# PHYSIOLOGICAL, BIOCHEMICAL AND MOLECULAR ANALYSES OF FRUIT DEVELOPMENT IN OIL PALM (*ELAEIS GUINEENSIS* JACQ.)

**By**

**Nurniwalis binti Abdul Wahab BSc., MSc.**

**Thesis submitted to The University of Nottingham Malaysia Campus, in fulfilment of the requirement for the degree of Doctor of Philosophy**

**2017**

## ABSTRACT

In Malaysia, oil palm is the main commodity crop and has strongly contributed to the country's economic development. However, the industry is facing several challenges including diminishing land resource for expansion. One of the targets to ensure oil palm sustainability is to increase yield per unit area without affecting its quality and with minimal impact to the environment. The non-synchronized ripening process and shedding of the fruits is among the factors that limit yield improvement and affect oil quality. Thus, the present study was conducted to deepen our understanding of the changes that occur within the fresh fruit bunch (FFB) during fruit development. This work also takes into account the various fruit positions (inner, middle and outer) and locations (apical, central and basal) within the bunch.

In the present study, physiological, biochemical and molecular analyses were performed using fruits from various stages of fruit development. Physiological and biochemical analyses included measurements of carotenoid, FAC (fatty acid composition) and ethylene production. UV-Vis spectrophotometer, gas chromatography mass spectrophotometry (GCMS) and GC equipped with flame ionization detector (FID) were used to estimate and quantify the β-carotene content, FAC and ethylene production from the fruit mesocarp from young until the ripening stages. Results showed that changes in the β-carotene content, FAC and ethylene production throughout the fruit developmental stages from young until ripening were significant ($p<0.001$). Both accumulation of the β-carotene and ethylene production increased as the

fruits developed and reached the maximum at the ripest stage used in this study. Similarly, changes in the FAC was observed especially in palmitic, oleic, stearic and linoleic acid throughout the fruit developmental process where the level of oleic acid was found to be the highest in the ripest fruit stage used in this study, surpassing that of palmitic acid.

In the present study, it was also observed that fruit positioned within the spikelets throughout the bunch at the various fruit developmental stage from young until ripening showed significant changes ($p<0.001$) in the β-carotene content and FAC. For β-carotene, the outer fruits have the highest content while the inner fruits have the least. For FAC, palmitic and stearic acids were higher in the outer fruits as compared to the inner fruits while the levels of oleic and linoleic acids were in contrast to that of palmitic and stearic. The inner fruits also recorded higher iodine value (IV) than the outer fruits.

Ethylene production in the ripe bunch was not influenced by the fruit location or position within the bunch. Thus, the molecular mechanism of ethylene perception during fruit developmental processes, including ripening and abscission were investigated. In the present study, with the use of the oil palm genome sequence data, the oil palm ethylene receptor family which comprised of seven putative ethylene receptors including three splice variants were identified. In addition, combination of *in silico* bioinformatics tools and laboratory bench work also resulted in the successful isolation and characterisation of the putative ethylene receptor genes and their corresponding promoters. The putative regulatory motifs in the promoters

provide a means to understand the possible mechanisms that drive the expression of the ethylene receptor genes. Expression analyses of the ethylene receptors in various oil palm tissues suggest multiple roles of the ethylene receptors in regulating many processes in oil palm which includes fruit development and fruit abscission.

Subsequently, the effect of gene(s) that influence oil quality was also investigated in this study. The full-length *FLL1* gene encoding a lipase class 3 and its corresponding promoter were successfully isolated and characterised. *FLL1* was expressed highly in the mesocarp tissues and at various developmental stages and at a much higher expression in the cold induced ripe fruits. Promoter analysis *via* transient expression assay using GUS as the reporter gene showed that the GUS expression in the mesocarp slices was targeted especially to the vascular bundles. Southern analysis of the *FLL1* gene revealed the gene to belong to a multigene family. With the availability of the oil palm genome sequence data, sixty-two predicted proteins with identities ranging from 26 to 83% to *FLL1* were identified indicating that there may be more lipase class 3 genes involved in the breakdown of TAGs in the oil palm thus affecting the oil quality.

From this study, it can be concluded that fruit development in oil palm is a complex process involving physiological, biochemical and molecular changes within the bunch throughout the ripening process and is affected by the position of the fruits within the spikelets. Since changes within the bunch were not influenced by the fruit location within the bunch, this may provide a

simpler method for bunch analysis where changes within the whole bunch can just be represented by the fruits from a mix of all fruits positioned within the spikelet from the apical region of the bunch.

The molecular studies on the identification, isolation and characterisation of the genes and promoter of the ethylene receptors and lipase class 3 gene families also would be useful for genetic manipulation of oil palm especially for modifying oil composition and production of higher value products.

## ACKNOWLEDGEMENT

Syukur Alhamdulillah, with the will of the Almighty Allah s.w.t, this thesis is finally completed. I would like to take this opportunity to firstly, thank my supervisors to whom I am greatly indebted. To Festo and Jerry, you guys are the best. Thank you Festo for being you and persevering with me as I take my time in writing the thesis. Your utmost support, advice and encouragement throughout the duration of the study especially when I needed it the most is very much appreciated. Jerry, despite the long distance, you have been a great help especially with your 'story telling' ability, constructive suggestions and ideas.

I would also like to take this opportunity to express my deepest gratitude and appreciation to Dr Jue, Kak Aminah, Mahadzir, Fadilla, Chan and Biha for their help, support and technical know-how given in so many ways. Without it, the journey would have been a lot rockier.

To my friends, colleagues, students and lecturers, I am grateful for your kind advice, help, assistance and support on my research and writing efforts throughout the years. There are so many of you, but I hope you know who you are. To the Gene Function Group members, thank you for lending lots of helping hands in expediting my lab works especially during the bunch processing. To the Breeding and Tissue Culture Team and Bioinformatians, thank you for the fruit bunches and sharing your working experiences and knowledge that has helped support and expand my own work tremendously.

To my parents, thank you for the constant dua', love and support. To my husband and family, your emotional support, patience and understanding has kept me on my toes and to push myself forward even when the going gets tough. Without it, I may not get see the light at the end of the tunnel. So thank you from the bottom of my heart.

Last but not least, to MPOB, especially the Director General, Deputy Director General (Service), Director of ABBC, Pn. Rubaah and Pn Suraya from the Training Unit, thank you for the scholarship and encouragement to improve and better myself. The journey has been bittersweet and the learning experience has taught me an invaluable lesson that the key to achieve happiness, satisfaction and success in life is to have a strong balance between family, work and friends.

## SUPERVISORY COMMITTEE

The members of the supervisory committee are as follows:

**Professor Festo Massawe, PhD**

Professor and Head of School

School of Biosciences

Faculty of Science

The University of Nottingham Malaysia Campus

(Principal Supervisor)

**Professor Jeremy Roberts, PhD**

Professor of Plant Biology

School of Biosciences

Faculty of Science

The University of Nottingham Sutton Bonington Campus

(Co-supervisor)

## TABLE OF CONTENT

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| aa | Amino acid |
| ABA | Abscisic acid |
| ABBC | Advanced Biotechnology and Breeding Centre |
| ACC | 1-aminocyclopropane-1-carboxylate |
| ACO | ACC oxidase |
| Acc. No. | Accession number |
| ACS | ACC synthase |
| ANOVA | Analysis of variance |
| Asp | Aspartate |
| ATG | Start codon |
| AZ | Abscission zone |
| BLAST | Basic Local Alignment Search Tool |
| bp | Base pair |
| CARE | *cis*-acting regulatory elements |
| cDNA | Complementary DNA |
| cm | Centimetre |
| CPO | Crude palm oil |
| CTAB | Cetylatrimethyl ammonium bromide |
| CTR | Constitutive triple response |
| D | Aspartic acid residue |
| D x P | *dura x pisifera* |
| DAA | Days after anthesis |
| DAP | Days after pollination |
| dATP | Deoxyadenosine triphosphate |
| dCTP | Deoxycytidine triphosphate |
| dGTP | Deoxyguanosine triphosphate |
| DIECA | Diethyldithiocarbonic acid |
| DNA | Deoxyribonucleic acid |
| dNTPs | Deoxyribonucleoside riphosphates |
| dTTP | Deoxythymidine triphosphate |
| *E. coli* | *Escherichia coli* |
| EDTA | Ethylenediaminetetraacetic acid |
| EIL | EIN3-like proteins |
| EIN | Ethylene-insensitive |
| ERF | Ethylene response factor |
| ERS | Ethylene response sensor |
| ER | Ethylene receptor |

| | |
|---|---|
| et al. | Et alia (and other, Latin) |
| ETR | Ethylene-resistant |
| EtBr | Ethidium bromide |
| EST | Expressed sequence tags |
| FAC | Fatty acid composition |
| FAME | Fatty acid methyl esters |
| FFA | Free fatty acid |
| FFB | Fresh fruit bunch |
| FID | Flame ionization detector |
| FL | Full-length |
| FW | Fresh weight |
| g | gram |
| GA | Gibberellin |
| GC | Gas Chromatography |
| GC-MS | Gas Chromatography Mass Spectrometry |
| GO | Gene ontology |
| GFP mix | Glyoxal/Formamide/Sodium Phosphate Mix |
| GUS | β-glucuronidase |
| h | Hour |
| H | Histidine residue |
| ha | Hectare |
| HCl | Hydrocloric acid |
| HMM | Hidden Markov Model |
| IAA | Indole-3-acetic acid |
| ID | Identity |
| IV | Iodine value |
| Jacq. | Jacquin |
| K | Lysine residue |
| KAS II | β-ketoacyl ACP synthase II |
| Kb | Kilobase |
| kDa | kilodalton |
| $KHCO_3$ | Potassium bicarbonate |
| L | Litre |
| LB | Luria and Bertani |
| LD-PCR | Long Distance Polymerase Chain Reaction |
| LiCl | Lithium chloride |
| LRE | Light responsive element |
| LTRE | Low temperature responsive element |
| *lycb* | Lycopene β-cyclase |
| *lyce* | Lycopene ε-cyclase |

| | |
|---|---|
| M | Molar |
| mg | Milligram |
| $MgCl_2$ | Magnesium chloride |
| min | Minute |
| ml | Millilitre |
| mm | Millimetre |
| mM | Milimolar |
| mmol | Milimole |
| MOPS | 3-(N-morpholino)propane-sulphonic acid |
| MPOB | Malaysian Palm Oil Board |
| mRNA | Messenger RNA |
| MUFA | Monounsaturated fatty acid |
| NaCl | Natrium chloride |
| NaOH | Natrium hydroxide |
| ng | Nanogram |
| NIST | The National Institute of Standards and Technology |
| nr | Non-redundant |
| nm | Nanometre |
| *Nr* | Never ripe |
| OD | Optical density |
| OER | Oil extraction rate |
| ORF | Open reading frame |
| PCR | Polymerase Chain Reaction |
| PG | Polygalacturonase |
| pers. comm. | Personal communication |
| PKC | Palm kernel cake |
| Poly A+ RNA | Polyadenylated RNA |
| PI | Isoelectric point |
| ppm | Parts per million |
| PUFA | Polyunsaturated fatty acid |
| R | Arginine residue |
| RACE | Rapid Amplification of cDNA Ends |
| REG | Regulatory element group |
| RNA | Ribonucleic acid |
| RNAi | RNA inteference |
| rpm | Revolution per minute |
| RT-PCR | Reverse Transcription Polymerase Chain Reaction |
| S | Serine residue |
| SAM | S-adenosyl-l-methionine |
| SDS | Sodium dodeocyl sulphate |

| | |
|---|---|
| sec | Second |
| SFA | Saturated fatty acid |
| SSC | Sodium Chloride-Sodium Citrate |
| t | tonnes |
| TAE | Tris acetate EDTA |
| TAG | Triacylglyceride |
| TAIR | The Arabidopsis Information Resource |
| TE | Tris EDTA |
| TSS | Transcription start site |
| U | Unit |
| USA | United States of America |
| UTR | Untranslated region |
| UV | Ultra violet |
| V | Volt |
| *VIR* | *virescens* |
| v/v | Volume per volume |
| w/v | Weight per volume |
| WAA | Weeks after anthesis |
| X-gal | 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside |
| Y-patch | Pyrimidine patch |
| yr | year |
| µl | Microlitre |
| µg | Microgram |
| µM | Micro Molar |
| $\alpha$ | Alpha |
| β | Beta |
| $\lambda$ | Lambda |
| % | Percentage |
| $^{o}$C | Degree celsius |
| & | And |

# CHAPTER 1

# CONTEXT, OBJECTIVE AND OVERVIEW

## 1.1 The Oil Palm – An introduction

Oil palm is a tropical perennial monoecious monocotyledon species that belongs to the genus *Elaeis,* a subfamily of *Cocoideae,* from the palm family, the *Arecaceae* (Hartley 1988). The genus comprises of two species, *Elaeis guineensis* or commonly known as the African oil palm *and Elaeis oleifera* or the American oil palm (Hartley 1988). *Elaeis guineensis* Jacq. originated from West Africa whereas *Elaeis oleifera* is from Central and South America (Hartley 1988). Both species are unique and have prominent features that distinguish one from another. *E. guineensis* is the commercial planting material as it has bigger fruits than *E. oleifera* and thus produces higher oil yield (Hartley 1988). *E. oleifera* also has lower height increments that is one-fifth of *E. guineensis* (Corley and Tinker 2003). It has root structures that contain tannins and a more lignified parenchyma believed to act as protection to several types of pest and diseases (Arnaud and Rabechault 1972). The mesocarp also produces oil that contains higher oleic acid, iodine value and higher carotene (Mohd Din et al. 2000; Corley and Thinket 2003).

To date, oil palm is grown across the equator of tropical regions in Asia, Africa and South America. It was introduced into South-east Asia in 1848 where four seedlings were planted in the Botanical Gardens at Bogor, Java (Whitmore 1979). Since then, the four seedlings have become the foundation of all oil

palm planted in Indonesia and Peninsular Malaysia (Whitmore 1979). In Malaysia, oil palm was first planted in 1875 as an ornamental but was later developed as a commercial crop when it was planted in Tanamaran Estate in 1917 (Hartley 1988). Despite the slow growth in the 1950s, more oil palms were planted on a large scale in the 1960s to diversify the agricultural sector and to reduce the high dependency on rubber (Yusoff et al. 2000; FELDA 2014). This act has prompted the growth of the oil palm industry which has now become Malaysia's main commodity crop (EPU 2014; MPOB 2016). In 2015, the total area planted with oil palm reached 5.64 million hectares, an increase of approximately 4.6% from 2014 (MPOB 2016). The majority of the planted area is owned by private estates (61%) and government agencies (23%) while the rest belongs to independent smallholders (16%) (MPOB 2016).

Malaysia is now the second largest palm oil producer in the world behind Indonesia with 19.96 million tonnes of crude palm oil (CPO) with total export revenues of RM 60.2 billion (MPOB 2016). The great success of the oil palm industry in Malaysia is partly due to the country's favourable weather, rain and soil conditions and good agricultural policies envisaged by the country's leaders (Kushairi 2009).

## 1.2 The importance of oil palm

Oil palm, *E. guineensis* is the highest yielding oil crop in the world. Palm oil (31%) derived from oil palm is the most consumed vegetable oil followed by soybean (*Glycine max*) (22%) in the world of vegetable oils and fats (Oil World

2016). The productivity of oil palm in terms of oil yield in comparison to other world's major oil crops is shown in Table 1.1. The crop produces 3.61 tonne (t) of CPO and 0.35 t palm kernel oil per hectare (ha) per year, 5-10 times more than other major oil crop (Oil World 2016). The land requirement for planting oil palm is also very much lower with only 0.25 ha for production of 1 t oil in comparison to 1.5, 1.6 and 1.3 ha by soybean, sunflower and rapeseed (*Brasicca napus*), respectively (Table 1.1) (Oil World 2016). With the support from the latest research and development especially in the areas of crop improvement and crop management, the oil palm yield is predicted to reach a maximum yield potential of 10-12 t/h/yr (Breure 2003; Wahid et al. 2005; Murphy 2014) but this has yet to be attained on a commercial basis.

**Table 1.1: Productivity of various major oil crops in the world**

| Oil Crop | Production (million t) | Oilseeds yield (t/ha/yr) | Total area (million ha) |
|---|---|---|---|
| Oil palm | 68.56* | 3.96* | 17.32 |
| Soybean | 48.23 | 0.40 | 120.15 |
| Rapeseed | 26.66 | 0.75 | 35.40 |
| Sunflower | 15.18 | 0.62 | 24.61 |

Source: Oil World 2016

* Combined tonnage of palm oil and palm kernel oil

The cost for palm oil production is also the lowest in comparison to other oil crops such as soybean, sunflower and rapeseed and with the high share of vegetable oil exports, the demand for palm oil is high (Abdullah and Wahid

2011). The global demand for palm oil will remain high and was predicted to increase between 93 million t and 156 million t by 2050 (Corley 2009).

The oils (palm oil and palm kernel oil) derived from the oil palm fruits are also versatile in both food (90%) and non-food (10%) sectors (Sambanthamurthi et al. 2000). Palm oil is mainly used for food while palm kernel oil is mainly used in the oleochemical industry. In the food industry, palm oil is mainly used as: cooking oils, Vanaspati (purified hydrogenated vegetable oil), shortenings and margarines (Miskandar et al. 2011). Through research and development new palm oil products are being produced such as premium vegetable oils like high SmartBalance™ and Novelin™, palm-based processed cheese, non-dairy products and as raw materials for confectionery products (Miskandar et al. 2011; Siew 2011). In the non-food sector, palm oil is mainly used as the raw material for the oleochemical industry for making candles, detergents, lubricants and for the cosmetic industries for making cosmetics and personal care products (Hazimah et al. 2011). With the demand for biofuels as an alternative source of energy, Malaysia has also developed palm oil as a raw material for biofuel production (Chin 2011). The use of B7-biodiesel, a blend of 7% palm oil (palm methyl esters) and diesel has been implemented throughout the country while the next 10% palm oil blend is expected to be launched in 2016 (Wahab 2015).

Besides the oil, many oil palm value-added products have been developed to maximize the potential of the oil palm. The oil palm trunk for example, provides significant biomass that can be transformed into many

value-added products. This include products such as laminated board, plywood, fibreboard, compressed wood and furniture which have gained much attention from consumers (Sulaiman et al. 2008; 2012). In comparison to wood, the use of the abundant oil palm biomass as raw material for production of value added products will not only reduce production costs but will also increase the economic returns (Sulaiman et al. 2012). Under Malaysia's National Biomass Strategy (NBS) 2020, the use of oil palm biomass as a raw material for the production of value added products such as biofuels and biobased chemicals is targeted to increase the country's economy income up to RM30 billion per year and could create 65,000 new jobs by 2020 (Economic Tansformation Programme 2013).

Another example is the palm kernel cake (PKC), a by-product from milling palm kernel oil (MPOB 2016). The PKC normally contains 14.6 to 16.0% crude protein and is used as the main ingredients in animal feed (Zahari and Alimon 2004; Alimon 2004). Malaysia exports 2.6 million tonnes of PKC generating RM 936 million in 2015 (MPOB 2016). With so many benefits, it is not surprising that oil palm production has increased tremendously over the years, from being a relatively unknown crop almost a century ago to now becoming a crop that produces the most consumed and traded oil commodity in the world (MPOB 2016; Oil World 2016).

Despite the many benefits the palm has to offer, land expansion for cultivating oil palm is becoming scarce due to the unavailability of suitable land (MPOB 2016). Environmental issues such as deforestation, biodiversity loss

and greenhouse gas emissions (GHG) in relation to oil palm cultivation is also becoming a concern  despite Malaysia's firm stand on keeping its rainforest and maintaining its status as a mega biodiversity country (Fitzherbert et al. 2008; http://bigcatrescue.org/malaysia-one-of-12-mega-biodiversity-countries-in-the-world/). Consequently, these issues have become one of the challenges faced by the oil palm industry. The oil palm sustainability thus needs to be maintain possibly through diversification without being dependent on land expansion to increase production. The advancement in the area of biotechnology, especially in genomics and genetic engineering, have provided opportunities to expedite crop improvement (Singh et al. 2013a; Parveez et al. 2015).

## 1.3 Fruits and the ripening process

Fruits are a good source of minerals, vitamins, fibers and antioxidants for human health. Soft fleshy fruits such as tomato (*Solanum lycopersicum)* and peaches (*Prunus persica)* have received much more attention due to their importance in human diet in comparison to oil based fruits such as oil palm (Osorio et al. 2013). The economic importance of fruits has prompted many researchers to investigate various aspects of fruit development including fruit organogenesis, differentiation, development, maturation and ripening. Studies have taken place at the molecular, genetic, biochemical and physiological levels, and have focused on hormone interactions as well as nutritional composition and enhancement (Giovannoni 2001; Adams-Phillips et al. 2004).

Tomato for example, is a model plant that has been widely use to understand fleshy fruit organ development, maturation, ripening, shelf-life, and nutritional quality (Giovannoni, 2001). The availability of advanced genetic and molecular toolkits, short generation time, extensive germplasm collections, well characterised mutant stocks, high-density genetic maps, immortalized mapping populations, efficient transient and stable transformation, expressed sequence tag (EST) resources and microarray data helps provide the necessary information (Mounet et al. 2009).

Fruit ripening is one of the areas that is widely studied due to its importance in fruit harvest and storage. The ripening process especially in fleshy fruits involves a wide range of biochemical, physiological and molecular changes that affect the appearance, color, texture, flavor and aroma of the fruit and as a result, the ripe fruit becomes more attractive, desirable and palatable for consumption (Giovannoni 2001; 2004; Bouzayen et al. 2010). The many changes that occur during fruit ripening are driven by the coordinated expression of ripening-related genes that code for enzymes that participate directly in the physiological and biochemical changes. They also encode regulatory proteins that are involved in the signalling pathways, and in the transcriptional machinery that regulate gene expression to trigger the ripening process (Bouzayen et al. 2010; Klee and Giovannoni 2011). For instance, the changes and softening in tomato cell wall are caused by genes that are expressed in a tissue and developmental-specific manner during ripening such as expansins and polygalacturonases (PG) that encode enzymes that disrupt

the hemicellulose-cellulose and pectin network in the cell wall (Brady 1987; Srivastava 2002).

Ripening of fleshy fruits can be historically grouped into two, climacteric and non-climacteric, based on the increased respiration rate and ethylene production that accompanies the ripening process (Abeles et al. 1992). In climacteric fruits, ripening is accompanied by a peak in respiration and a concomitant burst of ethylene, whereas in non-climacteric fruits, such as grapes (*Vitis vinifera*) and citrus, these dramatic changes are lacking (Abeles et al. 1992). The sharp increase in the climacteric ethylene production at the beginning of the ripening process is the controlling factor for the biochemical and physiological changes in the fruits (Lelievre et al. 1997).

## 1.4 The oil palm fruit

Oil palm is a versatile crop where almost all parts of the palm have been utilized economically and ecologically. The most valued component of the palm is the fruit with a national oil extraction rate (OER) of 20.46% (MPOB 2016). The fruits are borne on spikelets that are arranged in a spiral manner to form compact bunches of 10 to 30 kg with 1000-3000 fruits per bunch. The fruit is a drupe that varies in shape from almost spherical to ovoid or elongated and can reach up to 5 cm in length and 30 g in weight. Generally the fruit is dark purple, almost black before it ripens and turns orange red when ripe (Hartley 1988) (Figure 1.1).

**Figure 1.1: Structure of the oil palm fruits in a compact bunch.** Photo credit to Malaysian Palm Oil Board (MPOB) collection 2001.

The fruit is comprised of three layers consisting of the exocarp or external skin, the fleshy mesocarp or pulp and the endocarp or shell that surrounds the endosperm or kernel (Figure 1.2). Anatomical studies have shown that the fruit exocarp is made up of a layer of parenchyma cells whereas the mesocarp is made up of parenchyma cells that contain oil droplets. The mesocarp is also embedded with fibro-vascular and fibre bundles. The endocarp consists of stone cells and the kernel contains aleurone grains and oil droplets. Embedded within the kernel is the embryo (Latif 2000). The fleshy mesocarp produces palm oil whereas the kernel produces palm kernel oil, both

of which are biochemically distinct especially in their fatty acid compositions (Sambanthamurthi et al. 2000).

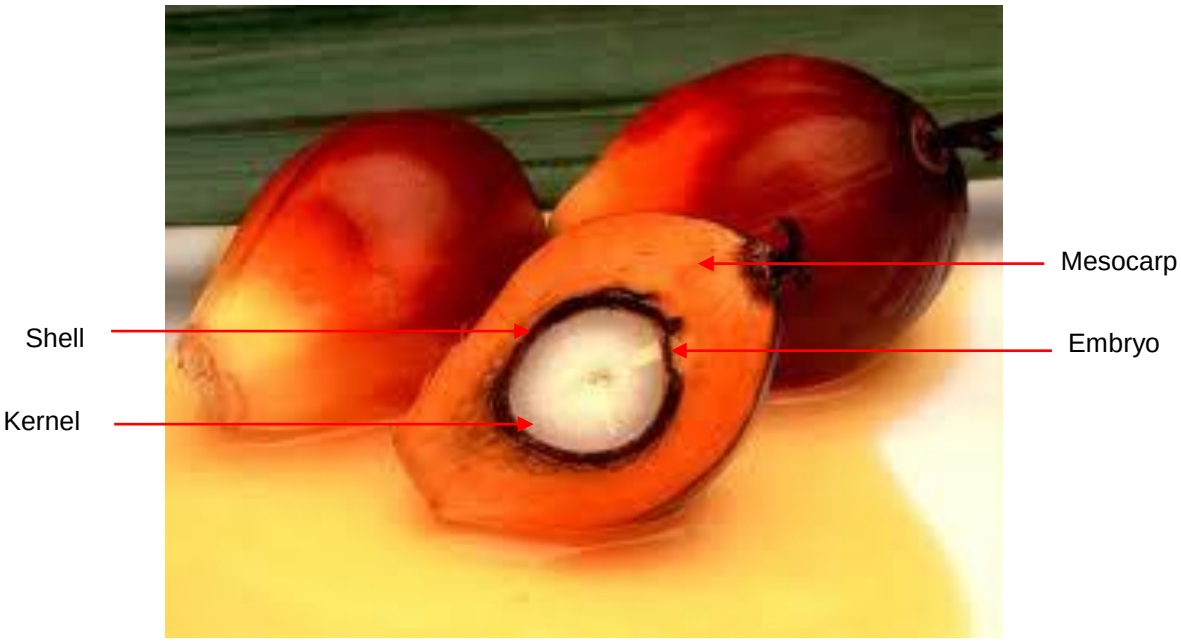



**Figure 1.2: Longitudinal section of the oil palm fruit to show the internal structure which consists of the mesocarp, kernel, shell and embryo.** Photo credit to MPOB collection 2001.

Palm oil derived from the mesocarp has an equal balance of saturated (50%) and unsaturated (39% unsaturated and 11% polyunsaturated) fatty acids. Its fatty acid composition contains a mixture of 44% of palmitic acid (16:0), 39% of oleic acid (18:1), 10% of linoleic acid (18:2), 5% of stearic acid (18:0) and a small mixture of myristic acid (14:0), linolenic acid (18:3), palmitoleic acid (16:1), lauric acid (12:0) and arachidic acid (20:0) (Tang 2000) (Table 1.2).

**Table 1.2: The fatty acid composition (% by weight) of palm oil and palm kernel oil** (Tang 2000).

| Fatty acid | Symbol | Palm Oil (%) | Palm Kernel Oil (%) |
|---|---|---|---|
| Caprylic | C8:0 | - | 4.2 |
| Capric | C10:0 | - | 3.7 |
| Lauric | C12:0 | - | 48.7 |
| Myristic | 14:0 | 1.0 | 15.6 |
| Palmitic | 16:0 | 44.3 | 7.5 |
| Palmitoleic | 16:1Δ9 | 0.1 | - |
| Stearic | 18:0 | 4.6 | 1.8 |
| Oleic | 18:1 Δ9 | 38.7 | 14.8 |
| Linoleic | 18:2 Δ9, 12 | 10.5 | 1.6 |
| Linolenic | 18:3 Δ9, 12, 15 | 0.3 | - |
| Arachidic | 20:0 | 0.3 | - |

Palm kernel oil is comprised of mostly medium chained fatty acids, predominantly 49% lauric acid (12:0) which is similar to coconut oil. Other fatty acid constituents include 16% myristic acid (14:0) followed by 15% oleic acid (18:1) and a mixture of other minor fatty acids (20%) (Tang 2000) (Table 1.2). Due to its high content of saturated fatty acid especially lauric acid and myristic acid, palm kernel oil is widely used in industries for the production of oleochemicals products such as soap and detergents (Hazimah et al. 2011).

The oil palm fruits vary widely and are normally classified based on their internal fruit structure (*dura, tenera, pisifera*), fruit morphology (normal, mantled) and pigmentation of the exocarp (nigrescens, virescens, alberscens). For commercial purposes, the internal fruit structure form is of economic importance (Kushairi and Rajanaidu 2000). The three types of internal fruit form namely *dura, pisifera* and *tenera* are distinguished by shell thickness that is controlled by a single gene (Singh et al. 2013a) (Figure 1.3).The *dura* ($sh^+$, $sh^+$) has a thick-shell between 2 mm - 8 mm while the *pisifera* ($sh^-$, $sh^-$) is shell-less. The cross between the *dura* and the *pisifera* produce the heterozygote *tenera* that has a thin shell of 0.5 mm - 4 mm surrounded by a ring of fibers in the mesocarp. As a result, the *tenera* has a high proportion (60% - 90%) of oil-bearing pulp, or mesocarp, and thus, the oil yield is 30% more as compared to *dura* (Kushairi and Rajanaidu 2000). Hence, the *teneras* have been widely used as the commercial oil palm planting material in Malaysia since 1960 replacing the *Deli dura* which was first used as planting materials for commercial purposes in 1917 (Yusof 2000).

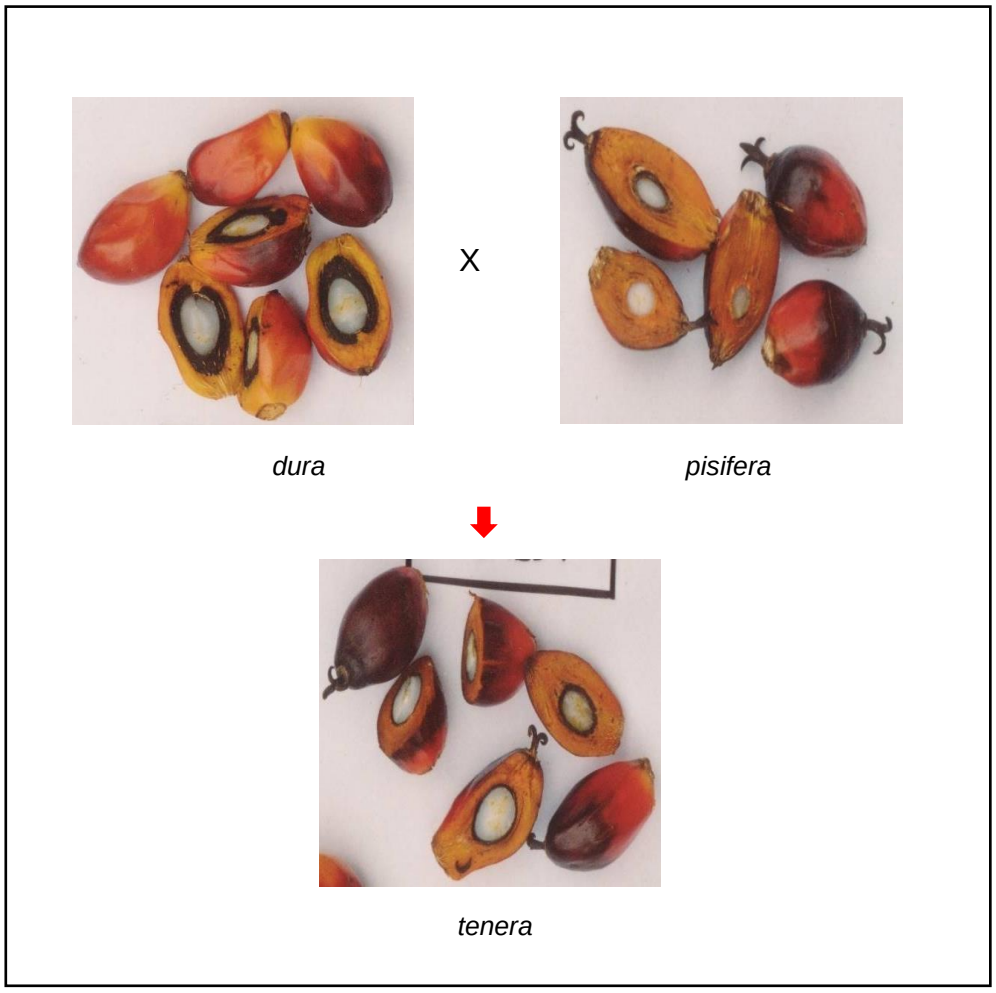



**Figure 1.3: The three fruit forms (*dura, pisifera* and *tenera*) of oil palm (*Elaeis guineensis*) characterised by their shell thickness.** Photo credit to Malaysian Palm Oil Board (MPOB) collection 2001.

### 1.4.1 Development, ripening and detachment of the oil palm fruits

Oil palm fruits are developed from the female inflorescence found between the 17$^{th}$ to 20$^{th}$ leaf axil from the central spear (Corley and Tinker 2003). The mature pistillate flowers are white in colour and anthesis normally occurs between 36-48 hours (h) to allow pollination to take place (Latif 2000; Adam et al. 2005; 2007). Upon pollination, the colour of the flower turns pink or violet and following fertilization, the oil palm fruit develops approximately two weeks later (Adam et al. 2005; Hartley 1988). Ripening of the oil palm fruit bunch normally takes about 6 months from pollination or 20 to 24 WAA (weeks after anthesis) depending on the genotype and environmental factors (Hartley 1988; Azis 1990a; Tranbarger et al. 2011; Mohd Hudzairi et al. 2012). Figure 1.4 shows examples of oil palm fruits at various developmental stages from young until the ripening stage.

Ripening of the fruits on a bunch is unsynchronized. The ripening process normally starts from the outer and top region of the bunch and moves downwards towards the basal (Corley and Thinker 2003). Sambanthamurthi et al. (2000) suggested that variation in the time of pollination is likely the reason for the unsynchronized fruit ripening. The ripening indicator within the fruit bunch is mainly based on the visual observation on the colour changes in the outer fruits (Corley and Thinker 2003; 2016). The ripe fruits on the outer bunch are deep orange in colour and are larger in size in comparison to the inner fruits which are paler and smaller in size (Hartley 1988; Corley and Thinker 2003).

4 WAA

8 WAA

12 WAA

16 WAA

20 WAA

22 WAA

**Figure 1.4: Oil palm fruits at different stages of development from young to the ripening phase.**

Within the inner bunch, parthenocarpic fruits that do not contain endosperm and embryo are also formed (Hartley 1988). The poor growth of the inner fruits is thought to be caused by the limited space for development of the mesocarp and a lack of nutrient uptake (Mohd Haniff and Mohd Roslan 2002).

Following the ripening process, abscission of the fruits will start to take place. The detachment of fruits from the bunch follows the fruit ripening pattern (Hartley 1988; Corley and Thinker 2003). The ripest fruit will start to fall off from the bunch starting from the top of the bunch and moves downwards towards the basal region. One loose fruit detaches from the bunch and is found at the base of the palm is one of the indicator that the fruit bunch is ready for harvest (Idris 2004). However, evidence to show that all the fruits especially those located within the inner bunch have equally matured and ripened is also lacking.

Information on the development of oil palm fruits, maturation and ripening process especially in terms of morphology, structure and biochemical changes has been described (Thomas et al. 1971; Hartley 1988; Azis et al., 1990; Sambanthamurthi et al. 2000; Tranbarger et al. 2011; Bille Ngalle et al. 2013). Here, the changes in oil palm fruits that are prominent and of significant importance are discussed below.

#### 1.4.1.1 Oil synthesis

In the developing oil palm mesocarp, oil is synthesized after 15 WAA and reaches its maximum when the fruit is fully ripened at approximately 20 –

24 WAA depending on the genotype and surrounding environment (Azis et al. 1990; Sambanthamurthi et al. 2000; Aziz et al. 1986; Oo et al. 1986). Morphology, cellular and biochemical studies have also been conducted to support these data (Azis et al. 1990; Sambanthamurthi et al. 2000; Tranbarger et al. 2011).

Within the oil palm fruit bunch, ripening of the fruits is an unsynchronized process (Hartley 1988; Corley and Thinker 2003). Ripening of the outer fruits followed by fruit detachment from the bunch and falling to the ground is a sign that the bunch is ready to be harvested. At this stage the outer fruits are fully ripe and at a maximum oil content unlike the inner fruits (Henderson and Osborne 1999). The inner fruits are less mature and are still capable of synthesizing oil. Although the ripening of the oil palm fruits is unsynchronized, there are contradictory opinions on the termination of oil synthesis in the bunch. Azis (1990) reported that a single loose fruit is the indicator for full bunch ripeness and that oil synthesis is terminated. Sambanthamurthi et al. (2000) reported that oil synthesis is complete at the first sign of fruit abscission and the following increase in oil content is caused by the desiccation process as the fruit ripens. In contrast, Corley and Law (2001) reported that despite the decrease in the moisture content in the mesocarp of ripe fruits, the weight of the oil continues to increase after the first fruit abscission. Oo et al. (1986) also reported that oil content continues to increase over the next 1-2 WAA after abscission of the first fruit. In addition, Mohanaraj and Donough (2013) also found that oil content increases as the number of loose fruits increase implying

that the maximum oil yield can be achieved when all fruits have shed from the bunch. This however is at the expense of reducing oil quality with the increase in free fatty acid (FFA) and the increase in cost and labour work to collect loose fruits (Oo et al. 1986; Mohanaraj and Donough, 2013; Corley and Tinker 2016).

#### 1.4.1.2 Lipid and fatty acid composition

Palm oil is mainly made up of triglycerides (TAG) (Sambanthamurthi et al. 2000). These are comprised of a glycerol backbone and three fatty acids. In the young fruits (8-12 WAA) of the commercial *tenera* and *dura* palms, TAG content is very low accounting for only 7.1% to 14% of the total lipids. During this stage, the polar lipids especially phopsholipids predominate with 54% to 72% of the total lipid content in the mesocarp tissues (Bafor and Osagie, 1988; Oo et al. 1986). As oil synthesis starts at approximately 16 WAA, the TAG content increases while the polar lipids decrease. At 20 WAA in *tenera* palms or 22 WAA in *dura* palms, the TAG reached its maximum content accounting for 76% to 97% of the total lipid while the polar lipids decrease below 2%. At this stage, glycolipids makes up the major content in the polar lipids (Bafor and Osagie, 1986; Oo et al. 1986). Changes in the fatty acid composition were also observed in the fruits from young until the fruit ripening stage. Oo et al. (1986) reported that in young fruits (8-12 WAA), polyunsaturated linoleic acid (18:2) and linolenic acid (18:3) are high as these are typical membrane and chloroplast lipids. During the final weeks of ripening when oil accumulation starts to increase, palmitic acid (16:0) predominates the fatty acid content with

44% followed by oleic acid (18:1) and linoleic acid (18:2) both 39% and 10%, respectively. (Oo et al. 1986; Aziz et al. 1986).

#### 1.4.1.3 Fruit colour

Generally, the common nigrescens fruit type is dark purple or almost black in colour especially those located on the outer bunch while fruits on the inner bunch are much paler (Azis 1984). As the fruits mature and ripens, the outer fruits change colour especially at the base into orange red while the dark purple or black coloured fruits at the apex undergo minimal changes. The inner fruits of the bunch, especially those positioned at the base, also remain pale (Azis 1984; Hartley 1988). A more prominent change in fruit colour is seen in the less common virescens type where immature fruits are green especially those on the outer bunch and turns reddish-orange when ripe (Hartley 1988). This has led to the discovery of the *VIRESCENS* gene that has been shown to encode a R2R3-MYB transcription factor that controls the fruit exocarp colour (Singh et al. 2014).

Biochemical analyses have demonstrated that chlorophyll and carotenoids are present in the developing mesocarp tissues (Azis 1984). The different pigment levels in the mesocarp tissues are related to the stage of fruit development (Ikemefuna and Adamson 1984). In tomato for example, the change in the fruit colour from green in unripe to red in ripe fruits is the result of chlorophyll degradation and carotenoid pigment accumulation (Grierson et al. 1987). In young fruits, the level of carotenoid is the lowest as opposed to the

highest chlorophyll content. In oil palm, as the fruits mature and ripen, the mesocarp turns from yellow-green to orange red as the strongly coloured carotenoids are being synthesized. The carotenoid content is highest in ripe fruits as opposed to the chlorophyll content (Ikemefuna and Adamson 1984; Azis 1984). The high level of carotenoids (normally between 500-700 parts per million [ppm]) which mostly comprises α- and β-carotene (90%) gives the CPO the bright orange-red colour (Jalani and Rajanaidu 2000; Choo and Ng 2011). In plants, carotene plays an important protective role against light and oxidative degradation of membrane lipids (Gruszecki and Strzałka 2005). Carotenoids are also the precursors for pro-vitamin A, a dietary supplement for human consumption (Sambanthamurthi et al. 2000; Benade 2003).

### 1.5 Harvesting and oil quality

The fruit detachment process starts from the tip of the bunch and move down to the basal region (Hartley 1988; Corley and Thinker 2003). One loose fruit found at the base of the palm and the observable colour change of the mesocarp to bright orange red are indicators that the fruit bunch is ready for harvest (Idris 2004). In laser-based fluorimetry and spectrophotometer analysis of CPO, chlorophylls were detected implying that there are variation in the ripeness of oil palm fruits processed in the mills (Tan et al. 1994; 1997). Chlorophyll promotes oxidation of the oil which leads to a reduction in storage stability (Diosady 2005). Thus, to minimize the variation in bunch ripeness, MPOB has implemented/outlined a manual grading standard for bunch

ripeness as a guideline for oil palm mills, plantation and smallholders to ensure high OER and oil quality in Malaysia (MPOB 2003). Despite the guidelines, grading mistakes can still happen due to human error which has prompted researchers to look at other means to help improve the grading system. Some of the methods include the use of a colour vision system (Abdullah et al. 2002; Saeed et al. 2012), artificial intelligence system (Ismail et al. 2009), magnetic resonance imaging and nuclear magnetic resonance (Shaarani et al. 2010) on the whole bunch and sensor that uses inductive frequency techniques to detect the maturity of the fruit (Harun et al. 2013).

Harvesting standard and practices also play an important role in determining oil quality. In Malaysia, due to the labour shortage in the oil palm plantation industry, the time interval between harvests is usually longer than recommended causing more fruit detachment from the bunch. Loose fruits left uncollected account for about 3-5% of the bunch weight and consequently, failure to collect and process the loose fruits contribute to a reduction in OER and a loss in profit (Mohd Solah and Rahim 2009). Harvesting under ripe bunches does not maximize the bunch yield while overripe bunches produces more oil but has a higher FFA content (Corley and Tinker 2003). FFA content in fats and oils is used as an index to determine their quality (Kardash and Turýan 2005). The high FFA content in CPO causes a reduction in oil quality, thus affecting its market price (Sambanthamurthi et al. 2000). The accepted standard for premium oil quality is that the FFA content in CPO should not exceed 5% (PORAM 2015).

FFA are released from the glycerol backbone of the triacylglycerides by the action of an active endogenous lipase enzyme found in abundance in the oil palm mesocarp (Sambanthamurthi et al. 1991; Henderson and Osborne, 1991). Ngando Ebongue et al. (2006) demonstrated that the highest lipase activity was detected at 35°C and at pH9. On the other hand, Sambanthamurthi et al. (1991) and Cadena et al. (2013) showed that by inducing the lipase activity at a cold temperature of 5$^{o}$C, a maximum FFA content was detected. Sambanthamurthi et al. (1995) also observed that the lipase activity is also induced by ripening oil palm fruits. Dessassis (1957) reported that an increase of up to 30% FFA occurred within 5 minutes in crushed mesocarp due to the activity of the lipase enzyme. The FFA content is further enhanced through bruising of the oil palm fruits; poor handling of fresh fruit bunches and fruits in the field, and also delayed processing of the fruits in the mills (Corley and Tinker 2003; Fatin and Rosnah 2014). All these factors, coupled with oxidation of FFA, cause oil rancidity and undesirable flavour and odour in the oil (Chong 2011). Removal of FFA is widely done using alkaline and distillation approaches, but both techniques cause high oil losses and consume high-energy (Azmir 2012).

One of the ways to minimize the losses of oil quality due to endogenous lipase activity is to use oil palm planting materials with lower lipase activities which could be obtained from screening the oil palm germplasm (Wong et al. 2005; 2015; Ngando Ebongue et al. 2008). In the commercial planting material, the *tenera,* the lipase content in the mesocarp varies among genotypes. The FFA level ranges from 22% to as high as 73% when exposed to cold

temperature, indicating that the lipase activity is genotype-dependent (Sambanthamurthi and Kushairi 2002). This discovery makes it possible to select oil palm genotypes with low lipase activity. Through breeding and selection programme, MPOB has developed the PS13 breeding population with low lipase activity (Maizura et al. 2008).

## 1.6 Hormone control of fruit development and ripening

Plant hormones are small signalling molecules present in low amounts to regulate plant growth and development and response to environmental stimuli (Santner and Estelle 2009). Their chemical structures are quite simple but small changes in the hormone concentration can cause considerable changes in plant physiology (Santner and Estelle 2009). In general, a hormone response begins through signal initiation by binding to its receptor, then the signal is transmitted to the downstream molecular cascade and the signal output results in metabolic and cellular changes (Xiong et al. 2009). Most of the time the plant's response is controlled by interaction of more than two hormones (Srivastava, 2002; McAtee et al. 2013; Seymour et al. 2013).

Many studies have shown that a dynamic interaction between multiple hormones have effects on fruit development, maturation and ripening (McAtee et al. 2013). The complex interaction or cross-talk between hormones can happen at many levels including their biosynthesis and transport as well as their signalling pathways but most of the interactions are still not fully understood (Kumar et al. 2014). Figures 1.5 and 1.6 represent the main

phytohormones and genes involved in biosynthesis, perception and signalling during fruit development and ripening of the tomato model plant.

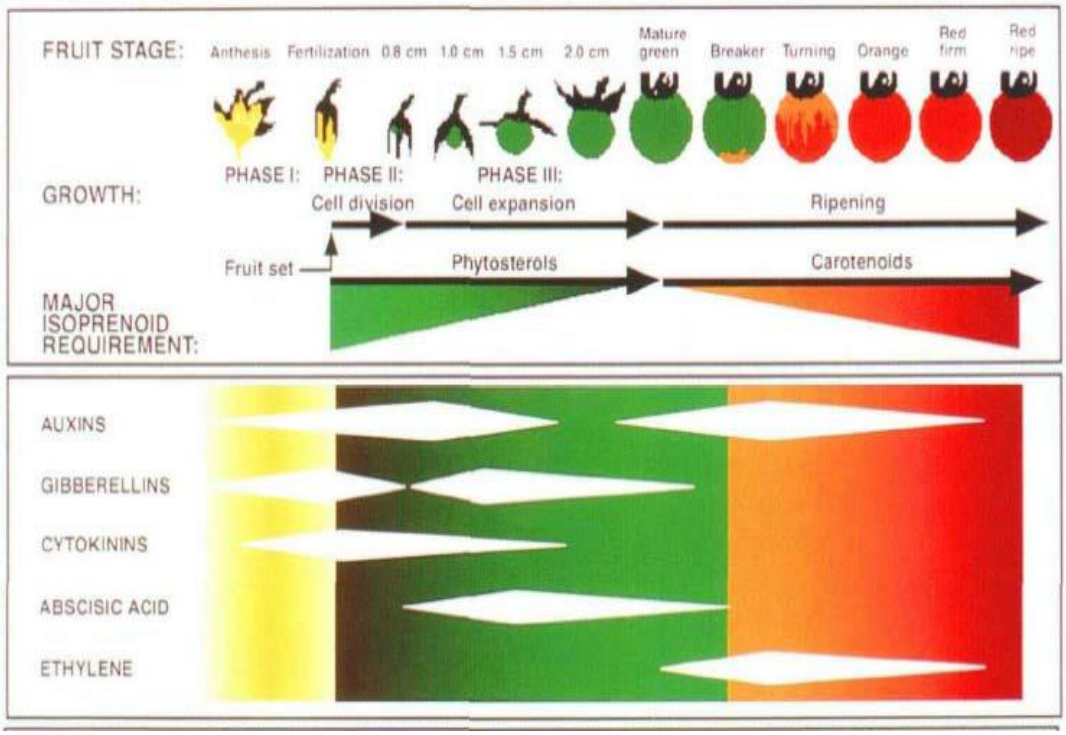



**Figure 1.5: Hormone changes during fruit development using tomato as a model for fleshy fruits** (adapted from Gillaspy et al. 1993).

In fruit development, fruit set represents the first stage of development after fertilization. The transition from fruit set to the fruit growth phase is mainly regulated by auxin and cytokinins (Gillaspy et al. 1993; Pesaresi et al. 2014). At this stage, rapid cell division and cell expansion takes place which lead to a significant increase in fruit weight and size (Pesaresi et al. 2014). In seeds, besides auxin and cytokinin, high levels of gibberellin (GA) have also been detected indicating their role as triggers for the continued growth of the

surrounding tissues *via* cell division and expansion and to determine fruit size (Crane 1964).

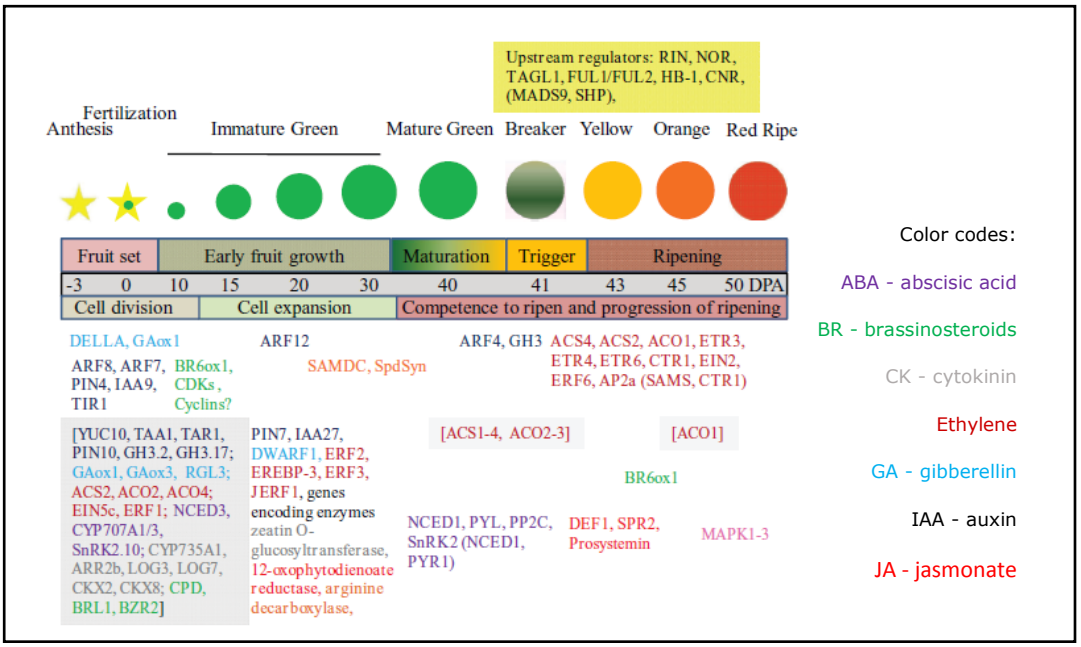



**Figure 1.6: Genes involved in biosynthesis, perception and signalling of various hormones in tomato fruit development and ripening** (adapted from Kumar et al. 2014).

Fruit maturation represents the developmental phase in which the fruit reaches its final size and its ability to ripen but has yet to undergo the ripening process (McAtee et al. 2013; Seymour et al. 2013; Kumar et al. 2014). This phase is also mainly regulated by auxin based on the high level of auxin in the tomato ripening inhibitor (*rin*) mutant at the breaker stage compared to the wild-type fruits (Davey and Van Staden 1978). Over expression of a capsicum *GH3* gene encoding IAA-amino synthase enzyme caused a reduction in auxin levels in tomato fruits which is thought to cause the early fruit ripening phenotype in

transgenic tomato (Liu et al. 2005). In grapes, the most abundant free auxin, indole-3-acetic acid (IAA) decreases while its conjugate form (IAA-Asp) increases concomitant with elevated ethylene synthesis for fruit ripening (Böttcher et al. 2010). The balanced ratio between IAA-and its IAA-conjugate has been suggested to trigger the ripening process (Böttcher et al. 2010).

Fruit ripening especially in climacteric fruits such as tomato, avocado (*Persea americana*), apples (*Malus domestica*) and bananas (*Musa acuminata*) is predominantly regulated by ethylene (Abeles et al. 1992). Application of exogenous ethylene and inhibitor of the gaseous hormone such as 1-methylcyclopropene (1-MCP) have been shown to be affect ripening in climacteric fruits such as apple (Yang et al. 2013) and peach (Rasori et al. 2002). For instance, exogenous ethylene applied to climacteric fruits at pre-climacteric stage helps speed up the ripening process while those treated with 1-MCP delayed fruit ripening (Rasori et al. 2002; Yang et al. 2013).

Abscisic acid (ABA) has also been demonstrated to play an important role in the fruit ripening process (McAtee et al. 2013). In both climacteric and non-climacteric fruits, there is a sharp ABA production peak at the beginning of fruit ripening and/or the ripening process (Leng et al. 2014). In tomato, the rise in ABA production triggers the production of ethylene and fruit ripening (Zhang et al. 2009). Application of exogenous ABA to grape fruits also enhances the production of several proteins involved in the ripening process (Giribaldi et al. 2010). ABA is normally associated with a plant's response to various abiotic stresses such as drought, high temperature, chilling, and salinity (Qin and

Zeevaart 2002; Seki et al. 2007). However, research on the involvement of ABA in regulating abiotic stress in fruits has yet to be reported.

Depending on the variety or genotype, the classified climacteric fruits such as melon (*Cucumis melo*) (Pech et al. 2008) and guava (*Psidium guajava*) (Azzolini et al. 2005) have shown both climacteric and non-climacteric characteristics during ripening. Based on the general fruit classification, ripening of non-climacteric fruit does not involve high respiration rate and a sharp ethylene production peak (McMurchie et al. 1972). However, at the molecular level, the expression level of ethylene receptors in strawberries (*Fragaria* × *ananassa*) (Trainotti et al. 2005) and grapes (Chervin and Deluc 2010) exposed to exogenous ethylene were increased suggesting an important role for ethylene in the ripening of non-climacteric fruits which has yet to be fully understood.

## 1. 7 Manipulation of ethylene perception and signalling

Ethylene is synthesized as a natural product of plant metabolism (Gane 1934). It is produced by all cells at low basal levels during plant development, but its production increases drastically in meristematic, stressed, senescence or ripening tissues (Abeles et al. 1992). Ethylene is synthesized through a two-step conversion from methionine to S-adenoysl-methionine (SAM) to 1-aminocyclopropane-1-carboxylic acid (ACC) and finally to ethylene by three key enzymes such as SAM synthetase; ACC synthase (*ACS*) and ACC oxidase

(*ACO*) (Yang and Hoffman 1984). The conversion of SAM to ACC by ACS is the rate-limiting step in ethylene biosynthesis (Yang and Hoffman 1984).

McMurchie et al. (1972) proposed two systems of ethylene biosynthesis in plants where in system 1, ethylene can inhibit its synthesis (autoinhibitory) while in system 2, it can induce/stimulate its own production (autoinduction/autostimulatory). Based on a tomato model system, the developmental control of ripening is based on the concept of these two systems. System 1 functions prior to ripening where ethylene is produced at low level in immature fruits and exogenous ethylene does not stimulate further synthesis. In contrast, system 2 functions during ripening based on the autocatalytic burst of ethylene (Klee 2004). Molecular studies in tomato demonstrated that both systems are regulated by coordinated developmental and ethylene-induced expression of individual genes such as *ACO, ACS* and ethylene receptor genes (Nakatsuka et al. 1998; Barry et al. 2000).

A lot of work have been performed to alter/delay natural processes such as ripening and senescence (Stearns and Glick 2003; Czarny et al. 2006). Apart from ethylene biosynthetic genes, the ethylene receptor genes are also the key target for altering the ethylene action in plants for crop improvement (Stearns and Glick 2003; Czarny et al. 2006; Little et al. 2009; Agarwal et al. 2012). Ethylene receptors and constitutive triple response 1 (*CTR1*) are negative regulators of ethylene responses in the signalling pathway (Hua and Meyerowitz 1998; Tieman et al. 2000). In the absence of ethylene, the receptors located at the endoplasmic reticulum actively inhibit/suppress

ethylene signalling by activating CTR1, a Raf-like serine/threonine kinase to phosphorylate EIN2 (Ethylene insensitive2), an endoplasmic reticulum Nramp-like transmembrane protein. This prevents EIN2 from transmitting a signal to evoke ethylene responses. EIN2 is a positive regulator of the ethylene signalling pathway where loss of function mutation in EIN2 results in ethylene insensitivity in Arabidopsis, which demonstrates a vital role in ethylene signalling (Stepanova and Alonso 2005). In the presence of ethylene, the receptors bind the ethylene through a copper-cofactor located in the transmembrane domain and suppression is relieved followed by the activation of the signalling pathway (Rodrıguez et al. 1999). Activation of the signalling pathway involves deactivation of the receptor-CTR1 complex causing a loss in the ability to phosphorylate EIN1. As a result, the activated EIN2 undergoes proteolytic changes to release the C-terminal domain and moves to the nucleus (Alonso et al. 1999). In the nucleus, the activated EIN2 initiates a transcriptional cascade involving the ethylene insensitive 3 protein (EIN3), EIN3-like proteins (EIL) and ethylene response factor (ERF) transcription factors and their homologues lead to the activation or suppression of ethylene-responsive target genes to evoke the ethylene responses (Figure 1.7) (Solano et al. 1998; Gallie 2015). Knowledge of the receptors and their signalling pathway components as well as the structural conservation among the plants provide a means to manipulate and regulate the ethylene sensitivity in plants especially to alter fruit ripening, abscission and senescence *via* genetic manipulation (Agarwal et al. 2012).

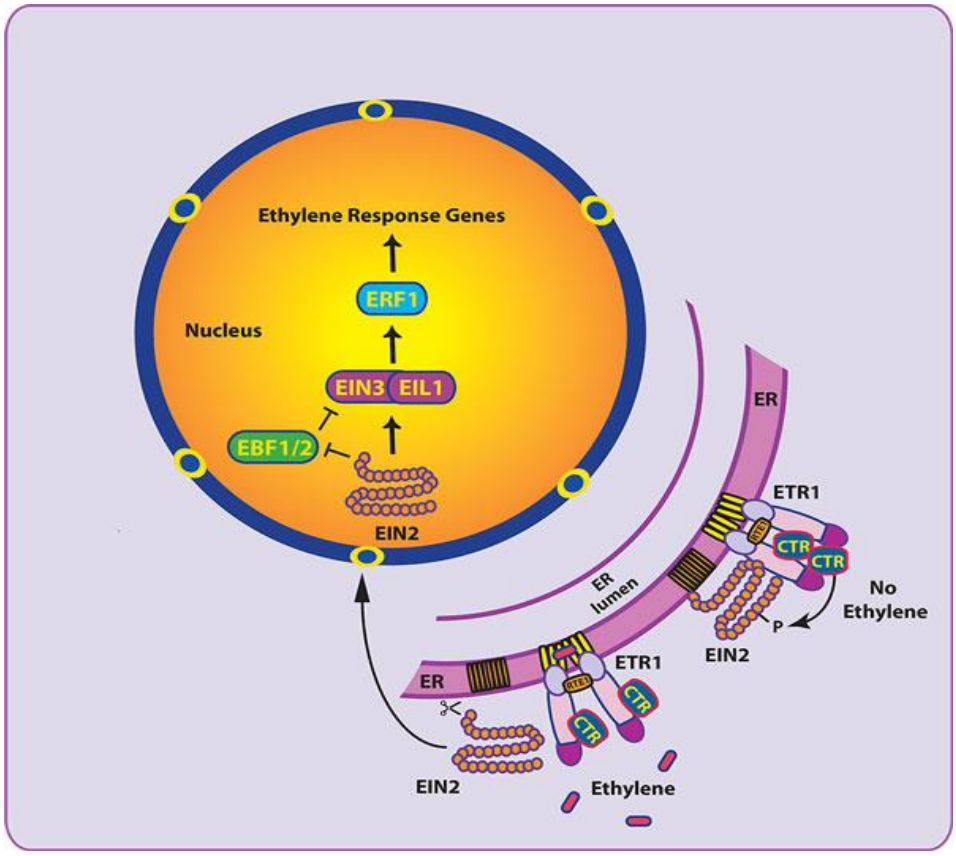



**Figure 1.7: Schematic diagram of the ethylene signalling pathway** (adapted from Gallie 2015).

One of the prominent strategies to produce transgenic plants with the desired trait such as delayed ripening is to use heterologous expression of mutated ethylene receptor genes (Czarny et al. 2006). Ethylene receptors contain a conserved ethylene-binding domain at the N-terminal that binds ethylene in a high affinity manner just as demonstrated in five receptor isoforms in Arabidopsis and tomato (O'Malley et al. 2005). Specific point mutations within this domain disrupt the ability of the receptors to bind ethylene thus creating a plant that is insensitive or confers a severe or reduced response to ethylene (Lanahan et al. 1994; Hua et al. 1995; 1998; Hua and Meyerowitz

1998; Hall et al. 1999). In Arabidopsis, the dominant ethylene insensitive *etr1-1* mutation caused the plant tissues to become insensitive to ethylene hence mutants did not exhibit the ethylene triple response phenotype (Lanahan et al. 1994). Insertion of the dominant *etr1-1* gene into tomato resulted in a reduced ethylene level in the transgenic tomatoes generated and these exhibited delayed fruit ripening (Wilkinson et al. 1997). In addition, a lower ethylene level or reduced sensitivity was also observed in petunia (*Petunia axillaris*) and carnation (*Dianthus caryophyllus*) transformed with the *etr1-1* gene resulting in a longer flower shell-life as the rate at which the flowers normally undergo the senescence and abscission processes were delayed (Wilkinson et al. 1997). Other components of the ethylene signalling pathway such as *CTR1, EIN2* and *EIN3* have also been isolated and used to further understand the role of ethylene in modulating ethylene responses in plants (Hua 2015). The use of these downstream signalling components to create transgenic plants that confer reduced sensitivity or complete insensitivity to ethylene *via* sense and anti-sense approaches has also been suggested (Arora 2005).

The genes responsible for ethylene biosynthesis, perception and signalling are encoded by multigene families (Hua et al. 1998; Barry et al. 2000). This includes genes such as *ACS, ACO* and the ethylene receptors where there are at least 12 *ASC*s, 5 *ACO*s and 5 ethylene receptor isoforms in Arabidopsis (Lin et al. 2009). The expression of these multiple gene isoforms are diverse and are differentially regulated by various developmental, hormonal or environmental conditions, where in many situations is in response to

ethylene production (Hua et al. 1998; Tieman et al. 2000). Thus, for genetic manipulation of ethylene sensitivity levels, a more stable system is needed to transfer the desired trait to produced transgenic plants with modified ethylene sensitivity. The use of a tissue and/or organ specific promoter that is induced developmentally and/or by external stimuli has been demonstrated to regulate the ethylene sensitivity level to obtain the desired trait(s). For instance, over-expression of mutated ethylene receptor genes (*Cm-ERS1/HS70A* and *Cm-ETR1/H69A)* under the control of a tapetum-specific promoter induced male sterility in tobacco (*Nicotiana tabacum*) without causing secondary effects to the transgenic plant unlike the use of a constitutive promoter (Takada et al. 2005). In addition, transformation of the *Cm-ERS1/HS70A* and *Cm-ETR1/H69A* genes can also induce sterility in heterologous transgenic plants such as lettuce (*Lactuca sativa*) (Takada et al. 2007). The use of the mutated ethylene receptor to induce male sterility systems also provides a good strategy to prevent transgene flow through pollen dispersal that can cause risk associated with transgenic plants to other plants (Takada et al. 2005).

### 1.7.1 Ethylene studies in oil palm

It is surprising that in an important crop such as oil palm, fruit development and ripening especially at the hormonal level has not received as much attention as the physiological and biochemical aspects. To date, hormonal changes at the development, maturation and ripening of oil palm fruits have been described by Tranbarger et al. (2011) and Teh et al. (2013).

Up until recently, earlier reports on hormone analyses on oil palm was performed to understand the mechanism of growth regulators in the development of immature inflorescences (Huntley et al. 2002). Habib et al. (2012) investigated the differences in the hormonal level in leaf materials between tissue culture derived truncated leaf and wild type oil palm seedlings. More studies have also emerged to understand hormonal regulation of oil palm height (Mohd. Nasriq et al. 2016) as well as to investigate the role of hormones in flower abnormality (Ooi, pers. comm).

Fruit ripening especially in climacteric fruits has always been associated with an outburst of ethylene (Abeles et al. 1992). The possible role of ethylene in triggering oil production and other changes associated with oil palm fruit ripening has been examined but it not in great detail (Mohd Haniff 1986; Sambanthamurthi et al. 1991). Genes as well as ESTs associated with ethylene synthesis and signal transduction pathway have been identified from the oil palm mesocarp indicating an important role for ethylene in development and ripening as well as the abscission processes of the oil palm fruits (Nurniwalis et al. 2008; Tranbarger et al. 2011). Earlier studies by Henderson and Osborne (1999) detected an increase in the ethylene production in the mesocarp during the late stages of fruit ripening (140-145 days after anthesis [DAA]). Tranbarger et al. (2011) categorized the oil palm fruit as climacteric based on the maximum 35 fold increase of autocatalytic burst of ethylene in the fruits at the ripening stage of 160 DAA. This is also supported by transcriptome sequence data showing the transcript profile of genes involved in the ethylene biosynthetic,

perception and signalling pathways similar to the two system control of fruit development and ripening in tomato (Nakatsuka et al. 1998; Barry et al. 2000; Tranbarger et al. 2011).

Ethylene has also been shown to induce the abscission process in the primary zone of oil palm fruits during the late stages of fruit ripening (Henderson and Osborne 1994). At a laboratory scale, application of 10 ppm exogenous ethylene to oil palm fruit spikelets aged 30, 120 and 160 days after pollination in a controlled environment has resulted in a synchronized fruit shedding within 24 h (Roongsattham et al. 2012). High PG activity at the base of the oil palm fruit-pedicel junction was observed just prior to the abscission process indicating an important role in the fruit shedding process (Henderson and Osborne 1994). Fourteen oil palm PG genes were identified and these genes showed differential gene expression profiles at the base of the fruit prior to the abscission process. The most prominent expression with an increase of 700-5000 fold was shown by the *PG4* gene when exposed to ethylene (Roongsattham et al. 2012).

Plants respond to ethylene through the action of the ethylene receptors, the first component in the ethylene signalling pathway (Binder et al. 2012). A full-length ethylene response sensor (ERS)-like ethylene receptor (*EgERD3*) has been isolated and characterised from the oil palm (Nurniwalis 2006; Nurniwalis et al. 2012). The expression pattern of *EgERD3* matched the development and ripening of the oil palm fruits. It is also similar to the pattern of oil synthesis as in the mesocarp tissues, possibly indicating that the

expression of the gene correlates closely with the pattern of oil synthesis in these tissues (Nurniwalis 2006). Similar transcript expression patterns were also found in the ethylene receptor *LeNR* in tomato (Payton et al. 1996) and the *ERS1*-type in peach (Rasori et al. 2002) during the ripening process of such fruits. These ethylene receptors are ethylene-dependent and the expression has also been shown to be up-regulated by ethylene. Fruit-specific suppression of *LeETR4*, another member of the ethylene receptor family in tomato demonstrated an early ripening characteristic, a different role to that of *LeNR* which suggests that the different members of the ethylene receptor family play non-overlapping roles in controlling the ripening in tomato fruits (Kevany et al. 2008).

## 1.8 Plant genome sequencing

A number of conventional approaches have been employed for gene discovery in plants. In oil palm for example, various methods have been used which include differential screening (Siti Nor Akmar et al. 1995), subtractive hybridization (Siti Nor Akmar et al. 1996), differential display (Nurniwalis and Siti Nor Akmar, 2000; 2001) and generation of ESTs (Nurniwalis, 2006; Nurniwalis et al. 2008; Low et al. 2008). These methods were very laborious and time consuming. However, the availability to access to the oil palm genome data can facilitate the discovery of genes on a much larger scale and in a shorter amount of time. Hence, allowing a rapid understanding on the role and function of the target gene in oil palm.

In plants, whole genome sequencing was first initiated in the late 1990s. Arabidopsis was the first flowering plant genome to be sequenced and published (The Arabidopsis Genome Initiative 2000). Due to its short generation time and small genome size (125 Mbp), Arabidopsis has become an important model system for identifying genes and determining their functions. The Arabidopsis genome can be accessed *via* The Arabidopsis Information Resource (TAIR) (http://www.arabidopsis.org) and has allowed researchers access to the entire plant DNA. It has also provided the foundation for a more comprehensive comparison of conserved processes in all eukaryotes, identifying a wide range of plant-specific gene functions and establishing rapid systematic ways to identify genes for crop improvement (The Arabidopsis Genome Initiative 2000). Following this success, sequencing of genomes from other plants has followed. This includes important crops such as rice *(Oryza sativa)* (Goff et al. 2002), maize *(Zea mays)* (Schnable et al. 2009), soybean (Schmutz et al. 2010) and potato *(Solanum tuberosum)* (The Potato Genome Consortium 2011) among other agriculturally significant crop. Genome sequencing of these crops serves as a platform to improve crop production especially those of critical importance as food and bioenergy security for the future (Rounsley et al. 2009).

For fruit/horticultural crops, at least 18 genomes have also been sequenced (Gapper et al. 2014) including tomato (The Tomato Genome Consortium 2012), apple (Velasco et al. 2010), strawberry (Shulaev et al. 2011), banana (D'Hont et al. 2012) and date palm (Al-Dous et al. 2011). The

use of genome sequence data coupled with the advancement of the latest 'omics' platforms such as transcriptomics, proteomics and metabolomics has also allowed tremendous understanding of fruit developmental and ripening in many fruit crops (Gapper et al. 2014). As a result, many important fundamental questions relating to important horticultural crops are being answered, and novel approaches with application to industry are in progress (Osorio et al. 2013, Gapper et al. 2014).

In 2009, MPOB's first draft version of the oil palm (*E. guineensis*, *pisifera* fruit form and *E. oleifera*) genome sequence was announced and later published in 2013 (Wahid 2009; Singh et al. 2013a). Apart from MPOB, the Asiatic Centre for Genome Technology Sdn. Bhd, Malaysia and Sime Darby Bhd., Malaysia have also sequenced the oil palm genome. Recently, the draft genome sequence of an elite *dura* palm was also published (Jin et al. 2016). The oil palm genome sequence is an important component that would contribute to the improvement of oil palm yields and add value to the oil characteristics for the production of food, animal feed and biofuel (Mohd Basri 2009). The availability of the oil palm genome data provides an opportunity for researchers to speed up their quest towards achieving sustainable oil palm production. The discovery of two important genes such as the *SHELL* gene that differentiates between the three oil palm fruit forms, hence controlling oil yield and the *VIRESCENS (VIR)* gene that controls fruit colour demonstrates the significance and advantages of oil palm genome sequencing (Singh et al. 2013b; 2014). Identification of these two genes allows selection of their

economic and colour traits even at the early-nursery stage prior to introduction in the field (Singh et al. 2013b; 2014). Other benefits of having genome sequence data is that members of multigene families can also be identified in a much shorter time. In plants such as Arabidopsis and tomato, members of the ethylene receptor family have been shown to play overlapping and non-overlapping roles (Kevany et al. 2008; Binder et al. 2012), making the identification and characterisation of individual genes within the multigene family important.

## 1.9 Oil palm genetic modification

The oil palm crop planted in Malaysia has spurred the economic and social development to the country. However, despite the many benefits it can offer, the oil palm industry still faces challenges in their effort to maintain sustainability. These included shrinking of arable land to cultivate oil palm, labour issues, and threats by pest and disease (MPOB 2016). Under the Malaysian Palm Oil National Key Economic Area (NKEA), improving FFB yield and generation of higher value products are some of the targets to increase the productivity and profit of the oil palm industry (Economic Transformation Programme, 2010). The increase in yield is possible through oil palm crop improvement *via* conventional breeding and good crop management in oil palm plantations (Breure 2003; Wahid et al. 2005; Murphy 2014). In addition, oil palm biotechnology also offer tools that can be used to expedite crop improvement

especially through genomics and genetic engineering (Sambanthamurthi et al. 2009).

At MPOB, the oil palm genetic engineering programme started in the late 1980s involving various research approaches such as biochemical studies, target gene(s) and promoter(s) isolation as well as genetic transformation (Cheah et al. 1995). The main target of the programme is to produce palms with higher oleic acid content (Cheah et al. 1995). Oleic acid is monounsaturated and can be used for diversification to penetrate into non-traditional markets such as liquid and salad oils. It can also serve as an attractive feedstock for further fatty acid modifications in the oleochemical industry in Malaysia (Cheah et al. 1995; Parveez et al. 2015). Subsequently, the target of the programme was expanded to increase stearic acid, palmitoleic acid, ricinoleic acid, lycopene (carotenoid) and biodegradable plastics content in the mesocarp (Parveez 2003; Parveez et al. 2015).

In any genetic engineering work, to ensure its success, several tools need to be made available. This include the availability of target gene(s), promoters (tissue-specific/constitutive) as well as a stable transformation system (Siti Nor Akmar et al. 2003; Parveez et al. 2015). For oil palm to be genetically engineered for production of improved oil quality and yield improvement, this requires modification of the fatty acid composition. To date, several key genes in lipid biosynthesis and ripening have been identified and subsequently isolated (Siti Nor Akmar et al. 2001; Parveez et al. 2015). This include genes encoding β-keto-acyl-ACP synthase II (*KAS-II*) (Umi Salamah et

al. 2012), palmitoyl-ACP thioesterase (Abrizah et al. 1999), stearoyl-ACP desaturase (Siti Nor Akmar et al., 1999), oleoyl-ACP desaturase (Syahanim et al. 2007) and lipase (Nurniwalis et al. 2007; 2015) among others.

Besides fatty acid biosynthesis genes, the possibility of genetic manipulation of other genes that can contribute towards yield improvement need to be looked into as well. Potential genes of interest with possible functions during the development and ripening processes of the oil palm fruits can be discovered from many sources which include ESTs(Nurniwalis 2006; Nurniwalis et al. 2008) and transcriptome data among others (Tranbarger at al. 2011). The ability to delay fruit ripening and shedding for example, without the expense of reducing oil quality may help facilitate harvesting, which may in time save manpower, cost and ultimately contribute to the increase in oil yield.

To produce transgenic plants with the desired trait, the transgene carrying the trait of interest has to be functionally expressed for production of heterologous protein (Juven-Gershon and Kadonagaa 2010). The expression of the transgene is highly dependent on the promoter. The promoter is responsible to initiate and regulate the transcription process of the transgene, hence allowing gene expression (Porto et al. 2013). Studies on the gene expression profile have served as background information to isolate the corresponding promoter for genetic manipulation (Porto et al. 2013). Tissue-specific promoter allows the expression of transgene to be directed to a particular tissue where the promoter is active (Zheng and Baum 2008). The most important step towards isolating tissue specific promoter is to isolate the

tissue-specific genes. Among the approaches there were employed to isolate these cDNAs include differential screening (Siti Nor Akmar et al. 1996) and subtractive hybridization (Siti Nor Akmar et al. 1995).

In oil palm, to genetically engineer the palm to modify the mesocarp oil composition, a mesocarp-specific promoter is required. This is to ensure that the transgene is expressed in the fruit mesocarp and to avoid accumulation of the targeted products in other parts of the plant that could be detrimental to the plant (Siti Nor Akmar et al. 2003; Parveez et al. 2015). MSP1, is the first mesocarp-specific promoter isolated from oil palm and is being used in the MPOB genetic engineering program to drive the specific expression of the transgene to the mesocarp (Siti Nor Akmar and Zubaidah, 2008). As different promoters have different strengths and specificity to drive the expression of transgenes (Que et al. 1997), it was envisaged that more mesocarp-specific promoters are needed to drive the expression of the targeted genes specifically to the mesocarp. This is also one of the strategies to avoid homology dependent gene silencing especially when interacting genes share sequence homology in the promoter region (Mette et al. 1999). A 90 bp homology in the promoter sequence is sufficient to cause the gene to be silent in the plant (Vaucheret 1993). Thus, effort to search for mesocarp-specific gene candidates have been progressing steadily using several approaches such as differential display and differential screening (Nurniwalis and Siti Nor Akmar 2001), RT-PCR (Nurniwalis et al. 2003) and generation of ESTs (Nurniwalis 2006; Nurniwalis et al. 2008).

The development and establishment of a stable transformation system for delivery of foreign DNA into host cell as well as selection of transformed plants is very important in any genetic engineering work to ensure regeneration into a complete plant (Birch 1997). For oil palm, the first successful regeneration of transgenic was achieved using biolistics or microprojectile-bombardment. This work involved a lot of optimization on the physical and biological parameters to maximize delivery of DNA into embryogenic calli, selecting the most effective constitutive promoter and selection agents and its concentration (Parveez et al. 1996; 1997; 1998: Chowdhury et al. 1997). The transformed oil palm embryogenic calli were regenerated on a selection medium containing the herbicide Basta (glufosinate ammonium as the active ingredient) and confirmed *via* molecular and protein analyses (Parveez 2000). As biolistic procedures tend to insert multiple copies of transformed gene(s) into the genome of transgenic plants, another oil palm transformation system mediated by Agrobacterium was developed (Masli et al. 2009).

In the efforts to produce the desired traits in oil palm, several transformation vectors containing the MSP1 promoter have been designed and constructed (Masani and Parveez 2008). These constructs were bombarded into oil palm embryogenic cultures using Agrobacterium-mediated transformation and generated glufosinate ammoinium-resistant embryogenic calli that was later regenerated into plantlets (Parveez et al. 2015). Some of the palms have already been transferred to soil and grown in the biosafety nursery. The success of the transformation procedure was confirmed *via* PCR

amplification and Southern hybridization where presence of transgenes were detected in a few of the plantlets (Nurfahisza et al. 2014). However, the presence of chimeric transgenic palms were also detected suggesting the need to make further improvements on the developed transformation protocols.

Recently, another method for oil palm transformation was reported using oil palm protoplast (Masani et al. 2014). The protocol for the stable transformation of oil palm protoplasts was established using DNA microinjection PEG-mediated transfection (Masani et al. 2014). This possibly allow the regeneration of transgenic oil palm from a single cell and can potentially eliminate the problem of producing chimeric plants as recently reported for oil palm (Nurfahisza et al. 2014).

## 1.10 Aims and objectives

### 1.10.1 Aims

Malaysian oil palm industry faces a number of challenges including the scarcity of suitable land for expansion of oil palm cultivation and labour shortage. Conventional breeding programme has made great contributions to yield improvement (Rival and Jaligot, 2010). However, unsynchronized ripening and subsequent shedding of the ripest fruits before harvest limit further yield improvement and affects oil quality. As an important crop, regulation of oil palm fruit ripening especially the unsynchronization of fruit ripening within the fruit bunches has not been given much attention. Therefore, this study is aimed at understanding the mechanism and changes that occur within the FFB during

fruit development. This study is focused on the biochemical, physiological and molecular changes in particular on the ethylene (fruit ripening) and lipase (lipid degradation) involvement throughout the fruit developmental process.

In this study, it was hypotesised that the location and position of fruits within the bunch from young until the ripening stages would affect and display significant changes in the production and accumulation of the biochemical compounds chosen for this study. The use of the oil palm genome data would also allow the quest and search to identify target genes (with potential use for genetic modification) that belong to multigene families to be performed in a shorter amount of time compared to the conventional method. It was also hyphothesised that there would also be differential gene expression of the target genes in various oil palm tissues especially during fruit development and abscission.

### 1.10.2 Objectives

The aim of this study was achieved using the following objectives:

- To evaluate the biochemical and physiological changes that occur in the oil palm fruits during fruit development based on their locations within the FFB (apical, central and basal) and positions (outer, middle and inner) within the spikelets. Measurements included quantification of carotenoid, fatty acid composition (FAC) and ethylene production.

- To identify, isolate and characterise members of the ethylene receptor family from oil palm genome using bioinformatics tools and laboratory approaches.
- To determine the expression patterns of the members of the oil palm ethylene receptor family in various oil palm tissues.
- To isolate and characterise a lipase class 3 gene and promoter from the oil palm.
- To identify other putative lipase class 3 gene family from the oil palm genome using bioinformatics tools.

