UNIVERSITI PUTRA MALAYSIA

CHARACTERIZATION OF EXPRESSED GENES ISOLATED FROM OIL PALM VEGETATIVE, NORMAL AND ABNORMAL INFLORESCENCE MERISTEMS USING EST APPROACH

LEE YANG PING

FSMB 2003 28

CHARACTERIZATION OF EXPRESSED GENES ISOLATED FROM OIL PALM VEGETATIVE, NORMAL AND ABNORMAL INFLORESCENCE MERISTEMS USING EST APPROACH

LEE YANG PING

MASTER OF SCIENCE
UNIVERSITI PUTRA MALAYSIA

2003

UPM

CHARACTERIZATION OF EXPRESSED GENES ISOLATED FROM OIL PALM VEGETATIVE, NORMAL AND ABNORMAL INFLORESCENCE MERISTEMS USING EST APPROACH

By

LEE YANG PING

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirement for the Degree of Master of Science

July 2003

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

CHARACTERIZATION OF EXPRESSED GENES ISOLATED FROM OIL PALM VEGETATIVE, NORMAL AND ABNORMAL INFLORESCENCE MERISTEMS USING EST APPROACH

By

LEE YANG PING

July 2003

Chairman : **Harikrishna a/l Kulaveerasingam, Ph.D.**

Faculty : **Food Science and Biotechnology**

Little is known about the function of genes expressed in oil palm vegetative meristem tissues, which are expressed during transition from vegetative to reproductive growth, and those expressed during the formation of abnormal inflorescence. This study was aimed at isolating sufficiently large numbers of expressed sequence tags (ESTs) from an oil palm vegetative meristem cDNA library so that the expression of genes in this tissue could be studied. The genes (or ESTs) that were specifically expressed in vegetative meristems at early stages of normal inflorescence meristem development, and in abnormal inflorescence meristems were isolated to study the transition from vegetative to reproductive growth and the formation of floral abnormalities from clonal palms.

A random EST approach has been used to obtain vast amounts of genes from many organisms. However, the random EST approach may result in the isolation of clones that are highly repeated in proportion to their abundance in the mRNA

population of the appropriate tissues. Therefore, cold-plaque screening and suppression subtractive hybridization (SSH) techniques were employed in the EST approach to isolate as much unique transcripts as possible in this study.

Based on the EST collections made from the vegetative meristem tissues, 1088 ESTs were isolated. The redundancy of the ESTs was reduced to about 18.9% where 81.1% out of 1088 oil palm vegetative meristem ESTs were unique due to the use of the cold-plaque screening approach. Classification of the putative function of ESTs provides an idea of the type of genes expressed in the vegetative meristem tissue. The expressed genes range from housekeeping genes to genes related to photosynthesis. About 44% of the ESTs were unknown and were not characterized in other species.

Stage specific expressed genes were isolated from vegetative meristem, inflorescence meristem, normal and abnormal inflorescence meristem subtraction libraries. About 601 contigs were identified from these four subtraction libraries. More than 57% of the ESTs encoded unknown functions. Through the use of digital differential display calculations, the ESTs were expressed in a tissue specific manner. Reverse northern analysis confirmed that the the majority of the EST-contigs were tissue specific.

This study reveals that the apical meristem devotes more cellular activity to the biosynthesis of cellular components than to photosynthesis which is predicted from the EST analysis and physiological experiments conducted on plants. This study also reveals profound changes in gene expression that are involved in the transition

from the vegetative meristem to an inflorescence meristem that is driven by the action of a series of genes expressed in a stage-specific manner. Significant differences in gene expression between normal and abnormal inflorescence meristems could not be detected with this EST approach. Transient changes in gene expression or epigenetic phenomenon have been proposed to explain the molecular basis of abnormal floral development.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

PENYIFATAN GEN-GEN TEREKSPRES TERPENCIL DARIPADA MERISTEM VAGETATIF, NORMAL DAN ABNORMAL INFLORESEN KELAPA SAWIT DENGAN MENGUNAKAN KAEDAH EST

