extraction kit (cat. no. 28704), following the manufacturers protocol. Eluted DNA was diluted to a concentration of 25 ng in 45 µl of Tris-EDTA buffer (pH 8.0) denatured at 95 °C for 10 min in a thermal cycler. The purified probe was labeled with ^{32}P-dCTP radio isotopes using Rediprime random primer labeling kit (cat. no. RPN1633, Amersham Biosciences) and was purified using Probequant G-50 column (Pharmacia Biotech, India) according to the manufacturer's instruction.

5.3 RESULTS AND DISCUSSION

5.3.1 Isolation and sequence analysis of two microsomal fatty acid desaturase genes from *P. pinnata.*

cDNA containing the complete coding regions of two genes sequences were screened by using probe (approximately 434 bp of partial *FAD2* gene radio labeled with ^{32}P-radio isotope), from the earlier constructed early immature seeds (90-DAF) of *Pongamia* cDNA library (Chapter-3). Based on the hybridization signals obtained, the clone numbers 137 & 284 were identified to be *FAD2* genes and were sequenced. Sequencing of the cDNA sequences obtained from the sequencing service (Macrogen, Korea) were subjected to nucleotide blast. Full-length cDNAs of two microsomal oleate desaturase were identified based on the homologous sequence alignment with gene sequences of other closely related species. The nucleotide sequence of PpFAD2-1 was 1,441 bp, containing 1158 bp of an open reading frame (ORF), a 5′-untranslated region (5′-UTR) of 71 bp, and a 3′-UTR of 212 bp. The A+T and G+C content in the sequence were 49.8 % and 50.2 %, respectively (Figure 5.1A). Similarly the full length cDNA sequence of PpFAD2-2 was 1508 bp in length and contained an ORF of 1152 bp, a 5′-UTR of 95 bp, and a 3′-UTR of 261 bp; the A+T content and G+C content in the sequence were 52.2 % and 48.8 %, respectively (Figure 5.1B). The ORFs in both the sequences began with an ATG start codon and ended with a stop codon that encodes

TH-1255_09610610

predicted proteins of 385 and 383 amino acid residues (calculated molecular masses 44.48 and 44.18 kDa; predicted p*I*s 8.81 and 8.70, respectively). The chemical formula of PpFAD2-1 & PpFAD2-2 nucleotide sequences are $C_{2097}H_{3104}N_{520}O_{533}S_{11}$, $C_{2065}H_{3062}N_{524}O_{536}S_{12}$ respectively and its half-life at *in vitro* conditions is predicted to be 30 h for both proteins. Amino acid composition analysis of PpFAD2-1 showed that Ala, Gly, Pro (each 6.5 %), Ile, Lys (5.5 %), Leu (10.9 %), Phe (5.7 %), Ser (6.2 %), Tyr (7.5 %) and Val (7.8 %) are the most abundant amino acids. Similarly PpFAD2-2 showed Ala (6.5 %), Gly (6.0 %), His, Ile, Lye, Phe, Thr (each 5.0 %), Leu (10.4 %), Pro (6.8 %), Ser (5.5 %), Tyr (7.8 %) and Val (8.1 %) are the most abundant residues. cDNA sequences encoding PpFADF2-1 and PpFAD2-2 were deposited with GenBank under accession numbers KF651985 & KF651986 respectively.

Analysis of the deduced amino acid sequences of two microsomal oleate desaturases from *P. pinnata* (PpFAD2-1 & PpFAD2-2) revealed that both sequences shows eight histidines distributed between three histidine boxes, HXXXH, HXXHH and HXXHH**.** According to Shanklin et al (1994, 1998), these histidine boxes, which are highly conserved among membrane-bound acyl-CoA and acyl-lipid desaturases have been proposed to act as potential ligands for non-heme iron atoms. These microsomal proteins probably use a common reactive center, and the histidine rich motifs in the sequence are ought to be involved in oxygen activation and substrate activation process by forming a part of di-iron center (Shanklin et al 1997). According to Kurdrid et al (2005), the Spirulina-Δ6 desaturase enzyme lost its activity due to the mutations that occurred in single histidine position of one of the three boxes. All three conserved histidine boxes were located in hydrophilic regions and according to topological model these domains are exposed towards the cytoplasmic side (Los & Murata 1998). The position of the histidine boxes and sequence similarity with other related species are mentioned in Table 5.2.

TH-1255_09610610

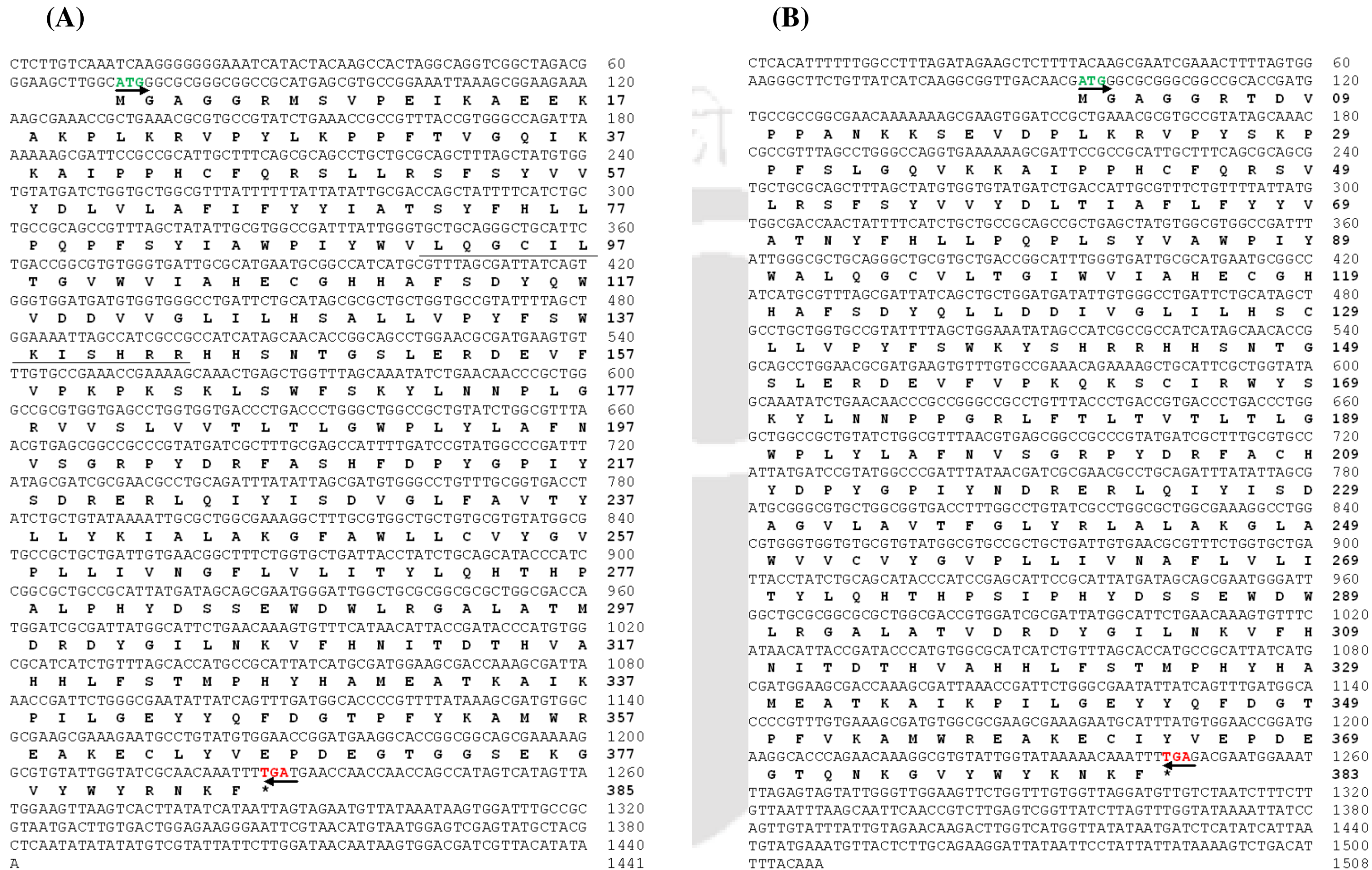


Figure 5.1: A complete nucleotide sequence and deduced polypeptide of (**A**) FAD2-1 and (**B**) FAD2-2 cDNA from *P. pinnata*. The start and stop codons are marked by arrows and also the stop codon denoted by asterisk (*). Horizontal lines indicate primers synthesized for expression studies. The 5' untranslated region (UTR) is 71 & 95 bp and the 3' UTR is 212 & 261 bp for *PpFAD2*-1 & *PpFAD2*-2 genes respectively.

TH-1255_09610610

Table 5.2: Sequence comparison of deduced amino acids and the conserved histidine boxes.

Sources	% identity and similarity of deduced amino acids	Sequences of conserved histidine boxes (H-Box)		
	Identity (similarity)	**H-Box 1**	**H-Box 2**	**H-Box 3**
PpFAD2-1 (KF651985)	100	105 HECGH	141 HRRHH	315 HVAHH
GmFAD2-1A (AAB00859)	84	109 HECGH	145 HRRHH	319 HVAHH
OeFAD2-1 (AAW63040)	85	105 HECGH	141 HRRHH	315 HVAHH
SiFAD2 (AAX11454)	83	105 HECGH	141 HRRHH	315 HVTHH
JcFAD2-2 (JN544422)	79	109 HECGH	145 HRRHH	319 HVTHH
PpFAD2-2 (KF651986)	100	105 HECGH	141 HRRHH	315 HVAHH
VfFAD2 (AAN87573)	91	105 HECGH	141 HRRHH	315 HVAHH
OeFAD2-2 (AAW63041)	81	105 HECGH	141 HRRHH	315 HVAHH
AtFAD2 (AAA32782)	81	105 HECGH	141 HRRHH	315 HVAHH
GmFAD2-2 (AAB00860)	88	105 HECGH	141 HRRHH	315 HVAHH

Number in parentheses denotes the GenBank accession numbers.

TH-1255_09610610

The deduced amino acid sequence alignment of PpFAD2- 1 and PpFAD2-2 with other related sequences revealed that 74 % sequence similarity exists and the accession numbers used to align the sequences are as follows: *Olea europaea* (Oe FAD2-1, AAW63040; OepFAD2-2, AAW63041), *Sesamum indicum* (SiFaD2, AAX11454), *Jatropha curcas* (JcFAD2-2, JN544422), *Glycine max* (GmFAD2-1A, AAB00859; GmFAD2-2, AAB00860), *Vernicia fordii* (VfFAD2, AAN87573), *Arabidopsis thaliana* (At-FAD2, AAA32782). The secondary structure analysis was performed by SOSUI software (http://bp.nuap.Nagoya-u.ac.jp/sosui) and revealed that there are total six transmembrane amino acid domains present in both sequences (Figure 5.2A-B). The same domains were also observed in *J. curcas* FAD2 & FAD3, *Tropaceolum majus* FAD2 and cotton FAD2-3 (Mietkiewska et al 2006; Pingzhi et al 2013; Pirtel et al 2001).

Alignment of the two deduced amino acid sequences revealed that PpFAD2-1 and PpFAD2-2 shared 83 % identity (Figure 5.3) among them. PpFAD2-1 and PpFAD2-2 amino acid sequences displayed significant homology to the reported plant FAD2 sequences (about 79-91 % identity) encoding microsomal oleate desaturases. However aromatic amino acid containing sequences have been found at the C-terminal end of both sequences which are thought to be mandatory for maintaining localization in the ER (McCarney et al 2004).

TH-1255_09610610

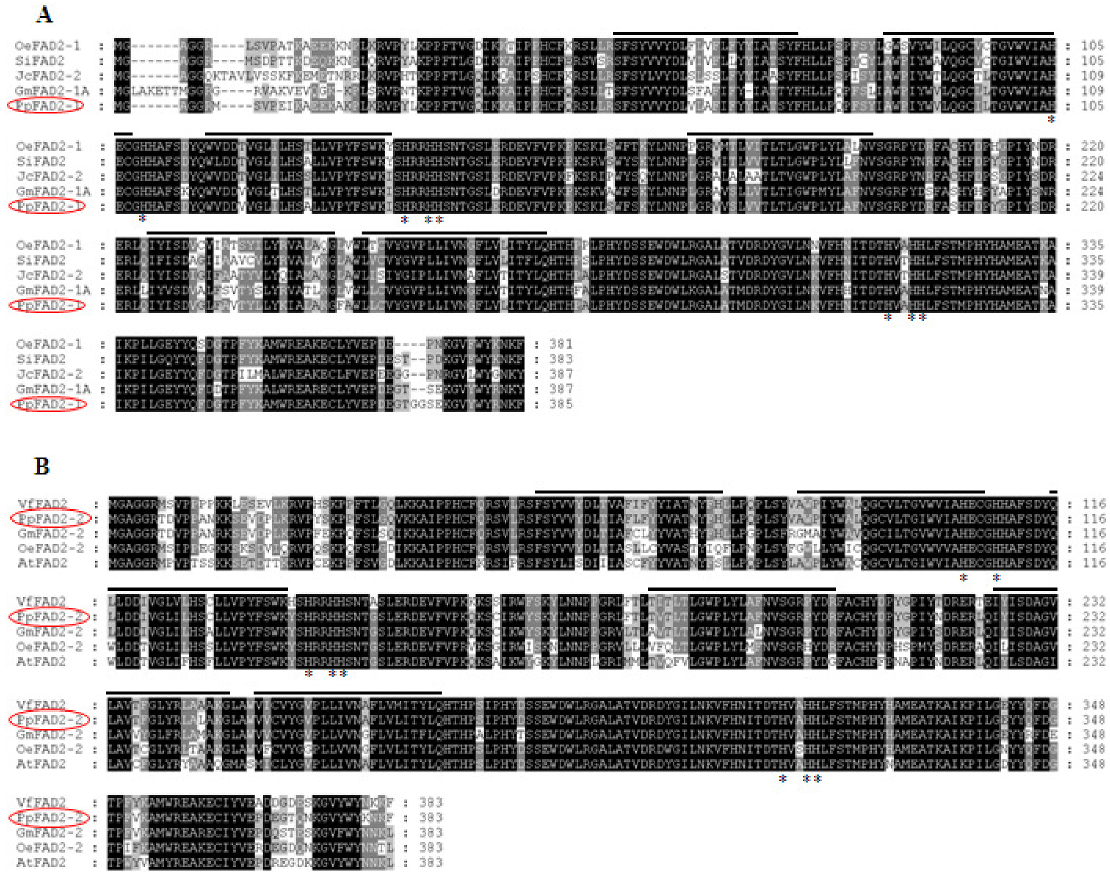


Figure 5.2: Comparison of the deduced amino acid sequences of two microsomal FA desaturases from *P. pinnata* with other characterized microsomal desaturases. (**A**) PpFAD2-1 with other seed specific FAD2 proteins. (**B**) PpFAD2-2 with other constitutive FAD2 proteins. Strictly identical residues are indicated by a black background and gray boxes indicate conservative changes (identity ≥75) between the desaturases. Eight conserved histidine residues are indicated by asterisks beneath the sequences. The potential transmembrane domains are indicated by horizontal lines up the sequences. Accession numbers of the desaturases included in the alignment: *O. europaea* (OeFAD2-1, AAW63040; OepFAD2-2, AAW63041), *S. indicum* (SiFaD2, AAX11454), *J. curcas* (JcFAD2-2, JN544422), *G. max* (GmFAD2-1A, AAB00859; GmFAD2-2, AAB00860), *V. fordii* (VfFAD2, AAN87573), *A. thaliana* (At-FAD2, AAA32782).

TH-1255_09610610

```
                        *        20         *        40         *        60         *        80         *       100         *
PpFAD2-1 : MGAGGRMSVPEIKAEEKAKPLKRVPYLKPPFTVGQIKKAIPPHCFQRSLLRSFSYVVYDLVLAFIFYYIATSYFHLLPQPFSYIAWPIYWVLQGCILTGVWVIAHECGHHAFSDYQ : 116
PpFAD2-2 : MGAGGRTDVPPANKKSEVDPLKRVPYSKPPFSLGQVKKAIPPHCFQRSVLRSFSYVVYDLTIAFLFYYVATNYFHLLPQPLSYVAWPIYWALQGCVLTGIWVIAHECGHHAFSDYQ : 116
           MGAGGR  VP         PLKRVPY KPPF36GQ6KKAIPPHCFQRS6LRSFSYVVYDL 6AF6FYY6AT YFHLLPQP SY6AWPIYW LQGC6LTG6WVIAHECGHHAFSDYQ
                                                                                                                   *   *

             120         *       140         *       160         *       180         *       200         *       220         *
PpFAD2-1 : WVDDVVGLILHSALLVPYFSWKISHRRHHSNTGSLERDEVFVPKPKSKLSWFSKYLNNPLGRVVSLVVTLTLGWPLYLAFNVSGRPYDRFASHFDPYGPIYSDRERLQIYISDVGL : 232
PpFAD2-2 : LLDDIVGLILHSCLLVPYFSWKYSHRRHHSNTGSLERDEVFVPKQKSCIRWYSKYLNNPPGRLFTLTVTLTLGWPLYLAFNVSGRPYDRFACHYDPYGPIYNDRERLQIYISDAGV : 232
            6DD6VGLILHS LLVPYFSWK SHRRHHSNTGSLERDEVFVPK KS 6 W5SKYLNNP GR6 3L VTLTLGWPLYLAFNVSGRPYDRFA H5DPYGPIY DRERLQIYISD G6
                                   *   **

                 240         *       260         *       280         *       300         *       320         *       340
PpFAD2-1 : FAVTYLLYKIALAKGFAWLLCVYGVPLLIVNGFLVLITYLQHTHPALPHYDSSEWDWLRGALATMDRDYGILNKVFHNITDTHVAHHLFSTMPHYHAMEATKAIKPILGEYYQFDG : 348
PpFAD2-2 : LAVTFGLYRLALAKGLAWVVCVYGVPLLIVNAFLVLITYLQHTHPSIPHYDSSEWDWLRGALATVDRDYGILNKVFHNITDTHVAHHLFSTMPHYHAMEATKAIKPILGEYYQFDG : 348
            AVT5 LY46ALAKG AW66CVYGVPLLIVN FLVLITYLQHTHP 6PHYDSSEWDWLRGALAT6DRDYGILNKVFHNITDTHVAHHLFSTMPHYHAMEATKAIKPILGEYYQFDG
                                                                                             *   *

           *       360         *       380
PpFAD2-1 : TPFYKAMWREAKECLYVEPDEGTGGSEKGVYWYRNKF : 385
PpFAD2-2 : TPFVKAMWREAKECIYVEPDEGTQNKGVYWYKNKF-- : 383
           TPF KAMWREAKEC6YVEPDEGT       Y  4
```

Figure 5.3: Comparison of the deduced amino acid sequences of *Pongamia FAD2* genes. Endoplasmic reticulum (ER) signal sequences are indicated by arrow.

TH-1255_09610610

The phylogenetic relationship of the *PpFAD2* genes and its deduced amino acid sequence with other characterized oleate desaturase sequences were done by using MEGA 5.1v and N-J tree was constructed (Figure 5.4). As reported earlier (Hernadez et al 2005), the plant Δ12 desaturase sequences were classified namely housekeeping type and seed type FAD2s. In phylogenetic tree PpFAD2-1 is closely grouping with *G. max* (GmFAD2-1A, AAB00859; GmFAD2-1B, BAD89861) which is a legume species whereas it shows distant relationship with *Helianthus annuus* (HaFAD2-1, AAL68981) and *O. europaea* (OepFAD2-1, AAW63040). Similarly PpFAD2-2 shows strong close relationship with *J. curcas* (JcFAD2-2, JN544422), *V. fordii* (VfFAD2, AAN87573) and *G. max* (GmFAD2-2, AAB00860; GmFAD2-3, ABF84063). But shows distant grouping with *Gossypium hirsutum* (GhFAD2-1, CAA65744; GhFAD2-2, CAA71199; GhFAD2-3, AAL37484), *G. max* (GmFAD2-2, AAB00860; GmFAD2-3, ABF84063), *O. europaea* (OepFAD2-1, AAW63040), *H. annuus* (HaFAD2-2, AAL68982; HaFAD2-3, AAL68983). Interestingly, PpFAD2-1 grouped in a cluster with *FAD2* genes which exhibit "seed specific" expression pattern whereas PpFAD2-2 clustered with *FAD2* genes that show "housekeeping" expression strategy throughout the tissues. Reports are available on seed specific FAD2s; which may have evolved independently after separation by duplication from housekeeping *FAD2* genes (Marinez-Rivas et al 2001).

Hydrophobicity plot analysis of the two FAD2 amino acid sequences were generated by the method of Kyte & Doolittle (1982). Four and two different hydrophobic regions were found in PpFAD2-1 and PpFAD2-2, respectively (Figure 5.5A & B). Among the four, two flanking hydrophobic domains are long enough to span the membrane twice and according to Shanklin et al (1994) the predicted membrane banding domains are part of structural model of fatty acid desaturases.

TH-1255_09610610

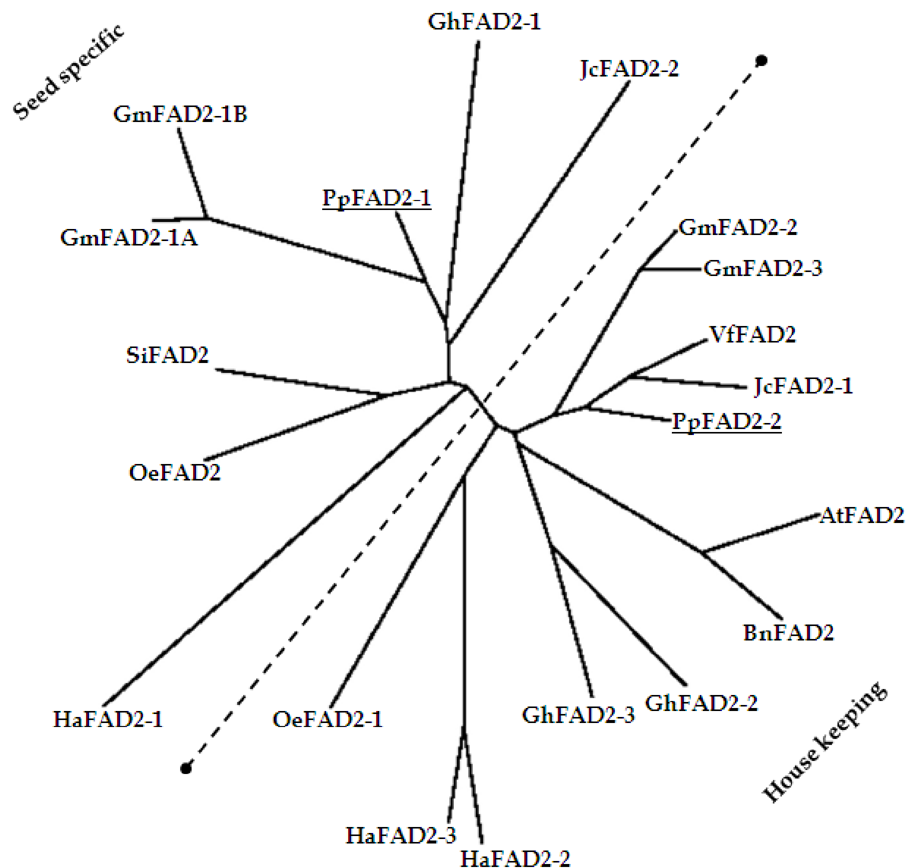


Figure 5.4: Phylogenetic analysis of microsomal FAD2 desaturases. Unrooted tree calculated using the neighbor-joining method in MEGA4.1. Positions of the two *P. pinnata* microsomal FA desaturases are underlined. Accession numbers of the desaturases included in the analysis: *J. curcas* (JcFAD2-1, JN544421; JcFAD2-2, JN544422), *G. hirsutum* (GhFAD2-1, CAA65744; GhFAD2-2, CAA71199; GhFAD2-3, AAL37484), *G. max* (GmFAD2-1A, AAB00859; GmFAD2-1B, BAD89861; GmFAD2-2, AAB00860; GmFAD2-3, ABF84063), *S. indicum* (SeFAD2, AAX11454), *O. europaea* (OepFAD2-1, AAW63040; OepFAD2, AAW63041), *H. annuus* (HaFAD2-1, AAL68981; HaFAD2-2, AAL68982; HaFAD2-3, AAL68983), *B. napus* (BnFAD2, AAF78778), *V. fordii* (VfFAD2, AAN87573), *A. thaliana* (AtFAD2, AAA32782).