## 1.11 Structure of the thesis

The thesis is comprised of five chapters.

- Chapter 1 presents the background information in relation to the research study including relevant literature reviews. In addition, the aims and objectives of the project and a brief outline of the structure of the thesis are also described here.
- Chapter 2 presents the biochemical and physiological analyses that included extraction of fatty acid and carotenoid as well as ethylene gas measurement from the various fruit locations (apical, central and basal) and positions (outer, middle and inner) within FFBs of varying stages of fruit development.
- Chapter 3 presents the use of bioinformatics tools to identify, mine and map the ethylene receptor gene family from the oil palm genome

followed by isolation and characterisation of the individual genes and its corresponding promoters.

- Chapter 4 presents the isolation and characterisation of the full-length *FLL1* cDNA encoding a lipase class 3, prediction of the protein structure, isolation of the promoter prior to *in vivo* transient promoter assay. In addition, the oil palm genome date was used to search for putative members of the lipase class 3 gene family.
- Finally, Chapter 5 summarises and concludes the overall research work. This chapter also describes the impact of the research findings and provides suggestions/recommendation for future research.

# CHAPTER 2

# BIOCHEMICAL AND PHYSIOLOGICAL CHANGES DURING CLONAL PALM FRUIT DEVELOPMENT

## 2.1 Introduction

The oil palm fruits are produced in thousands and are formed in a tight bunch. An understanding of bunch development and ripening is important if harvesting time and yield are to be optimised. For an important crop, regulation of oil palm fruit ripening especially on the non-synchronization of the fruit ripening process has not been given much attention. Thus, to further understand this process within the oil palm fruit bunch, this chapter investigates the biochemical and physiological changes on three traits namely carotenoid production, fatty acid accumulation and ethylene synthesis within the different age fruit bunches during clonal palm fruit development and the implications of the key results are discussed.

## 2.2 Materials and methods

### 2.2.1 Plant Material

Fresh fruit bunches (FFB) from five clonal oil palm (Clone P164, Trial 0.393) were used in this study. The FFBs were harvested from MPOB Research Station Ulu Paka, Terengganu, Malaysia. Independent inflorescences from each palm were tagged at anthesis at six to seven fruit development stages from young until ripening. Processing of the individual FFB involved dividing

the spikelets into three regions based on location within/on the bunch: bottom or basal, central and the apical or apical region of the bunch (Keshvadi et al. 2012). Within each region, the spikelets were randomly sampled and undamaged fruits were collected based on their position within the spikelets: the outer or dark coloured fruits, the middle layer containing a mixture of dark and pale coloured fruits and the inner and pale coloured fruits. Figure 2.1 illustrates a pictorial representation of the fruit bunch processing. For fatty acid and carotenoid extraction analyses, the mesocarp tissues were processed by separating the fruits from the bunch. The fruits were then rinsed with distilled water and after removal of the exocarp, the mesocarp tissues were frozen immediately in liquid nitrogen and kept at -70$^{o}$C prior to usage.

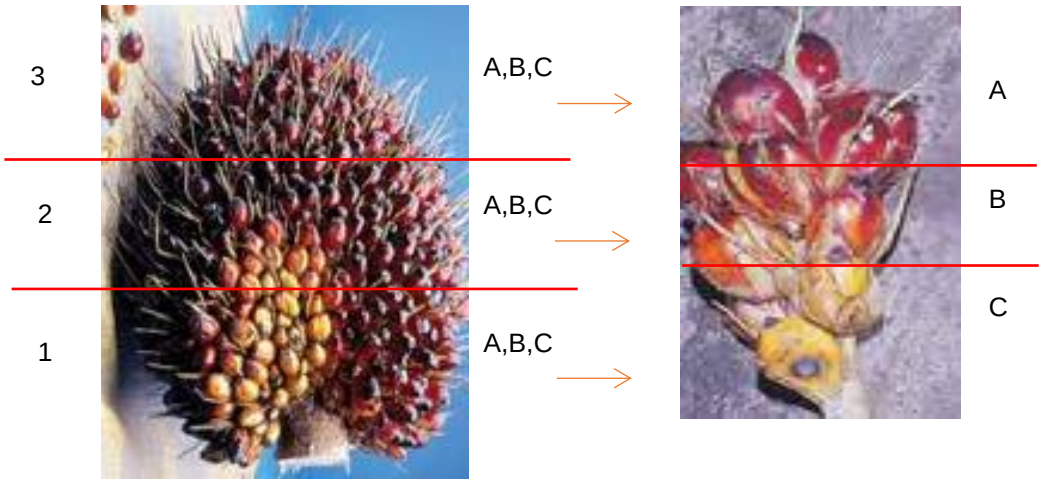



**Figure 2.1: Sampling of the individual bunch based on the fruit location within the FFB and their position within the fruit spikelets**. Numbers 1, 2 and 3 represents the basal/bottom, central and apical/top region of the FFB respectively, whereas A, B, C represents the outer, middle and inner positioned fruits respectively, within the spikelets.

### 2.2.2 Determination of β-carotene content

#### 2.2.2.1 Standard preparation for β-carotene

β-carotene standard (Sigma-Aldrich) was dissolved in chloroform (1 mg/ml) and diluted into a series of concentrations: 0.5, 1.0, 1.5, 2.0, 2.5 and 5.0 µg/ml. The increasing standard concentration was used to plot a standard curve for β-carotene quantification. Correlation between the β-carotene concentration and the spectrophotometry absorbance reading at $A_{446}$ was used to determine the β-carotene content in the samples.

#### 2.2.2.2 Carotenoid extraction

Carotenoid extraction was performed based on the method of Kaur et al. (2002) with some modifications. Five grams of frozen mesocarp tissues were ground in 10 ml 0.1M MES buffer, pH 6. An equal volume of acetone was added and the mixture was mixed with gentle shaking for 10 min. The solution was carefully removed and the remains were extracted twice with 10 ml of acetone. Diethyl ether (10 ml) was then added to the acetone pool extract and mixed gently. One ml of 3.3% (w/v) NaCl was added to the mixture to allow partitioning of the solvents. Once the solvent partitioning had settled, the upper layer was collected and transferred to a new tube. The combined solution was dried under nitrogen gas flow to produce the crude carotenoid extracts.

To remove unwanted lipids and TAGs from the crude carotenoid extracts, the cold saponification process was performed. Ten ml of 100% (v/v) ethanol was added to the crude carotenoid extract and mixed to dissolve the

dry extract. One ml of 60% (w/v) KOH was then added to the solution, mixed gently and the solution was left to crystalize at -20$^{o}$C for at least 1 h. Then, 10 ml of diethyl ether was added to the solution. Using a dropper, two to three volumes of sterile water was added to the solution to separate the organic layer containing the carotenoid extracts and the aqueous layer. The upper organic layer was collected, transferred into a new tube and dried under nitrogen gas flow.

2.2.2.3 Quantification of carotenoids

The total carotenoid extract content from oil palm mesocarp was quantified using UV-1800 UV-VIS Spectrophotometer (Shimadzu, Japan) equipped with the UV-probe 2.3 Software. Prior to spectrum quantification, the dried saponified carotenoid extracts from section 3.2.2.2 were dissolved in 1.5 ml diethyl ether and filtered through a 0.2 µm Ministart® filter (Sartorius). One hundred µl of the saponified carotenoid sample was dried under nitrogen gas flow and dissolved in 1.5 ml iso-octane. The UV absorbance was performed in the range of 200 nm to 750 nm. Total carotenoid was expressed as µg/g fresh weight (fw) or (ppm) of β-carotene content calculated based on the calibration curve of the corresponding β-carotene standard. The β-carotene content in each sample was then calculated using a linear regression $y = mx + c$, where $y$ = sample absorbance at 446 nm, $x$ = β-carotene content of the carotenoid sample, $c$ = linear through zero. The regression equation and correlation coefficient (R2) were obtained using Microsoft® Excel Version 2013.

### 2.2.3 Analysis of Fatty Acid Composition (FAC)

#### 2.2.3.1 Lipid extraction

Lipids were extracted from oil palm mesocarp according to the method of Christie (1989) with a slight modification. Two grams of frozen tissue were ground under liquid nitrogen until fine and powdery form, mixed with 40 ml NaCl (0.9%) then transferred to 100 ml chloroform:methanol (2:1). The mixture was shaken thoroughly and left to settle overnight. The bottom layer was collected and filtered and the lipids were obtained by removing the solvent using a rotary evaporator.

#### 2.2.3.2 Preparation of Fatty Acid Methyl Esters (FAME)

FAME was prepared using the acidic methylation method as described by Christie (1989). The dried lipid sample was dissolved in 2 ml toluene and transferred into a new test tube. Two ml of 1% sulphuric acid in methanol was added and the mixture was incubated at 80$^{o}$C for 2 h in a heating block set. Then, 5 ml of 5% NaCl and 2 ml hexane were added to the mixture and the sample mixture was centrifuged for 2 min at 2500 rpm. Using a glass Pasteur pipette, the upper layer (containing the methyl esters) was transferred into a new vial and washed with 5 ml of 2% potassium bicarbonate ($KHCO_3$) and centrifuged for 2 min at 2500 rpm. The upper layer was collected, filtered and removed *via* evaporation in a stream of nitrogen gas. The dried sample was dissolved in 0.5 ml hexane prior to injection in the Gas Chromatography Mass Spectrometry (GC-MS).

For oil samples especially from 20 WAA and 22 WAA fruits, after solvent removal, FAME preparation was performed as described by PORIM (1995). Fifty µl of oil was added to 1 ml hexane and mixed thoroughly. Sodium methylate (100 µl) was added to the mixture, vortexed for 1 min and centrifuged shortly. The upper layer (~ 800 µl) was collected and stored at -20$^{o}$C prior to injection in the GC-MS.

2.2.3.3 GC-MS Analysis

GC-MS was performed on a 6890N Agilent chromatograph connected to a 5973 Agilent mass selective detector (Agilent Technologies, USA). The chromatography was carried out using a 30 m x 0.25 mm HP-5MS capillary column with a film thickness of 0.25 µm. Helium was used as the carrier gas with a flow rate of 1 ml/min with constant flow mode. Sample injection (1 µl) was carried out using the split/split-less injector with split ratio of 100:1. The oven temperature was set from 100$^{o}$C to 160$^{o}$C at a rate of 6$^{o}$C per min, while the final temperature was set at 240$^{o}$C at a rate of 3$^{o}$C per min with a total runtime of 42.07 min. The system was pre-calibrated with RM-6 Standard (Supelco, USA). Mass spectral analysis was performed in electron ionization mode.

2.2.3.4 Data Analysis

Data analysis was performed as described by Bahariah et al. (2007). The fatty acids were identified partially by their retention time within the GC, but mainly

by their mass spectra. The National Institute of Standards and Technology (NIST) 98 Mass Spectral Library which contains over 150,000 standard reference spectra of high quality, including most FAMEs of interest was used to match spectra in oil palm to reference spectra. The amount of each fatty acid present in a sample was computed using Automated Mass Spectral Deconvolution and Identification System programme developed at NIST (Mallard and Reed, 1997). The quantity of each FA was estimated using this programme by integrating and calculating the correlation area under the GC total ion current peak associated with each FA. Percentage of each FA was obtained using a data integrator that converts the correlation area under each peak to a percentage of the total area under all the peaks. The iodine value (IV) was measured based on percentage calculation of four fatty acid components following the formula as described by PORIM (1995).

### 2.2.4 Measurement of Ethylene Production

#### 2.2.4.1 Fruit preparation

Randomly selected individual oil palm fruits from the different region and location within the FFBs were stripped from the spikelets, tepals (Van Heel et al. 1987; Henderson and Osborne, 1994) removed and surface sterilized prior to analysis. The weight of the fruits, the volume of the glass jars used for the experiment and the fruit volume in the jar were measured and recorded. To avoid wound ethylene, the individual fruits were left at room temperature for 1

– 2 hours (Henderson and Osborne, 1994) then placed in the individual glass jars and sealed prior to ethylene measurement after 24 h.

2.2.4.2 Gas chromatography (GC) analysis

The ethylene production from the individual fruits was measured by removing 1 ml gas sample from each sealed jars using a gas tight hypodermic syringe and injecting it into a GC (Clarus 500, Perkin Elmer, USA). The GC was equipped with a stainless steel column (GS carbonplot, 30 m x 0.53 mm, Agilent Technologies, USA) and a flame ionization detector (FID). Helium was used as the gas carrier with a flow rate of 6 ml / min. Temperatures for the oven, injector and FID was 240$^{o}$C, 150$^{o}$C and 220$^{o}$C, respectively. Calibration of the GC was carried out by injecting one ml of external ethylene gas standard (10 ppm) (Scotty Specialty Gases, Bellefonte, PA). The amount of ethylene was expressed in nl/kg/h.

2.2.5 Statistical analysis

All results were analysed using the SAS Version 9.2 (SAS Institute Inc, Cary, North Carolina) or/and Genstat 17 softwares (VSN International, 2011). Descriptive statistics of the different stages of fruit development, fruit location and position were analysed using the summary statistics procedure. Duncan multiple range test was performed for multiple-parameter analysis. An analysis of variance (ANOVA) was performed to test the significance and $P<0.05$ was

considered statistically significant. The interaction effects were performed using the generalised linear model procedure.

## 2.3 Results

### 2.3.1 Carotenoid extraction from oil palm mesocarp

The total carotenoid content from the clonal palm fruits were measured in the form of β- carotene extracted from the oil palm mesocarp tissues. A standard curve for β-carotene using a commercial β-carotene standard was plotted as shown in (Figure 2.2). The β-carotene content from the extracted carotenoids from the fruits were quantified using the standard curve. The β-carotene standard curve equation is $y = 374.69x + c$, where $c = 0$ (linear through zero). Table 2.1 presents the summary of the average β-carotene content at six different stages of fruit development. The six fruit stages represent the green (4 WAA and 8 WAA), matured green (12 WAA and 16 WAA) and ripe (20 WAA and 22 WAA) fruits harvested from clonal oil palm FFBs. At 4 WAA, the overall mean β- carotene content was found to be lowest with 60.2 ppm ($p<0.001$). The β-carotene content increased as the fruits develop, mature and ripen and reached the highest value at the recorded fruit ripening stage of 22 WAA with 532.6 ppm ($p<0.001$). The increases in the β-carotene content were not significant in young green fruits from 4 WAA until 12 WAA. However, at 16 WAA, the mean β-carotene content significantly ($p<0.001$) increased to 135.6 ppm, an amount twice the average content at 4 WAA. From this stage (16 WAA) onwards, the β-carotene content significantly increased ($p<0.001$) and by the

22 WAA, the average mean β-carotene content had increased fourfold above that at 16 WAA.

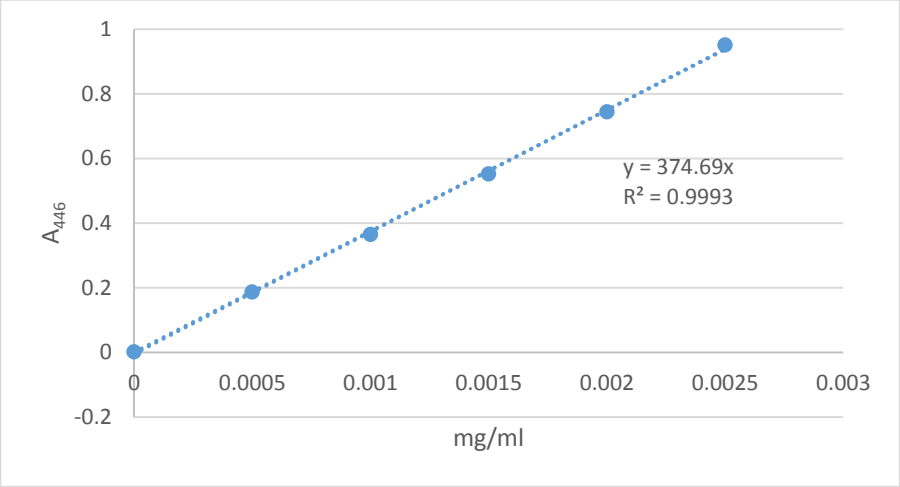



**Figure 2.2: Standard curve for the commercial β-carotene standard.** The β-carotene standard curve equation is y = 374.69x + c, where c = 0 (linear through zero).

**Table 2.1: Summary mean statistics of β-carotene extracted from oil palm fruits at six fruit developmental stages** based on absorbance (A) readings at 446 nm and β-carotene standard curve ($p<0.001$).

| Fruit development | β-carotene content (ppm) | | |
|---|---|---|---|
| Stage (WAA) | Overall Mean | Range of individual values | Max-Min |
| 4 | 60.2 ± 21.0[a] | 21.1 - 132.9 | 111.8 |
| 8 | 79.5 ± 24.7[ab] | 24.3 - 151.6 | 115.1 |
| 12 | 97.6 ± 32.7[ab] | 30.8 - 224.5 | 193.7 |
| 16 | 135.6 ± 59.3[b] | 36.5 - 321.8 | 285.3 |
| 20 | 332.5 ± 230.2[c] | 50.3 - 1084.0 | 1033.7 |
| 22 | 532.6 ± 392.1[d] | 64.1 - 1667.0 | 1602.9 |

* All data on each fruit developmental stages were based on five palms, each containing three fruit locations within the bunch (apical, central and basal) and each location containing three fruit positons within the spikelets (outer, middle and inner); where n = 45.

**Mean values with different superscripts are significantly different ($p<0.001$).

There were significant variations ($p<0.0.01$) in the amount of β-carotene content in all the six fruit development and ripening stages of the clonal palms (Table 2.1). At 4 WAA, the individual values ranged from 21.1 ppm to 132.9 ppm while at 22 WAA the values ranged from 64.1 ppm to 1667.0 ppm. As the fruits develop from young until the ripening stage, the variations become greater. Comparison between the individual young (4WAA) and ripe (22 WAA) fruits also showed a large difference in the β-carotene content that ranged from 43 ppm to 1534.1 ppm. This data is attributed to the different locations and positions of the fruits within each bunch (Figure 2.3).

Figure 2.3 presents an example of a spectral profile of oil palm carotenoid extracted from the mesocarp of clonal palm P161 from the various locations, positions and at different stages of fruit development. Absorption of carotenoids was based on the triple peaks that were found to be within the visible light range of 400 - 500 nm with a maximum absorption at 446 nm. Absorption at 446 nm corresponds to β-carotene which is one of the major carotenes in oil palm. In addition, the absorption of carotenoids was also detected within the UV light range (10 - 380 nm) with maximum absorption at 276 nm. At 4 WAA until 12 WAA, the profile of the absorbance reading at the various locations and positions within the FFBs were not easily distinguishable. At this stage, quantification of the β-carotene content were insignificant (Table 2.1). However, by the 16 WAA, the absorbance profile in each location and position became more segregated.

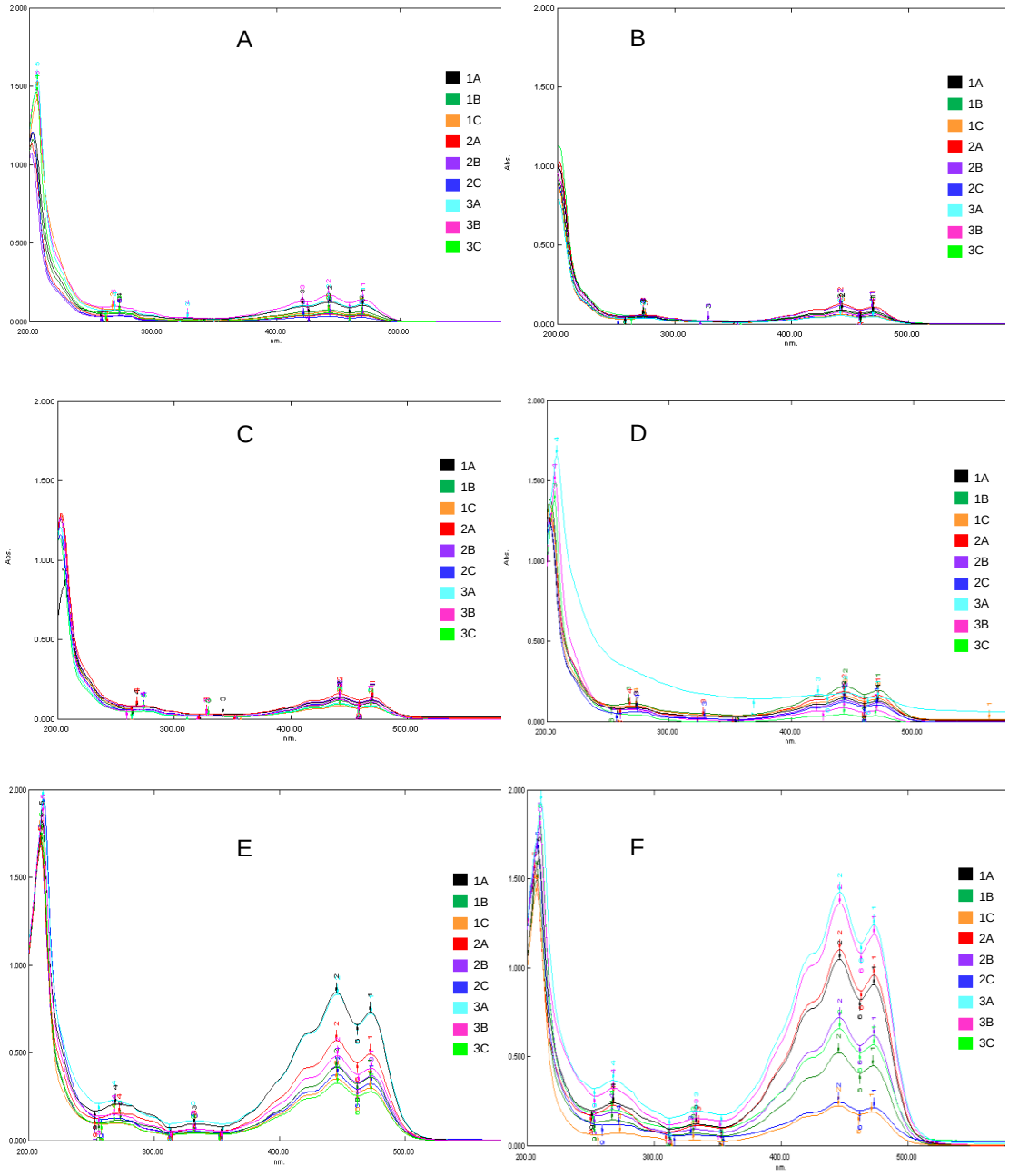



**Figure 2.3: Spectral profile of oil palm carotenoid extracted from the mesocarp of palm P161 at various locations, positions and stages of fruit development**. A - F represent the fruit developmental stages at 4, 8, 12, 16, 20 and 22 WAA, respectively. Coloured lines in each graph represents the location and position of the fruits collected within the FFB where 1 = basal FFB area, 2 = central FFB area, 3 = apical FFB area, A = outer fruits, B= middle fruits and C= inner fruits. The x-axis = absorbance reading from 200 nm to 600 nm and y-axis = absorbance value (Abs).

As the fruits ripens, the segregation and variation in the spectral profile/absorbance pattern of the nine positions within the bunch became more prominent especially at 22 WAA. As a result, changes in the β-carotene content were significant at these stages (Tables 2.1 and 2.2). It was also observed that the top four highest spectral profiles were found at the fruits positioned on the outer spikelets located throughout the whole bunch.

**Table 2.2: Mean statistics of β-carotene extracted from oil palm fruits located within the apical, central and basal area of the FFBs (p<0.001).**

| Fruit development and ripening stages | Fruit location within bunch | | |
|---|---|---|---|
| | Basal | Central | Apical |
| 4 WAA | 68.8 ± 21.6[a] | 50.2 ± 20.6[a] | 61.5 ± 34.1[a] |
| 8 WAA | 73.5 ± 21.8[a] | 85.9 ± 26.2[a] | 79.3 ± 22.9[a] |
| 12 WAA | 93.7 ± 43.6[a] | 101.0 ± 27.1[a] | 98.4 ± 27.0[a] |
| 16 WAA | 127.5 ± 53.4[a] | 144.4 ± 59.3[a] | 135.0 ± 67.1[a] |
| 20 WAA | 347.2 ± 273.5[bc] | 264.6 ± 151.9[b] | 385.8 ± 249.3 [c] |
| 22 WAA | 497.2 ± 355.3 [d] | 506.5 ± 452.4[d] | 593.9 ± 355.3 [d] |

* All data were based on five palms and each fruit location contains three fruit positons within the spikelets (outer, middle and inner)

**Mean values with different superscripts are significantly different (p<0.001).

Table 2.2 presents the comparison between the mean β-carotene content extracted from six stages of oil palm fruit development and the fruit location within the FFB. Variations in the mean β-carotene content were found within the apical, central and basal region of the six FFBs at the different fruit

development and ripe stages. However, ANOVA in general showed that the overall changes within the basal, central and apical locations in all six fruit developmental stages were not significant. Fruits located on the apical recorded the highest overall mean while those in the central region has the least mean β-carotene content. The ripest fruits at 22 WAA has the highest β-carotene content in all three locations within the bunch compared to the rest of the fruits from the different development stages. At 22 WAA, the highest β-carotene content was found to be at the apical area followed by the central and basal regions. At 20 WAA, the pattern of β-carotene accumulation within the bunch was dissimilar to the 22 WAA fruits. The highest β-carotene content was also found to be at the apical but the least β-carotene content was found to be in the central region. For the young and maturing stages, accumulation of the β-carotene within the bunch somehow follows the same trend. Most of the fruits recorded the highest content in the centre whereas the lowest was found at the basal region.

In comparison to the fruit locations within the bunch, interaction analysis between the fruit developmental stages and positions showed significant ($p<0.001$) changes in the β-carotene content. In general, accumulation of β-carotene on the outer, middle and inner fruits within the FFBs were similar in all the fruit developmental stages from young until ripening (Figure 2.4). The outer fruits have the highest content of β-carotene while the inner fruits that are located at the far end of the fruit spikelets recorded the least content of β-carotene. The rise in the β-carotene content was observed especially during

the ripening stages. Within the outer fruits, the lowest β-carotene content was detected at 4 WAA with 77.3 ppm while the highest was found in 22 WAA with 895.5 ppm. There is an almost 12-fold increase in the β-carotene content in the outer fruits as compared to the middle (8 fold) and inner (5 fold) fruits from young until the ripening stages (Figure 2.4). These results conform to visual observations of the fruits where outer fruits are also more coloured than the middle fruits that have a high mixture of white (on the fruit tail) and coloured fruits (on the fruit head) while those at the basal of the spikelets are almost entirely white or of a very pale colour.

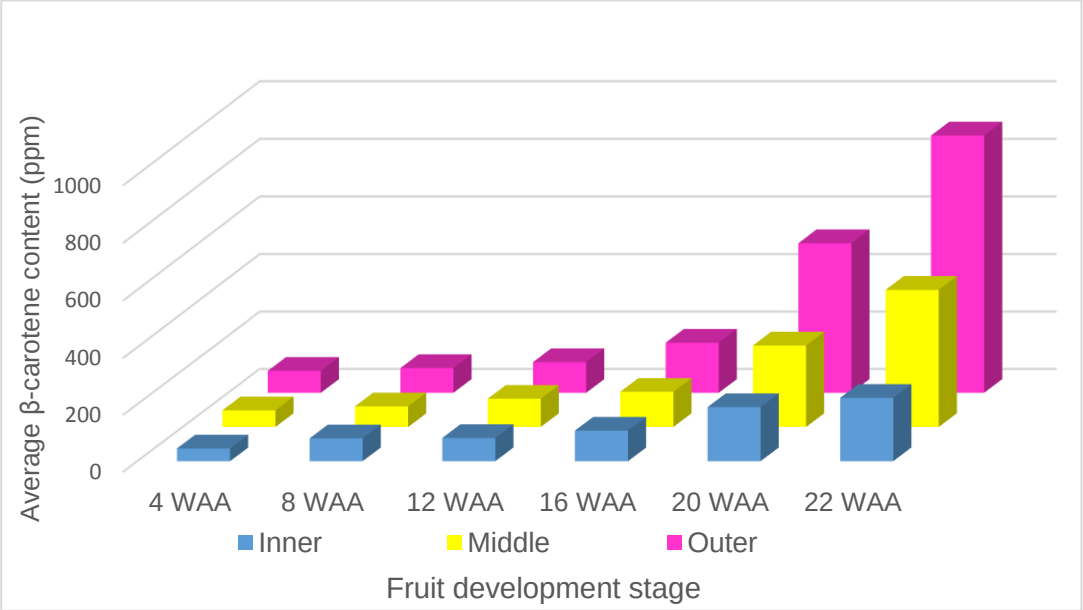



**Figure 2.4: Average content of β-carotene on the outer, middle and inner fruit positions within the bunch at six different stages of oil palm fruit development.**

2.3.2 FAC in oil palm mesocarp at various stages of fruit development

Apart from the six different fruit stages used in the carotenoid analysis, an additional fruit developmental stage was also included in this analysis. The 18 WAA bunch which represents the stage in between the mature green and ripening was included to look at the changes between these two stages. In this work, twelve fatty acids ranging from lauric acid (C12:0) to behenic acid (C22:0) were identified using the commercial standards and the list of fatty acids is shown in Table 2.3.

**Table 2.3: The list of fatty acid and their retention time identified using the FAME mix RM6 commercial standard *via* HP-88 capillary column.**

| Fatty acid | Retention Time (min) |
|---|---|
| Lauric acid, C12:0 | 10.6 |
| Myristic acid, C14:0 | 13.3 |
| Pentadecylic acid, C15:0 | 14.8 |
| Palmitic acid, C16:0 | 16.8 |
| Palmitoleic acid, C16:1 | 17.4 |
| Margaric acid, C17:0 | 18.2 |
| Stearic acid, C18:0 | 20.3 |
| Oleic acid, C18:1 | 21.1 |
| Linoleic acid, C18:1 | 22.3 |
| Linolenic acid, C18:2 | 23.7 |
| Arachidic acid, C20:0 | 23.3 |
| Behenic acid, C22:0 | 26.7 |

Table 2.4 presents the average summary of the FAC in the different fruit developmental stages in clonal palms. In general, significant ($p<0.001$) changes were observed in the FAC within the seven fruit developmental stages of the clonal palms. Four fatty acids were found to be predominant throughout all seven fruit development and ripening stages. The fatty acids are palmitic (C16:0), stearic, (C18:0), oleic (C18:1) and linoleic acid (C18:2), respectively. The rest of the fatty acids were either not detected in this experiment especially in young fruits or found to be at trace levels in particular in the maturing fruits from 16 WAA onwards.

The fatty acids can be grouped into two which consist of saturated (SFA) and unsaturated fatty acids based on the number of double bonds present/absent in their structure. The unsaturated fatty acid can be further classified into monounsaturated (MUFA) which contains a single double bond and polyunsaturated (PUFA) which contains two or more double bonds. The four predominant fatty acids were comprised of both the saturated (C16:0 and C18:0) and unsaturated fatty acids (C18:1 and C18:2). Figure 2.5 presents the overall view of the fatty acid groups during clonal palm fruit development. Comparison of all three groups showed that SFA represented the highest total FA in all the evaluated fruit developmental stages. For MUFA, the total FA increased as the fruits developed and ripens whereas the level of PUFA was found to be in contrast to MUFA.

**Table 2.4: The average fatty acid composition in oil palm mesocarp at various stages of fruit development in clonal palms (p<0.001).**

| WAA* | Fatty acid composition (% area) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | C12:0 | C14:0 | C15:0 | C16:0 | C16:1 | C17:0 | C18:0 | C18:1 | C18:2 | C18:3 | C20:0 | C22:0 |
| 4 | 0.00 ± 0.0 | 0.00 ± 0.0 | 0.00 ± 0.0 | 47.19 ± 9.8 | 0.00 ± 0.0 | 0.00 ± 0.0 | 5.08 ± 1.9 | 14.53 ± 2.1 | 33.19 ± 9.5 | 0.00 ± 0.0 | 0.0 ± 0.0 | 0.00 ± 0.0 |
| 8 | 0.00 ± 0.0 | 0.00 ± 0.0 | 0.00 ± 0.0 | 53.52 ± 10.7 | 0.00 ± 0.0 | 0.00 ± 0.0 | 5.91 ± 2.0 | 19.63 ± 2.9 | 20.94 ± 10.6 | 0.00 ± 0.0 | 0.0 ± 0.0 | 0.00 ± 0.0 |
| 12 | 0.00 ± 0.0 | 0.00 ± 0.0 | 0.00 ± 0.0 | 47.46 ± 6.2 | 0.00 ± 0.0 | 0.00 ± 0.0 | 6.28 ± 1.7 | 24.17 ± 3.4 | 22.08 ± 7.0 | 0.00 ± 0.0 | 0.0 ± 0.0 | 0.00 ± 0.0 |
| 16 | 0.02 ± 0.0 | 0.60 ± 0.4 | 0.04 ± 0.0 | 36.92 ± 4.2 | 0.32 ± 0.3 | 0.33 ± 0.6 | 4.23 ± 1.9 | 40.75 ± 8.6 | 15.85 ± 6.5 | 0.70 ± 0.6 | 0.23 ± 0.2 | 0.02 ± 0.0 |
| 18 | 0.02 ± 0.1 | 0.65 ± 0.2 | 0.03 ± 0.0 | 39.82 ± 3.8 | 0.13 ± 0.1 | 0.42 ± 1.1 | 5.33 ± 0.9 | 41.36 ± 5.4 | 12.01 ± 1.8 | 0.02 ± 0.0 | 0.39 ± 0.0 | 0.07 ± 0.1 |
| 20 | 0.03 ± 0.1 | 0.66 ± 0.5 | 0.01 ± 0.0 | 43.56 ± 2.1 | 0.04 ± 0.1 | 0.08 ± 0.1 | 4.66 ± 0.4 | 40.09 ± 2.3 | 10.68 ± 0.3 | 0.00 ± 0.0 | 0.17 ± 0.2 | 0.00 ± 0.0 |
| 22 | 0.01 ± 0.0 | 0.76 ± 0.3 | 0.04 ± 0.1 | 41.48 ± 2.8 | 0.08 ± 0.1 | 0.18 ± 0.4 | 4.36 ± 0.8 | 42.86 ± 2.7 | 9.11 ± 2.1 | 0.07 ± 0.1 | 0.41 ± 0.4 | 0.62 ± 3.2 |

*WAA = week after anthesis

** All data were based on five palms, each containing three fruit locations within the bunch (apical, central and basal) and each location containing three fruit positons within the spikelets (outer, middle and inner) n=45

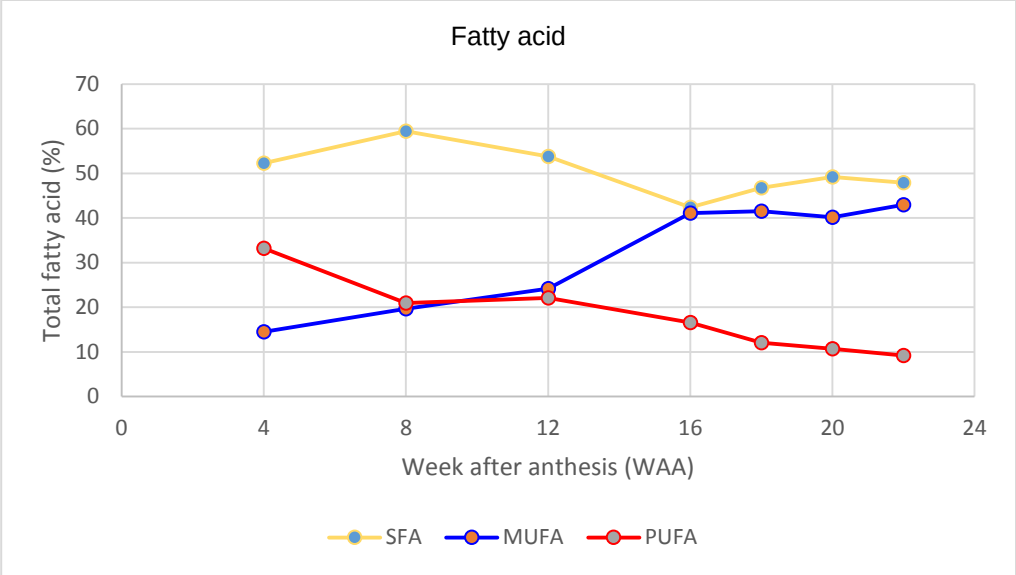



**Figure 2.5: Average changes in saturated fatty acid (SFA), monounsaturated fatty acid (MUFA) and polyunsaturated fatty acid (PUFA) at the different stages of oil palm fruit development.**

At the early stages of the evaluated fruit development especially during the green and mature green fruit developmental stages, the total SFA was slightly over 50% of the total fatty acids. At 4 WAA for example, SFA was found to represent the highest total FA with 52.3%. The high SFA was mainly contributed by palmitic (47.2%) and stearic acids (5.1%). At this stage PUFA content was higher that MUFA with 33.2% as compared to 14.5%. PUFA content was represented by 33.2% of linoleic acid (C18:2) wereas MUFA represented 14.5% of total FA solely from oleic acid. However, as the fruits developed into maturity and ripens from 16 WAA to 22 WAA, the SFA value decreases whereas the total unsaturated fatty acids (MUFA and PUFA) increases at these stages. At 16 WAA, SFA still represented the highest (42.4%) total FA as compared to MUFA (41.1%) and PUFA (16.6%) but the

percentage decreased by 1.2 fold fom 4 WAA. At this stage MUFA content increased almost 2.5X than that of PUFA. At the final evaluated fruit development stage of 22 WAA, MUFA content was at its highest level compared to the other fruit developmental stages while PUFA has the lowest.

Comparison of the fatty acid content showed that the major changes in the fatty acid levels was found to be in linoleic, oleic and palmitic acids. At the immature green fruit developmental stages especially at 4 WAA, linoleic acid was at its highest level and decreased significantly as the fruits ripens. In contrast, the level of oleic acid was the lowest at this stage and increased significantly as the fruits ripens. The palmitic acid level was also high in young fruits and dropped by 1.2 fold during the mature green stages at 16 and 18 WAA. At this stage, the level of oleic acid increased from the green stages and was 1.1 fold higher than that of palmitic acid. By the 22 WAA, the level of palmitic acid increased to 41.48% but the oleic acid composition remained the highest and is higher than that of palmitic acid by 1 fold.

There were also a large variation amongst the individual fatty acid profiles in all seven stages of fruit development. This is likely attributed by the various location and position of the fruits sampled from within the bunch. Table 2.5 presents a summary on the variations for each major fatty acids in individual palms from all seven stages of oil palm fruit development. Palmitic acid ranged from 21.72% - 68.47%, oleic acid ranged from 11.37% - 69.52%, linoleic acid ranged from 1.27% – 47.24% and stearic ranged from 0.28% –11.87%. The rest of the fatty acids were found at trace levels and were only detected in

certain location and position within the bunch and not in all the nine location and positions used in this study. Therefore, from here onwards, the results presented on FAC will focus on the four main fatty acids in the mesocarp, namely palmitic, oleic, linoleic and stearic acids.

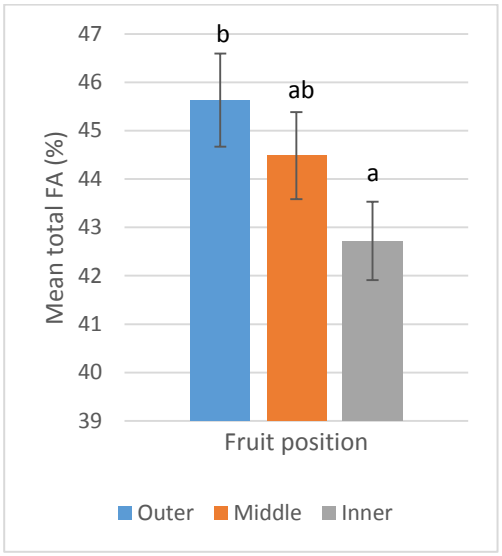
**Table 2.5: Summary statistics of the four main fatty acid profile within oil palm FFB at 7 fruit development stages based on FAME analysis (p<0.05).** Numbers 1, 2 and 3 in the fruit location and position column represents the basal, central and apical region within the bunch whereas alphabets A, B and C represents the outer, middle and inner fruit positions within fruit spikelets.

| Fruit development stage (WAA) | Fatty acid | Fatty acid composition (%) | | |
|---|---|---|---|---|
| | | Overall mean | Range of individual values | Max-Min |
| 4 | C16:0 | 47.19 | 36.59 - 64.96 | 28.37 |
| | C18:0 | 5.080 | 1.175 - 11.07 | 9.895 |
| | C18:1 | 14.53 | 11.37*- 20.76 | 9.392 |
| | C18:2 | 33.19 | 18.34 - 47.24** | 28.90 |
| 8 | C16:0 | 53.52 | 36.98 - 68.47** | 31.49 |
| | C18:0 | 5.909 | 2.530 - 11.87** | 9.337 |
| | C18:1 | 19.63 | 10.55 - 26.36 | 15.81 |
| | C18:2 | 28.62 | 8.816 - 37.44 | 33.59 |
| 12 | C16:0 | 47.46 | 39.07 - 66.66 | 27.59 |
| | C18:0 | 6.284 | 2.608 - 9.312 | 6.704 |
| | C18:1 | 24.17 | 19.38 - 34.96 | 15.58 |
| | C18:2 | 22.08 | 7.142 - 32.59 | 25.45 |
| 16 | C16:0 | 36.92 | 26.48 - 47.59 | 21.11 |
| | C18:0 | 4.232 | 0.277* - 10.4 | 10.16 |
| | C18:1 | 40.75 | 23.86 - 51.40 | 27.54 |
| | C18:2 | 15.85 | 7.767 - 34.01 | 26.24 |
| 18 | C16:0 | 39.82 | 21.73* - 48.09 | 26.36 |
| | C18:0 | 5.331 | 2.613 - 7.791 | 5.178 |
| | C18:1 | 41.36 | 33.18 - 69.52** | 36.35 |
| | C18:2 | 12.01 | 5.877 - 16.32 | 10.45 |
| 20 | C16:0 | 43.56 | 38.99 - 47.87 | 8.882 |
| | C18:0 | 4.659 | 3.460 - 5.981 | 2.521 |
| | C18:1 | 40.09 | 35.77 - 46.98 | 11.22 |
| | C18:2 | 9.797 | 1.272* - 14.72 | 13.45 |
| 22 | C16:0 | 41.48 | 32.49 - 48.33 | 15.84 |
| | C18:0 | 4.359 | 2.382 - 5.889 | 3.507 |
| | C18:1 | 42.86 | 32.26 - 48.97 | 16.71 |
| | C18:2 | 9.105 | 6.286 - 18.31 | 12.03 |

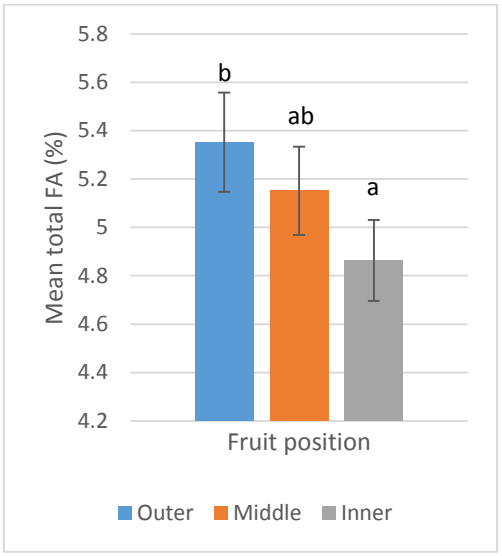
*lowest % of fatty acid, **highest % of fatty acid

2.3.2.1 Effect of fruit location (apical, central and basal) and position (outer, middle and inner) within the bunch

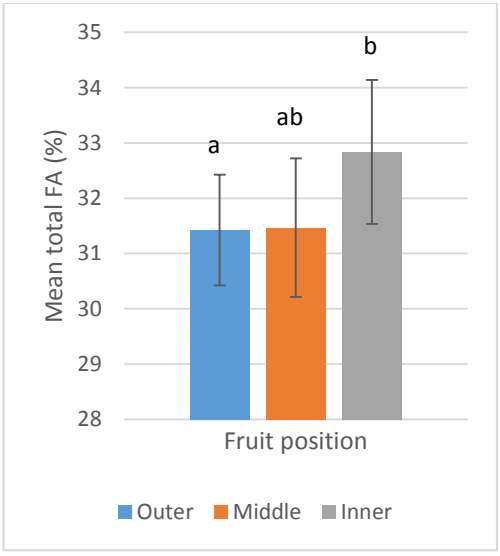
An interaction analysis between fruit developmental stages, fruit locations and positions within the bunch was performed to look at their effect on the fatty acid composition. Apart from the significant ($p<0.001$) changes between the different stages of fruit development and their FAC, a significant ($p<0.001$) interactions was also observed between the FAC and the fruits positioned on the outer, middle and inner fruits within the spikelets. Figure 2.6 represents the average total content of palmitic, stearic, oleic and linoleic acids from all seven fruit developmental stages in the outer, middle and inner fruits within the fruit bunch. For palmitic ($p<0.05$), stearic ($p<0.10$) and oleic ($p<0.05$) acids, the total mean fatty acids were significantly different in the fruits positioned within the fruit spikelets. Both saturated fatty acids consisting of palmitic and stearic acids contains the least content in the inner fruits. Both palmitic and stearic acid contents increased towards the middle fruits and is the highest in the outer fruits. The opposite pattern of fatty acids content was observed in the unsaturated fatty acids (oleic and linoleic acid). Both oleic and linoleic acids content are the highest in the inner fruits and decreased towards the middle and outer fruits where the outer fruits had the least content of both fatty acids. This pattern of FAC accumulation was demonstrated especially in the young fruits from the green to the mature green stages. As for the fruit location within the bunch, the pattern of FAC accumulation was found to be similar but not significantly different within the apical, central and basal region.



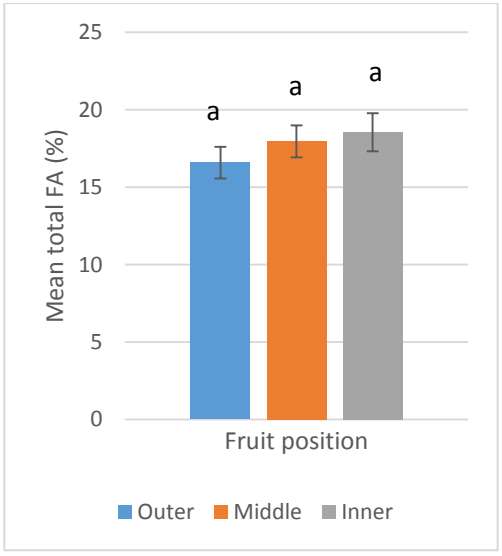

Palmitic acid (C16:0)




Stearic acid (C18:0)




Oleic acid (C18:1)




Linoleic acid (C18:2)

**Figure 2.6: Average total FAC of the 4 prominent fatty acids in oil palm fruits in all the developmental stages at the outer, middle and inner region of the fruit spikelets within the bunch.**

2.3.3 Iodine value (IV)

Table 2.6 represents the mean IV of the fruit developmental stages used in this study. In general, ANOVA showed highly significant ($p<0.001$) differences in the IV at the various stages of fruit development. The highest IV (68.7%) was found at 4 WAA, the youngest evaluated fruit developmental stage while the lowest (52.88%) was found at the ripest evaluated stage at 22 WAA. Interaction analysis was also performed to look at the effect of IV x fruit location x position. However these interactions were insignificant except for IV x fruits position. There were significant ($p<0.05$) differences in the mean IV and the fruits positioned within the inner, middle and outer spikelets in all the fruit developmental stages. The highest IV was found to be in the inner fruits followed by the middle and outer positioned fruits (Table 2.7).

**Table 2.6: The average IV in oil palm mesocarp at various stages of fruit development ($p<0.001$).**

| Developmental stage (WAA) | IV (%) |
|---|---|
| 4 | $68.65 \pm 17.8^{c}$ |
| 8 | $53.15 \pm 17.4^{a}$ |
| 12 | $59.04 \pm 11.0^{b}$ |
| 16 | $64.63 \pm 7.1^{c}$ |
| 18 | $56.55 \pm 2.9^{ab}$ |
| 20 | $53.02 \pm 2.4^{a}$ |
| 22 | $52.88 \pm 3.5^{a}$ |

**Table 2.7: IV of clonal palm fruits at various developmental stages based on their inner, middle and outer positions within the bunch fruit spikelets (p<0.001).**

| Developmental week (WAA) | Fruit position | | |
|---|---|---|---|
| | Outer | Middle | Inner |
| 4 | 64.83 ± 17.1 | 69.96 ± 15.6 | 71.17 ± 20.2 |
| 8 | 48.02 ± 17.6 | 54.35 ± 18.2 | 57.09 ± 16.2 |
| 12 | 53.06 ± 8.8 | 58.90 ± 11.0 | 65.15 ± 10.1 |
| 16 | 62.03 ± 5.6 | 65.66 ± 7.5 | 66.20 ± 7.7 |
| 18 | 56.51 ± 4.6 | 55.64 ± 0.6 | 57.50 ± 1.9 |
| 20 | 53.63 ± 2.0 | 52.85 ± 1.9 | 52.59 ± 3.2 |
| 22 | 52.88 ± 6.0 | 52.44 ± 1.9 | 53.32 ± 1.0 |

### 2.3.4 Ethylene production in oil palm fruits at various stages of fruit development

Table 2.8 presents the average changes in the ethylene production in the clonal oil palm fruits at different stages of fruit development. Overall, there was a highly significant difference ($p<0.001$) in the ethylene production at the various fruit developmental stages. In the young fruits, especially at the green and mature green stages, ethylene production was very low where it was almost at the basal level. At these stages, the changes in the ethylene production were insignificant. As the transition from the mature green to the ripening stage takes place, the mature fruit starts to produce more ethylene. At $20^{th}$ WAA, the increase in the ethylene level was almost twice than the amount produced at 16 WAA. Ethylene levels continued to rise significantly ($p<0.001$)

and recorded the maximum level at 22 WAA which represents the ripest stage used in this study.

**Table 2.8: The average ethylene production from the oil palm fruits at 6 fruit developmental stages (p<0.001).**

| Fruit development week (WAA) | Overall Mean (µl/kg/h) |
|---|---|
| 4 | 0.03 ± 0.01[a] |
| 8 | 0.03 ± 0.01[a] |
| 12 | 0.03 ± 0.01[a] |
| 16 | 0.03 ± 0.01[a] |
| 20 | 0.05 ± 0.04[ab] |
| 22 | 0.12 ± 0.18[b] |

* All data on each fruit developmental stages were based on five palms, each containing three fruit locations within the bunch (apical, central and basal) and each location containing three fruit positons within the spikelets (outer, middle and inner).

**Mean values with different superscripts are significantly different (p<0.001).

Within the 22 WAA bunch, variation between the ethylene produced by the individual fruits ranged from 0.01 ng/kg/h to 0.77 ng/kg/h. However, interaction analysis between the ethylene produced from the various fruits located within the apical, central and basal region as well as those positioned on the outer, middle and inner fruit spikelets of the bunch were not significantly different. Nevertheless, there is a similar trend of ethylene production from within the bunch and fruit spikelets. From the basal to the apical region, the ethylene production increased where the fruits within the apical region recorded

the highest level of ethylene produced at 0.11 nl/kg/h while the basal recorded the least with 0.08 nl/kg/h (Figure 2.7).

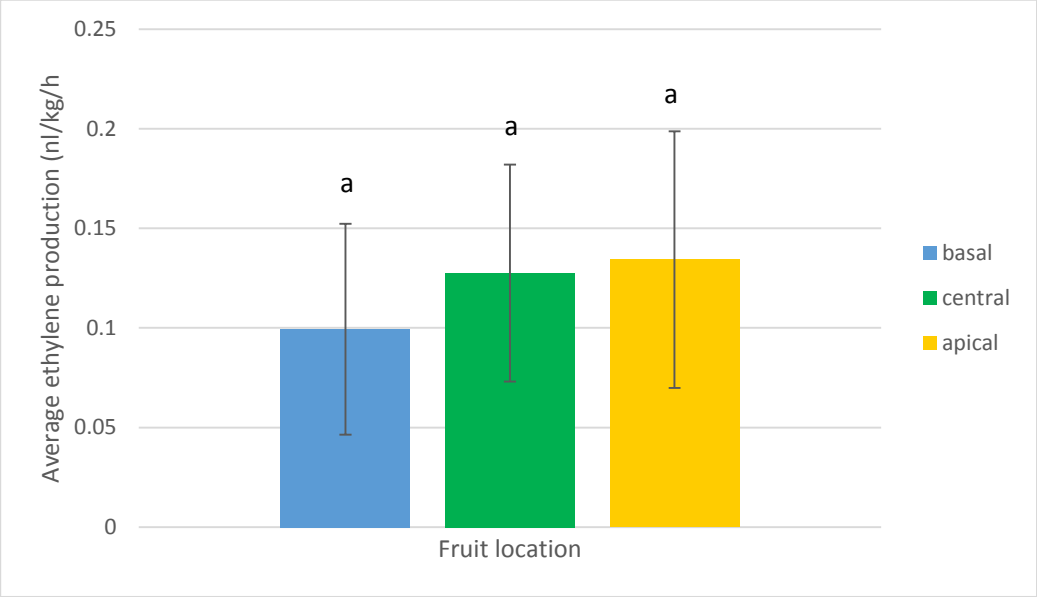



**Figure 2.7: Total ethylene production in oil palm fruits at 22 WAA at the basal, central and apical region of the fruit bunch.**

The ethylene production was also found to increase from the outer to the middle and inner positioned fruits within the spikelets (Figure 2.8). The inner fruits produced ethylene twice the amount of the outer fruits. Comparison of the average ethylene production from the fruit from the various location and positions within the 22 WAA bunch showed that the highest ethylene production in the inner fruits was detected in the fruits spikelets located within the apical bunch. However, in the central region, the highest ethylene level was detected in the outer fruits whereas the production of ethylene in the basal region is similar to the apical region.

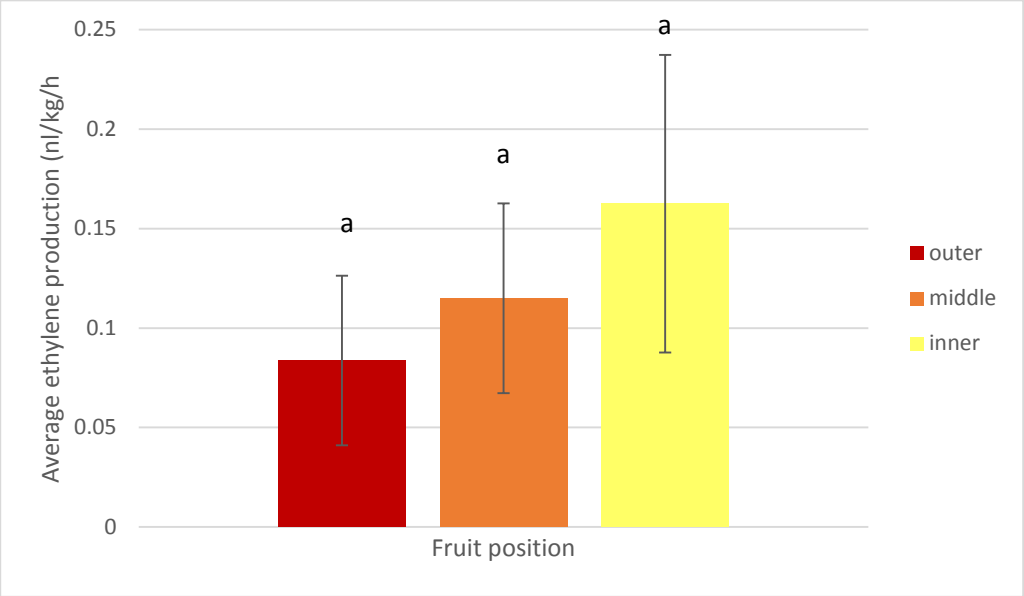



**Figure 2.8: Total ethylene production in oil palm fruits at 22 WAA at the outer, middle and inner region of the fruit spikelets within the bunch.**

## 2.4 Discussion

### 2.4.1 Accumulation of β-carotene, FAC and ethylene production in clonal palms

Clonal palms represent 0.9% of the annual world planting material (Kushairi et al. 2010). In Malaysia, clonal palm are produced about 3.5 million ramets or clonal plantlets per year (Soh et al. 2011). Clonal palms are produced through *in vitro* propagation of elite palm. Hardon et al. (1987) estimated that the yield of clonal palm can increase up to 30% if the clones are derived from best individual palms. With the improved standards of field agronomy, clonal oil palm also offers the potential for greater productivity. In addition, clonal palms allows the establishment of uniform tree stands comprising identical copies (clones) of a limited number of highly productive oil palms (Mutert and Fairhurst, 1999).

In this study, clone P164 was selected based on their best yield performance in Bungor Soil of Ulu Paka, Terengganu, Malaysia (Fadila et al. 2012). Biochemical and physiological analysis were performed on various fruit developmental stages comprising of the young and ripe fruits. This in particular was focused on the random fruit sampling approach of the various fruit locations (basal, central and proximal region) and positions (outer, middle and inner fruit) within the bunch to understand the changes that happen within the fruit bunch as it develops into maturity and ripens. Among the changes investigated were carotenoid or β-carotene content, FAC and ethylene production during the development and ripening process in clonal palm fruits.

Carotenoids in the oil palm fruits (90%) are mainly made up of α- and β-carotenes (Ikemefuna and Adamson, 1983; Tay and Choo, 2000). Absorption of carotenoids extracted from the oil palm mesocarp was observed with the presence of three peaks (Kaur et al. 2008). In the clonal fruits, changes in the average total content of β-carotene during fruit development is similar to the commercial *tenera* planting materials (Kaur et al. 2002). The level of the β-carotene is the lowest in young fruits, and reached the maximum in the ripe fruit (Ikemefuna and Adamson 1984; Azis 1984; Kaur et al. 2002). In the present work, accumulation of carotenoid increases as the fruits mature and ripens from 16 WAA onwards, concurrent to the accumulation of oil in the mesocarp of the oil palm fruits (Sambanthamurthi et al, 2000; Tranbarger et al. 2011). At 22 WAA, the average β-carotene content at the ripest stage is also equivalent to that of the commercial *tenera* palm fruits which contain about 500-700 ppm of carotenoids (Choo et al. 1997, Corley and Thinker, 2003, Kaur et al. 2008). The high level of carotenoids in ripe fruits results in the bright-orange colour in the CPO (Jalani and Rajanaidu 2000). It is also a precursor for pro-vitamin A, a dietary supplement for human consumption (Sambanthamurthi et al. 2000).

In the clonal ripe fruits at 20 and 22 WAA, changes in the FAC showed a similar trend to those described by Tan et al. (2000). Most of the fatty acids fall within the range of FAC observed in the typical Malaysian palm oil (Tan et al. 2000). In the commercial palms, palmitic acid has always been the most dominant fatty acid followed by oleic acid although variations within both fatty

acid content have been reported among *E. guineensis* populations (Tan et al. 2000; Noh et al. 2002).

In the present study, the highest oleic acid accumulation compared to the rest of the fatty acids in the ripe fruits could be attributed to the high oleoyl ACP and low palmitoyl ACP thoiesterase activities. Biochemical studies on CPO have shown that palmitic acid accumulates in the oil through the action of a palmitoyl ACP thioesterase, a *FatB* type thioesterase that cleaved palmitoyl ACP to release palmitic acid (Sambanthamurthi et al. 1990). In addition, a *FatA* type thioesterase encoding oleoyl ACP thioesterase is also present in oil palm where it cleaves oleoyl-ACP to release oleic acid (Abrizah and Sambanthamurthi, 1995). Both *FatA* and *FatB* type (Abrizah et al. 1999) thoiesterase genes have been identified in oil palm and have sequence diversity (Abrizah, pers. comm). Partial purification of oil palm thioesterase also showed that both palmitoyl-ACP and oleoyl-ACP thioesterase activities reside on different/separate enzymes (Abrizah et al. 1992). This finding showed that both enzyme activities are independent of each other, therefore an increase in oleic acid can be achieved at the expense of palmitic acid and vice versa. The anti-sense technology is being used to down regulate the palmitoyl ACP thoiesterase gene to reduce the accumulation of palmitic acid in palm oil *via* genetically engineering the oil palm for production of high oleic palms (Parveez et al. 2003).

The low activity of thoiesterase towards palmitoyl ACP most likely allowed chain elongation of palmitoyl-ACP to steoryl-ACP by the action of β-

ketoacyl ACP synthase II (*KASII*) and then desaturated/converted to oleoyl-ACP by the action of an active steoryl ACP desaturase (*SAD*). In addition, the KAS II activity is also likely high to allow the increase in the oleic acid content. A negative correlation between the level of palmitic acid and KAS II activity has been demonstrated by Sambanthamurthi et al (1999) which indicated that by increasing the KAS II activity, a reduction in the palmitic acid accumulation can be achieved in the oil palm mesocarp and vice versa. Figure 2.9 represent the pathway for fatty acid biosynthesis in plants.

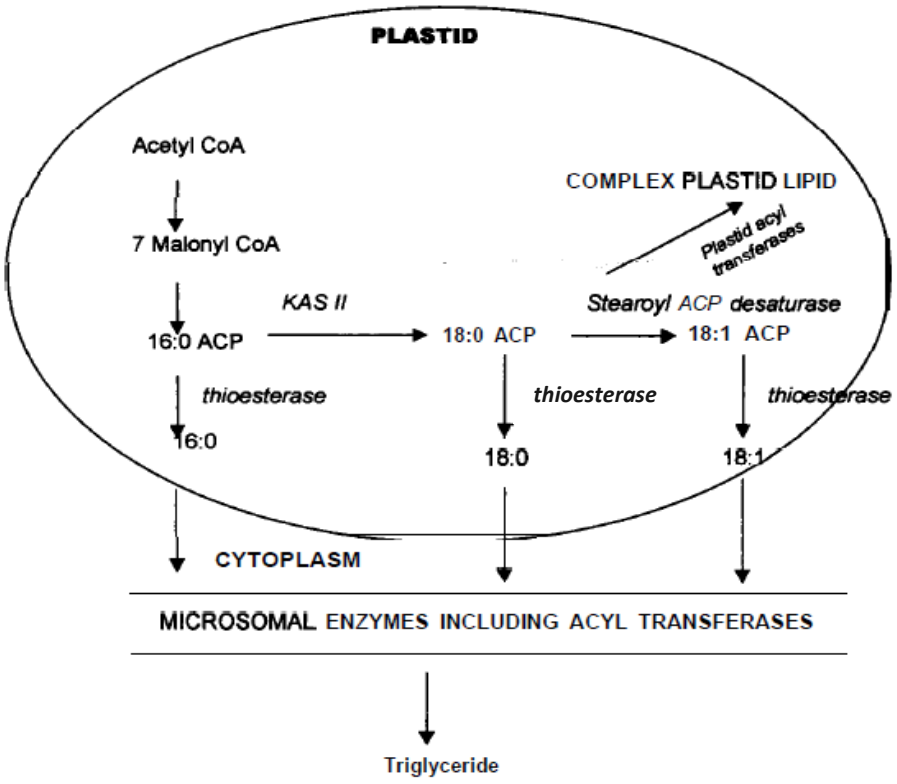



**Figure 2.9: Fatty acid biosynthesis in oil palm** (Adapted from Sambanthamurthi et al. 1999).

Environmental factors such as heat accumulation (Di Vaio et al. 2013), temperature, rainfall (Maffei, 1988) and water logging (Yasumoto et al. 2011) have also been shown to affect oil composition. In olive for example, the different heat accumulation in relation to environment changes influenced the drupe development, fruit colour as well as oil with higher oleic and stearic acid contents (Di Vaio et al. 2013). Lehrian et al. (1980) reported more saturated fats are produced in cocoa at high temperatures. In Malaysia, due to the favourable weather condition for growing oil palm, there is a lack of evidence to show that changes in the environment affect oil composition. However, throughout the duration of the study, the occurrence of El Nino and La Nina in Malaysia during sampling may affect the biochemical changes within the palm including the fruit within the bunches to adapt to the current environment.

In the young and ripe fruits, the IV represents the unsaturation level of the FAC. It is contributed by the unsaturated fatty acids content of oleic, linoleic, palmitoleic and linolenic acids (PORIM 1995). Changes in the IV of the clonal fruits showed that as the fruits developed from green to mature green and finally ripens, the unsaturation level decreases. The high IV in young fruits observed in this study is similar to the report shown by Bafor and Osagie (1986). The high IV content in young fruits is contributed by PUFA as they act as membrane and chloroplast lipids (Oo et al. 1986). On the other hand, Crombie and Hardman (1958) did not observed any changes in the IV in young, mature and ripe fruits. The IV in the ripe fruits is also comparable to that observed in the Malaysian palm oil (Tan and Oh, 1981).

Multiple hormones are known to be involved in the development and ripening process in fruits (McAtee et al. 2013). Ethylene for example are widely known to play many roles in plant growth and development process including ripening and abscission (Abeles et al. 1992). In oil palm, ethylene has been suggested to play a role in regulating the ripening process but at the point when the study started, the role of ethylene in the ripening process of oil palm fruits was unclear. Thus, the present study was set to understand the role of ethylene in regulating the oil palm fruit ripening process based on their production throughout the fruit developmental process from young until the ripening stages. Within the duration of the study, several reports on the ethylene production from oil palm fruits were published. Results from this study supports the findings from other researchers such as Tranbarger et al. (2011) and Cadena et al. (2012) where a rise in the ethylene peak was observed at the ripening stages. The highest level of ethylene production during ripening indicates that oil palm has a characteristic of a climacteric fruit. The level of ethylene measured however, was low in comparison to previous reports. This could also be the result of using a single oil palm fruit to measure ethylene production instead of fruit spikelets as used by Tranbarger et al (2011) and Cadena et al. (2012). In banana, Chillet et al. (2008) reported that very low detection of ethylene concetrations can occur due to the high concentrations of carbon dioxide in the sealed container (Chillet et al. 2008).

2.4.2 The effects of fruit location and position within the bunch

Earlier studies used the average random fruit sampling value to describe the biochemical and physiological changes during the development of oil palm fruits (Noh et al. 2002; Kaur et al. 2002; Tranbarger et al. 2011; Cadena et al. 2012; Prada et al. 2013; Montoya et al. 2014). However, in this work, to further understand the distribution of β-carotene, FAC and ethylene production during the oil palm fruit development within the complex fruit bunch, random sampling of the fruits were performed based on nine individual measurements to represents the outer, middle and inner positioned fruits located within the basal, central and apical region of the bunch.

In general, the fruits positioned within the spikelets in the FFBs affects the accumulation of carotenoid and FAC within the bunch. Fruits positioned on the outer spikelets throughout the bunch has the highest content of β-carotene while the inner fruits have the least. The high carotenoid content gives the deep orange colour in the ripe fruits thus has become the indicator for bunch ripening (Corley and Thinker 2003; 2016). Within the bunch, the accumulation of β-carotene decreases from the outer toward the middle and inner positioned fruits. The pattern of β-carotene accumulation was observed throughout the clonal palm fruit development from young until the ripening stage. The compact structure of the fruit bunch also caused the inner fruits on the spikelets to have the least exposure of sunlight as compare to the middle and outer fruits.

In clustered-type fruits like grapes, variation in the carotenoid content was observed in a grape cluster exposed to the sun compared to the shaded

grape bunches (Razungles et al. 1998). In addition, the compact grape vine canopies has also been shown to affect the light and temperature conditions of the grape bunch zone, which affects carotenoid synthesis in the fruits (Kamfer, 2009). This finding suggests that light is likely playing an important role in affecting and regulating carotenoid biosynthesis in the inner, middle and outer oil palm fruits positioned within the fruit bunch.

Light signalling has been shown to be involved in regulating carotenoid accumulation (Giovanonni 2004; Yuan et al. 2015). In tomato, functional analysis of two light signalling genes, *HY5* and *COP1LIKE* showed that suppressing the expression of *HY5* resulted in the decreased in carotenoid accumulation whereas downregulating *COP1LIKE* resulted in the increased in carotenoid levels, which indicates that both genes play positive and negative roles in controlling the fruit pigment accumulation (Liu et al. 2004). In addition, light also regulates the phytoene synthase (*PSY*) gene, one of the upstream genes in the carotenoid biosynthesis pathway to modulate carotenoid biosynthesis during deetiolation in Arabidopsis seedlings (Toledo-Ortiz et al. 2010).

Light exposure has also been shown to induce the expression of carotenogenic genes in vegetative (leaves) and reproductive tissues (flower and fruits). During these processes, photoreceptors are activated and translocated to the nucleus, leading to the induction of carotenogenic gene transcription (Pizarro and Stange 2009). Light intensities also effects the carotenoid levels where a five fold-increase in the ratio of lycopene β-cyclase

(*lcyb*) and lycopene ε-cyclase (*lcye*) gene expression was observed in Arabidopsis and tomato plants grown in low and strong light (Hirschberg 2001).

In oil palm, several carotenogenic genes such as *psy, lcyb* and *lcye* have been isolated and characterised (Rasid et al. 2007; 2008; 2009). These genes are targeted to modify the carotenoid content in oil palm to increase the lycopene content in the mesocarp *via* genetic modification (Parveez et al. 2003). The position of the fruits within the bunch based on the various β-carotene content can possibly be used as targets to formulate a strategy to increase carotenoid especially the lycopene content. The use of a mesocarp-specific promoter and light-positive and light-negative regulators (Liu et al. 2004; Davuluri et al. 2005) will possibly allow the production of modified palms with increased carotenoid content in the fruit mesocarp without affecting the carotenoid biosynthesis regulation in non-target plant organs.

Ethylene has also been shown to play an important role in carotenoid biosynthesis in fruits such as *Carica papaya* or papaya (Barreto et al. 2011), *Prunus armeniaca* or apricot (Marty et al. 2005) and tomato (Su et al. 2015) among others. The high ethylene production in climacteric fruits such as tomato and papaya at the onset of ripening is highly correlated with the rapid skin colour change and accumulation of β-carotene and lycopene (Lelièvre et al. 1997; da Silva et al. 2005). Hormone treatment studies showed that application of ethylene or ethylene precursor such as aminocyclopropane carboxylic acid (ACC) to climacteric fruits at the green stages promotes/accelerated the ripening process with the increase in total carotenoids (Barreto et al. 2011; Su

et al. 2015) In contrast, application of ethylene inhibitors such as 1-methylcyclopropene (1-MCP) suppresses ethylene responses which includes senescence and ripening in both climacteric and non-climacteric fruits (Huber 2008).

Differential carotenoid gene expression in ripe and green fruits have been widely documented (Hirschberg 2001). In tomato, gene expression analyses on carotenoid biosynthesis pathway genes showed that the expression of *PSY* and β-carotene hydroxylase-1 increased four and thirty-fold higher from breaker to red stage of tomato fruit ripening (Smita et al. 2013). The changes in the expression level of these genes were associated with a 40% increase in lycopene content at red-ripe stage as compared with breaker stage suggesting the role of specific carotenoid pathway-related genes in accumulation of high lycopene during the fruit ripening processes (Smita et al. 2013). *PSY* is also strongly induced by ethylene during ripening indicating an important role in controlling carotenoid accumulation (Lois et al. 2000). The formation of phytoene is known to be the central step to the whole carotenoid pathway with the expression of *PSY* gene being controlled by various factors including light, temperature, as well as drought (Bramley 2002; Hirchsberg 2001). Comparison of the expression of the *PSY1* gene in wild-type tomato fruits and ethylene-suppressed transgenic tomato showed that expression of *PSY1* gene was lower in ethylene-suppressed transgenic tomato plants where a decreased in carotenoid accumulation was also observed (Picton et al. 1993). The role of ethylene in carotenoids formation was further demonstrated by the

phenotype of the *Nr* mutant that showed a decrease in ethylene sensitivity and accumulates low levels of lycopene and β-carotene in ripened tomato fruits (Lanahan et al. 1994).

In relation to other hormones, ethylene has been shown to work antagonistically with auxin to modulate carotenoid accumulation during tomato fruit ripening (Su et al. 2015). Auxin treatment significantly reduces lycopene accumulation in tomato while ethylene greatly induced lycopene and β-carotene. At the transcriptional level, the lower transcript level of several upstream carotenoid genes such as *Psy, Ziso, Pds and Crtiso* matched to the slower fruit colour change and decreased in the lycopene accumulation after auxin treatment. On the other hand, ethylene induces lycopene and α-, β- and δ-carotene accumulation by inducing the transcript expression of *PSY1* gene which is also in agreement with the rapid fruit colour change and increased lycopene accumulation (Su et al. 2015). Application of auxin to tomato fruits also showed an increase in the neoxanthin and violaxanthin levels, a precursor for ABA production (Taylor et al. 2000).

ABA has also been demonstrated to control ripening in many fruits where it triggers ethylene biosynthesis to speed up the ripening process (Zhang et al. 2009). Tomato mutants with ABA deficiency as a result of low violaxanthin and neoxanthin levels has been shown to affect carotenoid content. For example, in the high pigment 3 (*hp3)* tomato mutant, the mutation in *hp3* occurred in the gene for zeaxanthin epoxidase (*Zep*), which converts zeaxanthin to violaxanthin resulting in an increased number of plastid

compartment, hence allowing more storage capacity for higher lycopene content during ripening (Galpaz et al. 2008).

Besides affecting the various content of carotenoid, the position of the outer, middle and inner oil palm fruits within the spikelets located throughout the bunch may also affect the accumulation of certain type of fatty acids. The complex nature of the fruit position within the bunch may play a role in such a way to protect the naturally occurring fatty acids in the fruits. Saturated fatty acids such as palmitic and stearic acid form straight chain carbon-hydrogen single chemical bonds, hence are chemically stable. Unlike unsaturated fatty acids such as oleic and linoleic acids, the chemical bond is less stable due to the presence of double bond(s) in the carbon and missing the hydrogen bond (FAQs-MPOB, 2016).

In vegetable oils, exposure to sunlight causes oxidation where oxygen reacts with the double bond in the fatty acid to form peroxides and/or free radicals. This reduces the quality of the oil hence affecting its value (Przybylski et al. 2005). Because of the chemical structure of the fatty acids, vegetable oil that contain polyunsaturated fatty acids (PUFA) are the most unstable and prone to oil rancidity. Analysis on several vegetable oils exposed to light showed that the relative density, viscosity, peroxide, refractive index and free fatty acids values were significantly increased especially in soybean oil (contained the highest PUFA) and the least in palm kernel oil in comparison to the unexposed oil samples (Fekarurhobo and Obomanu, 2009).

Due to the complex nature of the oil palm fruits within the bunch where thousand of fruits are compact together, the inner fruits positioned within the fruit spikelets/bunch have the least exposure to sunlight in comparison to the middle and outer positioned fruits. This possibly explains why the inner fruits contained the highest content of unsaturated fatty acids. The outer fruits on the other hand, are more exposed to sunlight, therefore accumulates palmitic and stearic acid, both of which are more stable in nature as compared to the unsaturated fatty acids. The findings on the fruits position either within the outer and middle region can possibly be manipulated to maximize the recovery of the targeted fatty acids from the fruit bunch. For example, segregating the inner fruits for higher unsaturated fatty acid especially oleic acids at the mills may not be practical but can possibly contribute towards the current global need for high-oleic oils that are less prone to oxidation (Montoya et al. 2014).

Unlike other fruits with high/rich FA content in the mesocarp like avocado, studies on the role of ethylene on fatty acid biosynthesis or vice versa is scarce in oil palm fruits. This is probably due to the limited studies on the effect of hormones especially ethylene and their inhibitors to the fruits on the changes at the biochemical and molecular level especially during fruit development, maturation and ripening. The pyrosequencing-based transcriptome approach for oil palm mesocarp fruit development demonstrated that key genes involved both ethylene transcriptional activity and fatty acid synthesis are differentially expressed and play important roles during fruit maturation and ripening (Tranbarger et al. 2011). Genes in each pathway are

co-regulated suggesting that changes in the biochemical and physiological processes are likely coordinated during the maturation and ripening phased in the mesocarp (Tranbarger et al. 2011).

In the present study, accumulation of both carotenoid and FAC were not affected by the location of the fruits within the apical, central and basal regions of the bunch. Several studies on the location of fruits within the oil palm fruit bunches also showed similar results. A study on the changes in peel colour of ethylene treated oil palm fruits were not affected within the apical, central and basal region of the bunch (Nualwijit and Lerslerwong, 2014). Another study conducted by Keshvadi et al. (2012) on the effect of moisture content within the oil palm fruit bunch during development and ripening also showed insignificant changes within the fruits located on the apical, central and basal regions. In banana for example, the hand position within the bunch did not have a large influence on variability during ripening or storage (Marin et al. 1996). Thus, the findings from the present study suggest that the changes that occur within the fruits located on the apical, central and basal bunch is similar throughout the bunch but is affected by the fruit position whether on the outer, middle or inner spikelets within the bunch.

The production of ethylene during the fruit ripening stages was also not significantly affected by the outer, mddle and innter fruit positions within the apical, central and basal region of the bunch. However, the trend in the ethylene production from the fruits located within the apical, central and basal region of the bunch is similar to those described by Henderson and Osborne,

1994). Analysis on ethylene production from different parts of a ripe oil palm fruit showed that the fruit apex produces the highest level of ethylene (Henderson and Osborne, 1994). The high ethylene level at the apical and fruit peak suggests that ethylene synthesis is likely to begin from the apical part of the ripe fruit and progresses downwards towards the fruit base. This collectively suggest that the ripening process in oil palm fruit bunch starts from the apical region and moves down towards the basal (Corley and Thinker, 2003; 2016). However, in this study, it is not clearly understood why the trend in ethylene production based on the fruit position (outer, middle and inner) showed that the inner fruits are producing more ethylene as compared to the outer fruits especially at the apical region of the bunch. In contrast to reports by Corley and Thinker (2003; 2016), Nualwijit and Lerslerwong (2014) reported that the ripening degree progresses from the basal to the central and finally to the top region within the fruit bunch. Whether the ripening signal is triggered from the inner fruits to the outer fruits at the apical region and then moves down to the basal of the bunch still remains to be understood.

Nevertheless, findings from this study clearly demonstrated that the fruit development and ripening within the oil palm fruit bunch is unsynchronized based on the variability of carotenoid accumulation and FAC in the fruit positioned on the outer, middle and inner bunch. The effect of other hormones as well as nutrients uptake may also affect the variability of these contents in the development of oil palm fruits. However, since there is not much study

focused in these areas, this causes difficulties and results in slow progress in understanding the unsynchronization of oil palm ripening within the bunch.

Grape berry is an example of a fruit that also has an uneven or asynchronized ripening pattern. Grape berries are developed within fruit clusters and all berries within a cluster do not ripen at the same rate. At the onset of ripening (veraison stage) some berries within a cluster are more developed than others but upon maturity the fruits will ripe uniformly (Gouthu et al. 2014). Metabolic studies showed that synchronization of berry ripening at the later stages of development towards maturity is dependent on the genotype, environmental conditions, and internal hormone dynamics (Böttcher et al. 2010; Pagay and Cheng 2010; Dai et al. 2011). These findings suggest that ripening programme is flexible and can be modulated. Hence, information from grape berry studies can possibly be used as a model to synchronize ripening in oil palm fruits within the bunch, which in time may contribute to the increase in oil palm yield.

Taken together the results of the biochemical and physiological changes involving FAC, carotenoid and ethylene production within the bunch have shown that the fruit developmental process within the bunch is a very complex process. Thus, to further understand the role of ethylene in the oil palm fruit developmental processes and its possible manipulation for genetic modification at the molecular level, the next chapter describes the use of oil palm genome to identify and characterise ethylene related genes.

# CHAPTER 3

# MINING THE ETHYLENE RECEPTOR FAMILY FROM THE OIL PALM GENOME

## 3.1 Introduction

Oil palm fruit is climacteric where the fruit ripening process is accompanied by an autocatalytic burst of ethylene. To further understand the role of this hormone, at the molecular level, in regulating the development, maturation and ripening of oil palm fruits, this chapter focuses on the identification and characterisation of the ethylene receptors genes. Ethylene receptors play a key role in sensing the gas and in triggering the signalling pathway that ultimately leads to an ethylene response. Previously, a full-length ethylene receptor gene designated *EgERD3* was isolated and characterized (Nurniwalis 2006). Southern blot analysis demonstrated that the gene belonged to a multigene family. Thus, with the availability of oil palm genome sequence data, the quest to discover genes that code for important traits as well as those that belong to multigene families can be achieved in a much shorter time as compared to conventional methods. In this chapter, the process of mining and identification of the ethylene receptor genes from the oil palm genome, isolation of the full-length cDNAs and corresponding promoters, prediction of their protein structure, gene mapping and characterisation of their expression pattern in various oil palm tissues including fruit tissues are described.

