UNIVERSITI PUTRA MALAYSIA

ISOLATION AND CHARACTERIZATION OF UPREGULATED FLORAL TRANSCRIPTS FROM MANGOSTEEN (*Garcinia mangostana* L.)

CHAN KAM LOCK

FBSB 2008 8

ISOLATION AND CHARACTERIZATION OF UPREGULATED FLORAL TRANSCRIPTS FROM MANGOSTEEN (*Garcinia mangostana* L.)

By

CHAN KAM LOCK

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

July 2008

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

ISOLATION AND CHARACTERIZATION OF UPREGULATED FLORAL TRANSCRIPTS FROM MANGOSTEEN (*Garcinia mangostana L.*)

By

CHAN KAM LOCK

July 2008

Chairman : Ho Chai Ling, PhD

Faculty : Biotechnology and Biomolecular Sciences

Mangosteen (*Garcinia mangostana L.*) is one of the slowest-growing and longest living tropical fruit trees. Besides long juvenile period, lack of profuse flowering and irregular fruiting during early maturing stage are some of the major problems associated with growing mangosteen as an export fruit or for fruit products. The initiation of flowering process, development and maturation of flower in mangosteen are largely unknown. The understanding of these processes is important to solve some of the problems associated with growing mangosteen as one of the major fruits. Thus, the objectives of this study were to isolate, identify and sequence the mangosteen transcripts that were upregulated in the floral tissues, and study the gene expression and gene copy number of the selected upregulated floral transcripts. In this study, NSTEP method was found to be the best total RNA isolation method for mangosteen tissues. A subtracted cDNA library was constructed to facilitate the isolation of upregulated transcripts from mangosteen flower. Reverse northern screening and sequence analysis revealed that 28.5 % (149/522) of

these transcripts were upregulated in mangosteen flower. Among these transcripts, 82 of them were assembled into 30 contigs whereas 67 were singletons. A total of 63.9 % of these unigenes had non-signifancant matches to sequences in the non-redundant protein database in GenBank, 19.6 % had significant matches to unknown proteins and the remaining 16.5 % had putative functions that were further classified into six categories according to their biological functions. A total of three transcripts were selected for further characterization by real time reverse-transcription polymerase chain reaction and southern hybridization analysis. They were GmAGmbp (protein with GATA-type zinc finger domain), GmHsa32 (phosphosulfolatate synthase related protein) and GmbZIP (bZIP transcription factor). The 3' untranslated region (UTR) of these three transcripts were isolated from a cDNA library constructed using flower of 0.5-1.0 cm. All of these transcripts were verified to be expressed predominantly in the mangosteen flower tissue. GmAGmbp and GmHsa32 were found to be single copy genes in the mangosteen genome. The subtracted cDNAs isolated in this study might be used as expression markers for crop improvement in the future. However, further characterization of expression patterns and functional analyses are required to gather more valuable information on how these transcripts function during the flowering process.

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

PEMENCILAN DAN PENCIRIAN TRANSKRIPT YANG DINAIK-ATURKAN DI BUNGA DARIPADA MANGGIS (*Garcinia mangostana L.*)

Oleh

CHAN KAM LOCK

July 2008

Pengerusi : Ho Chai Ling, PhD

Fakulti : Bioteknologi and Sains Biomolekular

Manggis (*G. mangostana L.*) adalah salah satu buah-buahan tropika yang mempunyai pertumbuhan yang sangat lambat dan hayat hidup yang lama. Selain daripada tempoh juvenil yang lama, antara masalah-masalah yang berkaitan dengan penanaman manggis untuk diekspot atau untuk penghasilan produk buah-buahan adalah pembungaan yang kurang dan ia jarang berbuah pada peringkat awal kematangan. Proses permulaan pembungaan, perkembangan dan kematangan bunga manggis adalah kurang diketahui. Pemahaman tentang proses-proses tersebut adalah penting untuk mengatasi masalah-masalah yang terlibat dalam penanaman manggis sebagai salah satu buah-buahan yang penting. Oleh itu, objektif-objektif untuk kajian ini ialah untuk memencil, mengenalpasti dan menjujuk transkript-transkript manggis yang dinaik-aturkan dalam tisu-tisu bunga, dan mengkaji penzahiran gen dan bilangan salinan gen transkript-transkript bunga dinaik-aturkan yang terpilih. Dalam kajian ini, kaedah NSTEP merupakan kaedah pemencilan keseluruhan RNA yang paling sesuai untuk tisu-tisu daripada manggis. Satu perpustakaan