TH-1255_09610610

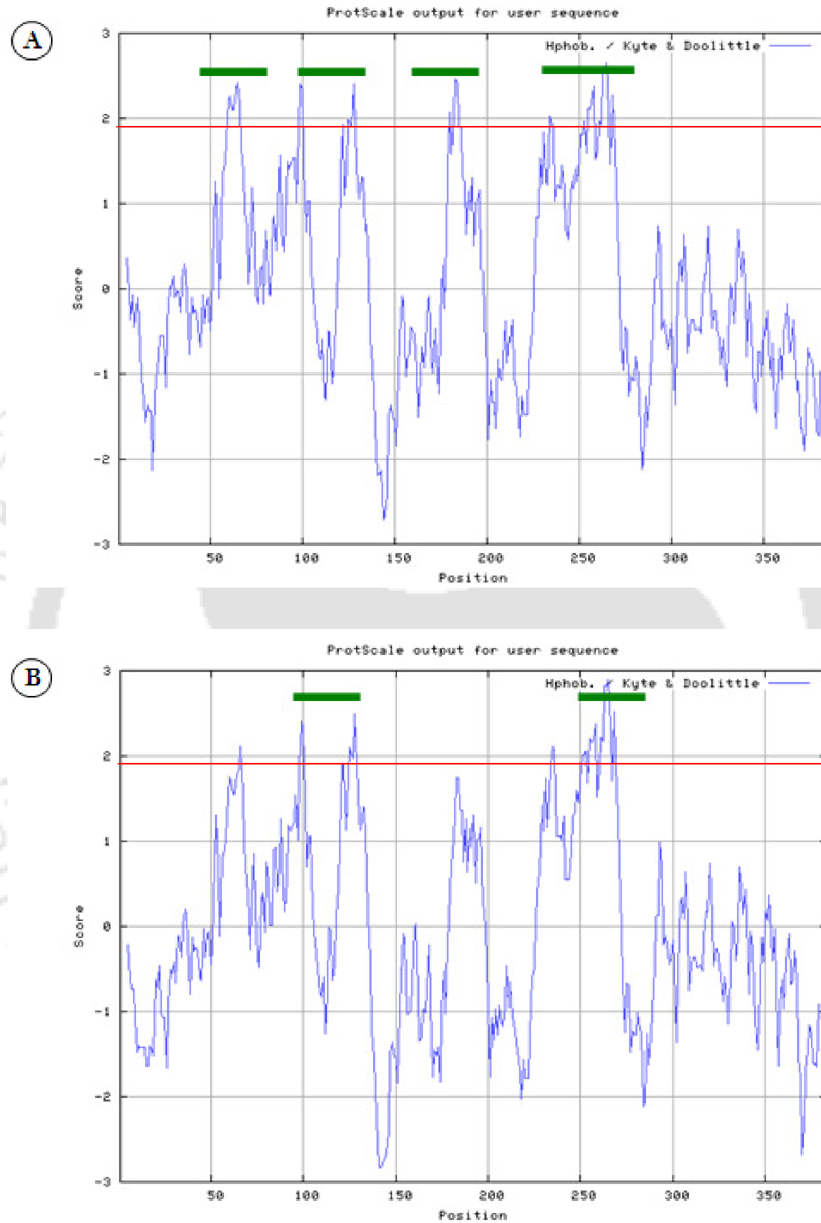


Figure 5.5: Hydropathy plot of *Pongamia* (**A**) PpFAD2-1; (**B**) PpFAD2-2. Numbers on the x-axis are amino acid residues. Numbers on the y-axis are hydrophobic means. Horizontal bars hydrophobic domains.

TH-1255_09610610

5.3.2 Genomic organization and tissue specific expression of *PpFAD2*

The 5' and 3' UTR sequences of two *FAD2* genes were unique, except the coding sequences. This makes us to prepare gene specific probes which corresponds to each *Pongamia FAD2* gene by PCR and suggested that *PpFAD2*-1 and *PpFAD2*-2 genes were two distinct members of the *Pongamia* microsomal oleate desaturase gene family.

To characterize the genomic organization of the *PpFAD2* gene in *Pongamia* oil seeds, approximately 10 μg of genomic DNA was extracted from the seeds. The $A_{260/280}$ ratio was found to be 1.7, which is free of contaminants and was subjected to restriction digestion. The product was analyzed on 1 % agarose gel to resolve digested products completely. The digested product shows complete smear which has fragments of *Pongamia* digested genomic DNA and was transferred to Hybond N^+ membrane. The PCR amplification gave approximately 424 bp of *PpFAD2* gene product, the amplified product was eluted and purified. The membrane with digested product was probed with *PpFAD2*-1 PCR gene product (~424 bp). Though there are more than 2 hybridization bands with high stringency conditions, there was one distinct band around 4.5 kb and a large week band of approximately 12 kb observed from the genomic DNA when digested with *Hind*III restriction enzyme. By digestion with *Eco*RI, two distinct bands of around 6.0 kb, 2.0 kb and one faint band approximately 8.0 kb was observed. Neither enzyme has restriction sites within the probe used. By taking the above fragments of *Pongamia* genomic DNA into consideration, the genome may have at least two copies of *FAD2* gene as observed in other related plant genome (Heppard et al 1996; Liu et al 1999; Nielsen et al 1997) (Figure 5.6A-C).

Sufficient amount of RNA was obtained and the purity ($A_{260/280}$) of the RNA samples were 1.9 (described in chapter-3, section 3.3.2) and used to study gene expression profiles. To study the tissue specific expression and developmental regulation of *PpFAD2*-1 and *PpFAD2*-2 genes at the transcriptional level, semi quantitative PCR was performed using gene specific primers with the total RNAs isolated from various developmental stages of *Pongamia* seeds (90-DAF, 180-DAF, 270-DAF and 360-DAF) and also from root, leaf tissues. To ensure equal amounts of RNA template, the experiment was conducted by incorporating the constitutive expressed gene *GAPDH*, which shows housekeeping pattern of expression.

TH-1255_09610610

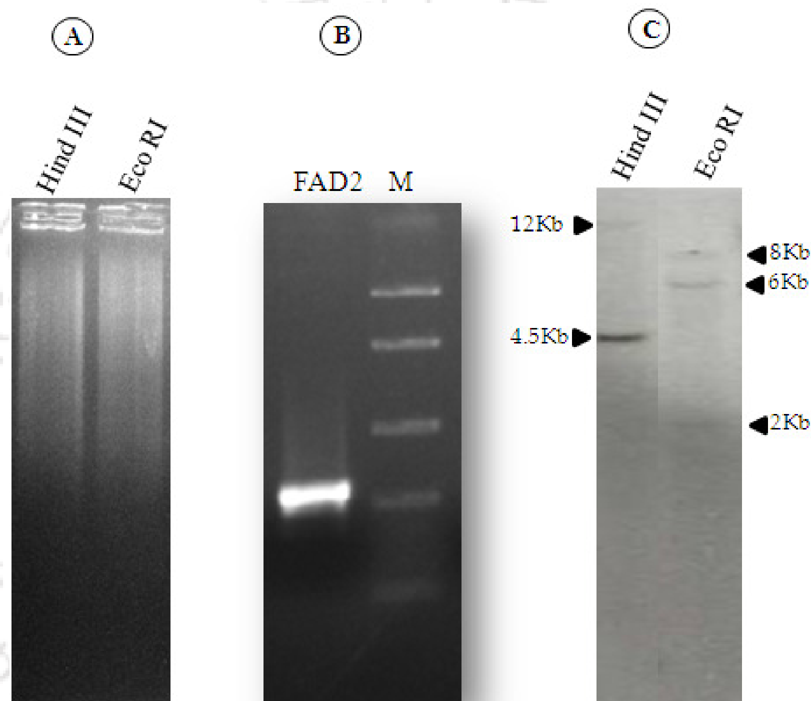


Figure 5.6: Southern blot hybridization. (**A**) Restriction digestion of *Pongamia* genome; (**B**) The partial sequence of fatty acid desaturase from *Pongamia* shows ~424 bp amplified product. Lane-1 amplified PCR product; Lane-2 Middle range DNA ladder (Fermentas). (**C**) Southern blot analysis of *Pongamia* genomic DNA digested with the indicated restriction enzymes, and probed (^{32}P-labeled) with partial sequence of PpFAD2-1.

TH-1255_09610610

The results revealed that PpFAD2-1 gives amplification with RNA extracted from various developmental stages of seeds with increase in expression levels but no amplification was observed with root and leaves, indicating that function of PpFAD2-1 are restricted to seed developmental stages. Although in case of PpFAD2-2 the PCR amplification was observed with all tissue types and the same was observed with GAPDH primers also (Table 5.3 and Figure 5.7).

Table 5.3: Semi-quantitative PCR analysis

Gene	Root	Leaf	90-DAF	180-DAF	270-DAF	350-DAF
PpFAD2-1	-	-	+	+	+	+
PpFAD2-2	+	+	+	+	+	+
GAPDH	+	+	+	+	+	+

To further confirm, the constitutive expression pattern of the gene *PpFAD2*-2 in tissues of *Pongamia* (leaves, roots, seed developmental stages), northern blot analysis was performed and the results revealed that the gene is showing expression levels in all the tissue types. The high intensity probe binding was observed with leaves and roots, whereas faint bands were observed with seed developmental stages when compared to other tissues (Figure 5.8). Though there is a house keeping pattern of gene expression with *PpFAD2*-2, the high level of expression profile was observed with leaves and roots; this is may be that the enzyme is involved in membrane formation and maintenance of its integrity.

TH-1255_09610610

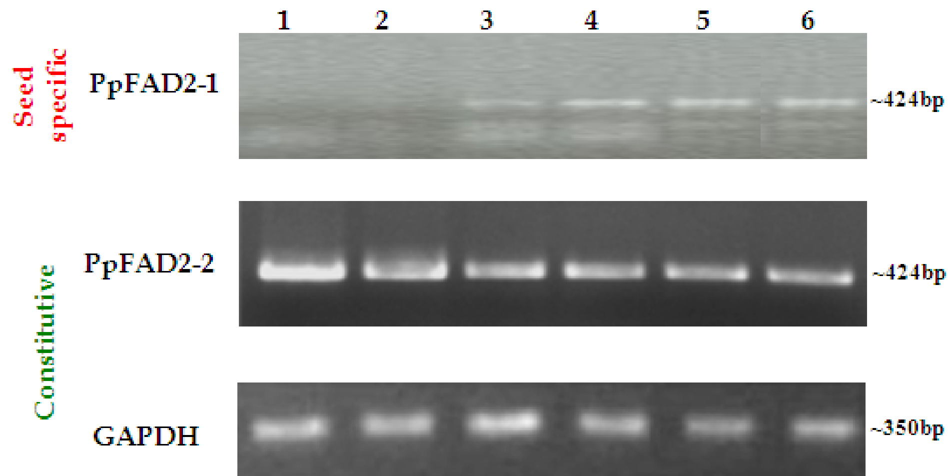


Figure 5.7: Semi quantitative PCR analysis of *PpFAD2*-1 and *PpFAD2*-2 genes. First strand cDNAs were synthesized from total RNA extracted from different tissues. To ensure equal amounts of template, GAPDH was used as a reference gene. Lane 1-root; 2-leaf; 3-90DAF; 4-180DAF; 5-270DAF and 6-350DAF.

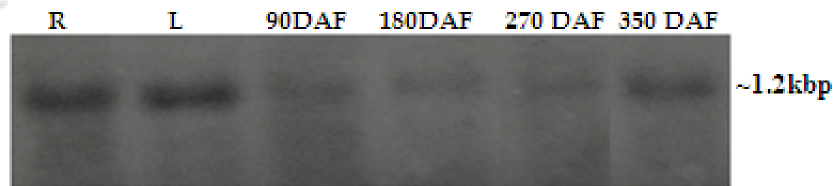


Figure 5.8: Northern hybridization analysis of *PpFAD2*-1 expression of *Pongamia* tissues. The blot prepared with 10µg of the total RNA isolated from various tissues of *Pongamia* was analyzed by RNA blot hybridization with ^{32}P-labelled (PpFAD2-2) cDNA fragment of *P. pinnata*.

TH-1255_09610610

Similarly real time PCR was performed by using gene specific primers to confirm the tissue specific expression pattern of *PpFAD2*-1 gene. Total RNA was isolated from developing seeds at different stages (90-DAF, 180-DAF, 270-DAF and 350-DAF). The expression analysis showed that the action of *PpFAD2*-1 starts in the early stage of seeds and the rates of fatty acid synthesis gradually increased proportionally with seed growth. The PpFAD2-1 expression levels in 350-DAF tissue is almost double when we compare with the expression levels in 90-DAF, but minimal difference was observe between 18-DAF, 270-DAF tissues (Figure 5.9). Different pattern of expression of both *FAD*2 genes was observed in which one gene shows predominant expression in developing seeds (*PpFAD2*-1), whereas the other shows constitutive expression (*PpFAD2*-2) in all tissues of *Pongamia*.

In general, some seed microsomal desaturase genes and other related gene, which are developmentally regulated by various factors (Holdsworth et al 1999; Murphy & Piffanelli 1998; Seffens et al 1990), play in important role during seed development in determining seed storage lipid composition. Heppard et al (1996) have isolated two different cDNA sequences, designated FAD2-1 and FAD2-2 cDANs encoding microsomal ω-6 desaturase in *G. max* and presented the molecular evidence that although both cDNAs are expressed in developing soybean seeds, their expression specificity and function in lipid biosynthesis are not identical. They suggested that the FAD2-1 gene, which has a speed-specific expression pattern, may be involved in desaturation of oleic acid of seed storage lipid by providing the experimental data that the *FAD2*-1 gene is expressed much greater than the *FAD2*-2 gene during biosynthesis of seed storage lipids. In contrast FAD2-2 gene, which is constitutively expressed in both vegetative and developing seeds, may be responsible for conversion of oleic acid to linoleic acid with membrane lipids. The current study is strongly supported by the existing reports by robustly matching *PpFAD2*-1 and *PpFAD2*-2 genes with *G. max* (leguminoceae) fatty acid desaturases, which implicated in biosynthesis of *G. max* storage lipids because of its expression and much higher sequence similarity. In addition the both genes could be utilizable in manipulating seed oil composition by current genetic engineering technology for a variety of applications in biofuel industries.

TH-1255_09610610

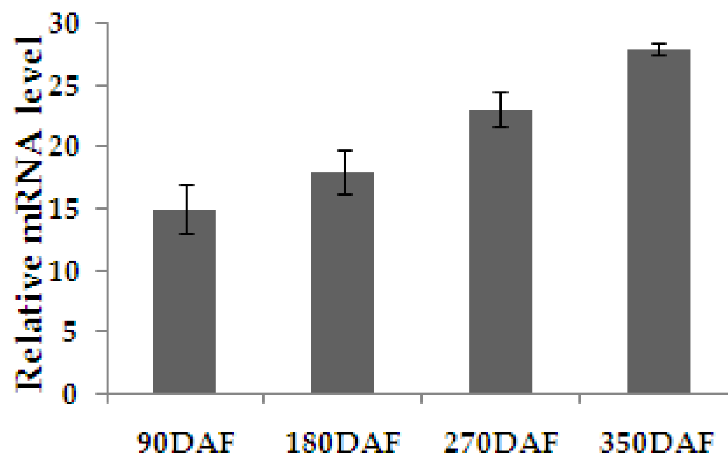


Figure 5.9: Expression of *PpFAD2*-1 gene: Real time PCR was performed with cDNA isolated from seeds of early immature (90-DAF), early mature (180-DAF), late immature (270-DAF) and late mature (350-DAF). Relative expression ratio of each sample is compared with the sample at 350-DAF. (Error bars indicate standard deviation values of three independent experiments).

TH-1255_09610610

Plant fatty acid desaturase is responsible for catalysis of oleic acid to linoleic acid through the eukaryotic pathway of plant lipid biosynthesis. In higher plants, this enzyme is located in the microsomal membrane of plant seeds and plays the major role in desaturation of microsomal membrane lipids during plant seed development (Ohlrogge 1995). Recently, the gene encoding the microsomal ω-6 fatty acid desaturase has been characterized form some plant species. In this chapter, two full length genes coding for a fatty acid desaturases have been screened that may be involved in synthesis of the seed storage linoleic acid. In addition, developmental regulation of the two *PpFAD2* genes expression at the transcriptional and quantitative level was investigated. The results demonstrated that the nucleotide sequences of the PpFAD2-1 and PpFAD2-2 cDNA contain one ORF corresponding to 385, 383 amino acid redidues and relatively high similarities (more than 74 %) to those of other plant fatty acid desaturase genes. Table 5.2 also reveals that the deduced amino acid sequences of the PpFAD2-1 and PpFAD2-2 polypeptide have much higher identities (79-91 %) with and similarities to the sequences of FAD2. Based on these observations we could assumed that the both *PpFAD2*-1 and *PpFAD2*-2 cDNA would code for a fatty acid desaturases in *Pongamia* oil seeds. Another evidence for this assumption was give by the presence of two important sequences. One is the presence of a transmembrane targeted signal peptide in N-terminal region of the both PpFAD2-1 and PpFAD-2 polyopeptide, which is supposed to be directed towards ER membrane (Nielsen et al 1997; Shell 1998; Caliebe & Soill 1999). The other is the presence of eight histidines in three boxes HXXXH, HXXHH and HXXHH which are highly conserved in the membrane bound proteins containing non heme iron motifs (Lidqvist et al 1996). The occurrence of these sequences in PpFAD2 polypeptide strongly suggests that the PpFAD2 is a fatty acid desaturase becasuuse the sequences are characteristic of the plant microsomal membrane bound fatty acid desaturases (Okuley et al 1994; Nielsen et al 1997, 1999; Lee et al 1998).

5.4 CONCLUSION

We have isolated and characterized two microsomal oleate desaturase genes from *Pongamia*, both containing the three histidine boxes typical of all membrane-bound desaturases. Standard sequence analysis of the two genes *PpFAD2*-1 and *PpFAD2*-2

TH-1255_09610610

indicates that they code for two microsomal oleate desaturase enzymes. To the best of our knowledge this is the first time that the cloning and expression studies of *FAD2* genes from *P. pinnata* being reported. Based on the results obtained on expression studies, it can be said that *PpFAD2*-1 is possibly involved in the desaturation of storage lipids during the seed developmental stages only. Whereas *PpFAD2*-2 may be mostly responsible for reserve lipids in leaves and roots along with the seed developmental stages (constitutive expression), which indicates its possible participation in the linoleate biosynthesis required for membrane formation. The genomic southern blot revealed that there are at least two copies of *FAD2* genes present in the whole genome of *P. pinnata*. The exact localization of expression of two genes (housekeeping/seed specific) in *Pongamia* tissues provide new information about the role of each gene product in regulating PUFA biosynthesis in membrane and storage lipids. The present study provides background information on future study of genetic regulation either at gene expression or lipid desaturation level in *Pongamia pinnata*. Further research should therefore be carried out to investigate FAD profiles at the post-transcriptional and translational levels in *Pongamia* tissues.

TH-1255_09610610

Chapter 6

Nuclear Genome Estimation and Chromosome number determination in *P. pinnata*

The chapter describes the nuclear genome estimation in CPT, various tissues of CPT, individuals collected from Assam and different location from India. Furthermore it also deals with somatic chromosome number, its correlation with nuclear genome.

TH-1255_09610610

NUCLEAR GENOME ESTIMATION AND CHROMOSOME NUMBER DETERMINATION IN *P. PINNATA*

6.1 INTRODUCTION

The depleting reserves of fossil fuel and increasing demand linked to growing cost of diesel combined with strict emission norms have necessitated the need for an alternative source of energy that can supplement or replace diesel. As an alternative fuel, in recent years biodiesel (methyl or ethyl esters of fatty acid) has received considerable attention (Jingura et al 2009; Kesari et al 2010). This is owing to the fact that it is biodegradable, renewable, non toxic and safe to store. In addition to low cost production, biodiesel has been reported to be non-toxic, clean and ecofriendly. Both edible and non-edible seeds have been explored for the production of biodiesel. Due to scarcity of edible oils in developing countries many researchers have been working on non-edible oil yielding plants like *Pongamia pinnata.* There is a growing interest in the use of *P. pinnata* to alleviate the energy crisis as the oil can be converted to biodiesel by chemical or biological transesterification (Kesari et al 2010; Kumar & Sharma 2008).

P. pinnata (Fabaceae) popularly referred to as 'Karanj' is a versatile legume having many potential applications and can be cultivated on the marginal lands without much attention (Kesari et al 2008). This plant is considered as strategic crop for countries like India, Australia and Brazil because of diverse uses and applications, of particular importance being the source of biodiesel. Although the plant looks promising, it lacks an improved germplasm. Consequently several pests and diseases have already been observed at industrial production level. While this potential biodiesel crops are being harvested all over the world with the purpose of crop improvement, little is known about its nuclear DNA content and chromosome number.

Since it became possible to measure the nuclear DNA content of a single nucleus, various researchers have reported variation among different species (Bennett & Leitch 1995, 2005). More recently it was found that variation in nuclear DNA content among species are predominantly associated with difference in amount of

TH-1255_09610610

repetitive sequences. Particularly retrotransposons play a dominant role in nuclear DNA content differences and most of the variation in nuclear DNA content of plants can be ascribed to its differential accumulation (Bennetzen 2007). The occurrence and extent of nuclear DNA content difference below species level are still controversial and not satisfactorily analyzed yet.

Flow cytometry (FCM), a fast and accurate method for the estimation of nuclear DNA content, has become the predominant technique for establishing plant genome size (Dolezel et al 2007; Greilhuber et al 2007). This technique has provided relevant information for biodiversity, ecological, evolutionary and taxonomy studies. Additionally precise estimation of nuclear DNA content has been considered important for breeding programs, sequencing projects and research using molecular markers (Bennett & Leitch 1995). During the last decade, one of the most challenging problems faced by FCM users has been adjusting the buffer composition to the requirements of specific species, especially if the material contains staining inhibitors such as phenols, caffeine, and other secondary metabolites in the cytosol of leaf cells (Loureiro et al 2007; Noirot et al 2003; Price et al 2000). Hence, the composition of the isolation buffer is critical for accurate FCM measurements. Besides facilitating the isolation of intact nuclei, the buffer should maintain stability, prevent aggregation, protect DNA from degradation, and provide an appropriate environment for specific and stoichiometric staining of nuclear DNA (Loureiro et al 2007). Phytochemical investigation of *P. pinnata* indicated the presence of abundant prenylated flavonoids which increases the complexities in addition to the phenolic compounds.

Chromosome number is important in cytotaxonomic studies and is useful in contributing to the delimitation of taxonomic units at different hierarchical levels. Chromosome count is generally constant for a species. Numerical variation such as polyploidy and aneuploidy among related species account for difference in the chromosome count. Recent studies showed that the polyploidization resulted in better harvest of the biodiesel crops (Kanellos 2009). Cytogenetic studies suggested the chromosome number of *P. pinnata* to be 2n=20 or 2n=22 (Atchinson 1951; Sarbhoy 1977). *Prosopis juliflora,* one of the biofuel plant, have chromosome number of 2n=28 and nuclear DNA content (2C) of 0.86 pg (Bukhari 1997). *J. curcas* another promising

TH-1255_09610610

biofuel crop has chromosome number of 2n=22 and the estimated nuclear DNA content of 2C=0.85 pg (Dahmer et al 2009).

In the current chapter our main objective was to measure nuclear DNA content by flow cytometry for *P. pinnata* with respect to *Z. mays* as standard by internal and pseudo-internal standardization. Along with that we performed experiments to detect for the presence of inhibitors and means to neutralize the effect of the PI inhibitors by addition of antioxidants in varying composition. The experiments were also performed to check the nuclear DNA content in various tissues [root, shoot, leaf, stem, flower and four stages of seeds (90-DAF, 180-DAF, 270-DAF and 350-DAF)] of elite genotype/candidate plus tree (CPT-NGPP46) of *Pongamia* and also between saplings collected from different locations of India. Nuclear DNA content calculation and comparison between CPT and *in vitro* raised plants were also checked. Also an attempt was made to see if any correlation exists between that of the nuclear DNA content and chromosome number in five individuals of *P. pinnata* collected from various locations of Assam. The knowledge will facilitate the use of the available germplasm resources, enabling the development of optimized breeding strategies focusing improvement in yield and production of this plant species particularly as it relates to the oil trait.

6.2 MATERIAL AND METHODS

6.2.1 Plant materials and culture

Seeds sampled for the current study were collected from five individuals occurring in their natural habitat (Assam) (Table 6.1). The seeds were rinsed in 25 % (v/v) ethanol for 10 min and washed with distilled water three times before placing in polythene bags containing sand and clay (ratio of 1:4) and hardened in mist chamber at 28 ^{o}C ± 2 ^{o}C and 70-80 % relative humidity. Saplings collected from various parts of India were also maintained at green house condition. Earlier characterized mature CPT of *P. pinnata* obtained from Sila Forest Range, North Guwahati, Assam (latitude 26º14’6”N & longitude 91º41’28”E) and acclimatized vitro plants (AVP) from CPT were obtained using nodal segments raised from seedlings of *P. pinnata* as described previously (Kesari et al 2012) and analyzed after 4 year cultivation period in an open field. Simultaneously *in vitro*-rooted plantlets (IRPs) of the CPT were sampled at the end of the rhizogenesis phase, immediately before transplanting to *ex vitro* condition. IRP

TH-1255_09610610

plants have been grown in trials for last five years. DNA content of nodal segments (NS) that were the starting material for micropropagation was also estimated. Seeds of reference standards [*Solanum lycopersicum* Stupicke polni tyckove rane (2C=1.96 pg), *Zea mays* `CE-777' (2C=5.43 pg), *Pisum sativum* Ctirad (2C=9.09 pg)] were gently supplied by Jaroslav Dolezel (Experimental Institute of Botany, Czech Republic) and were raised (Figure 6.1).

