

# OIL PALM ETHYLENE RECEPTOR GENE FAMILY: IDENTIFICATION, CHARACTERISATION AND EXPRESSION ANALYSIS

NURNI WALIS ABDUL WAHAB\*; CHAN KUANG LIM\*; ZUBAIDAH RAMLI\*; ROZANA ROSLI\*; FADILA AHMAD MALIKI\*; MOHAMAD ARIF ABD MANAF\*; JEREMY ROBERTS\*\* and FESTO MASSAWE†

## ABSTRACT

The ethylene receptor family is the first component of the ethylene signaling pathway to trigger the plant's response towards ethylene and consequently affect plant growth and developmental processes. In oil palm (*Elaeis guineensis* Jacq.), an ethylene receptor gene has been isolated and characterised. However, the full information about the gene family is unavailable. Thus, in this study, we focused on the identification, isolation and characterisation of the ethylene receptor gene family from oil palm. A total of seven genes were identified, isolated and characterised using a combination of *in silico* bioinformatics tools and experimental insights. Structural comparison and phylogenetic analyses showed the genes are classified into two subfamilies. Subfamily I consists of three members whereas subfamily II has four members. Expression analyses of the ethylene receptors in various oil palm tissues suggested multiple roles of the ethylene receptors in regulating many processes in oil palm growth and development such as fruit growth, ripening and abscission, and flower development. Therefore, all together, these results will be beneficial to further investigate the functions of ethylene and the potential use of the genes to manipulate ethylene responses in oil palm for yield improvement.

**Keywords:** ethylene signaling, fruit development, mesocarp.

**Date received:** 18 October 2019; **Sent for revision:** 20 October 2019; **Accepted:** 10 January 2020; **Available online:** 29 July 2020.

## INTRODUCTION

Ethylene (C<sub>2</sub>H<sub>4</sub>) is an endogenous plant hormone that influences numerous processes throughout the plant life cycle. This includes plant growth and developmental processes especially germination, root initiation, root hair development, flower development, sex determination, senescence,

abscission, fruit ripening as well as responses to biotic and abiotic stresses (Abeles *et al.*, 1992). Ethylene is a simple hydrocarbon structured molecule that have unique chemical and physical properties which enable it to be present in a gaseous form under ambient conditions (Lin *et al.*, 2009). A trace amount of ethylene is sufficient to cause major responses that are either desirable or harmful to the plants. Thus, the ability to control ethylene responses is significant in agriculture especially in areas that have substantial commercial importance (Chang, 2016).

In other plant species, genes involved in ethylene synthesis and signaling pathway in relations to altering ethylene levels for crop improvement have been widely studied (Czarny *et al.*, 2006). This is especially targeted to alter fruit ripening, abscission and senescence (Agarwal *et al.*, 2012). In oil palm, research on ethylene has not been conducted extensively but its role in promoting the ripening

\* Malaysian Palm Oil Board,  
6 Persiaran Institusi, Bandar Baru Bangi,  
43000 Kajang, Selangor, Malaysia.  
E-mail: nurni@mpob.gov.my

\*\* Office of Vice Chancellor,  
University of Plymouth, Drake Circus,  
Plymouth, Devon PL4 8AA, United Kingdom.

† School of Biosciences,  
University of Nottingham Malaysia,  
Jalan Broga, 43500 Semenyih,  
Selangor, Malaysia.

process of oil palm fruit has been established (Henderson and Osborne, 1999; Tranbarger *et al.*, 2011; Nurniwalis, 2017). Furthermore, the role of ethylene in triggering oil accumulation and other changes associated with fruit ripening has been raised (Sambanthamurthi *et al.*, 2000). At the molecular level, gene transcripts associated with ethylene synthesis and signaling pathway have been identified from oil palm mesocarp (Nurniwalis *et al.*, 2008; Tranbarger *et al.*, 2011). In addition, a full-length ethylene receptor gene, *EgERD3*, was previously isolated and characterised via conventional methods (Nurniwalis *et al.*, 2018). Northern analysis of *EgERD3* demonstrated that the gene was expressed in the mesocarp tissues following the ripening pattern of the oil palm fruit (Nurniwalis *et al.*, 2018). These findings indicate an important role for ethylene in the development and ripening process of oil palm fruits.

Ethylene receptors are the first components in the complex ethylene-signaling pathway and play a key role in sensing the ethylene gas. They negatively regulate the signaling pathway, where in the presence of ethylene the receptors become inactivated, which leads to the induction of ethylene signaling and ultimately, an ethylene response (Hua and Meyerowitz, 1998; Tieman *et al.*, 2000). Plant responses to ethylene are mediated by a family of receptors. However, the number of receptors in the gene family differs between species (Hua and Meyerowitz, 1998; Tieman *et al.*, 2000). Based on this information, the presence of other receptor isoforms in oil palm would also be expected. Subsequently, Southern blot analysis of the oil palm *EgERD3* gene confirmed the presence of other ethylene receptor isoforms (Nurniwalis *et al.*, 2018).