## 3.2 Materials and methods

### 3.2.1 List of Solutions, Buffers, Media and Reagents

The list and preparation of standard solutions, buffers, media and reagents used in this study are presented in Appendices 1 and 2.

### 3.2.2 Plant Materials

Various tissues from oil palm *Elaeis guineensis* variety *tenera* (progeny of *dura* x *pisifera*) were used in this study. This includes tissues such as mesocarp at various stages of fruit development, kernel, spear leaves, roots and germinated seedlings. The oil palm fruits and spear leaves were obtained from MPOB-UKM Research Station, Bangi, Selangor, Malaysia. The germinated seedlings were supplied by the MPOB Kluang Research Station, Johor. The 2-year old seedling roots from MPOB HQ nursery were kindly provided by Dr. Abrizah Othman.

Oil palm inflorescences were tagged at anthesis and the fruits were collected at different developmental stages from early fruit development until ripening. For RNA extraction, the fruits were separated from the bunch and rinsed with distilled water. After removal of the exocarp, the mesocarp and kernel tissues were collected and immediately stored at -70$^{o}$C until required. The spear leaves and roots were excised from plants, rinsed with distilled water before cutting to the size of 2 cm x 5 cm. For the germinated seedlings, the hypocotyls were removed from the seedlings. All tissues were frozen in liquid nitrogen and stored at -70$^{o}$C prior to usage.

This study also used the base of the oil palm fruits containing the primary and adjacent abscission zone (AZ) (Roongsattham et al. 2012). The AZ represents the place where cell separation at the junction of the fruit base and the pedicel takes place prior to fruit shedding from the bunch (Henderson and Osborne 1994). The AZ was collected from the the base of clonal oil palm fruits (P165) at three locations within the bunch (basal, central and apical) area at two ripening stages (20 WAA and 22 WAA) used in this study. Processing, handling and storage of the tissues is as described earlier.

### 3.2.3 Identification of genes and searches of conserved domains

Homology search and Hidden Markov Model (HMM) were employed to identify the ethylene receptor family members from the oil palm genome sequence data. Name search using ethylene receptor, *ETR, ERS* and *EIN4* as keywords in public databases helped in the identification of ethylene receptor genes from other plants and their nucleotide and/or protein sequences were downloaded from TAIR (http://www.arabidopsis.org), SIGnAL: (http://signal.salk.edu/cgi-bin/RiceGE), SGN (http://solgenomics.net/), NCBI (http://www.ncbi.nlm.nih.gov), supplementary data by Tranbarger et al. (2011), SWISSPROT, TrEMBL (UniProt Knowledgebase: http://www.expaxy.org/sprot/) and date palm draft sequence (http://qatar-weill.cornell.edu/research/datepalmGenome/), respectively. These sequences, especially those from Arabidopsis, rice, tomato and date palm *(Phoenix dactylifera)*, were used to search the oil palm transcriptome and genome

databases (Singh et al. 2013a) based on BLAST algorithm using Biowise V2 system.

Biowise V2 system is an in-house programming script developed by the Bioinfomatics Team at MPOB, Bangi, Malaysia. The programme is based on LINUX operating system and requires three major execution steps. The first execution was carried out to compare the gene sequences to the known sequences in the oil palm databases using the Basic Local Alignment Search Tool (BLAST) programme with the default parameters [processor number (-a) = 8, maximum number of hits output (-v and –b) = 20 and e-value (-e) = 0.00001]. The next execution was performed to annotate the blast report (such as list of hit subjects, fasta sequences of hit subjects, folder containing individual fasta files and folder containing individual Blast report) in the form of tab-delimited format and finally the last step allows the retrieval of the hit region on the oil palm database *via* extraction of the specific region of the sequence in the fasta files.

The HMM-profile was constructed using HMM algorithm to identify other members of the ethylene receptor family in oil palm. Firstly, a multiple sequence alignment of the downloaded ethylene receptor protein sequences was performed using Clustal W from Bioedit and/or Muscle from Mega 5.05. Those with incomplete or redundant sequences were omitted from further analyses. Next, the HMMbuild programme in the HMMER3 package (Eddy 2011) was used to construct the HMM-profile using the multiple sequence alignment information. The HMMbuild profile was then used to search against the Oil

Palm Gene Model using hmmsearch (Eddy 2011). The Oil Palm Gene Model is a gene finding system that was built based on the integration of multiple sources of programmes such as GlimmerHMM, SNAP, JIGSAW, and MAKER to predict the gene structures in oil palm (Singh et al. 2013a). The identified protein sequences were used to search the oil palm transcriptome and genome databases using BiowiseV2 as described above.

### 3.2.4 Total RNA Extraction

Total RNA from majority of the oil palm tissues were extracted according to the modified method of (Prescott and Martin 1987). Five grams of frozen tissue were ground and transferred to 15 ml extraction buffer (Appendix 3). Homogenization was carried out immediately using a polytron homogenizer with the addition of phenol:chloroform (50:50) followed by centrifugation of the extract at 13,000 x g for 30 min at 25$^{o}$C. The aqueous phase was transferred into a new tube and chloroform:isoamyl alcohol (24:1) was added, followed by centrifugation at 13,000 x g for 30 min at 25$^{o}$C. Precipitation of RNA with 2M LiCl was carried out at 4$^{o}$C overnight. The next day, the tube was centrifuged at 13,000 x g for 30 min at 25$^{o}$C. The supernatant was discarded and the pellet was washed again with 2M LiCl before dissolving in 5 ml ice-cold sterile water. Precipitation of RNA was repeated overnight with 2M LiCl at 4$^{o}$C. The pellet obtained was washed with 2M LiCl and finally dissolved in an ice-cold sterile water.

Total RNA from the fruit AZ was isolated using Qiagen RNeasy Plant Mini Kit following the manufacturer's protocol. The frozen tissue (0.1g) were ground, transferred to 450 µl buffer RLT, vortexed vigorously followed by incubation at 56$^{o}$C to disrupt the tissue. The lysate was then transferred to a QIAshredder spin column (lilac) and centrifuged for 2 min at full speed. Next, the supernatant of the flow-through was transferred to a new microcentrifuge tube and 0.5 volume of ethanol (96–100% v/v) was added to the cleared lysate. The mixture was then transferred to an RNeasy spin column (pink) and centrifuged for 15 sec at 10,000 rpm. The RNeasy spin column membrane was washed once with 700 µl buffer RW1 and twice with 500 µl buffer RPE for 15 sec each at 10,000 rpm. Finally, the total RNA was eluted in 50 µl of RNase-free water and kept at -20$^{o}$C for storage.

For total RNA from both oil palm male and female flowers tissues, the total RNAs were kindly provided by Dr Meilina Ong and Dr Ooi Siew Eng from Plant Development Laboratory, MPOB, Bangi, Malaysia.

### 3.2.5 3` and 5` Rapid Amplification of cDNA Ends (RACE)

RACE was carried out using SMART™ RACE cDNA Amplification and Advantage 2 PCR Kit from Clontech, USA.

#### 3.2.5.1 First strand cDNA synthesis

Single stranded 3` RACE-Ready cDNA was synthesized from 1 µg DNAse-treated 20 WAA mesocarp total RNA. The total RNA was first denatured at

70ºC for 3 min in a 4 µl reaction mixture containing 1 µl of 3'-CDS primer A and sterile water. Next, reverse transcription was carried out in a 10 µl reaction mixture containing the denatured RNA, 2 µl of 5X first strand buffer, 1 µl of 20 mM DTT, 1 µl of 10 mM dNTP mix and 1 µl of SmartScribe Reverse Transcriptase (100 U) at 42ºC for 90 min. The reaction mixture was then heat inactivated at 70ºC for 10 min and 250 µl of trycine-EDTA buffer was added to the first strand cDNA prior to long term storage at -20ºC until needed.

Synthesis of single stranded 5` RACE-Ready cDNA was performed similarly as described above for 3` RACE-Ready cDNA. For preparation of the 5` RACE-Ready cDNA, 1 µl of 5'-CDS primer A was used in the RNA denaturing mixture instead of 1 µl of 3'-CDS primer A. In addition, 1 µl of oligo SMARTer IIA oligo was added to the 10 µl reverse transcription mixture. The incubation temperature and time remained as described for 3` RACE-Ready cDNA synthesis.

#### 3.2.5.2 3` and 5' RACE PCR Reactions

Touchdown PCR amplification was carried out in a 50 µl reaction mixture containing 34.5 µl of sterile water, 5 µl of 10X Advantage 2 PCR buffer, 1 µl of 10 mM dNTP mix, 1 µl of 50X Advantage 2 Polymerase Mix, 5 µl of 10X universal primer A mix, 1 µl of 10 µM gene-specific primers (Appendix 4) and finally (100 ng) 2.5 µl of 3'-RACE-Ready cDNA with the following conditions: 94ºC (5 sec), 72ºC (3 min) for 5 cycles; 94ºC (5 sec), 70ºC (10 sec), 72ºC (3 min) for 5 cycles; 94ºC (5 sec), 68ºC (10 sec) and 72ºC (3 min) for 25 cycles

and a final extension at 72ºC for 10 min. Some adjustments (± 5ºC) were made to the annealing temperatures especially when the fragment amplified on the agarose gel did not generate a discrete band. The amplified PCR products were analysed, purified, cloned and the plasmid isolated, digested and sequence analysed as described in sections 3.2.6 until 3.2.9.

### 3.2.6 Recovery and purification of PCR fragments

The amplified PCR fragments were purified from agarose gel using Qiaquick Gel Extraction Kit from Qiagen, Germany following the manufacturer's protocol. DNA fragment with the expected size was excised from the gel and transferred into a microcentrifuge tube. After weighing the gel slice in the microcentrifuge tube, 300 µl of QG buffer was added to each 100 mg of gel slice and the tube was incubated at 50ºC for 10 min. When the entire gel slice had completely dissolved, 100 µl of isopropanol was added to the sample and mixed. The sample mixture was then transferred to the QIAquick spin column and centrifuged at 13,000 rpm for 1 min. The column was then washed with 750 µl of PE buffer and finally the DNA was eluted with 50 µl of EB buffer (10 mM Tris-Cl, pH 8.5).

### 3.2.7 Cloning of PCR fragments

Cloning of amplified PCR products were performed using TOPO TA Cloning® Dual Promoter Kit (Invitrogen, USA). The cDNA fragment (100 ng) was first concentrated to 4 µl using speed vacuum concentrator (Heto Holten Maxi

DryPlus VACUUM Concentrator System, Denmark). Cloning reaction was carried out in a 6 µl reaction mixture containing 100 µg of the concentrated DNA fragment, 1 µl of salt solution (1.2 M NaCl, 0.06 M $MgCl_2$) and 1 µl of TOPO® vector at 22.5ºC for 30 min and quickly chilled on ice for 5 min. Two µl of the cloning reaction was added into a vial containing One Shot® Chemically Competent *Escherichia coli* TOP10 and was mixed gently. The vial was incubated on ice for 25 min and then, the cells were heat-shocked for 45 sec at 42ºC without shaking and immediately placed on ice for 5 min. This was followed by the addition of 250 µl of SOC medium to the tube and incubation was carried out at 37ºC for 25 min followed by incubation for 45 min at 37ºC with shaking at 150 rpm. Finally, 100 µl of the mixture was spread on a pre-warmed LB agar plate containing 50 µg/ml of ampicillin and 40 µg/ml of X-Gal and incubated overnight at 37ºC.

### 3.2.8 Plasmid Isolation and Digestion

Plasmid DNA was isolated using Qiagen Plasmid Mini Kit from Qiagen, Germany. A single white bacteria colony was grown in LB broth containing 50 µg/ml of ampicillin overnight at 37ºC with shaking at 200 rpm. The next day, the cultures were centrifuged at 5000 rpm for 10 min and the supernatant was discarded. The bacterial cell pellet was resuspended in 250 µl P1 buffer and transferred into a microcentrifuge tube. P2 buffer (250 µl) was added and the tube was gently inverted 4-6 times to mix. This was followed by the addition of 350 µl N3 buffer and the tube was inverted immediately and gently for 4-6 times

followed by centrifugation at 13,000 rpm for 10 min. The supernatant was then transferred to the QIAprep spin column and centrifuged for 1 min. The column was washed by adding 750 µl PE buffer and finally the DNA was eluted with 50 µl EB buffer (10mM Tris-Cl, pH 8.5). Digestion of the plasmids were performed in a 20 µl reaction mixture containing 2 µl (1 µg) plasmid DNA, 2 µl of Buffer $O^+$, 2 µl of *EcoRI (*10 U/µl) from Fermentas (USA) and 14 µl of sterile water at 37$^o$C overnight.

### 3.2.9 DNA sequencing and Sequence Analyses

DNA sequencing of plasmids and amplified PCR products were carried out by service provider, Nextgene Sdn. Bhd, Malaysia. The plasmids were sequenced from the 5' and 3' ends of the cDNA using M13 reverse and forward universal primers. Nucleotide and protein sequences were analysed using Bioedit (http://www.mbio.ncsu.edu/bioedit/bioedit.html), Biology Workbench Version 3.2 (http://www.biologyworkbench.sdsc.edu) and/or Expasy toolkit (www.expasy.org). Protein domains and conserved motifs were identified using NCBI-CDD search programme (Marchler-Bauer et al. 2013), Interproscan5 or IPRScan (www. http://www.ebi.ac.uk/Tools/pfa/iprscan5/) and/or SMART software (http://smart.embl-heidelberg.de/). Sequence similarity searches were performed *via* NCBI Blast (http://blast.ncbi.nlm.mih/gov/Blast.cgi). Comparison of the amino acid sequences between each sequences were performed with BL2seq programme, multiple sequence alignment was performed with Clustal W, protein molecular mass and isoelectric point

determination (PI) were calculated using Biology Workbench. Hydropathy profiles were analyzed using Kyte-Doolittle Plot. Prediction of the transmembrane helices location and location of intervening loop regions were detected using TMHMM2.0 programme. Prediction of the gene structures were carried out using SPIDEY-NCBI (http://www.ncbi.nlm.nih.gov/SPIDEY) and exonerate command *via* LINUX operating system. *In silico* promoter analysis to search for regulatory motifs was performed using PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/), Plant *Cis*-acting Regulatory DNA Elements or PLACE (www.dna.affrc.go.jp/PLACE/) and/or Plant Promoter Database - Softberry or Plantprom (http://linux1.softberry.com/). Motif search in the promoter regions were performed using the motif-based sequence analysis tools or MEME suite (http://meme.nbcr.net/) with the following parameters: number of motifs = 20, optimum motif width = ≥ 5 and ≤ 10, optimum site number = ≥ 3 and ≤ 10 and number of repetition = any. The identified motifs were annotated using PlantCARE.

### 3.2.10 Long-Distance-PCR (LD-PCR)

PCR amplifications were carried out similarly to 5` RACE. The PCR cocktail and conditions were as described in section 3.2.5.2 using gene-specific primers designed from the far end of 5` and 3` RACE products. The primer pair combination is listed in Appendix 4. The amplified PCR products were

analysed, purified, cloned and the plasmid isolated, digested and sequence analysed as described in sections 3.2.6 until 3.2.9.

### 3.2.11 Amplification of the splice variants

PCR amplification was carried out in a 25 µl reaction mixture containing 1X Advantage 2 PCR buffer (Clontech, USA), 1X Advantage 2 Polymerase Mix, 1X dNTP mix and 0.2 µM gene-specific primer pair combinations (35424SVF1 & ET3542R2, 35424SVF2 & ET3542R2 and 34177SVF2 & 34177SVR1) (Appendix 4). The templates used were 50 ng DNA, 250 ng of 20WAA mesocarp cDNA and 20 ng of spear leaves cDNA. PCR amplification was performed under the following conditions: denaturation at 95$^{o}$C (1 min); 95$^{o}$C (30 s) and 61$^{o}$C (1 min) for 30 cycles and a final extension of 60$^{o}$C for 1 min.

### 3.2.12 Phylogenetic analysis

Phylogenetic analyses were performed using the Molecular Evolutionary Genetics Analysis software package, MEGA5 (http://www.megasoftware.net) (Tamura et al. 2011). Multiple protein sequences were aligned using ClustalW in the MEGA 5 software with default parameters (gap opening penalty = 10, gap extension penalty = 0.2, gap separation distance = 4, Blosum weight matrix with residue specific and hydrophilic penalties (Thompson et al. 1994). The phylogenetic relationship between the members of the ethylene receptor family were constructed by the neighbor-joining method using bootstrap analysis with 1000 replicates. Evolutionary protein distances were calculated in the units of

amino acid using Poisson parameter model with pairwise deletion of gaps in the alignment data. The constructed tree was viewed using Tree Explorer in the MEGA 5 software.

### 3.2.13 Mapping to oil palm chromosome and genetic linkage groups

The oil palm ethylene receptor gene sequences were mapped on to the oil palm chromosomes by identifying their scaffold positions on the EG6 build using exonerate search with a 60% self-score threshold. Genetic markers flanking the ethylene receptor genes were identified based on their mapped position to the scaffolds of the EG6-build (data not shown). The markers information on the T128 and P2 linkage groups (Singh et al. 2013a; 2013b) were used to predict the location of the ethylene receptors. Graphic representation of the linkage groups was visualized using MapChart 2.2 (Voorrips 2002).

### 3.2.14 Genomic DNA Extraction

The genomic DNA from oil palm leaves was isolated based on the modified Cetylatrimethyl Ammonium Bromide (CTAB) method of (Doyle and Doyle 1990). A total of 20 g tissues were ground using pestle and mortar in liquid nitrogen until fine and powdery form. Then, it was transferred into a tube containing 100 ml of modified CTAB buffer (Appendix 3) that had been pre-heated at 60$^{o}$C, followed by the addition of 0.5% (500 µl) of β-mercaptoethanol. The mixture was swirled gently to mix and left to incubate at 60$^{o}$C for 30 min, followed by cooling at room temperature. This was followed by the addition of

100 ml of chloroform:isoamyl alcohol (24:1) and the mixture was centrifuged at 10,000 rpm for 15 min at 4$^{o}$C. The aqueous phase was transferred into a new tube and 0.6 volume of isopropanol was added to the mixture and left at 4$^{o}$C overnight to allow precipitation of nucleic acids.

The next day, the tube was centrifuged at 12,000 rpm for 20 min at 4$^{o}$C. The supernatant was discarded and the pellet was washed twice with 76% (w/v) ethanol/10mM ammonium acetate and left to dry before dissolving in 2 ml of TE buffer, pH 8.0 at 50$^{o}$C for 4-5 h. RNase solution (2.5 µl) was added to the mixture followed by incubation at 37$^{o}$C for 30 min. This was followed by the addition of 1 ml of 7.5M ammonium acetate, pH 7.7 and the mixture was left to chill on ice for 30 min and centrifuged at 12,000 rpm for 20 min at 4$^{o}$C. The supernatant was transferred into a new tube and precipitation of DNA with 2.5 volume of ethanol was carried out at 4$^{o}$C overnight. The next day, centrifugation to pellet the DNA and washing of the pellet was carried as described above. Finally the pellet was dissolved in 2 ml of TE buffer, pH 8.0 and kept at 4$^{o}$C.

### 3.2.15 Promoter isolation

To determine the possible regulatory motifs driving the expression of the putative ethylene receptor genes, the putative promoters were amplified using gene-specific primers designed based on the reconstitution of both the 5`RACE and the oil palm genome sequence data. The PCR amplification was carried out in a 25 µl reaction mixture containing 1X Advantage 2 PCR buffer (Clontech, USA), 1X Advantage 2 Polymerase Mix, 1X dNTP mix and 0.2 µM

gene-specific primer pair combinations (Appendix 4) and 50 ng DNA. Most of the PCR amplifications were performed under the following conditions: denaturation at 95ºC (1 min); 95ºC (30 s) and 65ºC (3 min) for 30 cycles and a final extension of 65ºC for 3 min. For *EgERS3*, the promoter was amplified with the following conditions: denaturation at 94ºC (15 sec); 94ºC (5 sec), 59ºC (15 sec), 72ºC (3 min) for 30 cycle and a final extension at 72ºC for 7 min. *In silico* analysis of the promoters were performed as described in section 3.2.9.

### 3.2.16 Expression analyses *via* RT-PCR

#### 3.2.16.1 Synthesis of first strand cDNA for Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

Single stranded cDNA 5 was synthesized from 2 µg DNAse-treated total RNA using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA). Reverse transcription was carried out in a 20 µl reaction mixture containing 2 µg DNAse-treated total RNA, 2X RT buffer, 5 U/µl MultiScribe™ Reverse Transcriptase, 2X dNTP mix (100 mM) and 2X RT random primers with the following conditions: 25ºC (1 min); 37ºC (120 min) and 85ºC (5 min) to stop the reaction. The synthesized cDNA was kept at -20ºC for long term storage.

#### 3.2.16.2 PCR amplifications

PCR amplification was carried out in a 25 µl reaction mixture containing 0.1 µg cDNA, 1X Advantage 2 PCR buffer (Clontech, USA), 1X Advantage 2

Polymerase Mix, 1X dNTP mix, 0.2 µM forward and reverse gene-specific primers (Appendix 4) with the following conditions: denaturation at 95$^{o}$C (1 min); 95$^{o}$C (30 s) and 60$^{o}$C (1 min) for 30 cycles and a final extension of 60$^{o}$C for 1 min. Some adjustments (± 5$^{o}$C) were made to the annealing temperatures especially when the amplified fragment was not discrete on the agarose gel.

### 3.2.17 Amplification of internal control for RT-PCR

Amplification of the actin gene as internal control for expression analysis was performed in a 25 µl reaction mixture containing 0.1 µg cDNA, 1X Advantage 2 PCR buffer (Clontech, USA), 1X Advantage 2 Polymerase Mix, 1X dNTP mix, 0.2 µM actin F and 0.2 µM actin R primers (Appendix 4) with the following conditions: denaturation at 95$^{o}$C (1 min); 95$^{o}$C (15 sec), 55$^{o}$C (30 sec) and 72$^{o}$C (2 min) for 30 cycles and a final extension of 72$^{o}$C for 7 min.

### 3.2.18 Estimation of RNA and DNA Purity and Integrity

Quantification of the amount of RNA and DNA was performed using NanoDrop® Spectrophotometer ND-1000 (Thermo Fisher Scientific Inc.) at three different wavelengths (260 nm, 280 nm and 230 nm). The ratio between the readings at 260 nm and 280 nm ($A_{260}$:$A_{280}$) provides an estimate of the purity of the DNA and RNA. Integrity of RNA was determined *via* electrophoresis on 1% (w/v) agarose gel.

### 3.2.19 Agarose Gel Electrophoresis 1% (w/v)

Agarose gel was prepared by dissolving 1% (w/v) agarose in 1X TAE buffer (Appendix 1). The mixture was poured into the gel tank and left to solidify. A 1/5 volume of 6X DNA Loading Dye (Fermentas, USA) was added to the amplified PCR products and the mixtures were loaded into the wells of the submerged gel. Gel electrophoresis was carried out at 70V with 1X TAE as the electrophoresis buffer. After migration of the dye at 2/3 of the distance through the gel, the gel was removed from the running buffer and immersed in ethidium bromide (EtBr) (0.5 µg/ml) for 20 min at room temperature. The stained gel was viewed using a UV (ultra violet) transilluminator and the image was captured using SYNGENE Gel Imaging System, United Kingdom.

## 3.3 Results

### 3.3.1 Identification of oil palm ethylene receptor candidate genes

Figure 3.1 represents the flowchart of the process to identify the putative ethylene receptor genes from oil palm. For homology search, the sequence from a number of known ethylene receptor gene family members from a variety of plant species such as rice, date palm, maize, Arabidopsis and soybean as well as *EgERD3* (Nurniwalis 2006; Nurniwalis et al. 2012) were used to search against the oil palm transcriptome databases. The transcriptome databases are comprised of assembled short read transcripts originating from multiple tissues based on *E. guineensis tenera* fruit form as well as from *E. oleifera* (Singh et al. 2013a). A large number of sequences were identified *via* BLASTN and BLASTX but those with scores and e-values of more than 100 and $e^{-60}$, respectively were selected and translated into amino acid sequences for further analyses. Out of all the transcriptome databases, the *pisifera* transcriptome database generated the most reliable and largest coverage especially in terms of the assembled transcript length as well as the deduced open reading frame (ORF). Thus, in total 9 unique assembled transcripts or isotigs were identified using the homology search approach and the sequence information is shown in Table 3.1.

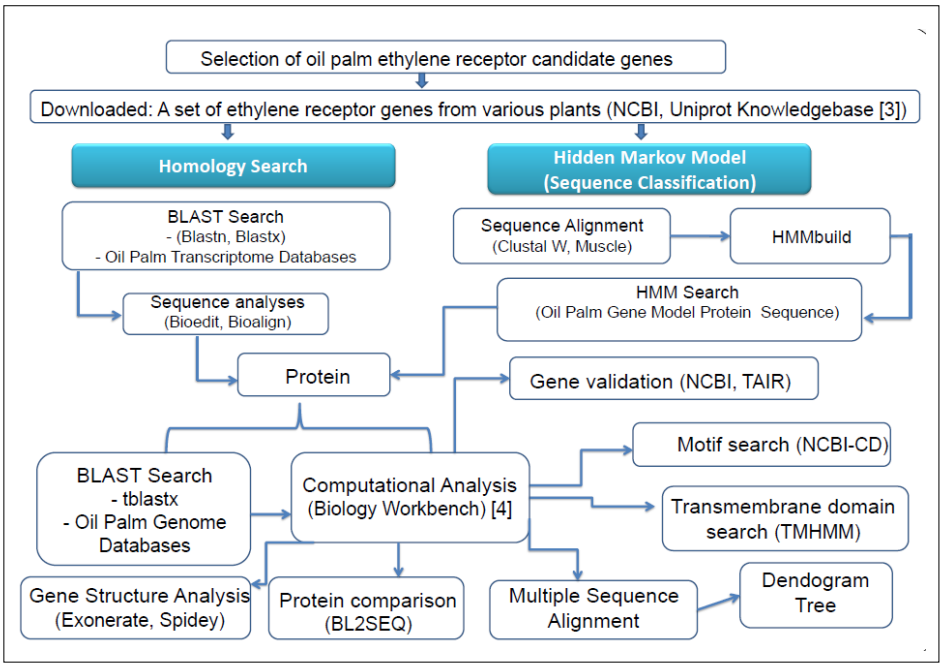



**Figure 3.1: A schematic diagram of the process to identify the ethylene receptor gene family from the oil palm genome.**

**Table 3.1: List of oil palm genome sequences with the longest size and highest hits (E value = 0.0) to the *pisifera* xscript transcriptome database and their sequence characteristics.**

| No. | Transcript ID | Gene name | Length (bp) | Protein (Open Reading Frame ) | | |
|---|---|---|---|---|---|---|
| | | | | Length (aa) | Mol. Mass (kDA)* | PI* |
| 1. | Isotig34176 | *EgERS1* | 2479 | 629 | 70.013 | 7.305 |
| 2. | Isotig34177 | *EgERS1-1* | 2441 | 617 | 68.787 | 7.305 |
| 3. | Isotig45619 | *EgERS2* | 2565 | 635 | 70.793 | 7.124 |
| 4. | Isotig46141 | | 2204 | 549 | 60. 820 | 7.661 |
| 5. | Isotig24120 | *EgETR2-1* | 3607 | 687 | 76.811 | 6.301 |
| 6. | Isotig24121 | *EgETR2* | 3051 | 687 | 76.811 | 6.301 |
| 7. | Isotig35423 | *EgETR3* | 3195 | 757 | 85.209 | 6.304 |
| 8. | Isotig35424 | *EgETR3-1* | 3032 | 757 | 85.209 | 6.304 |
| 9. | Isotig44932 | *EgETR4* | 4186 | 755 | 85.150 | 6.704 |

* Molecular mass (kDa) and isoelectric point (PI) of the putative proteins were calculated using Biology Workbench Biology Version 3.2, University of California

In addition, a HMM profile was also built as an alternative approach to identify other putative ethylene receptors from the oil palm. Multiple sequence alignments of two hundred protein sequences were performed and the resulting HMM build was used to search against the oil palm gene model. The output of the protein search with scores and e-values similar to that of the homology search were selected for further analysis while the rest were discarded. Comparison of the HMM protein sequences to the translated amino acid sequences generated from the homology search showed a similar result. All 9 isotigs that were identified earlier (Table 3.1) were successfully identified using this approach. In addition, a predicted protein sequence with protein identity (ID) designated as maker-p5_sc00025-augustus-gene-13.28-mRNA-1 was also identified (Table 3.2). It is comprised of 2469 bases and is 100% identical to isotig46141. It is however, 265 bp longer than isotig46141, especially 189 bp longer at the 5` upstream region. Figure 3.2 shows the nucleotide sequence alignment between isotig46141 and maker-p5_sc00025-augustus-gene-13.28-mRNA-1. Additionally, another predicted protein sequence that encodes 737 amino acid residues or 3187 nucleotides was also identified from the HMM build (Table 3.2). The protein designated as maker-p5_sc00069-augustus-gene-6.47-mRNA-1 corresponds to isotig53269, a 948 nucleotide assembled transcript in the transcriptome database. Consequently, this finding brings to a total of 10 unique isotigs identified from the oil palm genome.

**Table 3.2: Sequence information and characteristic (E value = 0.0) of the additional sequences identified from the oil palm gene model.**

| No. | Gene Model ID | Gene name | Length (bp) | Protein (Open Reading Frame) | | |
|---|---|---|---|---|---|---|
| | | | | Length (aa) | Mol. Mass (kDA) | PI |
| 1. | maker-p5_sc00025-augustus-gene-13.28-mRNA-1 | *EgERS3* | 2469 | 635 | 70.793 | 7.137 |
| 2. | maker-p5_sc00069-augustus-gene-6.47-mRNA-1 | *EgETR5* | 3187 | 737 | 82.364 | 6.763 |

* Molecular mass (kDa) and PI of the isotigs were calculated using Biology Workbench Version 3.2, University of California.

```
PiV2x_isot46141_(new_region=1to2      ------------------------------------------------------------
maker-p5_sc00025-augustus-gene-1     TTCCATTGTTTAAAAGGAATGTAGAGAGGAGAATATTCTGGCTGTGGCGGTTGCTGGGAG

PiV2x_isot46141_(new_region=1to2      ------------------------------------------------------------
maker-p5_sc00025-augustus-gene-1     ATGGAGGGCTGTGACTGCATTGAGCCGCAATGGCCTGCAGATGAGCTTTTGGTGAAGTAT

PiV2x_isot46141_(new_region=1to2      ------------------------------------------------------------
maker-p5_sc00025-augustus-gene-1     CAGTACATATCGGACTTCTTTATAGCACTTGCATATTTCTCGATACCCCTTGAGCTCATT

PiV2x_isot46141_(new_region=1to2      ---------AAGAAGTCTTCATTCTTCCCGTATAGATGGGTGTTAATACAGTTCGGTGCT
maker-p5_sc00025-augustus-gene-1     TATTTTGTGAAGAAGTCTTCATTCTTCCCGTATAGATGGGTGTTAATACAGTTCGGTGCT
                                              **************************************************

PiV2x_isot46141_(new_region=1to2      TTCATTGTTCTTTGTGGAGCAACCCACTTGATTAATTTATGGACTTTCACCGTGCACTCG
maker-p5_sc00025-augustus-gene-1     TTCATTGTTCTTTGTGGAGCAACCCACTTGATTAATTTATGGACTTTCACCGTGCACTCG
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      AGAACACTTGCCATAGTAATGACAGTAGCAAAGATTTCAACTGCCGTTGTGTCATGTGCC
maker-p5_sc00025-augustus-gene-1     AGAACACTTGCCATAGTAATGACAGTAGCAAAGATTTCAACTGCCGTTGTGTCATGTGCC
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      ACTGCATTGATGCTTGTGCATATAATTCCTGACCTGTTGAGTGTGAAAACAAGGGAGCTT
maker-p5_sc00025-augustus-gene-1     ACTGCATTGATGCTTGTGCATATAATTCCTGACCTGTTGAGTGTGAAAACAAGGGAGCTT
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      TTTCTGAAGAATAAGGCTGAGGAGCTTGATAGGGAGATGGGCCTTATAAGAACACAGGAA
maker-p5_sc00025-augustus-gene-1     TTTCTGAAGAATAAGGCTGAGGAGCTTGATAGGGAGATGGGCCTTATAAGAACACAGGAA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      GAGACAGGGAGGCATGTCCGAATGCTGACGCATGAAATCAGAAGCACACTTGACAGGCAC
maker-p5_sc00025-augustus-gene-1     GAGACAGGGAGGCATGTCCGAATGCTGACGCATGAAATCAGAAGCACACTTGACAGGCAC
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      ACTATATTGAAGACTACACTTGTTGAGCTAGGGAGGACTTTGGATTTGGCAGAATGTGCA
maker-p5_sc00025-augustus-gene-1     ACTATATTGAAGACTACACTTGTTGAGCTAGGGAGGACTTTGGATTTGGCAGAATGTGCA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      CTGTGGATGCCATCCAATAGTGGTTTAAATCTTCAGCTTTCTCATACTCTGCACCACCAA
maker-p5_sc00025-augustus-gene-1     CTGTGGATGCCATCCAATAGTGGTTTAAATCTTCAGCTTTCTCATACTCTGCACCACCAA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      ATACCTCTTGGATCCGTCGTGTCTATTAACCTACCAATTGTCAATCAAGTTTTCAGCAGT
maker-p5_sc00025-augustus-gene-1     ATACCTCTTGGATCCGTCGTGTCTATTAACCTACCAATTGTCAATCAAGTTTTCAGCAGT
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      AACCGTGCAATAAGGATTCCGCATACTTGTCCACTAGCAAGGATCCGGCCACTCACAGGA
maker-p5_sc00025-augustus-gene-1     AACCGTGCAATAAGGATTCCGCATACTTGTCCACTAGCAAGGATCCGGCCACTCACAGGA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      AGATATGTGCCACCAGAAGTTGTTGCTGTCCGTGTACCGCTATTACATCTTTCAAATTTC
maker-p5_sc00025-augustus-gene-1     AGATATGTGCCACCAGAAGTTGTTGCTGTCCGTGTACCGCTATTACATCTTTCAAATTTC
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      CAAGTAAATGATTGGCCTGAGCTATCTGCAAAAAGCTATGCAATAATGGTTTTGATTCTT
maker-p5_sc00025-augustus-gene-1     CAAGTAAATGATTGGCCTGAGCTATCTGCAAAAAGCTATGCAATAATGGTTTTGATTCTT
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      CCATCGAATAGTGCAAGAAAGTGGCATTTCCATGAACTGGAGCTTGTGGAGGTGGTTGCT
maker-p5_sc00025-augustus-gene-1     CCATCGAATAGTGCAAGAAAGTGGCATTTCCATGAACTGGAGCTTGTGGAGGTGGTTGCT
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      GATCAGGTAGCAGTTGCACTTTCTCATGCTGCAATTCTAGAAGAGTCTATGCAGGCACGC
maker-p5_sc00025-augustus-gene-1     GATCAGGTAGCAGTTGCACTTTCTCATGCTGCAATTCTAGAAGAGTCTATGCAGGCACGC
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      AACCTACTCATGGAGCAAAATGTTGCTCTAGATTTAGCTCGTCGGGAGGCAGAAATGGCA
maker-p5_sc00025-augustus-gene-1     AACCTACTCATGGAGCAAAATGTTGCTCTAGATTTAGCTCGTCGGGAGGCAGAAATGGCA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      ATTCGTGCTCGCAATGATTTTCTAGCTGTCATGAACCATGAAATGCGGACTCCTATGCAT
maker-p5_sc00025-augustus-gene-1     ATTCGTGCTCGCAATGATTTTCTAGCTGTCATGAACCATGAAATGCGGACTCCTATGCAT
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      GCAATCATTGCTCTGTCCTCCCTGCTTCTGGAAACCGAGCTAACGCCAGAGCAACGCTTG
maker-p5_sc00025-augustus-gene-1     GCAATCATTGCTCTGTCCTCCCTGCTTCTGGAAACCGAGCTAACGCCAGAGCAACGCTTG
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      ATGGTGGAAACTGTATTGAAGAGTAGTAACCTTCTCGCAACACTCATCAATGATGTTTTA
maker-p5_sc00025-augustus-gene-1     ATGGTGGAAACTGTATTGAAGAGTAGTAACCTTCTCGCAACACTCATCAATGATGTTTTA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      GATCTTTCTAAGCTTGAGGATGGGAGCCTTGAGCTGGAGATTGCAGCATTCAATCTTCAT
maker-p5_sc00025-augustus-gene-1     GATCTTTCTAAGCTTGAGGATGGGAGCCTTGAGCTGGAGATTGCAGCATTCAATCTTCAT
                                     ************************************************************
```

```
PiV2x_isot46141_(new_region=1to2      TCTGTTTTCAGAGAGGTCATTCATTTGATAAAGCCGATAGCAGCTGTAAAGAAGCTCTCA
maker-p5_sc00025-augustus-gene-1     TCTGTTTTCAGAGAGGTCATTCATTTGATAAAGCCGATAGCAGCTGTAAAGAAGCTCTCA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      GTATTGGTTACACTAGCACCTGATCTGCCTTTGTGTGCCATTGGTGATGAGAAGCGGCTT
maker-p5_sc00025-augustus-gene-1     GTATTGGTTACACTAGCACCTGATCTGCCTTTGTGTGCCATTGGTGATGAGAAGCGGCTT
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      ATGCAAGCTATTTTAAATATATCTGGTAATGCCGTTAAGTTCACAAAGGAGGGTCATATA
maker-p5_sc00025-augustus-gene-1     ATGCAAGCTATTTTAAATATATCTGGTAATGCCGTTAAGTTCACAAAGGAGGGTCATATA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      TCAATTATGGCTTCTGTTGCGAAACCAGATTCTCTAAGAGACTCTCGAGCTCCCGAATAT
maker-p5_sc00025-augustus-gene-1     TCAATTATGGCTTCTGTTGCGAAACCAGATTCTCTAAGAGACTCTCGAGCTCCCGAATAT
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      TATCCAATTACTAGTGACGGGCACTTCTACTTGCGTGTGCAGGTAAAGGATACTGGTTGT
maker-p5_sc00025-augustus-gene-1     TATCCAATTACTAGTGACGGGCACTTCTACTTGCGTGTGCAGGTAAAGGATACTGGTTGT
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      GGAATCAGTCCCCAAGATTTACCACACATCTTTACCAAATTTGCACGCTCTCAACATGGA
maker-p5_sc00025-augustus-gene-1     GGAATCAGTCCCCAAGATTTACCACACATCTTTACCAAATTTGCACGCTCTCAACATGGA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      GCAAACAAAGGTTATAGTGGAAGCGGACTCGGGCTCGCCATTTGTAAGAGGTTTATAAGC
maker-p5_sc00025-augustus-gene-1     GCAAACAAAGGTTATAGTGGAAGCGGACTCGGGCTCGCCATTTGTAAGAGGTTTATAAGC
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      CTCATGGAAGGACACATTTGGCTTGAGAGTGAAGGTGTGGGTAAAGGTTGCACAGCAACA
maker-p5_sc00025-augustus-gene-1     CTCATGGAAGGACACATTTGGCTTGAGAGTGAAGGTGTGGGTAAAGGTTGCACAGCAACA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      TTTATTGTCAAACTTGGTGTATGTGAAAAATCGAAGAGTCATCTTCAACAAATTGTTCCA
maker-p5_sc00025-augustus-gene-1     TTTATTGTCAAACTTGGTGTATGTGAAAAATCGAAGAGTCATCTTCAACAAATTGTTCCA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      GTGACTGGGTCAAATATTGGAGATGCTGATCTTTCTGGTCCAAGAGTACCCTTTCGAGAC
maker-p5_sc00025-augustus-gene-1     GTGACTGGGTCAAATATTGGAGATGCTGATCTTTCTGGTCCAAGAGTACCCTTTCGAGAC
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      GAGAGTGGTCTGCTTCCCTCCAGGTTCCGGTACCAGAGAAGTGTATAGATATTACAGAAG
maker-p5_sc00025-augustus-gene-1     GAGAGTGGTCTGCTTCCCTCCAGGTTCCGGTACCAGAGAAGTGTATAGATATTACAGAAG
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      AAAAATTTCTGCATGATGGAGCAATACAGCTTGAAAAGCAATGATCCATGGGTCCAATTT
maker-p5_sc00025-augustus-gene-1     AAAAATTTCTGCATGATGGAGCAATACAGCTTGAAAAGCAATGATCCATGGGTCCAATTT
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      TGTGACGGTTCAAGGCGACGGGCTGTTCAGCCAGCAACACAGGTTAAGAGGATTCTGATG
maker-p5_sc00025-augustus-gene-1     TGTGACGGTTCAAGGCGACGGGCTGTTCAGCCAGCAACACAGGTTAAGAGGATTCTGAT-
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      GTTTAGTGATGATCATGAGCTGAGGTGATTAGATTTGCAGCTAGGCTAGGCTGAATTGAA
maker-p5_sc00025-augustus-gene-1     -------------------------------GATTTGCAGCTAGGCTAGGCTGAATTGAA
                                                                     ****************************

PiV2x_isot46141_(new_region=1to2      CAGCAAGTTGCCTGATAGGATTTCTGTTTCACAAGTTCTATATTAGGAACACTTGCTTCC
maker-p5_sc00025-augustus-gene-1     CAGCAAGTTGCCTGATAGGATTTCTGTTTCACAAGTTCTATATTAGGAACACTTGCTTCC
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      CATCGAAAGCATTCTAAATTTGCTAACTGCAATAAATCCATTGGAGAAAGAGAATGGGTT
maker-p5_sc00025-augustus-gene-1     CATCGAAAGCATTCTAAATTTGCTAACTGCAATAAATCCATTGGAGAAAGAGAATGGGTT
                                     ************************************************************


PiV2x_isot46141_(new_region=1to2      TTCTTGAACTTGTCAAGCAGATTGGCATAGCGTAGGATGCATGCTTGTCTTGTATTTGGA
maker-p5_sc00025-augustus-gene-1     TTCTTGAACTTGTCAAGCAGATTGGCATAGCGTAGGATGCATGCTTGTCTTGTATTTGGA
                                     ************************************************************

PiV2x_isot46141_(new_region=1to2      GGAGCCATGAATCGTATATCTTTACTTATAGAGCATCAAACATGATCTTTTCA-------
maker-p5_sc00025-augustus-gene-1     GGAGCCATGAATCGTATATCTTTACTTATAGAGCATCAAACATGATCTTTTCATTTTTCC
                                     *****************************************************

PiV2x_isot46141_(new_region=1to2      ------------------------------------------------------------
maker-p5_sc00025-augustus-gene-1     TCTGAGTTCATACTTAGGTTGAATAGCTTTCACTATATATTTCTAACCATTGCAAATATT


PiV2x_isot46141_(new_region=1to2      ----------------------------------------
maker-p5_sc00025-augustus-gene-1     TGGGCTGATGTTTTGAATTTTAAATTGATCTTTCAAATGTG
```

**Figure 3.2: Multiple sequence alignment by MUSCLE (3.8) between isotig46141 and p5_sc00025-augustus-gene-13.28-mRNA-1.** The highlighted nucleotides represents the first start (atg) and stop codons (tga).

To ensure that all the 10 identified isotigs encode ethylene receptors, blastp (Altschul et al. 1997) searches against the non-redundant (nr) NCBI database was performed. At the protein level, all 10 isotigs have sequence similarities ranging from 37% to 94% to other plant ethylene receptors and the highest hits are found to be to date palms. Table 3.3 represents the top two hit list with the highest scores and e-value against the NCBI database. The 10 putative ethylene receptor genes encode polypeptides that have ORF of 617 - 757 amino acid residues. Deduced amino acid sequence comparison of the 10 predicted proteins to each other showed a 38 - 100% identity (Table 3.4). The lowest identity of 38% was found between isotig44932 and isotig34176, isotig44932 and isotig34177, isotig35423 and isotig45619, isotig35423 and maker-p5_sc00025-augustus-gene-13.28-mRNA-1, isotig35424 and isotig45619 as well as isotig34524 and maker-p5_sc00025-augustus-gene-13.28-mRNA-1. The highest identity (100%) was found between isotig24121 and isotig24120 as well as isotig35423 and isotig35424, respectively. Multiple sequence alignment between isotig24121 and isotig24120 as well as isotig35423 and isotig35424 at the nucleotide level confirmed the 100% identity within the ORF. However, the sequence identity at 5` region located upstream from the start codon is low (Figures 4.3 and 4.4). The difference within this region was further compared to the oil palm genome data and the result is presented in section 3.3.2.

**Table 3.3: Homology Search Results of the Putative Ethylene Receptor Genes from Oil Palm against the NCBI and TAIR databases.**

| Sequence ID | Database | Accession Number &Description | Score | E value | % Identity |
|---|---|---|---|---|---|
| **Isotig 34176** | TAIR | 1. AT1G66340.1 *ETR1, EIN1, ETR, AtETR1* | 914 | 0.0 | 71 |
| | | 2. AT2G40940.1 *ERS1*, ERS ethylene response sensor | 857 | 0.0 | 71 |
| | NCBI | 1. XP_008775899.1 PREDICTED: LOW QUALITY PROTEIN: ethylene receptor-like *[Phoenix dactylifera]* | 1210 | 0.0 | 94 |
| | | 2. XP_004982062.1 PREDICTED: ethylene receptor-like isoform X1 [*Setaria italica*] | 1095 | 0.0 | 85 |
| **Isotig 45619** | TAIR | 1. AT1G66340.1 *ETR1, EIN1, ETR, AtETR1* | 870 | 0.0 | 67 |
| | | 2. AT2G40940.1 *ERS1*, ERS ethylene response sensor | 815 | 0.0 | 67 |
| | NCBI | 1. XP_008810202.1 PREDICTED: LOW QUALITY PROTEIN: ethylene receptor 2-like *[Phoenix dactylifera]* | 1139 | 0.0 | 88 |
| | | 2. XP_008807380.1 PREDICTED: ethylene receptor 2-like *[Phoenix dactylifera]* | 1125 | 0.0 | 88 |
| **Isotig 46141** | TAIR | 1. AT1G66340.1 *ETR1, EIN1, ETR, AtETR1* | 913 | 0.0 | 70 |
| | | 2. AT2G40940.1 *ERS1,* ERS ethylene response sensor | 855 | 0.0 | 69 |
| | NCBI | 1. XP_008807380.1 PREDICTED: ethylene receptor 2-like *[Phoenix dactylifera]* | 1206 | 0.0 | 94 |
| | | 2. ABD66593.1 ethylene receptor [*Lilium formosanum x Lilium longiflorum*] | 1142 | 0.0 | 87 |
| **Isotig 35423** | TAIR | 1. AT3G04580.2 *EIN4* Signal transduction histidine kinase | 766 | 0.0 | 55 |
| | | 2. AT3G04580.1 *EIN4* \| Signal transduction histidine kinase | 766 | 0.0 | 55 |
| | NCBI | 1. XP_008794286.1 PREDICTED: protein *EIN4*-like isoform X1 *[Phoenix dactylifera]* | 1403 | 0.0 | 90 |
| | | 2. XP_008801307.1 PREDICTED: protein *EIN4*-like isoform X1 *[Phoenix dactylifera]* | 1306 | 0.0 | 83 |
| **Isotig 44932** | TAIR | 1. AT3G04580.2 *EIN4* Signal transduction histidine kinase | 758 | 0.0 | 54 |
| | | 2. AT3G04580.1 *EIN4* \| Signal transduction histidine kinase | 758 | 0.0 | 54 |
| | NCBI | 1. XP_008787409.1 PREDICTED: protein *EIN4 [Phoenix dactylifera]* | 1442 | 0.0 | 92 |
| | | 2. XP_008801307.1 PREDICTED: protein *EIN4*-like isoform X1 *[Phoenix dactylifera]* | 1098 | 0.0 | 71 |

| | | | | | |
|---|---|---|---|---|---|
| **Isotig 24121** | TAIR | 1. AT3G04580.2 *EIN4* Signal transduction histidine kinase | 661 | 0.0 | 55 |
| | | 2. AT3G04580.1 *EIN4* \| Signal transduction histidine kinase | 661 | 0.0 | 55 |
| | NCBI | 1. XP_008801310.1\| PREDICTED: protein *EIN4*-like isoform X1 *[Phoenix dactylifera]* | 1177 | 0.0 | 92 |
| | | 2. XP_008801307.1\| PREDICTED: protein *EIN4*-like isoform X2 *[Phoenix dactylifera]* | 1162 | 0.0 | 92 |
| **Maker 53269** | TAIR | 1. AT3G04580.2 \| Symbols: *EIN4* \| Signal transduction histidine kinase | 438 | e-123 | 37 |
| | | 2. AT3G04580.1 \| Symbols: *EIN4* \| Signal transduction histidine kinase | 438 | e-123 | 37 |
| | NCBI | 1. XP_008801620.1 PREDICTED: ethylene receptor 2-like *[Phoenix dactylifera]* | 1220 | 0.0 | 87 |
| | | 2. XP_009411143.1 PREDICTED: ethylene receptor 2-like [*Musa acuminata subsp. malaccensis]* | 788 | 0.0 | 58 |
| **Isotig 24120** | TAIR | 1. AT3G04580.2 *EIN4* Signal transduction histidine kinase | 661 | 0.0 | 55 |
| | | 2. AT3G04580.1 *EIN4* \| Signal transduction histidine kinase | 661 | 0.0 | 55 |
| | NCBI | 1. XP_008801310.1\| PREDICTED: protein *EIN4*-like isoform X1 *[Phoenix dactylifera]* | 1177 | 0.0 | 92 |
| | | 2. XP_008801307.1\| PREDICTED: protein *EIN4*-like isoform X2 *[Phoenix dactylifera]* | 1162 | 0.0 | 92 |
| **Isotig 35424** | TAIR | 1. AT3G04580.2 *EIN4* Signal transduction histidine kinase | 766 | 0.0 | 55 |
| | | 2. AT3G04580.1 *EIN4* \| Signal transduction histidine kinase | 766 | 0.0 | 55 |
| | NCBI | 1. XP_008794286.1 PREDICTED: protein EIN4-like isoform X1 *[Phoenix dactylifera]* | 1403 | 0.0 | 90 |
| | | 2. XP_008801307.1 PREDICTED: protein EIN4-like isoform X1 *[Phoenix dactylifera]* | 1306 | 0.0 | 83 |
| **Isotig 34177** | TAIR | 1. AT1G66340.1 *ETR1, EIN1, ETR, AtETR1* | 906 | 0.0 | 72 |
| | | 2. AT2G40940.1 *ERS1*, ERS ethylene response sensor | 856 | 0.0 | 71 |
| | NCBI | 1. XP_008775899.1 PREDICTED: LOW QUALITY PROTEIN: ethylene receptor-like *[Phoenix dactylifera]* | 1178 | 0.0 | 93 |
| | | 2. XP_004982062.1 PREDICTED: ethylene receptor-like isoform X1 [*Setaria italica*] | 1082 | 0.0 | 85 |

**Table 3.4: Protein correlation of the nine putative ethylene receptors from the oil palm genome to each other with BL2SEQ,** Biology Workbench using the default matrix selection parameter (Matrix: BLOSUM62, Gap Opening Penalty: 11, Gap Extension Penalty: 1 and Lambda Ratio: 0.5).

| Sequence ID | Protein size (aa) | Amino acid identity (%) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | EgERS1 | Eg ERS1-1 | EgERS2 | EgERS3 | EgETR3 | EgETR3-1 | EgETR2-1 | EgET2 | EgETR4 | EgETR5 |
| EgERS1 | 629 | 100 | | | | | | | | | |
| EgERS1-1 | 617 | 98 | 100 | | | | | | | | |
| EgERS2 | 635 | 84 | 80 | 100 | | | | | | | |
| EgERS3 | 635 | 81 | 83 | 87 | 100 | | | | | | |
| EgETR3 | 757 | 40 | 40 | 40 | 40 | 100 | | | | | |
| EgETR3-1 | 757 | 40 | 40 | 40 | 40 | 100 | 100 | | | | |
| EgETR2-1 | 687 | 40 | 40 | 38 | 38 | 85 | 85 | 100 | | | |
| EgETR2 | 687 | 40 | 40 | 38 | 38 | 85 | 85 | 100 | 100 | | |
| EgETR4 | 755 | 38 | 38 | 39 | 39 | 71 | 71 | 72 | 72 | 100 | |
| EgETR5 | 737 | 41 | 41 | 40 | 39 | 45 | 45 | 45 | 45 | 44 | 100 |

```
isocontig35424      ----------------------------------------------------GACAAGCT
isocontig35423      CCCTCCTCTTTGAGCGCTGAAATGCGAGCAGGGAGGCAAGGCCTCGGAGAAGGACGGGCG
                                                                     ***  **

isocontig35424      AATCTTTAAATCT----ATTTATTAGTT--TGGTATCTTCCAAAGACCAGAG-------T
isocontig35423      GAGGACGAGGTCTGAGGATTTCACTGTCACTGGTGGCCGAGGAGGGACGGACCCACTCCT
                     *     *  ***    ****    **   ****  *     * *  * **        *

isocontig35424      CTTGAGGTTCAGGTTAGGGTTCTTTTTCTGTTCTTCGCAAGGATCATTGAAATAGATAAC
isocontig35423      CTTGAGGT--GACTCGGGGCCCTTCGCGAAGCCTTTTCGGTGGGAGGCGGATTGGA-GAC
                    ********     *  ***  ***        ***  *   *      * * * **  **

isocontig35424      TACAAACT-GTGAT-GCTGGGGTTTTGTTTCACAAGAACG-TGGTTAGATACTCTTGTTC
isocontig35423      GACGGTTTTGTGGTTACTGGTTTTCGATCCCTCTCTGCTGCTGGTTTTGGGTATCTGGCT
                     **    * *** *  ****  **   *  * *      * *****         **

isocontig35424      TTGATTTTTCTCTCTT----CTTCTATTGATGAGATTTCTTTATATTCAATT--------
isocontig35423      GTACTGTTACTATCTTTTGACTTGGAAAGGTGGAGTCTTTGTGGTTTTCATGGGATATGG
                     *  * ** ** ****    ***  *  * **   * * * *   **  **

isocontig35424      --------------------------ACAGGGAACATGAAGGAATCTTTGAAAGTGGAGG
isocontig35423      AGGACGGGGGAATGGAAGGGAATAAAAAAGGGAACATGAAGGAATCTTTGAAAGTGGAGG
                                             * ********************************

                                                                        ^
isocontig35424      CAAAACTCTGGAAAACTTGAGATCTGCTGATGGATTGAATGAAGCTTTCTAGTTCAAATG
isocontig35423      CAAAACTCTGGAAAACTTGAGATCTGCTGATGGATTGAATGAAGCTTTCTAGTTCAAATG
                    ************************************************************
```

**Figure 3.3: Selected region of the nucleotide sequence alignment of isotig35423 and isotig35424**. The '*' represents a single, fully conserved residue whereas those without any symbol represents no consensus. The first putative 'atg' start codon is in bold and marked with (^) (CLUSTALW, Biology Workbench Version 3.2, University of California).


**Figure 3.4: Selected region of the nucleotide sequence alignment of isotig24120 and isotig24121.** The '*' denotes a single, fully conserved residue whereas those without any symbol denotes no consensus. The first putative 'atg' start codon is in bold and marked with (^). (CLUSTALW, Biology Workbench Version 3.2, University of California).

Multiple sequence alignment of isotig34176 and isotig34177 showed a 98% identity at the protein level (Table 3.4). Isotig34177 lacks twelve amino acids (Figure 3.5) or is 36 bases shorter (Figure 3.6) than isotig34176 within the coding region, possibly indicating the presence of an alternative splice variant of isotig34176. Protein motif search of the 12 amino acids did not result in any hits suggesting the possible function of this sequence is still unknown.

Isotig34176 is also 99.8% and 99.95% identical to *EgERD3* at both the protein and nucleotide levels (Figures 4.5 and 4.6). *EgERD3* is a putative ethylene receptor from oil palm that was isolated earlier based on RT-PCR using degenerate primers, 5` and 3` RACE approaches (Nurniwalis 2006). There is a single amino acid change between isotig34176 (R residue) and EgERD3 (K residue) caused by a base transition of G-A (Figure 3.5). The base substitution was also detected at the start of the 36 nucleotide gap in isotig34177 (Figure 3.6). The missense mutation in isotig34176/EgERD3 may or may not affect the function of the protein. For now it is not clear what caused the missense mutation but it may just have been a base mismatch due to sequencing error.

```
isotig34176      MEGCDCFEPQWPAEELLIKYQYISDFFIALAYFSIPLELIYFVKKSSFFPYRWVLVQFGA
EgERD3           MEGCDCFEPQWPAEELLIKYQYISDFFIALAYFSIPLELIYFVKKSSFFPYRWVLVQFGA
isotig34177      MEGCDCFEPQWPAEELLIKYQYISDFFIALAYFSIPLELIYFVKKSSFFPYRWVLVQFGA
                 ************************************************************

isotig34176      FIILCGATHLINLWTFTVHSKTVAIVMTVAKISTAAVSCATALMLVHIIPDLLSVKTREL
EgERD3           FIILCGATHLINLWTFTVHSKTVAIVMTVAKISTAAVSCATALMLVHIIPDLLSVKTREL
isotig34177      FIILCGATHLINLWTFTVHSKTVAIVMTVAKISTAAVSCATALMLVHIIPDLLSVKTREL
                 ************************************************************

isotig34176      FLKSKAEELDREMGLIRTQEETGRHVRMLTHEIRSTLDRHTILKTTLVELGRTLGLEECA
EgERD3           FLKSKAEELDREMGLIRTQEETGRHVRMLTHEIRSTLDRHTILKTTLVELGRTLGLEECA
isotig34177      FLKSKAEELDREMGLIRTQEETGRHVRMLTHEIRSTLDRHTILKTTLVELGRTLGLEECA
                 ************************************************************

isotig34176      LWMPSRSGSSLQLSHTLRHQITVGSTVPINHSIVNQVFSSSHAIIIPHACPLARIRPLAG
EgERD3           LWMPSRSGSSLQLSHTLRHQITVGSTVPINHSIVNQVFSSSHAIIIPHACPLARIRPLAG
isotig34177      LWMPSRSGSSLQLSHTLRHQITVGSTVPINHSIVNQVFSSSHAIIIPHACPLARIRPLAG
                 ************************************************************

isotig34176      RYVPPEVAAVRVPLLHLSNFQINDWPELSAKSYAVMVLMLPSDSARKWHIHELELVEVVA
EgERD3           RYVPPEVAAVRVPLLHLSNFQINDWPELSAKSYAVMVLMLPSDSARKWHIHELELVEVVA
isotig34177      RYVPPEVAAVRVPLLHLSNFQINDWPELSAKSYAVMVLMLPSDSARKWHIHELELVEVVA
                 ************************************************************

isotig34176      DQVAVALSHAAILEESMRARDLLMEQNIALDLARREAEMAIRARNDFLAVMNHEMRTPMH
EgERD3           DQVAVALSHAAILEESMRARDLLMEQNIALDLARREAEMAIRARNDFLAVMNHEMRTPMH
isotig34177      DQVAVALSHAAILEESMRARDLLMEQNIALDLARREAEMAIRARNDFLAVMNHEMRTPMH
                 ************************************************************

isotig34176      AIIALSSLLLETELTPEQRLMVETILKSSNLLATLINDVLDLSKLEDGSLELEIGPFNLH
EgERD3           AIIALSSLLLETELTPEQRLMVETILKSSNLLATLINDVLDLSKLEDGSLELEIGPFNLH
isotig34177      AIIALSSLLLETELTPEQRLMVETILKSSNLLATLINDVLDLSKLEDGSLELEIGPFNLH
                 ************************************************************

isotig34176      AVFREVMNLIKPIAAVKKLSVSVMLAPDLPLCAIGDEKRLMQTILNIAGNAVKFTKEGRI
EgERD3           AVFREVMNLIKPIAAVKKLSVSVMLAPDLPLCAIGDEKRLMQTILNIAGNAVKFTKEGRI
isotig34177      AVFREVMNLIKPIAAVKKLSVSVMLAPDLPLCAIGDEKRLMQTILNIAGNAVKFTKEGRI
                 ************************************************************

isotig34176      SLTASVAKPEYLRDIPDFCPVPSDRHFYLKVQVKDTGCGISPQDKAHLFTKFAQAQSGKN
EgERD3           SLTASVAKPEYLRDIPDFCPVPSDRHFYLKVQVKDTGCGISPQDKAHLFTKFAQAQSGKN
isotig34177      SLTASVAKPEYLRDIPDFCPVPSDRHFYLKVQVKDTGCGISPQDKAHLFTKFAQAQSGKN
                 ************************************************************

isotig34176      QGYSGSGLGLAICKRFVSLMEGHIWLESEGAGKGCTATFIVKLGTCENPIGFQQQVVPKA
EgERD3           QGYSGSGLGLAICKRFVSLMEGHIWLESEGAGKGCTATFIVKLGTCENPIGFQQQVVPKA
isotig34177      QGYSGSGLGLAICKRFVSLMEGHIWLESEGAGKGCTATFIVKLGTCENPIGFQQQVVPKA
                 ************************************************************

isotig34176      RPSHREADLSGPRALSKDEKGLARYQKSV
EgERD3           RPSHKEADLSGPRALSKDEKGLARYQKSV
isotig34177      RPSH------------KDEKGLARYQKSV
                 ****            *************
```

**Figure 3.5: Amino acid sequence alignment of isotig34176, isotig34177 and EgERD3**. The '*' denotes a single, fully conserved residue whereas those without any symbol denotes no consensus The highlighted region represents the location of the amino acid residues that are present in EgERD3 and isotig34176 but are absent in isotig34177 (CLUSTALW, Biology Workbench Version 3.2, University of California).

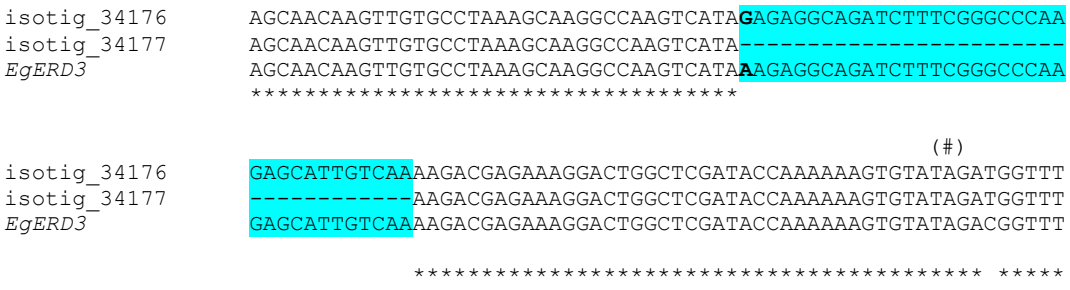



**Figure 3.6: Selected region of the nucleotide sequence alignment of isotig34176, Isotig34177 and *EgERD3*.** The '*' denotes a single, fully conserved residue whereas those without any symbol denotes no consensus. The highlighted region represents 36 nucleotides that are present in *EgERD3* and isotig34176 but are absent in isotig34177. The bold residue shows the only nucleotide that is not identical in *EgERD3* and isotig34176. The stop codon 'tag' is marked with (#). (CLUSTALW, Biology Workbench Version 3.2, University of California).

### 3.3.2 Genomic organization of the putative ethylene receptor genes

To further analyse the gene structure and to determine the location of the ten identified sequences at the genomic level, the nucleotide assembled transcripts were searched against the oil palm genome (*AVROS pisifera* fruit form) databases. Analyses were carried out on the various version of oil palm corresponds to the location of 10 putative genes have increased between 4 to 48 times more than the length in P3 scaffolds (Table 3.5). As the genome build progresses over the years, the sequence gaps found in between the scaffolds of the earlier genome build are narrowed down through the improved sequencing quality of all the sequence reads (Singh et al., 2013a). Isotig34176/isotig34177 and maker-p5_sc00069-augustus-gene-6.47-mRNA-1 which was initially located at different scaffolds in P3, P4 and P5 builds are now found on the same genetic scaffolds that is chromosome5 in the EG6 build (Table 3.5).

The overall analysis of the exon-intron structures performed using both NCBI-SPIDEY and exonerate detected the presence of introns flanked by at least two exons. Within the ORFs of the 10 assembles transcripts, the exon-intron structure can be clustered into two groups. The first group which is composed of isotig34176, isotig34177, isotig45619 and maker-p5_sc00025-augustus-gene-13.28-mRNA-1 contains 5 exons and 4 introns while the second group (isotig24120, isotig24121, isitig35423, isotig35424, isotig44932 and p5_sc00069-augustus-gene-6.47-mRNA-1) contains 2 exons and 1 intron (Figure 3.7).

**Table 3.5: Location of the putative ethylene receptor genes within the oil palm genome builds.**

| Transcript/Gene Model ID | Assembled Transcript size | Scaffold position, length (bold) and genomic coordinates [ ] of assembled transcript | | | |
|---|---|---|---|---|---|
| | | P3 _assembly (P3) | Pisifera4 (P4) | Pisifera5_orphan (P5) | Eguineensis6 (EG6) |
| Isotig34176 | 2479 | p3_sc00258 **1071840** [32876-19754] | p4_sc00525 **528422** [105033-91166] | p5_sc00102 **3028108** [1122640-1136682] | EG6_Chr5 **51430340** [14482643-14496169] |
| Isotig34177 | 2441 | p3_sc00258 **1071840** [32876-19754] | p4_sc00525 **528422** [105033-91166] | p5_sc00102 **3028108** [1122640-1136682] | EG6_Chr5 **51430340** [14482643-14496169] |
| Isotig45619 | 2565 | p3_sc00004 **8361249** [1560567-1553602] | p4_sc00006 **7978889** [2842007-2835040] | p5_sc00009 **6400178** [1467416-1460450] | EG6_Chr10 **31665614** [16922042-16929011] |
| maker-p5_sc00025-augustus-gene-13.28-mRNA-1 | 2469 | p3_sc00001 **9346513** [5420101-5426459] | p4_sc00016 **5625427** [3137375-3143732] | p5_sc00025 **5563611** [1328608-1322252] | EG6_Chr2 **64882328** [58959427-58953071] |
| Isotig24120 | 3607 | p3_sc00120 **1778787** [1209101-1214732] | p4_sc00069 **2960420** [1178295-1183926] | p5_sc00085 **3193001** [1188645-1194276] | EG6_Chr4 **56584025** [30947729-30942097] |
| Isotig24121 | 3051 | p3_sc00120 **1778787** [1208775-1214732] | p4_sc00069 **2960420** [1177969-1183926] | p5_sc00085 **3193001** [1188319 -1194276] | EG6_Chr4 **56584025** 30948055-30942097] |
| Isotig35423 | 3195 | p3_sc00336 **831682** [731238-728381] | p4_sc01706 **130282** [76459-73602] | p5_sc00193 **1359181** [289951-282566] | EG6_Chr11 **29613423** [5800949-5793557] |
| Isotig354234 | 3032 | p3_sc00336 **831682** [730439-728381] | p4_sc01706 **130282** [75660-73602] | p5_sc00193 **1359181** [289152-282566] | EG6_Chr11 **29613423** [5800150-5793557] |
| Isotig44932 | 4186 | p3_sc00024 **4081840** [907234-901900] | p4_sc00037 **4094324** [3191514-3196847] | p5_sc00134 **1851254** [944887-950220] | EG6_Chr12 **28572085** [24826754-24821419] |
| maker-p5_sc00069-augustus-gene-6.47-mRNA-1 | 3187 | p3_sc00048 **3098654** [2556873-2551091] | p4_sc00062 **3082920** [541688-547472] | p5_sc00069 **3153946** [579479-585261] | EG6_Chr5 **51430340** [580779-586567] |

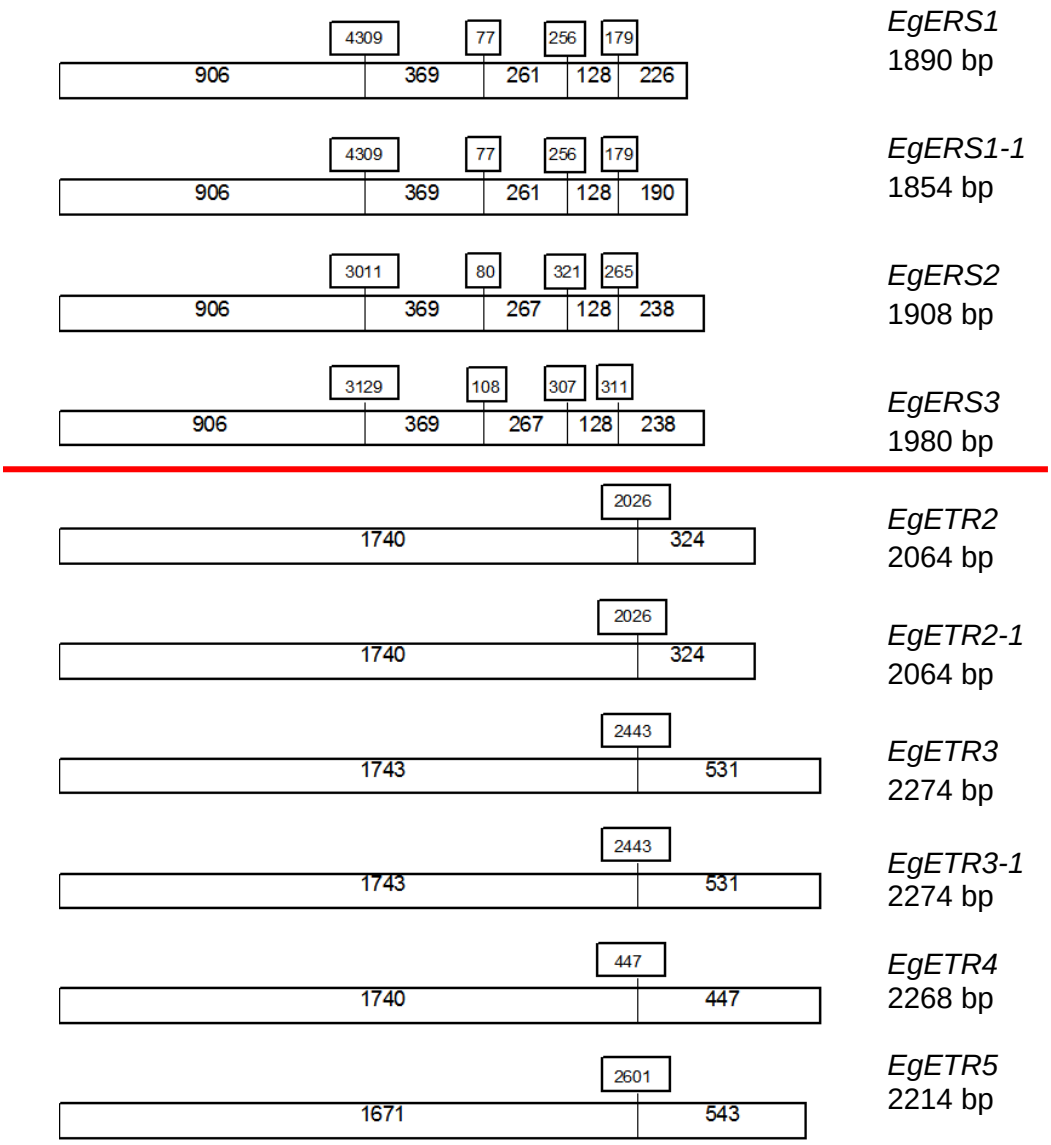



**Figure 3.7: A schematic representation of the genomic structure of the ethylene receptor gene family within the ORF**. Introns are represented by boxes that are flanked by exon boxes. Numbers represent the size of the exon/intron structure. The intron-exon structure of the ethylene receptor genes are divided into two groups based on the number of exons and introns within the ORF.

In both groups, the first exon in both subfamilies represents the largest exon size ranging from 906 to 1743, respectively (Figure 3.7). Presence of introns were also detected upstream from the start codon sites (ATG) in most of the assembled transcripts except for isotig24120, isotig35424, maker-p5_sc00025-augustus-gene-13.28-mRNA-1 and p5_sc00069-augustus-gene-6.47-mRNA-1. The largest intron size within this region was detected in isotig34176 and isotig34177 with an approximate size of 6226 bp whereas the smallest intron size of 634 bp was detected in isotig35423. Multiple sequence alignment analyses between isotig35423 and isotig35424 as well as isotig24121 and isotig 24120 suggests that the possible intron sequences were retained in isotig24120 and isotig35424 within the 5`untranslated region (UTR).

A schematic diagram to represent the uniqueness of the sequence and comparison to isotig24121 and isotig35423 is shown in Figure 3.8. Isotig35423 contains an intron with the size of 634 bases while isotig34524 retained a 224 bases intron. There is a gap of 410 nucleotides between the two introns. Similarly, the retained intron size in isotig24120 is 723 bases, the intron in isotig24121 is 882 bases and the gap between isotig24121 and isotig24120 is 159 bases. Nucleotide sequence translation of the retained intron only produced short proteins which is comprised between 2 up to 54 amino acid residues. Protein motif search of these amino acids did not result in hits to any conserved domains suggesting the possible function of this sequence is still unknown. These results also suggest that besides isotig34177, isotig24120

and isotig35424 are possibly splice variants of isotig24121 and isotig35423, respectively.

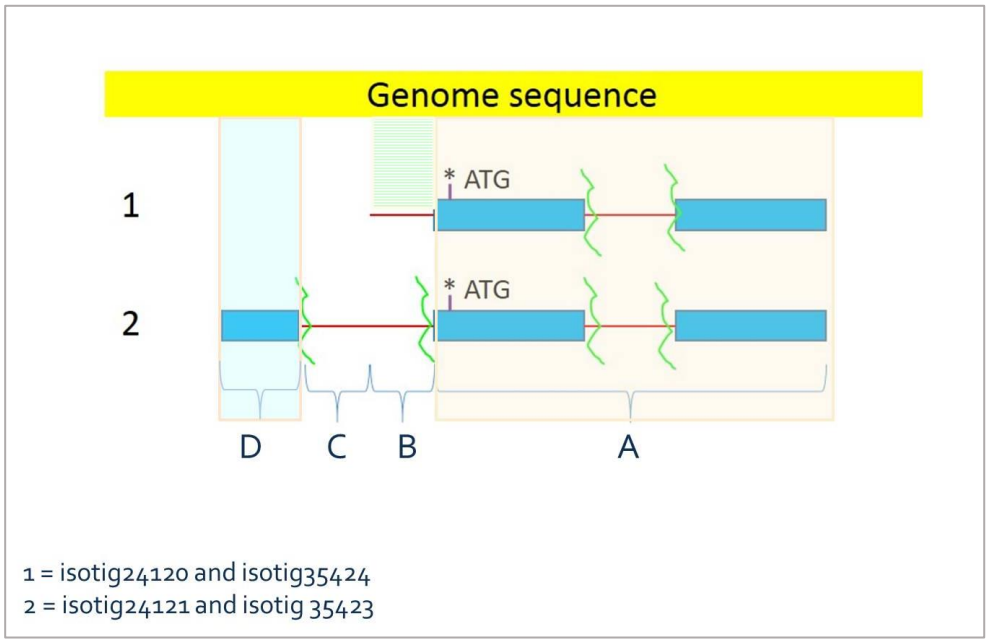



**Figure 3.8: A schematic gene structure representation at the 5`UTR region of a putative ethylene receptor and its splice variant based on their exon (▭) and intron (—) boundary region using the SPIDEY-NCBI programme.** ≀ = splice site, *ATG = start codon, 1 = isotig24120 and isotig35424, 2 = isotig24121 and isotig35423, A = exons and intron that are 100% identical at the nucleotide level in both number 1 and 2 set of genes, B = intron that is retained in isotig24120 and isotig35424 and is 100% identical to the genome sequence and C= gap between the introns in isotig24121 and isotig35423 and the retained intron in isotig24120 and isotig35424. D = exon region of isotig24121 and isotig 35423 that is 100% similar to the genome sequence.

Nucleotide sequence analyses of the assembled transcripts beyond the stop codon did not detect any presence of introns except for isotig35423 and isotig35424. A 1120 bp sized intron was detected 368 nucleotide downstream from the stop codon. In maker-p5_sc00025-augustus-gene-13.28-mRNA-1 and p5_sc00069-augustus-gene-6.47-mRNA-1, the presence or absence of intron within the 5` and 3` UTRs could not be justified based on the prediction of the ORF from the oil palm gene model. Thus, in order to explore these options, isolation of the genes were carried out to define the sequence information and the results are presented and discussed in section 3.3.3. With the genomic structure information of the 10 assembled transcripts, the transcripts from here onwards are referred to the designated gene name as listed in Tables 3.1 and 3.2, respectively.

### 3.3.3 Molecular cloning of the putative ethylene receptor genes

#### 3.3.3.1 5` Rapid amplification of cDNA ends

The amplified 5` RACE PCR products are shown in Figure 3.9. Table 3.6 represents the size information of the amplified 5` PCR products ranging from 194 bp - 697 bp. The amplified PCR products from the seven putative ethylene receptor genes are comprised of the 5` ORF including the start codon and the 5` UTRs. The size of the 5` UTRs of the seven putative genes ranged between 120 bp - 493 bp.

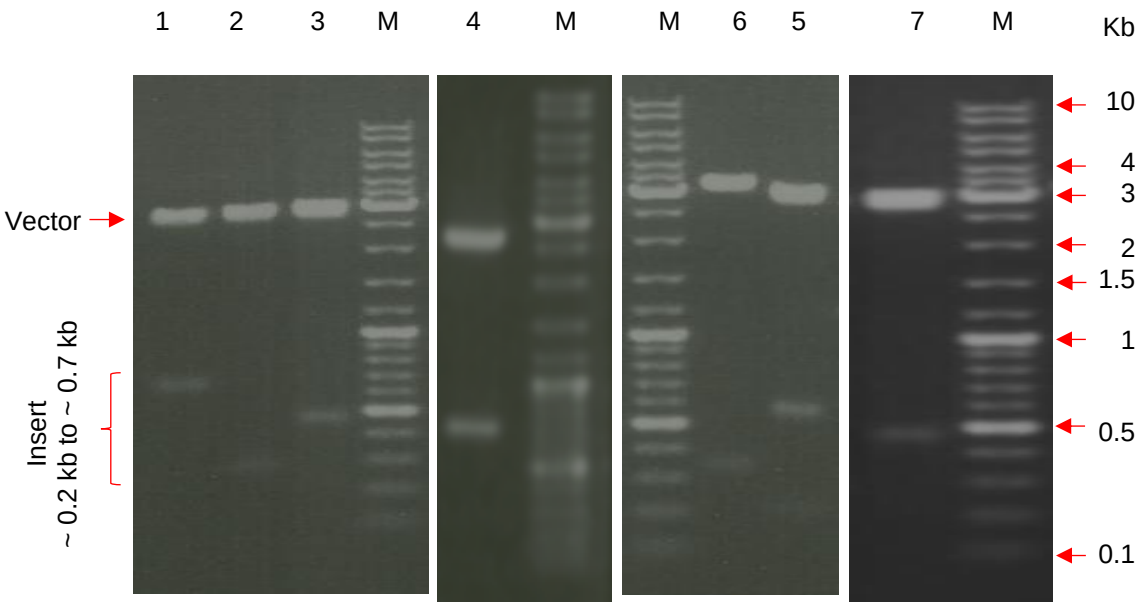



**Figure 3.9: Analysis of the digested plasmids carrying the 5` RACE PCR products by electrophoresis on 1% (w/v) agarose gel.** The plasmids were digested with *EcoRI*. Lanes 1 to 7 represent the digested products of pEgERS1, pEgERS2, pEgERS3, pEgETR2, pEgETR3, pEgETR4 and pEgETR5, respectively. M represents Gene Ruler DNA Ladder Mix (Fermentas).