Table 6.1: Study site characteristics-samples collected from various parts of Assam and India

S. No	Description	Coordinates	Temperature (°C)	Height from sea level (m)
01	Tinsukia	27.5000° N, 95.3700° E	22.2	117
02	Kokrajhar	26.4000° N, 90.2700° E	23	37
03	Guwahati	26.1838° N, 91.7633° E	23	55
04	Jorhat	26.7500° N, 94.2200° E	22	116
05	Tezpur	26.6300° N, 92.8000° E	23	73
06	Banaras	25.282°N 82.9563°E	26	80
07	Lucknow	26.847°N 80.947°E	26	123
08	Gorakhpur	26.7588°N 83.3697°E	26	88
09	Bhubaneswar	20.27°N 85.84°E	26	45
10	Hyderabad	17.366°N 78.476°E	26	536
11	Amravati	21.15°N 79.09°E	26	340

6.2.2 Sample preparation and isolation of nuclei

In preliminary experiments, five different buffers/procedures for the extraction and staining of the cells were tested namely Galbraith's buffer (Galbraith et al 1983), LB01 (Dolezel et al 1989), Propidium iodide/hypotonic Citrate (Krishan 1975), General Purpose Buffer (Loureiro et al 2007) and Woody plant buffer (Loureiro et al 2007) respectively. Out of all, Krishan's (1975) buffer was preferred because it gave clear histograms with minimum nuclear disruption and back ground noise. The youngest fully developed leaf sheath was used for the preparation of the suspension of intact nuclei. The leaves (1 cm^2) were chopped with a razor blade in 1 ml of modified Propidium iodide/hypotonic Citrate buffer by Awtar Krishan (1975) containing 0.1 % w/v sodium citrate (Sigma-Aldrich, cat. no. S4651), 0.3 % v/v detergent IGEPAL CA-630 (Sigma-Aldrich, cat. no. I3021). Staining was done with 25 mg/l PI (Sigma-Aldrich, cat. no. P4170) in a solution containing DNase-free RNase A (2 mg/ml) (Sigma Aldrich, cat. no. P4875). The suspension of nuclei was filtered through 30 µm nylon mesh (Swedesboro, NJ) and used for flow cytometric analysis (Figure 6.2).

TH-1255_09610610

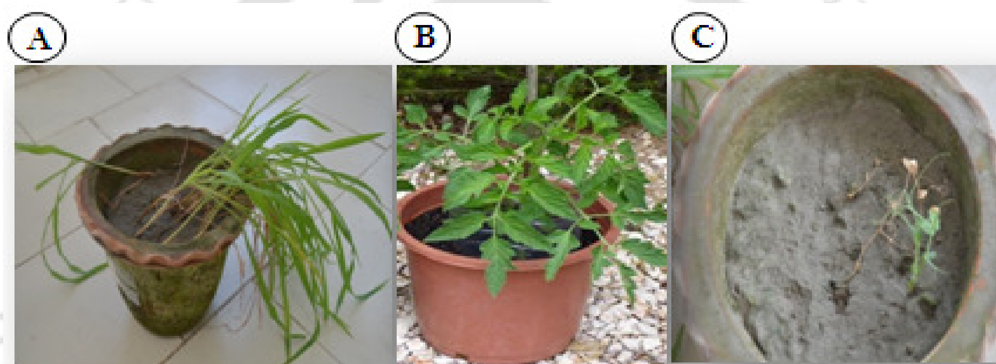


Figure 6.1: Collection and maintenance of standard germplasm (Reference plants). (**A**) *Z. mays*, (**B**) *S. lycopersicum*, (**C**) *P. sativum*.

TH-1255_09610610

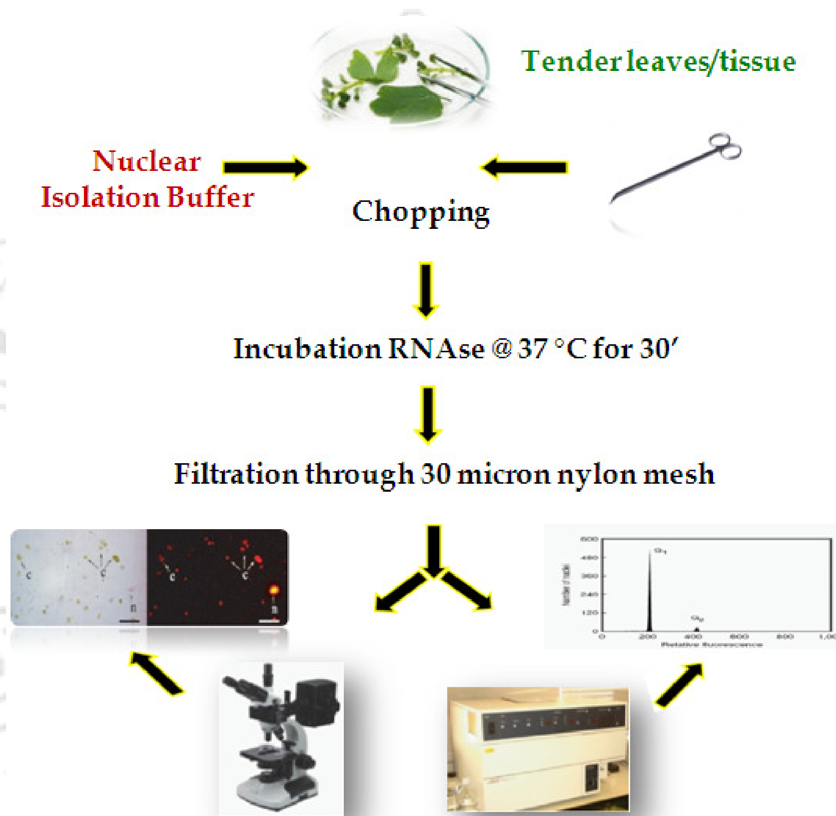


Figure 6.2: Flow diagram shows the process of sample in nuclear extraction buffer (NEB) and its analysis with Flow cytometer.

TH-1255_09610610

6.2.3 Flow cytometric analyses

Estimation of nuclear DNA content was performed with BD FACS Calibur flow cytometer (BD Biosciences, New Delhi, India). Samples stained with PI were excited with a 15 mW argon ion laser at 488 nm. PI fluorescence was collected through a 645 nm dichroic long-pass filter and a 620 nm band pass filter. The instruments settings i.e., voltage and gain were kept constant throughout the experiment. Three dot plots and one histogram plot was generated. Three dot plots were: (i) FSC vs. SSC to understand the size and granularity of the nucleus; (ii) SSC vs FL2A to spot intense fluorescence regions due to nuclei and to eliminate background fluorescence by drawing electronically a gate region around the signals due to intact nuclei; (iii) FL2A vs FL2W to discriminate between singlets and doublets. Frequency vs FL histogram was generated to compare the mean position of the sample peaks relative to the internal standard. For each sample at least 10,000 nuclei were analyzed at a rate of 20-50 nuclei per seconds. Internal and pseudo internal standardization procedures were followed. Test sample and standards were co-chopped in the internal standardization. The samples of standard plant nuclei and *P. pinnata* were chopped separately then mixed together just before acquiring in pseudo-internal standardization. The nuclear DNA content was estimated by comparing the mean fluorescence intensity of nuclei of the sample material with that of the reference standard and obtained by multiplying the nuclear DNA content of standard species by the ratio of their fluorescence intensities. The results were acquired using BD Cell Quest Pro software (version 6.0, BD Biosciences). The resulting histograms were analyzed using FlowJo v.7.6.5 (FlowJo, TreeStar Inc, Ashland, OR) for estimating mean fluorescent intensity (MFI), co-efficient of variation (CV). A suitable gating procedure was followed to resolve the data and to measure the parameters only for intact nuclei in a heterogeneous population. Auto fluorescence due to chloroplast was eliminated from analysis by selecting a region on fluorescence in FITC channel and PI fluorescence profile. To reduce the level of debris and disintegrated nuclei, the nuclei were gated in PI fluorescence channel vs. SSC dot plots. In this diagram the nuclei can clearly be identified by their defined fluorescence intensity/scattered light pattern. Doublets and clumps were eliminated by gating on fluorescence width and fluorescence area profiles. In this histogram, an "interest zone" was defined such that only single intact nuclei were included in the

TH-1255_09610610

fluorescence histogram which corresponds to G0/G1 and G2/M phases of the cell cycle. The mean fluorescence intensity of G0/G1 peak channel was taken into account to estimate genome size, as it resembles unreplicated haploid genome. The step-by-step gating procedure followed in this study described in Figure 6.3. The nuclear DNA content (picograms) was converted to base pairs by considering that 1 pg of DNA corresponds to 978 Mb (Dolezel et al 2003). The analysis was repeated if the coefficient of variation of the sample was >5.0. For accurate genome size estimation of *P. pinnata*, a total of 3 readings were taken on different days and an average of all was reported. Nuclear DNA content of *P. pinnata* was estimated using the linear relationship according to the following equation:

$$\text{Sample 2C nuclear DNA content (pg)} = \frac{\text{Sample 2C peak mean}}{\text{Reference standard 2C peak mean}} \times \text{Reference standard 2C peak mean}$$

6.2.4 Test for inhibitors

P. pinnata extracts were tested for unidentified compounds that reduce PI fluorescence of standard nuclei as follows. Nuclei were released from an approximate 40 mm x 20 mm *P. pinnata* leaf and one-half of a standard leaf that were simultaneously processed (co-chopped) and stained with PI (sample A). Sample B consisted of PI-stained nuclei from the independently processed and stained other half of the standard leaf used in sample A. After staining for 1 h, samples A and B were individually measured for mean PI fluorescence, after which they were mixed and measured up to 120 min. The experiment was repeated three times. Reduced fluorescence of nuclei from standard leaves simultaneously processed with *P. pinnata* leaves when compared to nuclei from independently processed standard leaves gave evidence of inhibitors in *P. pinnata* samples.

TH-1255_09610610

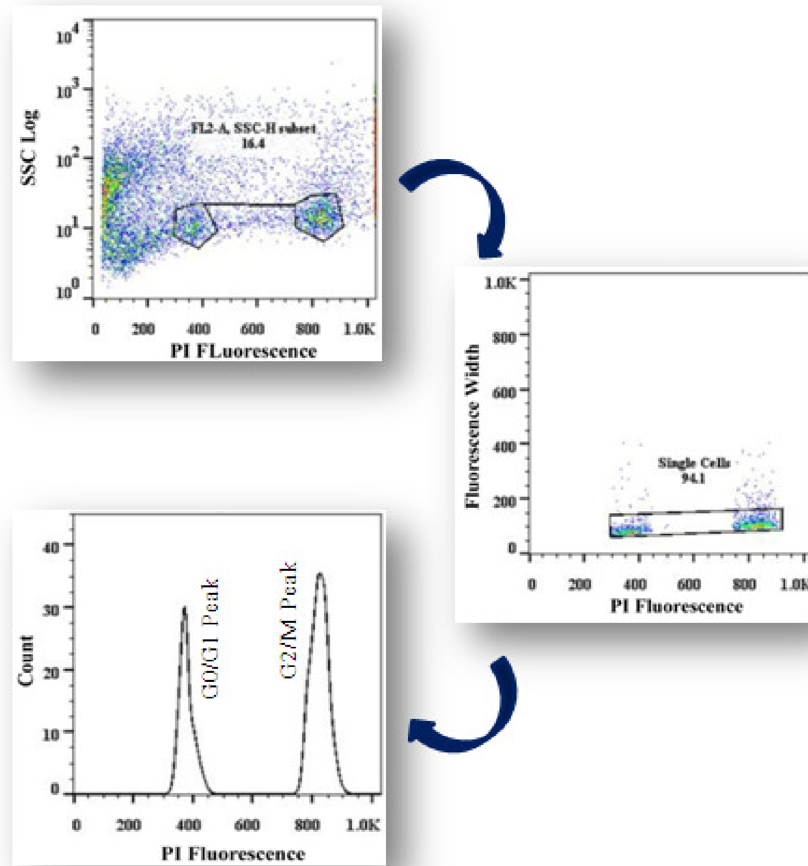


Figure 6.3: Step-by-step gating regime during standardization.

TH-1255_09610610

6.2.5 Test for concentration effect of inhibitors

The effect of buffer volume on PI fluorescence of nuclei from simultaneously processed *P. pinnata* and standard leaves was also measured. The experiment was designed to test whether or not the inhibitory effect of *P. pinnata* extracts on PI fluorescence of standard nuclei could be diluted out. In this experiment, approx. 13 mm square segments of *P. pinnata* and standard leaves were simultaneously processed in 0.1, 1.0 and 10.0 ml buffer. The samples were stained with PI and the mean fluorescence of *P. pinnata* and standard nuclei recorded. This experiment represented a 100-fold dilution of sample. In normal practice, 13 mm square leaf segments would typically be processed in about 1.0 ml buffer.

6.2.6 Effect of antioxidant on flow cytometric estimation of nuclear DNA content

An additional test was performed to study the effect of presence of antioxidant(s) in the buffer on the estimation of the nuclear DNA content of *P. pinnata*. For establishing the concentration and combination of antioxidants that stabilize PI fluorescence in leaf samples, a test for inhibitor was repeated using Propidium iodide/hypotonic Citrate method of Awtar Krishan (1975) with minor modification having 1.0 % w/v polyvinylpyrrolidone (PVP-40; buffer-A), 1.5 % w/v (buffer-B), 2 % w/v PVP and 15 mM 2-mercaptoethanol (buffer-C). Analysis was performed in replicates. After optimizing the nuclear DNA content estimation protocol for *P. pinnata,* the flow cytometric estimation of five individuals of *P. pinnata* were carried out. To indicate the presence of staining inhibitors in cytosol, six channels scale (0, 200, 400, 600, 800 and 1000) on histogram of flow cytometry was considered while analyzing the samples.

6.2.7 Statistical analysis

One-way analysis of variance (ANOVA) was carried out as described by Ott & Longnecker (1998) in order to test the effect of time on nuclear staining, presence of inhibitors, effect of slurry containing *P. pinnata* nuclei and effect of antioxidants on the nuclear DNA content estimation to know whether significant differences existed between them. Multiple comparison Tukey HSD and Duncan tests were applied where necessary to determine exactly which groups were different ($P < 0.05$).

TH-1255_09610610

6.2.8 Cytological analysis

6.2.8.1 Collection and storage of root tips

Mature (dried) seeds of *Pongamia* were allowed to sprout on moist soil in pots for 3-4 weeks in the departmental green house. When the plants were 3-5 cm tall, was transferred from soil to hydrophonics to obtain fresh root tips (Figure 6.4). After 3-4 days 1-1.5 cm of fresh roots tips were collected from the roots between 8.30-10.00 AM. The root tips were washed for few minutes with tap water and finally with distilled water and soaked on a filter paper. A total of 10 replicates of each sapling from different places from Assam were collected and fixed for the chromosome counts.

6.2.8.2 Pre-treatment and fixation

The collected roots of *P. pinnata* were pre-treated with saturated aqueous solution of ρ-dichloro benzene (PDB) [10 g of PDB was used to prepare saturated PDB solution] and kept at 16 ^{o}C for 3 h. The pre-treated material was washed with distilled water for several times, soaked on a filter paper. The root tips were then fixed in freshly prepared Carnoy's solution (Ethyl alcohol and glacial acetic acid in 3:1 ratio), incubated for 6 h at room temperature. After six hours, the root tips were washed with 70 % alcohol for 3-4 times and preserved in 70 % alcohol for future use.

6.2.8.3 Hydrolysis and staining

Fixed root tips were hydrolysed in 1 N HCL at 58 ^{o}C. Six different times (1, 2, 3, 4, 5 and 6 min) were tried for optimization of hydrolysis. Thereafter, the root tips were immersed for 2 h in staining solution [(Acetic acid-orcein 45 % (Darlington & Lacour 1976)]. The root tips turned black once stained.

6.2.8.4 Slide preparation, squash and observation

The root tips stained previously were taken out on a clean microscopic slide using a brush. The root was cut gently into small pieces towards the root tip, with a razor blade and dissecting needle. Few drops of 45 % glacial acetic acid was added onto it and macerated subsequently. A cover slip was placed over the tissue to overcome air bubbles. The cover slip was pressed down firmly by the thumb.

TH-1255_09610610

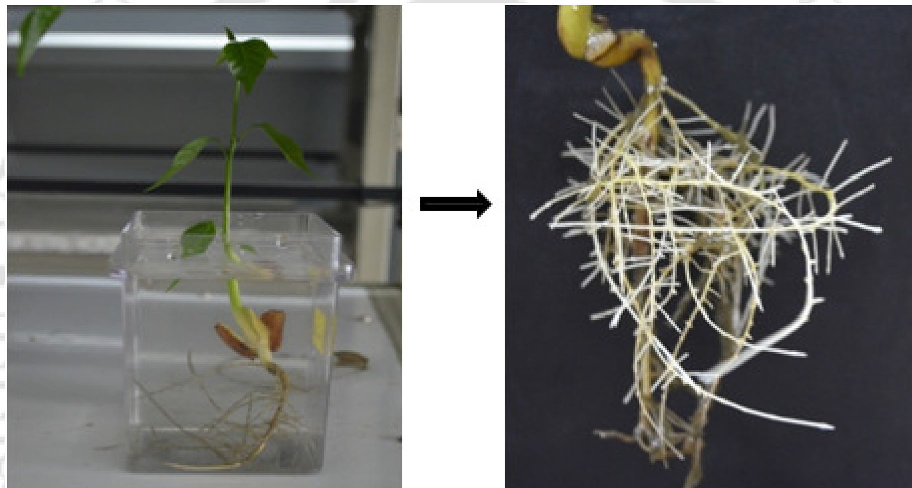


Figure 6.4: Maintenance of *Pongamia* plant in hydroponics culture.

TH-1255_09610610

Excess of acids were soaked with filter paper. The glass was heated over the flame for few times just to soften the tissue and then squashed by pressing the slide firmly between sheets of blotting paper. With a small cork stick or pencil headed eraser the cells were spread into a monolayer. After the cells were spread uniformly, the slide was observed under the compound light microscope at different resolutions. Three replicates of each slide for each species were tested to avoid the potential errors of chromosome counting as well as taxonomic ambiguity.

Each chromosome length was also measured in µM with Q-captured Pro-software, Olympus and average length of total chromosome in each sapling was compared with nuclear genome content.

6.3 RESULTS AND DISCUSSION

In this chapter we report the nuclear DNA content of five individuals of *P. pinnata,* from Assam, by using FCM. Efforts were also put to see the nuclear genome variation between different tissues types, saplings collected from various parts of India and also between CPT and micropropagated plants. The nuclear genome estimation followed here is relative but not absolute. It requires suspension of intact nuclei stained with DNA specific fluorochrome prior to analysis. Mitotic chromosome analysis along with chromosome length was also performed in five *Pongamia* individuals.

6.3.1 Test for robustness

Initial experiments were performed to confirm the robustness of the reference standards used. The reference standards procured from Dolezel lab was maintained under greenhouse conditions and the tender leaves were co-chopped for FACS analysis. The Mean fluorescence intensity (MFI) values (Figure 6.5) of both references were taken into consideration to calculate the nuclear genome content. The nuclear genome content obtained was matching with the existing reported results; hence the procedure followed was taken to be correct and preceded further to estimate nuclear genome content in *P. pinnata*.

TH-1255_09610610

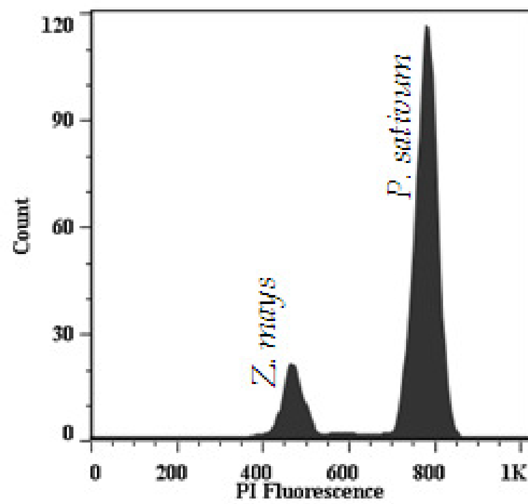


Figure 6.5: Test for robustness (Internal standard).

TH-1255_09610610

6.3.2 Reference standards for flow cytometry analysis

An ideal DNA reference standard should have a genome size close to the target species. This avoids the risk of nonlinearity and offset errors (Bagwell et al 1989; Vindelov et al 1983). The standard should be genetically stable with constant genome size, easy to use, and available in sufficient quantities. Similarly the genome size of the standard should be known with sufficient precision. These requirements are hard to satisfy and, as a result, different authors have used different standards, including human (Lysak et al 2000), domestic chicken (Galbraith et al 1983) as well as various plant species such as *Petunia integrifolia* (Marie & Brown, 1993), *Medicago sativa* (Martel et al 1997), *Avena sativa* (Morgan et al 1995), *Glycine max*, *Pisum sativum* and *Zea mays* (Dolezel & Greilhuber 2010). Clearly, there is a need for agreement on reference standards for DNA flow cytometry. As the genome size in plants ranges over 1000-fold (http://www.rbgkew.org.uk), a set of reference standards is needed with genome size distributed at appropriate intervals (Dolezel 2005).

In this experiment the plant reference standards like *S. lycopersicum* (2C=1.96 pg/2C), *Z. mays* (2C=5.43 pg/2C) and *P. sativum* (2C=9.09 pg) were taken into consideration. The main reason for choosing the above plants as reference standards was its short germination time and easily maintained at green house conditions. Among the three close reference plants mentioned above, *S. lycopersicum* gave overlapping peak with *P. pinnata*, suggesting that G0/G1 peak of tomato and *P. pinnata* are of the same order of magnitude (close to 2.0 pg). On account of overlapping nature, *S. lycopersicum* with *P. pinnata*, *S. lycopersicum* was ruled out as standard for estimation of nuclear DNA content of *P. pinnata*. Histogram G0/G1 peak of *Z. mays* appeared around 800 channel which is about two-folds to that of *P. pinnata* (G0/G1 peak at 400 channel) and appeared to be the most suitable internal reference standard plant for the estimation of nuclear DNA content in *P. pinnata* (Aadi et al 2013). The combination of *Z. mays* with *P. sativum* and *Z. mays* with *G. max* has been reported by Dolezel et al (1998), in his experiments use of *Z. mays* as reference standard strongly supports the current study. Although *P. sativum* gave clearly demarcated G0/G1, G2/M and standard G0/G1 peaks, which were well separated in the dynamic range still was not selected for this study as a reference standard since its nuclear DNA content is more than 4 fold of that of test sample (*P. pinnata*). According to the Suda & Leitch (2010) a reference

TH-1255_09610610

standard should have a genome size larger than that of the sample, but not more than four times larger. Inspite of overlapping nature of the G0/G1 peak of *Z. mays* and G2/M peak of *P. pinnata,* we have considered the histogram on account of the sharp, better resolution of peak, of high nuclei count of standard (*Z. mays*) for relative quantification of nuclear DNA amounts in pg (Figure 6.6A, B).

6.3.3 Test for inhibitors

P. pinnata leaves contain substance(s) that inhibit PI fluorescence of standard leaf nuclei. This is apparent in the flow histogram shown in Figure 6.7. Peaks i) and ii) represent nuclei of *P. pinnata* and standard leaves that were simultaneously processed prior to PI staining. Peak iii) is nuclei from independently processed and PI-stained standard leaves that were added and measured after 2 h. The differences between the two standard plant nuclei peaks, one inhibited (ii) and the other not inhibited (iii) were relatively stable over time. Table 2 presents a time course of the ratio of the mean fluorescence of nuclei of standard leaves, simultaneously processed and stained with *P. pinnata* leaves, to the mean fluorescence of added PI-stained nuclei from independently processed standard leaves. In 2 h after mixing, the ratio of mean fluorescence intensity was stable for the standard *Z. mays* and there was a progressive and significant decrease in the ratio from 0.876-0.818.

Table 6.2: Time course of the ratio of the mean (± SD) fluorescence of nuclei (standard 1) of standard leaves simultaneously processed and stained with *P. pinnata* leaves to the mean fluorescence of added PI stained nuclei (standard 2) from independently processed standard leaves.