The availability of the oil palm genome data provides the opportunity to identify the other putative members of the ethylene receptors family in a much shorter time as compared to the traditional methods. In *Arabidopsis* and tomato, the different receptor isoforms from the ethylene receptor family have been shown to play redundant and specific roles (Hua *et al.*, 1998; Binder *et al.*, 2012; Kevany *et al.*, 2008; Tieman *et al.*, 2000; Shakeel *et al.*, 2013). The overlapping and non-overlapping roles of the receptor isoforms make the identification and characterisation of the various receptor isoforms within the gene family important. Hence, as other members of the oil palm ethylene receptor family have yet to be discovered, this article reports the process of mining and identification of the ethylene receptor genes from the oil palm genome. To further understand and clarify the function of these genes in the oil palm, we also conducted isolation of the full-length complementary deoxyribonucleic acids, prediction of their protein structure, gene structure analysis and characterisation of their expression patterns in various oil palm tissues. The results can

help further our understanding on the role of the ethylene receptor genes during oil palm growth and development especially in the reproductive tissues. These findings are important to move forward in the ethylene research in oil palm with the aim to help improve oil palm yield.

## MATERIALS AND METHODS

### Plant Materials

Oil palm *Elaeis guineensis* var. *tenera* (progeny of *dura* x *pisifera*) tissues were used as research material. Inflorescences from MPOB-UKM Research Station, Bangi, Selangor, Malaysia were tagged at six developmental stages. Fresh fruit bunches were harvested and the fruits processed for their mesocarp and kernel tissues. Spear leaves from unopened leaf fronds, roots from two-year old seedlings and germinated seedlings of one- to two-week old were obtained from MPOB Kluang Research Station, Kluang, Johor, Malaysia.

### Mining of Oil Palm Ethylene Receptor Genes

Homology search and Hidden Markov Model (HMM) approaches were employed to identify oil palm ethylene receptor genes. Query keywords such as ethylene receptor, *ETR*, *ERS* and *EIN4* were used to download nucleotide and/or protein sequences from public databases such as 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.expasy.org/sprot/>) and date palm (*Phoenix dactylifera*) draft sequence (<http://qatar-weill.cornell.edu/research/datepalmGenome/>). Nucleotide and protein sequences especially those from *Arabidopsis thaliana*, *Oryza sativa*, *Solanum lycopersicum*, and *Phoenix dactylifera* and *EgERD3* from *E. guineensis* (Nurniwalis *et al.*, 2018) were used to search for similarities against an in-house *E. guineensis* transcriptome database (transcriptome assembly of GenBank BioProject PRJNA201497 and PRJNA345530) and P5 genome database (Singh *et al.*, 2013a) based on BLASTN/BLASTX/BLASTP programs. The assembled 1.535 Gb P5-build was constructed based on a combination of Roche/454 GS FLX Titanium (Roche/454) and Sanger bacterial artificial chromosome end sequencing and contains 74% predicted genes with RefSeq support (Singh *et al.*, 2013a). In addition, a HMM profile was constructed using HMMER3 package (Eddy, 2011). Multiple sequence alignment of the downloaded protein sequences was performed to construct the HMMbuild profile and used as a query to search

against the Oil Palm Gene Model (Singh *et al.*, 2013a; Chan *et al.*, 2017) using hmmsearch (Eddy, 2011). BLASTP search of the identified protein sequences against the in-house transcriptome and genome databases (Singh *et al.*, 2013a) was conducted.

### ***In silico* Analyses**

Nucleotide and amino acid (aa) sequence analyses were performed using Bioedit (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>) and/or ExPasy toolkit ([www.expasy.org](http://www.expasy.org)) and/or Biology Workbench Version 3.2 (<http://workbench.sdsc.edu>). Protein domains and conserved motifs were identified using NCBI-CDD search programme (Marchler-Bauer *et al.*, 2013), Interproscan5 or IPRScan (<http://www.ebi.ac.uk/Tools/pfa/iprscan5/>) and/or SMART software (<http://smart.embl-heidelberg.de/>). Subcellular localisation of the ethylene receptor genes was performed using Protcomp version 9.0 from softberry toolkit (<http://www.softberry.com>). BLAST (Altschul *et al.*, 1997) was used to search the GenBank/TAIR/RiceGD databases for similarities of the fragments generated from 5', 3' RACE and LD-PCR.