**Table 3.6: Size characteristics of the amplified RACE PCR products and introns in the UTRs of the putative ethylene receptors from the oil palm genome.**

| Gene name / ID | Amplified 5` PCR size (bp) | | | Intron size within 5` UTR (bp) | Amplified 3` PCR size (bp) (including poly A tail) | | | Intron size within 3` UTR (bp) |
|---|---|---|---|---|---|---|---|---|
| | ORF | 5` UTR | Total | | ORF | 3` UTR | Total | |
| *EgERS1* | 420 | 177 | 697 | 6226 | 714 | 265 | 1007 | None |
| *EgERS2* | 23 | 171 | 194 | 727 | 338 | 287 | 653 | None |
| *EgERS3* | 244 | 169 | 413 | 768 | 199 | 406 | 605 | None |
| *EgETR2* | 433 | 120 | 573 | 882 | 292 | 307 | 623 | None |
| *EgETR3* | 74 | 493 | 567 | 634 | 101 | 222 | 351 | None |
| *EgETR4* | 72 | 234 | 306 | 700 | 151 | 1314 | 1493 | None |
| *EgETR5* | 166 | 174 | 197 | None | 163 | 768 | 931 | None |

Manual sequence alignment and exonerate search of the 5` RACE PCR products against the oil palm EG6 build showed a 100% identity at the nucleotide level. In addition, the presence of introns was detected in six out of the seven putative sequences (Table 3.6). Five of the introns are present in *EgERS1, EgERS2, EgETR2, EgETR3* and *EgETR4* just as identified earlier in section 3.3.2 based on the overall analysis of the exon-intron structures using both NCBI-SPIDEY and exonerate. An additional 768 bp size intron which was not identified earlier due to insufficient nucleotide information is present in the 5` UTR of *EgERS3*. In all the six PCR fragments, the intron-exon boundaries follows the universal intron 'GT-AG' rule (Breathnach and Chambon 1981). No intron was found in the 5` UTR region of *EgETR5*, which was confirmed *via* 5` RACE. Illustration of the genomic structure of the ethylene receptor genes is shown in Figure 3.7. It has been reported that at least one intron is present in the 5` UTR of many plant ethylene receptors. For example, in Arabidopsis, the receptors contain a single intron in their 5` UTR (Chang et al. 1993; Hua et al. 1998; Hua and Meyerowitz 1998; Sakai et al. 1998) whereas the ethylene receptor *Nr* in tomato contains two introns in the 5' UTR sequence (Moussatche 2004). Apparently, intron occurrence within the UTR is common in plants. The number of introns present within the 5` UTR is normally two times higher than that in the 3` UTR (Hong et al. 2006). Intron presence and size within the 5`UTR have been shown to enhance the effect on gene expression as demonstrated in the *EF1α-A3* gene from Arabidopsis (Chung et al. 2006).

Sequence analysis of all seven putative genes showed that nucleotides flanking the ATG translation start site is adequate to the Kozak preferred nucleotide sequences (ANNATGG) (Nakagawa et al. 2008). It was also observed that the arrangement of the nucleotides flanking the start codon can be divided into two following the pattern of the ethylene receptor gene subfamilies (Table 3.7).

**Table 3.7: Nucleotides flanking the transcription start site (TSS) of the ethylene receptor genes from the oil palm genome.** The TSS is underlined and is indicated by the notation +1, is either A or G. Most of the nucleotides flanking the TSS are pyrimidines, just as found in plant genes.

| Gene ID | Transcription start site |
|---|---|
| *EgERS1* | T<u>G</u>A |
| *EgERS2* | C<u>A</u>T |
| *EgERS3* | C<u>A</u>C |
| *EgETR2* | C<u>A</u>G |
| *EgETR3* | T<u>A</u>C |
| *EgETR4* | G<u>G</u>A |
| *EgETR5* | T<u>G</u>A |

### 3.3.3.2 3` Rapid amplification of cDNA ends

Figure 3.10 presents the digested plasmids harboring the 3` end fragments from all seven putative ethylene receptor genes. Sequence analyses of the plasmids showed that the size of the inserts ranged from 351 bp up to 1493 bp (Table 3.6). The amplified PCR products contained the 3` ORF region including the stop codon and the 3` UTRs. In addition, the poly $(A)^+$ tails were detected in all fragments, confirming that the 3` ends were successfully obtained. The plant consensus putative polyadenylation signal (AATAAA) involved in facilitating the addition of poly$(A)^+$ tail to produce mature messenger mRNA (Joshi 1987a) were found in the 3` UTR in 5 of the putative genes except *EgERS1* and *EgETR3*. Both *EgERS1* and EgETR3 contained 4-5 of 6 bases match of the plant consensus signal/motif, usually found in the 3' UTR region in other plants (Hunt 1994). Sequence alignment of the amplified PCR products to the oil palm genome data showed 100% identity at the nucleotide level. Introns were not detected in the 3`UTRs in any of the putative ethylene receptor genes which mirrors the situation of other established ethylene receptor gene families such Arabidopsis (Schaller and Kieber 2002), tomato (Tieman and Klee 1999) and rice (Yau et al. 2004). This result contradicts the earlier analysis as described in section 3.3.2 where isotig35423 or *EgETR3* and isotig35424 or *EgETR3-1* contains an intron after the stop codon, presumably within the 3` UTR. As isotig35423 is comprised of assembled sequence data, a computational issue may have occurred and this might have possibly affected the final assembly of the transcript data.

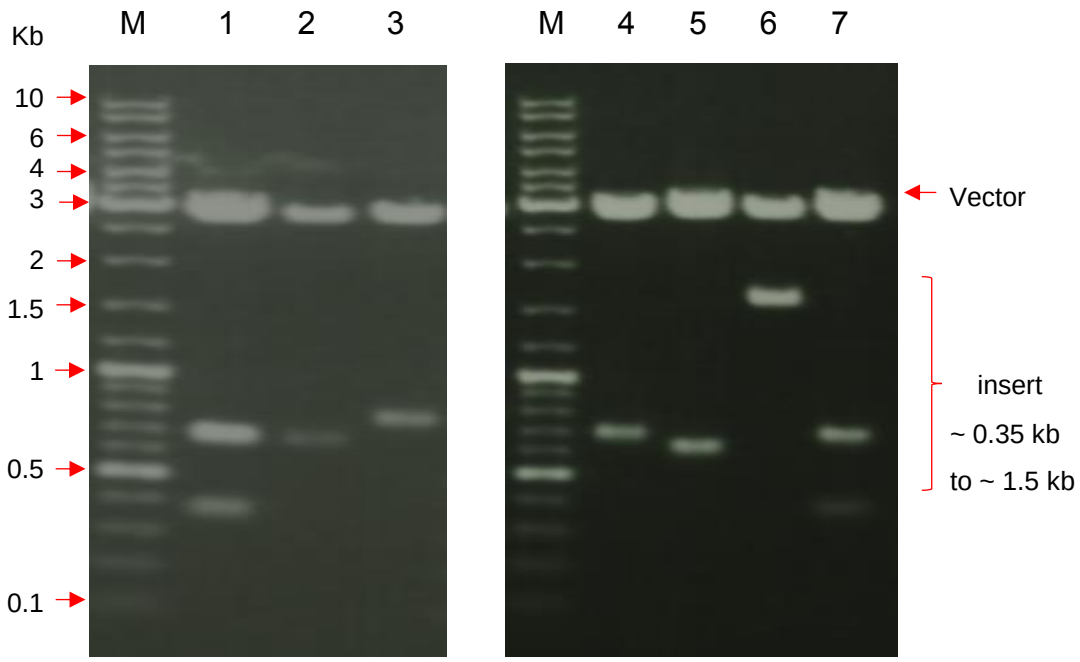



**Figure 3.10: Analysis of the digested plasmids carrying the 3` RACE PCR products on 1% (w/v) agarose gel electrophoresis**. The plasmids were digested with *EcoRI.* Lanes 1 to 7 represent the digested products of pEgERS1, pEgERS2, pEgERS3, pEgETR2, pEgETR3, pEgETR4 and pEgETR5, respectively. M represents Gene Ruler DNA Ladder Mix (Fermentas).

### 3.3.3.3 LD-PCR

The reconstitute sequence of the 5`, 3` PCR products, oil palm transcriptome and genome data indicated that the full-length cDNAs encoding the ethylene receptor genes were obtained. Sequence analyses confirmed the output of the combined sequence information from the earlier findings and this information is now available in Genbank with accession numbers XM_0190923292 = *EgERS1,* XM_010933391 = *EgERS2,* XM_010916199 = *EgERS3*, XM_010920466 = *EgETR2*, XM_010934797 = *EgETR3*, XM_010937464 = *EgETR4* and XM_010922209 = *EgETR5*, respectively. Information on the full-length (FL) putative ethylene receptor genes is presented in Table 3.8.

**Table 3.8: List of the full-length sequence information of the oil palm ethylene receptors**

| Gene ID | Total FL size (bp) | ORF (bp) | 5` UTR (bp) | 3`UTR (bp) | Poly A tail (bp) |
|---|---|---|---|---|---|
| *EgERS1* | 2360 | 1890 | 177 | 265 | 28 |
| *EgERS2* | 2394 | 1908 | 171 | 287 | 28 |
| *EgERS3* | 2507 | 1908 | 169 | 406 | 24 |
| *EgETR2* | 2728 | 2274 | 120 | 307 | 27 |
| *EgETR3* | 3007 | 2264 | 493 | 222 | 28 |
| *EgETR4* | 3844 | 2268 | 234 | 1314 | 28 |
| *EgETR5* | 3177 | 2214 | 174 | 768 | 21 |

### 3.3.4 Amplification of oil palm ethylene receptor splice variants

BLAST searches against the oil palm transcriptome, gene model and genome databases have identified three assembled isotigs (isotig35424, isotig24120 and isotig34177) that are possibly splice variants to three of the putative ethylene receptor genes. Thus, these splice variants are designated as *EgETR3-1, EgETR2-1* and *EgERS1-1* following the corresponding putative gene ID. Splice variants of the ethylene receptor genes have been detected in Arabidopsis (TAIR), [https://www. arabidopsis.org/servlets/Search?type =general&search_action=detail&method=1&show_obsolete=F&name=ethylene+receptor&sub_type=gene&SEARCH_EXACT=4&SEARCH_CONTAINS=1], on www.arabidopsis.org, (20$^{th}$ May 2013) and rice (Pareek et al. 2006). Thus, to distinguish the presence and expression of these splice variants, PCR amplification and RT-PCR was performed on two of the splice variants. The forward primers were designed to the unique regions of the variants and the reverse primers were designed to the common region present in the corresponding ethylene receptors.

For isotig35424, two primer pair combinations were used which resulted in the amplification of genomic DNA fragments with the size of ~350 bp and ~450 bp (Figure 3.11). The size of the RT-PCR products from mesocarp 20 WAA and spear leaves also matched to that of the genomic DNA. Similarly, the splice variant for *EgERS1*, namely isotig34177 or *EgERS1-1* was also amplified from both templates (genomic DNA and cDNA) to produce a PCR product with the size of ~290 bp (Figure 3.12). The success in amplifying both splice variants

indicates the accuracy of the database information and that the splice variants are expressed in both the tested oil palm tissues.

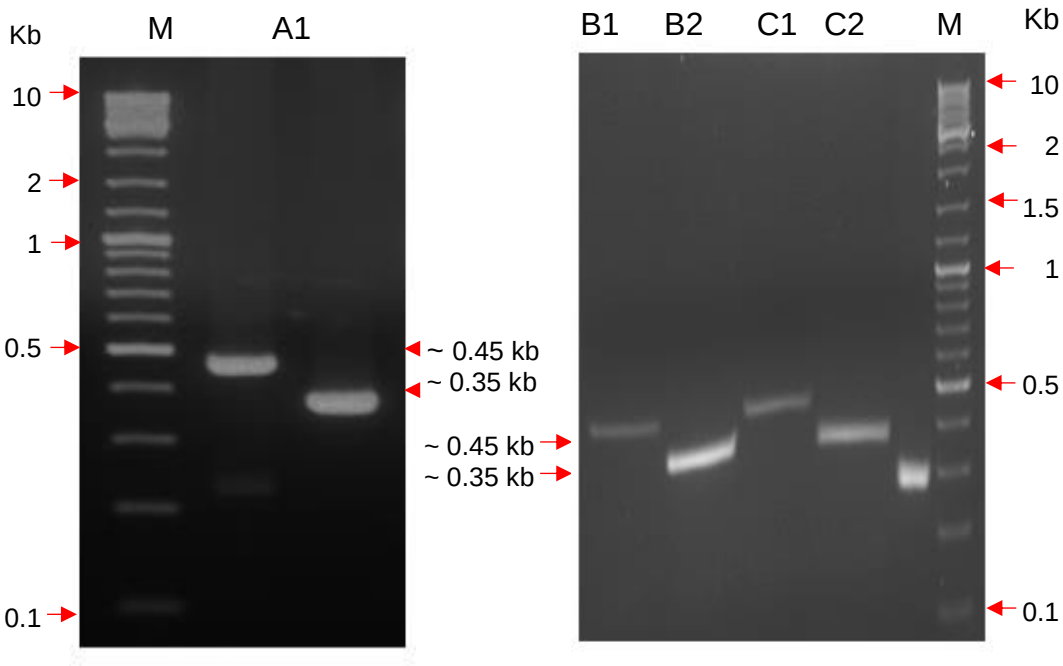



**Figure 3.11: Analysis of the amplified isotig35424 or *EgETR3-1*, a splice variant of *EgETR3* by on electrophoresis on 2.5% (w/v) agarose gel.** A (DNA), B (cDNA from mesocarp 20WAA) and C (cDNA from spear leaves) represents the templates used in the PCR amplifications. Numbers represent the primer combinations used in the PCR amplification where 1 = 3542SVF1 & 3542R2, 2 = 3542SVF2 & 3542R2 and M = GeneRuler DNA Ladder Mix (Fermentas).

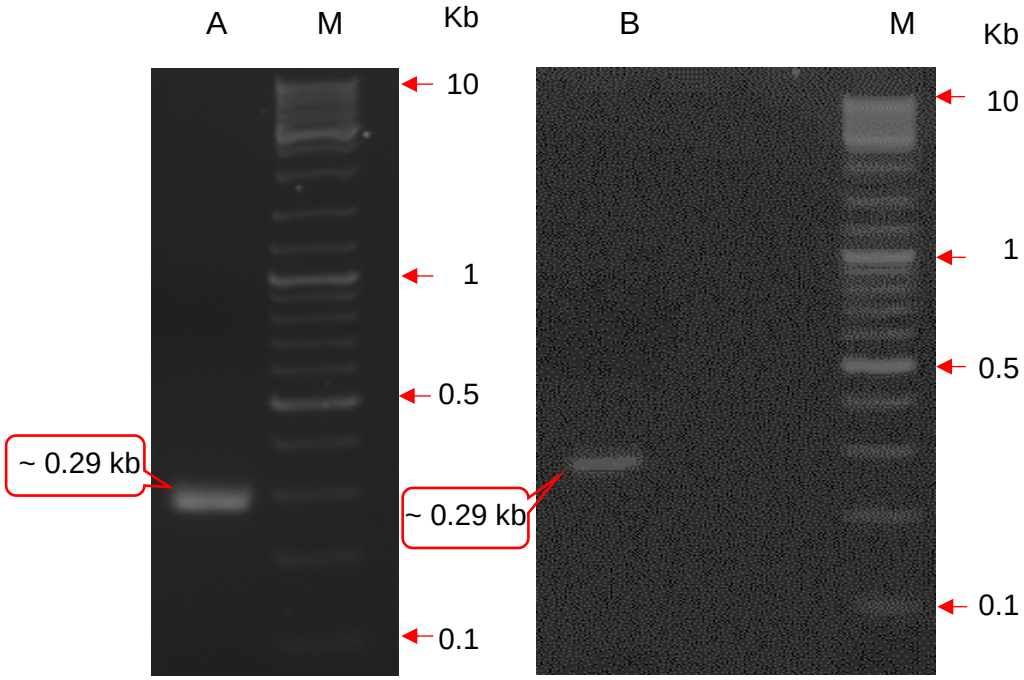



**Figure 3.12: Analysis of the amplified isotig34177 or *EgERS1-1*, a splice variant of *EgERS1* by electrophoresis on 2.5% (w/v) agarose gel electrophoresis.** A (DNA) and B (cDNA from mesocarp 20WAA) represents the templates used in the PCR amplifications. M = GeneRuler DNA ladder Mix (Fermentas).

### 3.3.5 Conserved domain, structural and sequence analysis of the ethylene receptor proteins

IPRScan programme and SMART software was used to identify the conserved domains present in the oil palm ethylene receptor genes. Five major domains spanning from the amino (N-) terminal until the carboxyl (C-) terminal were detected comprising transmembrane, GAF, Histidine kinase A (HisKA), histidine kinase-like ATPase (HATPase_c) and REC (cheY-homologous receiver). However, not all of the five conserved domains are present in all of the ethylene receptor proteins. The structural comparison of the seven ethylene receptor proteins is illustrated in Figure 3.13.

EgETR3 and EgETR4 are the only two proteins that contain all five conserved domains. The remaining five proteins contain four of the major conserved domains but only four proteins consist of the same conserved region that are the transmembrane, GAF, HisKA and HATPase_c. EgETR5 however contains the REC domain instead of the HisKA domain. The REC domain is only found at the C-terminal of EgETR3, EgETR4 and EgETR5. Ethylene receptors that contain the REC or receiver domain are normally called ETR-like and/or EIN4-like receptors while those that lack the receiver domains are normally called the ERS-like ethylene receptors (Chang et al. 1993; Hua et al. 1998; Yau et al. 2004). In comparison to ethylene receptors from other plants, those that carry the receiver domain are normally members of subfamily II of the ethylene receptor family (Hua et al. 1998; Yau et al. 2004).

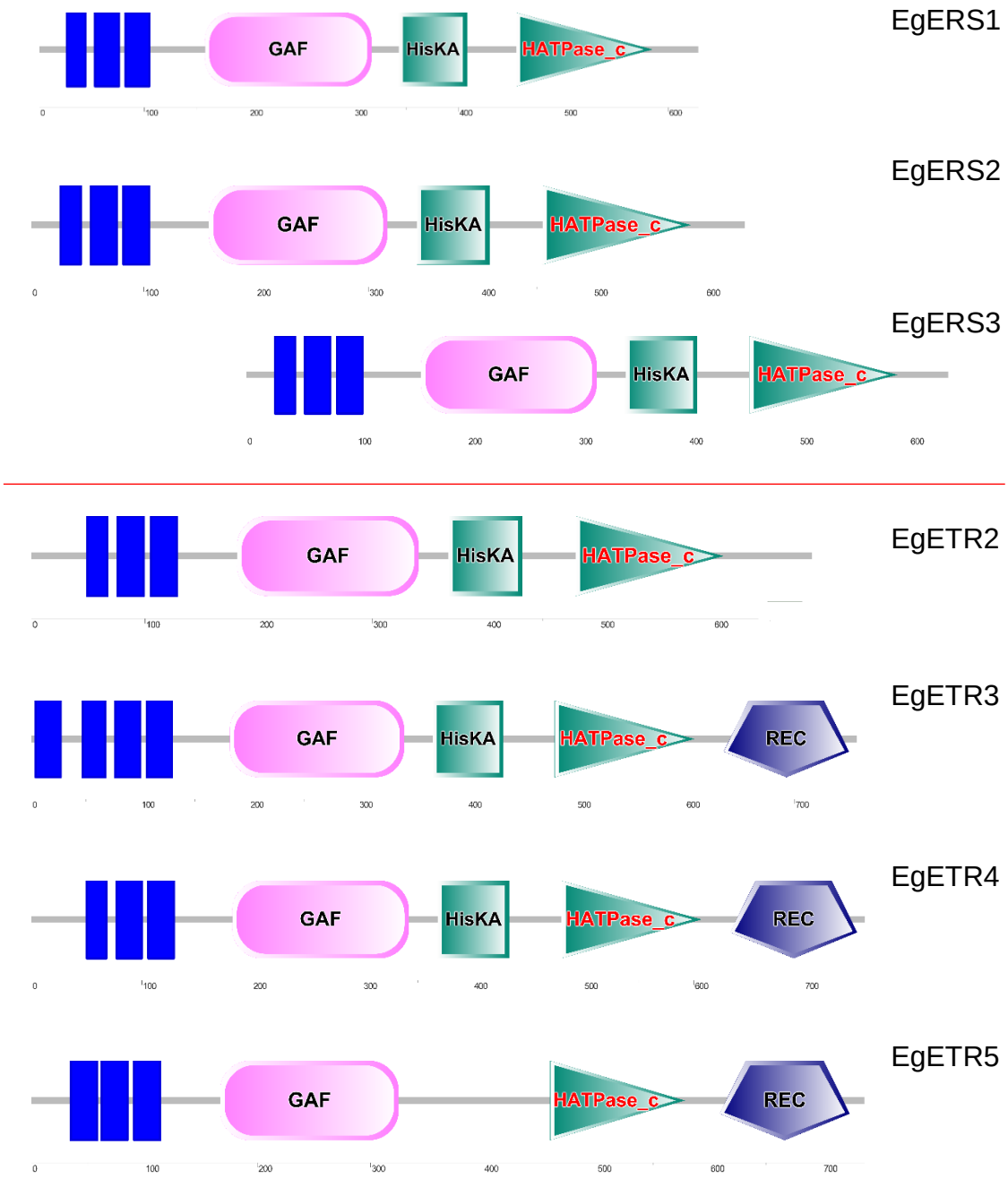



**Figure 3.13: Primary protein structure of the putative ethylene receptor genes in oil palm using the SMART software (http://smart.embl-heidelberg.de/).**

The transmembrane domain located at the N-terminal of the ethylene receptor proteins has three transmembrane-spanning segments. Hydropathy analysis using Kyte-Doolittle hydropathy profile also suggests that the triple membrane at the N terminal is hydrophobic (Figure 3.14). The transmembrane domain is also known as the sensor domain and is responsible for the binding of ethylene *via* the disulfide linkage and localization to the endoplasmic reticulum (Chen et al. 2002; Grefen et al. 2008). EgETR3 is the only protein that has a fourth membrane-spanning segment located at the far-end of the N terminal. A putative signal peptide was detected within the fourth membrane-spanning segment of EgETR3 and is thought to be important for endoplasmic reticulum (ER) membrane targeting (Ju and Chang 2012). The additional transmembrane segment within the transmembrane domain is also a characteristic of the subfamily II of the ethylene receptor family (Ju and Chang 2012). The detailed analysis of the location of the conserved regions within the ethylene receptor proteins is presented in Table 3.9.

Protein sequence correlation between the seven ethylene receptors is demonstrated in Table 3.4. It was observed that the ethylene receptors without the receiver domains are more identical to each other except for EgETR2. The identity between EgERS1, EgERS2 and EgERS2 is between 81% to 87% while the identity to EgETR2 is much lower at only 38% to 40%. Nonetheless, EgETR2 has a much higher identity between 72% and 85% to EgETR3 and EgETR4 although it lacks the receiver domain. Comparison of the three ethylene receptors with receiver domains (EgETR3, EgETR4, and EgETR5)

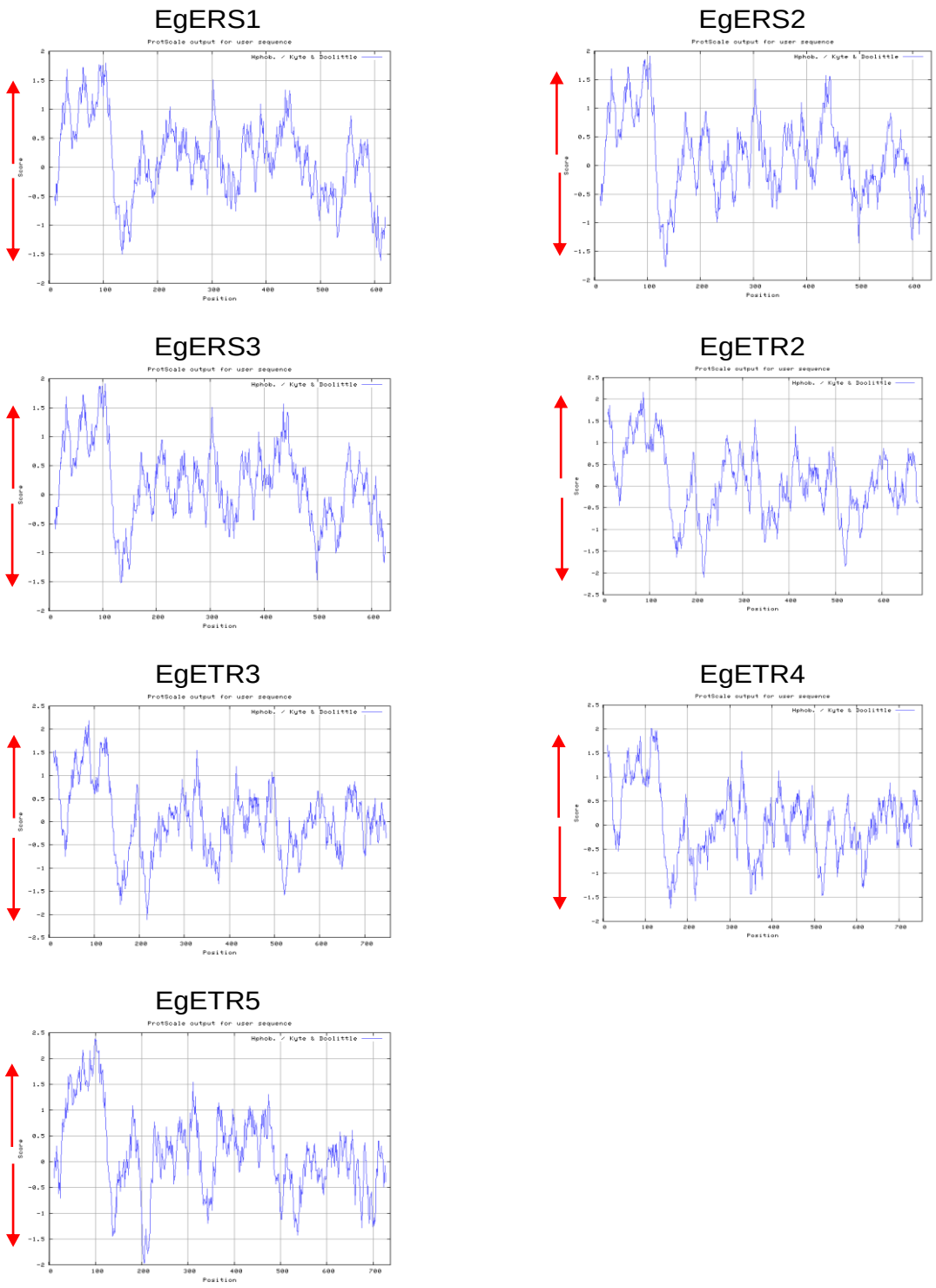



**Figure 3.14: Kyte-Doolittle hydropathy profile for the ethylene receptor predicted proteins from oil palm** (http://web.expasy.org/protscale). Arrows represent the level of hydropathy of the FLL1 protein ( ↑ = hydrophobic and ↓ = hydrophilic)

**Table 3.9: Analysis of the conserved domain locations in the ethylene receptor gene family in oil palm based on InterPro protein sequence analysis (http://www.ebi.ac.uk/interpro/).**

| Gene ID | Amino acid | Conserved domains | | | | |
|---|---|---|---|---|---|---|
| | | TM | GAF | HisKA | HATPase_c | REC |
| *EgERS1* | 629 | 26 – 43<br>53 – 75<br>82 – 104 | 158 – 317 | 343 - 408 | 455 - 585 | |
| *EgERS2* | 635 | 26 – 43<br>53 – 75<br>82 – 104 | 158 – 317 | 343 - 408 | 455 – 587 | |
| *EgERS3* | 635 | 26 – 43<br>53 – 75<br>82 – 104 | 158 - 317 | 343 - 408 | 455 – 589 | |
| *EgETR2* | 687 | 49 - 66<br>76 - 98<br>105 -127 | 181 - 341 | 367 - 432 | 479 - 609 | |
| *EgETR3* | 757 | 4 – 26<br>47 – 67<br>77 – 99<br>106 - 128 | 182 – 342 | 182 – 342 | 480 – 608 | 635 – 750 |
| *EgETR4* | 755 | 50 – 67<br>77 – 99<br>106 – 128 | 182 – 342 | 182 – 342 | 480 – 607 | 633 – 748 |
| *EgETR5* | 737 | 354 – 57<br>62 – 84<br>91 - 113 | 167 – 325 | | 459 – 578 | 609 – 724 |

* TM = transmembrane, HisKA = Histidine kinase A, HATpase_c = histidine kinase-like ATPase, REC = cheY-homologous receiver.

showed that EgETR3 and EgETR4 are more identical with 71% identity unlike EgETR5. The identity between EgETR2, EgETR3 and EgETR4 to EgER5 is between 44% and 45%. Despite the low identity, the identity of EgETR5 to EgERS1, EgERS2 and EgERS3, the receptors that do not contain the receiver domain, is even lower with only 39% to 41%.

Multiple sequence alignments of all seven putative ethylene receptor proteins were clustered into two groups as shown in Figure 3.15. The phylogenetic tree showed that EgERS1, EgERS2 and EgERS3 belong to the same group whereas EgETR2, EgETR3, EgETR4 and EgETR5 are clustered together in another group. The sub-division of the seven putative ethylene receptors into two groups conforms to the general pattern of ethylene receptor family in plants. In addition, the genomic structure of the receptors in each group also resembles that of the members within the subfamilies of the ethylene receptor family in other plants (Figure 3.7). The subfamily I member normally contains at least four introns and is flanked by five or more exons whereas subfamily II members have only one intron within the ORFs ( Hua et al. 1998; Yau et al. 2004). The homology between members of each subfamily is also higher (Table 3.4). As EgETR3 has a characteristics of a subfamily II member of the ethylene receptor family, therefore the rest of the receptors clustered together belong to the same subfamily. The remaining three receptors; EgERS1, EgERS2 and EgERS3 belong to the ethylene receptor subfamily I.

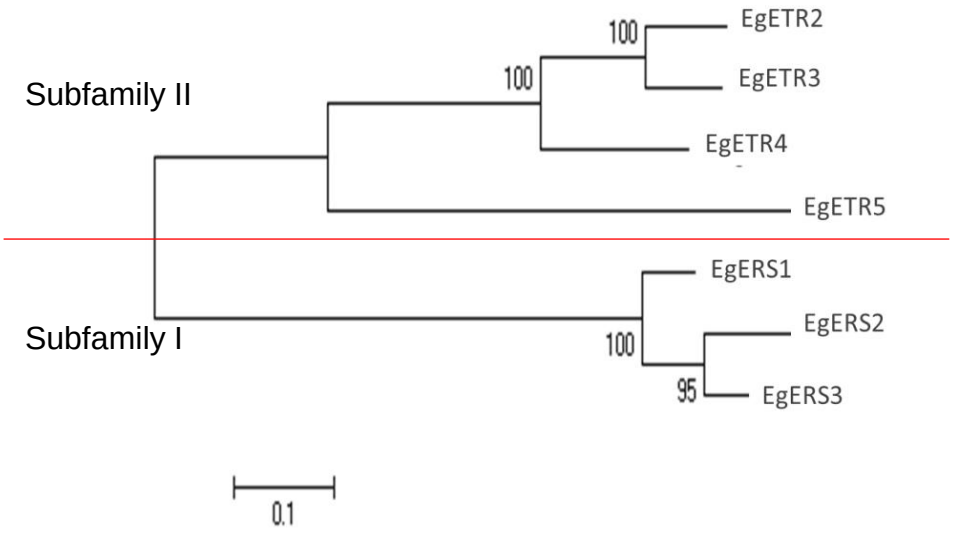


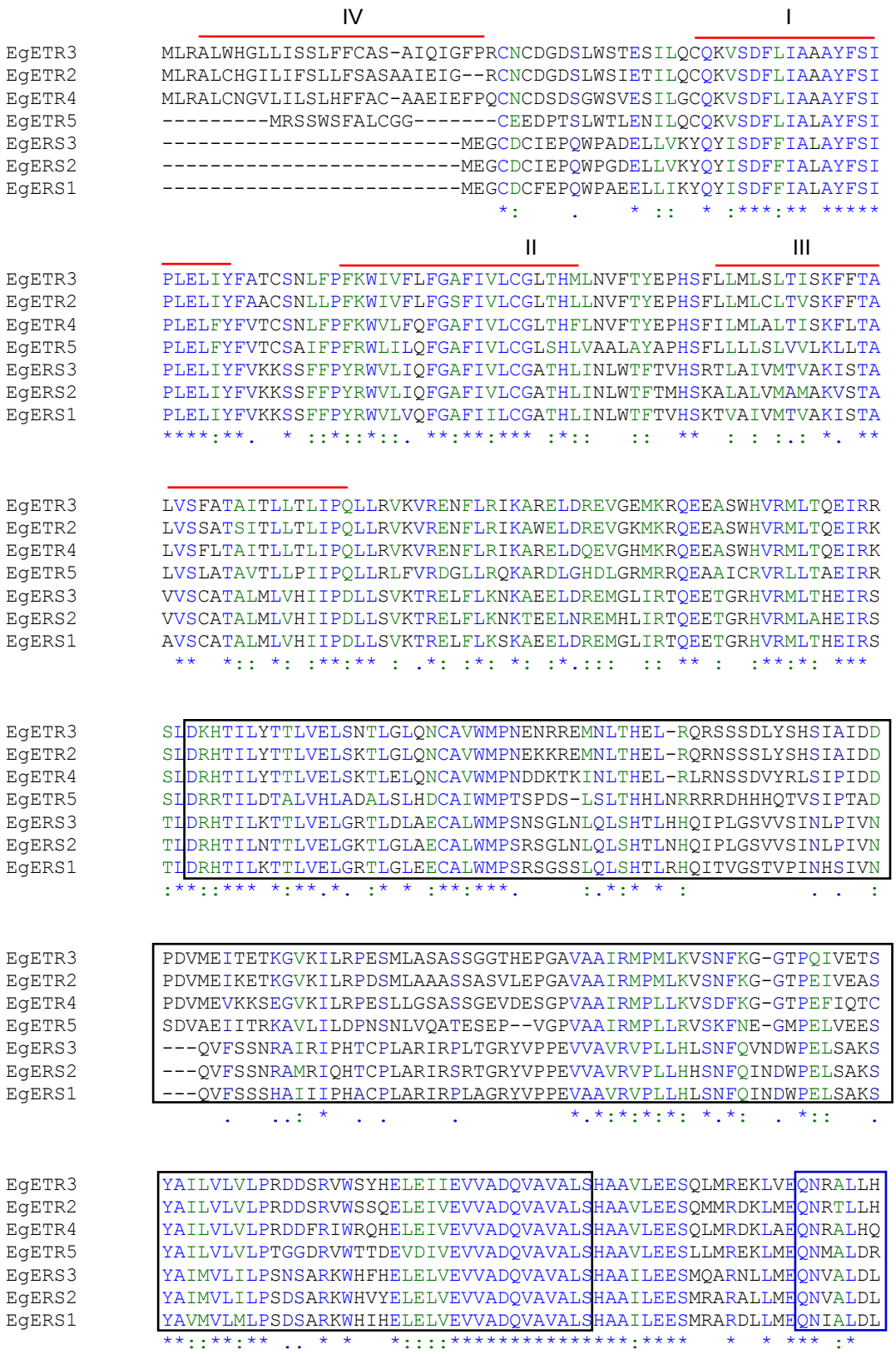

**Figure 3.15: Phylogenetic relationship of the seven putative ethylene receptor genes from oil palm.** Protein sequences were aligned based on multiple sequence alignment using Clustal W programme by Neighbour-joining method with 1000 bootstrap replicates. The ethylene receptor genes are classed into two groups. The bar indicates 0.1 substitutions per site.

The sequence alignments also showed that apart from the conserved transmembrane spanning domain, other domains are much more conserved within the members of each subfamily (Figure 3.16). The GAF domain located next to the transmembrane domain is thought to act as a connector between the sensor and the histidine kinase domain of the ethylene receptor gene (Chang et al. 1993). The name GAF signifies its presence in cGMP-specific phosphodiesterases, certain adenylate cyclases and the *E. coli* transcription factor, formate hydrogen lyase transcriptional activator (Fh1a) (Ciardi and Klee 2001). The exact role of the GAF domain in the function of the ethylene receptor

is still not clear but it is thought to facilitate protein-protein interactions among the receptors (Grefen et al. 2008).

The histidine kinase domain, located at the C-terminal of subfamily I members (EgERS1, EgERS2 and EgERS3) is highly conserved. All five conserved motifs responsible for the histidine kinase activity, especially the histidine (H) residue that is essential for the histidine kinase activity are present (Figure 3.16). It is similar to the bacterial two-component signalling system (Parkinson and Kofoid 1992) but it lacks the receiver domain. In contrast to the subfamily I receptors, the subfamily II receptors (EgETR2, EgETR3, EgETR4 and EgETR5) contained a diverged histidine kinase domain. The H residue that is critical for the histidine kinase activity is absent but the phosphate receiver residue, aspartate (D) and two other residues (D and lysine, K) conserved in the bacterial two-component regulators are present at the predicted positions within the respond regulator domains of EgETR3, EgETR4 and EgETR5.

In tobacco, *in vitro* analysis of the ethylene receptor *Nt-HK1* and *Nt-HK2* demonstrated that the kinase activity is not dependent on the conserved H residue (Xie et al. 2003; Zhang et al. 2004). Both receptors exhibit serine/threonine kinase activity and in the presence of Ca+, *Nt-HK2* can also exhibit histidine kinase activity. This finding suggests that these domains may have evolved into new kinases especially in the phosphorylation sites especially among various ethylene receptors (Xie et al. 2003; Zhang et al. 2004).


IV
I
II
III
EgETR3 MLRALWHGLLISSLFFCAS-AIQIGFPRCNCDGDSLWSTESILQCQKVSDFLIAAAYFSI
EgETR2 MLRALCHGILIFSLLFSASAAIEIG--RCNCDGDSLWSIETILQCQKVSDFLIAAAYFSI
EgETR4 MLRALCNGVLILSLHFFAC-AAEIEFPQCNCDSDSGWSVESILGCQKVSDFLIAAAYFSI
EgETR5 ---------MRSSWSFALCGG-------CEEDPTSLWTLENILQCQKVSDFLIALAYFSI
EgERS3 -------------------------MEGCDCIEPQWPADELLVKYQYISDFFIALAYFSI
EgERS2 -------------------------MEGCDCIEPQWPGDELLVKYQYISDFFIALAYFSI
EgERS1 -------------------------MEGCDCFEPQWPAEELLIKYQYISDFFIALAYFSI
EgETR3 PLELIYFATCSNLFPFKWIVFLFGAFIVLCGLTHMLNVFTYEPHSFLLMLSLTISKFFTA
EgETR2 PLELIYFAACSNLLPFKWIVFLFGSFIVLCGLTHLLNVFTYEPHSFLLMLCLTVSKFFTA
EgETR4 PLELFYFVTCSNLFPFKWVLFQFGAFIVLCGLTHFLNVFTYEPHSFILMLALTISKFLTA
EgETR5 PLELFYFVTCSAIFPFRWLILQFGAFIVLCGLSHLVAALAYAPHSFLLLLSLVVLKLLTA
EgERS3 PLELIYFVKKSSFFPYRWVLIQFGAFIVLCGATHLINLWTFTVHSRTLAIVMTVAKISTA
EgERS2 PLELIYFVKKSSFFPYRWVLIQFGAFIVLCGATHLINLWTFTMHSKALALVMAMAKVSTA
EgERS1 PLELIYFVKKSSFFPYRWVLVQFGAFIILCGATHLINLWTFTVHSKTVAIVMTVAKISTA
EgETR3 LVSFATAITLLTLIPQLLRVKVRENFLRIKARELDREVGEMKRQEEASWHVRMLTQEIRR
EgETR2 LVSSATSITLLTLIPQLLRVKVRENFLRIKAWELDREVGKMKRQEEASWHVRMLTQEIRK
EgETR4 LVSFLTAITLLTLIPQLLRVKVRENFLRIKARELDQEVGHMKRQEEASWHVRMLTQEIRK
EgETR5 LVSLATAVTLLPIIPQLLRLFVRDGLLRQKARDLGHDLGRMRRQEAAICRVRLLTAEIRR
EgERS3 VVSCATALMLVHIIPDLLSVKTRELFLKNKAEELDREMGLIRTQEETGRHVRMLTHEIRS
EgERS2 VVSCATALMLVHIIPDLLSVKTRELFLKNKTEELNREMHLIRTQEETGRHVRMLAHEIRS
EgERS1 AVSCATALMLVHIIPDLLSVKTRELFLKSKAEELDREMGLIRTQEETGRHVRMLTHEIRS
EgETR3 SIDKHTILYTTLVELSNTLGLQNCAVWMPNENRREMNLTHEL-RQRSSSDLYSHSIAIDD
EgETR2 SIDRHTILYTTLVELSKTLGLQNCAVWMPNEKKREMNLTHEL-RQRNSSSLYSHSIAIDD
EgETR4 SIDRHTILYTTLVELSKTLELQNCAVWMPNDDKTKINLTHEL-RLRNSSDVYRLSIPIDD
EgETR5 SIDRRTILDTALVHLADALSLHDCAIWMPTSPDS-LSLTHHLNRRRRDHHHQTVSIPTAD
EgERS3 TIDRHTILKTTLVELGRTLDLAECALWMPSNSGLNLQLSHTLHHQIPLGSVVSINLPIVN
EgERS2 TIDRHTILNTTLVELGKTLGLAECALWMPSRSGLNLQLSHTLNHQIPLGSVVSINLPIVN
EgERS1 TIDRHTILKTTLVELGRTLGLEECALWMPSRSGSSLQLSHTLRHQITVGSTVPINHSIVN
EgETR3 PDVMEITETKGVKILRPESMLASASSGGTHEPGAVAAIRMPMLKVSNFKG-GTPQIVETS
EgETR2 PDVMEIKETKGVKILRPDSMLAAASSASVLEPGAVAAIRMPMLKVSNFKG-GTPEIVEAS
EgETR4 PDVMEVKKSEGVKILRPESLLGSASSGEVDESGPVAAIRMPLLKVSDFKG-GTPEFIQTC
EgETR5 SDVAEIITRKAVLILDPNSNLVQATESEP--VGPVAAIRMPLLRVSKFNE-GMPELVEES
EgERS3 ---QVFSSNRAIRIPHTCPLARIRPLTGRYVPPEVVAVRVPLLHLSNFQVNDWPELSAKS
EgERS2 ---QVFSSNRAMRIQHTCPLARIRSRTGRYVPPEVVAVRVPLLHHSNFQINDWPELSAKS
EgERS1 ---QVFSSSHAIIIPHACPLARIRPLAGRYVPPEVAAVRVPLLHLSNFQINDWPELSAKS
EgETR3 YAILVLVLPRDDSRVWSYHELEIIEVVADQVAVALSHAAVLEESQLMREKLVEQNRALLH
EgETR2 YAILVLVLPRDDSRVWSSQELEIVEVVADQVAVALSHAAVLEESQMMRDKLMEQNRTLLH
EgETR4 YAILVLVLPRDDFRIWRQHELEIVEVVADQVAVALSHAAVLEESQLMRDKLAEQNRALHQ
EgETR5 YAILVLVLPTGGDRVWTTDEVDIVEVVADQVAVALSHAAVLEESLLMREKLMEQNMALDR
EgERS3 YAIMVLILPSNSARKWHFHELELVEVVADQVAVALSHAAILEESMQARNLLMEQNVALDL
EgERS2 YAIMVLILPSDSARKWHVYELELVEVVADQVAVALSHAAILEESMRARALLMEQNVALDL
EgERS1 YAVMVLMLPSDSARKWHIHELELVEVVADQVAVALSHAAILEESMRARDLLMEQNIALDL

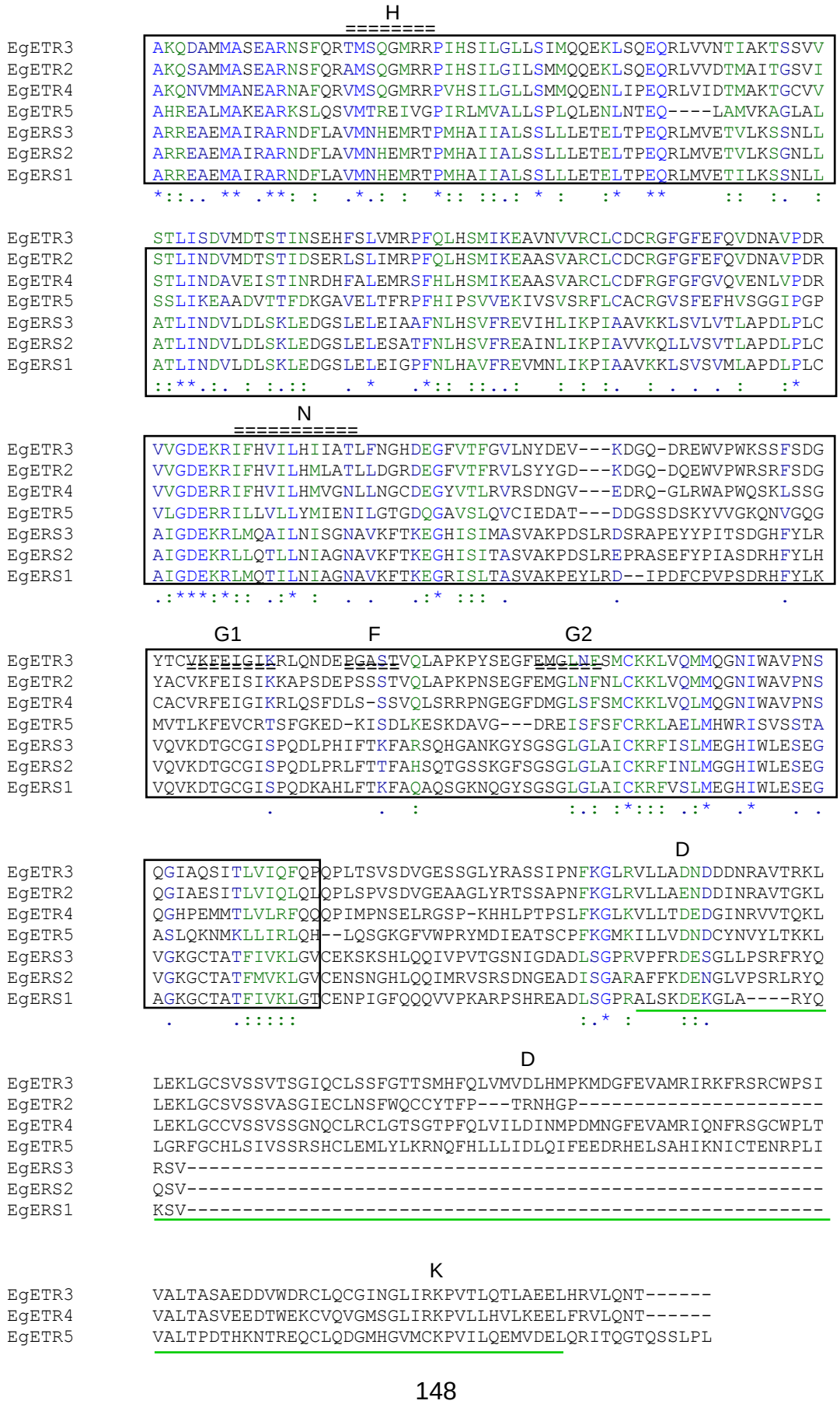

H
EgETR3 AKQDAMMASEARNSFQRTMSQGMRRPIHSILGLLSIMQQEKLSQEQRLVVNTIAKTSSVV
EgETR2 AKQSAMMASEARNSFQRAMSQGMRRPIHSILGILSMMQQEKLSQEQRLVVDTMAITGSVI
EgETR4 AKQNVMMANEARNAFQRVMSQGMRRPVHSILGLLSMMQQENLIPEQRLVIDTMAKTGCVV
EgETR5 AHREALMAKEARKSLQSVMTREIVGPIRLMVALLSPLQLENLNTEQ----LAMVKAGLAL
EgERS3 ARREAEMAIRARNDFLAVMNHEMRTPMHAIIALSSLLLETELTPEQRLMVETVLKSSNLL
EgERS2 ARREAEMAIRARNDFLAVMNHEMRTPMHAIIALSSLLLETELTPEQRLMVETVLKSGNLL
EgERS1 ARREAEMAIRARNDFLAVMNHEMRTPMHAIIALSSLLLETELTPEQRLMVETILKSSNLL
EgETR3 STLISDVMDTSTINSEHFSLVMRPFQLHSMIKEAVNVVRCLCDCRGFGFEFQVDNAVPDR
EgETR2 STLINDVMDTSTIDSERLSLIMRPFQLHSMIKEAASVARCLCDCRGFGFEFQVDNAVPDR
EgETR4 STLINDAVEISTINRDHFALEMRSFHLHSMIKEAASVARCLCDFRGFGFGVQVENLVPDR
EgETR5 SSLIKEAADVTTFDKGAVELTFRPFHIPSVVEKIVSVSRFLCACRGVSFEFHVSGGIPGP
EgERS3 ATLINDVLDLSKLEDGSLELEIAAFNLHSVFREVIHLIKPIAAVKKLSVLVTLAPDLPLC
EgERS2 ATLINDVLDLSKLEDGSLELESATFNLHSVFREAINLIKPIAVVKQLLVSVTLAPDLPLC
EgERS1 ATLINDVLDLSKLEDGSLELEIGPFNLHAVFREVMNLIKPIAAVKKLSVSVMLAPDLPLC
N
EgETR3 VVGDEKRIFHVILHIIATLFNGHDEGFVTFGVLNYDEV---KDGQ-DREWVPWKSSFSDG
EgETR2 VVGDEKRIFHVILHMLATLLDGRDEGFVTFRVLSYYGD---KDGQ-DQEWVPWRSRFSDG
EgETR4 VVGDERRIFHVILHMVGNLLNGCDEGYVTLRVRSDNGV---EDRQ-GLRWAPWQSKLSSG
EgETR5 VLGDERRILLVLLYMIENILGTGDQGAVSLQVCIEDAT---DDGSSDSKYVVGKQNVGQG
EgERS3 AIGDEKRLMQAILNISGNAVKFTKEGHISIMASVAKPDSLRDSRAPEYYPITSDGHFYLR
EgERS2 AIGDEKRLLQTLLNIAGNAVKFTKEGHISITASVAKPDSLREPRASEFYPIASDRHFYLH
EgERS1 AIGDEKRLMQTILNIAGNAVKFTKEGRISLTASVAKPEYLRD--IPDFCPVPSDRHFYLK
G1
F
G2
EgETR3 YTCVKFEIGIKRLQNDEPGASTVQLAPKPYSEGFEMGLNFSMCKKLVQMMQGNIWAVPNS
EgETR2 YACVKFEISIKKAPSDEPSSSTVQLAPKPNSEGFEMGLNFNLCKKLVQMMQGNIWAVPNS
EgETR4 CACVRFEIGIKRLQSFDLS-SSVQLSRRPNGEGFDMGLSFSMCKKLVQLMQGNIWAVPNS
EgETR5 MVTLKFEVCRTSFGKED-KISDLKESKDAVG---DREISFSFCRKLAELMHWRISVSSTA
EgERS3 VQVKDTGCGISPQDLPHIFTKFARSQHGANKGYSGSGLGLAICKRFISLMEGHIWLESEG
EgERS2 VQVKDTGCGISPQDLPRLFTTFAHSQTGSSKGFSGSGLGLAICKRFINLMGGHIWLESEG
EgERS1 VQVKDTGCGISPQDKAHLFTKFAQAQSGKNQGYSGSGLGLAICKRFVSLMEGHIWLESEG
D
EgETR3 QGIAQSITLVIQFQPQPLTSVSDVGESSGLYRASSIPNFKGLRVLLADNDDDNRAVTRKL
EgETR2 QGIAESITLVIQLQLQPLSPVSDVGEAAGLYRTSSAPNFKGLRVLLAENDDINRAVTGKL
EgETR4 QGHPEMMTLVLRFQQQPIMPNSELRGSP-KHHLPTPSLFKGLKVLLTDEDGINRVVTQKL
EgETR5 ASLQKNMKLLIRLQH--LQSGKGFVWPRYMDIEATSCPFKGMKILLVDNDCYNVYLTKKL
EgERS3 VGKGCTATFIVKLGVCEKSKSHLQQIVPVTGSNIGDADLSGPRVPFRDESGLLPSRFRYQ
EgERS2 VGKGCTATFMVKLGVCENSNGHLQQIMRVSRSDNGEADISGARAFFKDENGLVPSRLRYQ
EgERS1 AGKGCTATFIVKLGTCENPIGFQQQVVPKARPSHREADLSGPRALSKDEKGLA----RYQ
D
EgETR3 LEKLGCSVSSVTSGIQCLSSFGTTSMHFQLVMVDLHMPKMDGFEVAMRIRKFRSRCWPSI
EgETR2 LEKLGCSVSSVASGIECLNSFWQCCYTFP---TRNHGP----------
EgETR4 LEKLGCCVSSVSSGNQCLRCLGTSGTPFQLVILDINMPDMNGFEVAMRIQNFRSGCWPLT
EgETR5 LGRFGCHLSIVSSRSHCLEMLYLKRNQFHLLLIDLQIFEEDRHELSAHIKNICTENRPLI
EgERS3 RSV----------
EgERS2 QSV----------
EgERS1 KSV----------
K
EgETR3 VALTASAEDDVWDRCLQCGINGLIRKPVTLQTLAEELHRVLQNT------
EgETR4 VALTASVEEDTWEKCVQVGMSGLIRKPVLLHVLKEELFRVLQNT------
EgETR5 VALTPDTHKNTREQCLQDGMHGVMCKPVILQEMVDELQRITQGTQSSLPL

**Figure 3.16: Amino acid multiple sequence alignment of the oil palm ethylene receptor genes**. The '*' represents a single, fully conserved residue. The ':' represents conservation of strong groups. The '.' represents conservation of weak groups whereas those without any symbol represents no consensus (CLUSTAL W, Biology Workbench Version 3.2, University of California). The four hydrophobic domains region at the amino terminal are lined on top of the proteins and labeled I, II, III and IV in red. The GAF domain is boxed in black. The histidine kinase domain region is boxed in blue. The five conserved motifs in the histidine kinase domain found in bacteria protein kinase are double underlined and the conserved residues (H, N, G1, F and G2) are indicated on top of the protein sequence. The receiver domain is underlined in green and the conserved residues (D, D and K) in this domain is indicated on top of the protein sequence.

The receiver domain however was not detected in EgETR2, exhibiting a feature of an ERS-like ethylene receptors just like the ethylene receptor *AtERS2* in *Arabidopsis* (Hua et al. 1998). In Arabidopsis, the ETR and ERS-like ethylene receptors belong to both subfamilies I and II (Chang et al. 1993; Hua et al. 1998). On the other hand, comparison to other monocots such as rice (Yau et al. 2004) and maize (Chen and Gallie 2010) demonstrated that the ERS-like receptors belong to the ethylene receptor subfamily I while the ETR-like receptors belong to subfamily II. In date palm however, a predicted EIN4-like isoform 2 (Accession No.: XP00801310) with the highest score (1177) and homology (92%) to EgETR2 exhibits a protein structure similar to EgETR2 but lacks the HATPase_c domain. The results showed that an ERS-like receptors that belong to the ethylene receptor subfamily II does exist in monocots. Nevertheless, the division of the ethylene receptor family is not based on the absence of the receiver domain but more of phylogenetic relationships and some shared structural features among the receptors (Schaller and Kieber 2002).

### 3.3.6 Phylogenetic analysis of the ethylene receptor family

The phylogenetic tree was constructed to analyse the relationship between the seven oil palm ethylene receptors and fifty four ethylene receptors from other plants. The resulting tree demonstrated that the ethylene receptor family in plants is classified into two subfamilies, which confirmed the earlier findings. Due to the fairly large number of ER genes from other plants used in

constructing the tree, sub-groups within each sub-family were also observed just as described by (Niu et al. 2013). Each subfamily contains at least 3 groups and each group has at least one clade (Figure 3.17). It was also observed that each respective groups in both subfamilies contains ethylene receptors that are only present in either monocotyledons or dicotyledons, suggesting the possibility of sequence conservation among both types of plants. All seven putative ethylene receptors from oil palm are found to be in the same clade as the date palm ethylene receptors. The sequence homology between the ethylene receptors from both species is high ranging from 87% to 94% identity. This clearly demonstrates that each ethylene receptors from oil palm and date palm are more closely related to each other than any other members of the ethylene receptor family from the same species as well as from other plants species.

### 3.3.7 Mapping the putative ethylene receptor genes to the oil palm chromosome and genetic linkage group

Exonerate search on the latest EG6 build with a 60% self-score threshold showed that majority of the seven putative ethylene receptor genes are positioned on different chromosomes. The putative ethylene receptors *EgETR5* and *EgERS1* are positioned on the chromosome 5. The rest of the putative genes are confined to a separate chromosome (Table 3.10). Identification of the genetic markers flanking the putative ethylene receptors were determined based on their location on the corresponding EG6 build scaffolds. The genetic

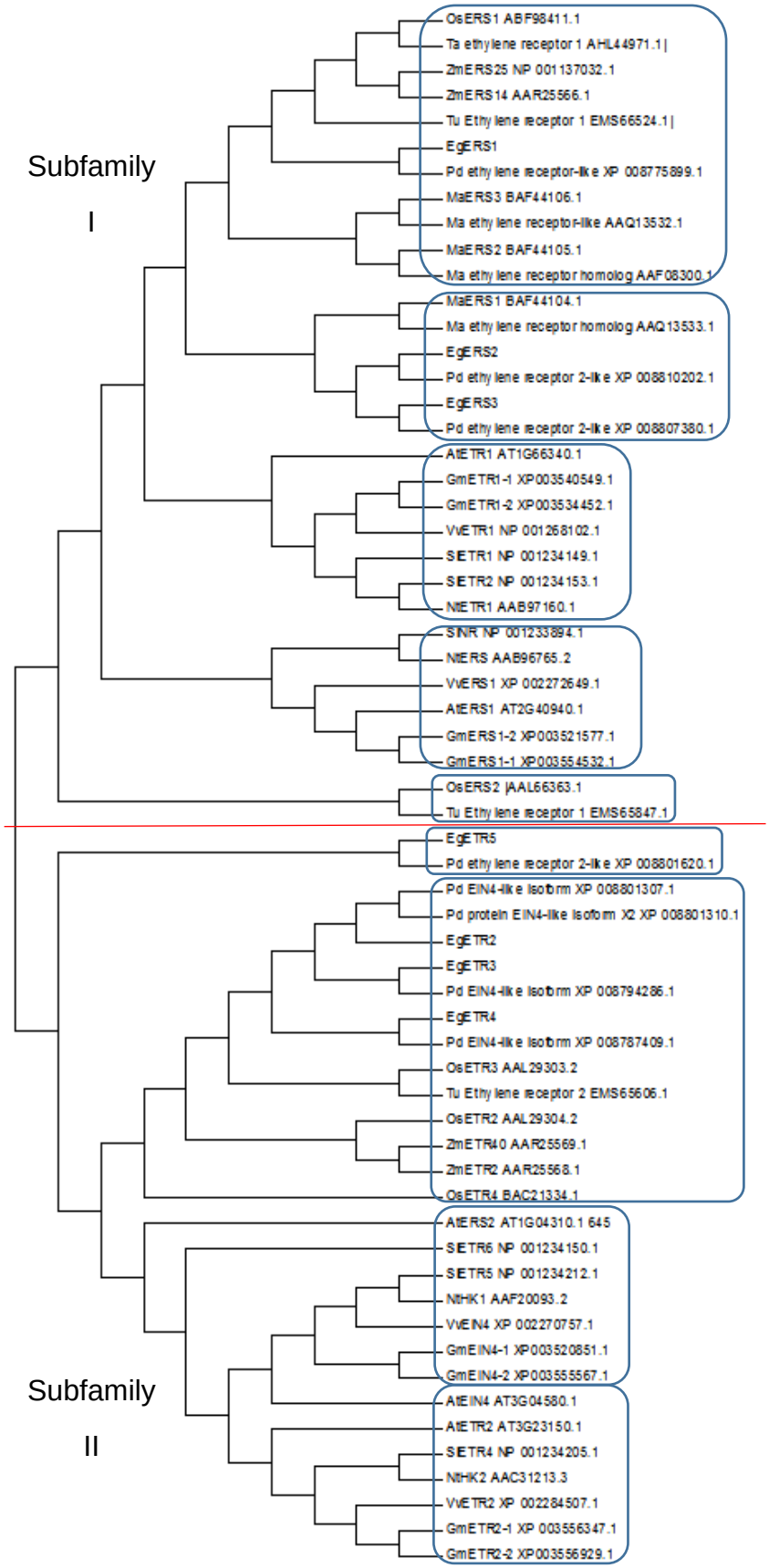



**Figure 3.17: Phylogenetic tree of ethylene receptor proteins in various dicot and monocot plants.** Protein sequences were aligned using Clustal W and the tree was constructed using MEGA 5.0 by neighbour-joining method with 1000 bootstrap replicates. The ethylene receptor proteins are labelled following gene ID or gene name from web searches. Abbreviation of the species are as follows: *Os = Oryza sativa, Ta = Triticum aestivum, Zm = Zea mays, Tu = Triticum urartu, Eg = Elaeis guineensis, Pd = Phoenix dactylifera, Ma = Musa acuminata, At = Arabidopsis thaliana, Gm = Glycine max, Vv = Vitis vinifera, Sl = Solanum esculentum, Nt = Nicotiana tabacum*

markers information is shown in Table 3.10. These markers were used to locate the position of the putative genes on the T128 and P2 genetic linkage maps. Results revealed that the distribution pattern of the seven putative genes on the genetic maps is similar to that of the oil palm chromosome. *EgERS1* and *EgETR5* are at least 34.9 cM apart on the T128 linkage group while the rest are separated in different linkage groups (Figure 3.18). However, the linkage group number is different from the chromosome number (Table 3.10).

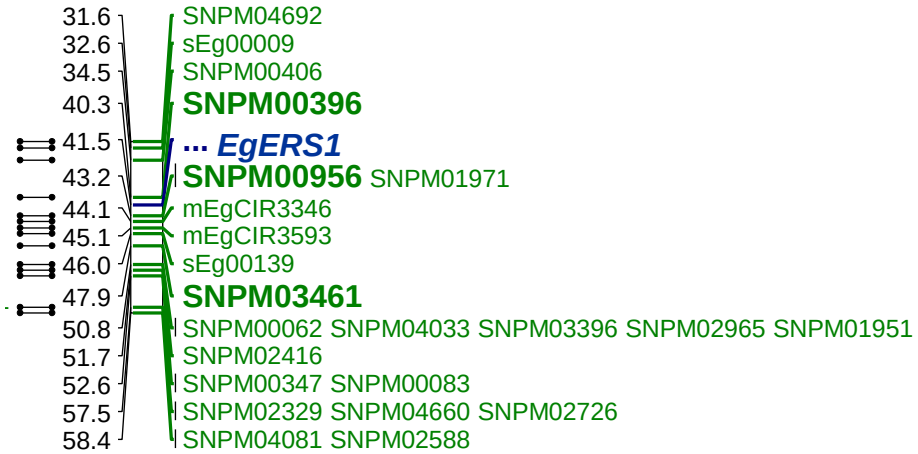
### 3.3.8 Promoter isolation and *in silico* promoter analysis

Figure 3.19 showed the analysis of the amplified putative ethylene receptor promoters on 1% agarose gel. *In silico* analysis using MEME on the nucleotide sequence 1500 bp upstream from the putative TSS identified twenty motifs that correspond to short sequences between 7 to 10 bp (Figure 3.20). Four of the motifs which include motifs 8, 9, 16 and 20 have a similarity greater than 0.60 with at least another motif. The list of the consensus motif sequences is shown in Table 3.11. The motifs found by MEME is not present on all the promoters except for motifs 1, 4, 5, 9 and 15. The rest of the motifs are found in at least 3 of the promoter regions. The p value however in most of the motifs is almost certainly not significant except for motif 1 with an e value of 5.2e-002 (Figure 3.20). The overall distribution of the motifs across the putative ethylene receptor promoter regions is shown in Figure 3.21.

**Table 3.10: Location of the putative ethylene receptor genes on oil palm chromosome and linkage groups**

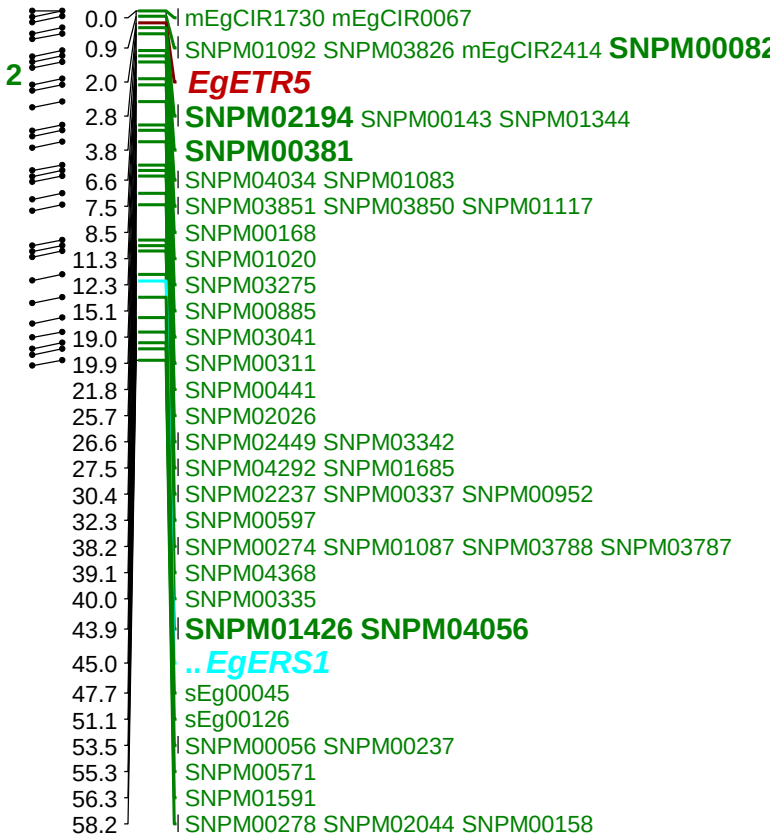
| Gene ID | Amino acid | Chromosome | Genetic linkage group (P2) | Genetic linkage group (T128) | Markers flanking the ethylene receptors (Ting et al. 2014) |
|---|---|---|---|---|---|
| *EgERS1* | 629 | 5 | 12 | 12 | SNPM02495, SNPM04056 |
| *EgERS2* | 635 | 10 | 15 | 13 | SNPM00396, SNPM00956 |
| *EgERS3* | 635 | 2 | 4 | 7 | SNPM01759, SNPM00363 |
| *EgETR2* | 687 | 4 | 11 | 6 | SNPM01119, SNPM01944 |
| *EgETR3* | 757 | 11 | 14 | 11 | SNPM03065, SNPM00308 |
| *EgETR4* | 755 | 12 | 13 | 14 | SNPM04739, SNPM00904 |
| *EgETR5* | 737 | 5 | 12 | 12 | SNPM00082, SNPM02194 |

LG15: *EgERS2*




LG4: *EgERS3*


LG12: *EgERS1* and *EgETR5*



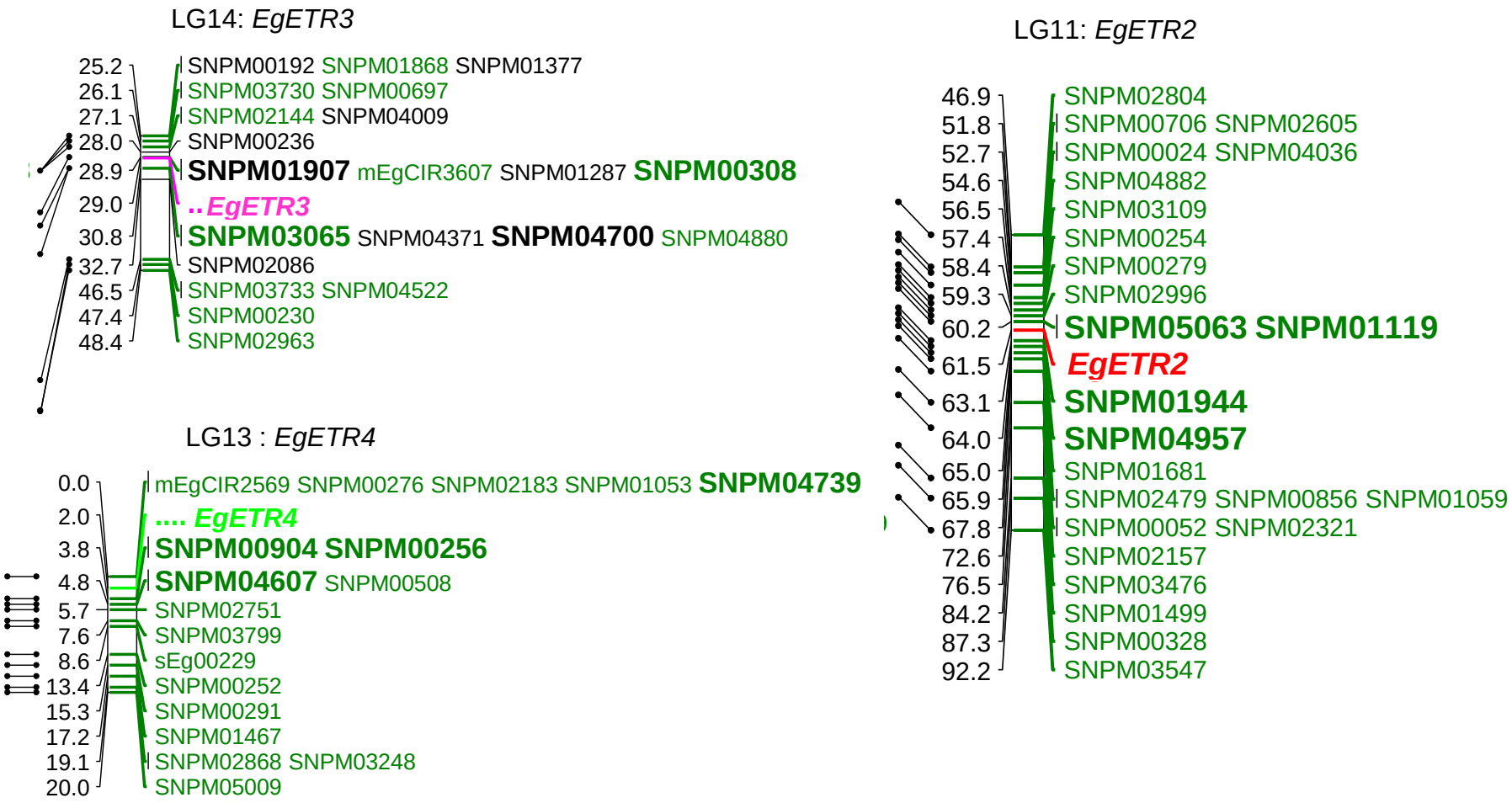



**Figure 3.18: Graphical representation/distribution of the seven putative ethylene receptor genes on oil palm DxP (P2) linkage maps (Ting et al. 2014) using MapChart software.**

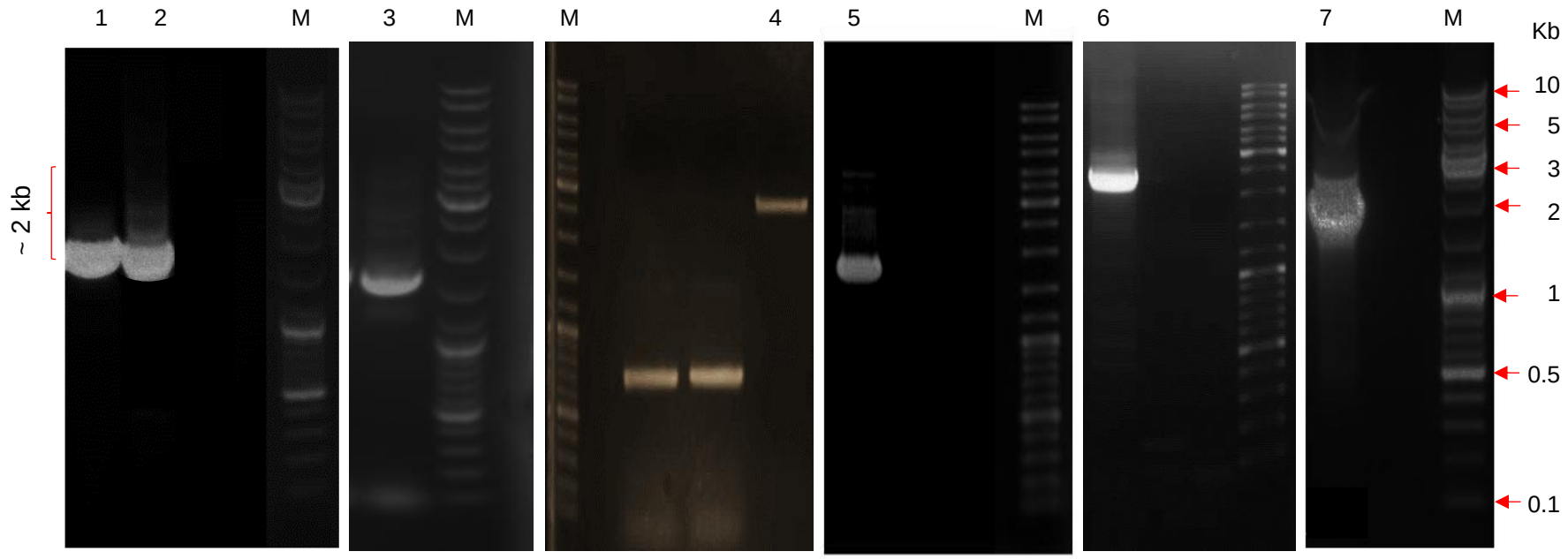



**Figure 3.19: Analysis of the amplified ethylene receptor promoters by electrophoresis on 1% agarose gel electrophoresis.** Lanes 1 – 7 represents the amplified promoters from *EgERS1*, *EgERS2*, *EgERS3*, *EgETR2*, *EgETR3*, *EgETR4* and *EgETR5*. M = GeneRuler DNA ladder Mix (Fermentas).

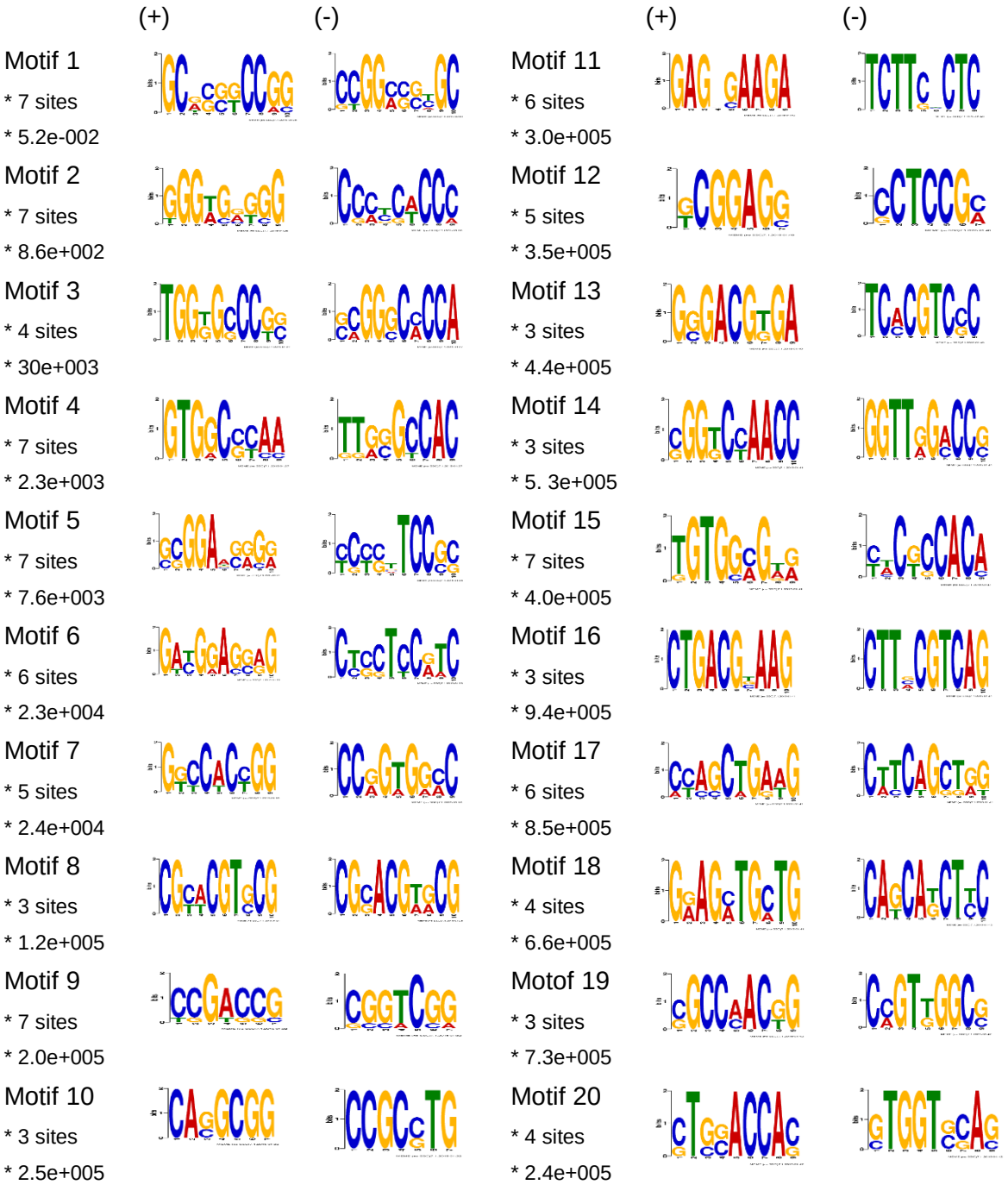



**Figure 3.20: Analysis of motifs present in the 1500 bp region of the putative ethylene receptor promoter**. Motifs are numbered from 1 to 20 and the presence of motifs in the promoter regions are indicated by the site number below the motif numbers, followed by the e values of the respective motifs. The graphic represents the nucleotide frequency distribution of each motifs in both + and – strands using WebLogo.

