UNIVERSITI PUTRA MALAYSIA

IDENTIFICATION AND ANALYSIS OF TWO GENES DIFFRENTIALLY EXPRESSED IN SPEAR LEAVES OF HIGH AND LOW YIELDING OIL PALM

ROSLINDA ABU SAPIAN

FBSB 2007 23

IDENTIFICATION AND ANALYSIS OF TWO GENES DIFFRENTIALLY EXPRESSED IN SPEAR LEAVES OF HIGH AND LOW YIELDING OIL PALM

By

ROSLINDA ABU SAPIAN

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

FEBRUARY 2008

Abstract of thesis presented to the Senate of University Putra Malaysia
in fulfillment of the requirement for the degree of Master of Science

IDENTIFICATION AND ANALYSIS OF TWO GENES DIFFERENTIALLY EXPRESSED IN SPEAR LEAVES OF HIGH AND LOW YIELDING OIL PALM

By

ROSLINDA ABU SAPIAN

February, 2008

Chairman : Associate Professor Suhaimi Napis, PhD

Faculty : Biotechnology and Biomolecular Sciences

Malaysia is the largest producer and exporter of palm oil with a 50% share of the global oil and fat production. Oil palm is the most productive oil crop with an average yield of about 3 to 4 t/ha/year. The productivity of oil palm is at least 3 to 8 times more compared to other oil-producing crops. Over the past few years, it has become clear that the possibilities for further expansion are now greatly reduced due to the low availability of suitable land. In addition, the increasing competitiveness of other vegetable oil crops, scarcity and cost of labour are some of the reasons driving the requirement to increase the productivity of existing planted land. This study was carried out with the objectives to isolate and analyse the differentially expressed genes in high yielding palms by using suppression subtraction hybridization (SSH) and annealing control primer (ACP) and confirming the differentially expressed gene candidates in high and low yielding palms using reverse northern and northern analysis. The SSH was performed using total RNAs were isolated from spear leaves of high and low yielding oil palm of population OxG and P312. In total, 250

subtracted clones were sequenced and 74.4% of them have significant matches with scores higher than 100 while 21.6% sequences have nonsignificant matches with sequences in the GenBank database. The remaining 4% sequences have no matches to the database. Majority of the genes that were differentially expressed in high yielding palm were associated with primary metabolism (48 sequences) such as glycolysis, oxidative pentose phosphate pathway, amino acid metabolism and acyl lipid metabolism including glucose-6-phosphate dehydrogenase and sucrose synthase. Besides, there were sequences encoding for enzymes in protein synthesis and processing (27 sequences), cell wall (8 sequences), gene expression and RNA metabolism (4 sequences), signal transduction and post-translational regulation (2 sequences), miscellaneous (2 sequences), secondary metabolism and hormone metabolism (2 sequences) and finally, defense and cell rescue (1 sequence). Fifteen clones, eight and seven from population OxG and P312, respectively; were selected for reverse northern analysis. Among these clones, five from population OxG (MAY39, MAY65, MAY79, MAY237 and MAY238) and six from population P312 (MAY133, MAY134, MAY144, MAY148, MAY154 and MAY240) were expressed only in high yielding palms and used as homologous probes in northern blot analysis. Northern analysis, demonstrated equal expression in high and low yielding palm of cytosolic aldolase from population OxG, beta galactosidase and pyruvate dehydrogenase from population P312. Clones encoding glucose-6-phosphate dehydrogenase and sucrose synthase were shown to have higher expression in high yielding palm of both populations, OxG and P312. These genes are closely related to photosynthesis and can be found almost in all plant tissues. They may be potentially useful as molecular markers for the screening of high yielding planting materials.