Time after mixing	*Z. mays*1/ *Z. mays*2[n]
0	0.876±0.017[a]
15	0.855±0.002[a,b]
30	0.852±0.014[a,b]
45	0.864±0.011[b]
90	0.832±0.007[c]
120	0.818±0.009[c]

[n]Significant difference was observed among the means of *Z. mays*1/*Z. mays*2. (Means with the same letters are not significantly different)

TH-1255_09610610

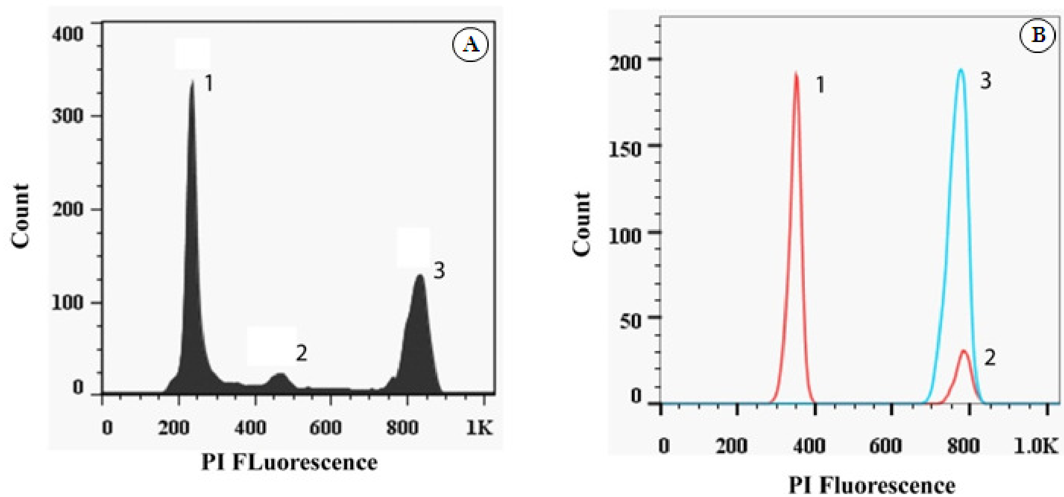


Figure 6.6: Appearance of G2 peak of *P. pinnata* in internal standardization. (**A**) *P. sativum* with *P. pinnata*; (**B**) *Z. mays* with *P. pinnata*. [1, 2 correspond to G0/G1 and G2 peak of *P. pinnata* and 3 correspond to G0/G1 peak of standards used.

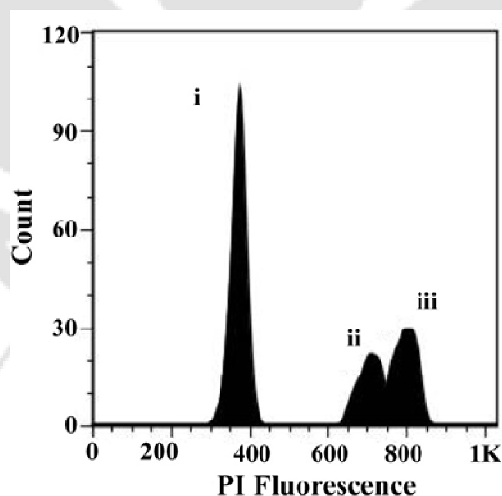


Figure 6.7: Histogram of PI stained nuclei from simultaneously processed *P. pinnata* (i) and standard (ii) leaves to which PI stained nuclei from independently processed standard leaves (iii) were added. PI fluorescence was measured 2 h after mixing the samples. The standard nuclei are Z*. mays* `CE-777'.

TH-1255_09610610

The current research provides evidence for the presence of inhibitor(s) of PI intercalation and/or fluorescence in *P. pinnata* leaves. The exact mode of action of the inhibitor(s) is unknown. Phenolic substances mostly possess active hydroxyl groups (providing free electron capable of forming hydrogen bonds). In the reduced state, these phenolics often show little or no colour and form hydrogen bonds (reversible) with the carboxyl group of DNA (Greilhuber et al 2007). But when hydroxyphenols are oxidized, a quinone structure is formed which often results in browning of the compound. Such quinines are highly reactive species themselves and form covalent bonds (irreversible) with carboxyl groups which affect the fluorescent dye accessibility to the DNA. β-mercaptoethanol, an antioxidant resists hydroxyphenols to form quinone structure (Endres 1961). These hydroxyphenols in turn form hydrogen bonds to DNA which can be maintained in reversible state and disrupted by the addition of a competitor PVP. Amide groups of PVP are available for binding with inhibitors, in competition with those of DNA. Combining PVP and antioxidants in nuclear isolation buffer helps to strip-off the phenolics from nucleus before they become oxidized and form covalent bonds as soon as they are released after chopping (Figure 6.8). In addition, the buffer contains citric acid that may help to maintain the nucleus integrity. The low average CV (coefficient variation) value associated with the fluorescence peaks of G0/G1 nuclei of this study indicates that a critical number of nuclei were reached during the extraction procedure with the buffers used. Fluorescence distributions wit high resolution and low CVs were also found using other buffers containing citrate (Dolezel & Bartos 2005).

TH-1255_09610610

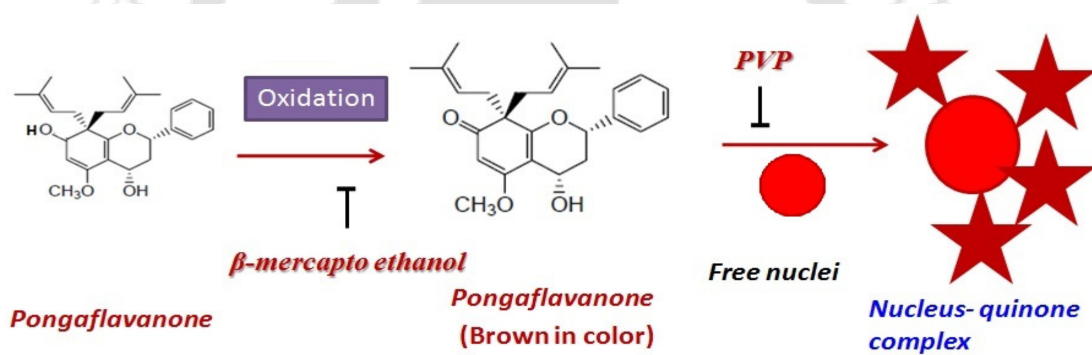


Figure 6.8: Effect of β-mercaptoethanol and PVP on PI binding.

TH-1255_09610610

6.3.4 Pseudo-internal standardization

The effect of concentration of inhibitor(s) was tested by processing same size standard leaf sections in varying concentrations (40 fold range) of buffer (containing independently processed *P. pinnata* leaves) prior to PI staining. With decreasing concentration of *P. pinnata* extract, the PI fluorescence of *Z. mays* increased. However, the fluorescence of standard *Z. mays* nuclei in the most diluted solution increased to only 1.71 % of that of the control (*Z. mays*) nuclei. There was a direct proportional increase in the nuclear DNA content estimated with increasing volume of *P. pinnata* extract by taking *Z. mays* as standard (Table 6.3). The variation of nuclear DNA content by pseudo-internal standardization was found to be 1.28 fold ($2C_{max}$ = 2.58 pg; $2C_{min}$ = 2.01 pg) (Figure 6.9).

Table 6.3: PI fluorescence of nuclei from mixtures of independently chopped *P. pinnata* and standard leaf samples.

Composition	Mean PI fluorescence		*P. pinnata*/ *Z. mays* ratio	*P. pinnata* DNA content (pg)[n]
	P. pinnata	*Z. mays*		
2 ml *P. pinnata* 0.5 ml Standard leaf nuclei	374.60±5.03	786.60±5.03	0.48	2.58[a]
0.5 ml *P. pinnata* 0.5 ml buffer 0.5 ml Standard leaf nuclei	364.30±6.11	873.30±6.10	0.42	2.26[b]
0.25 ml *P. pinnata* 1.75 ml buffer 0.5 ml Standard leaf nuclei	360.60±4.04	921.60±6.70	0.39	2.12[c]
0.05 ml *P. pinnata* 1.95 ml buffer 0.5 ml Standard leaf nuclei	339.60±8.54	917.00±8.50	0.37	2.01[d]
0 ml *P. pinnata* 2.00 ml buffer 0.5 ml Standard leaf nuclei		928.30±7.20		

[n]Multiple range test groupings (P<0.05) means with the different letters are significantly different

TH-1255_09610610

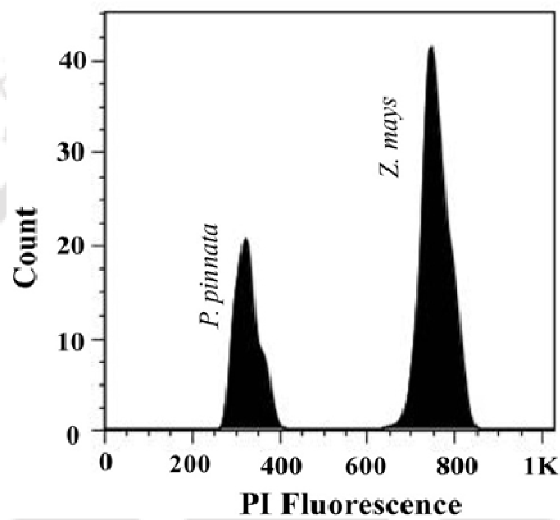


Figure 6.9: Pseudo internal standardization process, where samples are chopped individually and mixed while running with Flow cytometer.

TH-1255_09610610

6.3.5 Internal standardization

Testing the effect of inhibitor dilution on PI fluorescence involved co-processing of *P. pinnata* and standard leaves in buffer volumes ranging over a 100 fold range prior to PI staining (Figure 6.10). In this experiment the ratio of the means of *P. pinnata* and standard nuclei fluorescence remained similar when leaves were processed in 0.1, 1.0 and 10.0 ml buffer. On the contrary, the PI fluorescence of *Z. mays* increased significantly with the increasing buffer volume. But there was a progressive decrease of *P. pinnata* nuclear DNA content with increased buffer volume (Table 6.4).

Table 6.4: Mean PI fluorescence of nuclei from standard leaf nuclei and *P. pinnata* leaves simultaneously processed in different volumes of buffer.

Buffer volume	Mean PI fluorescence		***P. pinnata*/ *Z. mays* ratio**	***P. pinnata* DNA content (pg)**
	P. pinnata	***Z. mays***		
Z. mays + *P. pinnata*				
0.1 ml	307.6±2.5	648±2.6	0.47	2.57
1.0 ml	369.7±5.7	874±13.4	0.42	2.29
10.0 ml	376.0±5.0	892±8.8	0.42	2.29
Z. mays controls				
0.1 ml		838.0±2.6		
1.0 ml		859.0±2.0		
10 ml		894.0±6.2		

The variation of nuclear DNA content by internal standardization was found to be 1.13 fold ($2C_{max}$ = 2.58 pg; $2C_{min}$ = 2.29 pg). Taking into consideration minimum shift of MFI of the standard (*Z. mays*), the nuclear DNA content estimated for pseudo-internal standardization, internal standardization was 2.58 pg and 2.29 pg respectively (Table 6.5).

Table 6.5: DNA content of *P. pinnata* estimated by internal and pseudo-internal standard.

Standard species	***P. pinnata* genome size Internal standardization**	***P. pinnata* genome size Pseudo-internal standardization**
Z. mays	2.29±0.03	2.58±0.02

TH-1255_09610610

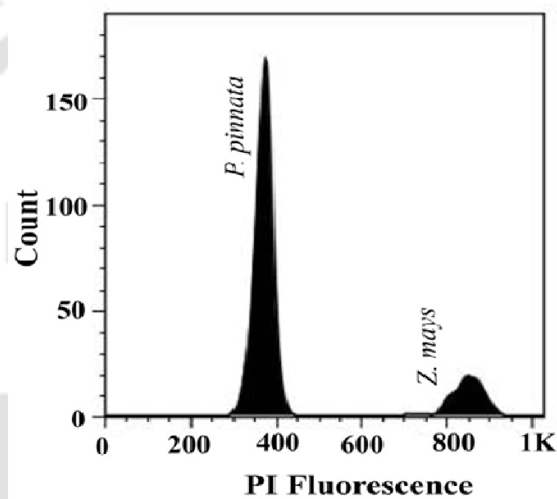


Figure 6.10: Internal standardization with *Z. mays.*

TH-1255_09610610

It is apparent that the simultaneous processing of tissue of the target and standard species is absolutely necessary to obtain reliable DNA content estimates. The study also establishes that compounds in *P. pinnata* leaves greatly inhibit the PI fluorescence of the standard nuclei. However, the proportion of plant species with natural inhibitors of PI intercalation or fluorescence remains to be determined. It is likely that naturally occurring inhibitors that decrease fluorochrome fluorescence of plant nuclei are common as has been found in our standardization methods of *P. pinnata*.

The presence of inhibitors compromises the reliability of estimated nuclear DNA content. Combined processing of the target and standard species may minimize but not entirely eliminate the effect of inhibitors on estimated nuclear DNA content. The practice of measuring mixed, independently-processed and stained nuclei from standard and target species (O'Brien et al 1996; Rayburn 1997) is reported to give greatly exaggerated differences in fluorochrome fluorescence that may not be due to differences in nuclear DNA amount (Price et al 2000). This is not evident in case of *P. pinnata* when comparing the small differences between estimated nuclear DNA amounts in internal, pseudo-internal standardization. The FL2A vs. SSC showed the presence of additional particle population similar to "tannic acid effect" (Loureiro et al 2006). They most probably resulted from aggregates of nuclei or parts of nuclei with unspecific particles and diverse particle devoid of nuclei (inhibitors, antioxidants) with PI stain.

6.3.6 Effect of antioxidants

Shift of G0/G1 peak of the internal standard at least six channels (0, 200, 400, 600, 800 and 1000) on the 1000 channel scale of histogram was considered as an indicative of the presence of staining inhibitors in the cytosol of the cell of target species. In the present research, a test for the presence of staining inhibitors clearly showed that the leaves of *Pongamia* contained compounds that biased the fluorescence of the internal standard nuclei. To avoid conducting instrument calibration between analyses of different samples, the buffers marked as A, B and C (containing either PVP alone or β-mercaptoethanol) was used throughout the experiments.

TH-1255_09610610

The percentage range of shift in the MFI of internal standardization of *Z. mays is* 0.4 %- 30 %. The percent shift of MFI of *Z. mays* channel was in the range of 5 % - 7 % by using antioxidants. So it was found that by usage of antioxidants, the *P. pinnata* nuclear DNA content varied from 2.72 pg (buffer A), 2.66 pg (buffer B) and 2.17 pg (buffer C) (Figure 6.11 and Table 6.6).

Table 6.6: Effect of antioxidants on flow cytometric estimation of 2C DNA content in *P. pinnata*.

Species name		***P. pinnata* 2C DNA in pg (mean ± SD) *Z. mays*[n]**
Buffer without antioxidants		2.29 ± 0.01[a]
Buffer with antioxidants	Buffer A	2.72 ± 0.01[b]
	Buffer B	2.66 ± 0.01[b]
	Buffer C	2.17 ± 0.00[c]

[n]Multiple range test groupings (P<0.05) means with the different letters are significantly different (Duncan one way ANOVA test, $P > 0.05$).

The fact that a straight line did not intersect the origin clearly showed that the *P. pinnata* cytosol effect of nuclei fluorescence differed between the test sample (*P. pinnata*) and that of controls (*Z. mays*). Theoretically, the nuclei fluorescence of *P. pinnata* should be strictly proportional to *Z. mays* nuclear fluorescence. But the theory failed as the regression between sample and standard fluorescence differed in terms of their slopes and did not go through the origin. In addition, the slope differences highlighted differential effect of *P. pinnata* cytosol on standard nuclei. The variation of nuclear DNA content in the regression equation of internal standardization ($y=0.278x+127.1$; $R^2=0.994$) and pseudo-internal standardization ($y=-0.169x+507.7$; $R^2=0.438$) is also shown in Figure 6.12. The regression value is higher for internal standardization.

Although nuclear DNA content estimated in the buffer supplemented with antioxidant was higher compared to that without supplementation, we stressed on the factor of the best fit curve in the internal standardization of *P. pinnata*, *Z. mays* ($R^2=0.99$) and the least shift of MFI (4.5 %) obtained by addition of antioxidant (buffer B) to the nuclear isolation buffer (Figure 6.13). So, the nuclear DNA content estimated by the buffer B would be the best for accurately determining the nuclear DNA content of *P. pinnata* considering *Z. mays* as standard.

TH-1255_09610610

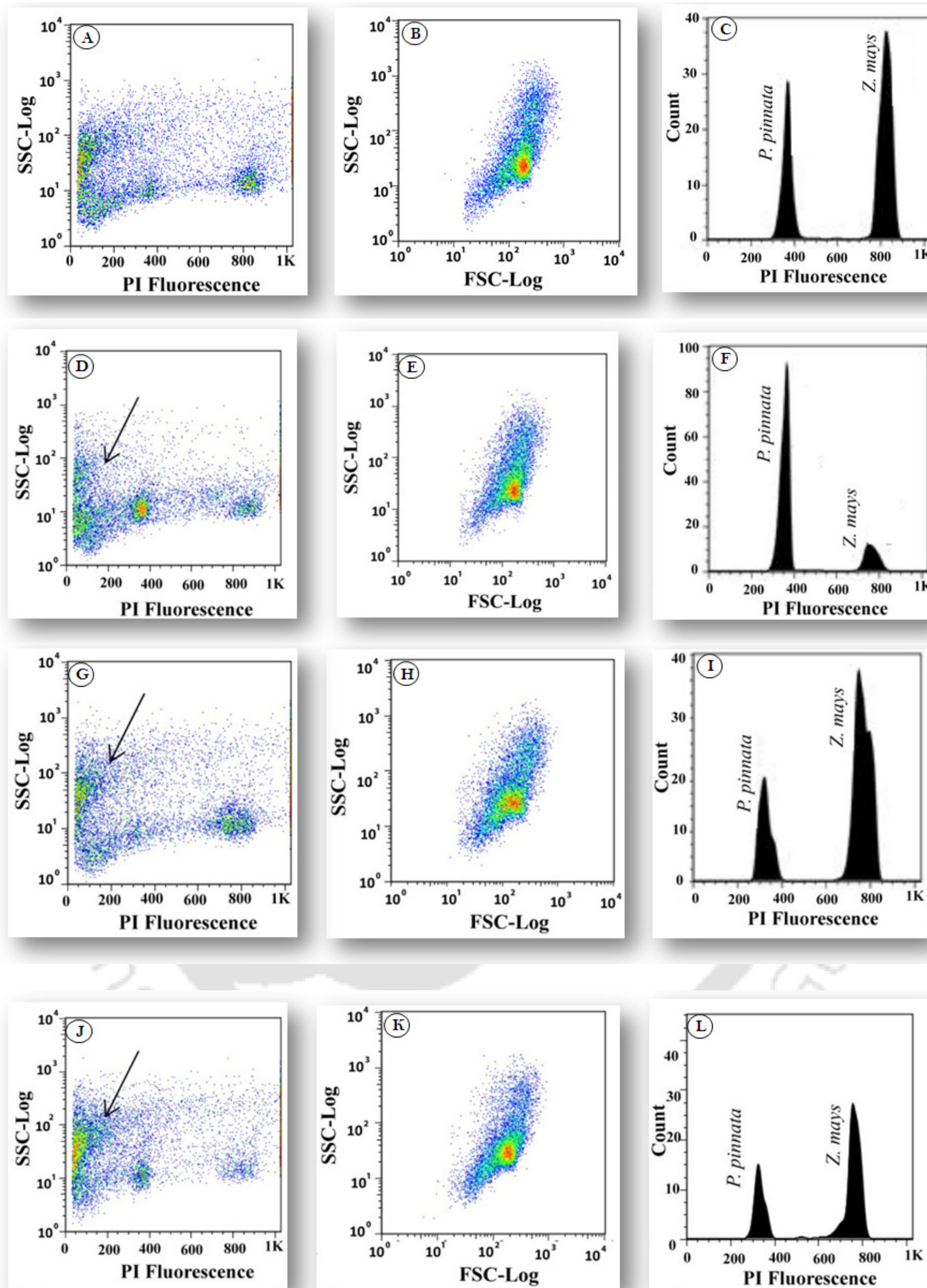


Figure 6.11: Effect of antioxidants on PI fluorescence and light scattering properties. Nuclei suspensions were obtained with PI/hypotonic citrate buffer of Awtar Krishan (A-C) and PI/hypotonic citrate buffer of Awtar Krishan supplemented with buffer-A [1 % PVP] (D-F), buffer-B [1.5 % PVP] (G-I) and buffer-C [2.0 % PVP and 15 mM β-mercaptoethanol] (J-L). Arrows indicate fluorescence particle without nuclei.

TH-1255_09610610

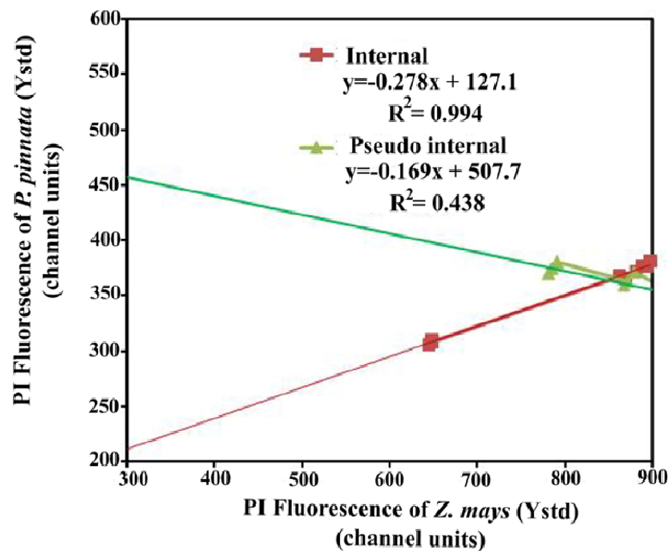


Figure 6.12: Comparison of the standard nuclei peak location and the *Pongamia* peak location. Each line shows different standardization methods of *P. pinnata* and *Z. mays*.

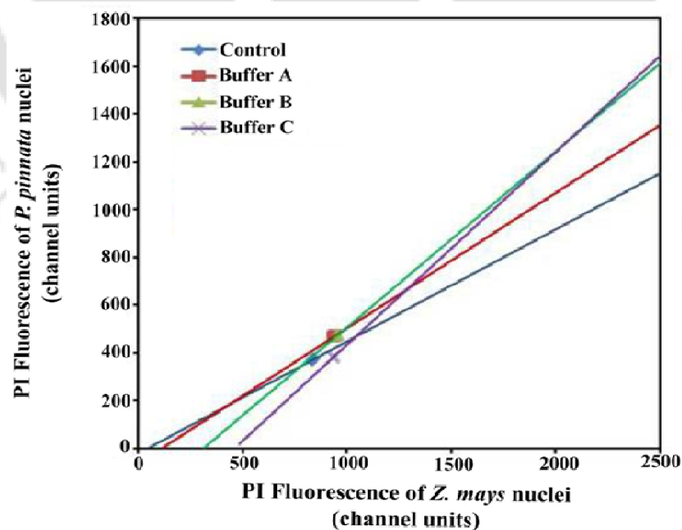


Figure 6.13: Effect of antioxidant on *Z. mays.*

TH-1255_09610610

In order to maintain the perfect quality and good stability of the cell nuclei, we tested five buffers [Galbraith's buffer (Galbraith et al 1983), LB01 (Dolezel et al 1989), Propidium iodide/hypotonic Citrate (Krishan 1975), General Purpose Buffer (Loureiro et al 2007) and Woody plant buffer (Loureiro et al 2007)] for preparation of the nuclei sample. The best results in the present work for the isolation of nuclei were obtained with sodium citrate buffer containing detergent resulting in low coefficient of variation (CV) (Krishan 1975) as other buffers fail to complete separation of nucleus from cytosol which lacks sodium citrate that helps in release and precipitating the nuclear DNA. The non-ionic detergent IGEPAL CA-630 was used to facilitate the release of nuclei from cell and prevent nuclei clumping and attachment of debris. The suspension of plant nuclei often contains high level of calcium oxalate, which blocks the fluidics system of flow cytometer. To prevent this, sodium citrate was used as a chelating agent. PI is a DNA selective fluorochrome without base dependent bias which intercalates into the double stranded DNA (ds RNA) and this is suitable for estimation of nuclear DNA content in absolute units provided the RNA is removed with RNase A (Dolezel et al 2003).

Earlier studies indicate that antioxidants such as PVP and 2-mercaptoethanol that are used in chopping buffers have no measurable effect on reducing the interference of secondary metabolites with the staining of DNA with PI (Bharathan et al 1994). On contrary, in the current study we have found that the usage of antioxidants resulted in consistency of nuclear DNA content of *P. pinnata* against *Z. mays* as an internal standard. A gradual increase in the MFI of *Z. mays* was evident with increase in the volume of the buffer. This unlikely observation may be due to decrease in DNA condensation of *Z. mays* with increasing buffer amount. Ideal internal standards display minimal variability, match as closely as possible to the configuration of sample DNA, and have a genome size larger than (not >4 times) the sample (Baranyi & Greilhuber 1995; Suda & Leitch 2010). Our best estimate for the nuclear DNA content (2C) of *P. pinnata* used in the current study is 2.66 pg by usage of antioxidants (buffer B) providing the best fit between the fluorescent intensity of *P. pinnata* and *Z. mays* on the regression curve.