Oleh

LEE YANG PING

Julai 2003

Pengerusi : **Harikrishna a/l Kulaveerasingam, Ph.D.**

Fakulti : **Sains Makanan dan Bioteknologi**

Pengetahuan mengenai fungsi gen-gen yang diekspres di dalam tisu meristem vagetatif kelapa sawit, pertumbuhan daripada fasa vagetatif kepada fasa pembiakan dan juga pembentukan infloresen abnormal adalah terhad. Tujuan kajian ini adalah untuk memencil bilangan tag urutan terekspres " expressed sequence tags (EST)" yang mencukupi daripada perpustakaan tisu meristem vagetatif kelapa sawit dan ini membolehkan gen-gen ekpres di dalam tisu ini dapat dikaji. Gen-gen (atau ESTs) yang diekspres secara spesifik di meristem vagetatif, awal pertumbuhan infloresen meristem normal dan di dalam infloresen meristem abnormal telah dipencil untuk pengajian pertumbuhan daripada fasa vagetatif kepada fasa pembiakan bunga dan juga pembentukan infloresen abnormal daripada klon-klon pokok kelapa sawit.

Kaedah EST rawak telah digunakan untuk mendapatkan jumlah gen-gen yang besar daripada organisma-organisma. Tetapi, kaedah ini menghasilkan pengulangan klon-klon yang tinggi yang berkadar langsung kepada perkadaraan populasi mRNA untuk tisu-tisu berkenaan. Oleh itu, kaedah penyalingan "cold-plaque" dan teknik

untuk tisu-tisu berkenaan. Oleh itu, kaedah penyalingan "cold-plaque" dan teknik "suppression subtractive hybridization (SSH)" digunakan di dalam kaedah EST untuk memencil sebanyak mungkin transkrip-transkrip yang unik.

Berdasarkan kepada koleksi EST dibuat daripada tisu meristem vagetatif, sebanyak 1088 EST telah dipencil. Redandensi EST telah dikurangkan kepada 18.9% di mana 81.1% daripada 1088 EST meristem vagetatif kelapa sawit adalah unik setelah menggunakan kaedah penyalingan "cold-plaque". Klasifikasi berkemungkinan fungsi EST dapat mempamirkan idea jenis gen-gen yang diekspres di dalam tisu meristem vagetatif. Jenis gen yang diekspres berlingkung daripada gen-gen pengekalrumahan kepada gen-gen berkaitan dengan fotosintesis. Lebih kurang 44% daripada fungsi EST masih tidak diketahui dan tidak dikaji di mana-mana spesis lain.

Gen-gen yang diekspres secara tahap spesifik telah dipencil daripada perpustakaan penolakan meristem vagetatif, normal dan abnormal meristem infloresen. Lebih daripada 57% EST mengekod protein yang masih tidak diketahui fungsinya. Dengan menggunakan penghitungan pembezaan pamir digital, EST adalah diekspres secara bentuk spesifik. Analisis "reverse northern" mengenalpasti kebanyakan EST diekspres secara tisu spesifik.

Kajian ini menunjukan bahawa aktiviti meristem pucuk lebih cenderong kepada aktiviti penghasilan komponen selular daripada fotosintesis yang dijangka daripada analisis EST dan eksperimen fisiologi yang dijalankan ke atas tumbuhan. Kajian ini juga menunjukan perubahan pengekspresian gen yang melibatkan perubahan

daripada meristem vagetatif kepada meristem infloresen adalah dijana oleh aksi sesiri gen-gen yang diekspres secara bentuk tahap-spesifik. Perbezaan yang siknifikan di dalam pengekspresian gen antara normal dan abnormalmeristem bunga tidak dapat dikesan dengan kaedah EST. Perubahan "transient" dalam pengekspresian gen atau fenomena "epigenetic" telah dibentang untuk menjelaskan pertumbuhan infloresen abnormal secara asas molekular.