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

PENGENALPASTIAN DAN ANALISIS DUA GEN EKSPRESI TERBEZA DALAM DAUN MUDA KELAPA SAWIT HASIL TINGGI DAN RENDAH

Oleh

ROSLINDA ABU SAPIAN

Februari, 2008

Pengerusi : Professor Madya Suhaimi Napis, PhD

Fakulti : Fakulti Bioteknologi and Sains Biomolekul

Malaysia adalah negara pengeluar dan pengeksport terbesar minyak sawit dengan menguasai 50% daripada pengeluaran minyak sawit dunia. Pokok sawit adalah tanaman minyak sayuran yang produktif dengan hasil purata 3 hingga 4 t/ha/tahun. Produktiviti pokok sawit adalah di antara 3 hingga 8 kali lebih lebih tinggi berbanding kebanyakan tanaman minyak sayuran lain. Kebelakangan ini, pertambahan tahunan keluasan kawasan tanaman sawit adalah jelas menurun disebabkan kurangnya kawasan yang sesuai. Persaingan yang kian meningkat daripada minyak sayuran lain, kekurangan tenaga kerja dan peningkatan kos buruh adalah antara faktor penyebab yang mendorong kepada keperluan untuk meningkatkan produktiviti daripada kawasan tanaman sedia ada. Kajian ini telah dijalankan dengan tujuan untuk memencil dan menganalisis gen-gen yang diekspres terbeza dalam pokok sawit hasil tinggi dengan menggunakan hibridisasi subtraksi penindasan (SSH) dan pencetus kawalan penyepuhan (ACP), dan menentukan calon gen-gen diekspres terbeza dalam pokok sawit hasil tinggi dan pokok sawit hasil

rendah munggunakan analisis northern berbalik dan northern. SSH telah dilakukan dengan menggunakn RNA keseluruhan yang dipencilkan daripada daun muda pokok hasil tinggi dan pokok hasil rendah daripada populasi OxG dan P312. Keseluruhannya, 250 klon subtraksi telah dijujuk. Berdasarkan analisis jujukan, 74.4% daripada jumlah jujukan mempunyai persamaan yang signifikan dengan skor melebihi 100 manakala 21.6% jujukan mempunyai persamaan yang tidak signifikan dengan jujukan dalam pengkalan data GenBank. Manakala, 4% selebihnya tidak mempunyai persamaan dalam pengkalan data. Kebanyakan gen yang mempunyai perbezaan ekspresi dalam pokok hasil tinggi berkait dengan metabolisme primer (48 jujukan) seperti glikolisis, tapakjalan oksidatif pentosa fosfat, metabolisme asid amino dan metabolisme asil lipid termasuk glukose-6-fosfat dehidrogenase dan sukrose sintase. Selain itu terdapat jujukan yang mengkodkan enzim yang terlibat dalam sintesis dan pemprosesan protein (27 jujukan), dinding sel (8 jujukan), pengekspresan gen dan metabolisme RNA (4 jujukan), transduksi isyarat dan regulasi post-translasi (2 jujukan), lain-lain (2 jujukan), metabolisme sekunder dan hormon (2 jujukan) dan pertahanan dan penyelamatan sel (1 jujukan). Lima belas klon, lapan dan tujuh daripada popuasi OxG dan P312, masing-masing, dipilih untuk analisis reverse northern. Di antara klon-klon ini, lima klon daripada populasi OxG (MAY39, MAY65, MAY79, MAY237 and MAY238) dan enam klon daripada populasi P312 (MAY133, MAY134, MAY144, MAY148, MAY154 and MAY240) diekspres hanya pada pokok hasil tinggi dan digunakan sebagai prob dalam analisis northern. Analisis northern seterusnya menunjukkan ekspresi yang sama dalam pokok sawit hasil tinggi dan pokok sawit hasil rendah oleh klon yang mengkodkan sitosolik aldolase dari populasi OxG, beta galactosidase dan pyruvate decarboxylase dari populasi P312. Klon yang mengkodkan glukose-6-fosfate dehidrogenase dan sukrose sintase menunjukkan ekspresi yang lebih tinggi dalam pokok sawit hasil tinggi kedua-dua populasi, OxG dan P312. Gen-gen ini adalah berkait rapat dengan fotosintesis dan boleh dijumpai dalam hampir kesemua tisu tumbuhan. Gen-gen ini berkemungkinan mempunyai potensi sebagai penanda molekul untuk pemilihan tanaman hasil tinggi.

ACKNOWLEDGEMENTS

Firstly, I would like to thank Malaysian Palm Oil Board (MPOB), the founder of the project and entrusting me with the project under MPOB-University Graduate Research Collaboration Fund. I would like to thank Assoc. Prof. Dr. Suhaimi Napis (main supervisor), Dr. Muhamad Arif Abd Manaf (MPOB) and Dr. Ho Chai Ling (UPM) for being in my supervisory committee and providing necessary support.