### **Phylogenetic Analysis**

The phylogenetic tree was constructed with MEGA5 software (Tamura *et al.*, 2011) by the Neighbour-joining method (Saitou and Neim, 1987) using bootstrap analysis with 1000 replicates. Multiple protein sequences were aligned using ClustalW (Thompson *et al.*, 1994). Evolutionary protein distances were calculated in aa units using Poisson parameter model with pairwise deletion of gaps in the alignment data.

### **Gene Structure and Chromosomal Localisation**

Prediction of the gene structure was carried out using exonerate program (Slater and Birney, 2005) with est2genome model parameter and visualised using Gene Structure Display Server 2.0 (<http://www.gsds.cbi.pku.edu.cn>). Mapping of the ethylene receptor genes to oil palm chromosomes was performed by identifying their scaffold positions on the P5 build genome (Singh *et al.*, 2013a) using exonerate search (Slater and Birney, 2005) with a 60% self-score threshold and protein2genome model parameters. Genetic markers flanking the ethylene receptor genes were identified based on their mapped positions to the scaffolds of the P5-build (Ting *et al.*, 2014). Markers information on the T128 and P2 linkage groups (Singh *et al.*, 2013a, b; Ting *et al.*, 2014) were used to predict the location of the ethylene receptors. Graphic representation of the linkage groups was visualised using MapChart 2.2 (Voorrips, 2002).

### **Ribonucleic Acid (RNA) Extraction, Complementary Deoxyribonucleic Acid (cDNA) Synthesis, Rapid Amplification of cDNA End (RACE) and Long-distance-polymerase Chain Reaction (LD-PCR)**

Total RNA isolated from the oil palm tissues were performed as described by Nurniwalis *et al.*, 2008; 2018). Total RNA from oil palm male and female flowers were kindly provided by Dr Meilina Ong-Abdullah and Dr Ooi Siew Eng from Biomarker Discovery Laboratory, MPOB, Bangi, Selangor, Malaysia. The first strand RACE-Ready cDNA was synthesised using SmartScribe Reverse Transcriptase (Clontech, USA) according to the manufacturer's protocol. The 5' and 3' RACE was carried out using SMARTer™ RACE cDNA Amplification and Advantage 2 PCR Kit (Clontech, USA) according to the manufacturer's protocol. Touchdown PCR amplification was carried out in a 50 µl reaction mixture containing 1X Advantage 2 PCR buffer, 200 µM dNTP mix, 1X Advantage 2 Polymerase Mix, 1X universal primer A mix, 200 µM gene-specific primers (*Table 1*) and finally 100 ng of the first strand RACE-Ready cDNA with PCR conditions as recommended by the manufacturer. Some adjustments ( $\pm 5^{\circ}\text{C}$ ) were made to the annealing temperatures especially when the PCR amplification did not generate a discrete band on the agarose gel. For LD-PCR, PCR amplifications were carried out similar to the 5' RACE, using gene-specific primer combinations (*Table 1*) designed from the far end of 5' and 3' RACE products.

### **DNA Manipulation and Sequence Analyses**

The amplified PCR products were analysed using 1% agarose gel electrophoresis, purified using Qiaquick Gel Extraction Kit (Qiagen, Germany) and cloned into pCR®II-TOPO® vector (TOPO TA Cloning Kit, Invitrogen, USA). The plasmids were isolated using Qiagen Plasmid Mini Kit (Qiagen, Germany) and digested with *EcoRI*. DNA sequencing of plasmids and amplified PCR products was carried out by a 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.

### **Expression Analysis via Reverse Transcription PCR (RT-PCR)**

First strand cDNA was synthesised from 2 µg of DNase-treated total RNA by using High-capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA) following the manufacturer's instructions. RT-PCR was carried out using Advantage 2 PCR Kit (Clontech, USA) in a reaction mixture containing 0.1 µg cDNA, 1X Advantage 2 PCR buffer (Clontech, USA), 1X Advantage 2

Polymerase Mix, 1X dNTP mix and 0.2  $\mu$ M forward and reverse gene-specific primers (Table 1). PCR was performed with the following procedure: denaturation at 95°C (1 min); 95°C (30 s) and 60°C (1 min) for 30 cycles and a final extension of 60°C for 1 min. Some adjustments ( $\pm$  5°C) were made to the annealing temperatures especially when a discrete band was not obtained on the agarose gel.

Amplification of the actin gene as internal control was performed in a reaction mixture containing 0.1  $\mu$ g cDNA, 1X Advantage 2 PCR buffer (Clontech, USA), 1X Advantage 2 Polymerase Mix, 1X dNTP mix, 0.2  $\mu$ M actin F and 0.2  $\mu$ M actin R primers (Table 1) 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 of 72°C for 7 min.