TH-1255_09610610

6.3.7 Test for homogeneity

The genome content in different tissue types [flower, root, leaf, shoot, stem and different seed development stages (90, 180, 270, 350-DAF)] were also estimated and were found to be in the range of 2.25-2.68 pg (Table 6.7 and Figure 6.14). The decrease in the genome content of flower may be due to the presence of floral pigments (anthocyanins) and other secondary metabolites which may interfere with nucleic acids that will reduce PI binding efficiency to the nucleic acids. Recently, Loureiro et al (2006) showed a strong effect of tannic acid in four isolation buffers on relative fluorescence intensity of PI stained nuclei of *P. sativum* and *Z. mays*, which is greatly reduced PI binding efficiency to the nucleus. The genome size in seed tissues is slightly higher than leaf tissues and the similar kind of results were reported in *Helianthus annuus* and *Brassica napus* as higher 2C value was observed in the seeds than those in the leaves (Sliwinska et al 2005). Probably the material studied here contained interfering staining inhibitors that they could not be completely eliminated by addition of antioxidants, and thus the overestimated values were due to the reduction of the fluorescence of the internal standard nuclei. Another explanation of the differences between the 2C values in different tissues could be due to a different chromatin structure in those organs. Differences in the chromatin condensation were previously observed in different tissues of *Z. mays* (Baluska 1990; Biradar & Rayburn 1994).

Table 6.7: Nuclear genome content of different tissues in *Pongamia*.

Tissue/organ type		**Nuclear genome content (pg) (Mean ± S.E)**	**DNA Index**	**Coefficient of variation**
Vegetative	Root	2.68±0.032	0.46	4.65
	Shoot	2.66±0.008	0.44	4.95
	Leaf	2.64±0.006	0.43	4.13
	Stem	2.62±0.007	0.42	4.20
Reproductive	Flower	2.25±0.027	0.41	4.40
	90-DAF	2.68±0.011	0.46	4.55
	180-DAF	2.67±0.007	0.46	4.23
	270-DAF	2.66±0.012	0.46	4.10
	350-DAF	2.64±0.006	0.43	4.30

Three replicate processing and measurements was noted for each plant.

TH-1255_09610610

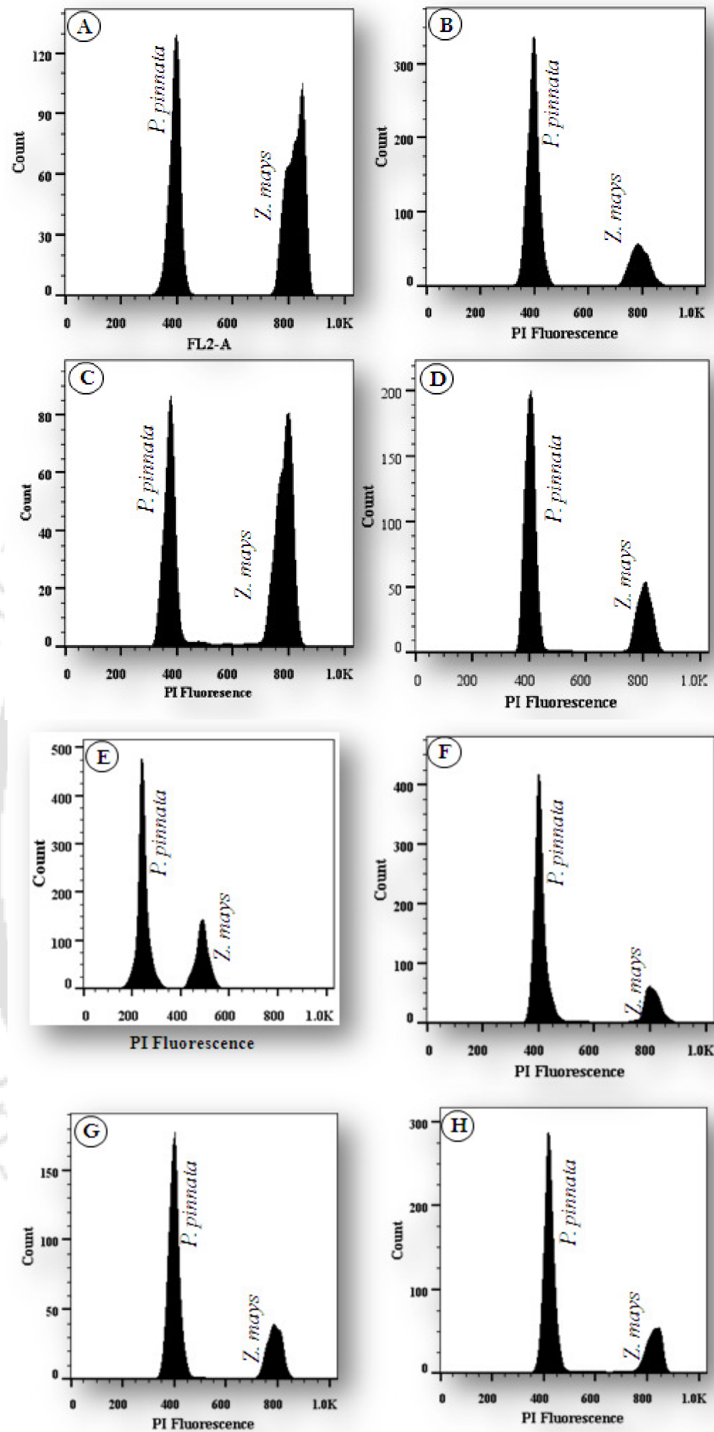


Figure 6.14: Nuclear genome content estimation in different tissue of *Pongamia*: A) root; B) shoot; C) leaf; D) stem; E) flower; F) 90-DAF; G) 180-DAF; H) 350-DAF.

TH-1255_09610610

The nuclear genome content of *Pongamia* collected from various parts of India (Banaras, Lucknow, Gorakhpur, Bhubaneswar, Hyderabad and Amravati) was also estimated and it was found to be in the range of 2.50 to 2.68 pg (Table 6.8 and Figure 6.15). The genome size between the tissues and among individuals collected from various locations was found to have slight variation. Intraspecific variation in nuclear genome content has been well studied in legume species (Walker et al 2006). Molecular investigations on plant nuclear DNA content have shown that mostly genome size variability is associated with increase in leaf furanocoumarins and differences in repetitive DNA content and chromatin condensation of the plant genome (Baluska 1990; Biradar & Rayburn 1994; Flavell et al 1974; Jeffrey et al 2005; Temsch et al 2010; Walker 2006).

Table 6.8: Genome content of *P. pinnata* saplings collected from various parts of India.

S. No	Location	Coordinates	Nuclear genome content (pg) (Mean ± S.E)	DNA index	Coefficient of variation
(a)	Banaras	25.282°N 82.9563°E	2.58±0.002	0.40	4.60
(b)	Lucknow	26.847°N 80.947°E	2.68±0.010	0.45	4.35
(c)	Gorakhpur	26.7588°N 83.3697°E	2.56±0.006	0.42	4.15
(d)	Bhubaneswar	20.27°N 85.84°E	2.62±0.032	0.44	4.28
(e)	Hyderabad	17.366°N 78.476°E	2.50±0.003	0.41	4.09
(f)	Amravati	21.15°N 79.09°E	2.60±0.022	0.44	4.34

Three replicate processing and measurements was noted for each plant.

After optimizing the protocol for nuclear DNA content estimation of *P. pinnata*, the experiment was carried out in three replicates for CPT (NGPP46) and individuals collected from different locations of India. Clearly defined histograms for accurate determination of nuclear DNA contents were obtained following flow cytometric analysis of intact leaf nuclei of *P. pinnata* individuals collected from different locations in India. According to the comparison of their G0/G1 peak values, the nuclear DNA content of *P. pinnata* was 0.49 times of that of *Z. mays*. Therefore it could be estimated that the nuclear DNA content of *P. pinnata* to be 2.66 pg and the 1-C value to be 1300 Mb in all the five individuals collected from various parts of India.

TH-1255_09610610

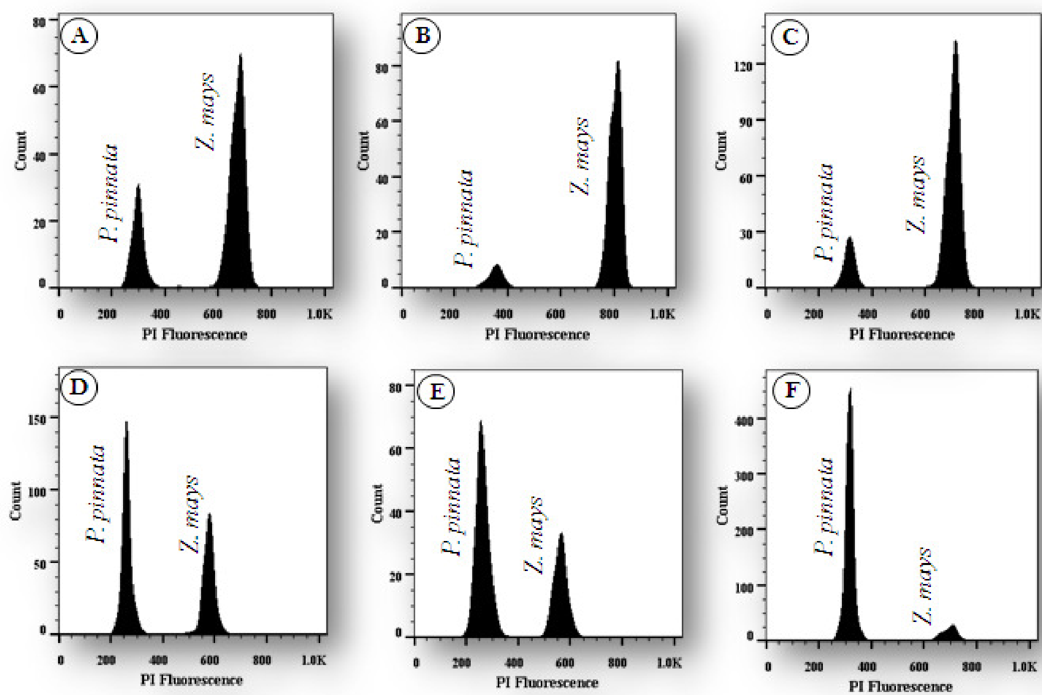


Figure 6.15: Histogram of mean fluorescence intensity in different individuals of *P. pinnata* collected from various parts of India. A) Banaras; B) Lucknow; C) Gorakhpur; D) Bhubaneswar; E) Hyderabad; F) Amravati.

TH-1255_09610610

In the same way the nuclear genome was also estimated in *Pongamia* germplasm collected from different parts of Assam (Jorhat, Kokrajhar, Guwahati, Tezpur, Tinsukia) and various parts of India. Total nuclear DNA content varied among the individuals (collected from Assam) from 2.61 pg to 2.68 pg respectively (Figure 6.16). Also the nuclear DNA content was negatively correlated with the average temperature of the city (Pearson's product moment correlation coefficient, r=-0.89) and the nuclear DNA content was positively correlated with the height from the sea level (r=0.86). The highest nuclear DNA content (2C=2.68 pg) was observed for the Tinsukia individuals which is the located 117 m above the sea level with the average temperature of 22.2 °C. The lowest nuclear DNA (2C=2.61 pg) content was observed for the individuals belonging to Kokrajhar and Tejpur districts which are located 37 and 73 m above the sea level with an average temperature of 23 °C (Table 6.9 and Figure 6.17). Nevertheless, as genome size estimation varied with temperature due to differences in chromatin condensation and arrangements of repetitive sequence elements (Jedrzejczyk & Sliwinska 2010). Walker et al (2006) reported true intraspecific variation in nuclear DNA contents for legume *Bituminaria bituminosa* populations, analyzed by flow cytometry, their studies suggest that the temperature plays significant role in nuclear genome determination. According to Walker et al (2006) the environmentally induced variation within populations may have been due to interference of leaf furanocoumarins, which are known to accumulate to a greater extent at higher temperatures.

Table 6.9: Correlation between genome size, temperature and height of *Pongamia* individuals from Assam.

S. No	**Description**	**Temperature (°C)**	**Height from sea level (m)**	2C DNA content (pg)
01	Tinsukia	22.2	117	2.68
02	Kokrajhar	23	37	2.66
03	Guwahati	23	55	2.62
04	Jorhat	22	116	2.61
05	Tezpur	23	73	2.61

Three replicate processing and measurements was noted for each plant.

TH-1255_09610610

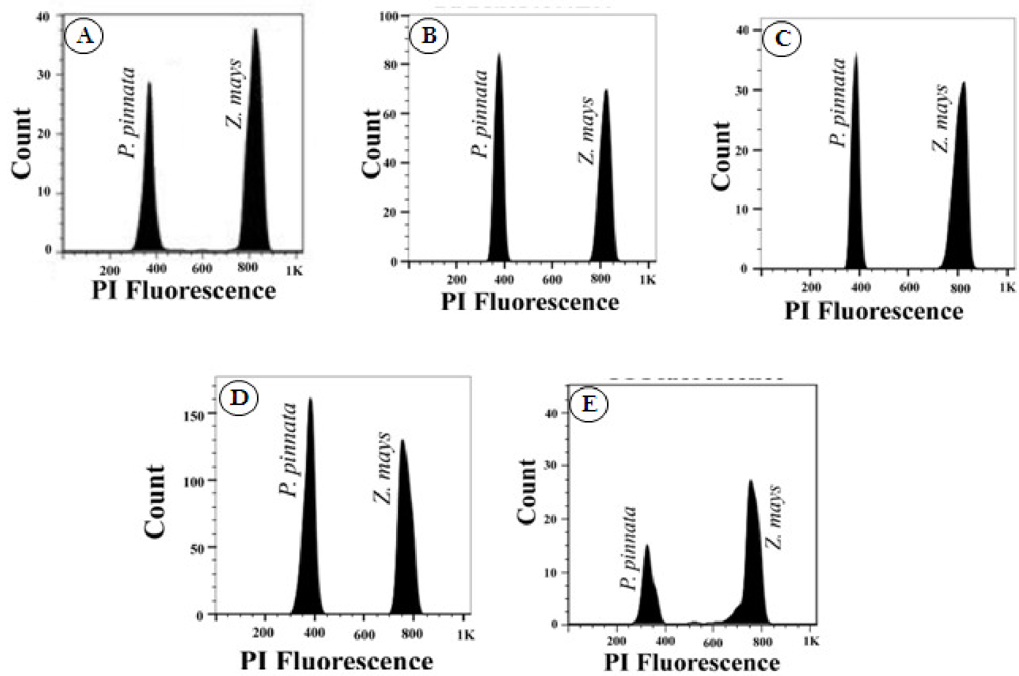


Figure 6.16: Histogram of mean fluorescence intensity in different individuals of *P. pinnata* (**A**) Tinsukia; (**B**) Kokrajhar; (**C**) Guwahati; (**D**) Jorhat; (**E**) Tezpur.

TH-1255_09610610

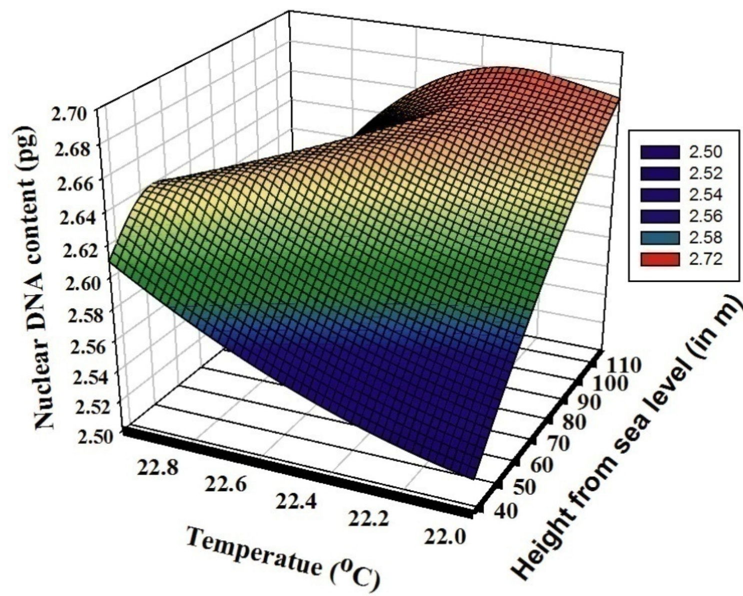


Figure 6.17: Correlation between genome size with geographical distribution of *P. pinnata* in Assam.

TH-1255_09610610

Efforts were also put to estimate the DNA content of candidate plus tree (CPT) and ploidy stability of *P. pinnata* plants grown *in vitro* conditions. The estimated nuclear DNA content of *P. pinnata* CPT was used for comparison with its *in vitro* raised plants (IRP) and *in vitro* raised field grown plants. Simultaneous analysis of nuclear DNA content of *P. pinnata* CPT, IRP, AVP (Acclimatized *in vitro* plants) and NS (Nodal segments) showed no significant difference at $\alpha=0.05$ (ANOVA: Standard *Z. mays*- $F_{3,32}=0.850$ and $P=0.477$) and were grouped together in the same group via TukeyHSD (Table 6.10). In addition, simultaneous analysis of mixtures of plant homogenates of CPT with each of the *in vitro* raised plants resulted in one single peak in the 2C range (Figure 6.18). This confirms that the little variation observed in the measurements is mainly due to sample handling and instrumental errors.

Table 6.10: 2C DNA content of *P. pinnata* estimated by using *Z. mays* as an internal reference standard.

Plant	***Z. mays* (Internal reference standard)**			
Tissue origin	N	*P. pinnata* DNA content (pg) (Mean ± S.E.*)	DNA Index	Coefficient of Variation
CPT	1	2.66±0.078[a]	0.45	4.95
In-vitro Rooted Plantlets	5	2.65±0.005[a]	0.44	4.50
Acclimatized *in vitro* plants	5	2.66±0.005[a]	0.44	4.12
Nodal Segments	1	2.66±0.006[a]	0.44	4.45

N= No. of plants analyzed (three replicate processing and measurements was noted for each plant).
* Multiple range test groupings (P<0.05) means with the different letters are significantly different.

TH-1255_09610610

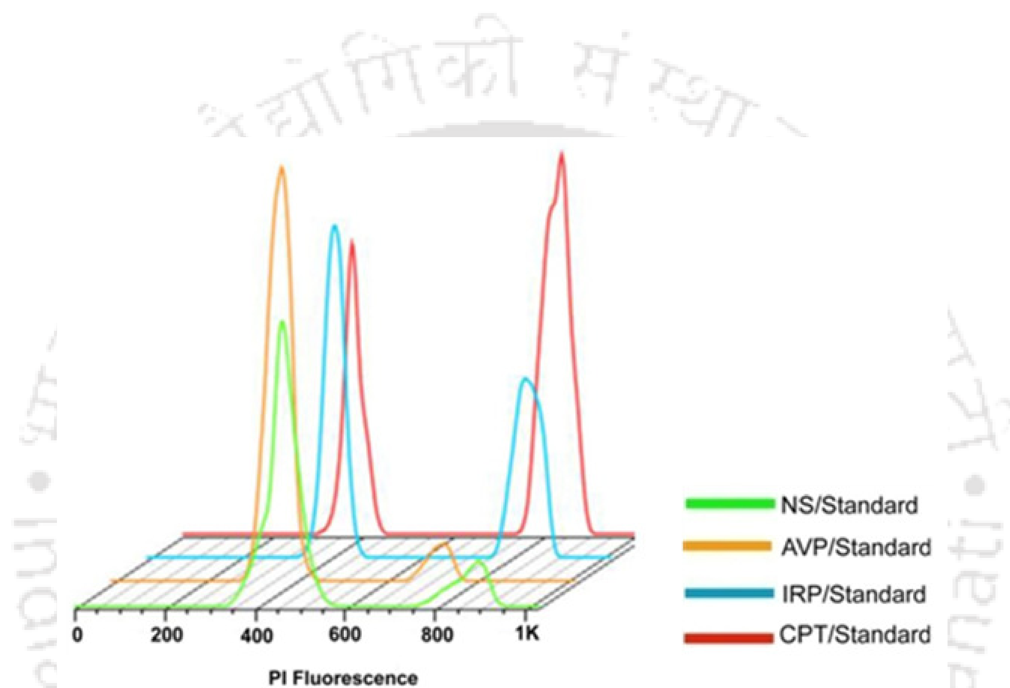


Figure 6.18: Histograms of relative DNA content of nuclei isolated from the nodal segments (NS), acclimatized *in vitro* plants (AVP), *in vitro* rooted plantlets (IRP) and NGPP 46 Candidate Plus Tree (CPT) with respect to internal standard (*Z. mays*).

TH-1255_09610610

The assessment of the genome size stability of *in vitro* derived clones is an essential step in the application of biotechnology for micropropagation of true-to-type clones. The explants source and mode of regeneration (somatic embryogenesis, organogenesis, and axillary bud multiplication) are known to play a major role in determining the presence or absence of variation. There are many reports of genome size stability studies in *in vitro* systems in various plants using FCM (Kubaláková et al 1997; Loureiro et al 2007; Pinto et al 2004; Thiem & Sliwinska 2003). We, in our earlier studies have used nodal segment as explants for micropropagation of *P. pinnata* because it lowers the risk of genetic instability (Kesari et al 2012). FCM analysis of 2C DNA content and ploidy level of *in vitro* raised plants of *P. pinnata* showed a profile similar to that of the CPT indicating genome size stability. Weber et al (2008) showed that cells of all *in vitro* cultures underwent more cycles of endoreduplication than the explants from which they were obtained. They also concluded that the ploidy profiles of the investigated *in vitro* cultures appeared to be influenced by the genome size, the growth factors applied, and the type of *in vitro* culture. In another study, Obae & West (2010) showed that *in vitro* regenerated (explant source was from wild plants) and wild plants of *H. canadensis* did not reveal any significant difference in nuclear DNA content. Sliwinska & Thiem (2007) reported genome size stability in six medicinal plant species that were propagated *in vitro*. The most noteworthy observation in the current study is that plants, whether derived from CPT or *in vitro* each showed the same ploidy level.

6.3.8 Cytological analysis

6.3.8.1 Collection of root tips

To check the ploidy level, chromosome counts in different individuals of *P. pinnata* collected from various parts of Assam was carried out. The best time of the day to collect root tips for karyological study was found to be 8.30-10.00 AM, where there are more cell metaphases. However the time was optimized for plants under study and accordingly growing root tips were collected between 8.30-9.00 AM. The collection period was found to give adequate numbers of metaphases in *Pongamia*.

TH-1255_09610610

6.3.8.2 Pretreatment

Pre-treatment is a necessary step in cytological studies for members of leguminous family. Different workers have used various pre-treatments and staining chemicals in their studies (Chen 1992; Das et al 1998; Joseph et al 1999; Nair & Sasikumar 2009; Rai et al 1997; West & Cowley 1993). This helps to block the mitosis cycle at metaphase stage by inhibiting spindle fiber formation revealing more number of cells in metaphase. PDB was effective in treating *Pongamia* samples to get metaphase chromosomes.

6.3.8.3 Hydrolysis and staining

Hydrolysis of the root tips is carried out to soften the root tissue. The roots of dicots are harder and larger size and therefore it is an important step to soften the pretreated roots for a specified period in 1N HCL. In the present study, the time was optimized as 5 min at 58 ^{o}C. Later the roots were allowed to cool down at room temperature and subjected to dye (2 % aceto-orcein or aceto carmine) in a test tube and heated over a flame for few seconds. The heating for a few seconds in the acid-dye mixture intensified the staining of the chromosomes and gave a clearer general picture of the chromosome structure.

6.3.8.4 Slide preparation, squash, and observation

The preparation of a good slide depends on the proper handling of the root material, cutting, maceration and squashing. Care was taken so that the tip of the root is not hurt. Adding of the acid onto the root tip slowly macerate it and also decolorize the cytoplasmic material leaving only the chromosomes stained. The squashing and spreading of the cells uniformly is a crucial step which results in proper visualization of the slide. The slides revealed well-spread metaphases under low power (10x) and high power (40x) magnifications in compound light microscope. Some cells were found to be in anaphase and prophase stages also. The ploidy level (2n=22) of *P. pinnata* matches with the findings of other workers (Patel & Narayan 1937; Sarbhoy 1977).

After successive collection, hydrolysis and squash preparation of the slide, the chromosome number was counted in all individuals collected from various locations of Assam and the somatic chromosome number was found to be 2n=22 (x/n=11), which is

TH-1255_09610610

consistent in all individuals. The chromosome number of *P. pinnata* matches with *Lablab niger* (2n= 20) (Bennett & Smith 1976), *Vigna lancifolia* (2n=22) (Parida et al 1990) belonging to family Leguminosae.