Most of the work in this thesis was carried out in the Gene Expression Laboratory, MPOB and thank you to Anita, Syafiq and Faizun for their technical assistant. My appreciation to Syahanim, Safiza, Nurhafizah, Zetty, Zainab, Khairul, Mahathir, Razwan, Zaini and Aminah for the wonderful hospitality throughout the project.

I would like dedicate this thesis to my beloved Mak and Abah. You both are always in my heart. My brothers and sisters, thank you for your support and loves.

Last but not least, to a very special someone who always there for me, to keep on encouraging and supporting me. Thank you for not giving up on me.

Alhamdulillah.

I certify that an Examination Committee met on 13th November 2007 to conduct the final examination of Roslinda Abu Sapian on her Master thesis entitled "Identification and Analysis of Two Genes Differentially Expressed in Spear Leaves of High and Low Yielding Oil Palm" in accordance with Universiti Putra Malaysia (Higher Degree) Act 1980 and Universiti Putra Malaysia (Higher Degree) Regulation 1981. The Committee recommends that the student be awarded the degree of Master of Science.

Members of Examination Committee were as follows:

Mohd Puad Abdullah, PhD
Lecturer
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Chairman)

Janna Ong Abdullah, PhD
Lecturer
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Internal Examiner)

Datin Siti Nor Akamar Abdullah, PhD
Associate Professor
Faculty of Agriculture
Universiti Putra Malaysia
(Internal Examiner II)

Ismanizan Ismail, PhD
Associate Professor
Faculty of Science and Technology
Universiti Kebangsaan Malaysia
(External Examiner)

____

HASANAH MOHD GHAZALI, Ph.D.
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 21st February, 2008

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirements for the degree of Master. The members of the Supervisory Committee were as follows:

Suhaimi Napis, **PhD**
Associate Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Chairman)

Ho Chai Ling, **PhD**
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Member)

Muhamad Arif Abd Manaf, **PhD**
Senior Research Officer
Malaysian Palm Oil Board
(Member)

AINI IDERIS, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 21st February, 2008

I declare that the thesis is my original work except for equations and citation, which have been duly acknowledged. I also declare that it has not been previously and is not concurrently, submitted for any other degree at University Putra Malaysia or at any other institution.

ROSLINDA ABU SAPIAN

Date: 21st February, 2008

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

%	percentage
α	alpha
β	beta
λ	lambda
ºC	degree Celcius
µg	microgram
µl	microliter
µM	microMolar
ACP	Annealing Control Primer^{TM}
BLAST	Basic Local Alignment Search Tool
BI	bunch index
bp	base pair
BWT	bunch weight (kg)
$CaCl_2$	calcium chloride
cDNA	complementary DNA
CsCl	cesium chloride
D x P	*Dura x Pisifera*
dATP	2’-deoxy-adenosine-5’-triphosphate
dCTP	2’-deoxy-cystidine-5’-triphosphate
DEPC	diethyl pyrocarbonate
dGTP	2’-deoxy-guanosine-5’-triphosphate
dH_2O	deionized water
DNA	deoxyribonucleic acid
Dnase 1	Deoxyribonuclease 1

dNTP	deoxynucleotide triphosphates
E. coli	*Esherichia coli*
EDTA	ethylenediaminetetraacetatic acid
EtBr	ethidium bromide
FTB	fruit to bunch (%)
g	gram
G6PDH	glucose-6-phosphate dehydrogenase
HCl	hydrochloride acid
HT	height (m)
IPTG	isoprophyl-β-D-thiogalactosidase
IV	iodine value
Jacq.	*Jacquin*
Kb	kilobase
KCl	potassium chloride
KTB	kernel to bunch (%)
KTF	kernel to fruit (%)
KPY	kernel per year (kg/palm/year)
L	liter
LB	Luria Bertani
LiCl	lithium chloride
M	molar
mA	miliAmpere
MABW	means average bunch weight (kg/palm/year)
MBNO	means bunch number
MC	moisture content (%)

MFFB	means fresh fruit bunch (kg/palm/year)
MFW	means fruit weight (g)
$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulphate
min	minute
ml	milliliter
mm	millimeter
mM	miliMolar
MMLV	Maurine Moloney Leukemia Virus
mmole	miliMole
MNW	mean nut weight (g)
MPOB	Malaysian Palm Oil Board
mRNA	messenger RNA
MTF	mesocarp to fruit (%)
MW	molecular weight
NaCl	sodium chloride
NaOAc	sodium acetate
NaOH	natrium hydroxide
ng	nanogram
nm	nanometer
nM	nanomolar
N-terminal	amino terminal
OD	optical density
OPY	oil per year (kg/palm/year)
ORF	open reading frame