TABLE 1. LIST OF PRIMERS FOR AMPLIFICATION OF ETHYLENE RECEPTOR GENES

Method	Gene ID	Primer name	Sequence (5'-3')
3' RACE	<i>EgERS1</i>	34176F3	CCA AGG ATA TAG CGG CAG TGG ACT T
	<i>EgERS2</i>	45619F2	TTG GGC TTG CCA TTT GTA
	<i>EgERS3</i>	46141F2	CTC GGG CTC GCC ATT TGT AAG AG GTT T
	<i>EgETR3</i>	3542F2	CTG GCA TCC AAT GTC TGA GCT CCT TC
	<i>EgETR2</i>	2412F2	AGC CGC AGG TCT CTA TCG GAC ATC TT
	<i>EgETR4</i>	44932F4	GGG AGA AAT GTG TGC AGG TGG GAA T
	<i>EgETR5</i>	53269UF1	GCC ACT GAT CGT TGC ACT CAC TCC T
5' RACE	<i>EgERS1</i>	34176R2	CTT CCT GTG TCC GTA TCA GGC CC
	<i>EgERS2</i>	45619R1	GCT GCC ACG GCC ACA ATA TTC TTC TC
	<i>EgERS3</i>	46141UR2	TTC TCG AGT GCA CGG TGA AAG TCC A
	<i>EgETR3</i>	35423R1	TGT ATG GCG GAT GCA CAG AAG AAG AGG
	<i>EgETR2</i>	24121R2	GTG AAC ACA TTG AGC AGG TGG GTC AG
	<i>EgETR4</i>	44932R3	TGA AGG GGA AGA GGT TGG AGC ATG T
	<i>EgETR5</i>	53269R4	CGG AGC ACG TGA CGA AGT AGA AGA GC
LD-PCR	<i>EgERS1</i>	ERS1F1 ERS1R1	GAA GAC GAC GAG GGT CTG CT TTT TGG TAT CGA GCC AGT CCT T
	<i>EgERS2</i>	ERS2F1 ERS2R1	GGA GTC GAG TCT CCT TTT CGA GCA CTG TCA TGC TCC ATT GTG
	<i>EgERS3</i>	ERS3F1 ERS3R1	GTA GAC GAG AAA GGC GAC AAA CTG TAT TGC TCC ATC ATG CAG A
	<i>EgETR2</i>	ETR2F1 ETR2R1	CGC CAT TGA GAT CGG ACG T GTA GCA GCA TTC GGA AAA GGA GT
	<i>EgETR3</i>	ETR3F1 ETR3R1	TGC ATC CGC CAT ACA GAT TGG ATT C TAT CCC ACA CTG CAA GCA CCT ATC C
	<i>EgETR4</i>	ETR4F1 ETR4R1	CTT TCC TTA CAT TTCTTGCTTGCGC ATT TCT CCC ACG TATCTTCCTCAAC
	<i>EgETR5</i>	ETR5F1 ETR5R1	ATG CGG TCG TCG TGG TCT TTC TTG GGT GAT CCT GTC CAA TTC GTC T
RT-PCR	<i>EgERS1</i>	ERS13F1 ERS13R1	GGT TTG AAT GTC TGG TTA GCA ATA CC CCG ATA GAA GGA TGT TGA TTA TAC AGT GCT TC
	<i>EgERS2</i>	ERS23F1 ERS23R1	CGT TGT GGA AGA AAA TGA TAT ACA CAA TGG AGG AAA CCC CTT CAC CAA TAG ATT TG
	<i>EgERS3</i>	ERS33F1 ERS33R1	CTG CAT GAT GGA GCA ATA CAG CTT GA AAG TAA AGA TAT ACG ATT CAT GGC TCC TCC
	<i>EgETR2</i>	ETR23F1 2412R3	CAT GCC CAA AAT GGA TGT TTT TGA GG TGA GGG GTA AAT GGA GAG TGG GAG CTA A
	<i>EgETR3</i>	ETR33F1 3524R3	GACAGCTCTCCATTATTCTCCAGG GGTGATTATCTGCATTTGATCTTCAATATTAAGTG
	<i>EgETR4</i>	ETR43F2 44932R5	GGTTCACCATATACAGTAATACTTGACAC CTC ACT TCG ACC AGT TTG TGA CAA CCA C
	<i>EgETR5</i>	ETR53F1 ETR53R1	CAAATATCACCCAAGCAATGGAGAG TGAAGGAAAATGGAGCTCATAGACCA