**Table 3.11: Identification of consensus motifs in the promoter regions of the seven ethylene receptors from oil palm using MEME software and their corresponding putative *cis*-regulatory elements from Plantcare database.**

| Motif No. | Width (nt) | Sequence composition of motifs | Putative *cis*-regulatory motifs | Description |
|---|---|---|---|---|
| Motif 1 | 10 | GCGCGGCCGG | - PE3 | - enhancer |
| | | | - Motifs I & IIa | - abscisic acid responsive element |
| Motif 2 | 9 | GGGTGGGGG | No match | |
| Motif 3 | 10 | TGGGGCCCGC | No match | |
| Motif 4 | 9 | GTGGCCCAA | CellCycle-1a | - transactivator in the cell cycle-dependent transcription |
| Motif 5 | 10 | GCGGAAGGGG | Chs-unit 1 | - part of a light responsive element |
| Motif 6 | 10 | GACGGAGGGG | - Chs-unit 1 | - part of a light responsive element |
| | | | - Chs-unit 1 m2 | - part of a light responsive element |
| Motif 7 | 9 | GGCCACCGG | - CE1 | - *cis*-acting element associated to ABRE, involved in ABA responsiveness |
| | | | - Chs-unit 1 | - part of a light responsive element |
| | | | - Chs-unit 1 m2 | - part of a light responsive element |
| Motif 8 | 10 | CGCACGTGCG | - ABRE | - *cis*-acting element involved in the abscisic acid responsiveness |
| | | | | - part of a light responsive element |
| | | | - Chs-CMA2c | - part of a light responsive element |
| | | | - RbcS-CMA6b | |
| Motif 9 | 7 | CCGACCG | - Chs-unit 1 | - part of a light responsive element |
| | | | - Chs-unit 1 m2 | - part of a light responsive element |
| Motif 10 | 7 | CAGGCGG | - JERE | - jasmonate and elicitor-responsive element |
| | | | - ABRE | - *cis*-acting element involved in the abscisic acid responsiveness |
| | | | | - *cis*-acting regulatory element involved in light responsiveness |
| | | | - G-box | |
| Motif 11 | 9 | GAGGGAAGA | No match | |

| Motif No. | Width (nt) | Sequence composition of motifs | Putative *cis*-regulatory motifs | Description |
|---|---|---|---|---|
| Motif 12 | 7 | GCGGAGG | - PE3 | - enhancer |
| | | | - ABRE | - *cis*-acting element involved in the abscisic acid responsiveness |
| | | | - Y-box | - part of a light responsive element |
| | | | - RbcS-CMA7c | - abscisic acid responsive element |
| | | | - Motif IIb | - abscisic acid responsive element |
| | | | Motifs I & IIa | - part of a light responsive element |
| | | | Chs-CMA2c | - part of a light responsive element |
| | | | Chs-unit 1 | - part of a light responsive element |
| | | | Chs-unit 1 m2 | |
| Motif 13 | 9 | GGGACGGGA | No match | |
| Motif 14 | 10 | CGGGCCAACC | - Unnamed__1 | - CG-1 factor binding site |
| | | | - G-box | - *cis*-acting regulatory element involved in light responsiveness |
| | | | | - module involved with light responsiveness |
| | | | - Sbp-CMA1 | - part of a light responsive element |
| | | | - Chs-unit 1 | |
| Motif 15 | 9 | TGTGGCGAG | GC-repeat | |
| Motif 16 | 10 | CTGACGCAAG | C-box | *cis*-acting regulatory element involved in light responsiveness |
| Motif 17 | 10 | CCAGCTGAAG | Chs-unit 1 | part of a light responsive element |
| Motif 18 | 10 | GGAGCTGCTG | HSE | *cis*-acting element involved in heat stress responsiveness |
| Motif 19 | 9 | CGCCCACGG | No match | |
| Motif 20 | 9 | CTGCACCAC | No match | |

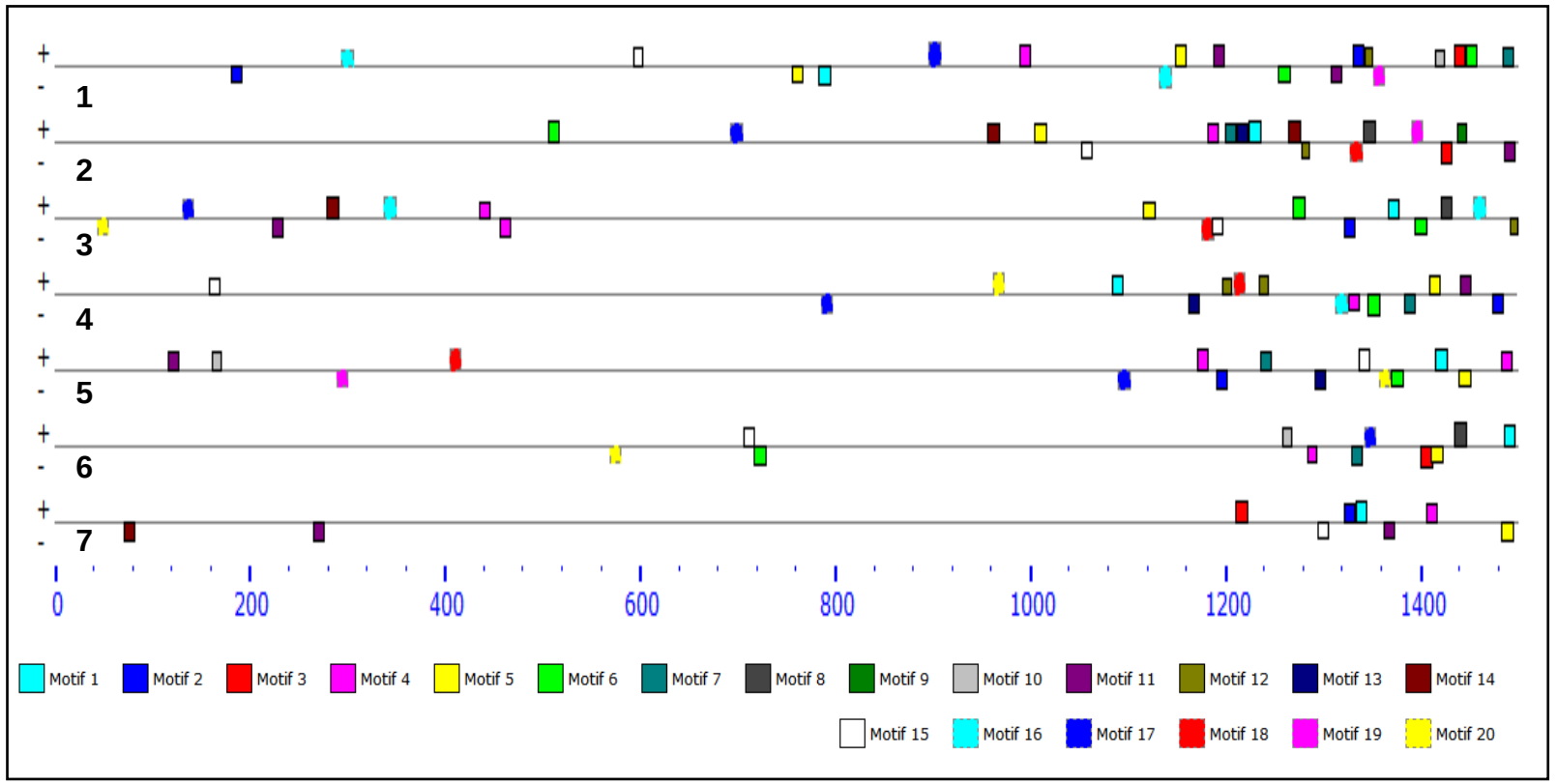



**Figure 3.21: Distribution of motifs in the promoter region of ethylene receptor genes from oil palm by MEME software**. Colored boxes represent the corresponding motifs. 1 = *EgETR4* (e-value = 1.1e-09 ), 2 = *EgERS2* (e-value = 1.6e-09 ), 3 = *EgETR5* (e-value = 3.5e-08 ), 4 = *EgERS1* (e-value = 4.8e-07 ), 5 = *EgETR2* (e-value = 3.5e-06), 6 = *EgERS3* (e-value = 0.00019 ) and 7 = *EgETR3* (e-value = 0.022).

Nonetheless, the 15 motifs were compared to known plant *cis*-acting regulatory elements in the Plantcare database (Table 3.11). Eight of the motifs matched at least two *cis*-regulatory elements excluding motifs 2, 3, 4, 5, 11, 13, 15, 16, 18, 19 and 20. Six motifs which consist of motifs 2, 3, 11, 13, 19 and 20 did not have any matching records to any of the known *cis*-regulatory elements in the PlantCare database. Most of the motifs are related to or part of the light responsive elements (LRE). Motif 10 contains a jasmonate and elicitor-responsive element, JERE that has a GCC-box similar to an ethylene responsive element. Motifs 1, 7 and 10 corresponds to ABRE, CE1 and motifs I and IIa, all of which are *cis*-acting elements that are associated with abscisic acid responsive elements. Motif 18 with a consensus sequence GGAGCTGCTG matched to a heat stress/shock responsive element, HSE.

PlantCare and PLACE Signal Scan analyses revealed the presence of the *cis*-acting regulatory elements in either all, some or just specific to the respective promoters. The majority of those present in a single promoter correspond to *cis*-acting light response elements (Tables 3.12 and 3.13). While the majority of the motifs are related to or part of the LRE, *cis*-acting regulatory motifs especially those in response to other hormones were also detected. In addition to those discovered from the MEME motifs, other *cis*-acting regulatory motifs in response to other hormones such as ethylene (ERE), auxin (AuxRR-core, TGA-box and TGA-element), gibberellin (P-box and TATC-box) and salicylic acid (TCA-element) were also detected in the PlantCARE database. The TATC-box was only detected on *EgERS1* whereas the rest of the hormone

related response elements are detected in at least 2 of the promoters. TC-rich repeats, MBS element and WUN-motifs, all related to stress related responsive elements such as defense, drought and wound were also detected. The summary of the *cis*-acting elements identified in the individual promoters is shown is Tables 3.12 and 3.13, respectively.

Both PlantCARE and PLACE Signal Scan programmes were also used to predict the TATA box of the putative promoters. The location of the TATA-boxes in these promoters were far from the 32 ± 7 distance from the TSS as suggested by (Joshi 1987b) except for *EgETR2* and *EgERS3.* Based on PlantCARE, the nearest TATA boxes were found to be in the range between 204 and 375 nucleotide from the TSS. As the TSS was determined by the furthest upstream nucleotide of the 5` RACE cDNA, there could be a possibility that the 5` UTR of these putative promoters are much longer. The interference of secondary structure in the mRNA may possibly prevent the amplification of a longer 5` RACE cDNA (Chenchik et al. 1996).

**Table 3.12: Summary list of single *cis*-acting elements identified by PlantCARE.**

| No. | Motifs | Promoter | Sequence | Function description |
|---|---|---|---|---|
| 1 | 4cl-CMA1b | EgERS1 | ATTCCGATAAACT | light responsive element |
| 2 | ACA-motif | EgERS1 | AATCACAACCATA | part of gapA in (gapA-CMA1) involved with light responsiveness |
| 3 | AC-II | EgERS1 | (C/T)T(T/C)(C/T)(A/C)(A/C)C(A/C)A(A/C)C(C/A)(C/A)C | |
| 4 | box II | EgERS1 | TCCACGTGGC | part of a light responsive element |
| 5 | C-box | EgERS1 | CTGACGTCAG | *cis*-acting regulatory element involved in light responsiveness |
| 6 | TATC-box | EgERS1 | TATCCCA | *cis*-acting element involved in gibberellin-responsiveness |
| 7 | TCCC-motif | EgERS1 | TCTCCCT | part of a light responsive element |
| 8 | ATGCAAAT motif | EgERS2 | ATACAAAT | *cis*-acting regulatory element associated to the TGAGTCA motif |
| 9 | Box III | EgERS2 | atCATTTTCACt | protein binding site |
| 10 | chs-CMA2b | EgERS2 | ATTGCAACTCAA | part of a light responsive element |
| 11 | rbcS-CMA7a | EgERS2 | GTCGATAAGG | part of a light responsive element |
| 12 | Unnamed__2 | EgERS2 | ATTAAATTTTAAATT/ CCCCGG/ AACCTAACCT | |
| 13 | Unnamed__5 | EgERS2 | TGTAATAATATATTTATATT | |
| 14 | plant_AP-2-like | EgERS3 | CGCGCCGG | |
| 15 | TCCACCT-motif | EgERS3 | TCCACCT | |
| 16 | TA-rich region | EgETR3 | TATATATATATATATATATATA | enhancer |
| 17 | WUN-motif | EgETR3 | TCATTACGAA | wound-responsive element |
| 18 | 3-AF1 binding site | EgETR4 | TAAGAGAGGAA | light responsive element |
| 19 | dOCT | EgETR4 | CaCGGATC | *cis*-acting regulatory element related to meristem specific activation |
| 20 | ELI-box3 | EgETR4 | AAACCAATT | elicitor-responsive element |
| 21 | AAAC-motif | EgETR5 | CAATCAAAACCT | light responsive element |
| 22 | Unnamed__15 | EgETR5 | CCTCTCCCGTC | |

**Table 3.13: Summary list of single *cis*-acting elements identified by PLACE Signal Scan Programme.**

| No | factor | Signal Sequence | SITE # | Promoter | Description |
|---|---|---|---|---|---|
| 1 | HEXAT | TGACGTGG | S000334 | EgERS1 | G-Box-like element |
| 2 | LRENPCABE | ACGTGGCA | S000231 | EgERS1 | LRE, light regulatory element |
| 3 | NRRBNEXTA | TAGTGGAT | S000242 | EgERS1 | negative regulatory region expression in all tissues within the stem internode, petiole and roots |
| 4 | OCSELEMENTAT | TGACGYAAGSRMTKACGYMM | S000158 | EgERS1 | ocs-like element |
| 5 | UPRMOTIFIAT | CCACGTCA | S000425 | EgERS1 | *cis*-acting element involved in unfolded protein response |
| 6 | 2SSEEDPROTBANAPA | CAAACAC | S000143 | EgERS1 | Conserved in many storage-protein gene promoters |
| 7 | ACGTABOX | TACGTA | S000130 | EgERS2 | responsible for sugar repression |
| 8 | ANAERO3CONSENSUS | TCATCAC | S000479 | EgERS2 | involved in the aenarobic fermentative pathway |
| 9 | DRE2COREZMRAB17 | ACCGAC | S000402 | EgERS2 | drought-responsive element |
| 10 | LTREATLTI78 | ACCGACA | S000157 | EgERS2 | low temperature responsive element (LTRE) |
| 11 | SITEIIBOSPCNA | TGGTCCCAC | S000217 | EgERS2 | meristematic tissue-specific expression |
| 12 | TATCCACHVAL21 | TATCCAC | S000416 | EgERS2 | part of the conserved *cis*-acting response complex (GARC) for a full GA response |
| 13 | RAV1BAT | CACCTG | S000315 | EgERS3 | Binding consensus sequence for transcription factor RAV |
| 14 | SBOXATRBCS | CACCTCCA | S000500 | EgERS3 | S-box Important for the sugar and ABA responsiveness |
| 15 | CACGTGMOTIF | CACGTG | S000042 | EgERS3 | |
| 16 | INTRONLOWER | TGCAGG | S000086 | EgERS3 | "3' intron-exon splice junctions |
| 17 | IRO2OS | CACGTGG | S000505 | EgERS3 | OsIRO2-binding core sequence |
| 18 | XYLAT | ACAAAGAA | S000510 | EgERS3 | *cis*-element identified among the promoters of the "core xylem gene set |
| 19 | MYB2AT | TAACTG | S000177 | EgETR2 | Binding site for MYB homolog |
| 20 | PYRIMIDINEBOXHVEPB1 | TTTTTTCC | S000298 | EgETR2 | Required for GA induction |
| 21 | TRANSINITDICOTS | AMNAUGGC | S000201 | EgETR2 | translational initiation codon in monocots |
| 22 | TRANSINITMONOCOTS | RMNAUGGC | S000202 | EgETR2 | translational initiation codon in dicots |

| | | | | | |
|---|---|---|---|---|---|
| 23 | GT1MOTIFPSRBCS | KWGTGRWAAWRW | S000051 | EgETR3 | light responsive element |
| 24 | MARABOX1 | AATAAAYAAA | S000063 | EgETR3 | A-box" found in SAR(scaffold attachment region; or matrix attachment region, MAR); |
| 25 | MYB26PS | GTTAGGTT | S000182 | EgETR3 | Myb26 binding site; |
| 26 | MYBATRD22 | CTAACCA | S000175 | EgETR3 | Binding site for MYB (ATMYB2) in dehydration-responsive gene |
| 27 | SORLREP3AT | TGTATATAT | S000488 | EgETR3 | Sequences Over-Represented in Light-Repressed Promoters |
| 28 | ABREATCONSENSUS | YACGTGGC | S000406 | EgETR4 | ABA-responsive elements |
| 29 | ABREOSRAB21 | ACGTSSSC | S000012 | EgETR4 | ABA-responsive elements |
| 30 | ABRERATCAL | MACGYGB | S000507 | EgETR4 | ABRE-related sequence" |
| 31 | BP5OSWX | CAACGTG | S000436 | EgETR4 | MYC protein) binding site |
| 32 | CTRMCAMV35S | TCTCTCTCT | S000460 | EgETR4 | CT-rich motif to enhance gene expression |
| 33 | GAGA8HVBKN3 | GAGAGAGAGAGAGAGA | S000427 | EgETR4 | GA octodinucleotide repeat |
| 34 | UP1ATMSD | GGCCCAWWW | S000471 | EgETR4 | *cis*-elements that regulate gene expression during initiation of axillary bud outgrowth |
| 35 | ACGTOSGLUB1 | GTACGTG | S000278 | EgETR5 | ACGT motif" required for endosperm-specific expression |
| 36 | ATHB6COREAT | CAATTATTA | S000399 | EgETR5 | Consensus binding sequence for homeodomain-leucine zipper protein |
| 37 | BS1EGCCR | AGCGGG | S000352 | EgETR5 | BS1 (binding site 1) Required for vascular expression |
| 38 | CMSRE1IBSPOA | TGGACGG | S000511 | EgETR5 | Carbohydrate Metabolite Signal Responsive Element 1) involved in the RT sucrose-inducible expression |
| 39 | PALBOXAPC | CCGTCC | S000137 | EgETR5 | Box A, one of the three putative *cis*-acting elements that conferred elicitor or light responsiveness |
| 40 | S1FSORPL21 | ATGGTATT | S000215 | EgETR5 | S1F box that play a role in downregulating promoter activity |
| 41 | SITEIIATCYTC | TGGGCY | S000474 | EgETR5 | Site II element" |
| 42 | UPRMOTIFIIAT | CCNNNNNNNNNNNNCCACG | S000426 | EgETR5 | Motif II" in the conserved UPR (unfolded protein response) |

### 3.3.9 Expression profiling of the putative ethylene receptor genes

The transcript accumulation profiles in the tested tissues are presented in Figure 3.22. In the various stages of mesocarp fruit development, it was observed that the expression profile of the ethylene receptor genes varies from one another. In *EgETR3*, an increase in expression was observed following the fruit developmental process from young until the ripening stage. Between 8 WAA to 12 WAA, the expression is low but the increase in expression is apparent at 15 WAA and is highest at 20 WAA. In *EgETR2* and *EgERS3*, the pattern of expression is similar to that of *EgETR3* especially with the increased expression level at 15 WAA and reaching the highest point at 20 WAA. In the young fruits, however, expression decreased from 8 WAA to 12 WAA. In *EgERS1*, there is a slight increase in the expression in the mesocarp tissue from 8 WAA until 17 WAA followed by a decreased in expression at 20 WAA, which represents the ripening stage of the oil palm fruits. Depending on the genotype, the oil palm ripening stages occur between 20-24 WAA where the fruit colour change, softening of tissue and high oil accumulation of oil is prominent. A different expression pattern was shown by *EgETR4*. The expression of *EgETR4* is highest at 10 WAA and decreased as the fruits developed and ripened. Contrary to what was observed with *EgETR2, EgETR3, EgETR4, EgERS1* and EgERS3, the expression of *EgETR5* and *EgERS2* is similar in all the mesocarp tissues at various fruit development stages with the highest expression being at 12 WAA in *EgETR5.*

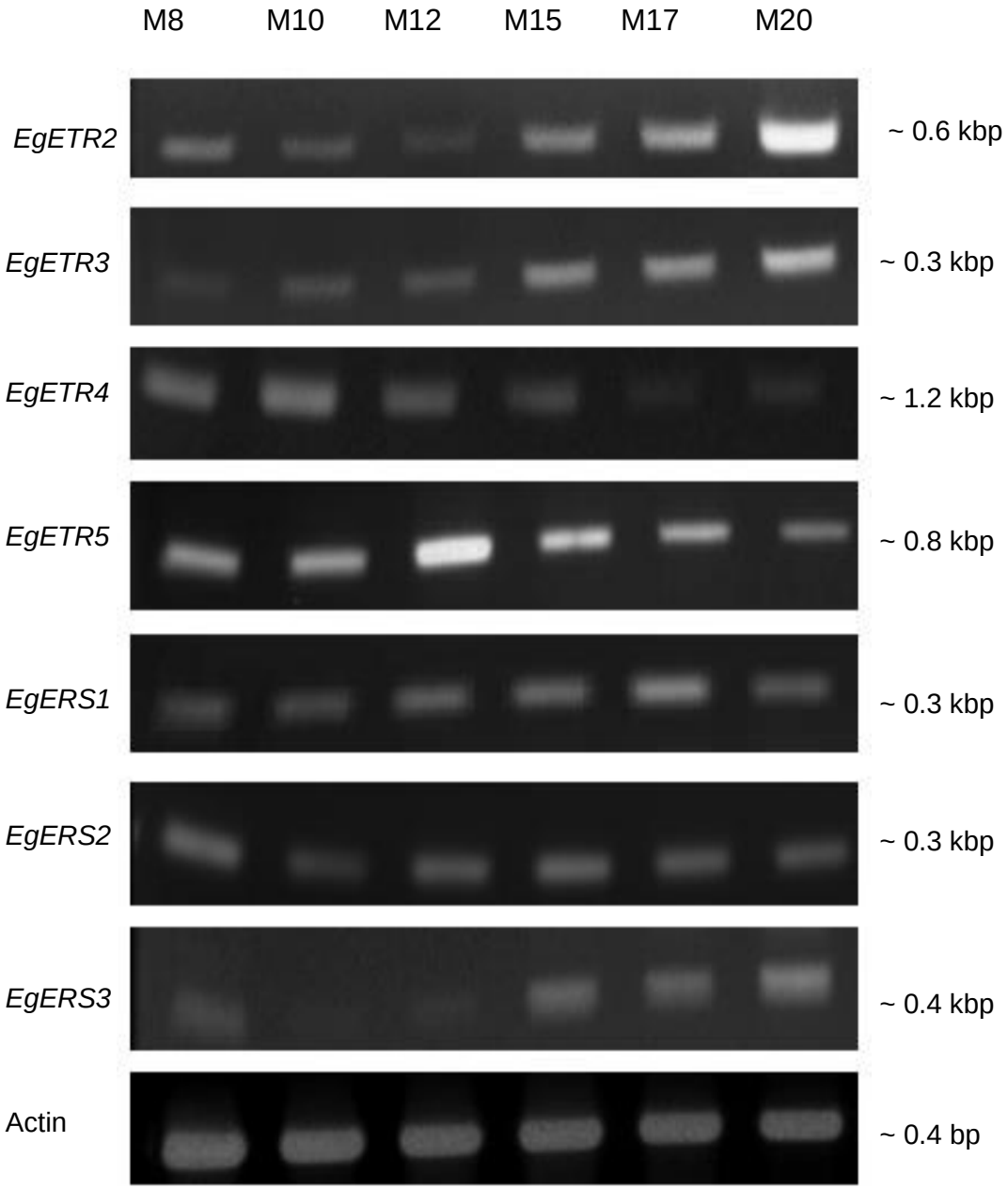



**Figure 3.22: Expression analysis of the oil palm ethylene receptor genes throughout fruit development in the oil palm mesocarp tissues *via* RT-PCR.** Actin was used to show approximate expression throughout the various tissues. M = mesocarp while numbers represent the WAA.

The expression pattern of the ethylene receptors genes at the fruit AZ were also investigated in three locations (apical, central and basal) within the oil palm fruit bunches at 20 and 22 WAA. Depending on the gene, the transcript level in the tested developmental stages and locations within the bunch varies where some of the ethylene receptors showed higher expression as compared to the rest (Figure 3.23). For instance, *EgETR2* had the highest expression in most of the tested AZ zones within the bunch followed by *EgERS3* whereas *EgERS2* had the lowest. An overview of the overall expression of the ethylene receptor genes showed that the expression at 20 WAA is slightly higher in all three locations within the fruit bunches in comparison to 22WAA fruit bunches. This in particular was especially observed at the centre in the 20 WAA bunches which represents the central region of the fruit bunches. It was also observed that the expression of *EgETR2* at 20 and 22 WAA at the apical region of the fruit bunch remained the same unlike *EgETR5*, where the transcript level was higher at 22WAA.

Besides the mesocarp and AZ tissues, the expression of the ethylene receptor genes were also determined by performing RT-PCR in different oil palm tissues which included kernel at 12 WAA, spear leaves, roots, germinated seedlings, female and male flowers (Figure 3.24). Gene expression profile of the ethylene receptors revealed that the genes were differentially expressed in various oil palm tissues. Overall, the expression of the ethylene receptor gene family in the tested tissues was low, some of which was only just detectable and the level of expression in most tissues differed from one another. The transcript accumulation was highest in the

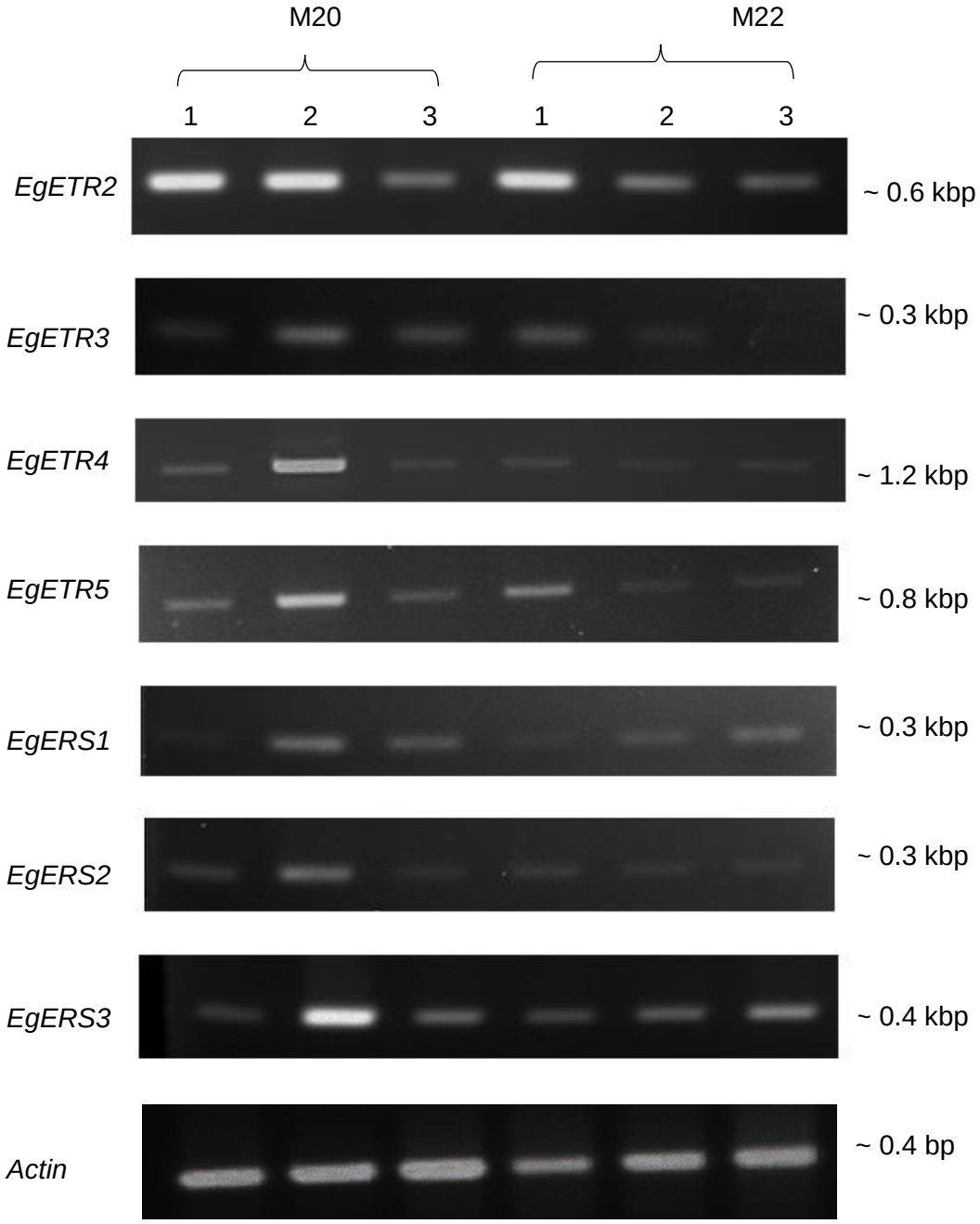



**Figure 3.23: Expression analysis of the oil palm ethylene receptor genes in the fruit AZ of oil palm fruits at two ripening stages *via* RT-PCR**. M20 = mesocarp at 20 WAA, M22 = mesocarp at 22 WAA, lanes 1 to 3 = apical, central and basal region of the oil palm fruit bunches. Actin represents approximate expression throughout the various tissues.

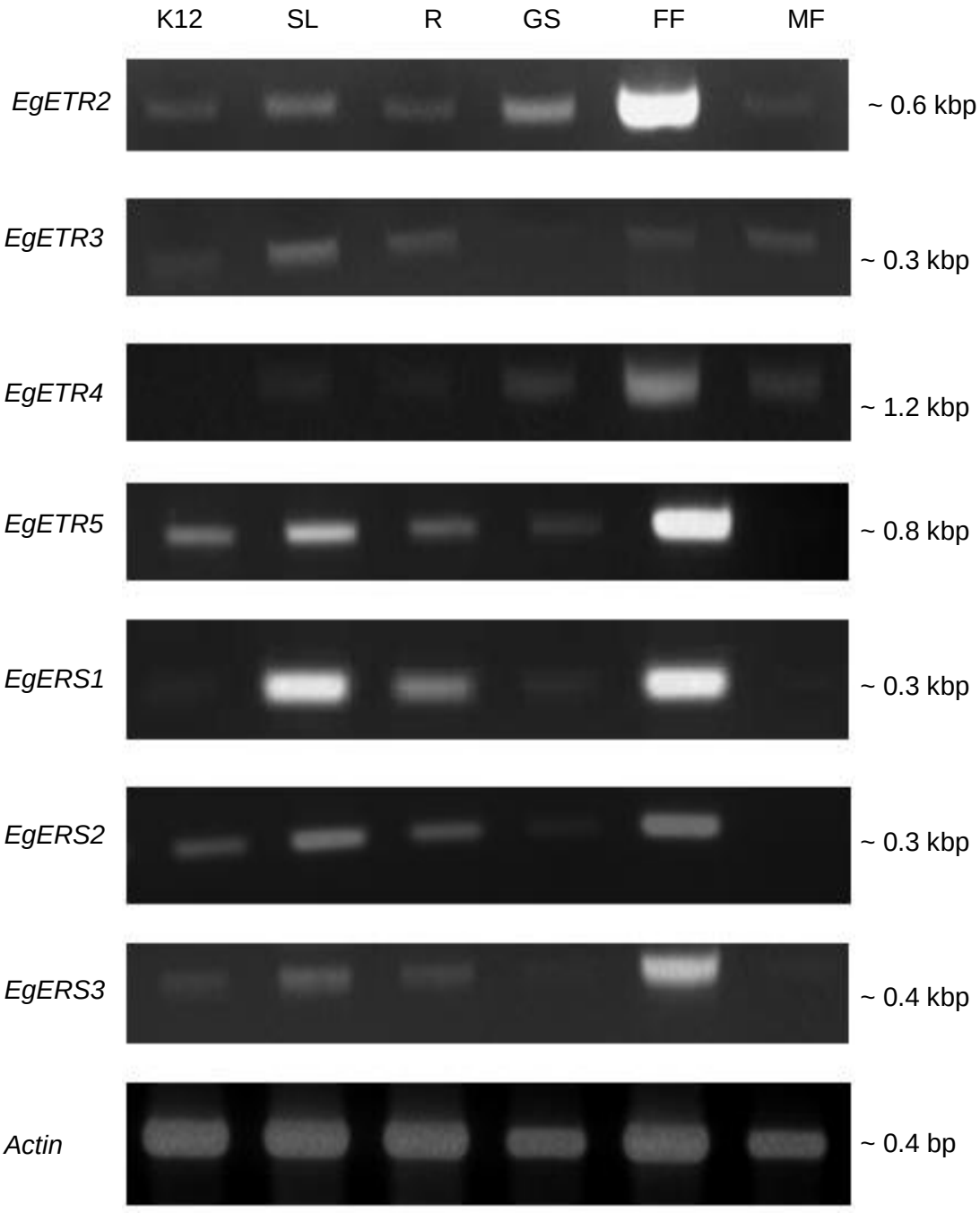



**Figure 3.24: Expression analysis of the oil palm ethylene receptor genes in various oil palm tissues *via* RT-PCR.** K12 = kernel at 12 WAA, SL = spear leaves, GS = germinated seedlings, R = roots, FF = female flower and MF = male flower. Actin represents approximate expression throughout the various tissues.

female flower in all of the tested tissues except for *EgETR3*. Within the ethylene receptor subfamily I, *EgERS1, EgERS2* and *EgERS3* had similar expression patterns but at different mRNA intensities. The transcripts were expressed in the kernel at 12 WAA, spear leaves and germinated seedlings, although the expression of *EgERS1* in kernel at 12 WAA is only at a trace level. In addition, the expression of these genes in roots is similar to that of *EgERS1* in kernel at 12 WAA and none was detected in the male flower. The subfamily II members on the other had exhibited a different expression pattern where some of the transcript accumulation was not detected in a certain tissue. *EgETR3* transcript was not expressed in roots, *EgETR4* was not detected in kernel at 12 WAA and *EgETR5* was not detected in the male flower.

## 3.4 Discussion

The genomes of many economically important crops have been sequenced with the aim of utilising the knowledge gained for improvement of crop quality and disease resistance. The sequencing technology has also been taken up /adapted by MPOB where the first draft sequences of both the African (*E. guineensis* var. *pisifera*) and American (*E. oleifera*) palms were announced in 2009 (Mohd Basri 2009). Four years later, the genome sequence was finally published and made publically available (Singh et al. 2013a). Accessibility to the oil palm genome data prior to it being published, has facilitated the work described in this chapter which has focused on the identification of genes involved in the perception of ethylene.

In this study, the ethylene receptors in oil palm were identified, isolated and characterised. The combination of these techniques elucidated the use of the oil palm genome sequence data. In total, ten ethylene receptors were identified from the oil palm genome, seven (*EgETR2, EgETR3, EgETR4*, *EgETR5, EgERS1, EgERS2* and *EgERS3*) of which represent the putative ethylene receptor genes. The remaining three (*EgETR2-1* [Accession No. XM_010920465], *EgETR3-1*, *EgERS1-1* [Accession No XM_010923295]) are alternative splice variants of three of the putative ethylene receptor genes. The receptors were identified by homology search and a HMM profile build based on known and established ethylene receptors from various plant species. The initial searches on the collection of oil palm transcriptome data from multiple tissue types and gene model proved to be a useful step prior to identifying the putative genes on the oil palm genome scaffold.

In parallel with the duration of this study, a lot of effort was made to improve the quality of the sequence reads thus closing the gaps between the genome scaffolds (Singh et al. 2013a). The use of genes isolated *via* experimental procedure such as *EgERD3* from the oil palm (Nurniwalis 2006) proved to be of much use in helping to identify the other putative oil palm ethylene receptor genes, and helped to minimize the gaps between the oil palm genome scaffolds. With the progress of the genome scaffolds from P3 to the current P6 build, the genetic markers also proved to be useful in identifying the location of the ethylene receptors on the oil palm genetic linkage maps, hence location on the oil palm chromosome. Genome sequencing data has facilitated the progress to unravel the roles of specific genes within a gene family. For instance, with the genome-wide identification of the MADS-box gene family in the rice genome, an understanding of the role of members of this gene family during reproductive development and stress in rice has been attained (Arora et al. 2007). In barley, the identification and mapping of genes and gene families involved in the ethylene biosynthesis and perception from the barley genome provides a means to understand the various roles of ethylene in plant growth and development and in response to biotic and abiotic stresses (Dahleen et al. 2012).

Earlier studies have shown that splice variants are present in certain tissues based on laboratory experiments such as the RT-PCR strategy (Pan et al. 2005), cold plaque screening and 5` RACE (Ooi et al. 2008), among others. With the current advancement in sequencing technologies, transcriptome information from numerous tissues and genomic sequenced

data provide the ability to identify splice variants more rapidly by aligning the exon positions from the transcriptome data to the genome.

A splice variant is a product of alternative splicing of pre-mRNA based on the different splice site that exists within the same mRNA resulting in the production of alternative transcripts and proteins (Lopez 1998). It is also an important post-translational mechanism that controls gene expression (Chen and Manley 2009). Many of the events in plants happen only in a specific tissue at a precise time during development and/or under certain physiological conditions (Graveley 2001) and different transcripts can play different roles in the plant's biological function (Black 2003). In the flower of gladiolus (*Gladiolus grandiflora hort.*), the ethylene receptor *GgERS1b* is a product of alternative splicing of the ERS1-type (*GgERS1a*) ethylene receptor (Arora et al. 2006). *GgERS1b* is almost identical to *GgERS1a* but lacks, by 636 bases or 212 amino acids, the significant first and second histidine kinase motifs. The expression of *GgERS1b* decreased in petal tissues with flower senescence and increased in anther tissues whereas *GgERS1a* was expressed highly and constitutively. *GgERS1b* may have been generated due to incomplete splicing but differential expression of *GgERS1b* in various tissue suggest that both *GgERS1a* and *GgERS1b* genes play important roles for the subfunctionalization process to provide ethylene insensitivity in gladiolus flower (Arora et al. 2006). In this study, although the oil palm *EgERS1-1* lacks 36 bases as was found in *EgERS1*, the possibility that it may play a specific function in oil palm tissues is very high which warrants further investigation in the future.

Intron retention is also a type of alternative splicing where intron is retained in the mature mRNA. Gene expression studies on intron-retained *cyclin A2* cDNA in humans has revealed that the truncated protein may play a different role to that of the full-length *cyclin A* gene which regulates cell proliferation (Honda et al. 2012). In Arabidopsis alternative splicing in introns present in the 5` or 3` UTR has been shown to modulate nonsense-mediated decay sensitivity of important regulatory genes (Kalyna et al. 2012). Two intron-retained splice variants of a single *PSY* gene known to play an important role to control carotene accumulation have been identified and characterized (Alvarez 2016). They are different in size and in the exon/intron retained composition of their 5` UTR and both have differential gene expression in responds to environmental signals. The differential regulation of both splice variants suggest the possibility that the *PSY* gene and its splice variants are working in concert to share the workload and to meet the requirements in different tissues as well as developmental and stress-related carotenoid biosynthesis processes (Alvarez 2016).

Unlike *EgERS1-1*, two of the splice variants, (*EgETR2-1* and *EgETR3 -1*) in the oil palm ethylene receptor gene family both of which belong to Subfamily II, retained an intron the 5` UTR region. *EgETR3-1* is also expressed in 20 WAA mesocarp and leaf tissues suggesting an important role in both or even more oil palm tissues. As of now it is not clear why these splice variants are present in oil palm but is possible that they have a specific/coordinated role to play in the physiological development of the oil palm especially that is associated with ethylene perception.

Amplification of two thirds of the splice variants in oil palm indicated a robust and efficient oil palm sequenced data.

Sixteen oil palm genetic scaffolds where each scaffold represents a chromosome were constructed by Singh et al. (2013a) based on the size of the sequence scaffolds in the EG5 linked build mapped with markers from the P2 and T128 genetic linkage maps. However, numbering of the T128 and P2 linkage groups follows the nomenclature as described by Billotte et al. (2005). Regardless of the different chromosome and linkage group numbers, the position of the genes correspond to the same location (Singh et al. 2013b).

The identification of the *cis*-regulatory motifs by PlantCARE and PLACE Signal Scan in the promoter region of the oil palm ethylene receptors are consistent to previously described and published information. The motifs present in the receptors relates to a variety of responses known to be associated with ethylene. This includes the interaction between ethylene and other phytohormones, which have been well reviewed and documented. For example: ethylene - ABA interaction in the control of fruit ripening (Setha 2012) and production of secondary metabolites such as carotenoids and anthocyanins (Chervin et al. 2009; McAtee et al. 2013), ethylene – auxin interaction in the control of reproductive organ abscission (Sawicki et al. 2015) and tomato fruit development (Wang et al. 2005); ethylene – GA plays a role in the elongation of submerged rice stems due to oxygen deficiency (Weiss and Ori 2007). Complex interactions between multiple hormones such as ethylene-ABA-gibberellin have also been shown to increase the production of secondary metabolites such as phenolic acids

in the roots (Liang et al. 2013) as well as the crosstalk between ethylene-salicylic acid-jasmonate, triggering the plant defense-signalling pathway (Leon-Reyes et al. 2012).

The *cis*-regulatory elements identified by these programmes in all the seven promoters are located at dissimilar locations across the promoters. Majority of the motifs are present in at least 2 of the promoters, which suggests a common regulatory mechanism that plays a similar role in the regulation of gene expression especially with regards to ethylene activity. The shared motifs across the ethylene receptor promoters may also suggest a combined interaction of the elements in regulation of the ethylene receptor gene expression. In rice for example, the GCN4 motif found in all the members of glutelin gene family was shown to be important for the seed-specific regulation (Takaiwa et al. 1996). Several protein-protein interactions and yeast hybrid studies have also shown that interactions between the members of the ethylene receptor family as well as with other components in the ethylene signaling pathway are important in modulating the ethylene signal (Shakeel et al. 2013; Niu et al. 2013).

In the oil palm ethylene receptor promoters, some of the motifs were found specifically in a single promoter which suggests that these small differences may affect the gene expression of the ethylene receptors. In soybean for example, the ethylene response element (ERE) present in the promoter of *GmERS1-2* and absent in *GmERS1-1* caused *GmERS1-2* to have a stronger response to ethylene treatment (Niu et al. 2013). In addition, the presence of repetitive *cis*-regulatory elements within a promoter can also cause differential expression pattern. The ethylene receptor *GmETR2-*

*2* has a stronger response than *GmETR2-1* to low temperature stress when the seedlings were placed at 4$^{o}$C for 12 hours (Niu et al. 2013). Analysis of the promoter regions suggests that perhaps the presence of two LTR elements in *GmETR2-2* and none in *GmETR2-1* is the cause of the stronger response in *GmETR2-2* (Niu et al. 2013).

The oil palm ethylene receptors family is classified into two subfamilies based on intron-exon structure, protein conserved domains, the primary structure as well as phylogenetic analyses. Subfamily I is comprised of 3 members (*EgERS1, EgERS2* and *EgERS3*) whereas subfamily II has 4 members (*EgETR2, EbETR3, EgETR4* and *EgETR5*). The ETR1-type ethylene receptor, has a characteristic of a subfamily I member and contains a response regulator domain that was not found in oil palm. Interestingly, the ETR1-type ethylene receptor is also not found in other monocots such as rice (Yau et al. 2004; Watanabe et al. 2004), wheat (*Triticum aestivum*) (Ma and Wang 2003), maize (Gallie and Young 2004) as well as date palm (Al-Mssallem et al. 2013), which explains the limited information on ethylene signalling in monocots (Ma et al. 2010). In other dicots, especially in model plant Arabidopsis, the *AtETR1* gene is the most exhaustively studied ethylene receptor (Alonso and Stepanova 2004; Chen et al. 2005; Agarwal et al. 2012). Genetic validation of ethylene-responsive mutants has provided extensive information in the dissection of the linear ethylene signalling pathway especially in dicots (Klee 2004). In monocots however, the ethylene signalling pathway is unclear because ethylene controls many processes that do not occur in Arabidopsis (Ma et al. 2010). In rice for example, it was proposed that besides the established components (*CTR1-*

*EIN2-EIN3/EIL*) in the ethylene signalling pathway, an additional feature that represents a novel component is likely to be present as well. The *MHZ1* component has been proposed, in the ethylene signalling pathway in rice, to regulate root growth based on the discovery of the maohuzi (*MHZ)1* mutant that exhibits kinase activity and is required for ethylene-induced root inhibition (Ma et al. 2013, 2014; Yang et al. 2015).

Phylogenetic analysis demonstrated that both subfamilies can be further divided into sub-groups whereas the ethylene receptors from monocots are grouped together and segregated from the dicots. The separation between the ethylene receptor gene family in monocots and dicots suggest that they once shared a common ancestor prior to separation in monocots and dicots about 100 million years ago (Binder er al. 2012). The ethylene receptor members in subfamily I is evolutionary more ancient as compared to members in subfamily II that have completely diverged as demonstrated by phylogenetic analysis in *Physcomitrella patens* and other angiosperms since the split between moss and angiosperms (Yasumura et al. 2012; Binder et al. 2012). Oil palm and date palm ethylene receptors in both subfamilies are highly homologous with 87% to 94% sequence identity. The strong conservation between oil palm and date palm suggest that the ethylene receptors from oil palm and date palm were separated from other monocots. Comparative genomics studies of the oil palm against date palm, Arabidopsis and banana indicated that that the oil palm genome matched the highest to date palm gene model with 94.4% identity (Singh et al. 2013a).

Gene expression analyses of the seven ethylene receptors *via* RT-PCR indicated that the receptors are differentially expressed in various oil palm tissues which includes reproductive as well as vegetative tissues. The ethylene receptors do not exhibit tissue-specific expression and this result is confirmed by the findings from other plant species such as kiwifruit *(Actinidia deliciosa*) (Yin et al. 2008), strawberries (Trainotti et al. 2005), muskmelon *(Cucumis melo*) (Sato-Nara et al. 1999) and tomato (Lashbrook et al. 1998; Tieman and Klee 1999) among others. There are those however that show specific expression pattern at particular developmental stages as observed in fruits such as *SlNR* in tomato (Lashbrook et al. 1998) and in different fruit parts like *CmETR1* and *CmERS1* in muskmelon (Sato-Nara et al. 1999).

The overall expression profile showed that the gene expression level of the oil palm ethylene receptor genes differed in each of the tested tissues including in the mesocarp at young and ripening stages as well as the AZ. The expression of the receptors seems to be higher in the reproductive tissues as opposed to the vegetative tissues. The highest expression of most of the oil palm ethylene receptors is in the female flower and is similar to *AtETR2* in Arabidopsis (Hua et al. 1998), *SlETR4* in tomato (Tieman and Klee 1999) and *CpERS1* in zucchini squash (*Cucurbita pepe L)* (Martínez et al. 2013). The constitutive expression of *EgETR5* also matched to *AtETR1* and *AtERS1* in Arabidopsis (Hua et al. 1998) and *SlETR1* in tomato (Lashbrook et al. 1998). The difference in the expression profile of the oil palm ethylene receptor genes may provide a clue as to the possible roles of the receptor(s) in the ethylene signalling pathway in different developmental

events. In particular, the receptor(s) response towards ethylene that plays multiple roles in regulating the physiological and developmental processes throughout the plant life cycle (Payton et al. 1996).

Just like other fruits, the development of the oil palm mesocarp is divided into a series of stages, starting with cell division in young fruits until cell separation in ripe fruits (Tranbarger et al. 2011). In the mesocarp fruit developmental stages from young until ripening, the expression pattern of the oil palm ethylene receptors was different. At 8 WAA, *EgETR2* and *EgERS3* were expressed at a slightly higher level than that of 10 and 12 WAA in the mesocarp tissues. This period represents active cell division and expansion which is associated with the beginning of fruit growth especially in terms of increase mass and size (Tranbarger et al. 2011). Similarly, this expression pattern was also observed in *AdETRS1* and *AdETR3* in kiwifruit (Yin et al. 2008), *VvERS1* and *VvEIN4* in grape (Chervin and Deluc 2010) and *DkERS1* in persimmon (*Diospyros kaki Thumb*) (Pang et al. 2007). At 12 WAA, EgETR5 was expressed the highest in comparison to the rest of the ethylene receptors. The 12 WAA or approximately 120 days after process is an active period of cell expansion that involves increases in fruit mass and size (Tranbarger et al. 2011). This is also a period where oil synthesis starts in the oil palm mesocarp (Sambanthamurthi et al. 2000). This finding suggest the possible of the involvement of *EgETR5* in ethylene perception during the cell expansion phase of the fruits particularly in the parenchymatous cells of the mesocarp tissue. The expression of three oil palm ethylene receptor genes follows the ripening pattern of the oil palm fruits. The expression of *EgETR2, EgETR3,*

*EgERS1* and *EgERS3* is highest at 20 WAA, which suggest the involvement of these receptors in the ripening process of the oil palm fruit mesocarp. The ripening-related gene expression is likely to be the consequence of increased ethylene production in the ripening mesocarp tissues.

The differential expression pattern demonstrated by the oil palm ethylene receptors somehow relates to the high ethylene levels in floral development and early stages of fruit set in tomato, decreases in later stages and increases again at the onset of fruit ripening (Srivastava and Handa 2005). The increase in the expression level of the ethylene receptors in tomato and other climacteric fruits is associated with the system 2 ethylene peak during the ripening phase (Klee 2004; Agarwal et al. 2012). In the oil palm fruits, the varying expression profiles of the oil palm receptor genes suggest their involvement in the two systems of ethylene production in climacteric plants. The high expression of the receptors in the young and ripening stages showed that the receptors are either associated with basal ethylene production in system 1 that plays a role in early vegetative growth or system 2 that is functional during ripening following the autocatalytic burst of ethylene in climacteric fruits (Alexander and Grierson 2002). Similarly, transcriptome analysis on oil palm fruits also showed that ethylene-related transcripts have expression profiles associated with system 1 and system 2 amount of ethylene production (Tranbarger et al. 2011). The various expression profiles of the ethylene receptors in the various oil palm tissues either during young, ripening or throughout the fruit developmental stages may also be due to the overlapping and non-overlappping roles of the receptor in modulating the ethylene signals during oil palm fruit

development. In tomato, suppression of the ethylene receptor *LeETR4* showed early ripening as opposed to the never-ripe *LeNR* demonstrating that the different members of the ethylene receptor family play non-overlapping roles in controlling fruit ripening (Kevany et al. 2008). Mutations in the the receptor ETR2, ERS1 and EIN4 demonstrated their overlapping role in controlling Arabidopsis growth especially on the well-characterized ethylene response effect such as the triple respond (Hua et al. 1995; Sakai et al. 1998).

Application of exogenous ethylene and ethylene antagonists such as 1-MCP are widely used to understand many ethylene-regulated developmental processes including fruit development and ripening (Blankenship and Dole 2003; Bouzayen et al. 2010). Depending on the receptors, the expression can either be up-regulated, down-regulated or made non-responsive by exposure to different combinations of ethylene and 1-MCP. For instance, exposure of propylene (an ethylene analogue) to peach at young (S1) and ripe (S4) stages increased the expression of *PpERS1* and the expression was reduced in 1-MCP treated fruits suggesting an important role in the peach ripening process (Rasori et al. 2002). In 8 h ethylene-treated tomato fruits, the expression of *SlETR6*, *SlETR4* and *SlNR* increased between 7 to 9 fold but at the protein level, rapid protein degradation was observed (Kevany et al. 2007). The *MdERS1* and *MdERS2* genes in apple also exhibited a similar pattern with an increase in the transcription and decrease in the protein level (Tatsuki et al. 2009). The differences between these two levels is due to the turnover of receptors in the presence or absence of ethylene which determines the

timing of fruit ripening (Kevany et al. 2007). There are also those such as *PpETR1* in peach fruits (Rasori et al. 2002) and *PsETR1* in passion fruit (*Passiflora edulis Sims)* (Mita et al. 1998) that remained unresponsive to both ethylene and 1-MCP treatments. *AdETR1* in kiwifruit on the other hand is down regulated after exposure to ethylene (Yin et al. 2008). The unresponsive or down regulation of some of the ethylene receptor genes in ethylene-induced fruits suggest that although these genes may not specifically be involved in fruit development and ripening they may play a role in sensing the ethylene signal and transmitting the response to other members of the receptor family thus activating the ethylene signalling pathway.

Core signalling genes located downstream of the ethylene receptors in the ethylene signalling pathway also exhibit differential gene expression patterns in response to ethylene during ripening. *AdCTR1* and *AdCTR2* in ripening kiwifruits showed little changes in expression when exposed to ethylene but responded differently to 1-MCP (Yin et al. 2008). *AdCTR2* was unresponsive to 1-MCP treatment, whereas the expression of *AdCTR1* was suppressed during fruit ripening. In ripening pear fruit, the expression of *PcCTR1* increases after exposure to ethylene (El-Sharkawy et al. 2003). In apple, ethylene treatment increased the expression of *CTR1*-like and *EIN2A* genes as well as ethylene biosynthetic genes such as *ACO2, ACS3* and *ACO3*, while 1-MCP treatment reduced the expression of *CTR1, EIN2A, EIL4* and ERFs genes as well as *ACS1, ACO1* and *ACO2* ethylene biosynthesis genes (Yang et al. 2013). This is some of the evidence that

demonstrates the complex interaction at the transcriptional level among the receptors especially in terms of their ethylene responsiveness.

Leaf, flower and fruit abscission is amongst the processes triggered by ethylene (Reid 1985). Previous studies in oil palm suggest that ethylene induces cell separation in the primary AZ of the oil palm fruit (Henderson and Osborne 1994). A *PG4* gene encoding polygalactoronase was highly expressed in the AZ of oil palm fruits exposed to ethylene suggesting that PGs play an important role in the separation of oil palm fruits (Roongsattham et al. 2012). Roongsattham et al. (2012) also demonstrated exposure of ethylene to oil palm fruit spikelets sampled from the centre of each fruit bunch resulted in synchronized fruit shedding. However, the time for fruits to detach from the spikelets depends on their developmental age. Looking at the overall gene expression profile of the oil palm ethylene receptors in the AZ from the apical, central and basal part of the ripe fruit bunches, differential expression patterns were observed. The transcript level in most of the receptors was higher in 20 WAA than the 22 WAA fruits which is similar to the total ethylene production in the clonal palm P165, used in the earlier study. This suggests that 20 WAA maybe associated with the peak of ethylene production in the tested oil palm fruit bunch although this may depend on the genotype as ripening of oil palm fruit bunches can take up to 24 WAA (Kaida and Zulkifly 1992; Tranbarger et al. 2011).

In the AZ of ethylene treated peach fruit the expression level of the ethylene receptor *ppERS1* increased up to four fold (Rasori et al. 2002). In mango (*Mangifera indica L.*) the expression of *MiERS1* is induced by ethephon in ripening stages but is further induced in the AZ suggesting a

more important role in the regulation of mango fruit abscission than fruit ripening (Ish-Shalom et al. 2011). With a much more comprehensive study on the various positions (inner, middle and outer fruits) within the spikelets from the various location (basal, central and proximal) in oil palm fruit bunches, it may be possible to enhance further our understanding of the role of ethylene receptors in the fruit abscission process.

The work presented in this chapter focuses on identification, isolation and characterisation of the ethylene receptor family, the first component involved in the ethylene receptor pathways. The result of the work presented here demonstrates the potential use of the oil palm genome sequence to be used as a resource for determining multigene families playing important roles during fruit development and shedding in oil palm.

# CHAPTER 4

# IDENTIFICATION, ISOLATION AND CHARACTERISATION OF LIPASE CLASS 3 GENE (*FLL1*) AND PROMOTER AND ITS CORRESPONDING GENE FAMILY

## 4.1 Introduction

Lipase class 3 is part of the triacylglycerol lipase family involved in lipid degradation, esterification, and transesterification processes in plants. Previously, a lipase class 3 EST gene was discovered from a 17-week-old mesocarp cDNA library (Nurniwalis 2006; Nurniwalis et al. 2008). The EST was found to be predominantly expressed in the tested mesocarp tissues and follows the fruit ripening process which also coincides with the oil synthesis period. Hence, 5` RACE was performed to amplify the 5` region of the gene and reconstitution of both 5` RACE and *O65EST* sequence information showed that the gene, designated as *FLL1* encodes 483 amino acid (aa) (Nurniwalis et al. 2007). To understand the potential role of this gene and its family in the degradation of palm oil hence, affecting oil quality, this chapter describes on the isolation and characterisation of the *FLL1* gene. Based on the high expression in the mesocarp tissues, the work was further extended to isolate and analyze its corresponding promoter *via in silico* and transient assay. This was targeted at the potential use of the promoter as a tool for oil palm genetic engineering. In addition, this chapter also describes the identification of other putative members of the lipase class 3 gene family based on the oil palm genome sequence data. Part of

the work described in this chapter has been published (Nurniwalis et al. 2015) and is as attached in the thesis (Appendix 6).

## 4.2 Materials and methods

### 4.2.1 Plants and treatments

Information on the various tissues used in this study is described in section 3.2.2. For cold induced treatment of the oil palm fruits, the fruits were incubated at 7°C for 5 h prior to RNA extraction which is described in section 3.2.4.

### 4.2.2 LD-PCR

The full-length cDNA of the lipase class 3 gene was generated *via* LD-PCR using a GeneRacer™ RACE Ready cDNA Kit (Invitrogen, USA). Touchdown PCR amplification was carried out in 50 µl of a reaction mixture containing 0.1 µg of cDNA template, 1x Advantage 2 PCR buffer (Clontech, USA), 1x Advantage 2 polymerase mix (Clontech, USA), 0.2 mM dNTP mix (Clontech, USA), 0.2 µM each of gene-specific primers LF6 and LAS10 (Appendix 4) with the following conditions: 94°C (5 sec) and 72°C (3 min) for 5 cycles; 94°C (5 sec), 70°C (10 sec), and 72°C (3 min) for 5 cycles; 94°C (5 sec), 68°C (10 sec), and 72 °C (3 min) for 22 cycles, and a final extension at 72°C for 7 min. The amplified PCR products were analysed, purified, cloned and the plasmid isolated and digested as described in sections 3.2.6 until 3.2.8.

### 4.2.3 Sequence analyses

DNA sequencing and sequence analyses were carried out as described in section 3.2.9. Additionally, prediction of subcellular localization was performed using the Target P programme v. 1.1 (http://www.cbs.dtu.dk/services/TargetP). The protein secondary structure prediction was carried out using the PsiPred programme (http://bioinf.cs.ucl.ac.uk/psipred). The protein 3D model was predicted using 3D-JIGSAW (http://bmm.cancerresearchuk.org/ ~3djigsaw/) and I-Tasser (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) and viewed using the RasMol latest version (http://www.umass.edu/microbio/rasmol/getras.htm).

### 4.2.4 Expression analyses *via* RT-PCR

RT-PCR was performed to determine the expression pattern of *FLL1* in the mesocarp of cold induced oil palm fruits. The single stranded cDNA was synthesized from 16 and 22 WAA mesocarp tissues as described in section 3.2.16.1. PCR amplification was carried out in 25 µl of a reaction mixture containing 0.1 µg of cDNA, a 1x Advantage 2 PCR buffer (Clontech, USA), a 1x Advantage 2 polymerase mix (Clontech, USA), a 1x dNTP mix (Clontech, USA), 0.2 µM LF1 and 0.2 µM LR1 primers (Appendix 4) with the following conditions: denaturation at 95$^{o}$C (1 min); 95°C (30 sec) and 60$^{o}$C (1 min) for 30 cycles; and a final extension at 60°C for 1 min. Amplification of the actin gene as internal control is described in section 3.2.17.

4.2.5 Genomic DNA amplification, generation and labelling of 3` UTR probe and Southern analyses

Genomic DNA was isolated as described in section 3.2.14. Amplification of the *FLL1* genomic region was carried out in a 50 µl reaction mixture containing 0.1 µg genomic DNA, a 1x Expand High Fidelity buffer with 15 ml $MgCl_2$ (Roche, Germany), a 0.7 U Expand High Fidelity enzyme mix (Roche, Germany), a 1x dNTP mix (Clontech, USA), and 0.2 µM LF6 and 0.2 µM LAS10 primers (Appendix 4) with the following two-step PCR conditions: an initial denaturation at 95°C (1 min); denaturation at 95°C (30 sec), simultaneous annealing and extension at 65°C (1 min) for 30 cycles, and a final annealing and extension at 65°C for 3 min.

The 3' UTR probe was amplified using gene specific primers LF3 and LR1 (Appendix 4) designed based on the 3' UTR region of the *FLL1* cDNA. The reaction was carried out in a 50 µl reaction mixture containing 5 µl of 10X BD Advantage 2 buffer, 1 µl of 50X dNTP (10mM), 1 µl each of LF3 and LR1 primers (10µM each), 1 µl of 50X BD Advantage 2 Polymerase Mix, 0.5 µg of genomic DNA and PCR grade water with the following conditions: 95°C (1 min); 95°C (30 sec), 60°C (3 min) for 30 cycles; a final extension at 60°C for 3 min. The probe was then labelled with $\alpha^{-32}$P dCTP by random primer reaction using a Megaprime DNA Labelling System Kit (Amersham, USA). A total of 28 µl mixture containing 30 ng of the respective fragment, 5 µl of random primer and sterile water was denatured at boiling point (100°C) for 10 min and left to cool down at room temperature for 10 min. Next, 10 µl of labelling buffer, 2 µl of 2U/µl Klenow enzyme and 5 µl of $\alpha$-$^{32}$P dCTP was added to the reaction mixture and incubated at 37°C for 30

min. After incubation, the probe was transferred to a chromaspin TE column (CLONTECH, USA) and centrifuged at 700 x g for 5 min to separate the labelled and unlabelled probe. In a microcentrifuge tube, the labelled probe was collected and 100 µg/ml of herring sperm DNA was mixed together prior to heating the mixture in boiling water for 5 min, followed by quick chilling on ice for 10 min. Once denatured, the labelled probe can be used immediately.

Southern blot analysis was performed to determine the presence of lipase class 3 gene(s) in the oil palm genome. For Southern blot preparation, the genomic DNA was digested individually with *BamHI, EcoRI, SpeI,* and *XbaI* (Fermentas, Germany), respectively. Digestion of the DNA was performed in a 20 µl reaction mixture containing 15 µg DNA, 0.4 µl of 0.1M spermidine, 2 µl of restriction buffer, 2 µl of enzyme (10U/µl) and sterile water at 37$^{o}$C overnight. The next day, 4 µl of loading dye was added to the digested DNA and the mixture was electrophoresed on 1% (w/v) agarose gel for 3 hrs at 100V in 1X TAE buffer, pH 7.9.

The transfer and fixation of the digested DNA to membrane was performed using the standard method of Sambrook and Russell (2001). Transfer of digested DNA to nylon charged membrane (Hybond-N$^{+}$, Amersham, USA) was carried out overnight by capillary blotting using 4N NaOH as the transfer buffer. The membrane was soaked in 2X SSC for 5 min and the DNA was fixed to the membrane by cross-linking through exposure to 1200 kJ UV light for 1 min using a UV crosslinker (CL-1000 Crosslinker, UVP, Inc.).

The membrane was pre-hybridized in a rotating hybridization tube containing 20 ml of hybridization buffer [5X SSC, 5X Denhardt, 0.5% (w/v) sodium dodeocyl sulphate (SDS)] and denatured herring sperm DNA (100 µg/ml) for 4 hrs at 65$^{o}$C. After removal of the pre-hybridization buffer, the membrane was then subjected to hybridization with the 3' UTR cDNA labelled probe in high stringency conditions (5x SSC, 5x Denhardt, 0.5% SDS at 65°C overnight) and then washed with 2x SSC and 0.1% SDS at 65°C for 10 min, 1x SSC and 0.1% SDS at 65°C for 15 min, and 0.5x SSC and 0.1% SDS at 65°C for 20 min. The membrane was exposed to X-ray film overnight at -80$^{o}$C.

4.2.6 Promoter isolation

The upstream genomic region of *FLL1* was further amplified to isolate its corresponding promoter using universal Genome Walker and Advantage 2 PCR kits (Clontech, USA). Four Genome Walker libraries were constructed *via* the digestion of genomic DNA with *DraI, EcoRV*, PvuII dan *StuI* prior to ligation with the Genome Walker adaptor. Primary PCR amplifications were performed using all four Genome Walker libraries as the DNA template. Each reaction mixture contained 100 ng of Genome Walker library, a 1x Advantage 2 PCR buffer (Clontech, USA), a 0.2 mM dNTP mix (Clontech, USA), 0.2 µM each of primers AP1 and LAS11 (Appendix 4), a 1x Advantage 2 polymerase mix (Clontech, USA), in a two-step cycle parameters: 94°C (25 sec) and 72°C (3 min) for 7 cycles; 94°C (25 sec) and 67°C (3 min) for 32 cycles, and a final extension at 67°C for 7 min. The primary PCR product was diluted 50x and used as DNA template for

secondary PCR. The secondary PCR mixture was similar to that of the primary PCR except the primers used were primers AP2 and LAS14 (Appendix 4). The PCR reaction was performed in 5 cycles of 94°C (25 sec) and 72°C (3 min), followed by 20 cycles of 94°C (25 sec) and 67°C (3 min), and a final extension at 67°C for 7 min.

4.2.7 Promoter-vector construct and transient assay analysis

Plasmid *FLL1*/*GUS* was constructed by replacing the CaMV 35S promoter contained in pBI221 with the *HindIII-Xba1* flanked *FLL1* promoter sequence from the position -664 to 83 bp. The amplification of the *HindIII-Xba1* flanked *FLL1* promoter region was performed in 50 µl of a reaction mixture containing 25 ng of plasmid DNA, a 1x Expand HF buffer with 15 mM $MgCl_2$ (Roche, Germany), a 0.2 mM dNTP mix (Clontech, USA), 0.1 µM each of gene-specific primers PLF2 and PLR1 (Appendix 4), and a 0.1 U HF Enzyme mix (Roche, Germany) under the following conditions: 94°C (3 min), 94°C (1 min), 57°C (1 min), and 72°C (90 sec) for 20 cycles, and a final extension at 72°C for 10 min. The cloning procedure was confirmed *via* *HindIII-Xba1* restriction enzyme analysis and the insert was verified by sequencing. The preparation of target materials for transformation, the bombardment parameters, and the GUS histochemical assay for transient promoter analysis were carried out as described by Zubaidah and Siti NorAkmar (2003).

### 4.2.8 Lipase class 3 gene family search

To search for other putative members of the lipase class 3 family, a HMM profile was build using the amino acid sequence of the conserved lipase class 3 domain (Pfam ID: PF01764) downloaded from Pfam (http://pfam.xfam.org) and searched against the translated transcriptome and gene model of oil palm (Singh et al. 2013a) using HMMER3 (Eddy 2011) hmm search programme. Sequence similarity searches were performed as described in section 3.2.9. Functional annotations of the putative proteins were performed using Blast2GO (https://www.blast2go.com).

## 4.3 Results

### 4.3.1 Isolation of the full-length *FLL1* gene

Reconstitution of the nucleotide sequences from the 5` RACE (Nurniwalis et al. 2007) and EST (Nurniwalis 2006; Nurniwalis et al. 2008) cDNAs have resulted in the amplification of the full-length *FLL1* cDNA *via* LD-PCR (Figure 4.1) . The full-length cDNA was 1721 bp long and contained an ORF of 1452 bp. The nucleotide and deduced amino acid sequence is shown in Figure 4.2. The detailed *FLL1* nucleotide sequence information and analysis has been described earlier by Nurniwalis et al. (2007). Nucleotide sequence alignment between *FLL1* and the oil palm transcriptome data from 15 WAA mesocarp tissues (Singh et al. 2013a) showed a 100% identity with an e-value of 0.0. Public database searches for homologies against the nucleotide and protein sequences showed that *FLL1* is 99 and 100%, respectively, identical to *EgLIP1* (Morcillo et al. 2013), a lipase class 3 cDNA (Accession No. AFV50601.1) from oil palm at the nucleotide and protein level (97% query coverage). The next closest and the following matches to the FLL1 deduced amino acid sequence share a much lower homology with a putative lipase (identity of 51% and an e-value of $6e^{144}$) from rice (Accession No. NP_001054678.2) followed by lipase (44% identity and an e-value of $3e^{-128}$) and triacylglycerol lipase (43% identity and an e-value of $2e^{-124}$) from castor bean (Accession Nos. AAV66577.1 and XP_002533321.1). The conserved domain and homology search showed that FLL1 contained the lipase consensus sequence [LIV]X[LIVAFY][LIAMVST]G[HYWV]SXG[GSTAC], that encoded the highly conserved lipase class 3 domain (Pfam ID: PF01764). The region of

sequence similarities detected all three putative amino acid residues that form the lipase catalytic triad. The serine (S) residue was detected at position 308 and the aspartic acid (D) residue at position 368. Histidine (H), the third residue to complete the catalytic triad, had more diverse flanking sequences, but the exact H position was predicted at location 464 based on consensus sequence alignments (Figure 4.3). In addition, putative active site lids at positions 223, 224, 226, 229-236 as well as nucleophilic elbow at positions 306 - 310 were detected in FLL1 (Figure 4.3).

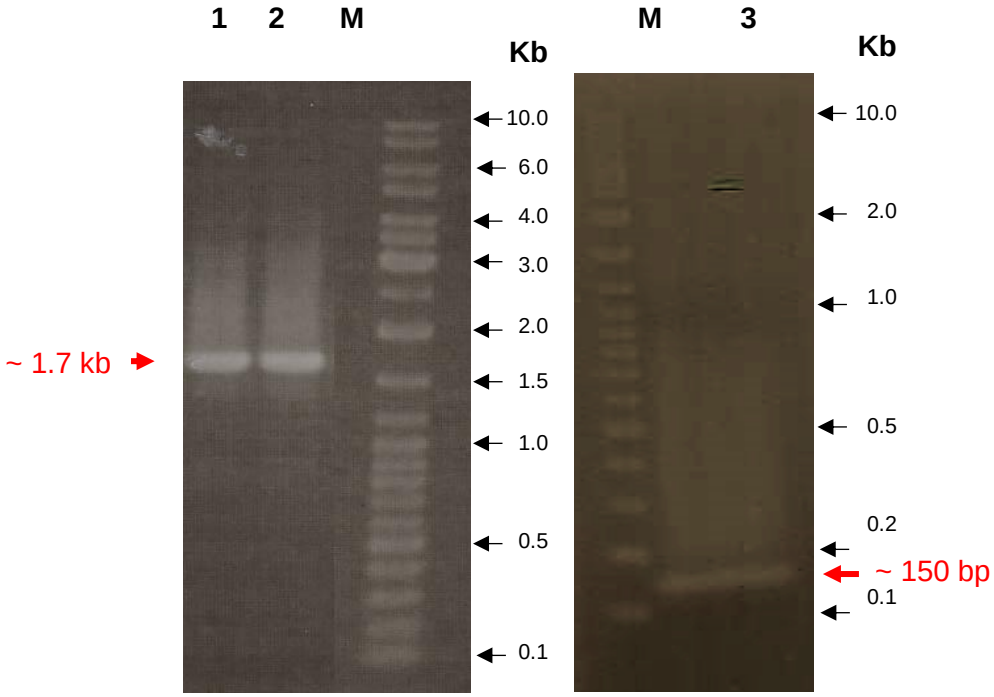



**Figure 4.1: Analysis of the amplified full-length *FLL1* and 3' UTR region by electrophoresis on 1% (w/v) agarose gel.** Lanes 1 and 2 = full-length *FLL1* fragment, Lane 3 = the 3' UTR fragment. M = Gene Ruler DNA Ladder Mix Marker (Fermentas).

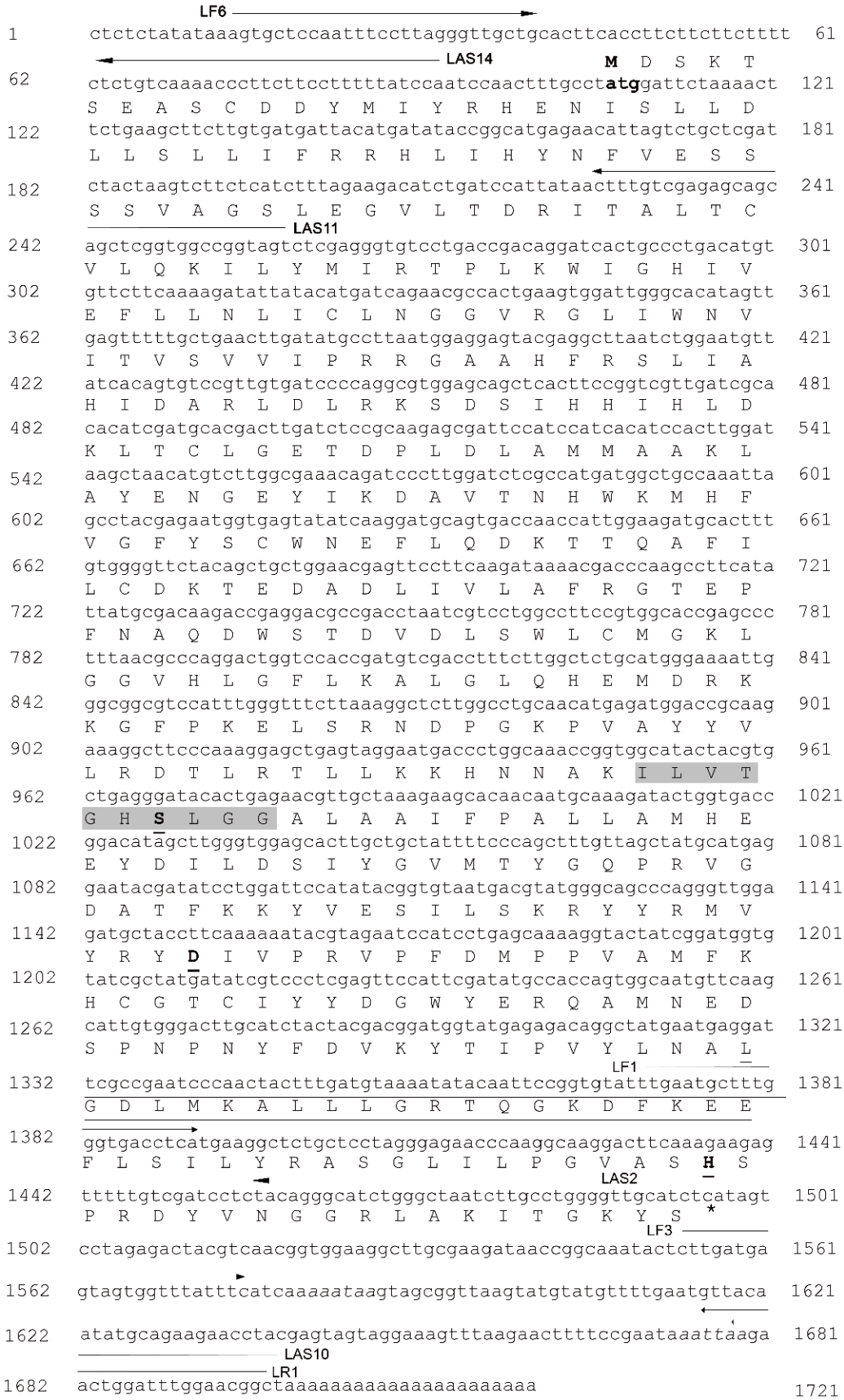

LF6
LAS14
LAS11
LF1
LAS2
LF3
LAS10
LR1

**Figure 4.2: Nucleotide and deduced amino acid sequences of *FLL1.*** The nucleotide sequence is shown in lower case letters, whereas the amino acid sequence is in capital letters. Numerals at both ends of the nucleotide row indicate a nucleotide number. The first start codon (atg) is bold and the stop codon (tga) is marked with an asterisk, respectively. The putative polyadenylation signals are in italics. The serine, aspartic acid, and histidine residues that form the catalytic triad are in bold and underlined. The lipase consensus motif surrounding the active serine residue is shaded. Arrows represent sense (LF6, LF1, and LF3) and antisense (LAS10, LAS2, LR1, LAS11, and LAS14) primer sequences designed for 5` RACE, LD-PCR, RT-PCR, and promoter isolation. The overlapping region between 5` RACE and O65EST cDNA is underlined.


**Figure 4.3: Multiple sequence alignment of FLL1 against the top four BLASTX hits in the NCBI sequence database.** The ':' denotes conservation of strong groups. The '.' denotes conservation of weak groups, whereas those without any symbol denote no consensus (CLUSTWALW, Biology Workbench Version 3.2, the University of California). The lipase consensus motif is boxed, whereas the serine, aspartic acid, and histidine residues that forms the putative catalytic triad is marked with the '^'.

### 4.3.2 FLL1 protein prediction

The deduced amino acid sequence of FLLI was predicted to have a secondary structure that consists of 12 β-sheets, 14 α-helixes, and 25 coils (Figure 4.4). FLL1 had more hydrophobic residues (47.4%), and a Kyte-Doolittle hydropathy profile showed the predicted protein was predominantly hydrophobic. The three-dimensional protein prediction using 3D-JIGSAW detected three family domains (PB073089, PF01764, and PB108722) in FLL1 but only PF01764 (the lipase class 3 domain) was homologous to FLL1. FLL1 had an identity of 50% and an e-value of $1e^{-55}$ to 3ngmA, a crystal structure of lipase from *Gibberella zeae* from location 64 to 221. The structural protein prediction of FLL1 using I-Tasser generated five 3-D output models, ranking model 1 as best predicted Zhang (2008) with a C-score more than -1.5 to indicate correct protein folding (Roy et al. 2010). Figure 4.5 represents the best predicted model 1 with a C-score value of 2.83 based on the highest structural alignment to a lipase protein from *Gibberella zeae* (PDB hit:3ngmA). The predicted EC number for FLL1 is 3.1.1.3, which represents a triacylglycerol lipase. Based on gene ontology (GO), the *FLL1* gene has a molecular function that is involved in triglyceride lipase activity (GO:0004806). No signal peptide or transmembrane domains were identified in FLL1 using various subcellular localization prediction searches, which would suggest that FLL1 is cytoplasmic.

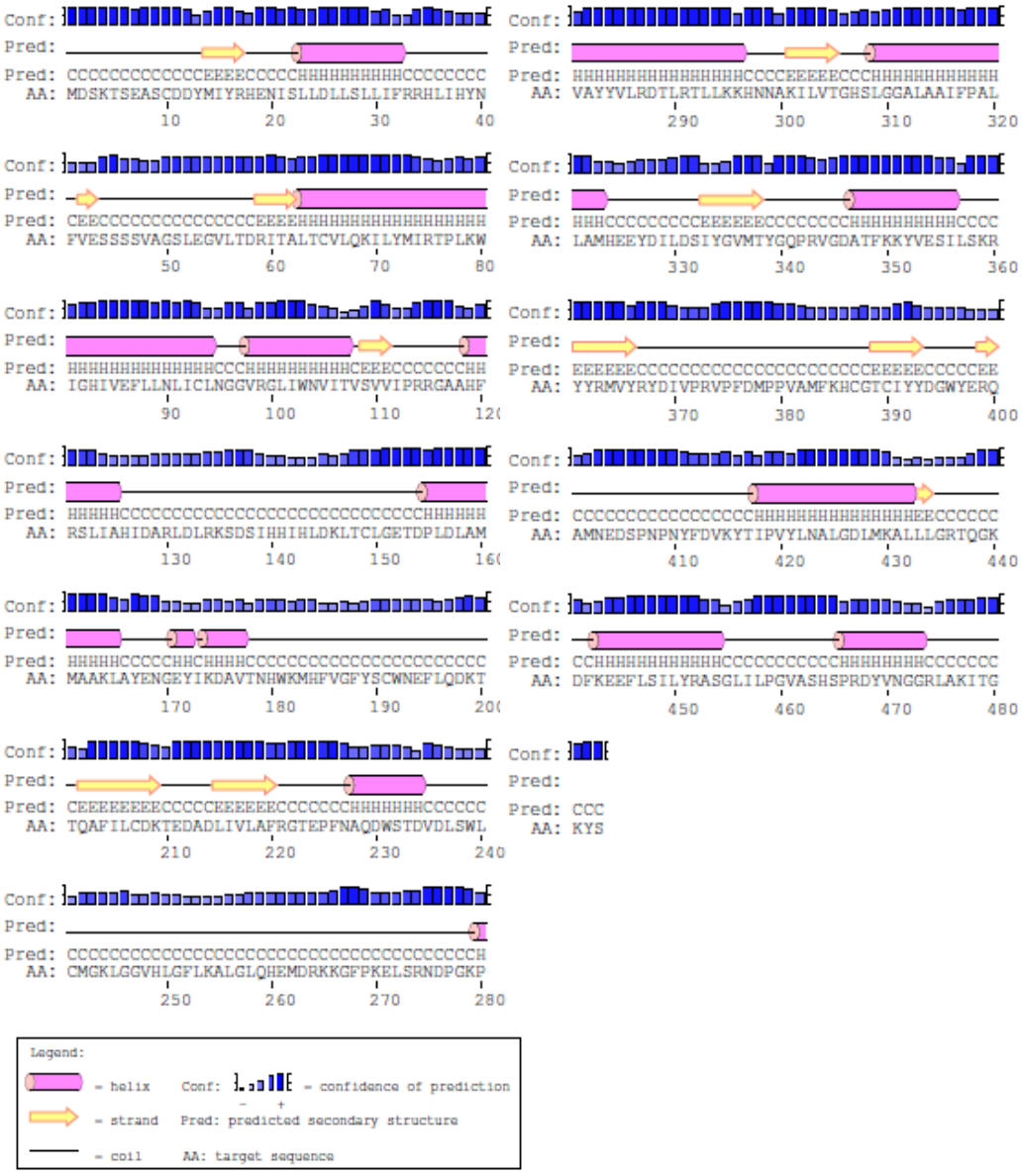



**Figure 4.4: Prediction of secondary structure for FLL1 using PsiPRED.**

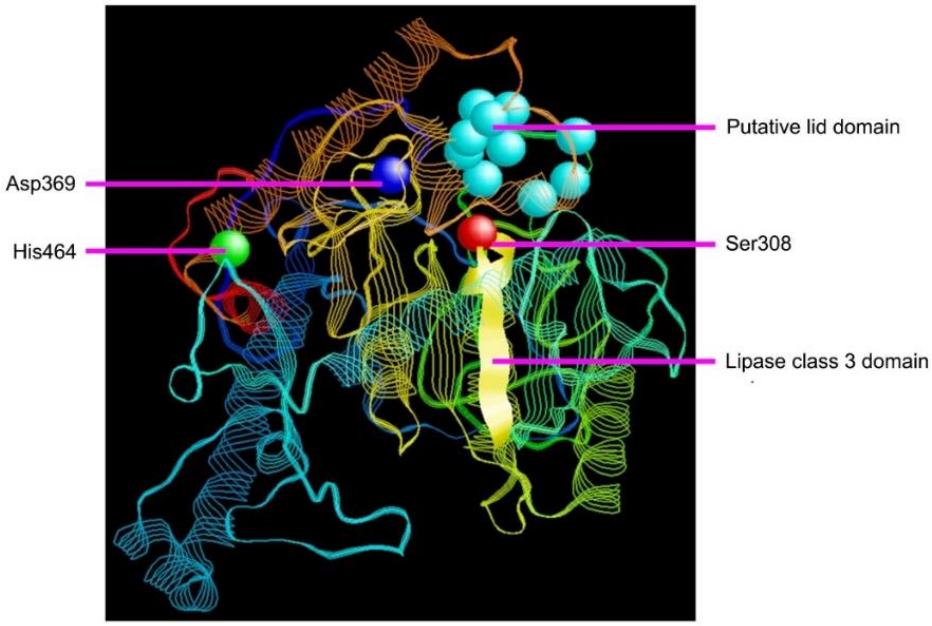



**Figure 4.5: A three-dimensional best model for FLL1 predicted protein selected based on a C-score using I-Tasser.** The putative lipase catalytic triad Ser308, Asp369, and His464 (represented as the balls), the lipase class 3 conserved region (represented as the ribbon) and the putative lid domain (represented as the balls) are indicated. The model is presented using RasMol v. 2.6.