OTB	oil to bunch (%)
OTDM	oil to dry mesocarp (%)
OTF	oil to fruit (%)
PCR	polymerase chain reaction
Poly A^+RNA	polyadenylated RNA
PTB	parthenocarpic to bunch (%)
RE	restriction enzyme
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolution per minute
rRNA	ribosomal RNA
RT	reverse transcription
SDS	sodium dodecyl sulphate
sec	second
SSH	Suppression Subtractive Hybridization
SuSy	sucrose synthase
TAE	tris-acetate-EDTA
TEP	total economic product (kg/palm/year)
Tm	annealing temperature
TOT	total oil (kg/palm/year)
UPM	University Putra Malaysia
v/v	volume per volume
w/v	weight per volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactospyranose

CHAPTER 1

INTRODUCTION

Oil is the generic term for fluids that are not miscible with water. This word originated from the Latin word "*oleum*" which means olive oil. There are a few types of oil such as cooking oil (olive oil and vegetable oil), painting oil, crude oil, petroleum or mineral oil and essential oil. Vegetable oil or vegoil is oil extracted from oilseed or other plant sources. Some vegetable oil, such as rapeseed, cotton seed and castor oil, are not fit for human consumption without further processing. Like all fats, vegetable oil consists of esters of glycerin, a varying blend of fatty acids that are insoluble in water but soluble in organic solvents.

Palm oil is a form of edible vegetable oil produced by fruit of the oil palm tree. There are two species of oil palm but the better known one is originated from Guinea, Africa. It was first illustrated by Nicholas Jacquin in 1763, hence its name, *Elaeis guinnensis Jacq*. The African oil palm is a member of the *Arecacea*, or palm family. Palm oil is the second largest source of edible oil in the world after soybean which is produced in the tropical countries (Scowcroft, 1990). In 2001, palm oil accounted for 23.6 million tones or 20% of the world oil and fat production (Khoo and Chandramohan, 2002). It is the most price competitive liquid cooking oil in many parts of the world. It is also used in the making of other food products like shortening, margarines (Sudin *et al.*, 1993) and spreads (Pantzaris, 1993).

The emphasis of oil palm breeding is on the yield improvement of palm oil and palm kernel oil as both are the major economic products from the palm. According to Chan (1999), palm oil has increased its share to 17% in 1998 from 14% of the world production of oil and fat and it is expected to equal the share of soybean oil (21%) by 2010. In term of export, palm oil will continue to lead with its world market share growing from the present 33 percent to about 40 percent by 2020 (Chan, 1999). Oil palm is more productive than other oil bearing crops and Malaysian oil palm currently yields an average of 3.66 tonnes/ha of oil per year, which is 7 and 2.5 times more than soybean and rapeseed, respectively (MPOB, 2001). In Malaysia, the area planted with oil palm has increased from 2.03 million hectares in 1990 to 3.5 million hectares in 2001 (Teoh, 2002).

The major goal of plant breeding and biotechnology of oil palm is to maximize yield. The yield potential of the palm is normally determined by the genetics of the crop and the site characteristic such as sunshine, canopy structure, rainfall, nutrient and floral inflorescence dissection (Corley, 1983). Despite its recognized position as the pillar of the economy, the Malaysian palm oil industry is at the crossroads. The contribution of the palm oil to the national economy will be stagnate unless it can achieve further growth and remain competitive. However, area expansion is limited by the increasing scarcity of land and labour. Thus, breeding efforts worldwide focused on yield improvements from existing areas rather than area expansion which can also reduce the production cost (Khoo and Chandramohan, 2002).

In order to get higher yield of palm oil, the application of modern biotechnological methods such as molecular marker make it feasible to improve important plant varieties such as yield, disease resistance, stress tolerance, seed quality, etc without expansion of the planted areas. DNA marker has accelerated the conventional breeding by providing easy, fast and automated assistance to scientists and breeders. A marker linked to the shell thickness gene in oil palm had been found by Mayes *et al.* (1997) and this will allow selection of specific fruit types in the nursery, before fruiting starts and some preliminary results from oil palm have been published by Jack *et al.* (1998). In general, markers will allow selection for characters which are not expressed, such as disease resistance in an environment where the disease is not present and selection for mature characters in immature plants. The latter possibility, particularly, could significantly accelerate breeding progress.