cDNA tertolak (subtracted cDNA library) telah dibina untuk memudahkan pemencilan transkript-transkript yang dinaik-aturkan daripada bunga manggis. Penyaringan northern 'berbalik' dan analisa jujukan mendapati 28.5 % (149/522) dari transkript-transkript tersebut dinaik-aturkan dalam bunga manggis. Di antara transkript-transkript ini, 82 daripada mereka terkumpul dalam 30 'contigs' manakala 67 adalah 'singletons'. Sejumlah 63.9 % unigen-unigen ini mempunyai padanan yang tidak sahih dengan jujukan-jujukan dalam pangkalan protein tidak bertindan di 'GenBank', 19.6 % mempunyai padanan yang sahih dengan protein yang tidak diketahui dan 16.5 % yang lain yang mempunyai fungsi ramalan telah dikelaskan selanjutnya kepada enam kategori mengikut fungsi biologikal mereka. Tiga daripada transkript-transkript ini telah dipilih untuk pencirian selanjutnya dengan menggunakan 'real-time reverse transcription polymerase chain reaction' dan analisa penghibridan 'southern'. Mereka ialah GmAGmbp (protein dengan domain jejari zink jenis GATA), GmHsa32 (protein yang berkaitan dengan phosphosulfolactate synthase) dan GmbZIP (faktor transkripsi bZIP). 3' 'untranslated region' (UTR) untuk tiga transkript ini telah dipencilkan daripada perpustakaan cDNA yang dibina dengan menggunakan bunga bersaiz 0.5-1.0 sm. Semua transkript ini telah disahkan adalah predominan di dalam tisu bunga manggis. GmAGmbp dan GmHsa32 didapati mungkin adalah gen-gen yang mempunyai salinan tunggal dalam genom manggis. Semua cDNA tertolak yang dipencilkan dalam kajian ini mungkin boleh digunakan sebagai penanda ekspresi untuk memajukan tanaman pada masa akan datang. Walaubagaimanapun, pencirian corak penzahiran dan analisa fungsi selanjutnya diperlukan untuk mengumpul maklumat yang lebih bermakna tentang bagaimana transkript-transkript tersebut berfungsi dalam proses pembungaan.

ACKNOWLEDGEMENTS

Firstly, I would like to thank Dr. Ho Chai Ling for giving me the opportunity to pursue my studies in the field of plant molecular biology and for her guidance and advice throughout this project. Besides that, I would like to express my appreciation to Dr. Suhaimi Napis and Dr. Parameswari for being my supervisory committee members and for their guidance and advice. I would also like to thank the Ministry of Science and Technology for funding this project. My sincere thanks to Mr. Adrian (MARDI) for his kindness and funding support.

I would like to acknowledge Faculty of Agriculture and Faculty of Design and Architecture for allowing me to collect samples for this project in 'Ladang 10' and 'Nursery Bahan Bakar Buah-buahan', respectively.

My deepest appreciation also goes to all the members of Genetic Lab (UPM) and special thanks are extended to Choong, Ky, Thuc, Sock Hwa and Siti. I would not have completed my project without their help and guidance.

Lastly, my deepest gratitude to my parents for their endless love and support.

I certify that an Examination Committee met on 11th July 2008 to conduct the final examination of Chan Kam Lock on his Master of Science thesis entitled "Isolation and Characterization of Upregulated Floral Transcripts From Mangosteen (*Garcinia mangostana* L.)" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree.

Members of the Examination Committee are as follows:

Norihan Mohd. Saleh, PhD
Associate Professor
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)

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

Rofina Yasmin Othman, PhD
Professor
Faculty of Science
University of Malaya
Malaysia
(Enternal Examiner)

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

Date:

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

Ho Chai Ling, PhD
Associate Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Chairman)

Suhaimi b. Napis, PhD
Associate Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Member)

Parameswari a/p Namasivayam, PhD
Lecturer
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Member)

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

Date: 11 September 2008

DECLARATION

I hereby declare that the thesis is based on my original work except for equations and citations, which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UPM or other institutions.