ACKNOWLEDGEMENTS

This thesis has been designed and developed with the aid of many people. First of all, I would like to thank Associate Professor Dr. K. Harikrishna, my supervisor for his valuable contribution to the work described in this thesis. Also to express my appreciation to Dr. Ho Chai Ling and Dr. Sharifah Shahrul Rabiah bte Syed Alwee, my supervisory committee members for their guidance and advise during the course of this project.

My appreciation is also extended to Mr. Ong, the Lab assistant of the plant molecular biology and genetics laboratory for helping me to process my purchase and ordering chemicals.

Besides that, I would like to thank Dr. Tan Siang Hee, Dr. Faridah, Mr. Choong Chieh Wen and Mr. Lee Weng Wah who always there to help me and guide me throughout my studies in UPM.

A thankful to my lab-mates: Pick Kuen, Pao Theen, Wai Har, Chee How, Yun Shin, Alfred Wong, Seddon, Chuen Yi, Yen Yen, Yeen Yee, Siew Mee, Adelene Au Yong, Tony Wong, Swee Sen, Kean Jin, Gaik Theng, Sock Hwa, Pak Kuan, Joey Ng, Tzer Ying, Zie, Rosilah and many others, for their advice and technical assistance.

Last but not least, to my family and house mates for their love, encouragement and understanding throughout my studies in UPM.

TABLE OF CONTENTS

		Page
ABSTRACT		**II**
ABSTRAK		**V**
ACKNOWLEDGEMENT		**VIII**
APPROVAL SHEETS		**IX**
DECLARATION FORM		**XI**
LIST OF TABLES		**XV**
LIST OF FIGURES		**XVI**
LIST OF ABBREVIATIONS		**XVIII**
CHAPTER		
1	**INTRODUCTION**	1
2	**LITERATURE REVIEW**	5
2.1	Developmental Stages of Oil Palm	5
2.2	Floral Abnormalities in Clonal Oil Palms	8
2.3	Shoot Apical Meristem	10
2.4	The Transition to Flowering	12
2.4.1	Environmental Signals	12
2.4.2	Endogenous Signals	14
2.5	Floral Organ Formation	16
2.6	Abundance of mRNA in Flowering Plants	19
2.7	The Regulation of Gene Expression In Plants	20
2.8	Genomics Research in Plants	25
2.9	Large-Scale Sequencing of Plant Genomes	28
2.9.1	Genome Sequencing	28
2.9.2	Expressed Sequence Tag	29
2.10	Functional Analysis of Plant Genes	32
2.10.1	DNA Microarray Technology	32
2.10.2	Serial Analysis of Gene Expression	33
2.10.3	Gene Trap System	34
2.11	Bioinformatics	38
3	**MATERIALS AND METHODS**	41
3.1	Plant Materials	41
3.2	Total RNA Extraction and Quantification	41
3.3	Purification of Poly A^+ mRNA from Total RNA	44
3.4	Oil Palm Vegetative Meristem cDNA Library Construction	45
3.4.1	Synthesizing First-Strand cDNA	45
3.4.2	Synthesizing Second-Strand cDNA	45
3.4.3	Blunting the cDNA Termini	46
3.4.4	Ligation of *Eco*RI-Adaptor	47
3.4.5	Phosphorylating the *Eco*R I Ends	47