The chromosome length was measured and length ranged from 2.39-5.16 µM, 2.34-4.65 µM, 1.91-4.35 µM, 2.13-3.99 µM and 1.95-3.15 µM for Tinsukia, Guwahati, Jorhat, Kokrajhar and Tezpur respectively. The highest average chromosome length was observed for Tinsukia individual (3.69 µM) and lowest size was observed in Tezpur individual (2.70 µM). The study also focused on the relationship between average chromosome length and the DNA contents of five individuals collected from Assam. Results indicate that, though the measured chromosome length differed from individual to individual, there is positive correlation between nuclear DNA content and averaged chromosome length in all five individuals. This indicates that DNA amount increases proportionately with increasing mean chromosome length and that the amount of chromosome number is remains uniform in all five individuals from Assam. A similar relationship has been shown by Nagl and Ehrendorefer (1974) in the *Anthemidea* (Asterasceae family) and in *Allium* between DNA content and chromosome volume (Jones & Rees 1968). The results in this study is also supporting by the Previous reports on chromosomal DNA content in *Allium cepa, Allium fistulosum* (Narayan 1988) and nuclear DNA variation in 42 *Allium* species (Labani & Elkington 1987) revealed that the DNA contents were directly proportional to the chromosome lengths (Table 6.11 and Figure 6.19, 6.20).

Table 6.11: Chromosome number and its length of *Pongamia* individuals collected from Assam.

Collection	Chromosome number	Chromosome length (µM)	Avg. length of chromosome (µM)
Tinsukia	22	2.39-5.16	3.69
Guwahati	22	2.34-4.65	3.30
Jorhat	22	1.91-4.35	3.20
Kokrajhar	22	2.13-3.99	2.77
Tezpur	22	1.95-3.15	2.70

Three replicate processing and measurements was noted for each plant.

TH-1255_09610610

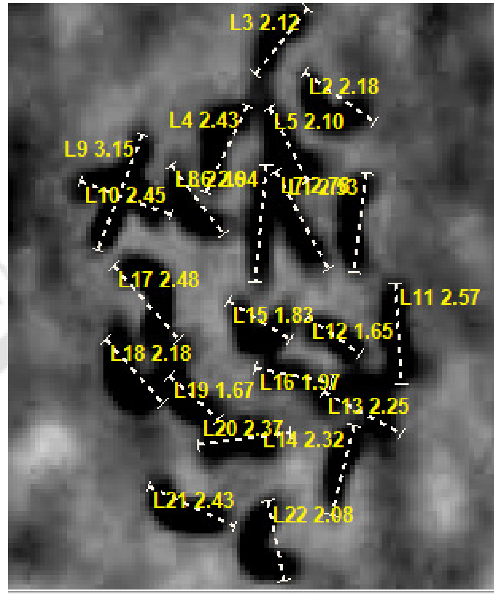


Figure 6.19: Somatic metaphase chromosome in *P. pinnata*.

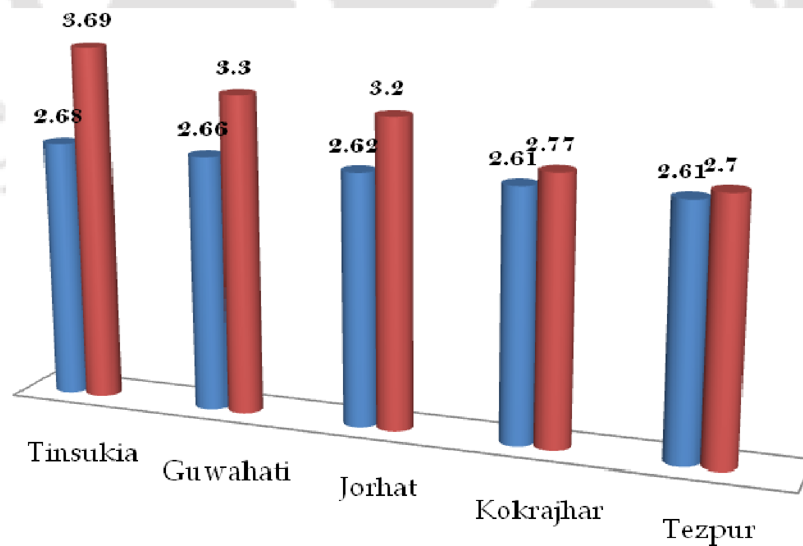


Figure 6.20: Correlation between nuclear genome size (pg) and length (µM) of *Pongamia* collected from different locations of Assam.

TH-1255_09610610

No intraspecific variation in the chromosome number was detected. Nuclear DNA content increased with increasing height from sea level. The nuclear DNA content was negatively correlated with an average temperature for studied locations. The nuclear DNA content estimated from the Kokrajhar individual is the same as the Tezpur individual due to similar environmental conditions. Intraspecific variations of nuclear DNA content below species level are supposed to be rare (Geilhuber 1998). Still 1.02 fold variation of genome size was observed among the individuals of *P. pinnata*. However the reported difference may be due to the chromosome polymorphisms and spontaneous aberrations, but also by technical shortcomings (Galbraith et al 1991; Geilhuber 1998). 1C-values of *Arabidopsis thaliana* and *Oryza sativa* are 157 and 490 Mb, respectively (Bennett et al 2003) whereas that of *Z. mays* is 2655 Mb (Dolezel et al 1998). Therefore, the estimated genome size of *P. pinnata* (1300 Mb) is 8.28 times that of *A. thaliana*, 5.41 times that of *O. sativa* and 0.49 times that of *Z. mays*. Based on records from plant DNA C-values database (Bennett & Leitch 2010) the genome size of *P. pinnata* falls in the upper end of C-value distribution of its family Fabaceae with 683 records ranging from 298 (*Leucaena macrophylla*) (Hartman et al 2000) to 26797 (*Vicia faba*) (Bennett et al 1982). It falls somewhere in the range among *Acacia nilotica* (1174 Mb) (Mukherjee & Sharma 1993), *Lupinus texensis* (1193 Mb) (Price et al 2005), *Onobrychis viciifolia* (1223 Mb) and *Acacia heterophylla* (1560 Mb) (Coulaud et al 1995).

6.4 CONCLUSION

The nuclear DNA content was estimated for *P. pinnata* despite the fact that it contains staining inhibitors in the leaf cytosol. The presence of staining inhibitors was tested for two reference standard used. Comparative estimation of nuclear DNA content was carried out by flow cytometry, and it was shown that the 2C value was consistent with repeated experiments. The addition of antioxidants helps in reducing the extent of inhibition of PI binding efficiency to the nucleus. The 2C value for *P. pinnata* was estimated to be 2.66 pg and it was found to be 2.25-2.68 pg with different tissue types, the less genome size in flower could be due to presence of floral pigments and secondary metabolites which may interfere in binding of PI to nucleus. The variation in genome size in *Pongamia* individuals collected from various parts of Assam, India and the positive correlation between the nuclear genome size and average chromosome

TH-1255_09610610

length could be due to the arrangements of repetitive sequence elements and chromatin condensation. The data presented in this chapter are noteworthy, as flow cytometry has been successfully used to determine the nuclear DNA content of *P. pinnata*. Also, the small nuclear DNA content, ease of vegetative manipulation and transformation are favorable features for the entry of other biodiesel crops as model system in the biotechnology era and will set the foundation for the needed sustainability in long-term breeding strategies for this valuable species.

TH-1255_09610610

Chapter 7

Summary

The chapter unveils the overall summery of the doctoral thesis and future directions.

TH-1255_09610610

SUMMARY

The depletion of fossil fuel coupled with soaring price hikes and adverse environmental impacts associated with the use of petro fuels, has triggered renewed interest in alternative fuels from renewable resources. In addition, the burning of fossil fuels produces green house gases which may lead to global warming. The only alternative way is to search for a renewable energy sources which are biodegradable. Plant sources for fuels that replace fossil fuels are a current subject. As an anticipatory research to meet the global demand for diesel, genotypes with high oil content need to be selected. *Pongamia pinnata*, oleaginous versatile legume tree comes first and foremost to researcher major attention because of the high oil content of the seeds. Although, these species are extensively characterized on the morphological aspects including candidate plus tree identification, micropropagation, genetic diversity study, protein profiling and symbiotic relationship; there is need to study extensively on molecular and cyto-genetic level in elite genotype of *Pongamia*.

In addition, information on fatty acid biosynthetic genes that involved in desaturation process and nuclear genome content is lacking in *Pongamia* plant. In this doctoral thesis the research is being carried out in earlier characterized elite genotype of *P. pinnata* occurring from North Guwahati, Assam that covered a wide range of molecular and cytogenetic topics such as germplasm collection (elite genotype), cDNA library construction from early immature seeds of elite genotype, screening and characterization of fatty acid biosynthetic genes, total nuclear genome content estimation and somatic chromosome number determination. The above works were achieved using tools of biology by understanding the basic molecular and cytological aspects of *P. pinnata* aimed at understanding the germplasm to improve its key traits.

The work was carried out in four distinct phases; **first phase** involved with the collection of earlier characterized germplasm (based on vegetative and reproductive characters) from Sila forest range, North Guwahati, Assam, India. The RNA extraction protocol was optimized to extract total RNA from the early immature seeds of elite genotype of *Pongamia* seed tissues which harbor polyphenolic compounds that interfere in RNA quality and quantity. With the best quality and quantity of the total

TH-1255_09610610

RNA, mRNA was successfully separated and preceded for cDNA library construction based on LD-PCR (long-distance PCR). Full length fatty acid biosynthetic genes were screened from the lawn of clones with gene specific partial sequences as probes. Long-term applications of this library would be in the conservation of *Pongamia* genes as a means of *ex situ* genetic conservation.

Second phase involved structural and functional characterization of fatty acid biosynthetic gene steroyl-ACP-desaturase screened from the cDNA library as described earlier. The enzyme mainly involved in conversion of saturated fatty acid to unsaturated fatty acid, in fact this is the first enzyme which decides the ratio of saturated to unsaturated fatty acids. The functional characters include the expression profile of the gene in different tissue types and copy number determination. Similarly structural characters includes the homology modeling based on the availability of template protein and its validation using different tools to confirm the quality of the structural model.

In the **third phase**, two full length fatty acid desaturases were screened from cDNA library and characterized for its transcript analysis through semi-quantitative RT-PCR. In addition the expression strategies of both genes (PpFAD2-1 & PpFAD2-2) were confirmed by real time and northern blot hybridization experiments. The findings of transcript analysis of fatty acid desaturase genes can be helpful in understanding the biosynthesis in *Pongamia*. However in future, expression analysis can be helpful in understanding the biosynthesis and in efforts to improve biosynthesis of TAGs (triacylglycerols).

In the **Fourth and final phase** an attempt was made to estimate total nuclear genome content in CPT of *Pongamia* and in various tissue types. Furthermore the nuclear DNA was estimated in *Pongamia* individuals collected from various parts of India from different locations of Assam. An attempt was also made to count somatic chromosome in different individuals from different locations and its chromosome length determination. The correlation between nuclear genome content and chromosome length were also studied in five individuals of Assam. The imminent study on nuclear genome content of a potential unexplored biodiesel crop even becomes a timelier step in these global warming providing new opportunities for comparative genomics study.

TH-1255_09610610

The work emphasized in this thesis has both theoretical and applied significance. The present study essentially includes germplasm collection, cDNA library construction, fatty acid biosynthetic gene characterization, nuclear genome estimation and somatic chromosome count of in elite genotype of *Pongamia*, an important legume tree occurring from North Guwahati, Assam. Such studies could also be extended for other biofuel crop. *P. pinnata* is regarded as one of the most useful plants for legume study and researchers having interests in biodiesel and cyto-genetic aspects can take the lead further from the current study.

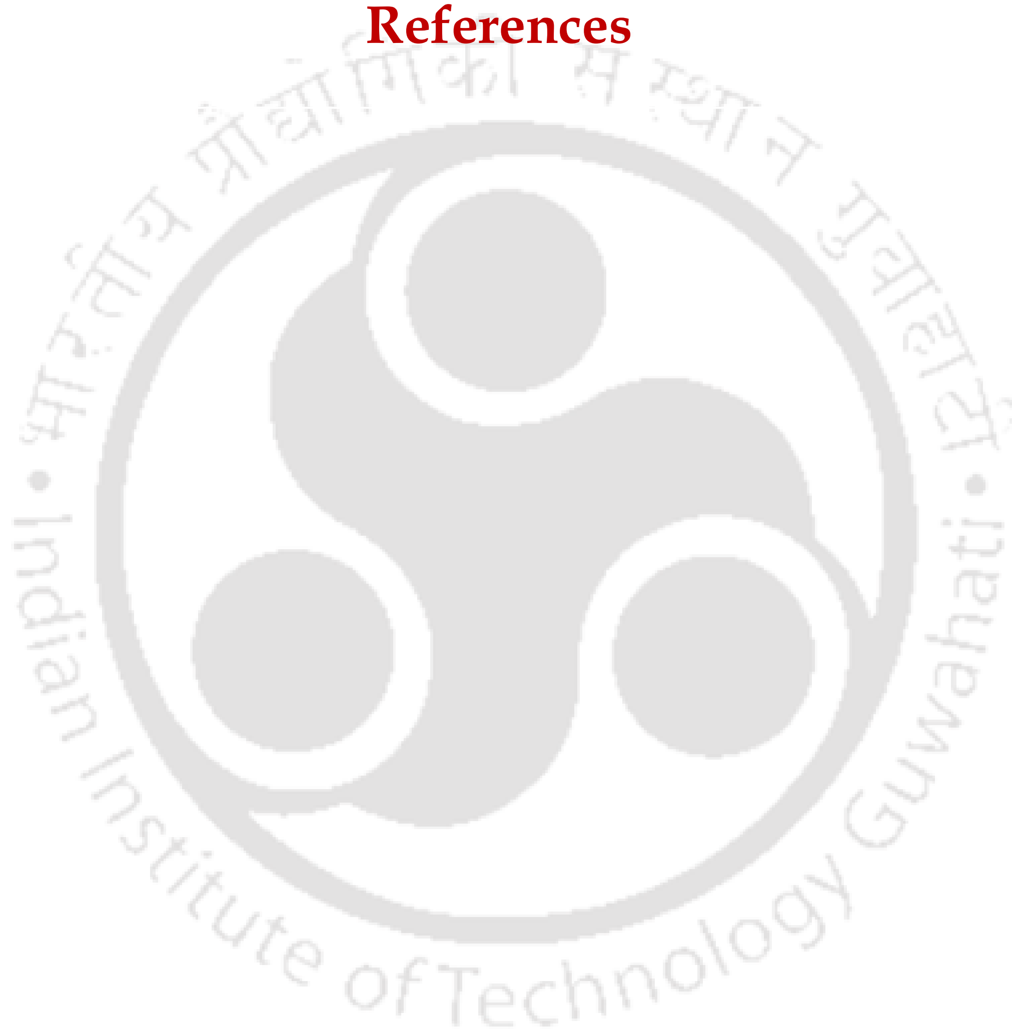
Future directions

Studies reported in this thesis have both theoretical & applied significance. The Potential works which could be explored on the basis of present research:

- The findings obtained in the current study provide the basis for future investigation of *Pongamia* fatty acid genes in terms of gene regulation and its role in increasing oil traits.
- Temperature stability of fatty acid desaturase gene to determine the level of poly unsaturated fatty acids (PUFA-Linoleic and α-linolenic acid) in heterologous (*Saccharomyces cerevisiae*) expression system.
- Characterization of the mechanisms that regulate *Pongamia* microsomal desaturase enzyme activity level could be carried out.
- Construction of fatty acid biosynthetic genes for desired traits into cloning and mobilizing vectors.
- Genetic characterization of any tree species is a challenging exercise particularly with legume plants that have no genetic maps and few identified sequences. Efficient utilization of *Pongamia* tree and availability of large-scale sequencing resources and detailed analysis rests on of its genome composition, especially with respect to repetitive DNA, which makes up the majority of the genome. Understanding plant genes and genome evolution is only possible if we comprehend the contributions of transposable elements. Since the genome content of *Pongamia* is small, investigation of repetitive sequence elements analysis and its classification could be done.

TH-1255_09610610

References



TH-1255_09610610

REFERENCES

Aadi MR, Basak S, Chowdhury RR, Rangan L (2013) Flow cytometric estimation of nuclear DNA content of *Pongamia pinnata* L. *Applied Biochemistry and Biotechnology*. DOI: 10.1007/s12010-013-0553-z.

Aam MM, Waris A, Nahar NM (2005) Prospects and potential of fatty acid methyl esters of some non-traditional seed oils for use as biodiesel in India. Biomass & Bioenergyvol. 29(4):293-302.

Ahmad S, Ashraf SM, Naqvi F, Yadav S, Hasnat A (2003) A polyesteramide from *Pongamia glabra* oil for biologically safe anticorrosive coating. Progress in Organic Coating. 47:95-102.

Akintayo ET (2004) Characteristics and composition of *Parkia biglobbossa* and *Jatropha curcas* oils and cakes. Bioresour Technol. 92:307-310.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.

Anderson GC (1991. Green scheme avoids quarrels. Nature 345:654.

Atchison E (1951) Studies in the Leguminosae. VI. Chromosome numbers among tropical woody species. Amer J Bot. 38:538-46.

Avelange-Macherel MH, Macherel D, Wada H, Murata N (1995) Site-directed mutagenesis of histidine in the v12 acyllipid desaturase of Synechocystis, FEBS Lett. 361:111-114.

Azevedo H, Lino-Neto T, Tavares RM (2003) An improved method for high-quality RNA isolation from needles of adult maritime pine trees. Plant Mol Biol Rep. 21:333-338.

Bablak Pm Draper J, Davey MR, Lynch PT (1995) Plant regeneration and micropropagation of Brachypodium ditachyon. Plant Cell Tiss. Org Cult. 42:97-107.

Bagwell CB, Baker D, Whetstoe S, Munson M, Hitchcox S, Ault K, Lovett EJ (1989) A simple and rapid method for determining the linearity of a flow cytometer amplification system. Cytometry. 10:689-694.

Baird WV, Estager AS, Wells JK (1994) Estimating nuclear DNA content in peach and related diploid species using laser flow cytometry and DNA hybridization. J Am Soc Hort Sc. 119:1312-1316.

Baluska F (1990) Nuclear size, DNA content, and chromatin condensation are different in individual tissues of the maize apex. Protoplasma. 158:1-2.

Bandaru VR, Tavva SS, Mohan DPV, Rap A, Rao MVS (2011) Image cytometric determinationof nuclear and chromosomal DNA contens in Allium sepa L. Nucleus. 54(2):71-75.

Bandel G (1974). Chromosome numbers and evolution in the Leguminosae. *Caryologia* 27:17-32.

Baranyi M, Greilhuber J (1995) Flow cytometric analysis of genome size variation in cultivated and wild *Pisum sativum* (Fabaceae). Pl Syst Evol. 194:231-239.

Bennett MD, Leitch IJ (1995) Nuclear DNA amounts in angiosperms. Ann Bot. 76:113-176.

Bennett MD, Leitch IJ (2005) Plant genome size research: a field in focus. Ann Bot. 95:1-6.

TH-1255_09610610

Bennett MD, Leitch IJ, Price HJ, Johnston JS (2003) Comparisons with *Caenorhabditis* (~100 Mb) and *Drosophila* (~175 Mb) using flow cytometry show genome size in arabidopsis to be ~157 Mb and thus ~25 % larger than the Arabidopsis Genome Initiative of ~125 Mb. Ann Bot. 91:547-557.

Bennett MD, smith JB (1976) Nuclear DNA amounts in angiosperms. Philosophical Transactions of the Royal Society of London B. 274:227-274.

Bennetzen JL (2007) Patterns in grass genome evolution. Curr Opin Plant Biol. 10:176-181.

Bharathan G, Lambert G, Galbraith DW (1994) Nuclear DNA content of monocotyledons and related taxa. American J Bot. 81:381-386.

Biradar DP, Rayburn AL (1994) Flow cytometric probing of chromatin condensation in maize diploid nuclei. New Phytologist. 126:31-35.

Birtic S, Kranner I (2006) Isolation of high-quality RNA from polyphenol, polysaccharide and lipid-rich seeds. Phytochem Analysis. 17:144-148.

Bloomfield DK (1960) Formation of Δ9-unsaturated fatty acids. J Biol Chem. 235:337-345.

Bobade SN, Khyade VB (2012) Detail study on the Properties of *Pongamia Pinnata* (Karanja) for the Production of Biofuel. Res J Chem Sci. 2(7):16-20.

Bottoms T (2000) Bama country-aboriginal homelands. In:McDonald G, Lane M (eds) Securing the wet tropics? Federation Press, Leichardt, NSW. pp.32-47.

Brijesh S, et al (2006) Studies on *Pongamia pinnata* (L.) Pierre leaves: understanding the mechanism (s) of action in infectious diarrhea. J Zhejiang Univ Science B. 7(8):665-674.

Brown TA (2001) Gene Cloning and DNA Analysis. A text book.

Browse J, Somerville C (1991) Glycerolipid synthesis: biochemistry and regulation. Annu Rev Plant Physiol. 42:467-506.

Bugos RC, Chiang VL, Campbell WH (1992) Characterization of bispecific caffeic acid/5-hydroxyferulic acid O-methyltransferase from aspen. Phytochemistry.31:1495-1498.

Bukhari YM (1997) Cytoevolution of taxa in Acacia and Prosopis (Mimosaceae). Austral J Bot. 45:879-891.

Bukhari YM (1997) Nuclear DNA amounts in *Acacia* and *Prosopis* (Mimosaceae) and their evolutionary implications. Hereditas. 126:45-51.

Caliebe A, Soill J (1993) News in chloroplast protein import. Plant Mol Biol. 39:641-645.

Catalan P, Shi Y, Armstrong L, Draper J, Stace CA (1995) Molecular Phylogeny of the grass genus *Brachypodium P. Beauv.* based on RFLP and RAPD analysis. Bo J Linn Scoc. 117:263-280.

Chang S, Puryear J, and Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. Plant Mol Biol Rep 11:113-116.

Chen ZY (1992) Cytology of Zingiberaceae. In: SirirugsaP, editor. A training report on cytotaxonomy of Zingiberaceae and some selected plants. Hat Yai: Department of Biology, Prince of Songkla University.19-29.

Chenchik A, Zhu YY, Diatchenko L, Li R, Hill J, Siebert PD (1998) in Gene cloning and analysis by RT-PCR, Generation and use of high quality cDNA from small amounts of total RNA by SMART™ PCR, eds Siebert P, Larrick J (Eaton, Natick, MA), pp.305-319.

TH-1255_09610610

Cheng B, Wu G, Vrinten P, Falk K, Bauer J, Qiu X (2010) Towards the production of high levels of eicosapentaenoic acid in transgenic plants: the effects of different host species, genes and promoters. Transgenic Res. 19:221-229.

Chopade VV, Tankar AN, Pande VV, Tekade AR, Gowekar NM, Bhandari SR, Khandake SN (2008) *Pongamia pinnata*: Phytochemical constituents, traditional uses and pharmacological properties: A review. Int J Green Pharm. 2(2):72.

Chsti Y (2007) Biodiesel from microalgae. Biotechnol Adv 25(3):294-306.

Clarindo WR, Carvalho CR (2009) Flow cytometric analysis using SYBR Green I for genome size estimation in coffee. Acta histochemica. 113:221-225.

Clarke L, Carbon J (1976) Cell 9:91-99

Coulaud J, Brown SC, Siljak-Yakovlev S (1995) First cytogenetic investigation in populations of *Acacia heterophylla*, endemic from La Reunion Island, with reference to *A. melanoxylon.* Ann Bot. 75:95-100.

Dahmer N, Wittmann MTS, Dos SDLA (2009) Chromosome numbers of *Jatropha curcas* L: an important agrofuel plant. Crop Brreding and Applied Biotechnology. 9:386-389.

Daniel M, et al (2001) Studies on the effect of leaf leachates of *Pongamia pinnata* on certain crops and weeds and the soil microflora. Natl Acad Sci Lett. 24:63-68.

Darlington CD, LaCour LF (1976) The handling of chromosomes. 6th ed. Wiley, New York.

Das AB, Rai S, Das P (1998) Estimation of 4C DNA and karyotype analysis in ginger (*Zingiber officinale Rosc*.) II. Cytologia. 63:133-9.

Davis RW, Botstein J, Roth J (1980) Adv Bacterial Gen Cold Spring Harbor New York.

Davydov R, Behrouzian B, Smoukov S, Stubbe J, Hoffman BM, Shanklin J (2005) Effect of substrate on the diiron (III) site in stearoyl acyl carrier protein delta 9-desaturase as disclosed by cryoreduction electron paramagnetic resonance/electron nuclear double resonance spectroscopy. J Biochem. 44:1309-1315.