CHAN KAM LOCK

Date: 31 July 2008

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFLP	-	Amplified Fragment Length Polymorphism
AIMS	-	Amplification of Insertion Mutagenised Sites
AMV	-	Avian Myeloblastosis Virus
AP1	-	*APETALA 1*
BAS	-	Bureau of Agricultural Statistic
BLAST	-	Basic Local Alignment Search Tool
bp	-	base pair
bZIP	-	basic leucine zipper
CaMV 35S	-	Cauliflower Mosaic Virus 35S
CCA1	-	*CIRCADIAN CLOCK ASSOCIATED1*
CDS	-	coding region
CI	-	chloroform: isoamyl alcohol
CO	-	*CONSTANS*
CoM	-	Coenzyme M
ComA	-	(2R)-phospho-3- sulfolactate synthase
COX 2	-	Cyclooxygenase 2
CRY	-	Cryptochromes
CsCl	-	caesium chloride
C_T	-	threshold cycle
CTAB	-	hexadecyl (or cetyl) trimethyl ammonium bromide
Cyp	-	*Cyclophilin*

DDRT-PCR	-	Differential Display Reverse Transcription Polymerase Chain Reaction
DEPC	-	diethylpyrocarbonate
DMSO	-	dimethyl sulphoxide
Dnase I	-	deoxyribonuclease I
EDTA	-	ethylene diamine tetracetate
ELF3	-	*EARLY FLOWERING3*
EST	-	Expressed Sequence Tag
EtBr	-	ethidium bromide
FLC	-	*FLOWERING LOCUS C*
FLO	-	*FLORICAULA*
FMs	-	floral meristems
FPF1	-	*FLOWERING PROMOTIVE FACTOR 1*
FRI	-	*FRIGIDA*
FT	-	*FLOWERING LOCUS T*
FWA	-	a late flowering gene
GA	-	gibberellin
GI	-	*GIGANTEA*
GmAGmbp	-	*Garcinia mangostana* AG motif binding protein
GmbZIP	-	*Garcinia mangostana* basic leucine zipper
GmCyP	-	*Garcinia mangostana* cyclophilin
GmHsa32	-	*Garcinia mangostana* heat stress associated 32
Gmαtubulin	-	*Garcinia mangostana* α-tubulin
GTC	-	guanidium thiocyanate

HIV	-	Human Immunodeficiency Virus
HS	-	heat shock
Hsa32	-	heat stress associated 32
HSP	-	heat shock protein
IBPGR	-	International Board of Plant Genetic Resources
IMs	-	inflorescence meristems
IPGRI	-	International Plant Genetic Resources Institute
LB	-	Luria-Bertani
LD	-	Long-day
LFY	-	*LEAFY*
LHY	-	*LATE ELONGATED HYPOCOTYL*
LiCl	-	lithium chloride
MADS	-	MCM1-AGAMOUS-DEFICIENS-SRF
$MgSO_4$	-	magnesium sulphate
MOPS	-	3-(N-morpholino)propanesulfonic acid
MRSA	-	Methicillin-Resistant *Staphylococcus aureus*
NaCl	-	sodium chloride
NaOAc	-	sodium acetate
NaOH	-	sodium hydroxide
NCBI	-	National Center for Biotechnology Information
NH_4OAc	-	ammonium acetate
NLS	-	nuclear localization signal
OD	-	optical density

ORFs	-	open reading frames
PBZ	-	paclobutrazol
PCI	-	phenol: chloroform: isoamyl alcohol
PCR	-	polymerase chain reaction
Pfr	-	Phytochromes of far red light-absorbing form
Pfu	-	plaque forming unit
PGE2	-	Prostaglandin E_2
PHY	-	Phytochromes
pI	-	isoelectric point
Poly (A)	-	polyadenylated (mRNA)
ppm	-	parts per million
PPO	-	polyphenol oxidase
Pr	-	Phytochromes of red light-absorbing form
PSL	-	phosphosulfolactate synthase-related protein
PVP	-	polyvinylpyrrolidone
R^2	-	correlation coefficient
RAPDs	-	randomly amplified polymorphic DNA markers
RFLP	-	restriction fragment length polymorphism
RT-PCR	-	Reverse transcription – PCR
SAGE	-	serial analysis of gene expression
SAM	-	shoot apical meristem
SD	-	Short-day
SDS	-	sodium dodecyl sulfate