3.4.6 Digesting With *Xho* I 48
3.4.7 Size Fractionation 48
3.4.8 Quantitating the cDNA With Ethidium Bromide Plate Assay 49
3.4.9 Ligating cDNA into the Uni-ZAP XR Vector 50
3.4.10 Preparing the Host Bacteria 50
3.4.11 Packaging the Ligated Product 50
3.4.12 Plating and Titering 51
3.4.13 Amplification of the Oil Palm Vegetative Meristem cDNA Library 51
3.5 Synthesis of Double Stranded cDNA for Probe Labeling 52
3.6 Non-Radioactive Nucleic Acid Labeling 55
3.7 Cold-Plaque Screening 55
3.7.1 Plating 55
3.7.2 Lifting 56
3.7.3 Hybridization and Primary Screening 57
3.7.4 Secondary Screening 59
3.8 Isolation of cDNA Clones 59
3.8.1 *In vivo* Excision 59
3.8.2 Plasmid Isolation 60
3.8.3 PCR Amplification 61
3.9 Dot-Blot 62
3.10 Reverse Northern Hybridization 63
3.11 Subtractive Library Construction 63
3.11.1 Vegetative Meristem Versus Inflorescence Meristem 65
3.11.1.1 *Rsa* I-Digestion 65
3.11.1.2 Adaptor Ligation 66
3.11.1.3 First Hybridization 67
3.11.1.4 Second Hybridization 68
3.11.1.5 PCR Amplification 69
3.11.1.6 Cloning of the Differentially Expressed cDNAs 70
3.11.1.7 Dot Blot and Reverse Northern Hybridization Screening 72
3.11.2 Normal Inflorescence Meristem Versus Abnormal Inflorescence Meristem 72
3.12 Sequencing of cDNA Clones 73
3.13 Computational Analysis of ESTs 74
3.14 Northern Blot and Hybridization 75

4 RESULTS 77

4.1 Oil Palm Vegetative Meristem cDNA Library Construction and Screening 77
4.1.1 Total RNA Isolation and mRNA Purification 77
4.1.2 Oil Palm Vegetative Meristem cDNA Library Construction 78

4.1.3 Cold-Plaque Screening 79
4.1.4 Reverse Northern Analysis 83
4.1.5 Computational Analysis of ESTs 84
4.2 Oil Palm Subtractive Library Construction 87
4.2.1 Oil Palm Vegetative Meristem Versus Early Young Inflorescence Meristems 87
4.2.2 Oil Palm Normal Versus Abnormal Inflorescence Meristem 90
4.2.3 Reverse Northern Analysis 92
4.2.4 Computational Analysis of ESTs 99
4.2.5 Northern Analysis 104

5 DISCUSSION 106

5.1 Cold-Plaque Screening 106
5.2 Subtraction Library Construction 108
5.3 Classification of Putative Function of ESTs 111
5.4 Gene Expression In Vegetative Shoot Apical Meristem 113
5.5 Stage-Specific Expressed Genes 123

6 CONCLUSION 132

BIBLOGRAPHY 134

APPENDICES 148
Appendix 1 148
Appendix 2 150

BIODATA OF THE AUTHOR 175

LIST OF TABLES

Table		Page
1	Phenotype of plants with reduced levels of DNA methylation or histone deacetylase activity (Finnegen, 2001)	26
2	The signal intensity of OPVMS, OPIMS, OPNIMS, and OPAIMS cDNA transcripts when hybridized with their corresponding pair of probes	93

LIST OF FIGURES

Figure		Page
1	Flow chart showing the basic steps used to enrich for differentially expressed sequences in the PCR-select procedure.	64
2	Isolation of total RNA from 6 month-old oil palm vegetative meristem tissues.	77
3	Size fractionation of synthesized cDNA.	78
4	Optimization of PCR cycles for the oil palm vegetative meristem cDNA synthesis.	80
5	Signal intensity generated from the hybridization of the labelled oil palm vegetative meristem cDNA probe with oil palm vegetative meristem cDNA library.	81
6	Amplification of cDNA inserts of opvm clones with T3 and T7 primers of pBluescript SK+ vector.	81
7	The distribution of the insert sizes of the analyzed cDNA clones from the oil palm vegetative cDNA library.	82
8	Example of a reverse Northern analysis of oil palm vegetative meristem cDNA inserts.	83
9	Functional categorization of 934 non-redundant oil palm vegetative meristem ESTs that were putatively identified.	86
10	PCR cycle optimization using SMART PCR cDNA Synthesis system..	89
11	PCR cycle optimization using the PCR-select cDNA subtraction system.	89
12	Amplification of cDNA inserts of OPVMS (A) and OPIMS (B) clones with M13 reverse and T7 primers of pCRBlunt vector.	90
13	PCR cycle optimization of normal and abnormal oil palm inflorescence meristem cDNA synthesis.	91
14	PCR cycle optimization using the PCR-select cDNA subtraction system. Optimum cycles were determined at 17 cycles for both normal and abnormal inflorescence meristem.	91
15	Amplification of cDNA inserts of OPNIMS (A) and OPAIMS (B) clones with M13 reverse and T7 primers.	92