TABLE 2. LIST OF PUTATIVE OIL PALM ETHYLENE RECEPTOR GENES AND THEIR SEQUENCE CHARACTERISTICS

Gene name	Length (bp)	Protein (open reading frame)						
		Length (aa)	MW (kDa)	PI	II	AI	GRAVY	SL
EgERS1	2 479	629	70.037	7.33	38.79 (S)	106.53	0.119	ER
EgERS2	2 565	635	70.779	7.10	41.06 (US)	107.65	0.132	ER
EgERS3	2 204	635	70.857	7.11	39.82 (S)	111.17	0.165	ER
EgETR2	3 051	687	76.806	6.31	39.98 (S)	99.91	0.082	ER
EgETR3	3 195	757	85.203	6.31	37.49 (S)	97.58	0.042	ER
EgETR4	4 186	755	85.145	6.71	39.01 (S)	101.55	0.067	ER
EgETR5	3 187	737	82.359	6.46	50.79 (US)	111.90	0.164	ER

Note: bp - base pair; aa - amino acid; MW - molecular weight; PI - isoelectric point; AI - aliphatic index; GRAVY - grand average of hydrophobicity; II - instability index; S - stable; US - unstable; SL - subcellular localisation; ER - endoplasmic reticulum.

## RESULTS

### Identification and Characterisation of Ethylene Receptor Genes

We have identified seven ethylene receptor genes in oil palm using homology search and HMMbuild profile. The seven ethylene receptor genes are designated as *EgERS1*, *EgERS2*, *EgERS3*, *EgETR2*, *EgETR3*, *EgETR4* and *EgETR5* and their corresponding encoded proteins are named EgERS1, EgERS2, EgERS3, EgETR2, EgETR3, EgETR4 and EgETR5, respectively. The nucleotide information conforms to those available in Genbank with NCBI Reference Sequence numbers XM\_010923292, XM\_010933391, XM\_010916199, XM\_010920466, XM\_010934797, XM\_010937466 and XM\_010922209.

The list of the oil palm ethylene receptor genes and their sequence characteristics is shown in Table 2. The identified ethylene receptor genes have nucleotide lengths that vary from 2204 base pair (bp) to 4186 bp. The size of the predicted polypeptides ranged from 629 aa to 757 aa with EgERS1 being the smallest with a predicted molecular weight of 70.04 kDa and EgETR3, the largest polypeptide of 85.20 kDa. The calculated PI ranged from 6.31 to 7.33, indicating that four of the ethylene receptor proteins may function better in a slightly acidic environment while the rest prefer a neutral working environment. The instability index (II) of the ethylene receptors varied from 37.49 to 50.79. EgERS2 and EgETR5 have II values above 40 and are predicted to be unstable due to the presence of certain dipeptides which can also be used as a measure of the *in vivo* half-life of the protein (Guruprasad *et al.*, 1990). The aliphatic index (AI) value of the ethylene receptors ranged from 97.58 to 111.90. AI indicates the thermostability of proteins where EgERS3 was found to be the most thermostable with the highest AI of 111.90. The grand average of hydrophobicity (GRAVY) values of the ethylene receptor genes varied from 0.042 to 0.165, indicating that all of the proteins have more hydrophobic features. The subcellular localisation of

all seven proteins were predicted to be extracellular, all of which are predominantly located in the endoplasmic reticulum.

### Isolation of the Ethylene Receptor Full-length Transcripts

RACE procedures were carried out to isolate and verify the 5' and 3' ends of the oil palm ethylene receptor genes and hence, to obtain the full-length sequence information (Figures 1 and 2). For 5' RACE, the size of the amplified PCR products ranged from 194 bp to 697 bp (Table 3). The amplified PCR products consisted of the 5' open reading frame (ORF) including the start codon and the 5' untranslated regions (UTR). The sizes of the 5' UTR of the seven ethylene receptor genes ranged from 169 bp to 493 bp. Manual sequence alignment and exon search of the 5' RACE PCR products against the oil palm P5 build showed a 100% identity at the nucleotide level. In addition, introns were detected in almost all the ethylene receptor genes except *EgETR5*. The largest intron size of approximately 6238 bp was found in *EgERS1* whereas the smallest intron of 635 bp was found in *EgETR3*. The intron-exon boundaries follow the universal intron 'GT-AG' rule (Breathnach and Chambon, 1981). Sequence analysis of all seven ethylene receptor genes showed that nucleotides flanking the ATG translation start site are adequate to the Kozak consensus sequences (ANNATGG), which is important to enhance the initiation of the translation process (Nakagawa *et al.*, 2008).