### 4.3.3 Expression analysis *via* RT-PCR

Comparison between normal and cold treated fruits at two stages of fruit development (16 and 22 WAA) showed that both the normal and cold treated fruits had a similar expression pattern in the mesocarp tissues. Both normal and cold treated fruits where expressed highly at 22 WAA than at 16 WAA. However, the expression of the *FLL1* gene was found to be even higher in the mesocarp of cold induced fruits than in the normal fruits at both stages of fruit development (Figure 4.6).

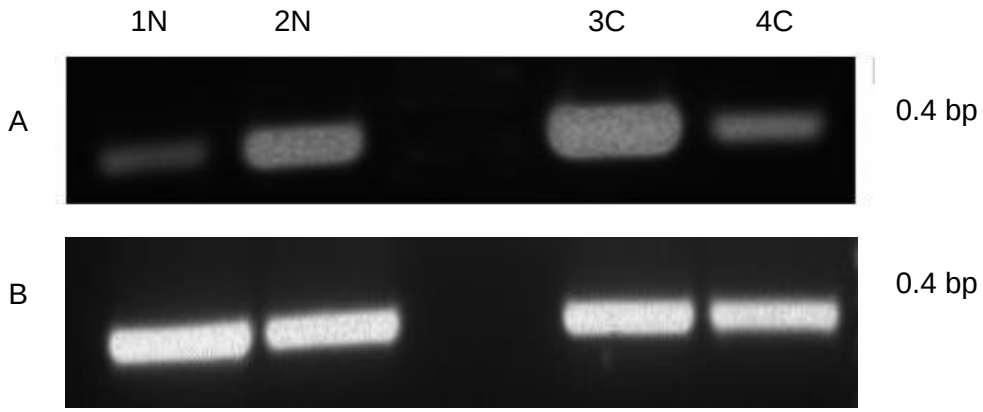



**Figure 4.6: Expression pattern of the A) *FLL1* gene and B) actin gene in the mesocarp of normal and cold induced fruits at two stages of fruit development *via* RT-PCR.** 1 and 4 = fruits at 16 WAA whereas 2 and 3 = fruits at 22 WAA. N = normal and C = cold treated fruits.

### 4.3.4 *FLL1* genomic amplification and Southern analysis

The amplification of the *FLL1* genomic sequence yielded a 5.5 kb fragment (Figure 4.7). The genomic organization and presence of *FLL1* in the oil palm genome were confirmed through a sequence alignment between the amplified *FLL1* genomic sequence and the p5-sc00064 of the fifth genome build [EG5-build] (Singh et al. 2013a). The result showed a 100% identity at the nucleotide level with an e-value of 0.0. An exonerated search with a 60% self-score threshold revealed *FLL1* at chromosome 3 in the EG5 build. The comparison of the *FLL1* genomic sequence and oil palm microsatellite genomic DNA by (Morcillo et al. 2013) (Accession No. HE661587.1) revealed a 97% identity suggesting that both regions are identical. The amplification of the *FLL1* genomic sequence and an exonerated search revealed the presence of four introns located in between 5 exons in *FLL1*. The size of the introns were 1424, 555, 1626, and 179 bp and the five exons 382, 224, 34, 620, and 251 bp, respectively. The intron-

exon boundaries in *FLL1* conformed to the universal GT-AG rule for introns starting with a GT dinucleotide and ending with an AG dinucleotide (Breathnach and Chambon 1981). A schematic representation of the *FLL1* gene structure is shown in Figure 4.8.

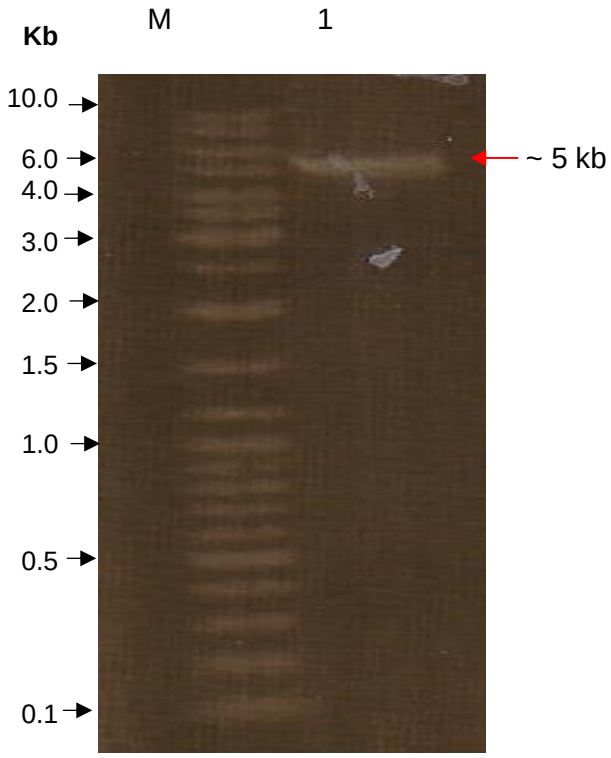



**Figure 4.7: Analysis of the amplified genomic *FLL1 gene* by electrophoresis on 1% (w/v) agarose gel.** Lane 1 = genomic *FLL1* fragment. M = Gene Ruler DNA Ladder Mix Marker (Fermentas).

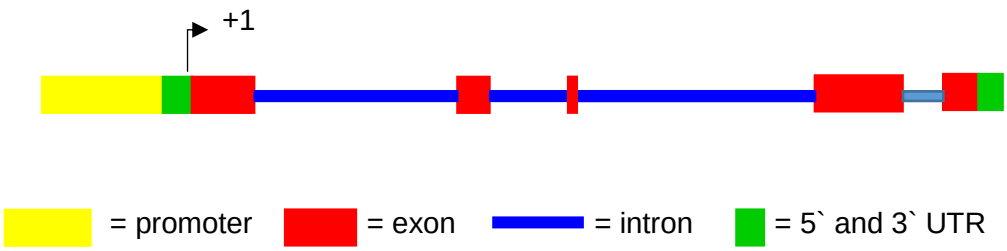



**Figure 4.8: Genomic structure of the *FLL1* gene.** Position 1 is assigned to the putative transcription start site.

For Southern blot analysis, the 3` UTR probe was generated *via* PCR amplification that produced a PCR fragment with the size of approximately 150 bp as shown in Figure 4.1. Hybridization was performed at high stringency conditions where more than two bands were detected in all the restriction enzyme-digested oil palm genomic DNAs (Figure 4.9). This result indicates that *FLL1* was not a single copy gene and that it belonged to a multi gene family.

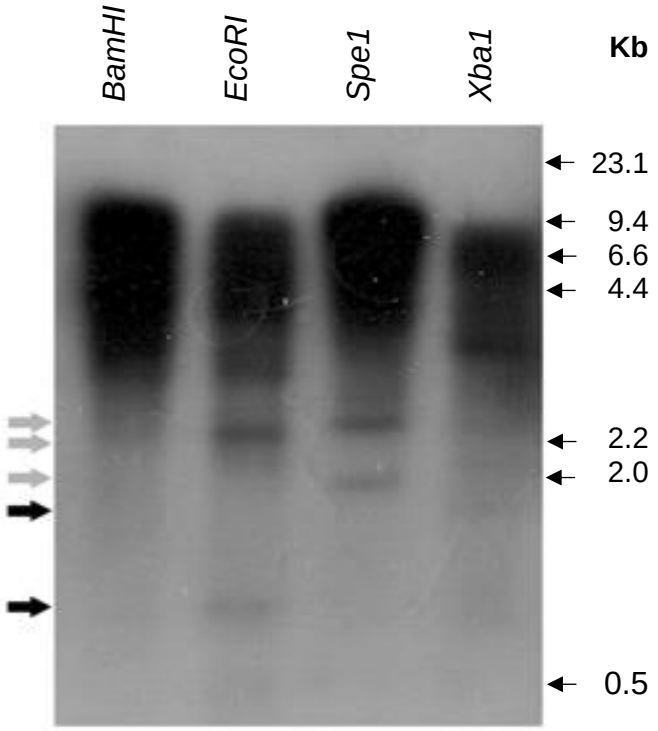



**Figure 4.9: The Southern analysis of oil palm genomic DNA. Blots were hybridized with a gene-specific sequence based on the 3`UTR of the *FLL1* gene**. Strong and faint signals are shown by the *grey* and *black arrows*.

### 4.3.5 Promoter amplification and transient assay

PCR amplification of the upstream genomic region of the *FLL1* gene resulted in the amplification of a 756 bp fragment. Figure 4.10 represents the nucleotide sequence of the amplified genomic fragment containing the putative *FLL1* promoter. This genomic sequence contained 671 bp of promoter sequence and 85 bp of 5' UTR. An adenine residue located furthest at the 5' terminal of *FLL1* was selected as TSS. The adenine residue is most likely the TSS as the pyrimidine sequences flanking the adenine residue are often the preferred TSS motif in plants {C/TAC/T} (Joshi 1987). The putative TSS motif also matches most of the TSS motifs in highly expressed genes in plants (Sawant et al. 1999).

The distribution of sequences in the *FLL1* promoter is divided into three main groups which contain the TATA-box, pyrimidine patch (Y-patch), and regulatory element group (REG) (Yamamoto et al. 2007). The putative TATA-box, TATATATTA, was present 37 bp upstream of the TSS, consistent with the distance of 32 ± 7 bp (Joshi 1987b). The Y-patch motif was located within 100 bp upstream of the TSS and downstream of the TATA-box. The role of the Y-patch motif, a plant-specific core element, is unknown but the local distribution of short sequence analysis showed that it makes up one of the general/important components in the core promoters of dicots and monocots (Yamamoto et al. 2007). The REG group contains *cis*-acting regulatory elements (CARE) that correspond to known transcriptional regulatory sequences. In the *FLL1* putative promoter, a number of putative CAREs were identified, which included those that respond to environment signals as well as those required to direct specific

expressions to specific tissues (Table 4.1).

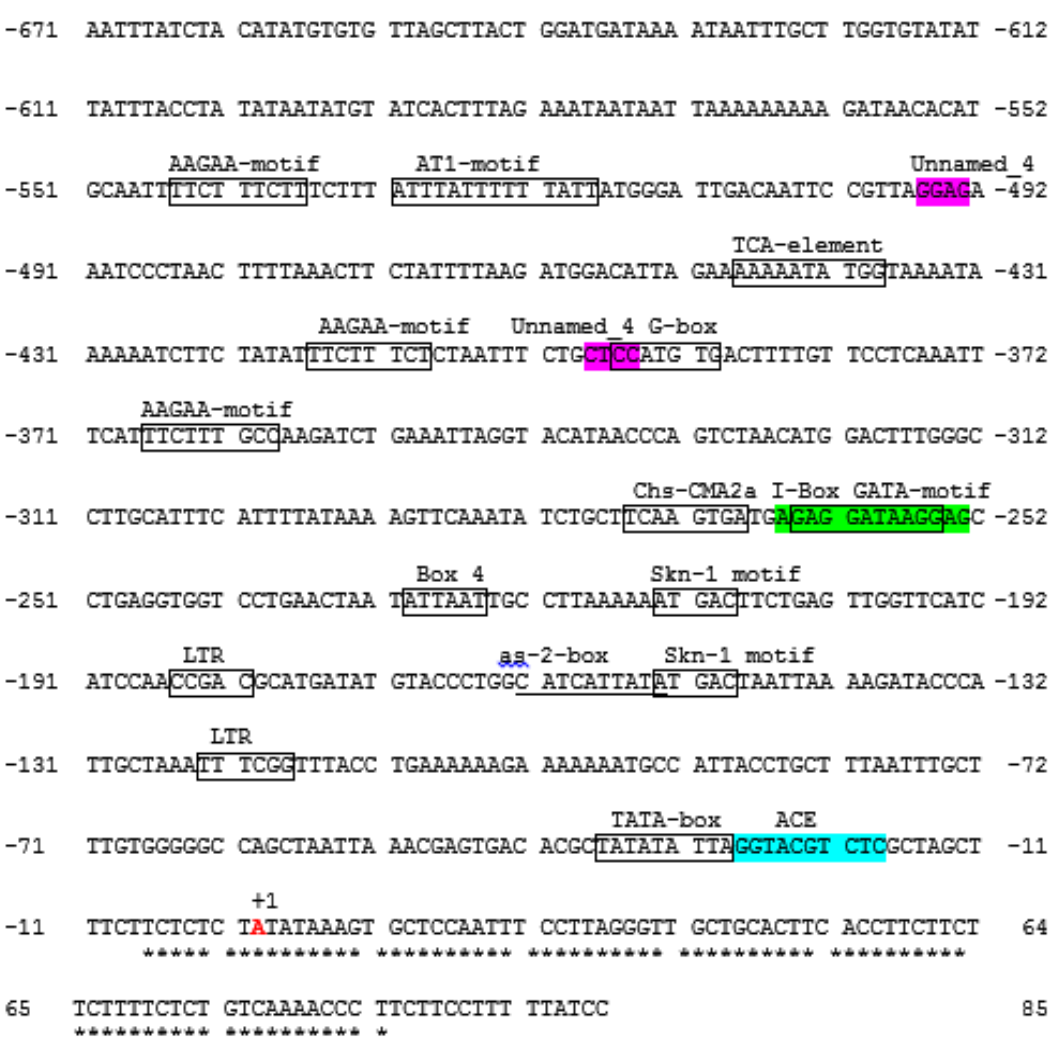



**Figure 4.10: Nucleotide sequence of the amplified genomic fragment containing the putative *FLL1* promoter.** The putative transcription start site is bold and marked with + 1. The putative regulatory elements include TATA-box, G-box, I-box, GATA-motif, ACE, LTR, TCA-elements, AT1-motif, AAGAA-motif, unnamed 4, as-2-box, chs-CMA2a and skn1-motif are either boxed, underlined or colour highlighted. The overlapping region of the 5`UTR obtained from the 5`RACE (Nurniwalis et al. 2007) are marked with '*".

**Table 4.1: Putative *cis* regulatory motifs in the *FLL1* promoter sequence.**

| No. | Name | Sequence 5` to 3` /Strand | Function description |
|---|---|---|---|
| 1 | 5`UTR Py-rich stretch | TTTCTTCTCT / (+) | *cis*-acting element conferring high transcription levels |
| 2 | AAGAA-motif | GAAAGAA / (-) | |
| 3 | ACE | GCGACGTACC / (-) | *cis*-acting element involved in light responsiveness |
| 4 | AT1-motif | AATTATTTTTTATT/(+) | part of a light responsive module |
| 5 | Box4 | ATTAAT / (+) | part of a conserved DNA module involved in light responsiveness |
| 6 | G-box | CACATGG / (-) | *cis*-acting regulatory element involved in light responsiveness |
| 7 | GATA-motif | AAGGATAAGG / (+) | part of a light responsive element |
| 8 | I-box | CCTTATCCT and gGATAAGGTG/(+/-) | part of a light responsive element |
| 9 | LTR | CCGAAA and CCGAC (-/+) | *cis*-acting element involved in low-temperature responsiveness |
| 10 | Skn-1_motif | GTCAT / (-) | *cis*-acting element required for endosperm expression |
| 11 | TCA-element | CCATCTTTTT / (-) | *cis*-acting element involved in salicylic acid responsiveness |
| 12 | Unnamed_4 | CTCC / (+/-) | |
| 13 | As-2-box | GATAatGATG / (-) | involved in shoot-specific expression and light responsiveness |
| 14 | chs-CMA2a | TCACTTGA / (-) | part of a light responsive element |

A schematic diagram of the *FLL1* promoter cloned into the pBI221 transformation vector carrying *GUS* as reporter gene that replaced the 35S promoter is shown in Figure 4.11. The activity of the *FLL1* promoter in oil palm mesocarp slices was analysed using a transient *GUS* expression analysis. The expression of the *GUS* gene driven by the *FLL1* promoter was detected in the mesocarp slices at 12 WAA but not in the leaves (control tissue) (Figure 4.12). These results support the expression analysis of *FLL1 via* Northern blot and RT-PCR (Nurniwalis et al. 2007) where the lipase class 3 gene was expressed throughout the fruit developmental stages. The mesocarp at 12 WAA was chosen as the target tissue to represent the expression of the promoter in comparison to the more mature and ripe fruit. As oil accumulation starts in the mesocarp approximately from 16 WAA onwards, it became more difficult to carry out transient expression assay especially in ripe fruits because of the high oil content. Longitudinal and crosssections of the mesocarp tissues demonstrated that the *GUS* expression was targeted to the vascular bundles.

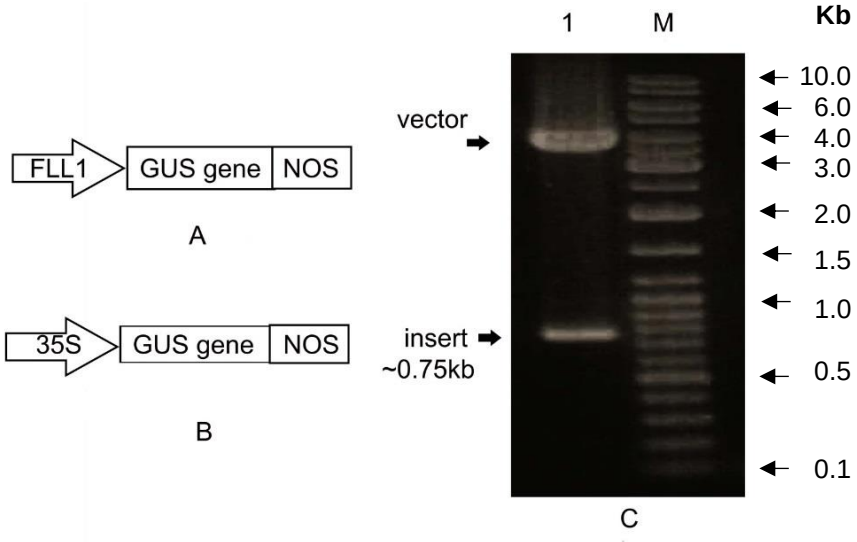



**Figure 4.11: A schematic diagram of the *FLL1GUS* plasmid (A), pBI221 plasmid (B), and *HindIII-Xba1* digestion analysis of a transformation vector pBI221 carrying the *FLL1* promoter by electrophoresis on 1% (w/v) agarose gel (C).**

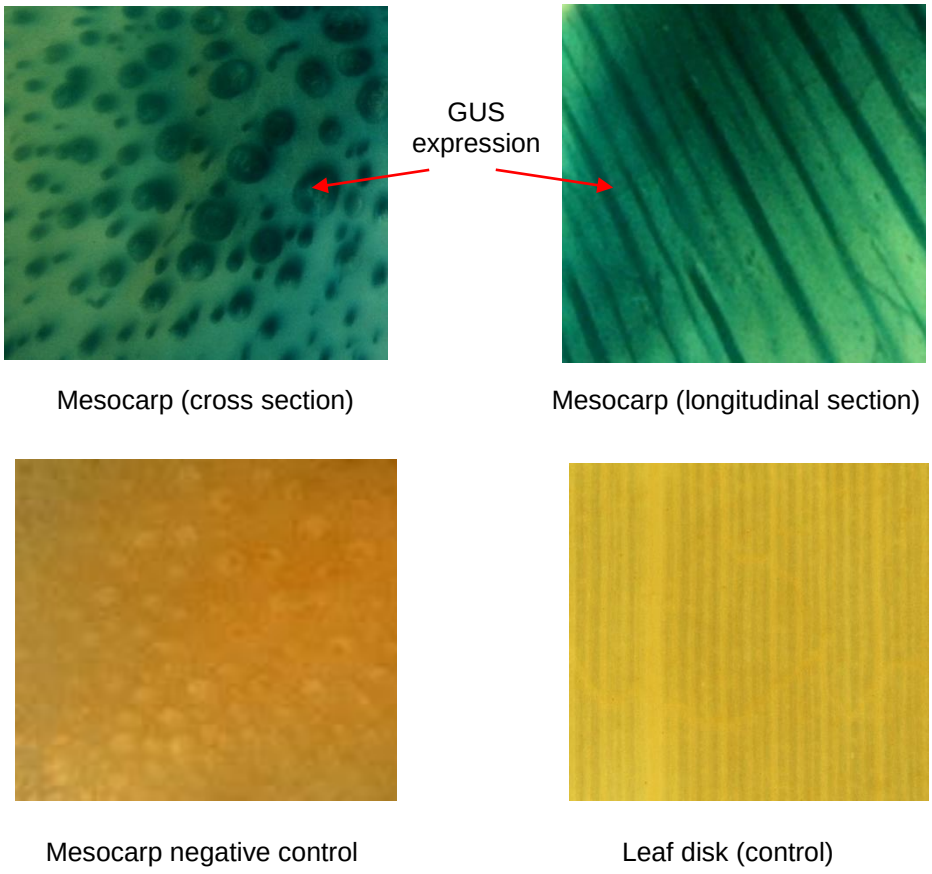



**Figure 4.12: An *FLL1pro:GUS* expression profile in the mesocarp and leaf tissues of oil palm.**

### 4.3.6 Searches for lipase class 3 gene family

Based on the result from Southern analysis, the search of other putative members of the lipase class 3 gene family was performed using the BLAST programme against the oil palm transcriptome and genome data (Singh et al. 2013a) by using Pfam ID: PF01764 as the query sequence.

A total of 307 nucleotide sequences were identified from the oil palm gene model and transcriptome databases, 80 sequences from gene model and 227 isotigs from the transcriptome database. These sequences were translated from start to stop codon to produce 89 proteins, 55 of which is present in both databases, 25 proteins in the gene model whereas 9 proteins

from the transcriptome database. Manual analyses on the protein sequences showed that 63% of the protein sequences contain all 10 amino acid lipase consensus sequences [LIV]X[LIVAFY] [LIAMVST]G[HYWV]SXG[GSTAC], that encodes the highly conserved lipase class 3 domain (Pfam ID: PF01764). Thirteen proteins contain at least 6-9 of the10 amino acid consensus sequence and another 20 proteins lack the conserved region altogether (Table 4.2). Protein motif scan performed on 33 of the proteins that do not contain the complete 10 consensus motifshowed that six of the proteins were found to contain the PF01764 motif. Thus, the remaining 27 proteins were excluded from further analyses.

**Table 4.2: General analysis on the oil palm putative members of the lipase class 3 family identified from the oil palm gene model (MP502.faa)**

| List of putative proteins subjected to analysis | No. of proteins |
|---|---|
| Total protein | 89 |
| Protein containing all 10 consensus amino acid PF01764 motif | 56 |
| Protein containing 9/10 amino acid PF01764 motif | 9 |
| Protein containing 8/10 amino acid PF01764 motif | 2 |
| Protein containing 7/10 amino acid PF01764 motif | 1 |
| Protein containing 6/10 amino acid PF01764 motif | 1 |
| Protein without any of the 10 amino acid PF01764 motif | 20 |

Sixty two putative proteins were then compared to FLL1 and the protein identities ranged from 26% to 100%. Multiple sequence analysis revealed that two of the proteins represent the truncated FLL1 where maker-p5_sc00064-snap-gene-3.39-mRNA-1_25 is 100% identical at position 242-483 and maker-p5_sc00064-snap-gene-3.39-mRNA-2_13 is 99% identical at position 193-483 of FLL1. Location for both prediction proteins in the genome scaffold also matched to that of FLL1. The next highest match to FLL1 with 83% identity is maker-p5_sc00064-snap-gene-3.37-mRNA-1_12.

Gene identification and annotation was performed based on blastp searches against the non-redundant (nr) NCBI database. The putative ID of the putative proteins based on the top hit list with the highest scores and e-value is shown in Appendix 7. Majority (79%) of the proteins were found to hit genes from date palms, indicating the closely related genome between date palm and oil palm (Singh et al. 2013a). The 62 predicted proteins were also functionally annotated based on GO classifications. Majority of the predicted proteins were assigned to the biological processes functional category. This category is further classified into several processes, predominantly by those involved in the lipid metabolic processes (Figure 4.13). Some of the predicted proteins are involved in more than one biological processes and some are assigned to both biological process and molecular functions. Twelve of the predicted proteins could not be functionally annotated based on the GO terms (Appendix 8).

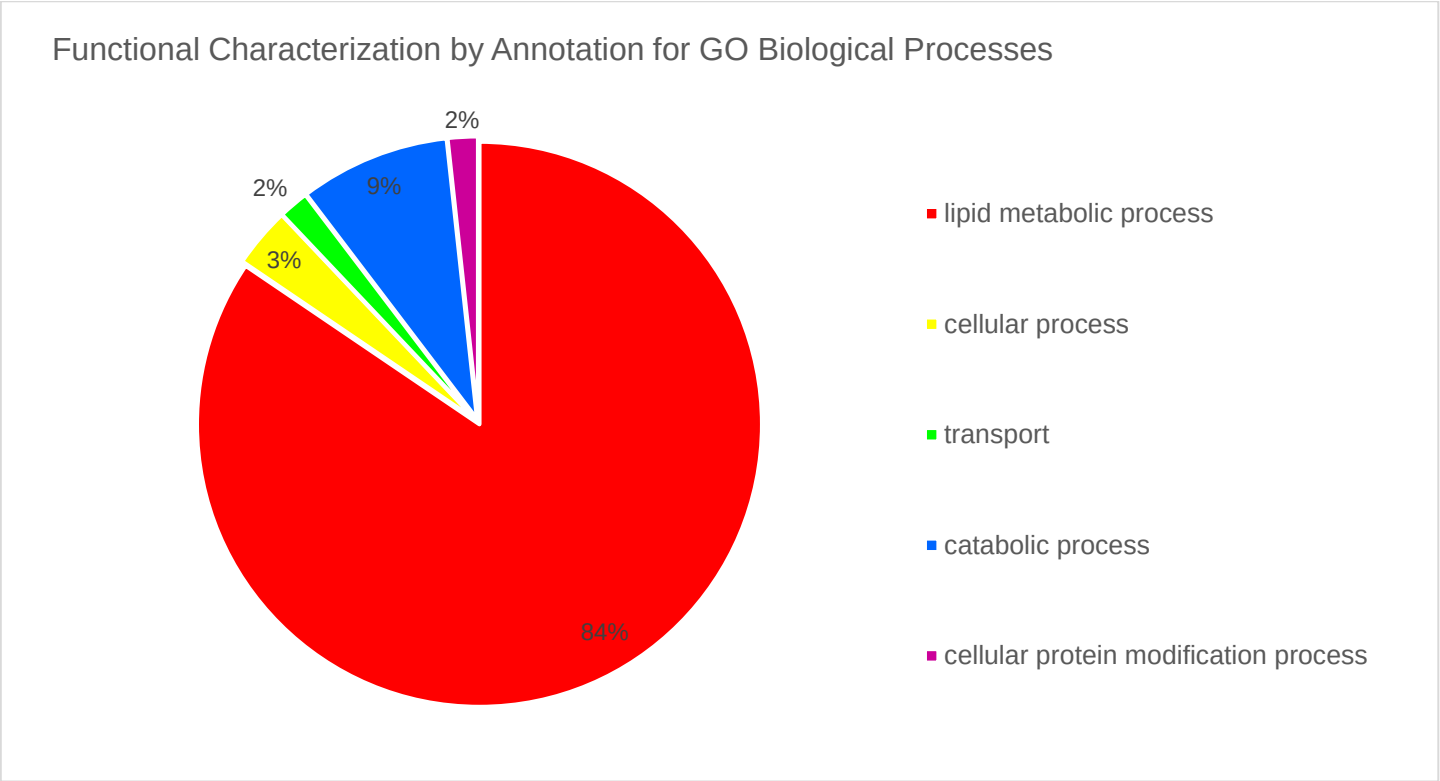



**Figure 4.13: Classification of predicted proteins of the lipase class 3 family in oil palm based on GO functional annotations for biological processes.**

The type of biological processes is represented by the different colored legend

## 4.4 Discussion

In this study, the isolation, characterisation of a full-length lipase class 3 gene (*FLL1*) and promoter and its corresponding gene family from oil palm was described, documented and discussed. FLL1 contains the lipase consensus sequence ILVTGHSLGG which includes the active serine residue ($S_{308}$) that forms the esterase box GxSxG motif. Together with the aspartic acid ($D_{368}$) and histidine ($H_{464}$) residues, they form the putative Ser-Asp-His catalytic triad that is essential for esterase and lipase activity to hydrolyze/de-esterify fatty acids from complex lipids (Li et al. 2012). This feature is not only common in lipases from plants but in animal, fungi, and bacteria as well (Patil et al. 2011). FLL1 is hydrophobic and the hydrophobic residues in the putative lid domain is likely important for the catalytic activity of FLL1.

FLL1 was 100% identical to the mesocarp transcriptome and EG5 build genome data at the nucleotide level. This result is in line to that of oil palm genome data where most of the oil palm genes involved in oil quality was derived from the paternal *pisifera* (Singh et al. 2013a). *FLL1* was also 99% identical to *EgLIP1* (Morcillo et al. 2013) (Accession No. AFV50601.1) at the nucleotide level and was 39 nucleotides longer at the 5` UTR possibly as result of a better cDNA amplification of only the full-length transcripts minus the truncated messages from the amplification process (Invitrogen Instruction Manual of GenerRacer™ Kit 2004). The sequence alignment of *EgLIP1* to the EG5 build genome data showed a 99% identity at the nucleotide level with an E value of 0.0. It was also located on the same p5-sc00064 scaffold of the p5-build and chromosome 3, which strongly suggests that both the genes are identical.

The Southern analysis showed that *FLL1* gene belong to a multigene family just like other lipase families such as the GDSL (Akoh et al. 2004) and the PR-lipase family (Szalontai 2012). With the availability of the oil palm genome sequence (Singh et al. 2013a), this has provided a rapid means to identify putative members of multigene family which was also shown in Chapter 3 with the discovery of the ethylene receptor gene family from oil palm. In this study, the preliminary search identified putative members of the lipase class 3 family in oil palm with identities ranging from 26 to 83% to *FLL1*. This is possibly due to the nature of the lipase class 3 family where polymorphisms were observed within the gene members. In Arabidopsis, thirty-eight putative lipase class 3 proteins were identified containing the conserved lipase class 3 domain (PF01764). However, there is a large divergence especially in their sequence homologies, gene structures, and expression patterns (Li et al. 2012). This result suggests that presence of other members of the lipase class 3 family in oil palm may also play specific or constitutive roles in plant growth and development especially with regards to TAGs.

The Northern and RT-PCR analyses showed that *FLL1* was highly expressed in the mesocarp tissues of oil palm (Nurniwalis et al. 2007). The *FLL1* transcripts were detectable in all the tested fruit developmental stages and the expression was stronger as the fruits reached maturity. The increased expression in the cold treated fruits is in agreement to what was reported by Sambanthamurthi et al. (1995). The activation of lipase at a low temperature in oil palm fruit bunches may provide an alternative to FFA production *via* fat splitting for use in the oleochemical industry (Sambanthamurthi and Kushairi

2002). *FLL1* is also present in germinated seedlings as well as in roots but at very much lower amounts just as reported by Morcillo et al. (2013). In addition, Oo and Stumpf (1983) have also detected an active lipase activity in oil palm germinated seedlings. Because lipase is the first enzyme involved in TAG breakdown, the low expression of *FLL1* in the germinated seedlings is logical for early seedling growth although there is no evidence to prove that lipase class 3 genes are responsible for this. Apart from the well-studied role of lipase in germinated seedlings (Murphy 1993), the role of lipase in fruits is still unclear. It has been suggested that the lipase in oil palm mesocarp serves to increase the palatability of fats (Morcillo et al. 2013). Another possible role of mesocarp lipase in trans-esterification during lipid synthesis has also been raised (Sambanthamurthi et al. 1995). At the molecular level, Morcillo et al. (2013) and Wong et al. (2015) have shown that the lipase class 3 gene may potentially be developed as a marker for oil quality where both *EgLIP1* (Morcillo et al. 2013) and *FLL1* (Wong et al. 2015) was able to distinguish between the high and low lipase genotypes. These finding can be used to select and breed low lipase genotypes which can eventually allow the ability to extend the fruit harvesting and processing without jeopardizing its quality.

At MPOB, the oil palm genetic engineering programme also targets the mesocarp as one of the tissue to introduce novel characteristics for the production of high and value-added products and improving oil quality (Cheah 1994, Parveez 2003). With the availability of tissue-specific promoters, the expression of the transgene carrying the trait of interest can be directed to the targeted tissue. With the high expression of the *FLL1* gene in the mesocarp and follows the ripening pattern of the oil palm fruits, isolation of its

corresponding promoter would be valuable for the genetic engineering programme in MPOB. This promoter has the potential to direct the expression of a desired gene carrying the trait of interest to the mesocarp for production of novel product(s) in oil palm.

*In silico* analysis using PlantCare shows that the *FLL1* promoter contained various *cis*-acting regulatory motifs which suggest a complex regulatory system to regulate the *FLL1* expression. Majority of the CAREs present in the *FLL1* promoter are associated with light responsive elements (LREs). The promoter of a fruit-specific gene, *GalUR* from strawberry whose activity is specifically targeted to the fruit, is dependent on irradiance. It contains a G-box motif that is essential for a fruit specific expression (Agius et al. 2005). In addition, an I-box motif, another LRE, is also present in the *GalUR* promoter from strawberry as well as in other fruit and in tissue-specific promoters in melon (Yamagata et al. 2002) and oil palm (Siti Nor Akmar and Zubaidah 2008). To date, the role of the LRE in the *FLL1* promoter in association with irradiance is uncertain, but the detection of multiple LREs suggests that the promoter activity can likely be enhanced by irradiance.

Two low temperature responsive element (LTRE) sequences associated with a low temperature responsiveness were identified in the *FLL1* promoter in the forward and reverse orientations. In *Arabidopsis*, the LTRE elements in the promoter region of the *COR15A* gene demonstrated an involvement in cold, drought, and ABA-regulated gene expressions (Baker et al. 1994). Transgenic potato containing the LTRE element exhibited an extensive cold inducibility in a tuber-specific manner (Liu et al. 1994). Integration of LTRE and LRE signals in *Arabidopsis* is required for full

response towards a cold acclimation (Catala et al. 2011). The LTR elements in the *FLL1* promoter suggests that it is likely to play a role in TAG hydrolysis possibly through the activation of its corresponding gene at a low temperature. This is somewhat consistent to that of the RT-PCR results as well as those reported by Sambanthamurthi et al. (1991; 1995) and Cadena et al. (2013) where lipases in oil palm are activated at a cold temperature.

The strength and activity of the promoter determines its effectiveness in controlling and regulating gene expression. The transient assay system using the *GUS* reporter gene is a common method in assessing the trait of a promoter. Testing the *FLL1* promoter activity *via* transient *GUS* expression using biolistic methods on oil palm mesocarp slices is preferred in comparison to generating oil palm transgenics. The assay saves a lot of time and is also the preferred choice to test other oil palm promoters (Zubaidah and Siti Nor Akmar 2010; Masura et al. 2011). With the high expression and specificity of the *FLL1* promoter observed in the mesocarp tissues, it has a high potential to be used as biotechnology tool to genetically engineer oil palm for production of novel oils and products. The expression of the *FLL1* promoter targeted to the vascular bundle in the mesocarp is also similar to that of the *MT3A* promoter of metallothionein gene, which directs a high expression in oil palm mesocarp tissues (Siti Nor Akmar and Zubaidah 2008; Zubaidah and Siti Nor Akmar 2010).

Phloroglucinol-stained mesocarp sections for lignin detection also showed a similar expression in the vascular bundles (Singh et al. 2013a). Lignins are normally associated with the vascular bundle in plants. They are hydrophobic in nature, which allows water transportation throughout the plant

(Xu et al. 2009). In the mesocarp tissues of oil palm (Sambanthamurthi et al. 1995) and olive (Panzanaro et al. 2010), lipase activities have been demonstrated to be associated with oil bodies and require a hydrophobic condition to function. Similar targeted expressions of the *GUS* gene and lignin to vascular bundles in mesocarp slices are unclear, but this raises the possibility that they could share a common trait or function that is still unknown.

The work presented in this chapter based on the isolation, characterisation and transient GUS assay of the *FLL1* promoter showed that the promoter can be used to direct the expression of transgene to the mesocarp tissues. Besides the MSP1 promoter (Siti Nor Akmar and Zubaidah, 2003; 2008) the *FLL1* promoter has become one of the valuable tool for the genetic engineering programme in MPOB to target transgene expression carrying the desired trait of interest to the mesocarp. The use of oil palm genome sequence data also provides the identification of the other lipase class 3 genes that may have possible novel function or can be used as markers for crop improvement as well as their corresponding promoters that may also be useful for genetic transformation to improve oil quality and production of value added products.

# CHAPTER 5
# GENERAL DISCUSSION

This study was undertaken to gain a deeper understanding of changes that occur during oil palm fruit development. The study involved physiological, biochemical and molecular analyses, with a particular emphasis on ethylene and lipase involvement in the fruit development. Clonal palms were selected for this study as they allow the establishment of uniform tree stands comprising identical copies (clones) of a limited number of highly productive oil palms (Mutert and Fairhurst, 1999). In chapter 2, changes throughout the fruit developmental stages from young until ripening were significant ($p<0.001$) in all the three (carotenoid, FAC and ethylene) measured parameters and are in agreement with previous studies (Ikemefuna and Adamson 1983; Bafor and Osagie 1986; Henderson and Osborne 1994; Tan et al. 2000; Sambanthamurthi et al. 2000; Kaur et al. 2002; Tranbarger et al. 2011). Accumulation of carotenoids especially β-carotene and the ethylene production in the fruits increased from the green stage and reached its maximum at the ripest stage. For β-carotene, the significant increase from the mature green (16 WAA) until the ripening stage also coincide with the oil synthesis period in the mesocarp tissue (Sambanthmurthi et al. 2000; Tranbarger et al. 2000). The characteristic of oil palm as a climacteric fruit was established based on the high level of ethylene production in the ripe fruits.

Changes in the major fatty acids such as palmitic, oleic, stearic and linoleic acid were observed throughout the fruit developmental process. In particular, oleic acid was found to be the highest fatty acid surpassing that of

palmitic acid. The reason for elevated levels of oleic acid is not known but the action of several biosynthetic genes in the fatty acid biosynthesis pathway has been suggested to play a role (Sambanthamurthi et al. 1990; Abrizah et al. 1999). It is also possible that environmental factors such as high rainfall, drought and longer daylight could be influencing the FAC (Hoong and Donough 1998; Prabowo and Foster 1998; Dumortier 1999). Evaluation of a large sample size would be needed to confirm reasons behind elevated levels of oleic acid. The high oleic acid trait in clonal palm would be advantageous for oil palm improvement programme.

In the present study, additional parameters such as fruit locations from the apical, central and basal bunch region as well as the fruit positioned within the outer, middle and inner spikelets located from the three regions within the bunch were included to evaluate the distribution of all three traits within the bunch. The fruit positioned within the outer, middle and inner spikelets throughout the bunch showed significant changes throughout the fruit developmental process. This was especially observed in the carotenoid and FAC. For β-carotene the highest accumulation was found in the outer fruits and the least in the inner fruits, which correspond to the colour of the fruits. The inner fruits are pale/whitish, the middle fruits have a mixture of yellow and black while the outer fruits are mostly deep-orangey in colour. The variation in the β-carotene content is likely due to the fruit exposure to sunlight based on the induced expression of carotenogenic genes during fruit ripening (Pizarro and Stange 2009). For FAC, saturated fatty acids comprising of palmitic and stearic acids were significantly higher in the outer fruits as compared to the inner fruits. As opposed to oleic and linoleic acids, which represents the

unsaturated fatty acids, the highest fatty acids were found in the inner fruits while the outer fruits have the least content. The high IV value which represents the high level of the unsaturated fatty acids in the youngest fruit development stage (4 WAA) could also indicate that the inner fruits within the bunch that contained highest amount of unsaturated fatty acids are probably less mature that the outer fruits. The inner fruits most probably are still capable for futher oil synthesis, demonstrating the unsynchronization in the development of the oil palm fruits within the bunch.

The role of ethylene in the ripening of the oil palm fruits was demonstrated by the significant increase in ethylene production. However, the ethylene production was not significantly affected by the outer, middle and inner fruits located within the apical, central and basal region of the ripe bunch. In contrast to earlier reports (Corley and Thinker 2003, 2016), ethylene production was found to be higher in the inner and apical part of the bunch suggesting a complexity of the ripening process within the bunch.

Ethylene is perceived by a family of receptors (Hua et al. 1998; Tieman and Klee 1999). Thus, to further understand the role of ethylene in the oil palm fruit developmental processes, chapter 3 focused on the identification, isolation and characterization ethylene receptor gene family, the first component of the ethylene signalling pathway. The ethylene signalling pathway is comprised of several components starting from the ethylene receptors until the transcription factors (Hall et al. 2007; Kendrick and Chang 2008). Ethylene receptors sense the ethylene signal, hence regulating the ethylene signalling pathway and activating the ethylene response (Kendrick and Chang 2008; Hua 2015). In Arabidopsis and tomato, the different roles of

each receptor in transmitting the ethylene signal has been demonstrated and interaction among the receptors are crucial in regulating the ethylene signal output (Hall et al. 2007; Gao and Schaller 2009). In fruit development, the receptors are either associated with basal ethylene production in early vegetative growth or are functional during ripening and abscission following the autocatalytic burst of ethylene in climacteric fruits (Alexander and Grierson 2002).

Seven ethylene receptor genes from the oil palm have been identified, isolated and characterized. In addition, three putative splice variants were also identified from the oil palm genome suggesting an important role in the physiological development of the oil palm with regards to the ethylene signal. Isolation and characterisation of the full-length putative ethylene receptor genes has also helped narrow down the sequence gaps between the oil palm genome scaffolds. Indirectly, this contributes towards the development and improvement of the oil palm genome sequence data. Mapping the ethylene receptors to the oil palm chromosomes is also useful to help improve the assembly of the oil palm genome. With the isolation and a comprehensive *in silico* analysis of the corresponding ethylene receptor promoters, the putative regulatory motifs have been identified. These motifs provide a means to understand the possible mechanisms that drive the expression of the ethylene receptor genes.

Generally, the expression of the seven ethylene receptor genes varies throughout the various oil palm tissues including reproductive and vegetative tissues. Most of the ethylene receptors are expressed in the tested tissues but at varying level of gene expression. During mesocarp fruit development from

young until the ripening stages, the expression of majority of the receptors were observed at the mature green until the ripening stage. The expression of some of the receptors increased as the mesocarp fruit develops (*EgETR2, EgETR3, EgERS1* and *EgERS3*) some are constitutively expressed throughout fruit development (*EgETR5 and EgERS2*) while some have decreasing expression as the fruit ripens (*EgETR4 and EgERS3*). The varying expression profiles of the oil palm receptor genes suggest their involvement in the two systems of ethylene production where some are involved in the basal ethylene production during early vegetative growth (system 1) and there are those that play a functional role during ripening (system 2).

In addition, gene expression analysis on the AZ of fruits located within the apical, central and basal region of the bunch also showed varying expression of the oil palm ethylene receptors. One of the ethylene receptors, *EgETR2* showed the highest expression at the apical region of the ripest bunch and decreased towards the central and basal region of the bunch. This expression is similar to the trend of ethylene production within the ripe bunch as described in Chapter 2 as well as earlier reports (Corley and Thinker 3003; 2016). This result may agree that fruit detachment from the bunch moves from the apical towards the basal region (Corley and Thinker 2003, 2016). However, since collection of the tissue samples were collectively based on fruit locations and not further segregated based on the position of the fruits within the spikelets, the findings on the fruit positions in Chapter 2 still needs further investigation.

The expression profile of the ethylene receptors in various oil palm tissues also indicated that each receptor may play different roles in regulating

various processes which include flower development, fruit development as well as fruit abscission. The various expression profiles of the ethylene receptors in the various oil palm tissues may also be due to the overlapping and non-overlappping roles of the receptor in modulating the ethylene signals throughout oil palm growth and development. These results are in agreement with ethylene receptors in other plants. For example, in Arabidopsis (Hall et al. 2007) and tomato (Lashbrook et al. 1998; Tieman and Klee 1999), different members of the ethylene receptor family play different roles in transmitting the ethylene signal (Hall et al. 2007) where each receptor exhibits different temporal and spatial expression patterns dependent on developmental stages and external stimuli (Lashbrook et al. 1998; Tieman and Klee 1999).

Based on the compact bunch structure, the outer fruits especially those positioned at the apical region can be easily detached from the bunch as compared to the inner fruits. However, the higher ethylene level within the inner fruits than the outer fruit (Chapter 2) in the ripe bunch may also suggest that the inner fruits may have separated from the bunch but still remained in place as it is held by the bracts of the fruit spikelets within the compact bunch. Harvesting the bunch with detached fruits causes scattering of the loose fruits and failure to collect and process these fruits leads to OER reduction and loss in profit (Mohd Solah and Rahim 2009). Bruised and detached fruits also undergo TAG hydrolysis to release FFA through the action of an endogenous lipase in the mesocarp (Henderson et al. 1991; Sambanthamurthi et al. 1991; 1995). The high FFA content affects the oil quality as it causes higher oxidation and rancidity level in the oil.

Subsequently, the effect of gene(s) that influence oil quality was investigated based on the earlier finding of a lipase class 3 expressed sequence tag that was expressed predominantly in the mesocarp tissues (Nurniwalis 2006; Nurniwalis et al. 2008; Nurniwalis et al. 2015) (Chapter 4). In this chapter, isolation and characterisation of the full-length *FLL1* gene encoding a lipase class 3 and its corresponding promoter were the focus of the study. As the mesocarp in ripe fruits contain high oil content, this makes it an abundant source for lipase activity. The expression of the *FLL1* gene was found to be induced in cold treated fruits, which is in agreement to the lipase activity in the oil palm mesocarp that is activated at cold temperature. The high and increased expression of the *FLL1* gene in the mesocarp tissues at various developmental stages especially as the ripe fruit makes it has been the forceful drive to isolate its corresponding promoter. Although *FLL1* gene is also expressed in the roots and germinated seedlings but at a much lower level, it is also appropriate to use a promoter that have great strength in the tissue of interest but have a less discrete site of expression in other tissues. Transient *FLL1* promoter assay directing the GUS gene expression into mesocarp slices showed that the *FLL1* promoter is an additional contribution to the genetic engineering programme in MPOB in the efforts to modify palm oil composition. The promoter has the potential to be used as a tool to target the expression of transgene to the mesocarp. Additionally, *FLL1* has been used as a marker to differentiate between low and high lipase genotypes (Wong et al. 2016). Similarly, a highly identical lipase class 3 gene, *Lip1* was also found to differentiate between high and low lipase genotypes (Morcillo et al. 2013). Southern analysis of the *FLL1* gene showed that it belongs to a

multigene family. Thus, with the success of identifying the putative members of the ethylene receptor gene family in the earlier work (Chapter 3), this has also prompted the use of the oil palm genome to identify members of the lipase class 3 gene family. Preliminary *in silico* analysis resulted in the identification of sixty-two predicted proteins with identities ranging from 26 to 83% to FLL1. Classification of the putative genes based on GO showed that majority of them were involved in biological processes especially those involved in the lipid metabolic processes. This finding indicates there is a possibility that more lipase class 3 genes may be involved in hydrolyzing TAGs thus affecting oil quality. These putative gene(s) can potentially be developed into validating markers for lipase selection at the nursery stage.

In summary, the present study:

1. Showed that fruit development in oil palm is a complex process involving physiological, biochemical and molecular changes within the bunch throughout the ripening process and is affected by the position of the fruits within the spikelets.
2. Revealed seven ethylene receptors and three putative splice variants in oil palm. These ethylene receptor genes and their corresponding promoters were isolated and characterised.
3. Revealed multiple roles of the ethylene receptors in regulating many processes in oil palm which includes fruit development and fruit abscission.
4. The *FLL1* gene and promoter was isolated and characterised paving the way for oil palm genetic engineering to modify the oil palm composition.

5. A number of putative lipase class 3 genes were also identified from the oil palm genome suggesting that there may be more lipases involved in the breakdown of TAGs in oil palm.

This is the first study undertaken to look at the physiological and biochemical changes during clonal palm fruit development especially within the fruit bunches. For the physiological and biochemical analyses, further work using a large sample size would be necessary to strengthen the findings of this work and to conclude on the ripening and abscission pattern within the bunch. However, it is necessary to note that since the changes within the apical, central and basal the findings on the study were not significant, this may provide a simpler means for bunch analysis. Fruits from the apical region with a mix of all the outer, middle and inner fruit may be enough to represent changes within the whole bunch.

The molecular studies on the ethylene receptor and lipase class 3 family also opens up a new possibility for genetic manipulation of oil palm to increase yield, improving oil quality as well as venturing into production of higher value products. To achieve this is it proposed that future work can also look into:

1. Fruit development and ripening related studies of non-coding RNAs

Limited work has been carried out in oil palm in comparison to other plants where modification of the ethylene synthesis and signalling pathway has been performed. This is especially targeted towards improving traits for food/fruits and horticulture applications. The expression profile of individual ethylene

receptors during ripening and abscission is interesting and could provide the foundation for a valuable programme to explore how the manipulation of responses to ethylene might provide a strategy to optimise oil palm yield. Information on the carotenoid accumulation and FAC within the bunch can be used as a guide to set up the research work. Additionally, application of current technologies such as transcriptome profiling based on non-coding RNAs from the oil palm fruits could provide more information of the involvement of these transcriptional unit with no protein coding potential in regulating the ripening and abscission processes in oil palm.

2. Modified palms for controlling oil quality

One of the ways to further study the effect of lipase on palm oil especially in high lipase genotypes is to inhibit the expression of the *FLL1* gene. Modification of the expression of the *FLL1* gene can be achieve using various approaches. One of the approach is to use the anti-sense RNA technology to down regulate/suppress the action of the lipase gene. The current genome editing technologies such as Clustered, Regularly Interspaced, Short Palindromic Repeats (CRISPR) and Transcription Activator-Like Effector Nucleases (TALEN) to knockout the *FLL1* gene can also be applied without the need to go through stringent protocols when dealing with transgenic materials.

3. Purification and overexpression of the lipase enzyme for downstream applications

*FLL1* and the identified putative lipase class 3 genes could potentially be developed for various applications. Studies on the oil palm lipase enzymatic properties and substrate specificity as well as overexpression of the gene in expression vectors can be explored further for scientific and commercial application as household cleaning agents, as well as applications for medical and health based industry.

## REFERENCES


Abdullah MZ, Guan LC, Mohamed AMD, Noor MAM (2002) Color vision system for ripeness inspection of oil palm *Elaeis guineensis*. *Journal of Food Processing and Preservation* **26**(3): 213-235

Abdullah R, Wahid, MB (2011) World Palm Oil Supply, Demand, Price and Prospects: Focus on Malaysian and Indonesian Palm Oil Industry. Kuala Lumpur, Malaysia, Malaysian Palm Oil Board Press: 1-30

Abeles, FB, Morgan, PW, Saltveit Jr, ME (1992) Ethylene in Plant Biology. Academic Press, San Diego, California, USA

Abrizah O, Othman O, Embi MN, Sambanthamurthi R (1992) Partial purification and characterization of acyl-ACP thioesterase from oil palm (*Elaeis guineensis*) Mesocarp. Presented at 18th Malaysian Biochemical Society Conference, 23-24 September 1992

Abrizah O, Sambanthamurthi R (1995) Acyl-ACP thioesterase(s) in the oil palm mesocarp (*Elaeis guineensis*). *Symposium on Biochemistry and Molecular Biology of Plant Lipids*. 1-4 June 1995, California, USA

Abrizah O, Lazarus CM, Stobart, AK (1999) Isolation of cDNA Clone Encoding an Acyl-Acyl Carrier Protein Thioesterase from the Mesocarp of Oil Palm (*Elaeis guineensis*). *Journal of Oil Palm Research* **Special Issue**: 81-87

P, Cheng H, De Grauwe L, Decat J, Schoutteten H, Moritz T, Straeten DVD, Peng J, Harberd NP (2006) Integration of plant responses to environmentally activated phytohormonal signals. *Science* **311**: 91-94

Adam H, Jouannic S, Escoute J, Duval Y, Verdeil J-L, Tregear JW (2005) Reproductive developmental complexity in the African oil palm (*Elaeis guineensis, Arecaceae*). *American Journal of Botany* **92**(11): 1836-1852

Adam H, Jouannic S, Morcillo F, Verdeil JL, Duval Y, Tregear JW (2007) Determination of Flower Structure in *Elaeis guineensis*: Do palms use the same homeotic genes as other species?. *Annals of Botany* **100**: 1-12

Adams-Phillips L, Barry C, Giovannoni J (2004) Signal transduction systems regulating fruit ripening. *Trends in Plant Science* **9**(7): 331-338

Agarwal G, Choudhary D, Singh VP, Arora A (2012) Role of ethylene receptors during senescence and ripening in horticultural crops. *Plant Signalling & Behavior* **7**(7): 827-846

Agius F, Amaya I, Botella MA, Valpuesta V (2005) Functional analysis of homologous and heterologous promoters in strawberry fruits using transient expression. *Journal of Experimental Botany* **56**(409): 37-46

Akoh CC, Lee GC, Liaw YC, Huang TH, Shaw JF (2004) GDSL family of serine esterases/lipases. *Progress in Lipid Research* **43**(6): 534-552

Alexander L, Grierson D (2002) Ethylene Biosynthesis and Action in Tomato: A Model for Climacteric Fruit Ripening. *Journal of Experimental Botany* **53**(377): 2039-2055

Alimon AR (2004) The nutritive value of palm kernel cake for animal feed. *Palm Oil Developments* **40**: 12-14

Al-Dous EK, George B, Al-Mahmoud ME, Al-Jaber MY, Wang H, Salameh YM, Al-Azwani EK, Chaluvadi S, Pontaroli AC, DeBarry J, Arondel V, Ohlrogge J, Saie IJ, Suliman-Elmeer KM, Bennetzen JL, Kruegger RR, Malek JA (2011) De novo genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). *Nature Biotechnology* **29**: 521-527

Al-Mssallem IS, Hu S, Zhang X, Lin Q, Liu W, Tan J, Yu X, Liu J, Pan L, Zhang T, Yin Y, Xin C, Wu H, Zhang G, Ba Abdullah MM, Huang D, Fang Y, Alnakhli YO, Jia S, Yin A, Alhuzimi EM, Alsaihati BA, Al-Owayyed SA, Zhao D, Zhang S, Al-Otaibi NA, Sun G, Majrashi MA, Li F, Tala, Wang J, Yun Q, Alnassar NA, Wang L, Yang M, Al-Jelaify RF, Liu K, Gao S, Chen K, Alkhaldi SR, Liu G, Zhang M, Guo H, Yu J (2013) Genome sequence of the date palm *Phoenix dactylifera L. Nature Communications* **4**: 2274

Alonso JM, Hirayama T, Roman G, Nourizadeh S, Ecker JR (1999) EIN2, a bifunctional transducer of ethylene and stress responses in Arabidopsis. *Science* **284**(5423): 2148-2152

Alonso JM, Stepanova AN (2004) The ethylene signalling pathway. *Science* **306**(5701): 1513-1515

Altschul SF, Madden TM, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipmann DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res*earch **25**(17): 3389-3402

Álvarez D, Voß B, Maass D, Wüst F, Schaub P, Beyer P, Welsch R (2016) Carotenogenesis is regulated by 5' UTR-mediated translation of phytoene synthase splice variants. *Plant Physiology* **172**(4): 2314-2326

Arnaud F, Rabechault H (1972) First observations on the cytohistochemical characters of resistance to 'sudden dieback' in the oil palm. *Oleagineux* **27**(11): 525-529

Arora A (2005) Ethylene receptors and molecular mechanism of ethylene sensitivity in plants. *Current Science* **89**(8): 1348-1361

Arora A, Watanabe S, Ma B, Takada K, Ezura H (2006) A novel ethylene receptor homolog gene isolated from ethylene-insensitive flowers of gladiolus (*Gladiolus grandiflora Hort). Biochemical and Biophysical Research Communications* **351**(3): 739-744

Azis AA (1984) The Biochemical aspects of ripeness standards. Proceedings of the (1984) *Symposium on Impact of the pollinating weevil on the Malaysian Oil Palm Industry*, 165–176

Azis AA (1990) A simple floatation technique to gauge ripeness of oil palm fruits and their maximum oil content. *Proceeding of the International Palm Oil Development Conference (PORIM'90)*, Kuala Lumpur, Malaysia: PORIM, 87–91

Aziz, A, Rosnah MS, Mohamadiah B, Wan Zailan WO (1986) Relationship of the formation of storage oil with fatty acid composition in developing oil palm fruit. *Proceeding of the Malaysian Biochemistry Society Conference*, 12: 147–51

Azis AA, Rosnah MS, Mohamadiah B, Wan Zailan, WO (1990) Morphological changes of the cellular component of the developing palm fruit (*Tenera: Elaeis guineensis*). *PORIM Bulletin* **21**: 30-34

Azmir J (2012) Removal of free fatty acid from crude palm oil using organic solvent nanofiltration (OSN). Johor, Malaysia: Universiti Teknologi Malaysia, Msc. Thesis

Azzolini M, Jacomino AP, Bron IU, Kluge RA, Schiavinto MA (2005) Ripening of 'Pedro Sato' guava: study on its climacteric or non-climacteric Nature*. Brazilian Journal of Plant Physiology* **17**(3): 299-306

Bafor ME, Osagie AU (1986) Changes in lipid class and fatty acid composition during maturation of mesocarp of oil palm (*Elaeis guineensis*) variety *dura*. *Journal of the Science of Food and Agriculture* **37**(9): 825-832

Bafor ME, Osagie, AU (1988) Changes in non-polar lipid composition of developing oil palm fruit (*Elaeis guineensis*) mesocarp. *Journal of the Science of Food and Agriculture* **45**(4): 325-331

Bahariah B, Parveez GKA (2007) Fatty acid analyses of transformed oil palm cultures and plantlets. *PIPOC Agriculture, Biotechnology and Sustainability Conference, Kuala Lumpur Convention Centre (KLCC)*, Kuala Lumpur, Malaysia, MPOB: 1071-1076

Baker SS, Wilhelm KS, Thomashow MF (1994) The 5′-region of *Arabidopsis thaliana* cor15a has cis-acting elements that confer cold-, drought- and ABA-regulated gene expression. *Plant Molecular Biology* **24**(5): 701-713

Barreto GPM, Fabi JP, De Rosso VV, Cordenunsi BR, Lajolo FM, do Nascimento JRO, Mercadante AZ (2011) Influence of ethylene on carotenoid biosynthesis during papaya postharvesting ripening. *Journal of Food Composition and Analysis* **24**: 620-624

Barry CS, Llop-Tous MI, Grierson D (2000) The regulation of 1-aminocyclopropane-1-carboxylic acid synthase gene expression during the transition from system-1 to system-2 ethylene synthesis in tomato. *Plant Physiology* **123**(3): 979-986

Benade AJ (2003) A place for palm fruit oil to eliminate vitamin A deficiency. *Asia Pacific Journal of Clinical Nutrition* **12(**3): 369-372

Bille Ngalle H, Bell JM, Franck G, Ngando-Ebongue N, Nyobe L, Ngangnou FC, Ntsomboh GN (2013) Morphogenesis of oil palm fruit (*Elaeis guineensis Jacq*) in mesocarp and endocarp development. *Journal of Life Sciences* **7**(2): 153-158

Billotte N, Marseillac N, Risterucci AM, Adon B, Brottier P, Baurens FC, Singh R, Herrán A, Asmady H, Billot C, Amblard P, Durand-Gasselin T, Courtois B, Asmono D, Cheah SC, Rohde W, Ritter E, Charrier A (2005) Microsatellite-based high density linkage map in oil palm (*Elaeis guineensis Jacq). Theoretical and Applied Genetics* **110**(4): 754-765

Binder BM, Chang C, Schaller GE (2012) Perception of ethylene by plants: ethylene receptors. *In* McManus MT (ed.) *Annual Plant Reviews Volume 44: The Plant Hormone Ethylene.* Oxford, UK: Wiley-Blackwell, 117-145

Birch RG (1997) Plant transformation: problems and strategies for practical application. *Annual Review of Plant Physiology and Plant Molecular Biology* **48**: 297-326

Black DL (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry* **72**(1): 291-336

Blankenship SM, Dole JM (2003) 1-methylcyclopropene: a review. *Postharvest Biology and Technology* **28**(1): 1-25

Böttcher C, Keyzers RA, Bos PK, Davies C (2010) Sequestration of auxin by the indole-3-acetic acid-amido synthetase GH3-1 in grape berry (*Vitis vinifera L*) and the

proposed role of auxin conjugation during ripening. *Journal of Experimental Botany* **61**(13): 3615-3625

Böttcher C, Harvey K, Forde CG, Boss PK, Davies C (2010) Auxin treatment of pre-veraison grape (*Vitis vinifera L*.) berries both delays ripening and increases the synchronicity of sugar accumulation. *Australian Journal of Grape and Wine Research* **17**: 1-8

Bouzayen M, Alain L, Pavendra N, Pech J-C (2010) Mechanism of fruit ripening. *In* Pua EC, Davey M (eds.) Plant Developmental Biology - Biotechnological Perspectives Vol 1. Heidelberg,Germany: Springer-Verlag, 319-339

Brady CJ (1987) Fruit Ripening. *Annual Review Plant Physiology* 38(1): 155-178

Bramley PM (2002) Regulation of carotenoid formation during tomato fruit ripening and development. *Journal of Experimental Botany* **53**: 2107-2113

Breathnach R, Chambon P (1981) Organization and Expression of Eucaryotic Split Genes Coding for Proteins. *Annual Review of Biochemistry* 50 (1): 349-383

Breure K (2003) The search for yield in oil palm: basic principles. *In* R Fairhurst R, Hardter T (eds.) *The Oil Palm-Management for Large and Sustainable Yields*. Potash & Phosphate Institute/Potash Institute of Canada and International Potash Institute, Singapore: 59-98

Cadena T, Prada F, Perea A, Romero HM (2013) Lipase Activity, Mesocarp Oil Content, Iodine Value in Oil Palm Fruits of *Elaeis guineensis, Elaeis oleifera*, the Interspecific Hybrid O×G (*E. Oleifera* × *E. Guineensis). Journal of the Science of Food and Agriculture* 93(3): 674-680

Catala R, Medina J, Salinas J (2011) Integration of low temperature and light signalling during cold acclimation response in Arabidopsis. *Proceedings of the National Academy of Sciences* **108**(39): 16475-16480

Chang C, Kwok SF, Bleecker AB, Meyerowitz EM (1993) Arabidopsis ethylene-response gene *ETR1*: Similarity of product to two-component regulators. *Science* **262**(5133): 539-544

Cheah SC (1994) Genetic engineering of oil crops for oil quality. *Palm Oil Developments* **20**: 28-34

Cheah SC, Sambanthamurthi R, Siti Nor Akmar A, Abrizah O, Mohamad Arif MA, Umi Salamah R and Parveez GKA (1995) Towards genetic engineering of oil palm. *In* Kader JC, Mazliak P (eds.) *Plant Lipid Metabolism*. Netherlands, Kluwer Academic Publishers, 570-571

Chen JF, Gallie DR (2010) Analysis of the functional conservation of ethylene receptors between maize and Arabidopsis. *Plant Molecular Biology* **74**(4): 405-421

Chen M, Manley JL (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nature Reviews Molecular Cell Biology* **10**(11): 741-754

Chen YF, Etheridge N, Schaller GE (2005) Ethylene signal transduction. *Annals of Botany* **95**(6): 901-915

Chen YF, Randlett MD, Findell JL, Schaller GE (2002) Localization of the ethylene receptor ETR1 to the endoplasmic reticulum of Arabidopsis. *Journal of Biological Chemistry* **277**(22): 19861-19866

Chenchik A, Diachenko L, Moqadam F, Tarabykin V, Lukyanov S, Siebert PD (1996) Full-length cDNA cloning and determination of mRNA 5' and 3' ends by amplification of adaptor-ligated cDNA. *BioTechniques* **21**(3): 526-534

Chervin C, Deluc L (2010) Ethylene signalling receptors and transcription factors over the grape berry development: gene expression profiling. *Vitis* **49**(3): 129-136

Chervin C, Tira-Umphon A, Chatelet P, Jauneau A, Boss PK, Tesniere C (2009) ethylene and other stimuli affect expression of the UDP glucose-flavonoid 3-O-glucosyltransferase in a non-climacteric fruit. *Vitis* **48**(1): 11-16

Chillet M, Bellaire LDLD, Hubert O, Mbeguie-A-Mbguie D (2008) Measurement of ethylene production during banana ripening. *Fruits*. **63**(4): 253-254

Chin M (2011) Biofuels in Malaysia: an analysis of the legal and institutional framework. Working Paper 64. CIFOR, Bogor, Indonesia

Chong CL (2011) Storage, handling and new products. *In* Wahid MBA, Choo, YM, Chan KW (eds.) Further Advances in Oil Palm Research (2000-2010), Vol 2. Bangi, Malaysia: Malaysian Palm Oil Board, 645-672

Choo YM, Ma AN, Yap SC (1997) Carotenes, vitamin E and sterols in oils from *Elaeis guineensis, Elaeis oleifera* and their hybrids. *Palm Oil Developments* **27**:1-9

Choo YM, Ng MH (2011) Phytonutrients from palm oil. *In* Wahid MBA, Choo, YM, Chan KW (eds.) Further Advances in Oil Palm Research (2000-2010), Vol 2. Bangi, Malaysia: Malaysian Palm Oil Board, 721-763

Chowdhury MKU, Parveez GKA, Saleh NM (1997) Evaluation of five promoters for use in transformation of oil palm (*Elaeis guineensis* Jacq.) *Plant Cell Report* **16**: 277-281

Chung BYW, Simons C, Firth AE, Brow CM, Hellens RP (2006) Effect of 5'UTR introns on gene expression in *Arabidopsis thaliana. BMC Genomics* **7**: 120

Ciardi J, Klee H (2001) Regulation of ethylene-mediated responses at the level of the receptor. *Annals of Botany* **88**(5): 813-822

Christie WW (1989) Gas chromatography and lipids: a practical guide. Oily Press Limited.

Corley RHV (2009) How much palm oil do we need?. *Environmental Science & Policy* **12**(2): 134-139

Corley RHV, Law IH (2001) Ripening, harvesting and oil extraction. *Planter* **77**: 507-524

Corley RHV, Tinker PB (2003) *The Oil Palm: Fourth Edition*. Oxford, UK: Blackwell Science Ltd

Corley RHV, Tinker PB (2016) *The Oil Palm: Fifth Edition.* Oxford, UK: Wiley Blackwell Ltd

Crane JC (1964) Growth substances in fruit setting and development. *Annual Review Plant Physiology* **15**(1): 303-326

Crombie WM, Hardman EE (1958) Fat metabolism in the west African oil palm (*Elaeis guineensis*): Part III. Fatty acid formation in the maturing exocarp. *Journal of Experimental Botany* **9**(2): 247-253

Czarny JC, Grichko VP, Glick BR (2006) Genetic modulation of ethylene biosynthesis and signalling in plants. *Biotechnology Advances* **24**(4): 410-419

Davey JE, Van Staden J (978) Endogenous cytokinins in the fruits of ripening and non-ripening tomatoes. *Plant Science Letters* **11**(3-4): 359-364

D'Hont A, Denoeud F, Aury JM, Baurens FC, Carreel F, Garsmeur O, Noel B, Bocs S, Droc G, Rouard M, Da Silva C et al. (2012) The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* **488**: 213-217

Da Silva MG, Oliveira JG, Vitoria AP, Corrêa SF, Pereira MG, Campostrini E, Santos EO, Cavalli A, Vargas H (2005) Correlation between ethylene emission and skin colour changes during papaya (*Carica papaya L*.) fruit ripening. *Journal de Physique IV (Proceedings)* **125**(1): 877-879

Dai ZW, Ollat N, Gomes E, Decroocq S, Tandonnet J-P, Bordenave L, Pieri P, Hilbert G, Kappel C, Leeuwen CV, Vivin P, Delrot S (2011) Ecophysiological, genetic, and molecular causes of variation in grape berry weight and composition: a review. *American Journal of Enology and Viticulture* **62**: 413-425

Davuluri GR, van Tuinen A, Fraser PD, Manfredonia A, Newman R, Burgess D, Brummell DA, King SR, Palys J, Uhlig J, Bramley PM, Pennings HM, Bowler C (2005) Fruit specific RNAi-mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes. *Nature Biotechnology* **23**: 890-895

Dessassis A (1957) Palm oil acidification. Oléagineux **12**: 525-534

Di Vaio C, Nocerino S, Paduano A, Sacchi R (2013) Influence of some environmental factors on drupe maturation and olive oil composition. *Journal of the Science of Food and Agriculture* **93**(5):1134-1139

Diosady LL (2005) Chlorophyll removal from edible oils. *International Journal of Applied Science and Engineering* **3**(2): 81-88

Doyle JJ, Doyle JL (1990) Isolation of DNA from small amounts of plant tissues. *Focus* **12**(1): 13-15

Dumortier F (1999) Performance and variability of oil palm DxP material in Papua New Guinea. *In* Rajanaidu N, Jalani BS (eds.) *Seminar on Performance of DxP Material, Clones and Interspecific Hybrids* (Baranquilla, Colombia, June 1995). Kuala Lumpur, Malaysia: Palm Oil Research Institute of Malaysia, 58-75

Economic Transformation Programme (2010) The NKEAs - Palm Oil & Rubber. http://etppemandugovmy/Palm_Oil-@-Palm_Oil_-%e2%97%98-_Rubberaspx. Accessed 4th January 2016

Economic Transformation Programme (2013) NKEA Palm Oil and Rubber. *ETP Annual Report*: 98–117

Eddy SR (2011) Accelerated Profile HMM Searches. *PLoS Computational Biology* 7(10): e1002195

El-Sharkawy I, Jones B, Li ZG, Lelièvre JM, Pech JC, Latché A (2003) Isolation and characterization of four ethylene perception elements and their expression during ripening in pears (*Pyrus communis L*) with/without cold requirement. *Journal of Experimental Botany* **54**(387): 1615-1625

EPU (2014) Malaysian Quality of Life 2004 Economic Planning Unit, Prime Minister's Department, Malaysia http://www.epu.gov.my. Accessed 5th November 2015

Fadila AM, Marhalil M, Mohd Din A (2012) Performance of four MPOB oil palm clones at two soil environment. *23rd Malaysian Society of Plant Physiology Conference: Plant Physiology in Addressing Green Economy*. 18-20 December 2012, Langkawi, Kedah, Malaysia.

FAQs-MPOB: What is the fatty acid composition of palm oil and its associated products? (2016) http://www.mpob.gov.my/en/faqs. Accessed 1st January 2016

FAQs-MPOB: How stable are unsaturated fatty acids? (2016) http://www.mpob.gov.my/en/faqs. Accessed 13th January 2016

Fatin SA, Rosnah S and Yunus R (2014) Effect of chopping oil palm fruits spikelets on the free fatty acids content release rate and its mechanical properties. *International Journal of Research in Engineering and Technology* **3**(1): 511-516

FELDA (2014) Federal Land Development Authority http://felda.net.my. Accessed 15th October 2015

Fekarurhobo GK, Obomanu FG (2009) Effects of short-term exposure to sunlight on the quality of some edible vegetable oils. *Research Journal of Applied Sciences* **4**: 152-156

Fitzherbert EB, Struebig MJ, Morel A, Danielsen F, Brühl CA, Donald PF, Phalan B (2008) How will oil palm expansion affect biodiversity?. *Trends in Ecology and Evolution* **23**(10): 538-545

Gallie DR (2015) Ethylene receptors in plants - why so much complexity? *F1000Prime Reports* **7**: 1-12

Gallie DR, Young TE (2004) The ethylene biosynthetic and perception machinery is differentially expressed during endosperm and embryo development in maize. *Molecular Genetics and Genomics* **271**(3): 267-281

Galpaz N, Wang Q, Menda N, Zamir D, Hirschberg J (2008) Abscisic acid deficiency in the tomato mutant high-pigment 3 leading to increased plastid number and higher fruit lycopene content. *Plant Journal* **53**(5): 717-730

Gao Z, Schaller GE (2009) The role of receptor interactions in regulating ethylene signal transduction. *Plant Signal Behavior* **4**(12): 1152-1153

Gapper NE, Giovannoni JJ, Watkins CB (2014) Understanding development and ripening of fruit crops in an 'Omics' era, *Horticulture Research* **1**: 14034

Gillaspy G, Ben-David H, Gruissem W, Darwin C (1993) Fruits : a developmental perspective. *The Plant Cell* **5**: 1439-1451

Giovannoni JJ (2001) Molecular Biology of fruit maturation and ripening. *Annual Review Plant Physiology Plant Molecular Biology* **52**: 725-749

Giovannoni JJ (2004) Genetic regulation of fruit development and ripening. *The Plant Cell* **16**: S170–S180

Giribaldi, M, Giuffrida MG (2010) Heard it through the grapevine: proteomic perspective on grape and wine. *Journal of Proteomics* **73**(9): 1647-1655

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D et al. (2002) A draft sequence of the rice genome (*Oryza sativa L ssp japonica*). *Science* **296**(5565): 92-100

Gouthu S, O'Neil ST, Di Y, Ansarolia M, Megraw M, Deluc LG (2014) A comparative study of ripening among berries of the grape cluster reveals an altered transcriptional

programme and enhanced ripening rate in delayed berries. *Journal of Experimental Botany* **65**(20): 5889-5902

Graveley BR (2001) Alternative splicing: increasing diversity in the proteomic world. *Trends in Genetics* **17**(2): 100-107

Gray J, Picton S, Shabbeer J, Schuch W, Grierson D (1992) Molecular biology of fruit ripening and its manipulation with antisense genes. *Plant Molecular Biology* **19**: 69-87

Grefen C, Städele K, Růžička K, Obrdlik P, Harter K, Horák J (2008) Subcellular localization and *in vivo* interactions of the *Arabidopsis thaliana* ethylene receptor family members. *Molecular Plant* **1**(2): 308-320

Grierson D, Purton M, Knapp J, Bathgate B (1987) Tomato ripening mutants. *In* Grierson D, Thomas H (eds.) *Developmental Mutants in Higher Plants.* London, UK: Cambridge Univ Press, 73-94

Gruszecki WI, Strzałka K (2005) Carotenoids as modulators of lipid membrane physical properties. *Biochimica et Biophysica Acta* **1740**(2): 108-115

Habib SH, Ooi SE, Nova´k O, Tarkowska´ D, Rolcˇı´k J, Dolezˇal K, Namasivayam P, Syed-Alwee SSR, Ho CL (2012) Comparative mineral and hormonal analyses of wild type and TLS somaclonal variant derived from oil palm ( *Elaeis guineensis Jacq var. tenera*) Tissue Culture. *Tissue Culture Plant Growth Regulation* **68**(2): 313-317

Hall AE, Chen QG, Findell JL, Schaller GE, Bleecker AB (1999) The relationship between ethylene binding and dominant insensitivity conferred by mutant forms of the ETR1 ethylene receptor. *Plant Physiology* **121**(1): 291-300

Hall BP, Shakeel SN, Schaller GE (2007) Ethylene receptors: ethylene perception and signal transduction. *Journal of Plant Growth Regulation* **26**(2): 118-130

Hardon JJ, Corley RHV, Lee CH (1987) Breeding and selecting the oil palm. *In* Abbot AJ, Atkin RK (eds.) *Improving Vegetatively Propagated Crops*. London, UK: Academic Press, 63-81

Hartley CWS (1988) *The Oil Palm 3rd Edition.* London, UK: Longman Scientific and Technical

Harun NH, Norhisam M, Roslina MS, Ishak A, Desa A, Wakiwaka H, Tashiro K (2013) Investigations on a novel inductive concept frequency technique for the grading of oil palm fresh fruit bunches. *Sensors* **13**: 2254-2266

Hazimah, AH, Zainab I, Rosnah I, Loh SK (2011) From palm oil to non-food products: green chemistry. *In* Wahid MBA, Choo, YM, Chan KW (eds.) *Further Advances in Oil Palm Research (2000-2010), Vol 2.* Bangi, Malaysia: Malaysian Palm Oil Board, 791-895

Henderson J, Osborne DJ (1994) Inter-tissue Signalling during the two-phase abscission in oil palm fruit. *Journal of Experimental Botany* **45**(276): 943-951

Henderson J, Osborne DJ (1991) Lipase activity in ripening and mature fruit of the oil palm stability *in vivo* and *in vitro*. *Phytochemistry* **30**(4): 1073-1078

Henderson J, Osborne DJ (1999) Ethylene as the initiator of the inter-tissue signalling and gene expression cascades in ripening and abscission of oil palm fruit. *In* Kanellis AK (ed.) *Biology and Biotechnology of the Plant Hormone Ethylene II*. Doedrecht, Netherlands: Kluwer Academic Publishers, 129-136

Hirschberg J (2001) Carotenoid Biosynthesis in Flowering Plants. *Current Opinion in Plant Biology* **4**: 210-218

Hoong HK, Donough CR (1998) Recent trends in oil extraction rate (OER) and kernel extraction rate (KER) in Sabah. *Planter* **74**: 181-202

Honda A, Valogne Y, Nader MB, Bréchot C, Faivre J (2012) An intron-retaining splice variant of human cyclin A2, expressed in adult differentiated tissues, induces a G1/S cell cycle arrest *in vitro. PLoS ONE* **7**(6): e39249

Hong X, Scofield DG, Lynch M (2006) Intron size, abundance and distribution within untranslated regions of genes*. Molecular Biology and Evolution* **23**(12): 2392-2404

Hua J (2015) Isolation of components involved in ethylene signalling. *In* Wen C-K (ed.) *Ethylene in Plants*. Dordrecht, Netherlands: Springer Science, 27-44

Hua J, Chang C, Sun Q, Meyerowitz EM (1995) Ethylene insensitivity conferred by Arabidopsis ERS gene. *Science* **269**(5231): 1712-1714

Hua J, Meyerowitz EM (1998) Ethylene responses are negatively regulated by a receptor gene family in *Arabidopsis thaliana. Cell* **94**: 261-271

Hua J, Sakai H, Nourizadeh S, Chen QG, Bleecker AB, Ecker JR, Meyerowitz EM (1998) EIN4 and ERS2 are members of the putative ethylene receptor gene family in Arabidopsis. *The Plant Cell* **10**(8): 1321-1332

Huber DJ (2008) Suppression of ethylene responses through application of 1-methylcyclopropene: a powerful tool for elucidating ripening and senescence mechanisms in climacteric and nonclimacteric fruits and vegetables*. Horticulture Science* **43**(1): 106-111

Hunt AG (1994) Messenger RNA 3' end formation in plants*. Annual Review Plant Physiology and Plant Molecular Biology* **45**(1): 47-60

Huntley RP, Jones LH, Hanke DE (2002) Cytokinins and gibberellins in sap exudate of the oil palm. *Phytochemistry* **60**(2): 117-127

Idris O (2004) Penuaian Buah Tandan Segar. *Dalam:* Perusahaan Sawit Di Malaysia - Satu Panduan (Edisi Melinium). Lembaga Minyak Sawit Malaysia: Bangi, Malaysia

Ikemefuna J, Adamson I (1984) Chlorophyll and carotenoid changes in ripening palm fruit, *Elaeis guineensis. Phytochemistry* **23**(7): 1413-1415 D

Ish-Shalom M, Dahan Y, Maayan I, Irihimovitch V (2011) Cloning and molecular Characterization of an ethylene receptor gene, *MiERS1*, expressed during mango fruitlet abscission and fruit ripening. *Plant Physiology and Biochemistry* **49**(8): 931-936

Ismail, WIW, Razali MH, Ramli, AR, Sulaiman, MN, Harun, MHB (2009) Development of imaging application for oil palm fruit maturity prediction. *Engineering E-Transaction* **4**(2): 56-63

Jalani S, Rajanaidu N (2000) Improvement in oil palm: yield, composition and minor components. *Lipid Techology* **12**: 5-8

Jin J, Lee M, Bai B, Sun Y, Qu J, Rahmadsyah, Alfiko Y, Lim CH, Suwanto A, Sugiharti M, Wong L, Ye J, Chua N-H Yue GH (2016) Draft genome sequence of an elite *Dura* palm and whole-genome patterns of DNA variation in oil palm. *DNA Research* **0**(0): 1-7

Joshi CP (1987a) Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. *Nucleic Acids Research* **15**(23): 9627-9640

Joshi CP (1987b) An inspection of the domain between putative TATA box and translation start site in 79 plant genes. *Nucleic Acids Research* **15**(16): 6643-6653

Ju C, Chang C (2012) Advances in Ethylene Signalling: Protein complexes at the endoplasmic reticulum membrane. *AoB Plants* **2012**: pls031

Kaida K, Zulkifly A (1992) A Microstrip Sensor for determination of harvesting time for oil palm fruits (*Tenera: Elaeis guineensis). Journal of Microwave Power and Electromagnetic Energy* **27**(1): 3-10

Kalyna M, Simpson CG, Syed NH, Lewandowska D, Marquez Y, Kusenda B, Marshall J, Fuller J, Cardle L, McNicol J, Dinh HQ, Barta A, Brown JW (2012) Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. *Nucleic Acids Research* **40**(6): 2454-2469

Kamffer Z (2009) Carotenoid and chlorophyll content of *Vitis vinifera cv. Merlot* grapes during ripening with reference to variability in grapevine water status and vigour. Stellenbosch, South Africa: Stellenbosch University, MSc Thesis

Kardash E, Turýan YI (2005) Acid value determination in vegetable oils by indirect titration in aqueous-alcohol media. *Croatica Chemica Acta* **78**(1): 99-103

Kaur JS, Sambanthamurthi R (2008) Carotenoid profiles and preliminary investigation on carotenoid biosynthesis in the oil palm (*Elaeis guineensis Jacq*.) mesocarp. *Journal of Oil Palm Research* **20**: 108-117

Kendrick MD, Chang C (2008) Ethylene signalling: new levels of complexity and regulation. *Current Opinion in Plant Biology* **11**(5): 479-485

Keshvadi A, Endan J, Harun H, Ahmad D, Saleena F (2012) The reflection of moisture content on palm oil development during the ripening process of fresh fruits. *Journal of Food, Agriculture and Environment***10**(1): 203-209

Kevany BM, Taylor MG, Klee HJ (2008) Fruit-specific suppression of the ethylene receptor *LeETR4* results in early-ripening tomato fruit. *Plant Biotechnology Journal* **6**(3): 295-300

Kevany BM, Tieman DM, Taylor MG, Cin VD, Klee HJ (2007) Ethylene receptor degradation controls the timing of ripening in tomato fruit. *Plant Journal* **51**(3): 458-467

Klee HJ (2004) Ethylene signal transduction moving beyond Arabidopsis. *Plant Physiology* **135**(2): 660-667

Klee HJ, Giovannoni JJ (2011) Genetics and control of tomato fruit ripening and quality attributes. *Annual Review Genetics* **45**(1): 41-59

Kumar R, Khurana A, Sharma AK (2014) Role of plant hormones and their interplay in development and ripening of fleshy fruits. *Journal of Experimental Botany* **65**(16): 4561-4575

Kushairi A (2009) Role of oil palm breeding in wealth creation and quality of life. Keynote paper presented at the 8th Malaysian Genetics Congress Role of Genetics in Wealth Creation and Quality of Life: 4–6 August 2009 Awana Genting, Pahang, Malaysia

Kushairi A, Rajanaidu N (2000) Breeding population, seed production and nursery management. *In* Jalani BS, Chan KW, Yusof B (eds.) *Advances in Oil Palm Research (Vol 1)*. Bangi, Malaysia: Malaysian Palm Oil Board, 39-96

Kushairi A, Tarmizi AH, Samsuri I, Ong-Abdullah M, Samsul Kamal R, Ooi SE, Rajanaidu N (2010) Production, performance and advances in oil palm tissue culture. Paper presented at the International Seminar on Advances in Oil Palm Tissue Culture, Indonesia.