Ding JQ, Mack RN, Lu P, Ren MX, Huang HW (2008) China's booming economy is sparking and accelerating biological invasions. BioScience. 58:317-324.

Dmitry A, Los a, Murata N (1998) Structure and expression of fatty acid desaturases. Biochimica et Biophysica Acta.1394:3-15.

Dolezˇel J, Greilhuber J (2010) Nuclear genome size: are we getting closer?. Cytometry. 77A:635-642.

Dolezel J, Bartos J (2005) Plant DNA flow cytometry and estimation of nuclear genome size. Ann Bot. 95:99-110.

Dolezel J, Bartos J, Voglmayr H, Greilhuber J (2003) Nuclear DNA content and genome size of trout and human. Cytometry. 51:127-128.

Doležel J, Binarová P, Lucretti S (1989) Analysis of nuclear DNA content in plant cells by flow cytometry. Biologia Plantarum 31:113-120.

Dolezel J, Greilhuber J, Lucretti S, Meister A, Lysak MA, Nardi L, Obermayer R (1998) Plant genome size estimation by flow cytometry: Inter-laboratory comparison. Ann Bot. 82(Suppl. A):17-26.

TH-1255_09610610

Dolezel J, Greilhuber J, Suda J (2007) Estimation of nuclear DNA content in plants using flow cytometry. Nat Prot. 2:2233-2244.

Draper J, Mur LA, Jenkins G, Ghosh-Biswas GC, Bablak P, Hasterok R, Routledge AP (2001). *Brachypodium distachyon*. A new model system for functional genomics in grasses. Plant Physiol. 127: 1539-1555.

Duke JA (1983) Handbook of Energy Crops.

Endres H (1961) Die Gerbwirkung niedermolekularer Polyhydroxyphenole, Leder. 12:294-297.

Eswar N, John B, Mirkovic N, Fiser A, Ilyin V A, Pieper U, Stuart AC, Marti-Renom MA, Madhusudhan MS, Yerkovich B, Sali A (2003) Tools for comparative protein structure modeling and analysis. Nucleic Acids Res. 31:3375-3380.

Falcone DL, Gibson SG, Lemieux B, Somerville C (1994) Identification of a gene complements ana Arabidopsis mutant deficient in chloroplast ω-6desaturase activity. Plant Physiol. 106:1453-1459.

Farrell R (2005) RNA Methodologies. 3rd edition. Amsterdam: Elsevier Academic Press.

Fay MF, Cowan RS, Leitch IJ (2005). The effects of nuclear DNA content (C-value) on the quality and utiligy of AFLP fingerprints. Ann Bot. 95:237-246.

Fedorov AMA (1969) Chromosome Number of Flowering Plants. Komarov Botanical Institute, Leningrad.

Flavell RB, Bennett MD, Smith JB, Smith DB (1974) Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochem Genet. 12:257–269.

Fofana B, Duguid S, Cloutier S (2004) Cloning of fatty acid biosynthetic genes β-ketoacyl CoA synthase, fatty acid elongase, stearoyl-ACP desaturase, and fatty acid desaturase and analysis of expression in the early developmental stages of flax (*Linum usitatissimum* L.) seeds. Plant Sci. 166(6):1487-1496.

Fox BG, Shanklin J, Somerville CR, Mu«nck E (1993) Stearoylacyl carrier protein v9 desaturase from Ricinus communis is a diiron-oxo protein. Proc Natl Acad Sci USA. 90:2486-2490.

Fox BG, Shanklin J, Ai J, Loehr TM, Sanders-Loehr J (1994) Resonance raman evidence for a Fe-O-Fe center in stearoyl-ACP desaturase primary sequence identity with other diiron-oxo proteins. J Biochem. 33:12776-12786.

Fulco AJ (1974) Metabolic alterations of fatty acids. Ann Rev Biochem 43:215-240.

Galbraith DW, Harkins K, Maddox J, Ayres N, Sharma D, Firrozabady E (1983) Rapid Flow Cytometric Analysis of the Cell Cycle in Intact Plant Tissues. Science. 220(4601):1049-1051.

Galbraith DW, Harkins KR, Knapp S (1991) Systemic endoploidy in *Arabidopsis thaliana*. Plant Physiol. 96:985-989.

Galbraith DW, Lambert GM, Macas J, Dolezel J (1998) Analysis of nuclear DNA content and plody in higher plants. In: Robinovitch et al., eds. Current protocols in cytometry. New York: Yohn Wiley & Sons, pp. 7.6.1-7.6.22.

Gao J, Liu J, Li B, Li Z (2001) Isolation and purification of functional total RNA from blue-grained Wheat endosperm tissues containing high levels of starches and flavonoids. Plant Mol Biol Rep.19:185a-185.

TH-1255_09610610

Geilhuber J (1998) Intraspecific variation in genome size: a critical reassessment. Ann Bot. 82(Suppl. A):27-35.

Gill BS, Singhal VK (1998). Chromosomes, chromosomal techniques and chromosomal evolution. In: Forest Genetics and Tree Breeding. 168-193.

Gilman EF, Watson DG (1994) *Pongamia pinnata*. Fact Sheet ST-498, Forest service, department of Agriculure, Southern Group of State Foresters.

Graef G, LaVallee BJ, Tenopir P, Tat M, Schweiger B, Kinney AJ, Van Gerpen JH, Clemente TE (2009) A high-oleic-acid and low-palmitic-acid soybean: agronomic performance and evaluation as a feedstock for biodiesel. Plant Biotech J. 7:411-421.

Greilhuber J, Temsch E, Loureiro J (2007) Nuclear DNA content measurement in Flow Cytometry with Plant Cells, In: J. Dolezel, J. Greilhuber, J. Suda (Eds.), Analysis of Genes, Chromosomes and Genomes, Wiley-VCH, Weinheim, Germany, pp. 67-101.

Greilhuber J, Temsch E, Loureiro J (2007) Nuclear DNA content measurement in Flow Cytometry with Plant Cells, In: J. Dolezel, J. Greilhuber, J. Suda (Eds.), Analysis of Genes, Chromosomes and Genomes, Wiley-VCH, Weinheim, Germany, pp. 67-101.

Guex N, Peitsch MC (1997) Swiss-Model and the Swiss-Pdb Viewer: an environment for comparative protein modeling. Electrophoresis. 18:2714-2723.

Hanahan D, Meselson M (1980) Plasmid screening at high colony density. Gene. 10:63-67.

Hanahan D, Meselson M (1983) Plasmid screening at high colony density. Methods in Enzymology.100:333-341.

Harborne JB (Ed.) (1994) The flavonoids: advances in research since 1986. CRC Press. 4.

Hartman TPV, Jones J, Blackhall NW, Power JB, Cocking EC, Davey MR (2000) Cytogenetics, molecular cytogenetics and genome size in *Leucaena* (Leguminosae, Mimosoideae), In : H. Guttenberger, Z. Borzan, S.E. Schlarbaum, T.P.V. Hartman (Eds.), Cytogenetic studies of forest trees and shrubs-review, present status and outlook for the future, For Genet, Special Issue, pp. 57-70.

Haunli W, Fuliang C, Wangxiang Z, Guibin W, Wanwen Y (2013) Cloning and expression of steryol-ACP desaturase and tow oleate desaturases genes from Ginkgo biloba L. *Plant Mol Biol Rep*. 31, 633-648.

Heppard EP, Kinney AJ, Stecca KL, Miao G-H (1996) Developmental and growth temperature regulation of two different microsomal ω-6 desaturase genes in soybean. Plant Physiol. 110:311-319.

Hernandez ML, Mancha M, Martinez Rivas JM (2005) Molecular cloning and characterization of genes encoding two microsomal oleate desaturases (FAD2) from olive. Phytochemistry. 66:1417-1426.

Higashi S, Murata N (1993) An in vivo study of substrate specificities of acyl-lipid desaturases and acyltransfereses in lipid synthesis in lipid synthesis in *Synechocystis* PCC 6803. Plant Physiol. 102:1275-1278.

Hill J, Nelson E, Tilman D, Polasky S, Tiffany D (2006) Environmental, economic and energetic costs and benefits of biodiesel and ethanol biofuels. Proc Natl Acad Sci USA. 103:11206-11210.

TH-1255_09610610

Holdsworth M, Kurup S, McKibbin R (1999) Molecular and genetic mechanisms regulating the transition from embryo development to germination. *Trends Plant Sci.* 4:275-280. http://www.madsci.org/posts/archives/1997-08/869070443.Ge.r.html

Hu XY, Mandy SG, Gupta M, Steven A (2006) Mapping of the loci controlling oleic and linoleic acid contents and development of *fad2* and *fad3* allele-specific markers in canola *(Brassica napus* L.). Teor Appl Genet. 113:497-507.

Irsch C, Takamiya-Wik M, Reinold S, Hahlbrock K, Somssich IE (1997) Rapid, transient, and highly localized induction of plastidial g3 fatty acid desaturase mRNA at fungal infection sites in *Petroselinum crispum*. Proc Natl Acad Sci USA. 94:2079-2084.

Jana LS, S, Vlasta J, Mamyil S, Tomas F, Pavel T, Jan S (2007) Chromosome Numbers and Genome Size Variation in Indian Species of *Curcuma* (Zingiberaceae). *Ann Bot*. 100(3):505-526.

Jatinder Singh Sangha, Keyu Gu, Jatinder Kaur and Zhongchao Yin (2010) SAhonrt Riempoprt roved method for RNA isolation and cDNA library construction from immature seeds of *Jatropha curcas* L Sangha *et al. BMC Research Notes* 3, 126.

Jedrzejczyk I, Sliwinska E (2010) Leaves and Seeds asMaterials for Flow Cytometric Estimation of the Genome Size of 11 RosaceaeWoody Species Containing DNA-Staining Inhibitors. J Bot. doi:10.1155/2010/930895.

Jeffrey L, Bennetzen MAJX, Devos KM (2005) Mechanisms of recent genome size variation in flowering plants. Annals of Botany 95:127-132.

Jena PK, Liu AH, Smith DS, Wysocki LJ (1996) Amplification of genes, single transcripts and cDNA libraries from one cell and direct sequence analysis of amplified products derived from one molecule. J immunol Methods. 190:199-213.

Jingura RM, Musademba D, Matengaifa R (2009) An evaluation of utility of *Jatropha curcas* L. as a source of multiple energy carriers. Int J Eng Sci Technol. 2(7):115-122.

Jones RN, Rees H (1968) Nuclear DNA variation in *Allium*. Heredity. 23:591-605.

Joseph R, Joseph T, Joseph J (1999) Karyomorphological studies in the genus Curcuma Linn. Cytologia. 64:313-317.

Kachroo A, Lapchyk L, Fukushige H, Hildebrand D, Klessig D, Kachroo P (2003) Plastidial fatty acid signaling modulates salicylic acid-and jasmonic acid-mediated defense pathways in the *Arabidopsis* ssi2 Mutant. The Plant Cell.15:2952-2965.

Kanellos M (2009) Boosting biofuels with extra chromosomes. www.greentechmedia.com/articles/read/boosting-biofuels-with-extra-chromosomes.

Karmee SK, Chadha A (2005) Preparation of biodiesel from crude oil of *Pongamia pinnata.* Bioresource Technology. 96(13):1425-1429.

Kaushik N, Kumar S, Kumar K, Beniwal RS, Kaushik N, Roy S (2007)Genetic variability and association studies in pod and seed traits of *Pongamia pinnata* (L.) Pierre in Haryana, India. Genetic Resources and Crop Evolution. 54(8):1827-1832.

Kesari V, Satyanarayan M, Parida A, Rangan L (2010) Molecular marker-based characterization in candidate plus trees of *Pongamia pinnata*, a potential biodiesel legume. AoB Plants.

TH-1255_09610610

Kesari V, Das A, Rangan L (2010) Effect of genotype and auxin treatments on rooting response in stem cuttings of CPTs of *Pongamia pinnata*, a potential biodiesel legume crop. Curr Sci. 98(9):1234-1237.

Kesari V, Das A, Rangan L (2010) Physico-chemical characterization and microbial assay from seed oil of *Pongamia pinnata*, potential biofuel crop. Biomass Bioenergy. 34(1):108-115.

Kesari V, Krishnamachari A, Rangan L (2008) Systematic characterization and seed oil analysis in candidate plus trees of biodiesel plant, *Pongamia pinnata*. Ann Appl Biol. 152:397-404.

Kesari V, Krishnamachari A, Rangan L (2009) Effect of auxins on adventitious rooting from stem cuttings of candidate plus tree *Pongamia pinnata* (L.), a potential biodiesel plant. Trees. 23(3):597-604.

Kesari V, Ramesh AM, Rangan L (2012) High frequency direct organogenesis and evaluation of genetic stability for in vitro regenerated *Pongamia pinnata*, a valuable biodiesel plant. *Biomass Bioenergy* 44:23-32.

Kesari V, Ramesh AM, Rangan L (2013) *Rhizobium pongamiae* sp. nov. from root nodules of *Pongamia pinnata*. BioMed research international. doi.org/10.1155/2013/165198

Kesari V, Rangan L (2010b) Development of *Pongamia pinnata* as an alternative biofuel crop-current status and scope of plantations in India. J Crop Sci Bio 13(3):127-137.

Kesari V, Sudarshan M, Das A, Rangan L (2009) PCR amplification of the genomic DNA from the seeds of Ceylon ironwood, *Jatropha*, and *Pongamia*. Biomass Bioenergy. 33(12):1724-1728.

Kim N, Kim W, Lee S, Lee H, Rhee H, Choi Y, Jung Y, Cha S (1999) Isolation of rice allergenic cDNA clones from a rice cDNA library by immunoscreening with a polyclonal antibody specific to 16 kD rice allergenic protein. Exp Mol Med. 31(4):185-90.

Kimmel AR, Berger SL (1987) Preparation of cDNA and the generation of cDNA libraries. Methods Enzymol. 152:307-316.

Knothe G (2005) Dependence of biodiesel fuel properties on the structure of fatty acid alkyl esters. Fuel Process. Technol. 86(10):1059-1070.

Knutzon DS, Bleibaum JL, Nelsen J, Kridl JC, Thompson GA (1992) Isolation and characterization of two safflower oleoyl-acyl carrier protein thioesterase cDNA clones. Plant Physiol.100:1751-1758.

Knutzon DS, Thompson GA, Radke SE, Johnson WB, Knauf VC, Kridl JC (1992) Modification of *Brassica* seed oil by antisense expression of a stearoyl-acyl carrier protein desaturase gene. P Nat Acad Sci USA. 89:2624-2628.

Kodama H, Horiguchi G, Nishiuchi T, Nishimura M, Iba K (1995) Fatty acid desaturation during chilling acclimation is one of the factors involved in conferring low temperature tolerance to young tobacco leaves. Plant Physiol. 107:1177-1185.

Kohchi T, Fujishige K, Ohyama K (1995) Construction of an equalized cDNA library from *Arabidopsis thaliana*. Plant J. 8: 771-776.

Kpikpi WM (2000) *Jatropha curcus as vegetable source of renewable energy*, Paper presented at ANSTI subnetwork meeting on Renewable energy.18-22.

TH-1255_09610610

Krishan A (1975) Rapid flow cyto fluorometric analysis of mammalian cell cycle by propidium iodide staining. J Cell Biol. 66:188-193.

Kubalakova M, Macas J, Dolezel J (1997) Mapping of repeated DNA sequences in plant chromosome by PRINS and C-PRINS. Theor Appl Genet. 94:758-763.

Kudrid P, Dubudhi S, Hongsthong A, Ruengjitchatchawalya M, Tanticharoen M (2005) Functional expression of spirulina-Δ6 desaturase gene in yeast, Saccharomyces cerevisiae. Mol Biol Rep. 32:215-226.

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

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol.157(1):105–32.

Labani RM, Elkington TT (1987) Nuclear DNA variation in the genus *Allium* L. (Liliaceae). Heredity. 59:119-128.

Lambert KN, Williamson VM (1993) cDNA library construction from small amounts of RNA using paramagnatic beads and PCR. Nucleic Acids Res. 21:775-776.

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

Laskowski RA, MasArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr. 26:283-291.

Lea M (2003) Stress on avocado fruits regulates 9-stearoyl ACP desaturase expression, fatty acid composition, antifungal diene level and resistance to Colletotrichum gloeosporioides attack. Physiol Mol Plant Pathology. 62:277-283.

Lee H, Lin T (2005) Isolation of plant Nuclei suitable for flow cytometry from recalcitrant tissue by use of filtration column. Plant Mol Bio Reporters. 23:53-58.

Lee VK, Feder JN, Penny DM, Irrinki A, Lee VK, Lebron JA, Watson N, Tsuchihashi, Z, Sigal E, Bjorkman PJ, Schatzman RC (1998) The Hemochromatosis Gene Product Complexes with the

Leitch IJ, Beaulieu JM, Chase MW, Leitch AR, Fay MF (2010) Genome size dynamics and evolution in Monocots. J Bot. Articl ID 862516, 18 pages.

Lin L, Cunshan Z, Vittayapadung S, Xiangqian S, Mingdon D (2011) Opportunities and challenges for biodiesel fuel. Appl. Energ. 88(4):1020-1031.

Lindqvist Y,Huang W,Schneider G, Shanklin J (1996) Crystal structure of Δ9 stearoyl-acyl carrier protein desaturase from castor seed and its relationship to other di-iron proteins. The EMBO J. 15:4081-4092.

Lingyong L, Xianolin W, Junyi G, Deyue Y (2007) Molecular cloning and characterization of a novel microsomal oleate desaturase gene from soybean. J. Plant Physiol. 164:1516-1526.

Liu Q, Singh SP, Brubaker CL, Sharp PJ, Green AG, Marshall DR (1999) Molecular cloning and expression of a cDNA encoding a microsomal ω-6 fatty acid desturase from cotton (Gossypium hirsutum). Aust J Plant Physiol. 26:101-106.

TH-1255_09610610

Loomis WD (1974) Overcoming problems of phenolics and quinones in the isolation of plant enzymes and organelles. Meth Enzymol. 31:528-545.

Los DA, Marata N (1999) Structure and expression of fatty acid desaturase. Biochim BIophys Acta. 1394:3-15.

Los DA, Murata N (1998) Structure and expression of fatty acid desaturases. Biochimica et Biophysica Acta.1394:3-15.

Loureiro J, Rodriguez E, Dolezel J, Santos C (2006) Flow cytometric and microscopic analysis of the effect of tannic acid on plant nuclei and estimation of DNA content. Ann Bot. 98(3):515-527.

Loureiro J, Rodriguez E, Dolezel J, Santos C (2007) Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. Ann Bot. 100(4):875-888.

Lu C, et al (2011) cDNA Libraries: Methods and Applications. Methods in Molecular Biology. 729.

Luo T, Peng SM, Deng WY, Ma DW, Xu Y, Xiao M, Chen F (2006) Characterization of a new stearoyl–Acyl carrier protein desaturase gene from Jatropha curcas. Biotechnol Lett. 28:657-662.

Lusa MH, Mancha M, Martinez-Rivas JM (2005) Molecular cloning and characterization of genes encoding tow microsomal oleate desaturastes (FAD2) from olive. Phytochemistry 66:1417-1426

Lysak MA, Rostkova A, Dixon JM, Rossi G, Dolezel J (2000) Limited genome size variation in *Sesleria albicans*. Ann Botany. 86:399-403.

Maniatis T, Fritsch Ef, Sambrook J (1982) Molecular Cloning: A Laboratory Manual. New York: Cold Spring Harbor Laboratory.

Marie D and Brown SC (1993) A cytometric exercise in plant DNA histograms, with 2C values for 70 species. Biol Cell 78:41-51.

Martel E, De ND, Siljak-Yakovlev S, Brown S, Sarr A (1997) Genome size variation and basic chromosome number in pearl millet and fourteen related *Pennisetum* species. J Heredity. 88:139-143.

Martinez-Rivas JM, Sperling P, Luhs W, Heinz E (2001) Spatial and temporal regulation of three different microsomal oleate desaturase genes (FAD2) from normaltype and high-oleic varieties of sunflower (*Helianthus annuus* L.). Mol Breed. 8:159-68.

Martini N, Shell JS (1998) Plant oils as fuels-present state of science and future development. Berlin: Springer. 276.

McCarrey JR, Williams SA (1994) Construction of cDNA libraries from limiting amounts of material. Curr Opin Biotechnol. 5(1):34-9.

McCartney AW, Dyer JM, Dhanoa PK, Kim PK, Andrews DW, McNew JA, Mullen RT (2004) Membrane-bound fatty acid desaturases are inserted co-translationally into the ER and contain different ER retrieval motifs at their carboxy termini. Plant J. 37:156-73.

Meagher RB, McLean M, Arnold J (1988) Recombination within a subclass of restriction fragment length polymorphisms may help link classical and molecular genetics. Genetics 120:809-818.

Meera B, et al (2003) A review of the chemistry and biological activity of *Pongamia pinnata*. J Medicinal Aromatic Plant Sci. 25:441-465.

TH-1255_09610610

Mekhedov S, De Ilarduya OM, Ohlrogge J (2000) Toward a functional catalog of the genome. A Survey of genes for lipid biosynthesis. Plant Physiol. 122:389-401.

Michaelson MJ, Price HJ, Ellison JR, Johnston JS. (1991) *a.* Comparison of plant DNA contents determined by Feulgen microspectrophotometry and laser flow cytometry.American Journal of Botany.78*:*183-188.

Mietkiewska E, Brost JM, Giblin EM, Francis T, Wang S, Reed D (2006) A Tropaeolum majus FAD2 cDNA complements the fad2 mutation in transgenic Arabidopsis plants. Plant Sci.171:187-93.

Mikkilineni V, Rocheford TR (2003) Sequence variation and genomic organization of fatty acid desaturase-2 (FAD2) and fatty acid desaturase-6 (FAD6) cDNAs in *maize*.Theor Appl Genet. 106(1):326-1332.

Mochida K, Uehara-Yamaguchi Y, Takahashi F, Yoshida T, Sakurai T, Shinozaki K (2013) Large-Scale Collection and Analysis of Full-Length cDNAs from *Brachypodium distachyon* and Integration with Pooideae Sequence Resources. PLoS One. 8(10):752-765.

Morgan ER, Burge GK, Seelye JF, Grant JE, Hopping ME (1995) Interspecific hybridization between *Limonium perigrinum Bergius* and *Limonium purpuratum* L. Euphytica. 83:215-24.

Mukherjee I (1970) Chromosome studies of some species of Hedychium. Botanical Magazine Tokyo. 83:237–241.

Mukherjee S, Sharma AK (1993) *In situ* nuclear DNA content in perennial fast and slow growing acacias from arid zones. Cytobios. 75:33-36.

Munnik T, Irvine RF, Musgrave A (1998) Phospholipid signaling in plants. Biochim Biophys Acta. 1389:222-272.

Murata N, Wada H (1995) Acyl-lipid desaturases and their importance in the tolerance and acclimatization to cold of cyanobacteria. Biochem J. 308:1-8.

Murillo I, Raventos D, Jaeck E, San Segundo B (1995) Isolation of total RNA and mRNA from plant tissues, Promega Notes Mag. 54:2-7.

Murphy DJ, Piffanelli P (1998) Fatty acid desaturases: structure, mechanism and regulation. In: Plant Lipid Biosynthesis: Fundamentals and Agricultural Applications (Harwood, J.L. ed.), Cambridge University Press, pp 95-130.

Mustafa B (2011) Potential alternatives to edible oils for biodiesel production-A review of current work. Energ Convers Manage. 52(2):1479-1492.

Nagaraj G, Mukta N (2004) Seed composition and fatty acid profile of some tree borne oilseeds. Journal of Oilseed Research. 21: 117-120.

Nagl W, Ehrendorfer F (1974) DNA content, heterochromatin, mitotic index and growth in perennial and annual *Anthemideae* (Asteraceae). Plant Syst Evol. 123:35-54.

Nair RR, Sasikumar B (2009) Chromosome number variation among germplasm collections and seedling progenies in Turmeric, *Curcuma longa* L.Cytologia. 74:153-157.

Nandi A, Krothapalli K, Shah J (2003) Arabidopsis sfd Mutants Affect Plastidic Lipid Composition and Suppress Dwarfing, Cell Death, and the Enhanced Disease Resistance Phenotypes Resulting from the Deficiency of a Fatty Acid Desaturase. Plant Cell.15(12):3051.

TH-1255_09610610

Narayan RKJ (1988) Constraints upon the organization and evolution of chromosomes in Allium. Theor Appl Genet. 75:319-329.

Natanam R, Kadirvel R, Viswanathan K (1989) The effects of karanj (*Pongamia glabra* vent) cake on the performance of white leghorn pullets. Anim. Feed Sci. Technol. 27:89-93.

Nielsen H, Brunak S, von Heijne G (1999) Machine learning approaches for the prediction of signal peptides and other protein sorting signals. Protein Eng. 12(1):3-9.

Nielsen H, Engelbrecht J, Brunak S, Von Heinjne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10:1-6.