For 3' RACE, the size of the amplified PCR products ranged from 351 bp to 1493 bp (Table 3). The amplified PCR products contained the 3' ORF region including the stop codon and the 3' UTR. In addition, the poly(A)<sup>+</sup> tail was detected in all fragments, confirming that the 3' ends of the mRNA were successfully obtained. The plant consensus putative polyadenylation signal (AATAAA) involved in facilitating the addition of poly(A)<sup>+</sup> tail to produce the mature messenger mRNA (Joshi,

1987) was found in the 3' UTR in five of the ethylene receptor genes, except *EgERS1* and *EgETR3*. Both *EgERS1* and *EgETR3*, however, contained an adequate base match to the plant consensus signal found in the 3' UTR region. Sequence alignment of the amplified PCR products to the oil palm genome showed 100% similarity at the nucleotide level. No introns were detected within the 3' UTR in any of the ethylene receptor genes.

LD-PCR was performed to amplify the far end of the 5' and 3' regions of the seven ethylene receptor genes. The final assembled sequence from the 5', 3', LD-PCR products, oil palm transcriptome, gene model and genome data indicated that the full-length cDNA encoding the ethylene receptor genes were obtained (Table 4). An illustration of their genomic structures is shown in Figure 3. Sequence analyses also showed that the assembled sequences conform to the corresponding predicted genes in

Genbank, except for *EgETR2*. Sequencing analysis of the PCR products showed that *EgETR2* encoded a polypeptide of 687 aa whereas XM\_010920466 is 756 aa in size. Sequence comparison between these two nucleotide sequences showed that there is possibly an indel mutation in XM\_010920466. One nucleotide is missing from the sequence in XM\_010920466 which possibly have caused a frameshift mutation, producing a polypeptide that is 69 aa residues longer than *EgETR2*. However, this could also be due to an error or mismatch in the sequence read which can occur during genome sequence assembly (Pop, 2009). *EgERS1* is also 99.8% and 99.95% identical to *EgERD3* (Nurniwalis *et al.*, 2018) at the protein and nucleotide levels, respectively. A single aa change at location 605 in *EgERS1* (R residue) and in *EgERD3* (K residue) was caused by a base transition of G-A. The missense mutation in *EgERS1*-*EgERD3* may or may not affect the function of the protein.

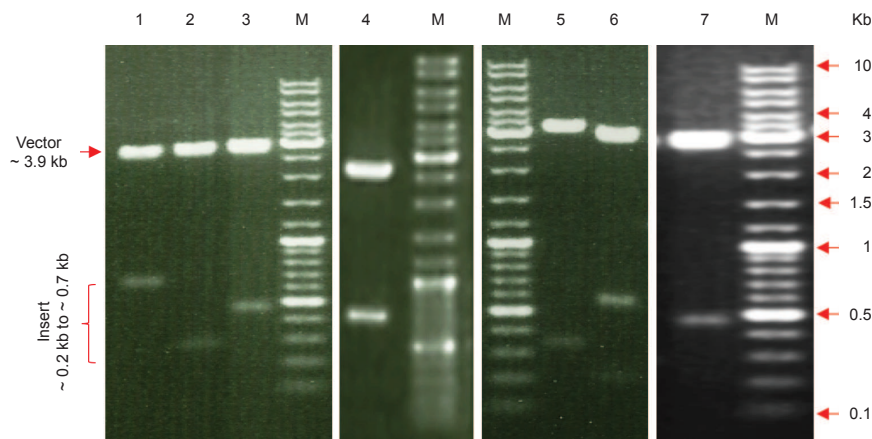


Figure 1. Analysis of the digested plasmids carrying the 5' rapid amplification of cDNA and polymerase chain reaction (RACE PCR) products by electrophoresis on 1% (w/v) agarose gel. The plasmids were digested with *EcoRI*. Lanes 1 to 7 represent the digested products of *pEgERS1*, *pEgERS2*, *pEgERS3*,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