16 Dot-blot reverse northern analysis of OPVMS gene fragments. 95

17 Dot-blot reverse northern analysis of OPIMS gene fragments. 96

18 Dot-blot reverse northern analysis of OPNIMS gene fragments. 97

19 Dot-blot reverse northern analysis of OPAIMS gene fragments. 98

20 Comparison of expressed genes between vegetative meristem and inflorescence meristem tissues. 103

21 Comparison of expressed genes of normal and abnormal Inflorescence meristem tissues. 103

22 Figure 22: Northern blot analysis with (A) shaggy-related protein kinase (OPNIMS028) cDNA (gene specific) probe. Weak signal were detected from each type of tissue. (B) SocE (OPNIMS166) cDNA (gene specific) probe. Signal undetected from each type of tissue(C) alpha -tubulin cDNA probe as control for RNA loading for young leaf (YL), vegetative meristem (VM), normal inflorescence meristem (NIM), and abnormal inflorescence meristem (AIM) RNA. 105

LIST OF ABBREVIATIONS

α	Alpha
β	Beta
λ	Lambda
µg	Microgramme
µl	Microliter
°C	Degree centrigrade
%	Percentage
Amp	Ampicillin
bp	Base pair
BLAST	Basic local alignment search tool
BSA	Bovine Serum Albumin
cm	Centimeter
DNA	Deoxyribonucleic acid
Dnase I	Deoxyribonuclease 1
cDNA	Complementary DNA
dNTPs	Deoxynucleotides
dATP	2'-deoxy-adenosine-5'-triphosphate
dCTP	2'-deoxy-cytidine-5'-triphosphate
dGTP	2'-deoxy-guanosine-5'-triphosphate
dTTP	Thymidine-5'-tryphosphate
dH_2O	Distilled water
DEPC	Dietyl pyrocarbonate
DTT	Dithiothreitol

EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol bis-(β-aminoethyle ether)
EtBr	Ethidium bromide
g	Gramme
GTE	Glucose-Tris-EDTA
hr	Hour
Jacq.	Jacquin
kb	Kilobase-pair
LB	Luria-Bertani
LiCl	Lithium Chloride
M	Molar
MADS	MCM1-AGAMOUS-DEFICIENS-SRF
mg	Milligram
min	Minute(s)
mm	Millimeter
mM	Millimolar
MOPS	3-(N-morpholino) propanesulfonic acid
MPOB	Malaysian Palm Oil Board
mRNA	Messenger RNA
NaCl	Sodium Chloride
NaOH	Sodium Hydroxide
NCBI	National Center for Biotechnology Information
ng	Nanogramme
OD	Optical Density
ORF	Open reading frame

PCI	Phenol : chloroform : isoamyl
PCR	Polymerase Chain Reaction
pfu	plaque forming units
Poly $(A)^+$	polyadenylated (mRNA)
PVP	Polyvinylpyrrolidone
PVPP	Polypolyvinylpyrrolidone
RNA	Ribonucleic Acid
RNase	Ribonuclease
rpm	Revolution Per Minute
SDS	Sodium Dodecyl Sulphate
SSC	Sodium Chloride-Sodium Citrate buffer
TAE	Tris Acetate EDTA
TE	Tris-HCL-EDTA
UV	Ultraviolet
v/v	Volume per volume
w/v	Weight per volume

CHAPTER 1

INTRODUCTION

Oil palm constitutes one of our most important natural resources. It provides not only fibres and wood but also many biochemicals, such as oils. It is one of the economically important oil bearing crops, the highest oil producer per unit land area in the world. At present, palm oil production is second only to that of soybean oil in terms of world vegetable oil production.