The general objective in this study is to isolate the differentialy expressed gene(s) in high yielding palms which can be further characterized for molecular marker development. The genes will be further studied to be used as markers. These markers will be used for the selection of high yielding palms at an early stage in order to maximize the oil palm yield without area expansion. For that reason, spear leaves have been used in this study as starting materials, with the hope that the high yield palms can be identified and selected in the nursery instead of after the onset of bunch production.

Subtractive hybridization is a method for enriching differentially expressed genes which designed to selectively amplify differentially expressed transcripts while suppressing the amplification of abundant transcripts. Differentially expressed genes were isolated by hybridizing cDNA representing one sample (tester) to an excess of mRNA representing a second sample (driver). Transcripts expressed in both the tester and driver would form mRNA/cDNA hybrid molecule, whereas a cDNA sequence unique to the tester would remain single-stranded. Annealing control primer system is based on two principles: the unique tripartite structure of the primers, which have distinct 3′- and 5′-end regions that are separated by a polydeoxyinosine [poly(dI)] linker, and the interaction of each region during two-stage PCR. It has the high annealing specificity which generates reproducible, authentic and PCR products that are detectable on agarose gel. The specific objectives of this study are firstly, to isolate and analyse the differentially expressed genes in high yielding palms by using suppression subtractive hybridization and annealing control primer methods. Secondly, is to confirm the differentially expressed gene candidates in high and low yielding palms by using reverse northern and northern analysis.

CHAPTER 2

LITERATURE REVIEW

2.1 The Oil Palm

The oil palm is the most productive oil crop compared to soy bean, sunflower and rapeseed. The oil palm gives the highest yield of oil per unit area of any crop and produces two distinct oils, palm oil and palm kernel oil. The fruit is unarmed except for short spines on the leaf base and within the bunch (Hartley, 1988). This monoecious plant produces fruit on a large compact bunch. The fruit has fleshy pulp (mesocarp) which provides the oil and surrounds a nut; whose shell encloses the palm kernel (Corley and Tinker, 2004).

2.1.1 Origin and distribution

The oil palm (*Elaeis guineensis*) has spread from Senegal to Angola, along the coast and interior of the Congo River, from its home in the tropical rain forest region of west and central Africa (Hartley, 1977). The palm fruits were taken by early slave traders in the 16th century to the new world where it became established first in Bahia, Brazil. *E. guineensis* seedlings were grown in European conservatories in the 18th century, and in the following century, it was brought to Calcutta as ornamental plants (Corley and Tinker, 2004).

From its home in West Africa, the oil palm (*E. guineensis* Jacq.) has spread throughout the tropic and is now grown in 16 or more countries. Nowadays, palm oil industries have widely developed and its products are an important component in the world of vegetable oil production. A hectare of oil palm can produce 5 tonnes of crude palm oil. This is five times more than the yield of any commercially grown oil crop (Teoh, 2002).

2.1.2 Botany

Oil palm is classified in the genus *Elaeis*, a subfamily of *Cocoidea*. A palm family is normally known as *Palmae* or Arecaceae. Oil palm is an important member of the monocotyledonous group. *Elaeis* in Greek means elaion or oil, while the specific name *guineensis* shows the palm origin, which is the Guinea Coast of Africa (Hartley, 1988). The morphology of the oil palm fruits varies widely and the classifications are based on variations of the internal structure of the fruits. The examples of these classifications are Dura (shell 2 to 8 mm thick), Pisifera (shell less) and Tenera (shell 0.5 to 4 mm thick). Tenera is a hybrid of Dura and Pisifera, is often designated as DxP and most commercially planted. The fruit is reddish, about the size of a large plum and grows in large bunches. Each fruit contains a single seed (the palm kernel) surrounded by a soft oily pulp. Oil is extracted from both the pulp of the fruit (palm oil, edible oil) and the kernel (palm kernel oil, used mainly for soap manufacture).

The South American oil palm, *Elaeis oleifera,* is another species in the genus *Elaies*. Due to its low oil yield, it is of little economic importance except for use in interspecific hybridization programs, for the purpose of introgression of interesting characters such as resistance to bud root disease, low vertical growth and oil fluidity (Le Guen *et al*., 1991).