Nishida I, Murata N (1996) Chilling sensitivity in plants and cyanobacteria: the crucial contribution of membrane lipids. Annu Rev Plant Physiol Plant Mol Biol. 47:541-568.

Noirot M, Barre P, Duperray C, Louarn J, Hamon S (2003) Effects of caffeine and chlorogenic acid on propidium iodide accessibility to DNA: consequences on genome size evaluation in coffee tree. Ann. Bot. 92(2):259-264.

Obae GS, West TP (2010) Nuclear DNA content of *Hydrastis canadensis* L. and genome size stability of in vitro regenerated plantlets. Plant Cell Tissue Organ. 102:259-263.

O'Brien IE, Smith DR, Gardner RC, Murray BG (1996) Flow cytometric determination of genome size in Pinus. Plant Sci. 115:91-99.

Ohlrogge J (1997) Regulation of fatty acid synthesis. Ann Rev Plant Physiol Plant Mol Biol. 48:109-136.

Ohlrogge J, Browse J (1995) Lipid biosynthesis. Plant Cell. 7:957-970.

Ohri D, Kumar A (1986) Nuclear DNA amounts in some tropical hardwoods. Caryologia. 39:303-307.

Okuley J, Lighatner J, Geldmann K, Yadav N, Lark E (1994) Arabidopsis FAD2 gene encode enzyme that is essential for polyunsaturated lipid synthesis. Plant Cell. 6:147-158.

Old RW, Primrose SB (1994) Principles of gene manipulation: an introduction to genetic engineering. 5th ed. Oxford ; Boston : Blackwell Scientific Publications.

Olszewska MJ, Osiecka R (1984) The relationship between 2C DNA content, systematic position and the level of nuclear DNA endoreplication during differentiation of root parenchyma in some dicotyledonous shrubs and trees. Comparison with herbaceous species. Biochem Physiol Pflanzen. 179:641-657.

Onken (1997) What is a cDNA library.

Ott L, Longnecker M (1998) An introduction to statistical methods and data analysis, Boston-MA: PWS-KENT.

Parida A, Raina SN, Narayan RKJ (1990) Quantitative DNA variation between and within chromosome complements of *Vigna* species (Fabaceae). Genetica. 82:125-133.

Patel JS, Narayana GV (1937) Chromosome numbers in some economic flowering plants. Curr Sci. 5:479.

Pedro JS, Mario RM, Claudio PR (2006) Chromosome numbers of chileanpteridopytes: First contribution. *Gayana Bot*. 63 (1):115-118.

Pingzhi Wu et al (2013) Functional characterization of two microsomal fatty acid desaturases from *Jatropha curcas* L. Journal of Plant Physiology. 170(15):1360-1366.

TH-1255_09610610

Pingzhi Wu, Sheng Z, Lin Z, Yaping C, Meiru L, Huawu J, Guojiang (2013) Functional characterization of two microsomal fatty acid acid desaturases from *Jatropha curcas* L. J Plant Physiol. 170:1360-1366.

Pinto G, Loureiro J, Lopes T, Santos C (2004) Analysis of the genetic stability of *Eucalyptus globulus* Labill. somatic embryos by flow cytometry. Theor Appl Genet. 109(3):580-587.

Pirtle IL, Kongcharoensuntorn W, Nampaisansuk M, Knesek JE, Chapman KD, Pirtle RM (2001) Molecular cloning and functional expression of the gene for a cotton Delta-12 desaturase fatty acid desaturase (FAD2). Biochim Biophys Acta. 1522:122-129.

Polhill RM, Raven PH (Ed.) (1981) Advances in legume systematics: parts 1 & 2. Kew, Royal Botanic Gardens. 2:1-7.

Price HJ, Dillon SL, Hadnett G, Rooney WL, Ross L,Johnston JS (2005) Genome evolution in the genus Sorghum (Poaceae). Ann Bot. 95:219-227.

Price HJ, Hodnett G, Johnston JS (2000) Sunflower (*Helianthus annuus*) leaves contain compounds that reduce nuclear propidium iodide fluorescence. Ann Bot. 86(5):929-934.

Raghavan S, Arora MC (1958) Chromosome numbers in Indian medicinal plant-II,"Proc. Ind Acad Sci. 47:352.

Raheman H, Phadatare AG (2004) Diesel engine emissions and performance from blends of Karanja methyl ester and diesel. Biomass Bioenergy. 27:393–397.

Rai S, Das AB, Das P (1997) Estimation of 4C DNA and Karyotype Analysis in Ginger (Zingiber officinale Rosc.) I. Cytologia. 62:133-41.

Rayburn AL, Birdari DP, Bullock DG, Nelson RL, Gourmet C, Wentzel JB (1997) Nuclear DNA content diversity in Chinese soybean introductions. Ann Bot. 80:321-325.

Rival A, Beule T, Barre P, Hamon S, Duval (1997) Comparative flow cytometric estimation of nuclear DNA content in oil palm (*E. guineensis*) tissue cultures and seed-derived plants. Plant cell reports. 16: 884-887.

Ruuska SA, Ohlrogge JB (2001) Protocol for small-scale RNA isolation and transcriptional profiling of developing Arabidopsis seeds. Biotechniques. 31:752-758.

Sali A, Blundell TL (1993) Comparative protein modeling by satisfaction of spatial restraints. J Mol Biol. 234:779-815.

Salzman RA, Fujita T, Zhu-Salzman K, Hasegawa PM, Bressan RA (1999)An improved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates. Plant Mol Biol Rep.17:11-17.

Sambrook J, Russel D (2001) Molecular Cloning: A Laboratory Manual, volume2. 3rd edition. New York: Cold Spring Harbor Laboratory Press.

Sarbhoy RK (1977) Cytogenetical studies in *Pongamia pinnata* (L.) Pierre. *Cytologia* 42:415-423.

Sarma AK, Konwer D, Bordoloi PK (2005) A comprehensive analysis of fuel properties of biodiesel from Koroch seed oil. Energy Fuels. 19:656-7.

Sasaki T, Matsumoto T, Yammoto K, Sakata K, Baba T, Katayse (2002) The genome sequence and structure of rice chromosome. Nature. 420:312-316.

TH-1255_09610610

Schauser L, Christensen L, Borg S, PouIsen C (1995). PZF, a cDNA lsolated from Lotus japonicus and Soybean Root Nodule Libraries, Encodes a New Plant Member of the RI NG-Finger Family of Zinc-Binding Proteins. Plant Physiol.107: 457-1458.

Scheetz TE, Trivedi N, Roberts CA, Kucaba T, Berger B, Robinson NL, Birkett CL, Gabin AJ, Oleary B, Braun TA, Bonaldo MF, Robinson JP, Sheffield VC, Soares MB, Casavant TL (2003) ESTprep: Prprocessing cDNA sequence reads. Bioinformatics. 19:1318-1324.

Schneider G, Lindqvist Y, Shanklin J, Somerville C (1992) Preliminary crystallographic data for stearoyl-acyl carrier protein desaturase from castor seed. J Mol Biol. 225:561-564.

Scott PT, et al (2008) *Pongamia pinnata*: An untapped resource for the biofuels industry of the future. Bioenerg Res. 1:2-11.

Seffens WS, Almoguera D, Wilde HD, Vonder HRA, Thomas TL (1990) Molecular analysisof a phylogenetically conserved carrot gene: Developmental and environmental regulation. Dev Genet 11:65-76.

Shah S, Sharma A, Gupta MN (2004) Extraction of oil from *Jatropha curcas* L. seed kernels by enzyme assisted three phase partitioning. Indian Crop Prod. 20:275-279.

Shanklin J, Achim C, Schmidt H, Fos BG, Munck E (1997) Mossbauer studies of alkaline-hydroxylase: evidence for a diiron cluster in an integral membrane enzyme. Proc Natl Acad Sci USA. 94:2981-2986.

Shanklin J, Cahoon EB (1998) Desaturation and related modifications of fatty acids. Annu Rev Plant Physiol Plant Mol Bio. 49:611-641.

Shanklin J, Whittle E, Fox BG (1994) Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. Biochemistry. 33:12787-12794.

Shapiro MH. Practical Flow Cytometry: Text book

Sharma AD, Gill PK, Singh P (2002) RNA isolation from plant tissues rich in polysaccharides. Anal Biochem. 314:319-321.

Sharma YC, Singh B (2008) Development of biodiesel from karanja: a tree found inrural India. Fuel. 87(8-9):1740-1742.

Sharmin E, Ashraf SM, Ahmad S (2007) Synthesis, characterization, antibacterial and corrosion protective properties of epoxies, epoxy-polyols and epoxy-polyurethane coatings from linseed and *Pongamia glabra* seed oils. Int J Biol Macromol 40:407-422.

Shweta S, Sharma A, Gupta MN (2004) Extraction of oil from *Jatropha curcas* L. seed kernels by enzyme assisted three phase partitioning. Indian Crop Prod. 20:275-279.

Singh G, Kumar S, Singh P (2003) A quick method to isolate RNA from wheat and other carbohydrate-rich seeds. Plant Mol Biol Rep. 21:93-93.

Sliwinska E, Thiem B (2007) Genome size stability in six medicinal plant species propagated in vitro. Biol Plant. 51:556-558.

Sliwinska E, Zielinska E, Jedrzejczyk I (2005) Are seeds suitable for flow cytometric estimation of plant genome size?. Cytometry. 64:72-79.

TH-1255_09610610

Slocombe SP, Piffanelli P, Fairbairn D, Bowra S, Hatzopoulos P, Tsiantis M, Murphy DJ (1994) Temporal and tissue-specific regulation of a *Brassica napus* stearoyl-acyl carrier protein desaturase gene. Plant Physiol. 104:167-1176.

Somerville C, Browse J, Jaworski JG, Ohlrogge JB (2000) Lipids. In Biochemistry and Molecular Biology of Plants, B. Buchanan, W. Gruissem, and R. Jones, eds (Rockville, MD: American Society of Plant Physiologists). pp. 456-527.

Srivastava A, Prasad R (2000) Triglycerides-based diesel fuels. Renew Sust Energ Rev. 4(2):111-133.

Stoutjesdijk PA, Singh SP, Liu Q, Hurlstone CJ, Waterhouse PA, Green AG (2002) hpRNA-mediated targeting of the Arabidopsis *FAD2* gene gives highly efficient and stable silencing. Plant Physiol. 129:1723-1731.

Suda J, Leitch IJ (2010) The Quest for Suitable Reference Standards in Genome Size Research. Cytometry Part A. 77A:717-720.

Suda J, Travnicek P (2006) Reliable DNA ploidy determination in dehydrated tissues of vascular plants by DAPI flow cytometry-New prospects for plant research. Cytometry Part A. 69A:273-280.

Suzuki Y, Sugano S (2003) Construction of a full-length enriched and a 5′-end enriched cDNA library using the oligo-capping method. Methods Mol Biol. 221:73–91.

Tasseva G, de Virville JD, Cantrel C, Moreau F, Zachowski A (2004) Changes in the endoplasmic reticulum lipid properties in response to low temperature in *Brassica napus*. Plant Physiol Biochem. 42:811-822.

Temsch E, Greilhuber J, Krisai R (2010) Genome size in liverworts. Preslia 82:63-80.

Thiem B, Sliwinska E (2003) Flow cytometric analysis of nuclear DNA content in cloudberry (*Rubus chamaemorus* L.) in vitro cultures. Plant Science.164(1):129-134.

Thompson GA, Scherer DE, Foxall-Van Aken S, Kenny JW, Young HL, Shintani DK, Kridl JC, Knauf VC (1991) Primary structure of the precursor and mature forms of stearoyl-acyl carrier protein desaturase from saZower embryos and requirement of ferredoxin for enzyme activity. Proc Nat Acad Sci. USA. 88:2578-2582.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25:4876-4882.

Tong L, Dan-wei MA, Ying XU, Wu-yuan D, Meng X, Ren-wei Q, Fang C (2007) Cloning and characterization of a stearoyl-ACP desaturase gene from *Jatropha curcas*. J Shanghai Univ. 11(2):182-188.

Transferrin Receptor and Lowers its Affinity for Ligand Binding. Proc Natl Acad Sci. 95:1472-1477.

Tuna M, Vogel KP, Arumuganathan K, Gill KS (2001) DNA content and ploidy determination of Bromagrass Germplasm Accessions by flow cytometry. Crop sci. 41:1629-1634.

Vicient CM, Delseny M (1999) Isolation of total RNA from Arabidopsis thaliana seeds. Anal Biochem. 268:412-413.

TH-1255_09610610

Vindelov LL, Christensen IJ, Jensen GJ, Nissen NI (1983) Limits of detection of nuclear DNA abnormalities by flow cytometric DNA analysis. Results obtained by a set of methods for sample-storage, staining and interval standardization. Cytometry. 3:332-339.

Vrinten P, et al (2005) Two FAD3 desaturase genes control the level of linolenic acid in flax seed. Plant Physiol. 139:79-87.

Walker DJ, Monino I, Correal E (2006) Genome size in *Bituminaria bituminosa* (L.) C.H. Stirton (Fabaceae) populations: separation of 'true' differences from environmental effects on DNA determination. Env Env Botany. 55:258-265.

Walker EA (2006) The genomic rate of adaptive evolution. Trends Ecol Evol. 21:569-575.

Wallis JG, Browse J (2002) Mutants of Arabidopsis reveal many roles for membrane lipids. Progress in Lipid Research. 41(3):254-278.

Weber H (2002) Fatty acid-derived signals in plants. Trends Plant Sci. 7:217-224.

Weber J, Georgiev V, Pavlov A, Bley T (2008) Flow cytometric investigations of diploid and tetraploid plants and in vitro cultures of *Datura stramonium* and *Hyoscyamus niger*. Cytometry A. 73:931-939.

West JP, Cowley EJ (1993) Floristic notes and chromosome numbers of some Chinese *Roscoea* (Zingiberaceae). Kew Bulletin. 48:799-803.

Xiao G, Zhang ZQ, Liu RY, Yin CF, Wu XM, Tan TL, Guan CY (2013) Molecular cloning and characterization of a novel gene involved in fatty acid synthesis in Brassica napus L. J Integrative Agri. 12(6):962-970.

Zaborowska Z, Starzycki M, Femiak I, Swiderski M, Legocki AB (2002) Yellow lupine gene encoding stearoyl-ACP desaturase-organization, expression and potential application. Acta Biochimica Polonica. 49:29-42.

New references included:

Bandana B, Paul TS, Peter MG (2011) Tree legumes as feedstock for sustainable biofuel production: Opportunities and challenges. J. Plant Physiol.168:1877-1884.

Bandana B, Stephen HK, Qunyi J, Sharon S, Peter MG, Paul TS (2013) Genetic and Genomic Analysis of the Tree Legume *Pongamia pinnata* as a Feedstock for Biofuels. The Plant Genome 6, doi: 10.3835/plantgenome2013.05.0015.

Kazakoff SH, Imelfort M, Edwards D, Koehorst J, Biswas B (2012) Capturing the Biofuel Wellhead and Powerhouse: The Chloroplast and Mitochondrial Genomes of the Leguminous Feedstock Tree *Pongamia pinnata*. PLoS ONE 7(12):e51687. doi:10.1371/journal.pone.0051687.

Sharon S, Paul TS, Peter MG (2013) Nodulation in the Legume Biofuel Feedstock Tree *Pongamia pinnata*. Agricultural Res. 2(3):207-214.

TH-1255_09610610

Publications and Biography

TH-1255_09610610

PUBLICATIONS

Journals Published:

- Aadi Moolam Ramesh, Vigya Kesari, Latha Rangan (2013). Characterization of a stearoyl-acyl carrier protein desaturase gene from potential biofuel plant, *Pongamia pinnata*. L. *Gene*. 542:113-121
- Aadi Moolam Ramesh, Supriyo Basak, Rimjhim Roy Chowdhury, Latha Rangan (2013). Development of Flow Cytometric Protocol for Nuclear DNA Content Estimation and Determination of Chromosome Number in *Pongamia pinnata* L., a Valuable Biodiesel Plant. Applied Biochemistry and Biotechnology. 172(1):533-548.
- Rimjhim Roy Choudhury, Supriyo Basak, Aadi Moolam Ramesh and Latha Rangan (2013). Nuclear DNA content of *Pongamia pinnata* L. and genome size stability of in vitro regenerated plantlets. *Protoplasma*. 251:703-709.

Other Publications:

- Vigya Kesari, Aadi Moolam Ramesh and Latha Rangan (2013). *Rhizobium pongamiae* sp. nov., from Root Nodules of *Pongamia pinnata. BioMed Research International* Volume 2013, http://dx.doi.org/10.1155/2013/165198.
- Vigya Kesari, Aadi Moolam Ramesh and Latha Rangan (2012). High frequency direct organogenesis and evaluation of genetic stability for in vitro regenerated *Pongamia pinnata*, a valuable biodiesel plant. *Journal of Biomass and Bioenergy*. 44:23-32.
- Mohana Subramanian Bhaskaran, Lakshmipriya Venkatesan, Ramesh Aadimoolam, Harikrishnan Tirunelveli Jayagopal and Rajan Sriraman (2010). Sequence diversity of internal transcribed spacer-1 (ITS-1) region of *Eimeria* infecting chicken and its relevance in species identification from Indian field samples. *Parasitol Res*.106:513-521.

Book chapter:

- Vigya Kesari, Aadi Moolam Ramesh and Latha Rangan (2013). "*Pongamia pinnata*: The biodiesel plant biology, tissue culture and genetic enhancement" *Biotechnological Applications for Environmental Protection*, In: Abhilash P.C. Springer (Accepted).

GenBank submissions:

- Ramesh, A.M., Kesari, V. and Rangan, L. NCBI GenBank Accession number KF192317-*Millettia pinnata* steroyl-ACP desaturase mRNA, complete cds.
- Ramesh, A.M., Kesari, V. and Rangan, L. NCBI Geneank Accession number KF651985-*Millettia pinnata* microsomal omega-6-desaturase (FAD2-1) mRNA, complete cds; nuclear gene for microsomal product.
- Ramesh, A.M., Kesari, V. and Rangan, L. NCBI Geneank Accession number KF651986-*Millettia pinnata* microsomal omega-6-desaturase (FAD2-2) mRNA, complete cds; nuclear gene for microsomal product.
- Ramesh, A.M., Kesari,V. and Rangan, L. NCBI GenBank Accession number HM626172-*Rhizobium* sp. VKLR01 *Atp*D gene, partial cds.
- Kesari, V., Ramesh, A.M. and Rangan, L. NCBI GenBank Accession number HM626171-*Rhizobium* sp. VKLR01 *rec*A gene, partial cds.

TH-1255_09610610

Participations

Training and Workshops

- 13th Indo-US workshop on "Applications of Flow cytometry in Plant genomics and nanotechnology" at IIT Guwahati, from 8th to 10th October, 2012. Organized by Dr. Latha Rangan (IITG) along with University of Miami, USA.
- RT PCR (Applied Biosystems) Training undergone from 2-4 Aug, 2010; conducted by Lab India, Gurgoan, India.

Conference proceedings:

- Aadi Moolam Ramesh, Vigya Kesari and Latha Rangan (2014). Characterization of a stearoyl-acyl carrier protein desaturase gene from potential biofuel plant, *Pongamia pinnata*, National conference on SCIENCE OF OMICS FOR AGRICULTURAL PRODUCTIVITY: FUTURE PERSPECTIVES. 4-6th March, Silver Jubilee year Pantnagar Biotechnology Programme, G.B. Pant University of Agriculture & Technology, Pantnagar (Uttarakhand, India). Under the aegis of Society for Plant Biochemistry & Biotechnology, New Delhi, pp 181 (**Awarded as Young Scientist**)
- Aadi Moolam Ramesh, Tushar and Latha Rangan (2012). Next generation sequencing and selective analysis of non-coding DNA in *Pongamia pinnata*. *International conference on Industrial Biotechnology (ICIB-2012)*, 21-23 November, Patiala University, pp 210. (Poster)
- S Basak, A M Ramesh, L Rangan (2012). Ploidy levels of potential biodiesel crops determined by flow cytometry. *World Congress on In Vitro Biology*, Seattle, USA, 3-7 June 2012, pp 3001. (Poster)
- A M Ramesh, V Kesari, L Rangan (2010). Characterization of *Rhizobium pongamiae* sp. nov., isolated from root nodules of Biodiesel plant *Pongamia pinnata*. *International Conference on Genomic Sciences-Recent Trends* (ICGS-2010), 12-14th November, 2010, MKU, Madurai, pp 112. (Poster)
- Kota Sathish, B Mohana Subramanian, Aadi Moolam Ramesh, Nirujogi Hanumantha Rao, Rajan Sriraman and VA Srinivasan (2009). Plant expressed EtMIC-1 is an effective immunogen in conferring protection against *Eimeria tenella* infection in chickens. *Plant Expression Systems for Recombinant Pharmacologics*, University of Verona, Italy (15-17 June). (Poster).
- A M Ramesh, V Kesari and L Rangan (2011). *Rhizobium pongamiae* sp. nv., from root nodlules of a versatile legume *Pongamia pinnata* L. National conference on *Emerging trends in Biotechnology* & Annual meeting of Society for Biotechnology. 24-26th September, ANU, Guntur, pp 97. (Oral)
- A M Ramesh, A Das, N Kasoju, U Bora and L Rangan (2011). Chemical and biological investigation of Essential oil of *Z. moran*, an endemic species from North East India. National conference on *Emerging trends in Biotechnology* & Annual meeting of Society for Biotechnology. 24-26th September, ANU, Guntur, pp 162. (Oral)
- V Kesari, A M Ramesh, L Rangan (2010). Characterization in candidate plus trees (CPTs) of *Pongamia pinnata* (L.) Pierre, a versatile legume from North Guwahati. A review. 33rd Conference of *Indian Botanical Society and International Symposium on the New Horizons of Botany*, 10-12th Nov, Shivaji University, Kolhapur, Maharashtra. p 217. (Poster)
- Vigya Kesari, R Sarath, Aadi Moolam Ramesh, MS Vinod, A Parida and L Rangan (2010). Morphological and Bio-molecular Approach in CPTs of *P. pinnata*, a promising crop from North Guwahati. *Energy and environmental sciences section*, at IYSC2010, 26-28th June, SRM University, Chennai. (Poster)

TH-1255_09610610

Contact:
Department of Biotechnology
Indian Institute of Technology Guwahati
Guwahati-781039
Mobile: +91-9678554474
Email:
aadi@iitg.ac.in
ramesh.santu@gmail.com

AADI MOOLAM RAMESH

Mr. Aadi Moolam Ramesh joined as a doctoral student in the Department of Biotechnology, Indian Institute of Technology Guwahati (IITG), India in July, 2009. He has carried out a multidisciplinary research during his doctoral study period under the joint supervision of Dr. Latha Rangan and Dr. Bithiah G Jaganathan at IITG. His research interests mainly include Plant Molecular Biology and Biotechnology.

Mr. Ramesh has completed B.Sc. (Microbiology) with specialization in Chemistry and Botany in 2002 from Govt. Degree college, Gajwel (Affiliated to Osmania University, Hyderabad), Andhra Pradesh, India. After that he has perceived his Master's (M.Sc) in Biotechnology from SRTM University, Nanded, Maharashtra. After his M.Sc, immediately he joined for Agriculture department, Govt. of Andhra Pradesh, where he worked as Multipurpose extension officer for one and half year. Later he secured all India rank 445 with 95.54 percentile in Graduate Aptitude Test in Engineering (GATE) organized by IIT Kharagpur and got the scholarship from 2006-2008 to pursue his master's degree (M.Tech) in Biotechnology from JNTUH campus, Hyderabad, India. During his post-graduation (M. Tech), he has developed a recombinant vaccine against chicken Coccidiosis at Indian Immunologicals Ltd., (IIL) Hyderabad, India and was published in peer reviewed journal.

Before joining IITG for his doctoral studies, he has worked on BCIL (Biotech consortium India Ltd., Govt. of India) funded project "Evaluation of sub-unit vaccine against chicken Coccidiosis" at Plant Molecular Biology Lab, Research & Development section, IIL, Hyderabad, India.

Currently he has completed the doctoral study with dedication and interest in multidisciplinary research. His doctoral research mainly focused on cyto-genetic studies in biofuel crop *Pongamia* towards understanding the germplasm to improve its key traits. During his doctoral studies he has undergone 13th Indo-US workshop on "Applications of Flow cytometry in Plant genomics and Nanotechnology, organized by IITG & University of Miami at IITG during October, 2012. Furthermore he has undergone a training programme on real-time PCR related to his doctoral research sponsored by Applied Biosystems, India in 2010. During his doctoral studies he has got fellowships from various funding bodies like DIT, DST, MHRD and DBT, Govt. of India. A part of his doctoral research has been published in international peer reviewed journals and some are under review. Since 2012, he is a student member in the International Society for the Advancement of Cytometry (ISAC). He is keenly interested to continue further research in the area of Plant Molecular Biology, particularly on genome characterization of important biofuel plants.

TH-1255_09610610