The challenge that the oil palm industry in Malaysia will face in the 21st century is its ability to maintain competitive and remain profitable while facing labour shortages and limited land resources. Furthermore, in the years to come, the demand for palm oil is expected to increase. In order to fulfil the increasing demand for palm oil, an improvement in yield is required. Floral development is an important introductory step that leads to fruit formation.

Flowering can be defined as the transition from vegetative to reproductive growth. This transition is manifested as a change in properties of the shoot apical meristem, which stops producing leaves and instead starts producing the floral meristems that give rise to flowers (Hong, 1998). Therefore, a deep understanding of the mechanisms regulating gene expression during the transition from vegetative stage to inflorescence or flowering stage will facilitate the genetic improvement of oil palm.

Very little is known about the genes that mediate the early phase of meristem progression. It is this stage of meristem conversion that defines the boundary between vegetative and reproductive growth (Colasanti and Sundarresan, 1996). Therefore, several approaches could be taken to study floral development and to identify various genes that are expressed in the various tissues.

An important factor that affects floweing and oil production is abnormalities in flower development. Since palm oil is derived from fruits that are products of flowering, some basic knowledge of the molecular aspects of abnormalities at an early stage of development would be useful.

One of the experimental approaches used to examine the molecular basis of flower development was to isolate single genes from various stages of floral tissues by means of conventional molecular biology methods such as differential screening of appropriate cDNA libraries (Shahrul, 1998) and a mRNA fingerprinting approach to isolate flower-specific cDNA in oil palm (Singh and Cheah, 2000). These techniques, although useful in directly identifying genes of interest, are however, time consuming, laborious and require large amounts of starting materials.

In recent years, a random EST approach has been used to obtain vast amounts of gene resources and has revolutionized the means by which functional genes can be identified and isolated from many organisms. After the first report on this approach by Adams *et al.* (1991), EST information has been rapidly accumulated (Adams *et al.*, 1992; Okubo *et al.*, 1992; Cooke *et al.*, 1996; Ablett *et al.*, 2000) and has become a major constituent of gene expression studies, and is also seen as an inexpensive and

rapid mean to identify large numbers of expressed genes from several developmental stages of flowering.

However, the random EST approach may result in the isolation of clones that are highly repeated in proportion to their abundance in the mRNA population of tissues. Thus, genes expressed at very low levels are not likely to be found within EST clones unless large numbers of ESTs are sequenced. These low abundance genes may be involved in important functions during floral development. To address this "expression bias", a cold-plaque screening technique (Hodge *et al.*, 1992) and a suppression subtractive hybridization method (Diatchenko *et al.*, 1996) have been used to reduce the prevalence of cDNAs corresponding to abundant transcripts as well as those expressed constitutively.

Bioinformatics plays a central role in EST data management and facilitates gene identification in a high throughput manner. Bioinformatics software such as BLAST (Altschul *et al.*, 1997) and Spotfire (Spotfire, USA), relate to the acquisition, archiving, deposition, retrieval, analysis, interpretation and display of biologically important data so as to answer biological problems such as, what a newly discovered gene does and what place it has in the complex mechanisms of the organisms life cycle.

The objective of this study was to isolate sufficiently large numbers of ESTs derived from an oil palm vegetative meristem cDNA library that can provide preliminary information about the expression of genes of that specific tissue. The differentially expressed genes between two set of tissues: "vegetative meristem and