SOC1	-	*SUPPRESSOR OF OVEREXPRESSION OF CO1*
spy	-	*spindly*
SQDG	-	sulfoquinovosyl diacylglycerol
SQUA	-	*SQUAMOSA*
SSC	-	standard saline citrate
SSH	-	suppression subtractive hybridization
TAE	-	Tris-acetate-EDTA
T-DNA	-	Transferred-DNA
TE	-	Tris-EDTA
T_m	-	melting temperature
TOC1	-	*TIME OF CHLOROPHYLL A/B BINDING PROTEIN1*
Tris	-	tris[hydroxymethyl]aminomethane
Tris-HCl	-	tris-hydrochloride
U	-	unit
USD	-	U.S. Dollars
UTR	-	untranslated region
UV	-	ultraviolet
V	-	volt
VRN2	-	*VERNALISATION2*
YAC	-	yeast artificial chromosome
ZIM	-	Zinc-finger protein expressed in Inflorescence Meristem

CHAPTER 1

INTRODUCTION

The mangosteen (*Garcinia mangostana* L.) is thought to have originated from Peninsular Malaysia and early cultivation of this crop was limited to Southeast Asia. The mangosteen spread to other tropical regions during the past few centuries. Mangosteen trees grow naturally as understorey plants in forest communities and are usually propagated by apomictic seeds. Mangosteen is one of the slowest-growing and longest living tropical fruit tree. It has been considered as the most delicious fruit of the tropics and has been named the 'queen of fruits'. In Southeast Asia, the fruit pericarp has been used traditionally as medicine for inflammation, diarrhoea, dysentery, wounds and skin infections. The mangosteen pulp contains high amounts of energy, vitamins and minerals, hence it can greatly improve food quality of low-income rural households especially children. Aside from being a source of fresh and processed food, the fruit rind contains 7-14 % catechin tannin and is used for tanning of leather and it also produces a natural black dye (http://www.civil.soton.ac.uk/icuc/factsheets.html).

The demand for mangosteen fruits usually exceeds the supply as mangosteen trees are rarely planted in commercial quantities. However, in recent years, the mangosteen has been subjected to renewed interest and it has gained increased recognition in the international markets. Thailand and Malaysia are the major commercial producers and suppliers of mangosteen to United Kingdom, Hong Kong, Singapore, Taiwan and Japan.

In 1990, the export quantity of mangosteen in Malaysia was 1, 544 tons valued at USD 456, 000, and about a decade later, the export quantity has increased to 1, 961 tons valued at USD 1, 127, 000 (Mohamad and Abd Rahman, 2006). In 2002, the total cultivation area for mangosteen in Thailand was 48, 000 hectares which yielded 160, 000-190, 000 tons fruits. The export value of mangosteen was USD 10 millions with overseas sales growing at an average rate of 102 % (Office of Agricultural Economics, 2003). It is expected that Thailand and Malaysia will maintain to be the major suppliers of mangosteen in the world market as both countries are still expanding their mangosteen production areas.

Nevertheless, mangosteen is still not cultivated on large scale despite tremendous consumer acceptance, good transport infrastructure and long shelf life. It is because of its long pre-bearing period resulting from the extremely slow growth of the developing seedlings, unusually long juvenile phase, low fruit yield, biennial bearing and short viability of seeds. Much efforts are needed to solve these problems in order to expand the mangosteen fruit industry to be one of the major ones in Malaysia. Research and development activities must be carried out intensively in order to solve the problems and to fulfill the increasing demand. Flowering is a fundamental process in plant development that leads to fruit formation. The molecular mechanisms underlying flowering, development and maturation of flower in mangosteen are poorly understood. Therefore, it is of paramount importance to study the flowering process in mangosteen to solve problems such as its long juvenile phase by genetic controlling of its flowering time.

The objectives of this study are:

1. To isolate, identify and sequence the mangosteen transcripts that are upregulated in the floral tissues with the aim to further understand the flowering process in this fruit tree.
2. To study the gene expression of the selected upregulated floral transcripts in floral bud and young shoot of mangosteen.
3. To determine the gene copy number of the selected upregulated floral transcripts in mangosteen.