Lanahan MB, Yen HC, Giovannoni JJ, Klee HJ (1994) The never ripe mutation blocks ethylene perception in tomato. *The Plant Cell* **6**: 521-530

Lashbrook CC, Tieman DM, Klee HJ (1998) Differential regulation of the tomato ETR gene family throughout plant development. *Plant Journal* **15**(2): 243–-252

Latif A (2000) The biology of the genus Elaeis. *In* Jalani BS, Chan KW, Yusof B (eds.) *Advances in Oil Palm Research (Vol 1)*. Bangi, Malaysia: Malaysian Palm Oil Board, 19-38

Lehrian DW, Keeney PG, Butler DR (1980) Triglyceride characteristics of cocoa butter from cacao fruit matured in a microclimate of elevated temperature. *Journal of American Oil Chemists' Society* **57**(2): 66-69

Lelievre JM, Latche A, Jones B, Bouzayen M, Pech JC (1997) Ethylene and fruit ripening. *Physiologia Plantarum*. **101**(4): 727-739

Leng P, Yuan B, Guo Y, Chen P (2014) The role of abscisic acid in fruit ripening and responses to abiotic stress. *Journal of Experimental Botany* **65**(16): 4577-4588

Leon-Reyes A, Du Y, Koornneef A, Proietti S, Körbes AP, Memelink J, Pieterse CMJ, Ritsema T (2012) Ethylene signalling renders the jasmonate response of Arabidopsis insensitive to future suppression by salicylic acid. *Molecular Plant-Microbe Interactions* **23**(2): 187-197

Li W, Ling H, Zhang F, Yao H, Sun X, Tang K (2012) Analysis of Arabidopsis genes encoding putative class III lipases. *Journal of Plant Biochemistry and Biotechnology* **21**(2): 261-267

Liang Z, Ma Y, Xu T, Cui B, Liu Y, Guo Z, Yang D, Xu W (2013) Effects of abscisic acid, gibberellin, ethylene and their interactions on production of phenolic acids in *Salvia miltiorrhiza Bunge* hairy roots. *PLoS ONE* **8**(9): e72806

Lin Z, Zhong S, Grierson D (2009) Recent advances in ethylene research. *Journal of Experimental Botany* **60**(12): 3311-3336

Little HA, Grumet R, Hancock JF (2009) Modified ethylene signalling as an example of engineering for complex traits: secondary effects and implications for environmental risk assessment. *Hortscience* **44**(1): 94-101

Liu Y, Roof S, Ye Z, Barry C, van Tuinen A, Vrebalov J, Bowler C, Giovannoni J (2004). Manipulation of light signal transduction as a means of modifying fruit nutritional quality in tomato. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 9897-9902.

Liu K, Kang BC, Jiang H, Moore SL, Li H, Watkins CB, Setter TL, Jahn MM (2005) A GH3 -like gene, *CcGH3*, isolated from *Capsicum chinense L* fruit is regulated by auxin and ethylene. *Plant Molecular Biology* **58**: 447-464

Liu ZB, Ulmasov T, Shi X, Hagen G, Guilfoyle TJ (1994) Soybean GH3 promoter contains multiple auxin-inducible elements. *The Plant Cell* **6**(5): 645-657

Lois LM, Rodríguez-Concepción M, Gallego F, Campos N, Boronat A (2000) Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. *Plant Journal* **22**: 503-513

Lopez AJ (1998) Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. *Annual Review of Genetics* **32**: 279-305

Low LE-T, Halimah A, Boon SH, Elyana MS, Tan CYA, Ooi LC-L, Cheah SC, Raha AR, Wan KL and Singh R (2008) Oil palm (*Elaeis guineensis Jacq*) tissue culture ESTs: identifying genes associated with callogenesis and embryogenesis. *BMC Plant Biology* **48**: 62

Ma B, Chen H, Chen S-Y, Zhang J-S (2014) Roles of ethylene in plant growth and responses to stress. *In* Tran L-SP, Pal S (eds.) *Phytohormones: A Window to Metabolism, Signalling and Biotechnological Applications*. New York, USA: Springer, 323-350

Ma B, Chen SY, Zhang SY (2010) Ethylene signalling in rice. *Chinese Science Bulletin* **55**(21): 2204-2210

Ma B, He S-J, Duan K-X, Yin C-C, Chen H, Yang C. Xiong Q, Song Q-X, Lu X, Chen H-W, Zhang W-K, Lu T-G, Chen S-Y, Zhang J-S (2013) Identification of rice ethylene-response mutants and characterization of MHZ7/OsEIN2 in distinct ethylene response and yield trait. *Regulation Molecular Plant* **6**(6): 1830-1848

Ma QH, Wang XM (2003) Characterization of an Ethylene receptor homologue from wheat and its expression during leaf senescence. *Journal of Experimental Botany* **54**(386): 1489-1490

Maizura I, Kushairi AD, Mohd Din A, Noh A, Marhalil M, Wong YT, Sambanthamurthi R (2008) PS13: Breeding population selected for low Llpase. *MPOB Information Series* No. 425

Mallard WJ, Reed J (1997) Automated Mass Spectral Deconvolution & Identification System AMDIS — User Guide Manual.

Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, Zhang D, Bryant SH (2013) CDD: Conserved domains and protein three-dimensional structure. *Nucleic Acids Research* **41**(D1): 1-5

Marin DH, Blankenship SM, Sutton TB, Swallow WH (1996) Physiological and chemical changes during ripening of Costa Rican bananas harvested in different seasons. J*ournal of the American Society for Horticultural Science* **121**(6):1157-1161

Martínez C, Manzano S, Megías Z, Garrido D, Picó B, Jamilena M (2013) Involvement of ethylene biosynthesis and signalling in fruit set and early fruit development in zucchini squash (*Cucurbita pepo L). BMC Plant Biology* **13**: 139

Masani MYA, Parveez GKA (2008) Development of transformation vectors for the production of potentially high oleate transgenic oil palm. Electronic Journal of Biotechnology 11(3)

Masani MYA, Noll G, Parveez GKA, Sambanthamurthi R, Pruefer D (2014) Efficient transformation of oil palm protoplasts by PEG-mediated transfection and DNA microinjection. *PLoS ONE* **9**(5): e96831

Marty I, Bureau S, Sarkissian G, Gouble B, Audergon JM, Albagnac G (2005) Ethylene regulation of carotenoid accumulation and carotenogenic gene expression

in colour-contrasted apricot varieties (*Prunus armeniaca). Journal of Experimental Botany* **56**(417): 1877-1886

Masli DIA, Kadir APG, Yunus AMM (2009) Transformation of oil palm using Agrobacterium tumefaciens. *Journal of Oil Palm Research* **21**: 643-652

Massimo M (1988) Environmental factors affecting the oil composition of some Mentha species grown in northwest Italy. *Flavour and Fragrance Journal* **3**(2): 79-84

Masura SS, Parveez GKA, Low LE-T (2011) Isolation and characterization of an Oil palm constitutive promoter derived from a translationally control tumor protein (TCTP) gene. *Plant Physiology and Biochemistry* **49**(7): 701-708

Matzke MA, Primig M, Tranovsky J, Matzke AJM (1989) Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. *EMBO Journal* **7**:643-649

McAtee P, Karim S, Schaffer R, David K (2013) A dynamic interplay between phytohormones is required for fruit development, maturation and ripening. *Frontiers in Plant Science* **4**: 79

McMurchie EJ, McGlasson WB, Eaks IL (1972) Treatment of fruit with propylene gives information about the biogenesis of ethylene. *Nature **237***(5352): 235-236

Mette MF, Winden VD, Matzke J, Matzke, MA (1999) Production of aberrant promoter transcripts contributes to methylation and silencing of unlinked homologous promoters in trans. *EMBO Journal* **18**: 241-248

Miskandar MS, Karimah A, Noor Lida MD, Rosnani AW, Salmi Yati S, Sivaruby K, Zaida Z (2011) Palm oil and palm kernel Oil for Low Trans Food Products. *In* Wahid MBA, Choo, YM, Chan KW (eds.) Further Advances in Oil Palm Research (2000-2010), Vol 2. Bangi, Malaysia: Malaysian Palm Oil Board, 685-720

Mita S, Kawamura S, Yamawaki K, Nakamura K, Hyodo H (1998) Differential expression of genes involved in the biosynthesis and perception of ethylene during ripening of passion fruit (*Passiflora edulis Sims). Plant and Cell Physiology* 39(11): 1209-1217

Mohanaraj SN, Donough CR (2013) Harvesting practices for maximum yield in oil palm: results from a reassessment at IJM plantations, Sabah. *Proceeding of International Palm Oil Congress 2013-ABS Agriculture* AP6: 23-27

Mohd Din A, Rajanaidu N, Jalani BS (2000) Performance of *Elaeis oleifera* from Panama, Costa Rica, Colombia and Honduras in Malaysia. *Journal of Oil Palm Research*. **12**(1): 71-80

Mohd Haniff H, Mohd Roslan MN (2002) Fruit set and oil palm bunch components. *Journal of Oil Palm Research* **14**(2): 24-33

Mohd Haniff, H (1986) PORIM Annual Research Report 1986. Palm Oil Research Institute of Malaysia. Bangi, Malaysia

Mohd Hudzari R, MaHalim AS, Roslan S (2012) A review on crop plant production and ripeness forecasting. *International Journal of Agriculture and Crop Sciences* **4**(2): 54-63

Mohd Nasriq HAR, Noor Azmi S, Nurniwalis AW, Puteri Edaroyati MW, Meilina O-A, Nur Ashikin PA, Parveez GKA, Roberts JA, Zubaidah R (2016) Impact of paclobutrazol on the growth and development of nursery grown oil palm (*Elaeis guineensis Jacq.)* seedlings. *Journal of Oil Palm Research.* In press.

Mohd Solah D, Rahim AS (2009) Roller-type oil palm loose fruit picker. MPOB TT **419**: 3-4

Montoya C, Cochard B, Flori A, Cros D, Lopes R, Cuellar T, Espeout S, Syaputra I, Villeneuve P, Pina M, Ritter E, Leroy T, Billotte N (2014) Genetic architecture of palm oil fatty acid composition in cultivated oil palm (*Elaeis guineensis Jacq.*) compared to its wild relative *E. oleifera* (H.B.K) Cortés. *PLoS ONE* **9**(5): e95412

Morcillo F, Cros D, Billotte N, Ngando-Ebongue GF, Domonhédo H, Pizot M, Cuéllar T, Espéout S, Dhouib R, Bourgis F, Claverol S, Tranbarger TJ, Nouy B, Arondel V (2013) Improving palm oil quality through identification and mapping of the lipase gene causing oil deterioration. *Nature Communications* **4**: 2160

Mounet F, Moing A, Garcia V, Petit J, Maucourt M, Deborde C, Bernillon S, Le Gall G, Colquhoun I, Defernez M, Giraudel JL, Rolin D, Rothan C, Lemaire-Chamley M (2009) Gene and metabolite regulatory network analysis of early developing fruit tissues highlights new candidate genes for the control of tomato fruit composition and development. *Plant Physiology* **149**(3): 1505-1528

Moussatche P (2004) The ethylene receptor multigene family: insights on expression, localization and function in Arabidopsis and tomato. Florida, USA: University of Florida, PhD Thesis

MPOB (2003) Oil Palm Grading Manual Second Edition. Malaysian Palm Oil Board. Bangi, Malaysia

MPOB (2016) Malaysian Oil Palm Statistics 2015. Malaysian Palm Oil Board. Bangi, Malaysia

Murphy DJ (1993) Plant lipids: their metabolism, function and utilization. *In* Lea RC, Leegood PJ (sds.) Plant Biochemistry and Molecular Biology. New York, USA: John Wiley and Sons, 113-128

Murphy DJ (2014) The future of oil palm as a major global crop: opportunities and challenges. *Journal of Oil Palm Research* **26**(1): 1-24

Mutert E, Fairhurst TH (1999) Oil palm clones: productivity enhancement for the future Malaysia. *Better Crops International* **13**(1): 45-47

Nakagawa S, Niimura Y, Gojobori T, Tanaka H, Miura KI (2008) Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. *Nucleic Acids Research* **36**(3): 861-871

Nakatsuka A, Murachi S, Okunishi H, Shiomi S, Nakano R, Kubo Y, Inaba A (1998) Differential expression and internal feedback regulation of 1-aminocyclopropane-1-carboxylate synthase, 1-aminocyclopropane-1-carboxylate oxidase, ethylene receptor genes in tomato fruit during development and ripening. *Plant Physiology* **118**(4): 1295-1305

Ngando Ebongue GF, Nouy B, Zok S, Carrière F, Zollo PH, Arondel V (2008) Identification of oil palm breeding lines producing oils with low acid values. *European Journal of Lipid Science and Technology* **110**: 506-509

Niu YY, Chen M, Chen XP, Ma YZ, Xu ZS, Li LC (2013) Characterization of ethylene receptors and their interactions with GmTPR - a novel tetratricopeptide repeat protein (TPR) in soybean (*Glycine max L). Journal of Integrative Agriculture* **12**(4): 571-581

Noh A, Rajanaidu N, Ahmad Kushairi D, Mohd Rafii Y, Mohd Din A, Mohd Isa ZA, Saleh G (2002) Variability in fatty acid composition, iodine value and carotene

content in the MPOB oil palm germplasm collection from Angola. *Journal of Oil Palm Research* **14**(2): 18-23.

Nualwijit N, Lerslerwong L (2014) Post harvest ripening of oil palm fruit is accelerated by application of exogenous ethylene. *Songklanakarin Journal of Science and Techology* **36**(3): 255-259

Nurfahisza AR, Rafiqah MA, Masani MYA, Hanin AN, Rasid OA, Parveez GKA, Ismail I (2014) Molecular analysis of transgenic oil palm to detect the presence of transgenes. *Journal of Oil Palm Research* **26**: 73-80

Nurniwalis AW (2006) Isolation of ethylene response sensor gene and generation of expressed sequenced tags from oil palm (*Elaeis guineensis Jacq.*) mesocarp. Selangor, Malaysia: University Putra Malaysia, MSc. Thesis

Nurniwalis AW, Siti Nor Akmar A (1999) Identification of differentially expressed gene(s) in the oil palm leaf and mesocarp tissue using differential display. *Proceedings of the 12th National Biotechnology Seminar*, 1999. Perak, Malaysia, 487-489

Nurniwalis AW, Siti Nor Akmar A (2001) Isolation of Differentially Expressed Genes in the Oil Palm Mesocarp by Differential Display and Screening. *Proceedings of 2001 International Palm Oil Congress, Agriculture Conference.* Kuala Lumpur, Malaysia: MPOB, 450-455

Nurniwalis AW, Siti Nor Akmar A, Tan SH (2003) Isolation of mesocarp-specific cDNA clone from oil palm *via* RT-PCR. *Proceedings of 2003 International Palm Oil Congress, Agriculture Conference*. Kuala Lumpur, Malaysia: MPOB, 766-772

Nurniwalis AW, Siti Nor Akmar A, Chan PL, Mohamad Arif AM (2007) Isolation of a cDNA encoding lipase class 3 family protein from oil palm (*Elaeis guineensis*). *Proceedings of the 2007 MPOB International Palm Oil Congress, Agriculture Conference.* Kuala Lumpur, Malaysia: MPOB, 1011-1020

Nurniwalis AW, Suhaimi N, Siti Nor Akmar A, Aminah S, Mohamad Arif MA (2008) Gene discovery *via* expressed sequence tags from the oil palm (*Elaeis guineensis Jacq)* mesocarp. *Journal of Oil Palm Research* **2**: 87-96

Nurniwalis, AW, Suhaimi, N, Siti Nor Akmar, A, Mohamad Arif, AM (2012) Isolation and characterization of an ethylene receptor (ERS-Type) cDNA from the oil palm (*Elaeis guineensis*) and Its Expression in the mesocarp during fruit development. *The IX$^{th}$ International Conference on the Plant Hormone Ethylene. Rotoroa, New Zealand:* 104

O'Malley RC, Rodriguez FI, Esch JJ, Binder BM, O'Donnell P, Klee HJ, Bleecker AB (2005) Ethylene-binding activity, gene expression levels, receptor system output for ethylene receptor family members from Arabidopsis and tomato. *Plant Journal* **41**(5): 651-659

Oil World (2016) World production (1000T), yields (T/ha) and harvested Area (1000ha) 2013/14-2015/16. *Oil World Annual Vol.1*

Oo KC, Lee KB, Ong ASH (1986) Changes in fatty acid composition of the lipid classes in the developing oil palm mesocarp. *Phytochemistry* **25**: 405-407

Oo KC, Stumpf PK (1983) Some enzymic activities in the germinating oil palm (*Elaeis guineensis*) seedling. *Plant Physiology* **73**(4): 1028-1032

Osorio S, Scossa F, Fernie AR, Bennett M, Giovannoni J, Seymour G (2013) Molecular regulation of fruit ripening. *Frontiers in Plant Science* **4**: 198

Pagay V, Cheng L (2010) Variability in berry maturation of concord and cabernet franc in a cool climate. *American Journal of Enology and Viticulture* **61**: 61-67

Pang JH, Ma B, Sun HJ, Ortiz GI, Imanishi S, Sugaya S, Gemma H, Ezura H (2007) Identification and characterization of ethylene receptor homologs expressed during fruit development and ripening in persimmon (*Diospyros kaki Thumb). Postharvest Biology and Technology* **44**(3): 195-203

Panzanaro S, Nutricati E, Miceli A, Bellis LD (2010) Biochemical characterization of a lipase from olive fruit (*Olea europaea L). Plant Physiology and Biochemistry* **48**(9): 741-745

Parkinson JS, Kofoid EC (1992) Communication modules in bacterial signalling proteins. *Annual Review of Genetics* **26**(1): 71-112

Parveez GKA, Chowdhury MKU, Saleh NM (1996) Determination of minimal inhibitory concentration of selection agents for oil palm (*Elaeis guineensis* Jacq.) transformation. *Asia-Pacific Journal of Molecular Biology and Biotechnology* **4**: 219-228

Parveez GKA, Chowdhury MKU, Saleh NM (1997) Physical parameters affecting transient GUS gene expression in oil palm using the biolistics device. *Industrial Crops and Products* **6**: 41-50

Parveez GKA, Chowdhury MKU, Saleh NM (1998) Biological parameters affecting transient GUS gene expression in oil palm embryogenic calli *via* microprojectile bombardment. *Industrial Crops and Products* **8**:17-27

Parveez GKA (2000) Production of transgenic oil palm (Elaeis guineensis Jacq.) using biolistic techniques. *In* Jain SM, Minocha, SC (eds.) *Molecular Biology of Woody Plants (Vol. 2)* Netherlands, Kluwer Academic Publishers, 327-350

Parveez GKA (2003) Novel products from transgenic oil palm. *AgBiotechNet* **5**:1-8

Parveez GKA, Rasid OA, Masani AMY, Haliza HF, Na`imatulapidah AM, Kushairi AD, Tarmizi AH, Zamzuri I (2003) Transgenic oil palm: where are we?. *In* Murata N, Yamada M, Nishida I, Okuyama H, Sekiya J and Wada H (eds.) *Advanced Research on Plant Lipids.* Netherlands: Kluwer Academic Publishers, 415-418

Parveez GKA (2003) Novel products from transgenic oil palm. *Ag Biotech Net* **5**(ABN113): 1-8

Parveez GKA, Rasid OA, Masani MYA, Sambanthamurthi R (2015) Biotechnology of oil palm: strategies towards manipulation of lipid content and composition. *Plant Cell Report* **34**(4): 533-543

Patil KJ, Chopda MZ, Mahajan RT (2011) Lipase biodiversity. *Indian Journal of Science and Technology* **4**(8): 971-982

Payton S, Fray RG, Brown S, Grierson D (1996) Ethylene receptor expression is regulated during fruit ripening, flower senescence and abscission. *Plant Molecular Biology.* **31**(6): 1227-1231

Pech JC, Bouzayen M, Latché A (2008) Climacteric fruit ripening: ethylene-dependent and independent regulation of ripening pathways in melon fruit. *Plant Science* **175**(1-2): 114-120

Pesaresi P, Mizzotti C, Colombo M, Masiero S, Franks RG, Wu X (2014) Genetic regulation and structural changes during tomato fruit development and ripening. *Frontiers in Plant Science* **5**: 124

Picton S, Barton SL, Bouzayen M, Hamilton AJ, Grierson D (1993) Altered fruit ripening and leaf senescence in tomatoes expressing an antisense ethylene-forming enzyme transgene. *The Plant Journal* **3**: 469-481

Pizarro L, Stange C (2009) Light-dependent regulation of carotenoid biosynthesis in plants. *Ciencia e Investigación Agraria* **36**:143-162

PORIM (1995). PORIM Test Methods. Palm Oil Research Institute of Malaysia, Bangi, Malaysia

PORAM (2015) Standard specifications for crude palm oil. http://econmpobgovmy/economy/exporters/PORAMSTANDARDhtm. Accessed 5th November 2015

Porto MS, Pinheiro MP, Batista VG, dos Santos RC, Filho PdeA, Lima LM (2013) Plant promoters: an approach of structure and function. *Molecular Biotechnology* **56**(1): 38-49

Prabowo NE, Foster HL (1998) Variations in oil and kernel extraction rates of palms in North Sumatera due to nutritional and climate factors. *Proceeding of the 1998 International Oil Palm Conference Commodity of the Past, Present and the Future*. Medan, Indonesia: 275-286

Prada F, Ayala-Diaz IM, Delgado W, Ruiz-Romero R, Romero HM (2013) Effect of fruit ripening on content and chemical composition of oil from three oil palm cultivars (*Elaeis guineensis Jacq.)* grown in Colombia. *Journal of Agricultural and Food Chemistry* **59**(18): 10136-10142.

Prescott A, Martin C (1987) A rapid method for the quantitative assessment of levels of specific mRNAS in plants. *Plant Molecular Biology Reporter* **4**(4): 219-224

Przybylski R, Mag T, Eskin N, McDonald B (2005) Canola oil. Bailey's Industrial Oil and Fat Products. **2**: 2

Qin X, Zeevaart JA (2002) Overexpression of a 9-cis-epoxycarotenoid dioxygenase gene in *Nicotiana plumbaginifolia* increases abscisic acid and phaseic acid levels and enhances drought tolerance. *Plant Physiology* **128**(2): 544-551

Rasid OA, Nor Hanin A, Masura SS, Masani AMY, Singh R, Ho CL, Suhaimi N, Parveez GKA, Sambanthamurthi R (2007) Lycopene: genetic engineering of potential nutraceutical. *Proceedings of the PIPOC 2007*. Kuala Lumpur, Malaysia, MPOB: 248-273

Rasid OA, Wan Nur Syuhada WS, Nor Hanin A, Masura SS, Ho CL, Suhaimi N, Sambanthamurthi R (2008) RT-PCR amplification and cloning of partial DNA sequence coding for oil palm (*Elaeis oleifera*) phytoene synthase. *Asia Pacific Journal of Molecular Biology and Biotechnology* **16**(1): 17-24

Rasid OA, Parveez GKA, Ho CL, Sambanthamurthi R, Suhaimi N (2009) Plant carotenoids: molecular genetics and regulation. *Journal of Oil Palm Research* **21**: 588-601

Rasori A, Ruperti B, Bonghi C, Tonutti P, Ramina A (2002) Characterization of two putative ethylene receptor genes expressed during peach fruit development and abscission. *Journal of Experimental Botany* **53**(379): 2333-2339

Razungles AJ, Baumes RL, Dufour C, Sznaper CN, Bayonove CL (1998) Effect of sun exposure on carotenoids and C13-norisoprenoid glycosides in Syrah berries (*Vitis vinifera L.)*. *Sciences Des Aliments* **18**: 361-373

Reid MS (1985) Ethylene and abscission. *HortScience* **20**(1): 45-50

Rival A, Jaligot E (2010) Oil palm biotechnologies are definitely out of infancy. *OCL.* 17(6): 368-374

Rodrıguez FL, Esch JJ, Hall AE, Binder BM, Schaller GE, Bleecker AB (1999) A copper cofactor in the ethylene receptor *ETR1* from Arabidopsis. *Science* **283**(5404): 996-998

Roongsattham P, Morcillo F, Chatchawan, J, Pizot, Moussu S, Jayaweera D, Collin M, Gonzalez-Carranza, ZH, Amblard P,, Tregear JW,1 Tragoonrung S, Verdeil J-L and Tranbarger TJ (2012) Temporal and spatial expression of polygalacturonase gene family members reveals divergent regulation during fleshy fruit ripening and abscission in the monocot species oil palm. *BMC Plant Biology* **12**(1): 150

Rounsley S, Marri PR, Yu Y, He R, Sisneros N, Goicoechea JL, Lee SJ, Angelova A, Kudrna D, Luo M, Affourtit J, Desany B, Knight J, Niazi F, Egholm M, Wing RA (2009) De novo next generation sequencing of plant genomes. *Rice* **2**(1): 35-43

Roy A, Kucukural A, Zhang Y (2010) I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocols* **5**(4): 725-738

Saeed OMB, Sankaran S, Shariff ARM, Shafri HZM, Ehsani R, Alfatni MS, Hazir MHM (2012) Classification of oil palm fresh fruit bunches based on their maturity using portable four-band sensor system. *Computers and Electronics in Agriculture* **82**: 55-60

Sakai H, Hua J, Chen QG, Chang C, Medrano LJ, Bleecker AB, Meyerowitz EM (1998) ETR2 is an ETR1-like gene involved in ethylene signalling in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **95**(10): 5812-5817

Sambanthamurthi R, Chong CL, Oo K-C, Yeo KH, Rajan P (1991) Chilling-Induced lipid hydrolysis in the Oil Palm (*Elaeis guineensis)* mesocarp. *Journal of Experimental Botany **42***(9): 1199-1205

Sambanthamurthi R, Kushairi AD (2002) Selection for lipase activity in the oil palm. *MPOB Information Series* No 158

Sambanthamurthi R, Oo K-C, Siti Hasnah P (1995) Factors affecting lipase activity in *Elaeis guineensis* mesocarp. *Plant Physiology and Biochemistry.* **33**(3): 353-359

Sambanthamurthi R, Sundram K, Tan Y (2000) Chemistry and biochemistry of palm oil. *Progress in Lipid Research* **39**: 507-558

Sambanthamurthi R, Abrizah O, Umi Salamah R (1999) Biochemical factors that control oil composition in the oil palm. *Journal of Oil Palm Research* **Special Issue**: 24-33

Sambanthamurthi R, Siti Nor Akmar A, Parveez GKA (2002) Genetic manipulation of the oil palm: challenges and future prospects. *The Planter* **78**(919): 547-562.

Sambrook J, Russell DW (2001) Molecular Cloning: A Laboratory Manual New York: Cold Spring Harbor Laboratory Press, New York, USA

Santner A, Estelle M (2009) Recent advances and emerging trends in plant hormone signalling. *Nature* **459**(7250): 1071-1078

Sato-Nara K, Yuhashi KI, Higashi K, Hosoya K, Kubota M, Ezura H (1999) Stage- and tissue-specific expression of ethylene receptor homolog genes during fruit development in muskmelon. *Plant Physiology* **119**: 321-329

Sawant SV, Singh PK, Gupta SK, Madnala R, Tuli R (1999) Conserved nucleotide sequences in highly expressed genes in plants. *Journal of Genetics* **78**(2): 123-131

Sawicki M, Ait Barka E, Clement C, Vaillant-Gaveau N, Jacquard C (2015) Cross-talk between environmental stresses and plant metabolism during reproductive organ abscission. *Journal of Experimental Botany* **66**(7): 1707-1719

Schaffer RJ, Friel EN, Souleyre EJ, Bolitho K, Thodey K, Ledger S, Bowen JH, Ma JH, Nain B, Cohen D, Gleave AP, Crowhurst RN, Janssen BJ, Yao JL, Newcomb RD (2007) A genomics approach reveals that aroma production in apple Is controlled by ethylene predominantly at the final step in each biosynthetic pathway. *Plant Physiology* **144**(4): 1899-1912

Schaller GE, Kieber JJ (2002) Ethylene. The Arabidopsis book. 1:e0071

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D et al. (2010) Genome sequence of the palaeopolyploid soybean. *Nature* **463**(7278): 178-183

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P et al. (2009) The B73 maize genome: complexity, diversity and dynamics. *Science* **326**(5956): 1112-1115

Seki M, Umezawa T, Urano K, Shinozaki K (2007) Regulatory metabolic networks in drought stress responses. *Current Opinion in Plant Biology* **10**(3):296-302

Setha S (2012) Roles of abscisic acid in fruit ripening. *Walailak Journal of Science and Technology* **9**(4): 297-308

Seymour GB, Østergaard L, Chapman NH, Knapp S, Martin, C (2013) Fruit development and ripening. *Annual Review of Plant Biology* **64**: 219-241

Shaarani SM, Cárdenas-Blanco A, Amin MHG, Soon MG, Hall LD (2010) Monitoring development and ripeness of oil palm fruit (*Elaeis guneensis*) by MRI and bulk NMR. *International Journal of Agriculture and Biology* **12**(1): 101-105

Shakeel SN, Wang X, Binder BM, Schaller GE (2013) Mechanisms of Signal Transduction by Ethylene: overlapping and non-overlapping signalling Roles in a receptor family. *AoB PLANTS* **5**: plt010

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P et al. (2011) The genome of woodland strawberry (*Fragaria vesca*). *Nature Genetics* **43**: 109-116

Siew WL (2011) Composition and properties of new palm oil products. *In* Wahid MBA, Choo, YM, Chan KW (eds.) *Further Advances in Oil Palm Research (2000-2010), Vol 2.* Bangi, Malaysia: Malaysian Palm Oil Board, 673-684

Singh R, Low ET, Ooi LC, Ong-Abdullah M, Nookiah R, Ting NC, Marjuni M, Chan PL, Ithnin M, Manaf MA, Nagappan J, Chan KL, Rosli R, Halim MA, Azizi N, Budiman MA, Lakey N, Bacher B, Van Brunt A, Wang C, Hogan M, He D, MacDonald JD, Smith SW, Ordway JM, Martienssen RA, Sambanthamurthi R (2014) The oil palm VIRESCENS gene controls fruit colour and encodes a R2R3-MYB. *Nature Communications.* **5**: 4106

Singh R, Ong-Abdullah M, Low ET, Manaf MA, Rosli R, Nookiah R, Ooi LC, Ooi SE, Chan KL, Halim MA, Azizi N, Nagappan J, Bacher B, Lakey N, Smith SW, He D, Hogan M, Budiman MA, Lee EK, DeSalle R, Kudrna D, Goicoechea JL, Wing RA, Wilson RK, Fulton RS, Ordway JM, Martienssen RA, Sambanthamurthi R (2013a)

Oil palm genome sequence reveals divergence of interfertile species in old and new worlds. *Nature* **500**(7462): 335-339

Singh R, Low ET, Ooi LC, Ong-Abdullah M, Ting NC, Nagappan J, Nookiah R, Amiruddin MD, Rosli R, Manaf MA, Chan KL, Halim MA, Azizi N, Lakey N, Smith SW, Budiman MA, Hogan M, Bacher B, Van Brunt A, Wang C, Ordway JM, Sambanthamurthi R, Martienssen RA (2013b) The oil palm SHELL gene controls oil yield and encodes a homologue of SEEDSTICK. *Nature* **500**(7462): 340-344

Siti Nor Akmar A, Shah FH, Cheah SC (1995) Construction of oil palm mesocarp cDNA library and the isolation of mesocarp-specific cDNA clones. *Asia Pacific Journal of Molecular Biology and Biotechnology* **3**: 106-111

Siti Nor Akmar A, Murphy DJ, Cheah SC (1996) Isolation of mesocarp-specific cDNA clones from the oil palm by subtractive hybridization. *Proceedings of 1996 PORIM International Palm Oil Congress, Agriculture Conference*, 1996. Kuala Lumpur, Malaysia: MPOB, 552-556

Siti Nor Akmar A, Cheah SC, Aminah S, Leslie CLO, Sambanthamurthi R, Murphy DJ (1999) Characterization and regulation of the oil palm (*Elaeis guineensis*) stearoyl-ACP desaturase genes. *Journal of Oil Palm Research* **Special Issue**: 1-17

Siti Nor Akmar A, Sambanthamurthi R, Parveez, GKA (2001) Genetic modification of oil palm for producing novel oils. *Proceedings of the PIPOC International Palm Oil Congress, Agriculture Conference*, 2001. Kuala Lumpur, Malaysia, MPOB: 18-30

Siti Nor Akmar A, Zubaidah R (2008) Mesocarp-specific metallothionein-like gene promoter for genetic engineering of oil palm. *Journal of Oil Palm Research* **2**: 1-8

Smita S, Rajwanshi R, Lenka SK, Katiyar A, Chinnusamy V and Bansal KC (2013) Expression profile of genes coding for carotenoid biosynthetic pathway during ripening and their association with accumulation of lycopene in tomato fruits. *Journal of Genetics* **92**: 363-368

Soh AC, Wong G, Tan CC, Chem PS, Chong SP, Ho YW, Wong CK, Choo CN, Nor AH, Kumar K (2011) Commercial-scale propagation and planting of elite oil palm clones: research and development towards realization. *Journal of Oil Palm Res*earch **23**: 935-952

Solano R, Stepanova AN, Chao Q, Ecker JR (1998) Nuclear Events in Ethylene Signalling: A transcriptional cascade mediated by ethylene-insensitive3 and ethylene-response-factor1. *Genes and Development* **12**(23): 3703-3714

Srivastava A, Handa AK (2005) Hormonal regulation of tomato fruit development: A molecular perspective. *Journal of Plant Growth Regulation* **24**: 67-82

Srivastava LM (2002) Plant Growth and Development: Hormones and Environment. Sydney, Australia: Academic Press

Stearns JC, Glick BR (2003) Transgenic plants with altered ethylene biosynthesis or perception. *Biotechnology Advances* **21**(3): 193-210

Stepanova AN, Alonso JM (2005) Arabidopsis ethylene signalling pathway. *Science Signalling* **276**: cm4

Su L, Diretto G, Purgatto E, Danoun S, Zouine M, Li Z, Roustan J-P, Bouzayen M, Giuliano G, Chervin C (2015) Carotenoid accumulation during tomato fruit ripening is modulated by the auxin-ethylene balance. *BMC Plant Biology* **15**: 114

Sulaiman O, Hashim R, Wahab R, Samsi HW, Mohamed AH (2008) Evaluation on some finishing properties of oil palm plywood. *Holz Roh Werkst* **66**: 5-10

Sulaiman O, Nurjannah S, Noor Afeefah N, Rokiah H, Mazlan I, Sato M (2012) The potential of oil palm trunk biomass as an alternative source for compressed wood. *BioResources* **7**(2): 2688-2706

Syahanim S, Abrizah O, Siti Nor Akmar A, Mohamad Arif AM, Ho CL (2007) Cloning of an oleoyl-coa desaturase from oil palm. *Proceedings of the PIPOC 2007 International Palm Oil Congress (Agriculture, Biotechnology & Sustainability)*, 2007. Kuala Lumpur, Malaysia, MPOB: 1001-1009

Szalontai B (2012) Functional analysis reveals an expansive evolution of the PRLIP (pathogenesis related lipase) gene family in plants. Pécs, Hungary: University of Pécs, PhD Thesis

Takada K, Kamada H, Ezura H (2005) Production of male sterile transgenic plants. *Plant Biotechnology* **22**(5): 469-476

Takada K, Watanabe S, Sano T, Ma B, Kamada H, Ezura H (2007) Heterologous expression of the mutated melon ethylene receptor gene *Cm-ERS1/H70A* produces stable sterility in transgenic lettuce (*Lactuca sativa). Journal of Plant Physiology* **164**(4): 514-520

Takaiwa F, Yamanouchi U, Yoshihara T, Washida H, Tanabe F, Kato A, Yamada K (1996) Characterization of common cis-regulatory elements responsible for the endosperm-specific expression of members of the rice glutelin multigene family. *Plant Molecular Biology* **30**(6): 1207-1221

Juven-Gershon T, Kadonagaa JT (2010) Regulation of gene expression *via* the core promoter and the basal transcriptional machinery. *Developmental Biology* **339**(2): 225-229

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, maximum parsimony methods. *Molecular Biology and Evolution* **28**(10): 2731-2739

Tan YA, Chong CL, Low KS (1997) Crude palm oil characteristics and chlorophyll content. *Journal of the Sciences of Food Agriculture* **75**: 281-288

Tan YA, Low KS, Chong C-L (1994) Rapid determination of chlorophylls in vegetable oils by laser-based fluorometry. *Journal of the Sciences of Food Agriculture* **66**: 479-484

Tan BK, Oh CH (1981). Malaysian palm oil chemical and physical characteristics. *PORIM Technology* **3**: 5

Tan YA, Kuntom A, Siew WL, Yusof M, Chong CL (2000) Crude palm oil survey 97/98 quality and identity characteristics. *PORIM Tech.* 22

Tang TS (2000) Composition and properties of palm oil products. *In* Jalani BS, Chan KW, Yusof B (eds.) *Advances in Oil Palm Research (Vol 1).* Bangi, Malaysia: Malaysian Palm Oil Board, 845-895

Tatsuki M, Hayama H, Nakamura Y (2009) Apple ethylene receptor protein concentrations are affected by ethylene, differ in cultivars that have different storage life. *Planta* **230**(2): 407-417

Tay BYP, Choo YM (2000) Valuable minor constituents of commercial red palm olein: carotenoids, vitamin E, ubiquinones and sterols. *Journal of Oil Palm Research* **12**:14-24

Taylor IB, Burbidge A, Thompson AJ (2000) Control of abscisic acid synthesis. *Journal of Experimental Botany* **51**:1563-1574

Teh HF, Neoh BK, Hong MPL, Low JYS, Ng TLM, Ithnin N, Thang YM, Mohamed M, Chew FT, Yusof HM, Kulaveerasingam H, Appleton DR (2013) Differential metabolite profiles during fruit development in high-yielding oil palm mesocarp. *PLoS ONE* **8**(4): e61344

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana. Nature* **408**(6814): 796-815

The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. *Nature* **475**(7355): 189-195

The Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**(7400): 635-641

Thomas RL, Sew P, Mok CK, Chan KW, Easau PT, Ng SC (1971) Fruit ripening in the Oil palm *Elaeis guineensis. Annals of Botany* **35**(5): 1219-1225

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**(22): 4673-4680

Tieman DM, Klee HJ (1999) Differential expression of two novel members of the tomato ethylene receptor family. *Plant Physiology* **120**(95): 165-172

Tieman DM, Taylor MG, Ciardi JA, Klee HJ (2000) The tomato ethylene receptors *NR* and *LeETR4* are negative regulators of ethylene response and exhibit functional compensation within a multigene family. *Proceedings of the National Academy of Sciences of the United States of America* **97**(10): 5663-5668

Ting N-C, Jansen J, Mayes S, Massawe F, Sambanthamurthi R, Ooi C-L, Chin CW, Arulandoo X, Seng T-Y, Sharifah SRSA, Maizura I, Singh R (2014) High density SNP and SSR-based genetic maps of two independent oil palm hybrids. *BMC Genomics.* **15**:309

Toledo-Ortiz G, Huqb E and Rodríguez-Concepcióna M (2010) Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. *Proceedings of the National Academy of Sciences of the United States of America* **107**(25): 11626-11631

Trainotti L, Pavanello A, Casadoro G (2005) Different ethylene receptors show an increased expression during the ripening of strawberries: does such an increment imply a role for ethylene in the ripening of these non-climacteric fruits?. *Journal of Experimental Botany* **56**(418): 2037-2046

Tranbarger TJ, Dussert S, Joët T, Argout X, Summo M, Champion A, Cros D, Omore A, Nouy B, Morcillo F (2011) Regulatory mechanisms underlying oil palm fruit mesocarp maturation, ripening, functional specialization in lipid and carotenoid metabolism. *Plant Physiology* **156**(2): 564-584

Umi Salamah R, Sambanthamurthi R, Omar AR, Parveez GKA, Mohamad Arif AM, Abrizah O, Abdul Masani Y, Cheah SC, Sharifah SRSA, Siti Nor Akmar A, Mohd Basri W. (2012) The isolation and characterisation of oil palm (*Elaeis guineensis*

*Jacq*.) β-ketoacy-acyl carrier protein (ACP) synthase (*KAS) II* cDNA. *Journal of Oil Palm Research* **24**: 1480-1491

Van Heel WA, Breure CJ, Menendez T (1987) The early development of inflorescences and flowers of the oil palm (*Elaeis guineensis Jacq.)* seen through the scanning electron microscope. *Blumea* **32**:67-78

Vaucheret, H (1993) Identification of a general silencer for 19S and 35S promoters in transgenic tobacco plant: a 90 bp of homology in the promoter sequence are sufficient for trans-inactivation. *Comptes Rendus de l'Académie des Sciences Paris* **317**: 1471-1483

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S et al. (2010) The genome of the domesticated apple (*Malus 3 domestica Borkh*.). *Nature Genetics* **42**: 833-839

Voorrips RE (2002) Map chart: software for the graphical presentation of linkage maps and QTLs. *The Journal of Heredity* **93**(1): 77-78

VSN International (2011). GenStat for Windows 14th Edition. VSN International, Hemel Hempstead, UK. Web page: GenStat.co.uk

Wahab AG (2015) Global Agriculture Information Network Report: *Biofuels Annual 2015* - MY5012

Wahid MBA, Siti Nor Akmar A and Henson IE (2005) Oil palm – achievements and potential. *Plant Production Science* **8**(3): 1-13

Wahid, AWB (2009) Palm oil board achieves another first in oil palm genomes development. http://bizthestarcommy/news/storyasp?file=/2009/11/16/business/ 5108 011&sec=business. Accessed 5th November 2015

Wan Zahari M, Alimon AR (2004) Use of palm kernel cake and oil palm by-products in compound feed. *Palm Oil Developments* **40**: 5-9

Wang H, Jones B, Li Z, Frasse P, Delalande C, Regad, F, Chaabouni S, Latché A, Pech J-C, Bouzayen M (2005) The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. *The Plant Cell* **17**(10): 2676-2692

Watanabe H, Saigusa M, Hase S, Hayakawa T, Satoh S (2004) Cloning of a cDNA encoding an ETR2-like protein (Os-ERL1) from deep water rice (*Oryza sativa L*) and increase in its mRNA level by submergence, ethylene, gibberellin treatments. *Journal of Experimental Botany* **55**(399): 1145-1148

Weiss D, Naomi O (2007) Mechanisms of cross talk between gibberellin and other hormones. *Plant Physiology* **144**: 1240-1246

Whitmore TC (1979) *Palms of Malaya*. Repeat print Kuala Lumpur, Malaysia: Oxford University Press

Wilkinson JQ, Lanahan MB, Clark DG, Bleeker AB, Chang C, Meyerowitz ME, Klee HJ (1997) A dominant mutant receptor from Arabidopsis confers ethylene insensitivity in heterologous plants. *Nature Biotechnology* **15**: 444-447

Wong YT, Kushairi AD, Mohamad O, Sambanthamurthi R (2005) Assay of lipase activity in *Elaeis guineensis* in various germplasms. *Proceedings of the 2005 Conference on Biotechnology of Plantation Commodities,* 2005. Kuala Lumpur, Malaysia: Ministry of Plantation Industries and Commodity, 441-444

Wong YT, Kushairi AD, Rajanaidu N, Osman M, Wickneswari R, Sambanthamurthi R (2015) Screening of wild oil palm (*Elaeis guineensis*) germplasm for lipase activity. *Journal of Agricultural Science* 1-12

Xie C, Zhang JS, Zhou HL, Li J, Zhang ZG, Wang DW, Chen SY (2003) Serine/threonine kinase activity in the putative histidine kinase-like ethylene receptor NTHK1 from Tobacco. *Plant Journal* **33**(2): 385-393

Xiong AS, Yao QH, Peng RH, Li X, Han PL, Fan HQ (2005) Different effects on ACC oxidase gene silencing triggered by RNA interference in transgenic tomato. *Plant Cell Reports* **23**(9): 639-646 D

Xiong GS, Li JY, Wang YH (2009) Advances in the regulation and crosstalks of phytohormones. *Chinese Science Bulletin* **54**: 4069-4082

Xu Z, Zhang D, Hu J, Zhou X, Ye X, Reichel KL, Stewart NR, Syrenne RD, Yang X, Gao P, Shi W, Doeppke C, Sykes RW, Burris JN, Bozell JJ, Cheng MZ, Hayes DG, Labbe N, Davis M, Stewart CN Jr, Yuan JS (2009) Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom. *BMC Bioinformatics* **10** Suppl 11: S3

Yamagata H, Yonesu K, Hirata A, Aizono Y (2002) TGTCACA motif is a novel cis-regulatory enhancer element involved in fruit-specific expression of the cucumisin gene *Journal of Biological Chemistry* **277**(13): 11582-11590

Yamamoto YY, Ichida H, Matsui M, Obokata J, Sakurai T, Satou M, Seki M, Shinozaki K, Abe T (2007) Identification of plant promoter constituents by analysis of local distribution of short sequences. *BMC Genomics* **8**(8): 67

Yang C, Lu X, Ma B, Chen S-Y, Zhang J-S (2015) Ethylene signalling in rice and Arabidopsis: conserved and diverged aspects. *Molecular Plant* **8**: 495-505

Yang SF, Hoffman NE (1984) Ethylene biosynthesis and its regulation in higher plants. *Annual Review of Plant Physiology* **35**(1): 155-189

Yang X, Song J, Campbell-Palmer L, Fillmore S, Zhang Z (2013) Effect of ethylene and 1-MCP on expression of genes involved in ethylene biosynthesis and perception during ripening of apple fruit. *Postharvest Biology and Technology* **78**: 55-66

Yasumoto S, Terakado Y, Matsuzaki M, Okada K (2011) Effects of High water Table and short-term flooding on growth, yield, and seed quality of sunflower. *Plant Production Science* **14**(3):233-248

Yau CP, Wang L, Yu M, Sze YZ, Yip WK (2004) Differential expression of three genes encoding an ethylene receptor in rice during development, in response to indole-3-acetic acid and silver ions. *Journal of Experimental Botany* **55**(397): 547-556

Yuan H, Zhang J, Nageswaran D, Li L (2015) Carotenoid metabolism and regulation in horticultural crops. *Horticulture Research* **2**: 15036

Yin XR, Chen KS, Allan AC, Zhang RMB, Lallu N, Ferguson IB (2008) ethylene-induced modulation of genes associated with the ethylene signalling pathway in ripening Kiwifruit. *Journal of Experimental Botany* **59**(8): 2097-2108

Yuki Y, Pierik R, Fricker MD, Voesenek LA, Harberd NP (2012) Studies of *Physcomitrella patens* reveal that ethylene mediated submergence responses arose relatively early in land-plant evolution. *Plant Journal* **72**(6): 947-959

Yusof B (2000) Techno-economic aspects of research and development in the Malaysian oil palm industry. *In* Jalani BS, Chan KW, Yusof B (eds.) *Advances in Oil Palm Research (Vol 1).* Bangi, Malaysia: Malaysian Palm Oil Board, 1-18

Yusoff MB, Fauziah AH, Suhaila AJ (2000) Globalisation, economic policy and equity : the case of Malaysia. Paper presented at OECD workshop on 'Poverty and Income Inequality in Developing Countries: A Policy Dialogue on the Effects of Globalisation', 30 Nov–1 Dec, Paris

Zhang M, Yuan B, Leng P (2009) The role of ABA in triggering ethylene biosynthesis and ripening of tomato fruit. *Journal of Experimental Botany* **60**(6): 1579-1588

Zhang Y (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9: 40

Zhang Z-G, Zhou H-L, Chen T, Gong Y, Cao W-H, Wang Y-J, Zhang J-S and Chen S-Y (2004) Evidence for serine / threonine and histidine kinase activity in the tobacco ethylene receptor protein NtHK2. *Plant Physiology* **136**: 2971-2981

Zhang M, Yuan B, Leng P (2009) The role of ABA in triggering ethylene biosynthesis and ripening of tomato fruit. *Journal of Experimental Botany* **60**:1579-1588

Zheng C, Baum BJ (2008) Evaluation of promoters for use in tissue-specific gene delivery. *Methods in Molecular Biology* **434**: 205-219

Zubaidah R, Siti Nor Akmar A (2010) Functional Characterisation of the oil palm type 3 metallothionein-like gene (MT3-B) promoter. *Plant Molecular Biology Reporter* **28**(3): 531-541

# APPENDICES

## APPENDIX 1: List and preparation of standard solutions and buffers

i) 1M Tris-HCl, pH 7.5, 8.0, 9.0 (500 ml)

The solution was prepared by adding 60.57 g Tris-base and 300 ml sterile water. The pH was adjusted to the appropriate pH value by adding 10M HCl. Sterile water was added to make up a final volume of 500 ml. The solution was autoclaved and kept at 4$^{o}$C.

ii) 8M LiCl (500 ml)

The solution was prepared by adding 169.56 g LiCl and sterile water to a final volume of 500 ml. The solution was filtered, autoclaved and kept at 4$^{o}$C.

iii) 20% SDS (200 ml)

The solution was prepared by adding 40 g SDS and sterile water to a final volume of 200 ml and kept at room temperature.

iv) 0.5M EDTA, pH 8.0 (500 ml)

The solution was prepared by adding 93.05 g EDTA to 300 ml of sterile water. The pH was adjusted to 8.0 by adding NaOH pellets with constant stirring and the final volume was adjusted to 500 ml with sterile water. The solution was autoclaved and kept at 4$^{o}$C.

v) 0.1M Aurin trycarboxylic acid (50 ml)

The solution was prepared by adding 2.1 g aurin tricarboxylic acid and sterile water to a final volume of 50 ml and kept at 4$^{o}$C.

vi) 1M HCl

The solution was prepared by adding 82.6 ml of concentrated HCl (37% fuming) and sterile water to a final volume of 1 L and kept at room temperature.

vii) TE Buffer, pH 8.0 (100 ml)

The solution was prepared by adding 1 ml of 1M Tris-HCl, pH 8.0, 1 ml 0.5M EDTA and sterile water to a final volume of 100 ml and kept at 4$^{o}$C.

viii) 5M NaCl (500 ml)

The solution was prepared by adding 146.1 g NaCl to 300 ml sterile water. After adjusting the final volume to 500 ml with sterile water, the solution was autoclaved and kept at room temperature.

ix) 3M Sodium acetate, pH 4.8 (500 ml)

The solution was prepared by adding 86.25 ml glacial acetic acid and sterile water. The pH was adjusted to 4.8 with the addition of 10N NaOH to a final volume of 500 ml. The solution was autoclaved and kept at 4$^{o}$C.

x) 0.5M Ascorbic Acid (500 ml)

The solution was prepared by adding 44.025 g ascorbic acid and sterile water to a final volume of 500 ml. The solution was filtered and kept at 4$^{o}$C.

(x) 0.4M DIECA (500 ml)

The solution was prepared by adding 34.26 g DIECA and sterile water to a final volume of 500 ml. The solution was filtered and kept at 4$^{o}$C.

xi) TAE Buffer, pH 7.9.

The 50X TAE stock buffer was prepared by adding 242g Tris-base, 57.1 ml glacial acetic acid, 100 ml 0.5M EDTA and 800 ml of sterile water. After mixing thoroughly, the final volume was adjusted to 1 L with sterile water and kept at room temperature for use. The 1X TAE buffer was prepared via dilution of the 50X TAE stock buffer.

## APPENDIX 2: List and preparation of media and reagents

i) LB Agar

The LB agar was prepared by adding 37 g Luria agar (Miller's LB Agar) and 800 ml deionized water to a final volume of 1 L prior to autoclaving. The media was then kept at 4$^{o}$C until needed.

ii) LB Broth

The media was prepared by adding 25 g Luria broth base (Miller's LB Broth Base) and sterile water to a final volume of 1 L, autoclaved and kept at 4$^{o}$C prior to usage.

iii) Ampicillin (50 mg/ml)

The solution was prepared by adding 0.5 g ampicillin 10 ml sterile water, filter sterilized using a 0.22 µm filter and stored in 1 ml volumes at -20$^{o}$C.

iv) X-gal (250 mg/ml)

The solution was prepared by adding 1 g X-gal and 4 ml dimetilformamide, mixed and stored in alluminium foil wrapped eppendorf tubes at -20$^{o}$C.

## APPENDIX 3: Preparation of Total RNA and DNA Extraction Buffers

i) RNA extraction buffer

The buffer was prepared by adding 50 ml 1M Tris-HCl, pH 9.0, 75 ml 2M LiCl, 250 ml 20% SDS, 10 ml 0.5M EDTA, 20 ml 0.1M aurin tricarboxylic acid and sterile water to a final volume of 1 L. Four ml of β-mercaptoethanol was also added to the extraction buffer prior to usage.

ii) 2X CTAB Buffer (200 ml)

The buffer solution was prepared by adding 2 g CTAB, 10.0 ml 1M Tris-HCl, pH 8.0, 4 ml 0.5M EDTA, pH 8.0, 8.2 g NaCl and sterile water to a final volume of 200 ml. The solution was autoclaved and kept at room temperature.

iii) Modified CTAB Method

The modified buffer was prepared by adding 250 µl 0.5M ascorbic acid, 250.0 µl 0.4M DIECA and 0.5 g 2% PVP-40 (w/v) to 25 ml of 2X CTAB buffer prior to usage.

## APPENDIX 4: List of primers

i) Amplification of actin as internal control

| Primer name | Sequence (5`-3`) |
|---|---|
| Actin F | GTGCTAACGAATACAGTTCACG |
| Actin R | CCAGCAGATGTGGATTTCAAAG |

ii) Amplification of putative ethylene receptor genes and promoters

| Method | Gene ID | Primer name | Sequence (5`- 3`) |
|---|---|---|---|
| 3`RACE | *EgERS1* | 34176F3 | CCA AGG ATA TAG CGG CAG TGG ACT T |
| | *EgERS2* | 45619F2 | TTG GGC TTG CCA TTT GTA |
| | *EgERS3* | 46141F2 | CTC GGG CTC GCC ATT TGT AAG AG GTT T |
| | *EgETR3* | 3542F2 | CTG GCA TCC AAT GTC TGA GCT CCT TC |
| | *EgETR2* | 2412F2 | AGC CGC AGG TCT CTA TCG GAC ATC TT |
| | *EgETR4* | 44932F4 | GGG AGA AAT GTG TGC AGG TGG GAA T |
| | *EgETR5* | 53269UF1 | GCC ACT GAT CGT TGC ACT CAC TCC T |
| 5`RACE | *EgERS1* | 34176R2 | CTT CCT GTG TCC GTA TCA GGC CC |
| | *EgERS2* | 45619R1 | GCT GCC ACG GCC ACA ATA TTC TTC TC |
| | *EgERS3* | 46141UR2 | TTC TCG AGT GCA CGG TGA AAG TCC A |
| | *EgETR3* | 35423R1 | TGT ATG GCG GAT GCA CAG AAG AAG AGG |
| | *EgETR2* | 24121R2 | GTG AAC ACA TTG AGC AGG TGG GTC AG |
| | *EgETR4* | 44932R3 | TGA AGG GGA AGA GGT TGG AGC ATG T |
| | *EgETR5* | 53269R4 | CGG AGC ACG TGA CGA AGT AGA AGA GC |
| LD-PCR | *EgERS1* | ERS1F1 | GAA GAC GAC GAG GGT CTG CT |
| | | ERS1R1 | TTT TGG TAT CGA GCC AGT CCT T |
| | *EgERS2* | ERS2F1 | GGA GTC GAG TCT CCT TTT CGA |
| | | ERS2R1 | GCA CTG TCA TGC TCC ATT GTG |
| | *EgERS3* | ERS3F1 | GTA GAC GAG AAA GGC GAC AAA |
| | | ERS3R1 | CTG TAT TGC TCC ATC ATG CAG A |
| | *EgETR2* | ETR2F1 | CGC CAT TGA GAT CGG ACG T |
| | | ETR2R1 | GTA GCA GCA TTC GGA AAA GGA GT |
| | *EgETR3* | ETR3F1 | TGC ATC CGC CAT ACA GAT TGG ATT C |
| | | ETR3R1 | TAT CCC ACA CTG CAA GCA CCT ATC C |
| | *EgETR4* | ETR4F1 | CTT TCC TTA CAT TTCTTTGCTTGCGC |
| | | ETR4R1 | ATT TCT CCC ACG TATCTTCCTCAAC |

| | | | |
|---|---|---|---|
| | *EgETR5* | ETR5F1 | ATG CGG TCG TCG TGG TCT TTC |
| | | ETR5R1 | TTG GGT GAT CCT GTC CAA TTC GTC T |
| Splice variant | *Isotig35424* | 3542SVF1 | GTT TGG TAT CTT CCA AAG ACC AGA GTC TTG |
| | | ET3542R2 | GCG TAC TTC T CCA TTC CCC TCG AGC TC |
| | | 3542SVF2 | ACT GTG ATG CTG GGG TTT TGT TTC ACA |
| | *Isotig34177* | 34177SVF2 | AGG CCA AGT CAT AAA GCG AGA |
| | | SV34177R1 | TCA AGC CGA TAG AAG GAT GTT |
| Promoter | *EgERS1* | 34176P3F1 | CCA AGG GGT GTC TGG TGC ATT ATG |
| | | 34176P3R1 | TCT TCA CGC GAC TCG CTT GGA CCA |
| | *EgERS2* | 45619P3F | TCA TTT GGT AAT CCA ACC TTG CAG G |
| | | 45619P3R | CTT CAC AAG ACT CCT TTT GGA AGC A |
| | *EgERS3* | 46141P07F | TCG AGG TGA AGT ATA TCA TCT TGT TCG |
| | | 46141P07R | GGA TTA TTA TGG GTT GGT TTG GAG GT |
| | *EgETR2* | 24121P4F | TGT GCT CTG TAT GGA AAG TGC AAG |
| | | 24121P4R | GAG GAG GAG AAG AAC GGT GTA GGA A |
| | *EgETR3* | 35423P3RCF | GAT CAT CCT AAT CCT CTT TCG CCC C |
| | | 35423P3RCR | TTT GAG TCG ATC CAC CTC TTC TCT CG |
| | *EgETR4* | 44932P5F1 | CAC CAT GCG GTA CCG TAT GAT AGG A |
| | | 44932P5R1 | CAC CAT CTC GAC CGC GCA TCA A |
| | *EgETR5* | 53269PF1 | GTT CGG GCT TTG TTA GAT ATG CTC AGA |
| | | 53269R4 | CGG AGC ACG GTG ACG AAG TAG A |
| RT-PCR | *EgERS1* | ERS13F1 | GGT TTG AAT GTC TGG TTA GCA ATA CC |
| | | ERS13R1 | CCG ATA GAA GGA TGT TGA TTA TAC AGT GCT TC |
| | *EgERS2* | ERS23F1 | CGT TGT GGA AGA AAA TGA TAT ACA CAA TGG |
| | | ERS23R1 | AGG AAA CCC CTT CAC CAA TAG ATT TG |
| | *EgERS3* | ERS33F1 | CTG CAT GAT GGA GCA ATA CAG CTT GA |
| | | ERS33R1 | AAG TAA AGA TAT ACG ATT CAT GGC TCC TCC |
| | *EgETR2* | ETR23FI | CAT GCC CAA AAT GGA TGT TTT TGA GG |
| | | 2412R3 | TGA GGG GTA AAT GGA GAG TGG GAG CTA A |
| | *EgETR3* | ETR33F1 | GACAGCTCTCCATTATTTCTCCAGG |
| | | 3524R3 | GGTGATTATCTGCATTTGATCTTCAATATTAAGTG |
| | *EgETR4* | ETR43F2 | GGTTCACCATATACAGTAATACTTGCAC |
| | | 44932R5 | CTC ACT TCG ACC AGT TTG TGA CAA CCA C |
| | *EgETR5* | ETR53F1 | CAAATATCACCCAAGCAATGGAGAG |
| | | ETR53R1 | TGAAGGAAAATGGAGCTCATAGACCA |

iii) Amplification of lipase class 3 gene and promoter

| Primer name | Sequence (5`-3`) |
|---|---|
| LAS2 | CCAGGCAAGATTAGCCCAGATGCCCTGTAG |
| LF6 | GTGCTCCAATTTCCTTAGGGTTGCTG |
| LAS10 | TAGCCGTTCCAAATCCAGTTCTTA |
| LF3 | CTTGATGAGAGTGGT |
| LR1 | GCCGTTCCAAATCCAGTTC |
| LF1 | GAATGCTTTGGGTGACCTC |
| LAS11 | CCGGCCACCGAGCTGCTGCTCTCGACAAA |
| LAS14 | GGATAAAAAGGAAGAAGGGTTTTGACAGAG |
| PLF2 | CGGAAGCTTCTACATATGTGTGTTAGC |
| PLR1 | GGGTCTAGAATAAAAAAGGAAGAAGGG |

## APPENDIX 5: Preparation of solutions for Southern analysis

i) 10X MOPS Buffer

The solution was prepared by adding 41.9 g MOPS, 3M sodium acetate, pH 4.8, 20.0 ml 0.5M EDTA, pH 8.0 and 800 ml sterile water. The pH was adjusted to 7 with 10N NaOH to a final volume of 1 L. The solution was autoclaved and kept at room temperature.

ii) 20X SSC, pH 7.0 (1 liter)

The buffer was prepared by adding 175.3 g NaCl, 88.2 g trisodium citrate.$2H_2O$ and 800 ml sterile water. The pH was adjusted to 7.0 by adding 1M HCl. Sterile water was adjusted to a final volume of 1 L, autoclaved and kept at room temperature.

iii) 50X Denhardt's (100 ml)

The solution was prepared by adding 1 g Ficoll, 1 g polyvinyl pyrolidon, 1 g bovine albumin fraction V and sterile water to a final volume of 100 ml, filtered sterilized using a 0.22 µm filter and kept at -20$^{o}$C.

**APPENDIX 6:** Lipase class 3 paper as attached

DOI: 10.1007/s10535-015-0500-7 BIOLOGIA PLANTARUM 59 (2): 227-236, 2015

# Genomic structure and characterization of a lipase class 3 gene and promoter from oil palm

A.W. NURNIWALIS[1,3]*, R. ZUBAIDAH[1], A. SITI NOR AKMAR[2], H. ZULKIFLI[1], M.A. MOHAMAD ARIF[1], F.J. MASSAWE[3], K.L. CHAN[1], and G.K.A. PARVEEZ[1]

*Malaysian Palm Oil Board, No 6, Persiaran Institute, Bandar Baru Bangi, 43000 Kajang Selangor, Malaysia*[1]
*Faculty of Agriculture and Institute of Tropical Agriculture, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*[2]
*School of Biosciences, University of Nottingham Malaysia Campus, Jalan Broga, 43500 Semenyih, Selangor, Malaysia*[3]

## Abstract

Lipase class 3 is part of the triacylglycerol lipase family involved in lipid degradation, esterification, and transesterification processes in plants. In this study, a lipase class 3 gene and promoter from oil palm (*Elaeis guineensis* Jacq.) were isolated and characterized by Northern blot, Southern blot, oil palm genome sequence, and transient expression GUS assay. The full-length lipase class 3 (FLL1) deduced polypeptide encoded 483 amino acids and was identical to that deduced from lipase (EgLip1) cDNA (GI: 409994625). It contained the lipase consensus sequence, GxSxG motif, and a putative catalytic triad and had a 3-dimensional protein model similar to that of a lipase from *Giberella zeae* with a 50 % identity. The Northern blot and reverse transcription polymerase chain reaction (RT-PCR) show that *FLL1* was predominantly expressed in the mesocarp and the expression increased as fruits reached maturity. A lower expression was detected in germinated seedlings and especially in roots. The expression of *FLL1* was also enhanced in the mesocarp of cold treated fruits. A high oil accumulation in the mesocarp during fruit development makes this tissue a suitable target for a genetic modification, hence the isolation of the *FLL1* promoter. The transient expression of the β-glucuronidase (*GUS*) gene driven by the *FLL1* promoter detected the *GUS* expression in mesocarp slices, especially in vascular bundles. This suggests the potential role of using the promoter as tool to direct the expression of a transgene to the mesocarp of transgenic oil palm.

*Additional key words*: fatty acids, *GUS* gene, mesocarp, RACE, transient assay, triacylglycerides.

## Introduction

Oil palm, *Elaeis guineensis* Jacq. is the most productive oil crop in the world. The most treasured component of the oil palm is the fruit which produces two different types of oil: the palm oil derived from the mesocarp, and the palm kernel oil from the kernel. Both oils differ in their fatty acid composition, hence are used in different applications. Palm oil contains 50 % saturated, 39 % unsaturated and 11 % polyunsaturated fatty acids (Tang *et al.* 2000), and the oil is mainly used as food (Miskandar *et al.* 2011). The palm kernel oil contains mostly medium-chain fatty acids and it is normally used for industrial purposes (Hazimah *et al.* 2011).

Triacylglycerides (TAGs) make up the main component of palm oil and are found in abundance in mesocarp tissues especially in mature fruits. TAGs, however, are prone to hydrolysis, which causes the

*Submitted* 21 April 2014, *last revision* 28 October 2014, *accepted* 30 October 2014.
*Abbreviations*: CARE - *cis*-acting regulatory elements; FFA - free fatty acids; FLL1 - full-length lipase class 3; GUS - β-glucuronidase; LD-PCR - long distance polymerase chain reaction; LRE - light responsive elements; LTRE - low temperature responsive elements; ORF - open reading frame; RACE - rapid amplification of cDNA ends; REG - regulatory element groups; RT-PCR - reverse transcription polymerase chain reaction; SDS - sodium dodecyl sulphate; SSC - sodium chloride-sodium citrate; TAG - triacylglycerides; TSS - transcription start site; UTR - untranslated region; WAA - weeks after anthesis; Y-patch - pyrimidine patch.
*Acknowledgments*: We thank the Director General of MPOB for permission to publish this paper. We would also like to thank Pn Aminah Shuib for a technical help, Chan Pek Lan for a constructive assistance, Dr. Meilina Ong-Abdullah for the supply of oil palm male and female inflorescences total RNA, all personnel in the Gene Function Laboratory, MPOB, for their support and continuous help, and Dr Ariffin Darus, Dr. Ooi Siew Eng, and Dr. Abrizah Othman for technical editing the paper.
* Corresponding author; fax: (+60) 387694496, e-mail: nurni@mpob.gov.my

release of free fatty acids (FFA) from the glycerol backbone. FFA content and its oxidation products cause an undesirable flavour and odour of oil (Chong 2011), which affects oil quality. The increase in FFA content in palm oil is also induced by an endogenous lipase that is present in the mesocarp (Sambanthamurthi *et al.* 1995). Thus, one of the ways to minimize the losses of oil quality is to use oil palms with lower lipase activities. Those could be obtained from screening the oil palm germplasm (Wong *et al.* 2005, Ngando Ebongue *et al.* 2008).

The lipase activity assay in the mesocarp has been reported by several researchers. Henderson and Osborne (1991) and Sambanthamurthi *et al*. (1991) demonstrated that the oil palm mesocarp contains an active endogenous lipase. Ngando Ebongue *et al.* (2006) demonstrated that the highest activity was detected at 35 °C and at pH 9. On the other hand, Sambanthamurthi *et al.* (1991) and Cadena *et al.* (2013) showed that by inducing the lipase activity at a cold temperature of 5 °C, a maximum FFA content is detected. Sambanthamurthi *et al.* (1995) also observed that the lipase activity is induced by ripening oil palm fruits.

Two lipase genes coding for a putative lipase homolog (P87EST) and a lipase class 3 family protein (O65EST) were discovered from a 17-week-old mesocarp cDNA library (Nurniwalis 2006, Nurniwalis *et al.* 2008). Both genes show differential expression patterns: The *P87EST* expression is constitutively low but in all tested tissues, whereas the expression of *O65EST* is high in mesocarp tissues. The increase in the expression of *O65EST* follows the ripening pattern of the oil palm fruit which coincides with the oil synthesis period. In this paper, we report the isolation of full-length cDNA encoding lipase class 3 (*FLL1*), the prediction of protein structure, the characterization of *FLL1* expression, the isolation of its promoter, and the transient promoter assay.

## Materials and methods

**Plants and treatments:** Various tissues from commercial oil plant (*Elaeis guineensis* Jacq.) fruits were used. Mesocarp and kernel tissues were harvested from fresh fruit bunches at various weeks after anthesis (WAA) for RNA extraction. Spear leaves were obtained from unopened leaf fronds. Roots from 2-year-old and 1-week-old seedlings were also used. For a cold induced treatment, oil palm fruits were incubated at 7 °C for 5 h prior to RNA extraction.

**Total RNA and DNA extraction:** Total RNA from different oil palm tissues was extracted by a modified method of Prescott and Martin (1987). Frozen tissues (5 g) were ground and transferred to 15 $cm^3$ of an extraction buffer [50 mM Tris-HCl, pH 9.0, 150 mM LiCl, 5 % (m/v) SDS, 5 mM $Na_2EDTA$, 2 mM aurin tricarboxylic acid, and 0.4 % (v/v) β-mercaptoethanol] and followed by phenol:chloroform (50:50, v/v) and chloroform:isoamyl alcohol (24:1, v/v) extraction. The precipitation of RNA with 2 M LiCl was carried out at 4 $^oC$ overnight. The pellet obtained was washed with 2 M LiCl, dissolved in ice-cold sterile water, and used for Northern analyses and first strand cDNA synthesis. Genomic DNA from oil palm leaves was isolated as described previously (Nurniwalis 2006).

**Rapid amplification of cDNA ends (RACE) and identification of transcription start site (TSS):** The 5` region of the lipase class 3 gene (*FLL1*) was amplified and TSS identified using a *GeneRacer*™ kit (*Invitrogen*, Carlsbad, USA) which specifically targets the 5` capped mRNA structure to produce full-length cDNA. First strand cDNA was synthesized from the 17 WAA oil palm mesocarp total RNA. Touchdown PCR amplification was carried out in a 0.05 $cm^3$ reaction mixture containing a *RACE-Ready* cDNA template, a 1× *Expand HF* buffer with 15 mM $MgCl_2$ (*Roche*, Mannheim, Germany), 0.2 mM dATP, 0.2 mM dCTP, 0.2 mM dTTP, 0.2 mM dGTP (*Applied Biosystems*, Foster City, USA), a *GeneRacer* 5` primer, a 0.2 μM LAS2 primer, and a 0.1 U *Expand HF* PCR enzyme (*Roche*). The PCR was performed at 94 ºC (2 min), 94 °C (30 s), and 72 ºC (3 min) for 5 cycles; 94 °C (30 s) and 70 °C (3 min) for 5 cycles; 94 °C (30 s), 65 °C (30 s), and 68 °C (3 min) for 25 cycles, and a final extension at 68 °C for 10 min. The PCR products were purified using a *Qiaquick* gel extraction kit and cloned into a *pCR®II-TOPO®* vector (a *TOPO TA* cloning kit, *Invitrogen*). The plasmids were isolated using a *Qiagen* plasmid mini kit, digested with EcoRI, and sent for automated sequencing using *M13* reverse and forward universal primers.

**Long distance (LD)-PCR:** Full-length cDNA of the lipase class 3 gene was generated *via* LD-PCR using a *GeneRacer*™ kit (*Invitrogen*). Touchdown PCR amplification was carried out in 0.05 $cm^3$ of a reaction mixture containing 0.1 μg of a *RACE-Ready* cDNA template, a 1× *Advantage 2* PCR buffer (*Clontech*, Palo Alto, USA), a 1× *Advantage 2* polymerase mix, a 0.2 mM dNTP mix, 0.2 μM each of gene-specific primers LF6 and LAS10 with the following conditions: 94 °C (5 s) and 72 °C (3 min) for 5 cycles; 94 °C (5 s), 70 °C (10 s), and 72 °C (3 min) for 5 cycles; 94 °C (5 s), 68 °C (10 s), and 72 °C (3 min) for 22 cycles, and a final extension at 72 °C for 7 min.

**Sequence analyses:** Nucleotide and protein sequences were analysed using *Biology Workbench*, *v. 3.2* (http://workbench.sdsc.edu) and *Bioedit* tools (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Sequence similarity searches were performed *via NCBI BLAST* (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The gene structure was analyzed using *SPIDEY-NCBI*

(http://www.ncbi.nlm.nih.gov/SPIDEY). Protein motifs/ domains were analyzed using *CDD Search* (Marchler-Bauer *et al.* 2013), *Inter-Pro Scan* (http://www.ebi. ac.uk/Tools/pfa/iprscan/), and *MEME* (http://meme.sdsc. edu/meme/meme.html). Regulatory motifs in the promoter sequences were identified using *PlantCARE* (http:// bioinformatics.psb.ugent.be/webtools/plantcare/html/). The prediction of subcellular localization was performed using the *Target P program v. 1.1* (http://www.cbs. dtu.dk/services/TargetP). Hydropathy profiles were analyzed using *Kyte-Doolittle Plot*. The protein secondary structure prediction was carried out using the *PsiPred* program (http://bioinf.cs.ucl. ac.uk/psipred). The protein 3D model was predicted using *3D-JIGSAW* (http://bmm.cancerresearchuk.org/ ~3djigsaw/) and *I-Tasser* (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) and viewed using the *RasMol* latest version (http://www. umass.edu/microbio/rasmol/getras.htm). Additional analyses and predictions not stated above were also performed using various programs in *Biology Workbench*, *Bioedit*, and *Expasy toolkit* (http://web. expasy.org/tools).

**Expression analyses *via* Northern blot and RT-PCR:** For Northern blot, 7 µg of total RNA was denatured at 55 $^{o}$C for 15 min in a solution containing 78 % (v/v) deionized formamide, 16 % (v/v) deionized glyoxal and a 10 mM sodium phosphate buffer. The denatured RNA was separated on a 1.2 % (m/v) agarose gel for 3 h at 100 V in a 40 mM Tris-acetate buffer, pH 7.6, and transferred to a nylon charged membrane (*Hybond-N*$^{+}$, *Amersham*, Buckinghamshire, UK) overnight by capillary blotting (Sambrook and Russell 2001). The 3` untrans-lated region (UTR) was generated *via* PCR using gene-specific primers LF3 and LR1 and was used as probe for Northern hybridization. The probe was labelled with $\alpha^{-32}$P dCTP by a random primer reaction using a *Megaprime* DNA labelling system kit (*Amersham*) following the manufacturer's instructions. The membrane was hybridized in 5× saline-sodium citrate (SSC), 5× *Denhardt*, 0.5 % (m/v) sodium dodecylsulphate (SDS) with the probe at 65 °C overnight, and the next day washed with 2× SSC and 0.1 % SDS at 65 °C for 10 min , 1× SSC and 0.1 % SDS at 65 °C for 15 min, and 0.5× SSC and 0.1 % SDS at 65 °C for 20 min. Finally, the membrane was exposed to an X-ray film at -70 °C overnight. For RT-PCR, a single stranded cDNA was synthesized from 2 µg of DNAse-treated total RNA using a high-capacity cDNA reverse transcription kit (*Applied Biosystems*). PCR amplification was carried out in 0.025 cm$^3$ of a reaction mixture containing 0.1 µg of cDNA, a 1× *Advantage 2* PCR buffer, a 1× *Advantage 2* polymerase mix, a 1× dNTP mix, 0.2 µM LF1, and 0.2 µM LR1 primers with the following conditions: denaturation at 95 $^{o}$C (1 min); 95 °C (30 s) and 60 °C (1 min) for 30 cycles; and a final extension at 60 °C for 1 min. The amplification of the actin gene as internal control was performed in 0.025 cm$^3$ of a reaction mixture containing 0.1 µg of cDNA, a 1× *Advantage 2* PCR buffer, a 1× *Advantage 2* polymerase mix, a 1× dNTP mix, and 0.2 µM actin F and 0.2 µM actin R primers with the following conditions: denaturation at 95 °C (1 min); 95 °C (15 s), 55 °C (30 s), and 72 °C (2 min) for 30 cycles; and a final extension at 72 °C for 7 min.

**Genomic DNA amplification, Southern analyses and gene family search:** The amplification of the FLL1 genomic region was carried out in 0.05 cm$^3$ of a reaction mixture containing 0.1 µg of genomic DNA, a 1× *Advantage 2* PCR buffer, a 0.7 U *Expand HF* PCR system enzyme, a 1× dNTP mix, and 0.2 µM LF6 and 0.2 µM LAS10 primers with the following two-step PCR conditions: an initial denaturation at 95 °C (1 min); denaturation at 95 °C (30 s), simultaneous annealing and extension at 65 °C (1 min) for 30 cycles, and final annealing and extension at 65 $^{o}$C for 3 min. A total of 15 µg of genomic DNA was digested individually with BamHI, EcoRI, SpeI, and XbaI (*Fermentas*, Germany), respectively. Southern analyses were carried out following a standard procedure described by Sambrook and Russell (2001). The membrane was subjected to hybridization with a 3' UTR cDNA probe in high stringency conditions (5× SSC, 5× Denhardt, 0.5 % SDS at 65 °C overnight) and then washed with 2× SSC and 0.1 % SDS at 65 °C for 10 min, 1× SSC and 0.1 % SDS at 65 °C for 15 min, and 0.5× SSC and 0.1 % SDS at 65 °C for 20 min. To search for other putative members of the lipase class 3 family, a *Hidden Markov* model (HMM) profile was build using the amino acid sequence of the conserved lipase class 3 domain (*Pfam ID*: PF10764) downloaded from *Pfam* (http://pfam. xfam.org) and searched against the translated transcriptome and gene model of oil palm (Singh *et al.* 2013) using the *HMMER3* (Eddy 2011) hmm search programme.