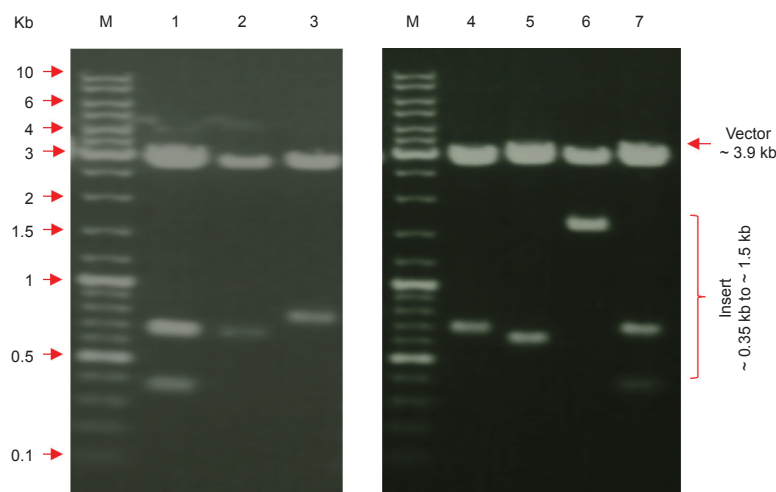


Figure 2. Analysis of the digested plasmids carrying the 3' rapid amplification of cDNA and polymerase chain reaction (RACE PCR) products on 1% (w/v) agarose gel. The plasmids were digested with *EcoRI*. Lanes 1 to 7 represent the digested products of *pEgERS1*, *pEgERS2*, *pEgERS3*, *pEgETR2*, *pEgETR3*, *pEgETR4* and *pEgETR5* plasmids, respectively. M represents Gene Ruler DNA Ladder Mix (Fermentas).

**TABLE 3. SIZE CHARACTERISTICS OF THE AMPLIFIED RAPID AMPLIFICATION OF cDNA END (RACE) POLYMERASE CHAIN REACTION (PCR) PRODUCTS AND INTRONS IN THE UNTRANSLATED REGION (UTRS) OF OIL PALM ETHYLENE RECEPTOR GENES**

Gene name/ID	Amplified 5' PCR size (bp)			Intron size within 5' UTR (bp)	Amplified 3' PCR size (bp)			Intron size within 3' UTR (bp)
	ORF	5' UTR	Total		ORF	3' UTR	Total	
<i>EgERS1</i>	420	177	697	6 238	714	265	1 007	None
<i>EgERS2</i>	23	171	194	727	338	287	653	None
<i>EgERS3</i>	244	169	413	768	199	406	605	None
<i>EgETR2</i>	290	433	723	882	292	307	623	None
<i>EgETR3</i>	74	493	567	635	101	222	351	None
<i>EgETR4</i>	72	234	306	700	151	1 314	1 493	None
<i>EgETR5</i>	166	174	197	None	163	768	931	None

Note: ORF - open reading frame; bp - base pair.

**TABLE 4. LIST OF THE FULL-LENGTH SEQUENCE INFORMATION OF THE OIL PALM ETHYLENE RECEPTOR GENES**

Gene ID	Total FL size (bp)	ORF (bp)	5' UTR (bp)	3' UTR (bp)	Poly A tail (bp)
<i>EgERS1</i>	2 360	1 890	177	265	28
<i>EgERS2</i>	2 394	1 908	171	287	28
<i>EgERS3</i>	2 507	1 908	169	406	24
<i>EgETR2</i>	2 728	2 064	433	307	27
<i>EgETR3</i>	3 007	2 274	493	222	28
<i>EgETR4</i>	3 844	2 268	234	1 314	28
<i>EgETR5</i>	3 177	2 214	174	768	21

Note: FL - full-length; bp - base pair; ORF - open reading frame; UTR - untranslated region.

**TABLE 5. NUCLEOTIDE (open reading frame) AND AMINO ACID (protein) CORRELATION BETWEEN SEVEN OIL PALM ETHYLENE RECEPTORS WITH BL2SEQ, BIOLOGY WORKBENCH (matrix selection parameter: BLOSUM62, gap opening penalty: 11, gap extension penalty: 1 and lambda ratio: 0.5)**

Gene ID	Protein ID	ORF size (bp)	Protein size (aa)	Nucleotide coding region identity (%)						
				<i>EgERS1</i>	<i>EgERS2</i>	<i>EgERS3</i>	<i>EgETR2</i>	<i>EgETR3</i>	<i>EgETR4</i>	<i>EgETR5</i>
Amino acid identity (%)	<i>EgERS1</i>	1 890	629	-	80	80	48	48	49	43
	<i>EgERS2</i>	1 908	635	82	-	90	49	49	49	44
	<i>EgERS3</i>	1 908	635	86	88	-	50	49	49	44
	<i>EgETR2</i>	2 064	687	37	37	37	-	88	76	53
	<i>EgETR3</i>	2 274	757	38	36	37	84	-	75	54
	<i>EgETR4</i>	2 268	755	38	37	38	71	72	-	53
	<i>EgETR5</i>	2 214	737	34	33	33	45	45	45	-

Note: The upper right region of the table represents the % identity of nucleotide coding sequence between two ethylene receptor genes whereas the lower left bottom represents the % identity of amino acid (aa) sequence between two ethylene receptor proteins. bp - base pair.