**Promoter isolation:** The isolation of the *FLL1* promoter was carried out using universal *Genome Walker* and *Advantage 2* PCR kits (*Clontech*). Four *Genome Walker* libraries were constructed *via* the digestion of genomic DNA with DraI, EcoRV, and PvuII dan StuI prior to ligation with the *GenomeWalker* adaptor. Primary PCR amplifications were performed using all four *Genome Walker* libraries as the DNA template. Each reaction mixture contained 100 ng of *Genome Walker* library, a 1× *Advantage 2* PCR buffer, a 0.2 mM dNTP mix, 0.2 µM each of primers AP1 and LAS11, a 1× *Advantage 2* polymerase mix, in a two-step cycle parameters: 94 °C (25 s) and 72 °C (3 min) for 7 cycles; 94 °C (25 s) and 67 °C (3 min) for 32 cycles, and a final extension at 67 °C for 7 min. The primary PCR product was diluted 50× and used as DNA template for secondary PCR. The secondary PCR mixture was similar to that of the primary PCR except the primers used were primers AP2 and LAS14. The PCR reaction was performed in 5 cycles of 94 °C (25 s) and 72 °C (3 min), followed by 20 cycles of 94 °C (25 s) and 67 °C (3 min), and a final extension at 67 °C for 7 min.

**Promoter-vector construct and transient assay analysis:** Plasmid *FLL1/GUS* was constructed by replacing the CaMV 35S promoter contained in pBI221 with the HindIII-Xba1 flanked *FLL1* promoter sequence from the position -664 to 83 bp. The amplification of the HindIII-Xba1 flanked *FLL1* promoter region was performed in 0.05 $cm^3$ of a reaction mixture containing 25 ng of plasmid DNA, a 1× *Expand HF* buffer with 15 mM $MgCl_2$ (*Roche*), a 0.2 mM dNTP mix (*Clontech*), 0.1 µM each of gene-specific primers PLF2 and PLR1, and a 0.1 U *HF Enzyme* mix (*Roche*) under the following conditions: 94 °C (3 min), 94 °C (1 min), 57 °C (1 min), and 72 °C (90 s) for 20 cycles, and a final extension at 72 °C for 10 min. The cloning procedure was confirmed *via* restriction enzyme analysis and the insert was verified by sequencing. The preparation of target materials for transformation, the bombardment parameters, and the GUS histochemical assay were carried out as described by Zubaidah and Siti Nor Akmar (2003).

## Results and discussion

The primers used for PCR amplifications are listed in Table 1 Suppl. A 1448 bp cDNA fragment designated as 5`4LAS2 was amplified to represent the 5` region of the FLL1 cDNA *via* 5` RACE. The complete sequence of the *GeneRacer*™ RNA oligo primer was detected at the 5` region confirming that the full-length message from the 5` end was successfully obtained and that the PCR product was very specific to the respective gene (Invitrogen Instruction Manual of *GenerRacer*™ kit, 2004). A translation start codon (ATG) was detected at the 5` end and the sequence adjoining the start codon is adequate to the Kozak (1999) initiation motif (ANNATGG). The O65EST cDNA isolated from the 17-WAA mesocarp cDNA library contains a 350 bp cDNA fragment including a 22 bp poly $(A)^+$ tail (Nurniwalis 2006, Nurniwalis *et al.* 2008). By assembling the 5` RACE and O65EST cDNA sequences, an overlap of 112 nucleotides or 37 amino acid residues with a 100 % identity at the nucleotide and protein levels were observed (Fig. 1 Suppl.). The reconstitution of both the sequences resulted in the amplification of the full-length FLL1 cDNA *via* LD-PCR. The full-length cDNA was 1 721 bp long and contained an open reading frame (ORF) of 1 452 bp (Fig. 1 Suppl.). The ORF was flanked by 106 bp of 5` UTR followed by 141 bp of 3` UTR and a 22 bp poly $(A)^+$ tail. At the 3` UTR, two putative polyadenylation sites were found to differ slightly from the plant consensus sequence, *i.e.*, AATAAA (Joshi 1987a). Nevertheless, it contained four to five out of six base matches of the conserved sequence found in most plants (Hunt 1994). One polyadenylation site was located 20 bp upstream the poly $(A)^+$ tail. This matched an expected regulatory region (11 - 24 bp) for 3` end pre-mRNA cleavage and polyadenylation, but often, an accurate position and a detailed mechanism of 3` mRNA processing remains unclear (Jin and Bian 2004).

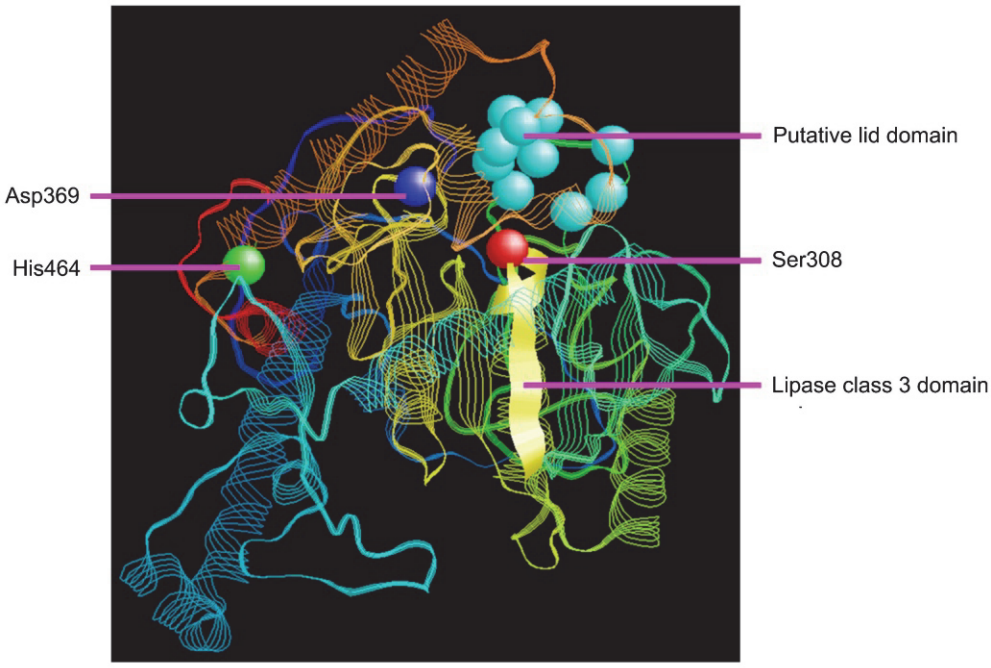



Fig. 1. A three-dimensional best model for FLL1 predicted protein selected based on a C-score using *I-Tasser*. The putative lipase catalytic triad Ser308, Asp369, and Hid464 (represented as the *balls*), the lipase class 3 conserved region (represented as the *ribbon*) and the putative lid domain (represented as the *balls*) are indicated. The model is presented using *RasMol v. 2.6*.

The nucleotide sequence alignment between *FLL1* and the oil palm transcriptome data from 15 WAA mesocarp tissues (Singh *et al.* 2013) showed a 100 % identity with an e-value of 0.0. The database searches for homologies against the nucleotide and protein sequences showed that *FLL1* is 99 and 100 %, respectively, identical to *EgLIP1*, a lipase class 3 cDNA (accession No. AFV50601.1) from oil palm at the nucleotide and protein level (97 % query coverage). The next closest and the following matches to the FLL1 deduced amino acid sequence share a much lower homology with a putative lipase (identity of 51 % and an e-value of $6e^{-144}$) from rice (acc. No. NP_001054678.2) followed by lipase (44 % identity and an e-value of $3e^{-128}$) and triacylglycerol lipase (43 % identity and an e-value of $2e^{-124}$) from castor bean (acc. Nos. AAV66577.1 and XP_002533321.1). The conserved domain and homology search showed that FLL1 contained the lipase consensus sequence [LIV]X[LIVAFY][LIAMVST]G[HYWV]SXG[GSTAC] that encoded the highly conserved lipase class 3 domain (Pfam ID: PF10764). The region of sequence similarities detected all three putative amino acid residues that form the lipase catalytic triad. The serine (S) residue was detected at position 308 and the aspartic acid (D) residue at position 368. Histidine (H), the third residue to complete the catalytic triad, had more diverse flanking sequences, but the exact H position was predicted at location 464 based on consensus sequence alignments (Fig. 2 Suppl.). In addition, putative active site lids at positions 223, 224, 226, and 229-236 as well as nucleophilic elbow at positions 306 - 310 were detected in FLL1 (Fig. 1 Suppl.).

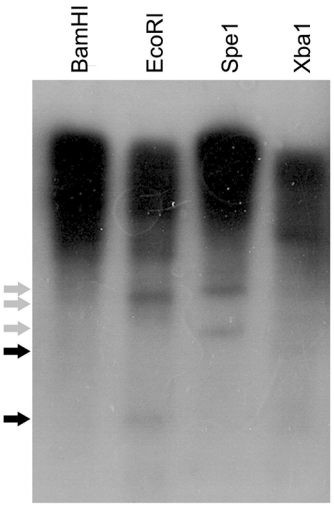



Fig. 2. The southern analysis of oil palm genomic DNA. Blots were hybridized with a gene-specific sequence based on the 3`UTR of the *FLL1* gene. Strong and faint signals are shown by the *grey* and *black arrows*.

The deduced amino acid sequence of FLLI was predicted to have a secondary structure that consists of 12 β-sheets, 14 α-helixes, and 25 coils. FLL1 had more hydrophobic residues (47.4 %), and a Kyte-Doolittle hydropathy profile showed the predicted protein was predominantly hydrophobic. The three dimensional protein prediction using *3D-JIGSAW* detected three family domains (pb073089, pf01764, and pb108722) in FLL1 but only pf01764 (the lipase class 3 domain) was homologous to FLLI. FLL1 had an identity of 50 % and an e-value of $1e^{-55}$ to 3ngmA, a crystal structure of lipase from *Gibberella zeae* from location 64 to 221. The structural protein prediction of FLL1 using *I-Tasser* generated five 3-D output models, ranking model 1 as best predicted (Zhang 2008) with a C-score more than -1.5 to indicate correct protein folding (Roy *et al.* 2010). Fig. 1 represents the best predicted model 1 with a C-score value of 2.83 based on the highest structural alignment to a lipase protein from *Gibberella zeae* (PDB hit:3ngmA). The predicted EC number for FLL1 is 3.1.1.3, which represents a triacylglycerol lipase. Based on a gene ontology (GO), the *FLL1* gene has a molecular function that is involved in triglyceride lipase activity (GO:0004806). No signal peptide or transmembrane domains were identified in FLL1 using various subcellular localization prediction searches, which would suggest that FLL1 is cytoplasmic.

Southern blot was used to determine the presence of lipase gene(s) in the oil palm genome. The oil palm genomic DNA was digested with BamHI, EcoRI, XbaI, and SpeI, respectively, followed by hybridization using the 3` UTR *FLL1* cDNA as probe. At high stringency conditions, more than two bands were detected in all the restriction enzyme-digested oil palm genomic DNAs (Fig. 2). This result indicates that *FLL1* was not a single copy gene and that it belonged to a multi gene family. Thus, the search of other putative members of this family was carried out using the *BLAST* program and a Pfam ID: PF10764 as query sequence on the oil palm transcriptome and genome data (Singh *et al.* 2013). To date, we have identified 27 putative genes encoding the lipase class 3 domains (data not shown) with protein identities ranging from 26 to 83 % to *FLL1*.

Northern blot and semi-quantitative RT-PCR were carried out using RNA from various oil palm tissues, such as the mesocarp at different developmental stages, namely 8, 10, 12, 15, 17, and 20 WAA, the kernel, spear leaves, roots, and germinated seedlings. Northern blot was performed using a pair of gene-specific primers that correspond to the 3` UTR of the *FLL1* sequence where it hybridized to a transcript with the size of about ~1.7 kb (Fig. 3). The size of the hybridized transcript was similar to that of the full-length *FLL1* cDNA which is approximately 1 721 bp long (including the poly $(A)^+$ tail). The expression of the mRNA transcript in the mesocarp tissue was found to gradually increase from the early fruit developmental stages up to when the fruit was fully ripened. This pattern of expression could possibly correlate closely with the pattern of oil synthesis in the mesocarp, and this result is in agreement to that found from the dot blot analysis (Nurniwalis 2006, Nurniwalis *et al.* 2008). The expression of the 3` UTR probe in the other tested tissues, however, appeared to be indistinct.

Hence, RT-PCR was conducted to examine the expression of *FLL1* in various oil palm tissues. For RT-PCR, two primers specific to the corresponding sequence of the lipase class 3 EST sequence, O65EST cDNA (ORF + 3` UTR), and *actin* as internal control were used. PCR products with the size ~400 bp were amplified and the size corresponded well to O65EST cDNA. The PCR products were amplified from all the mesocarp tissues from young fruits until the ripening period and the band intensities appeared to be strong in these tissues in comparison to the *actin* gene (Fig. 3). As RT-PCR is a more sensitive technique for mRNA detection and quantitation, trace amounts of PCR products were also detected in the germinated seedlings and even lower in the roots. No PCR product was detected in the rest of the tested tissues (the kernel, spear leaves, male and female flowers). RT-PCR was also performed to determine the expression pattern of *FLL1* in

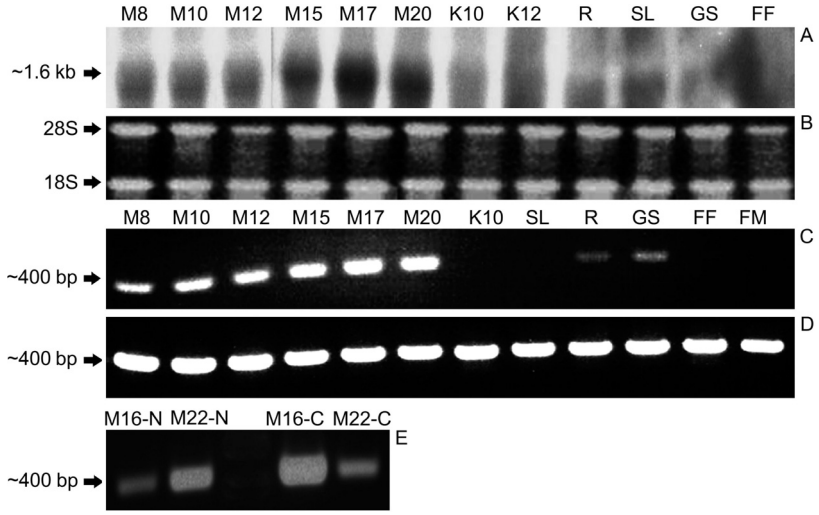



Fig. 3. Expression patterns of *FLL1* gene in various oil palm tissues as shown by the Northern analysis and RT-PCR. The blot was probed with a gene-specific sequence based on the 3`UTR (*A*); an ethidium bromide stained gel that shows approximate equal loading total RNA (*B*); the RT-PCR amplification of the 3`region of *FLL1* using gene-specific sequences (*C*); the amplification of the *actin* control gene (*D*); and the comparison of the *FLL1* cDNA in normal and cold treated fruits (*E*). M - mesocarp, K - kernel, R - roots, SL - spear leaves, GS - germinated seedlings, FF - female inflorescence, FM - male inflorescence, N - normal, and C - cold treated fruits. Numbers in the mesocarp and kernel - developmental stages at WAA.

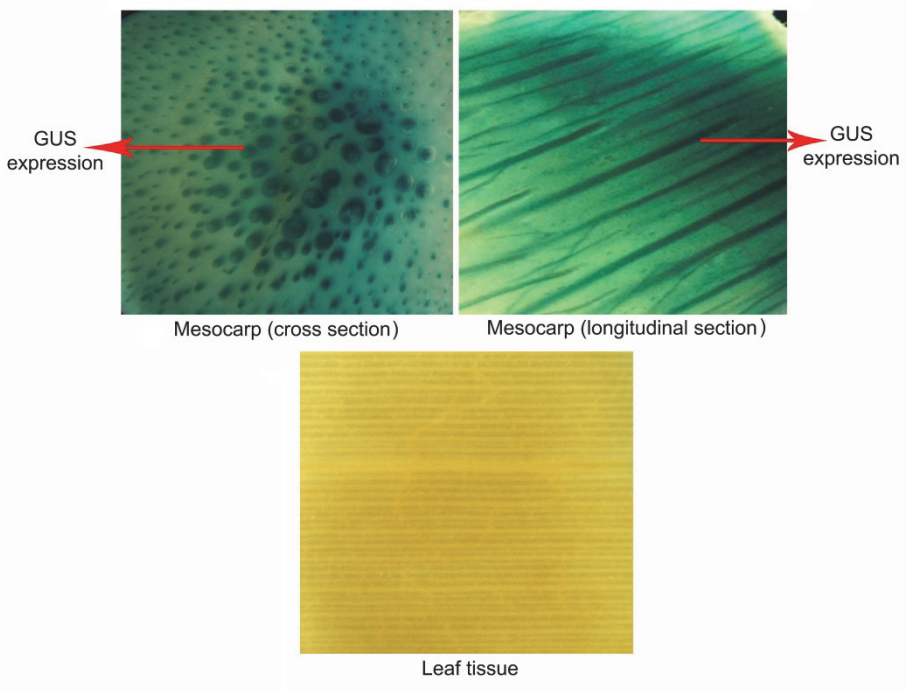



Fig. 4. An *FLL1pro:GUS* expression profile in the mesocarp and leaf tissues of oil palm.

the mesocarp of cold induced oil palm fruits. Comparison was carried out in normal and cold treated fruits at two stages of fruit development, *i.e*., 16 and 22 WAA. The results show that both the normal and cold treated fruits had a similar expression pattern. However, the expression of the *FLL1* gene was higher in the cold induced fruits than in the normal fruits (Fig. 3).

The amplification of the *FLL1* genomic sequence yielded a 5.5 kb fragment. The genomic organization and presence of *FLL1* in the oil palm genome were confirmed through a sequence alignment between the amplified *FLL1* genomic sequence and the scaffold p5-sc00064 of the fifth genome build [p5-build] (Singh *et al.* 2013). The result shows a 100 % identity at the nucleotide level with an e-value of 0.0. An exonerated search with a 60 % self-score threshold revealed *FLL1* at chromosome 3 in the EG5 build. The comparison of the *FLL1* genomic sequence and oil palm microsatellite genomic DNA by Morcillo *et al.* 2013 (acc. No. HE661587.1) revealed a 97 % identity suggesting that both regions are identical. The amplification of the *FLL1* genomic sequence and an exonerated search revealed the presence of four introns located in between five exons in *FLL1*. The size of the introns were 1 424, 555, 1 626, and 179 bp and the five exons 382, 224, 34, 620, and 251 bp, respectively. The intron-exon boundaries in *FLL1* conformed to the universal GT-AG rule for introns starting with a GT dinucleotide and ending with an AG dinucleotide (Breathnach and Chambon 1981).

The upstream genomic region of *FLL1* was further amplified to isolate its corresponding promoter using the Genome Walker approach. The PCR amplification resulted in the amplification of a 756 bp fragment. This genomic sequence contained 671 bp of promoter sequence and 85 bp of 5' UTR. An adenine residue located furthest at the 5' terminal of *FLL1* was selected as TSS. The adenine residue is most likely the TSS as the pyrimidine sequences flanking the adenine residue are often the preferred TSS motif in plants {C/TAC/T} (Joshi 1987b). The putative TSS motif also matches most of the TSS motifs in highly expressed genes in plants (Sawant *et al.* 1999).

The distribution of sequences in the FLL1 promoter is divided into three main groups which contain the TATA-box, pyrimidine patch (Y-patch), and regulatory element group (REG) (Yamamoto *et al.* 2007). The putative TATA-box, TATATATTA, was present 37 bp upstream of the TSS, consistent with the distance of 32 ± 7 bp (Joshi 1987b). The Y-patch motif was located within 100 bp upstream of the TSS and downstream of the TATA-box. The role of the Y-patch motif, a plant-specific core element, is unknown but the local distribution of short sequence analysis showed that it maked up one of the general/important components in the core promoters of dicots and monocots (Yamamoto *et al.* 2007). The REG group contains *cis*-acting regulatory elements (CARE) that correspond to known transcriptional regulatory sequences. In the *FLL1* putative promoter, a number of putative CAREs were identified, which included those that respond to environment signals as well as those required to direct specific expressions to specific tissues (Table 2 Suppl.).

A schematic diagram of the *FLL1* promoter cloned into the pBI221 transformation vector carrying *GUS* as reporter gene that replaced the 35S promoter is shown in Fig. 3 Suppl. The activity of the *FLL1* promoter in oil palm mesocarp slices was analysed using a transient *GUS* expression analysis. The transient *GUS* expression was driven by the *FLL1* promoter fused to the *GUS* reporter gene construct using optimized bombardment parameters (Zubaidah and Siti Nor Akmar 2003). The expression of the *GUS* gene was detected in the mesocarp slices at 12 WAA but not in the leaves (control tissue) (Fig. 4). These results support the expression analysis of *FLL1 via* Northern blot and RT-PCR. The lipase gene is expressed throughout the fruit developmental stages. Due to difficulty in handling the ripe fruit for a transient expression assay because of a high oil content, the 12 WAA stage was chosen as the target tissue to represent the expression of the promoter. Longitudinal and cross sections of the mesocarp tissues demonstrated that the *GUS* expression was targeted to vascular bundles. This result is similar to that of the *MT3A* promoter of metallothionein gene which directs a high expression in mesocarp tissues (Siti Nor Akmar and Zubaidah 2008, Zubaidah and Siti Nor Akmar 2010). Phloroglucinol-stained mesocarp sections for lignin detection also showed a similar result (Singh *et al.* 2013). Lignins are normally associated with the vascular bundle in plants. They are hydrophobic in nature which allows water transportation throughout the plant (Xu *et al.* 2009). In the mesocarp tissues of oil palm (Sambantamurthi *et al.* 1995) and olive (Panzanaro *et al.* 2010), lipase activities have been demonstrated to be associated with oil bodies and require a hydrophobic condition to function. Similar targeted expressions of the *GUS* gene and lignin to vascular bundles in mesocarp slices are unclear, but this raises the possibility that they could share a common trait or function that is still unknown.

We describe here the isolation and characterization of the full-length lipase class 3 gene (*FLL1*) and its corresponding promoter from oil palm. FLL1 contains the lipase consensus sequence ILVTGHSLGG which includes the active serine residue ($S_{308}$) that forms the esterase box GxSxG motif. Together with the aspartic acid ($D_{368}$) and histidine ($H_{464}$) residues, they form the putative Ser-Asp-His catalytic triad that is essential for esterase and lipase activities to hydrolyze/de-esterify fatty acids from complex lipids (Li *et al.* 2012). This feature is not only common in lipases from plants but in animal, fungi, and bacteria as well (Patil *et al.* 2011). FLL1 is hydrophobic and the hydrophobic residues in the putative lid domain is likely important for the catalytic activity of FLL1.

FLL1 was 100 % identical to the mesocarp trans-criptome and p5 build genome data at the nucleotide level. This result is in line to that of oil palm genome data where most of the oil palm genes involved in oil quality

were derived from the paternal pisifera (Singh *et al.* 2013). *FLL1* was also 99 % identical to *EgLIP1* (acc. No. AFV50601.1) at the nucleotide level and was 39 nucleotides longer at the 5` UTR possibly as result of a better cDNA amplification of only the full-length transcripts minus the truncated messages from the amplification process (Invitrogen Instruction Manual of GenerRacer™ Kit 2004). The sequence alignment of *EgLIP1* to the p5 build genome data shows a 99 % identity at the nucleotide level with an E-value of 0.0. It was also located on the same p5-sc00064 scaffold of the p5-build and chromosome 3, which strongly suggests that both the genes are identical.

The recent release of the oil palm genome sequence provides rapid means to identify genes of interest from oil palm (Singh *et al.* 2013). The preliminary search identified putative members of the lipase class 3 family in oil palm with identities ranging from 26 to 83 % to *FLL1*. This is possibly due to the nature of the lipase class 3 family where polymorphisms were observed within the gene members. Thirty-eight putative lipase class 3 proteins were identified in *Arabidopsis* containing the conserved lipase class 3 domain (PF01764) but there is quite a divergence in terms of their sequence homologies, gene structures, and expression patterns (Li *et al.* 2012). This result suggests that presence of other members of the lipase class 3 family in oil palm is possible and they may play specific or constitutive roles in plant growth and development especially with regards to TAGs.

The Northern and RT-PCR analyses indicated that *FLL1* was highly expressed in the mesocarp tissues of oil palm. The *FLL1* transcripts were detectable in all the tested fruit developmental stages and the expression was stronger as the fruits reached maturity. It was possibly induced in the cold treated fruits in agreement to what was reported by Sambanthamurthi *et al*. (1995). The activation of lipase at a low temperature in oil palm fruit bunches may provide an alternative to FFA production *via* fat splitting for use in the oleochemical industry (Sambanthamurthi and Kushairi 2002). FLL1 is also present in germinated seedlings as well as in roots but at very much lower amounts just as reported by Morcillo *et al.* (2013). Oo and Stumpf (1983) have detected a lipase activity in oil palm germinated seedlings. Because lipase is the first enzyme involved in TAG breakdown, the low expression of *FLL1* in the germinated seedlings is logical for early seedling growth although there is no evidence to prove that lipase class 3 genes are responsible for this. Apart from the well-studied role of lipase in germinated seedlings (Murphy 1993), the role of lipase in fruits is still unclear. It has been suggested that the lipase in the oil palm mesocarp serves to increase the palatability of fats (Morcillo *et al.* 2013). Another possible role of mesocarp lipase in trans-esterification during lipid synthesis has also been raised (Sambantahmurthi *et al*. 1995). The oil palm mesocarp tissue is also a suitable target for introducing novel characteristics for the production of high and value-added products and improving oil quality. Thus, from the molecular/genetic engineering aspect, the high expression of the *FLL1* gene encoding the lipase class 3 prompted the isolation of its corresponding promoter. This promoter has the potential to direct the expression of a desired gene carrying the trait of interest to the mesocarp for production of novel product(s) in oil palm.

The *in silico* analysis using *PlantCare* shows that the *FLL1* promoter contained several different *cis*-acting regulatory motifs, which suggest a complex regulatory system to regulate the *FLL1* expression. Majority of the CAREs present in the *FLL1* promoter are associated with light responsive elements (LREs). The promoter of a fruit-specific gene, *GalUR* from strawberry whose activity is specifically targeted to the fruit, is dependent on irradiance. It contains a G-box motif that is essential for a fruit specific expression (Agius *et al.* 2005). In addition, an I-box motif, another LRE, is also present in the *GalUR* promoter from strawberry as well as in other fruit and in tissue-specific promoters in melon (Yamagata *et al.* 2002) and oil palm (Siti Nor Akmar and Zubaidah 2008). To date, the role of the LRE in the *FLL1* promoter in association with irradiance is uncertain, but the detection of multiple LREs suggests that the promoter activity is most likely enhanced by irradiance.

Two low temperature responsive element (LTRE) sequences associated with a low temperature responsiveness were identified in the *FLL1* promoter in the forward and reverse orientations. In *Arabidopsis*, the LTRE elements in the promoter region of the *COR15A* gene demonstrated an involvement in cold, drought, and ABA-regulated gene expressions (Baker *et al.* 1994). Transgenic potato containing the LTRE element exhibited an extensive cold inducibility in a tuber-specific manner (Liu *et al.* 1994). Integration of LTRE and LRE signals in *Arabidopsis* is required for a full response towards a cold acclimation (Catalá *et al.* 2011). The LTR elements in the *FLL1* promoter suggests that plant lipases, especially *FLL1* encoding the lipase class 3, play a role in TAG hydrolysis possibly through the activation of its corresponding gene at a low temperature. This is somewhat consistent to that of the RT-PCR results as well as the findings reported by Sambanthamurthi *et al.* (1995) where lipases in oil palm are activated at a cold temperature.

The strength and activity of the promoter determines its effectiveness in controlling and regulating gene expression. The transient assay system using the *GUS* reporter gene is a common method in assessing the trait of a promoter. Testing the *FLL1* promoter activity *via* a transient *GUS* expression using biolistic methods on oil palm mesocarp slices is preferred in comparison to generating oil palm transgenics. The assay saves a lot of time and is also the preferred choice to test other oil palm promoters (Zubaidah and Siti Nor Akmar 2010, Masura *et al.* 2011). With the high expression and specificity of the *FLL1* promoter observed in the mesocarp tissues, it has a high potential to be used as biotechnology tool to genetically engineer oil palm for production of novel oils and products.

## References


Agius, F., Amaya, I., Botella, M.A., Valpuesta, V.: Functional analysis of homologous and heterologous promoters in strawberry fruits using transient expression. - J. exp. Bot. **56**: 37-46, 2005.

Baker, S.S., Wilhelm, K.S., Thomashow, M.F.: The 5'-region of *Arabidopsis thaliana* cor15a has *cis*-acting elements that confer cold-, drought- and ABA-regulated gene expression. - Plant mol. Biol. **24**: 701-713, 1994.

Breathnach, R., Chambon, P.: Organization and expression of eucaryotic split genes coding for proteins. - Annu. Rev. Biochem. **50**: 349-83, 1981.

Cadena, T., Prada, F., Perea, A., Romero, H.M.: Lipase activity, mesocarp oil content, and iodine value in oil palm fruits of *Elaeis guineensis, Elaeis oleifera*, and the interspecific hybrid O×G (*E. oleifera* × *E. guineensis*). - J. Sci. Food Agr. **93**: 674-680, 2013.

Catalá, R., Medina, J., Salinas, J.: Integration of low temperature and light signalling during cold acclimation response in *Arabidopsis*. - Proc. nat. Acad. Sci. USA **108**: 16475-16480, 2011.

Chong, C.L.: Storage, handling and new products. - In: Mohd Basri, W., Choo, Y.M., Chan, K.W. (ed.): Further Advances in Oil Palm Research (2000-2010), Vol. II. Pp. 645-672. Malaysian Palm Oil Board, Bangi 2011.

Eddy, S.R.: Accelerated profile HMM searches. - PLoS Comput. Biol. **7:** e1002195, 2011.

Hazimah, A.H., Zainab, I., Rosnah, I., Loh, S.K.: From palm oil to non-food products: green chemistry. - In: Mohd Basri, W., Choo, Y.M., Chan, K.W. (ed.). Further Advances in Oil Palm Research (2000-2010). Vol. II. Pp. 791-895. Malaysian Palm Oil Board, Bangi 2011.

Henderson, J., Osborne, D.J.: Lipase activity in ripening and mature fruit of the oil palm - stability *in vivo* and *in vitro*. - Phytochemistry **30**: 1073-1078, 1991.

Hunt, A.G.: Messenger RNA 3' end formation in plants. - Annu. Rev. Plant Physiol. Plant mol. Biol. **45**: 47-60, 1996.

Jin, Y., Bian, T.: Nontemplated nucleotide addition prior to polyadenylation: a comparison of *Arabidopsis* cDNA and genomic sequences. - RNA **10**: 1695-1697, 2004.

Joshi, P.C.: Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. - Nucl. Acids Res**.** **15**: 9627-9640, 1987a.

Joshi, P.C.: An inspection of the domain between putative TATA box and translation start site in 79 plant genes. - Nucl. Acids. Res. **15**: 6643-6653, 1987b.

Kozak, M.: Initiation of translation in prokaryotes and eukaryotes. - Gene **234**: 187-208, 1999.

Li, W., Ling, H., Zhang, F., Yao, H., Sun, X., Tang, K.: Analysis of *Arabidopsis* genes encoding putative class III lipases. - J. Plant Biochem. Biol. **21**: 261-267, 2012.

Liu, Z-B., Ulmasov, T., Shi, X., Hagen, G., Guilfoyle, T.J.: Soybean GH3 promoter contains multiple auxin-inducible elements. - Plant Cell **6**: 645-657, 1994.

Marchler-Bauer, A., Zheng, C., Chitsaz, F., Derbyshire, M.K., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Lu, S., Marchler, G.H., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Bryant, S.H.: CDD: conserved domains and protein three-dimensional structure. - Nucl. Acids Res. **41**(Suppl.): D348-D352, 2013.

Masura, S.S., Parveez, G.K.A., Low, L.E.T.: Isolation and characterization of an oil palm constitutive promoter derived from a translationally control tumor protein (TCTP) gene. - Plant Physiol. Biochem. **49**: 701-708, 2011.

Miskandar, M.S., Karimah, A., Noor Lida, M.D., Wan Rosnani, A.I., Salmi Yati, S., Sivaruby, K., Zaida, Z.: Palm oil and palm kernel oil for low trans food products. - In: Mohd Basri, W., Choo, Y.M., Chan, K.W. (ed.). Further Advances in Oil Palm Research (2000-2010). Vol. II. Pp. 685-720. Malaysian Palm Oil Board, Bangi 2011.

Morcillo, F., Cros, D., Billotte, N., Ngando-Ebongue, G.F., Domonhédo, H., Pizot, M., Cuéllar,T., Espéout, S., Dhouib, R., Bourgis, F., Claverol, S., Tranbarger ,T.J., Nouy, B., Arondel, V.: Improving palm oil quality through identification and mapping of the lipase gene causing oil deterioration. - Nat. Commun. **4**: 2160, 2013.

Murphy, D.J.: Plant lipids: their metabolism, function and utilization. - In: Lea, P.J., Leegood, R.C. (ed.): Plant Biochemistry and Molecular Biology. Pp 113-128. John Wiley and Sons, New York 1993.

Ngando Ebongue, G.F., Dhouib, R., Carriere, F., Amvam Zollo, P.H., Arondel, V.: Assaying lipase activity from oil palm fruit (*Elaeis guineensis Jacq*.) mesocarp. - Plant Physiol. Biochem. **44**: 611-617, 2006.

Ngando Ebongue, G.F., Nouy, B., Zok, S., Carrière, F., Zollo, P.H., Arondel, V.: Identification of oil palm breeding lines producing oils with low acid values. - Eur. J. Lipid Sci. Technol. **110**: 506-509, 2008.

Nurniwalis, A.W.: Isolation of ethylene response sensor gene and generation of expressed sequence tags from the oil palm (*Elaies guineensis Jacq*.) mesocarp. - MSc. Thesis, Universiti Putra Malaysia, Serdang 2006.

Nurniwalis, A.W., Suhaimi, N., Siti Nor Akmar, A., Aminah, S., Mohamad Arif, M.A.: Gene discovery *via* expressed sequence tags from the oil palm (*Elaeis guineensis Jacq*.) mesocarp. - J Oil Palm Res. Special Issue **2**: 87-96, 2008.

Oo, K.C., Stumpf, P.K.: Some enzymatic activity in the germinating oil palm (*Elaeis guineensis*) seedling. - Plant Physiol. **73**: 1028-1032, 1983.

Panzanaro, S., Nutricati, E., Miceli, A., De Bellis, L.: Biochemical characterization of a lipase from olive fruit (*Olea europaea L.*). - Plant Physiol. Biochem. **48**: 741-745, 2010.

Patil, K.J., Manojkumar, Z.C., Raghunath, T.M.: Lipase biodiversity. - Indian J. Sci. Technol. **4**: 971-982, 2011.

Prescott, A., Martin, C.: A rapid method for the quantitative assessment of levels of specific mRNAs in plant. - Plant mol. Biol. Rep. **4**: 219-224, 1987.

Roy, A., Kucukural, A., Zhang, Y.: I-TASSER: a unified platform for automated protein structure and function prediction. - Nat. Prot. **4**: 725-738, 2010.

Sambanthamurthi, R., Kushairi, A.D.: Selection for lipase activity in the oil palm. - MPOB TT No. 141. P. 4. Malaysian Palm Oil Board, Kajang 2002.

Sambanthamurthi, R., Chong, C.L., Oo, K.C., Yeo, K.H., Rajan, P.: Chilling-induced lipid hydrolysis in the oil palm (*Elaeis guineensis*) mesocarp. - J. exp. Bot. **42**: 1199-1205, 1991.

Sambanthamurthi, R., Oo, K.C., Parman, S.H.: Factors affecting lipase activity in *Elaeis guineensis* mesocarp. - Plant Physiol. Biochem. **33**: 353-359, 1995.

Sambrook, J., Russell, D.W.: Molecular Cloning: A Laboratory Manual. 3rd Ed. - Cold Spring Harbor Press, New York 2001.

Sawant, S.V., Singh, P.K., Gupta, S.K., Madnala, R., Tuli, R.: Conserved nucleotide sequences in highly expressed genes in plants. - J. Genet. **78**: 123-131, 1999.

Singh, R., Low, L.E.T., Ooi, L.C.-L., Ong-Abdullah, M., Ting, N.-C., Nagappan, J., Nookiah, R., Mohd Din, A., Rozana, R., Mohamad Arif, A.M., Chan, K.L, Mohd Amin, H., Norazah, A., Lakey, N., Smith, S.W., Budiman, M.A., Hogan, M., Bacher, B., Brunt, A.V., Wang, C., Ordway, J.M., Sambanthamurthi, R., Martienssen, R.A.: The oil palm *SHELL* gene controls oil yield and encodes a homologue of *SEEDSTICK*. - Nature **500**: 340-344, 2013.

Siti Nor Akmar, A., Zubaidah, R.: Mesocarp-specific metallothionein-like gene promoter for genetic engineering of oil palm. - J Oil Palm Res. Special Issue. **2**: 1-8, 2008.

Tang, T.S.: Composition and properties of palm oil products. - In: Yusof, B., Jalani, B.S., Chan, K.W. (ed.): Advances in Oil Palm Research. Vol. 1. Pp. 845-895. Malaysian Palm Oil Board, Bangi 2000.

Wong, Y.T., Kushairi, A.D., Mohamad, O., Sambanthamurthi, R.: Assay of lipase activity in *Elaeis guineensis* in various germplasms. - In: Proceedings of the 2005 Conference on Biotechnology of Plantation Commodities. Pp. 441-444. Malaysian Palm Oil Board (MPOB), Ministry of Plantation and Commodities, Kajang 2005.

Xu, Z., Zhang, D, Hu, J., Zhou, X., Ye, X., Reichel, K.L., Stewart, N.R., Syrenne, R.D., Yang, X., Gao, P., Shi, W; Doeppke, C., Sykes, R.W., Burris, J.N., Bozell, J.J., Cheng, M.Z., Hayes, D.G., Labbe, N., Davis, M., Stewart, C.N.J., Yuan, J.S.: Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom. - BMC Bioinformatics **10** (Suppl 11): S3, 2009.

Yamagata, H., Yonesu, K., Hirata, A., Aizono, Y.: TGTCACA motif is a novel *cis*-regulatory enhancer element involved in fruit specific expression of the cucumisin gene. - J. biol. Chem. **277**: 11582-11590, 2002.

Yamamoto, Y.Y., Ichida, H., Matsui, M., Obokata, J., Sakurai, T., Satou, M., Seki, M., Shinozaki, K., Abe, T.: Identification of plant promoter constituents by analysis of local distribution of short sequences. - BMC Genomics **8**: 67, 2007.

Zhang, Y.: I-TASSER server for protein 3D structure prediction. - BMC Bioinformatics **9**: 40, 2008.

Zubaidah, R., Siti Nor Akmar, A.: Development of a transient promoter assay system for oil palm. - J Oil Palm Res. **15**: 62-69, 2003.

Zubaidah, R., Siti Nor Akmar, A.: Functional characterisation of the oil palm type 3 metallothionein-like gene (MT3-B) promoter. - Plant mol. Biol. Rep. **28**: 531-541, 2010.

Table 1 Suppl. The list of primers for PCR amplifications.

| Primer name | Sequence (5`-3`) |
|---|---|
| LAS2 | CCAGGCAAGATTAGCCCAGATGCCCTGTAG |
| LF6 | GTGCTCCAATTTCCTTAGGGTTGCTG |
| LAS10 | TAGCCGTTCCAAATCCAGTTCTTA |
| LF3 | CTTGATGAGAGTGGT |
| LR1 | GCCGTTCCAAATCCAGTTC |
| LF1 | GAATGCTTTGGGTGACCTC |
| Actin F | GTGCTAACGAATACAGTTCACG |
| Actin R | CCAGCAGATGTGGATTTCAAAG |
| LAS11 | CCGGCCACCGAGCTGCTGCTCTCGACAAA |
| LAS14 | GGATAAAAAGGAAGAAGGGTTTTGACAGAG |
| PLF2 | CGGAAGCTTCTACATATGTGTGTTAGC |
| PLR1 | GGGTCTAGAATAAAAAAGGAAGAAGGG |

Table 2 Suppl. Putative *cis* regulatory motifs in the *FLL1* promoter sequence.

| No. | Name | Sequence 5` to 3` /Strand | Function description |
|---|---|---|---|
| 1 | 5`UTR Py-rich stretch | TTTCTTCTCT / (+) | *cis*-acting element conferring high transcription levels |
| 2 | AAGAA-motif | GAAAGAA / (-) | |
| 3 | ACE | GCGACGTACC / (-) | *cis*-acting element involved in light responsiveness |
| 4 | AT1-motif | AATTATTTTTTATT/(+) | part of a light responsive module |
| 5 | Box4 | ATTAAT / (+) | part of a conserved DNA module involved in light responsiveness |
| 6 | G-box | CACATGG / (-) | *cis*-acting regulatory element involved in light responsiveness |
| 7 | GATA-motif | AAGGATAAGG / (+) | part of a light responsive element |
| 8 | I-box | CCTTATCCT and gGATAAGGTG/(+/-) | part of a light responsive element |
| 9 | LTR | CCGAAA and CCGAC (-/+) | *cis*-acting element involved in low-temperature responsiveness |
| 10 | Skn-1_motif | GTCAT / (-) | *cis*-acting element required for endosperm expression |
| 11 | TCA-element | CCATCTTTTT / (-) | *cis*-acting element involved in salicylic acid responsiveness |
| 12 | Unnamed_4 | CTCC / (+/-) | |
| 13 | As-2-box | GATAatGATG / (-) | involved in shoot-specific expression and light responsiveness |
| 14 | chs-CMA2a | TCACTTGA / (-) | part of a light responsive element |

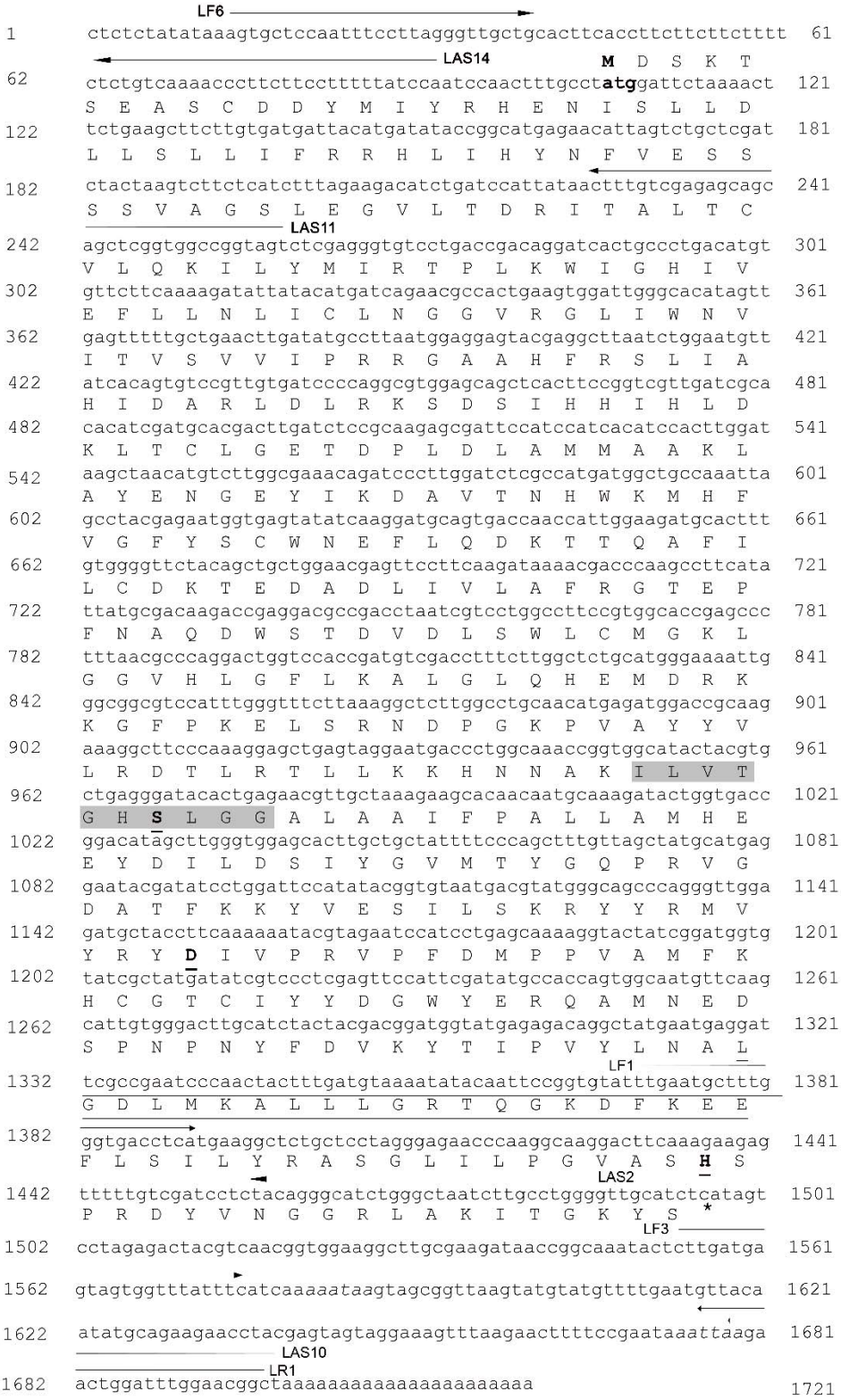



Fig. 1 Suppl. Nucleotide and inferred amino acid sequences of *FLL1*. The nucleotide sequence is shown in lower case letters, whereas the amino acid sequence is in capital letters. Numerals at both ends of the nucleotide row indicate a nucleotide number. The first start codon (atg) is *bold* and the stop codon (tga) is marked with an *asterisk*, respectively. The putative polyadenylation signals are in *italics*. The serine, aspartic acid, and histidine residues that form the catalytic triad are in *bold* and *underlined*. The lipase consensus motif surrounding the active serine residue is *shaded*. *Arrows* represent sense (LF6, LF1, and LF3) and antisense (LAS10, LAS2, LR1, LAS11, and LAS14) primer sequences designed for 5` RACE, LD-PCR, RT-PCR, and promoter isolation. The overlapping region between 5` RACE and O65EST cDNA is *underlined*.

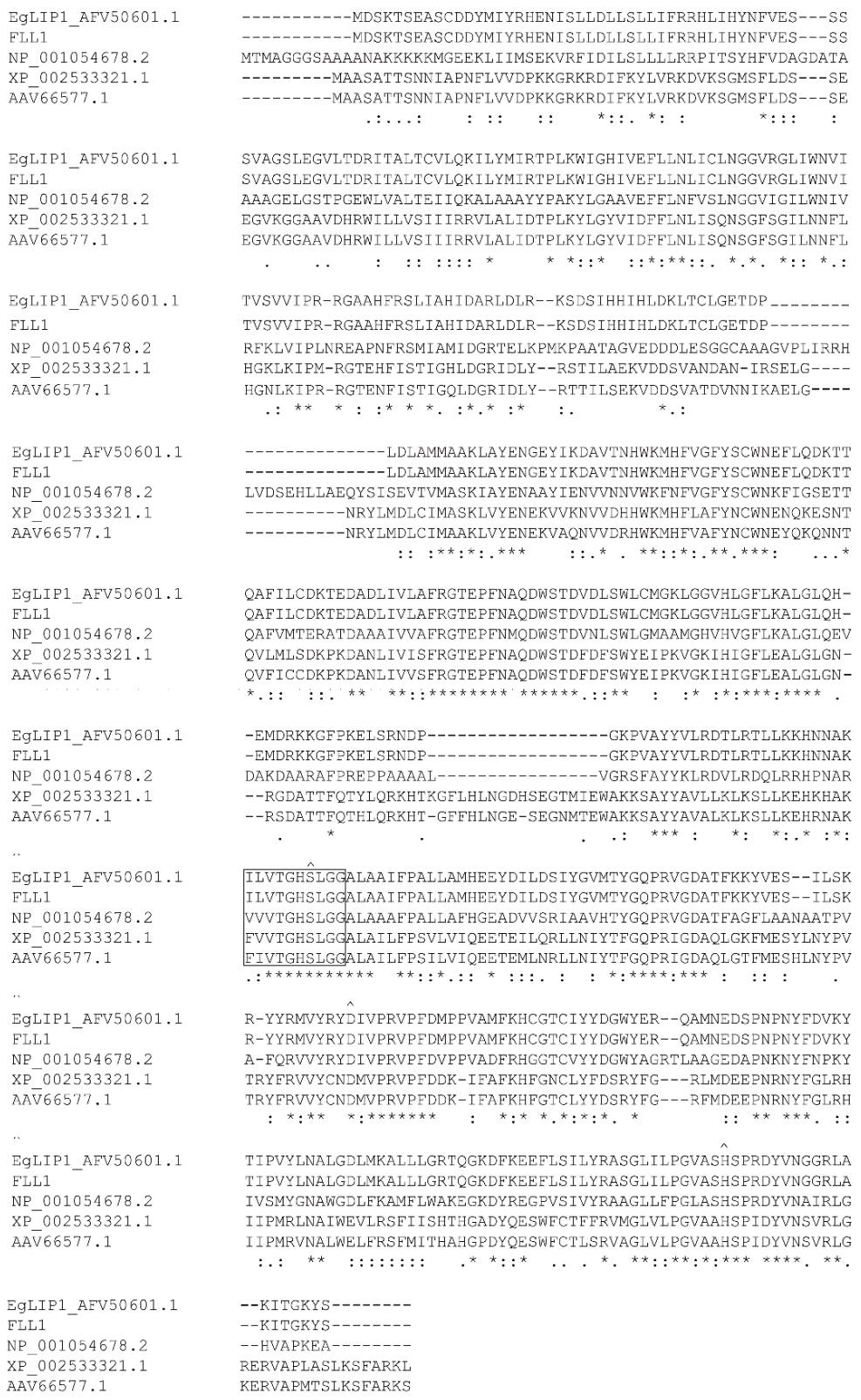



Fig. 2 Suppl. Multiple sequence alignment of FLL1 against the top four *BLASTX* hits in the *NCBI* sequence database. The ':' denotes conservation of strong groups. The '.' denotes conservation of weak groups, whereas those without any symbol denote no consensus (*CLUSTWALW*, Biology Workbench Version 3.2, the University of California). The lipase consensus motif is boxed, whereas the serine, aspartic acid, and histidine residues that forms the putative catalytic triad is marked with the '^'.

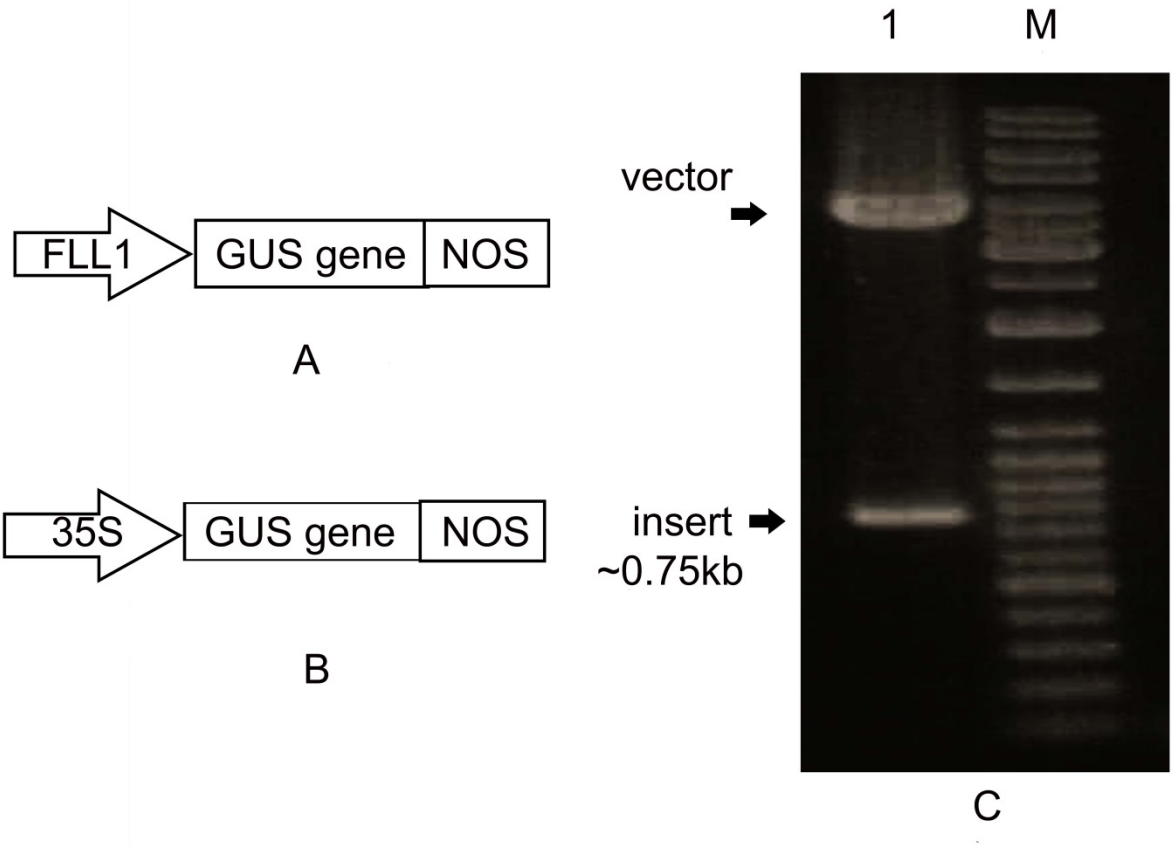



Fig. 3 Suppl. A schematic diagram of the FLL1GUS plasmid (*A*), pBI221 plasmid (*B*), and HindIII-Xba1 digestion of a transformation vector pBI221 carrying the FLL1 promoter on 1 % (m/v) agarose gel electrophoresis (*C*).

**APPENDIX 7: List of sequence identity of the putative members of the oil palm lipase class 3 family based on the top hit blastp searches against NCBI database.**

| No | Sequence name | Sequence identity/description | Sequence length | Hit ACC | E-Value | Bit-Score | Alignment length | Positives | Similarity (%) | Plant Hit |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | maker-p5_sc00101-snap-gene-19.43-mRNA-1_1 | lipase class 3-like | 396 | XP_008778985 | 0 | 747.273 | 396 | 373 | 94 | *Phoenix dactylifera* |
| 2 | maker-p5_sc00064-augustus-gene-2.38-mRNA-1_2 | lipase class 3 family protein | 516 | XP_009380716 | 0 | 615.15 | 507 | 398 | 78 | *Musa acuminata subsp. Malaccensis* |
| 3 | augustus_masked-p5_sc00028-abinit-gene-44.7-mRNA-1_3 | phospholipase a1-iigamma-like | 401 | XP_008812184 | 0 | 699.508 | 401 | 380 | 94 | *Phoenix dactylifera* |
| 4 | augustus_masked-p5_sc03770-abinit-gene-0.1-mRNA-1_4 | galactolipase chloroplastic-like | 437 | XP_008794177 | 0 | 711.064 | 394 | 372 | 94 | *Phoenix dactylifera* |
| 5 | augustus_masked-p5_sc00034-abinit-gene-42.4-mRNA-1_5 | galactolipase chloroplastic-like | 374 | XP_008791829 | 0 | 658.677 | 374 | 361 | 96 | *Phoenix dactylifera* |
| 6 | snap_masked-p5_sc00032-abinit-gene-11.16-mRNA-1_6 | phospholipase a1-chloroplastic-like | 412 | XP_002285367 | 0 | 633.254 | 403 | 344 | 85 | *Vitis vinifera* |
| 7 | snap_masked-p5_sc00013-abinit-gene-8.31-mRNA-1_7 | phospholipase a1-chloroplastic-like | 499 | XP_008801786 | 0 | 716.072 | 426 | 380 | 89 | *Phoenix dactylifera* |
| 8 | augustus_masked-p5_sc00062-abinit-gene-16.6-mRNA-1_8 | phospholipase a1-chloroplastic-like | 514 | XP_008776009 | 0 | 670.233 | 392 | 358 | 91 | *Phoenix dactylifera* |
| 9 | augustus_masked-p5_sc00023-abinit-gene-29.1-mRNA-1_9 | phospholipase a1-chloroplastic-like | 528 | XP_008798807 | 0 | 855.514 | 528 | 494 | 93 | *Phoenix dactylifera* |
| 10 | augustus_masked-p5_sc00006-abinit-gene-26.1-mRNA-1_10 | galactolipase chloroplastic-like | 449 | XP_008783953 | 0 | 666.766 | 435 | 393 | 90 | *Phoenix dactylifera* |
| 11 | augustus_masked-p5_sc00062-abinit-gene-17.1-mRNA-1_11 | phospholipase a chloroplastic-like | 398 | XP_008805240 | 0 | 584.334 | 346 | 323 | 93 | *Phoenix dactylifera* |
| 12 | maker-p5_sc00064-snap-gene-3.37-mRNA-1_12 | lipase | 499 | AFV50601 | 0 | 812.757 | 490 | 441 | 90 | *Elaeis guineensis* |
| 13 | maker-p5_sc00064-snap-gene-3.39-mRNA-2_13 | lipase | 298 | AFV50601 | 0 | 607.446 | 291 | 291 | 100 | *Elaeis guineensis* |

| No. | Sequence name | Sequence identity/description | Sequence length | Hit ACC | E-Value | Bit-Score | Alignment length | Positives | Similarity (%) | Plant Hit |
|---|---|---|---|---|---|---|---|---|---|---|
| 14 | maker-p5_sc00231-augustus-gene-3.41-mRNA-1_14 | triglyceride lipases isoform 1 | 989 | XP_008775136 | 0 | 1112.44 | 826 | 670 | 81 | *Phoenix dactylifera* |
| 15 | snap_masked-p5_sc00022-abinit-gene-50.17-mRNA-1_15 | phospholipase a1-chloroplastic-like | 485 | XP_008783171 | 0 | 735.717 | 400 | 383 | 95 | *Phoenix dactylifera* |
| 16 | snap_masked-p5_sc00035-abinit-gene-29.41-mRNA-1_16 | lipase | 254 | AFV50601 | 4.34E-63 | 216.468 | 170 | 130 | 76 | *Elaeis guineensis* |
| 17 | augustus_masked-p5_sc00396-abinit-gene-5.10-mRNA-1_17 | phospholipase a1-chloroplastic-like | 505 | XP_008784887 | 0 | 868.226 | 501 | 471 | 94 | *Phoenix dactylifera* |
| 18 | augustus_masked-p5_sc00013-abinit-gene-7.0-mRNA-1_18 | phospholipase a chloroplastic-like | 408 | XP_008801803 | 0 | 670.233 | 352 | 339 | 96 | *Phoenix dactylifera* |
| 19 | maker-p5_sc00032-augustus-gene-8.14-mRNA-1_19 | phospholipase a1-chloroplastic-like | 524 | XP_008802654 | 0 | 863.603 | 524 | 478 | 91 | *Phoenix dactylifera* |
| 20 | maker-p5_sc00032-snap-gene-48.40-mRNA-1_20 | lipase-like | 371 | XP_008796218 | 0 | 711.064 | 371 | 351 | 94 | *Phoenix dactylifera* |
| 21 | maker-p5_sc00021-augustus-gene-26.17-mRNA-1_21 | lipase-like isoform x2 | 421 | AEQ94180 | 0 | 682.559 | 389 | 337 | 86 | *Elaeis guineensis* |
| 22 | maker-p5_sc02332-augustus-gene-0.17-mRNA-1_22 | lipase-like isoform x2 | 386 | XP_008804573 | 0 | 618.616 | 331 | 316 | 95 | *Phoenix dactylifera* |
| 23 | p5_sc00018.path1.mRNA553_23 | phospholipase a1-ii 5 | 440 | XP_008807980 | 0 | 664.455 | 436 | 392 | 89 | *Phoenix dactylifera* |
| 24 | p5_sc00018.path1.mRNA556_24 | phospholipase a1-ii 5 | 449 | XP_008807980 | 0 | 752.666 | 450 | 417 | 92 | *Phoenix dactylifera* |
| 25 | maker-p5_sc00064-snap-gene-3.39-mRNA-1_25 | lipase | 242 | AFV50601 | 3.71E-175 | 503.442 | 242 | 242 | 100 | *Elaeis guineensis* |
| 26 | maker-p5_sc00064-augustus-gene-4.38-mRNA-1_26 | lipase | 238 | AFV50601 | 1.95E-111 | 340.887 | 237 | 192 | 81 | *Elaeis guineensis* |
| 27 | maker-p5_sc00007-augustus-gene-79.23-mRNA-1_27 | alpha beta-hydrolases superfamily protein | 711 | XP_008781074 | 0 | 935.636 | 741 | 646 | 87 | *Phoenix dactylifera* |
| 28 | maker-p5_sc00017-augustus-gene-35.54-mRNA-1_28 | sn1-specific diacylglycerol lipase beta | 551 | XP_008779417 | 0 | 1020.38 | 551 | 525 | 95 | *Phoenix dactylifera* |

| o. | Sequence name | Sequence identity/description | Sequence length | Hit ACC | E-Value | Bit-Score | Alignment length | Positives | Similarity (%) | Plant Hit |
|---|---|---|---|---|---|---|---|---|---|---|
| 29 | maker-p5_sc00126-snap-gene-9.42-mRNA-1_30 | PREDICTED: uncharacterised protein LOC103698231 isoform X1 | 715 | XP_008778442 | 0 | 828.55 | 691 | 609 | 88 | *Phoenix dactylifera* |
| 30 | snap_masked-p5_sc00064-abinit-gene-4.29-mRNA-1_31 | lipase class 3 family protein | 304 | XP_009391192 | 3.88E-86 | 278.1 | 358 | 216 | 60 | *Musa acuminata subsp. malaccensis* |
| 31 | p5_sc00064.path1.mRNA467_32 | sn1-specific diacylglycerol lipase alpha | 594 | XP_008792859 | 0 | 723.391 | 497 | 471 | 94 | *Phoenix dactylifera* |
| 32 | maker-p5_sc00037-augustus-gene-4.38-mRNA-1_33 | enhanced disease susceptibility 1 | 629 | XP_008799122 | 0 | 1133.24 | 628 | 602 | 95 | *Phoenix dactylifera* |
| 33 | p5_sc00018.path1.mRNA345_34 | triacylglycerol lipase 1 | 316 | XP_008791495 | 2.36E-144 | 436.032 | 287 | 256 | 89 | *Phoenix dactylifera* |
| 34 | p5_sc00197.path1.mRNA189_35 | calmodulin-binding heat-shock protein | 243 | XP_008787502 | 3.79E-154 | 448.743 | 228 | 220 | 96 | *Phoenix dactylifera* |
| 35 | p5_sc00044.path1.mRNA502_36 | enhanced disease susceptibility 1 | 648 | XP_008779531 | 0 | 1150.96 | 648 | 591 | 91 | *Phoenix dactylifera* |
| 36 | p5_sc00845.path1.mRNA37_37 | lipase-like pad4 | 652 | XP_008799057 | 0 | 1142.49 | 607 | 579 | 95 | *Phoenix dactylifera* |
| 37 | maker-p5_sc00045-augustus-gene-24.51-mRNA-1_38 | lipase-like pad4 | 615 | XP_008781628 | 0 | 912.138 | 624 | 549 | 87 | *Phoenix dactylifera* |
| 38 | p5_sc00004.path1.mRNA1026_39 | calmodulin-binding heat-shock | 259 | XP_008803902 | 2.58E-141 | 416.772 | 233 | 223 | 95 | *Phoenix dactylifera* |
| 39 | snap_masked-p5_sc00368-abinit-gene-7.23-mRNA-1_40 | uncharacterised loc101222656 | 237 | XP_008803902 | 4.21E-144 | 422.935 | 233 | 215 | 92 | *Phoenix dactylifera* |
| 40 | p5_sc00102.path1.mRNA118_41 | lipase class 3-like protein | 612 | XP_008795800 | 0 | 932.554 | 622 | 553 | 88 | *Phoenix dactylifera* |

| No. | Sequence name | Sequence identity/description | Sequence length | Hit ACC | E-Value | Bit-Score | Alignment length | Positives | Similarity (%) | Plant Hit |
|---|---|---|---|---|---|---|---|---|---|---|
| 41 | augustus_masked-p5_sc00001-abinit-gene-146.4-mRNA-1_42 | sn1-specific diacylglycerol lipase alpha-like | 371 | XP_008775446 | 1.41E-114 | 353.214 | 218 | 208 | 95 | *Phoenix dactylifera* |
| 42 | maker-p5_sc00001-snap-gene-108.39-mRNA-1_44 | gdsl esterase lipase at4g10955-like | 377 | XP_010267634 | 3.40E-170 | 491.115 | 314 | 268 | 85 | *Nelumbo nucifera* |
| 43 | p5_sc00116.path1.mRNA132_45 | gdsl esterase lipase at4g10955-like | 340 | XP_008795453 | 0 | 664.07 | 340 | 330 | 97 | *Phoenix dactylifera* |
| 44 | p5_sc00166.path1.mRNA129_46 | gdsl esterase lipase at4g10955-like | 339 | XP_008795453 | 2.79E-93 | 293.508 | 288 | 201 | 69 | *Phoenix dactylifera* |
| 45 | maker-p5_sc00166-snap-gene-8.52-mRNA-1_47 | gdsl esterase lipase at4g10955-like | 262 | XP_008795453 | 3.86E-110 | 333.569 | 216 | 187 | 86 | *Phoenix dactylifera* |
| 46 | maker-p5_sc00011-augustus-gene-36.58-mRNA-1_89 | lipase rog1 isoform x2 | 225 | XP_008785708 | 1.01E-121 | 363.614 | 200 | 182 | 91 | *Phoenix dactylifera* |
| 47 | maker-p5_sc01111-snap-gene-1.30-mRNA-1_48 | lipase class 3 family protein | 331 | XP_008811656 | 7.93E-104 | 338.961 | 257 | 220 | 85 | *Phoenix dactylifera* |
| 48 | p5_sc00166.path1.mRNA136_50 | gdsl esterase lipase at4g10955-like | 384 | XP_008795453 | 4.84E-118 | 358.607 | 344 | 246 | 71 | *Phoenix dactylifera* |
| 49 | augustus_masked-p5_sc00045-abinit-gene-1.6-mRNA-1_51 | protein phosphatase methylesterase 1 | 186 | XP_008786423 | 1.68E-52 | 182.185 | 94 | 92 | 97 | *Phoenix dactylifera* |
| 50 | maker-p5_sc00123-snap-gene-9.34-mRNA-1_52 | alpha beta-hydrolases superfamily | 457 | XP_008785708 | 0 | 583.178 | 402 | 323 | 80 | *Phoenix dactylifera* |
| 51 | maker-p5_sc00032-snap-gene-43.80-mRNA-1_61 | s-formylglutathione hydrolase | 281 | XP_008796163 | 0 | 563.148 | 281 | 274 | 97 | *Phoenix dactylifera* |
| 52 | maker-p5_sc00683-snap-gene-2.28-mRNA-1_64 | salicylic acid-binding protein 2-like | 311 | XP_009395210 | 3.07E-118 | 353.984 | 256 | 201 | 78 | *Musa acuminata subsp. malaccensis* |
| 53 | maker-p5_sc00070-augustus-gene-29.39-mRNA-1_69 | probable esterase pir7a | 270 | XP_008789941 | 5.38E-172 | 488.419 | 270 | 251 | 92 | *Phoenix dactylifera* |
| 54 | augustus_masked-p5_sc00037-abinit-gene-5.5-mRNA-1_70 | triacylglycerol lipase 2-like | 418 | XP_008788885 | 0 | 744.962 | 418 | 396 | 94 | *Phoenix dactylifera* |

| | Sequence name | Sequence identity/description | Sequence length | Hit ACC | E-Value | Bit-Score | Alignment length | Positives | Similarity (%) | Plant Hit |
|---|---|---|---|---|---|---|---|---|---|---|
| 55 | p5_sc00021.path1.mRNA572__72 | probable esterase pir7a | 216 | XP_008810789 | 3.22E-101 | 304.294 | 216 | 173 | 80 | *Phoenix dactylifera* |
| 56 | maker-p5_sc00021-snap-gene-48.16-mRNA-1_73 | probable esterase pir7a | 294 | XP_008800037 | 1.75E-117 | 350.517 | 254 | 198 | 77 | *Phoenix dactylifera* |
| 57 | maker-p5_sc00019-snap-gene-3.34-mRNA-1_74 | gpi inositol-deacylase isoform x1 | 819 | XP_008783007 | 0 | 1344.33 | 750 | 695 | 92 | |
| 58 | p5_sc00045.path1.mRNA113_75 | methylesterase chloroplastic | 345 | XP_008786473 | 0 | 620.542 | 345 | 331 | 95 | *Phoenix dactylifera* |
| 59 | maker-p5_sc00263-augustus-gene-1.30-mRNA-1_76 | methylesterase chloroplastic | 349 | XP_008802927 | 0 | 585.104 | 348 | 327 | 93 | *Phoenix dactylifera* |
| 60 | maker-p5_sc00263-augustus-gene-1.30-mRNA-1 | methylesterase chloroplastic | 349 | XP_008802927 | 0 | 585.104 | 348 | 327 | 93 | *Phoenix dactylifera* |
| 61 | maker-p5_sc00018-augustus-gene-39.43-mRNA-1_V2 | phospholipase a1-iidelta | 257 | XP_008807980 | 1.67E-137 | 407.142 | 253 | 233 | 92 | *Phoenix dactylifera* |
| 62 | maker-p5_sc00064-snap-gene-4.42-mRNA-1 | lipase | 161 | AFV50601 | 5.29E-54 | 188.734 | 147 | 109 | 74 | *Elaeis guineensis* |

**APPENDIX 8: List of annotations of the 62 proteins by sequence based on GO terms.**

| No. | Sequence name | Putative ID | GO-ID | Term |
|---|---|---|---|---|
| 1 | maker-p5_sc00101-snap-gene-19.43-mRNA-1_1 | lipase class 3-like | GO:0006629 | lipid metabolic process |
| 2 | maker-p5_sc00064-augustus-gene-2.38-mRNA-1_2 | lipase class 3 family protein | GO:0006629 | lipid metabolic process |
| 3 | augustus_masked-p5_sc00028-abinit-gene-44.7-mRNA-1_3 | phospholipase a1-iigamma-like | GO:0006629 | lipid metabolic process |
| 4 | augustus_masked-p5_sc03770-abinit-gene-0.1-mRNA-1_4 | galactolipase chloroplastic-like | GO:0006629 | lipid metabolic process |
| 5 | augustus_masked-p5_sc00034-abinit-gene-42.4-mRNA-1_5 | galactolipase chloroplastic-like | GO:0006629 | lipid metabolic process |
| 6 | snap_masked-p5_sc00032-abinit-gene-11.16-mRNA-1_6 | phospholipase a1- chloroplastic-like | GO:0006629 | lipid metabolic process |
| 7 | snap_masked-p5_sc00013-abinit-gene-8.31-mRNA-1_7 | phospholipase a1- chloroplastic-like | GO:0006629 | lipid metabolic process |
| 8 | augustus_masked-p5_sc00062-abinit-gene-16.6-mRNA-1_8 | phospholipase a1- chloroplastic-like | GO:0006629 | lipid metabolic process |
| 9 | augustus_masked-p5_sc00023-abinit-gene-29.1-mRNA-1_9 | phospholipase a1- chloroplastic-like | GO:0006629 | lipid metabolic process |
| 10 | augustus_masked-p5_sc00006-abinit-gene-26.1-mRNA-1_10 | galactolipase chloroplastic-like | GO:0006629 | lipid metabolic process |
| 11 | augustus_masked-p5_sc00062-abinit-gene-17.1-mRNA-1_11 | phospholipase a chloroplastic-like | GO:0006629 | lipid metabolic process |
| 12 | maker-p5_sc00064-snap-gene-3.37-mRNA-1_12 | lipase | GO:0006629 | lipid metabolic process |
| 13 | maker-p5_sc00064-snap-gene-3.39-mRNA-2_13 | lipase | GO:0006629 | lipid metabolic process |
| 14 | maker-p5_sc00231-augustus-gene-3.41-mRNA-1_14 | triglyceride lipases isoform 1 | GO:0005515 | protein binding |
| | maker-p5_sc00231-augustus-gene-3.41-mRNA-1_14 | triglyceride lipases isoform 1 | GO:0006629 | lipid metabolic process |
| 15 | snap_masked-p5_sc00022-abinit-gene-50.17-mRNA-1_15 | phospholipase a1- chloroplastic-like | GO:0006629 | lipid metabolic process |
| 16 | snap_masked-p5_sc00035-abinit-gene-29.41-mRNA-1_16 | lipase | GO:0006629 | lipid metabolic process |
| 17 | augustus_masked-p5_sc00396-abinit-gene-5.10-mRNA-1_17 | phospholipase a1- chloroplastic-like | GO:0006629 | lipid metabolic process |
| 18 | augustus_masked-p5_sc00013-abinit-gene-7.0-mRNA-1_18 | phospholipase a chloroplastic-like | GO:0006629 | lipid metabolic process |
| 19 | maker-p5_sc00032-augustus-gene-8.14-mRNA-1_19 | phospholipase a1- chloroplastic-like | GO:0006629 | lipid metabolic process |
| 20 | maker-p5_sc00032-snap-gene-48.40-mRNA-1_20 | lipase-like | GO:0006629 | lipid metabolic process |
| 21 | maker-p5_sc00021-augustus-gene-26.17-mRNA-1_21 | lipase-like isoform x2 | GO:0006629 | lipid metabolic process |
| 22 | maker-p5_sc02332-augustus-gene-0.17-mRNA-1_22 | lipase-like isoform x2 | GO:0006629 | lipid metabolic process |
| 23 | p5_sc00018.path1.mRNA553_23 | phospholipase a1-ii 5 | GO:0006629 | lipid metabolic process |

| No. | Sequence name | Putative ID | GO-ID | Term |
|---|---|---|---|---|
| 24 | p5_sc00018.path1.mRNA556_24 | phospholipase a1-ii 5 | GO:0006629 | lipid metabolic process |
| 25 | maker-p5_sc00064-snap-gene-3.39-mRNA-1_25 | lipase | GO:0006629 | lipid metabolic process |
| 26 | maker-p5_sc00064-augustus-gene-4.38-mRNA-1_26 | lipase | GO:0006629 | lipid metabolic process |
| 27 | maker-p5_sc00007-augustus-gene-79.23-mRNA-1_27 | alpha beta-hydrolases superfamily protein | GO:0006629 | lipid metabolic process |
| 28 | maker-p5_sc00017-augustus-gene-35.54-mRNA-1_28 | sn1-specific diacylglycerol lipase beta | GO:0006629 | lipid metabolic process |
| 29 | maker-p5_sc00126-snap-gene-9.42-mRNA-1_30 | PREDICTED: uncharacterised protein LOC103698231 isoform X1 | GO:0006629 | lipid metabolic process |
| 30 | snap_masked-p5_sc00064-abinit-gene-4.29-mRNA-1_31 | lipase class 3 family protein | GO:0006629 | lipid metabolic process |
| 31 | p5_sc00064.path1.mRNA467_32 | sn1-specific diacylglycerol lipase alpha | GO:0009056 | catabolic process |
| | p5_sc00064.path1.mRNA467_32 | sn1-specific diacylglycerol lipase alpha | GO:0006629 | lipid metabolic process |
| 32 | maker-p5_sc00037-augustus-gene-4.38-mRNA-1_33 | enhanced disease susceptibility 1 | GO:0006629 | lipid metabolic process |
| 33 | p5_sc00018.path1.mRNA345_34 | triacylglycerol lipase 1 | GO:0006629 | lipid metabolic process |
| 34 | p5_sc00197.path1.mRNA189_35 | calmodulin-binding heat-shock protein | | |
| 35 | p5_sc00044.path1.mRNA502_36 | enhanced disease susceptibility 1 | | |
| 36 | p5_sc00845.path1.mRNA37_37 | lipase-like pad4 | GO:0006629 | lipid metabolic process |
| 37 | maker-p5_sc00045-augustus-gene-24.51-mRNA-1_38 | lipase-like pad4 | GO:0006629 | lipid metabolic process |
| 38 | p5_sc00004.path1.mRNA1026_39 | calmodulin-binding heat-shock | GO:0009056 | catabolic process |
| | p5_sc00004.path1.mRNA1026_39 | calmodulin-binding heat-shock | GO:0006629 | lipid metabolic process |
| 39 | snap_masked-p5_sc00368-abinit-gene-7.23-mRNA-1_40 | uncharacterised loc101222656 | GO:0009056 | catabolic process |
| | snap_masked-p5_sc00368-abinit-gene-7.23-mRNA-1_40 | uncharacterised loc101222656 | GO:0006629 | lipid metabolic process |
| 40 | p5_sc00102.path1.mRNA118_41 | lipase class 3-like protein | GO:0009056 | catabolic process |
| | p5_sc00102.path1.mRNA118_41 | lipase class 3-like protein | GO:0006629 | lipid metabolic process |
| 41 | augustus_masked-p5_sc00001-abinit-gene-146.4-mRNA-1_42 | sn1-specific diacylglycerol lipase alpha-like | GO:0009056 | catabolic process |
| | augustus_masked-p5_sc00001-abinit-gene-146.4-mRNA-1_42 | sn1-specific diacylglycerol lipase alpha-like | GO:0006629 | lipid metabolic process |
| 42 | maker-p5_sc00001-snap-gene-108.39-mRNA-1_44 | gdsl esterase lipase at4g10955-like | GO:0006629 | lipid metabolic process |

| No. | Sequence name | Putative ID | GO-ID | Term |
|---|---|---|---|---|
| 43 | p5_sc00116.path1.mRNA132_45 | gdsl esterase lipase at4g10955-like | GO:0006629 | lipid metabolic process |
| 44 | p5_sc00166.path1.mRNA129_46 | gdsl esterase lipase at4g10955-like | GO:0006629 | lipid metabolic process |
| 45 | maker-p5_sc00166-snap-gene-8.52-mRNA-1_47 | gdsl esterase lipase at4g10955-like | GO:0006629 | lipid metabolic process |
| 46 | maker-p5_sc01111-snap-gene-1.30-mRNA-1_48 | lipase class 3 family protein | GO:0006629 | lipid metabolic process |
| 47 | maker-p5_sc00011-augustus-gene-36.58-mRNA-1_89 | lipase rog1 isoform x2 | | |
| | maker-p5_sc00032-snap-gene-43.80-mRNA-1_61 | s-formylglutathione hydrolase | GO:0009987 | cellular process |
| | maker-p5_sc00032-snap-gene-43.80-mRNA-1_61 | s-formylglutathione hydrolase | GO:0009056 | catabolic process |
| 50 | augustus_masked-p5_sc00037-abinit-gene-5.5-mRNA-1_70 | triacylglycerol lipase 2-like | GO:0016787 | hydrolase activity |
| | augustus_masked-p5_sc00037-abinit-gene-5.5-mRNA-1_70 | triacylglycerol lipase 2-like | GO:0006629 | lipid metabolic process |
| 51 | maker-p5_sc00019-snap-gene-3.34-mRNA-1_74 | gpi inositol-deacylase isoform x1 | GO:0016787 | hydrolase activity |
| | maker-p5_sc00019-snap-gene-3.34-mRNA-1_74 | gpi inositol-deacylase isoform x1 | GO:0009987 | cellular process |
| | maker-p5_sc00019-snap-gene-3.34-mRNA-1_74 | gpi inositol-deacylase isoform x1 | GO:0006629 | lipid metabolic process |
| | maker-p5_sc00019-snap-gene-3.34-mRNA-1_74 | gpi inositol-deacylase isoform x1 | GO:0006810 | transport |
| 52 | maker-p5_sc00018-augustus-gene-39.43-mRNA-1_V2 | phospholipase a1-iidelta | GO:0006629 | lipid metabolic process |
| 53 | maker-p5_sc00064-snap-gene-4.42-mRNA-1 | lipase | GO:0006629 | lipid metabolic process |
| 54 | maker-p5_sc00683-snap-gene-2.28-mRNA-1_64 | salicylic acid-binding protein 2-like | | |
| 55 | maker-p5_sc00070-augustus-gene-29.39-mRNA-1_69 | probable esterase pir7a | | |
| 56 | p5_sc00021.path1.mRNA572__72 | probable esterase pir7a | | |
| 57 | maker-p5_sc00021-snap-gene-48.16-mRNA-1_73 | probable esterase pir7a | | |
| 58 | p5_sc00045.path1.mRNA113_75 | methylesterase chloroplastic | | |
| 59 | maker-p5_sc00263-augustus-gene-1.30-mRNA-1_76 | methylesterase chloroplastic | | |
| 60 | maker-p5_sc00064-snap-gene-4.42-mRNA-1 | lipase | | |
| 61 | p5_sc00166.path1.mRNA136_50 | gdsl esterase lipase at4g10955-like | | |
| 2 | maker-p5_sc00123-snap-gene-9.34-mRNA-1_52 | alpha beta-hydrolases superfamily | | |