Similarity comparison amongst the ORF of the ethylene receptor genes showed that pairwise similarity ranged from 43% to 90% (Table 5). The lowest similarity was found between *EgERS1* and *EgETR5* and the highest similarity was between *EgERS2* and *EgERS3*. As for the aa sequences, the pairwise similarities ranged from 33% to 88%. Unlike the ORF, the lowest aa sequence similarity is between *EgERS2*-*EgETR5* and *EgERS3*-*EgETR5* while the highest aa sequence similarity was between *EgERS2*-*EgERS3*, as the nucleotide pairwise similarity of the ORF.

### *In silico* Analyses of Conserved Domain

Five major domains spanning from the N-terminal to the C-terminal were detected and comprised of transmembrane, GAF, histidine kinase A (HisKA), histidine kinase-like ATPase (HATPase\_c) and REC (cheY-homologous receiver)

domains (Figure 4). However, not all five conserved domains are present in all of the ethylene receptor proteins. *EgETR3* and *EgETR4* contain all five conserved domains whereas the rest of the proteins contain only four. *EgERS1*, *EgERS2*, *EgERS3* and *EgETR2* consist of the same conserved regions (transmembrane, GAF, HisKA and HATPase\_c) whereas *EgETR5* lacks the HisKA domain but contains the REC domain instead.

At the N-terminal of the ethylene receptor proteins, the transmembrane domain consists of three transmembrane segments (Figure 4). Hydropathy analysis using Kyte-Doolittle hydropathy profile suggests that the triple transmembrane segments are hydrophobic. Additionally, a fourth transmembrane segment or a signal peptide was detected at the far-end N-terminal of *EgETR2*, *EgETR3* and *EgETR4* (Figure 4). The GAF domain was detected next to the transmembrane domain and is present in all seven ethylene receptor proteins (Figure 4).

Following the GAF domain is the histidine kinase domain which is comprised of HisKA and HATPase\_c (Figure 4). Multiple aa sequence alignment showed that apart from the conserved transmembrane and GAF domains, the histidine kinase domain is also conserved amongst EgERS1, EgERS2 and EgERS3 (Figure 5). Five conserved motifs responsible for the histidine kinase activity found in bacteria protein kinase (Parkinson and Kofoed, 1992) are also present in all three receptors. The H, N, G1, F and G2 motifs are arranged in the same order to the bacterial histidine kinases with similarities ranging from 67% to 100%. The putative autophosphorylated histidine, H residue, in the histidine kinase domain is also present in

EgERS1, EgERS2 and EgERS3. In contrast, multiple sequence alignment of EgETR2, EgETR3, EgETR4 and EgETR5 showed that the conserved domains are more diverged especially within the histidine kinase domain (Figure 6). Only motifs H and N can be detected in all four ethylene receptors. Motifs G1, F and G2 could not be detected as the amino acids are more diverged than the consensus motif. In the H motif, the putative autophosphorylated H residue is absent and replaced by arginine, R; in EgETR5 and glutamine, Q; in EgETR2, EgETR3 and EgETR4. Sequence in the N motif in all four ethylene receptors is similar to the N motif in bacterial protein kinases. However, only one asparagine, N is present while the other N residue is absent.

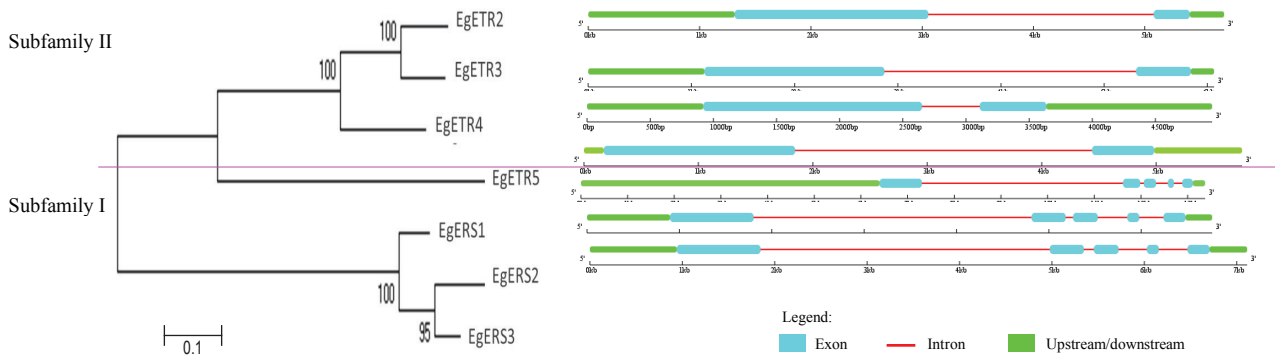


Figure 3. Schematic representation of oil palm ethylene receptor genomic structures and the division of oil palm ethylene receptor gene family into two subfamilies. The bar indicates 0.1 substitutions per site.

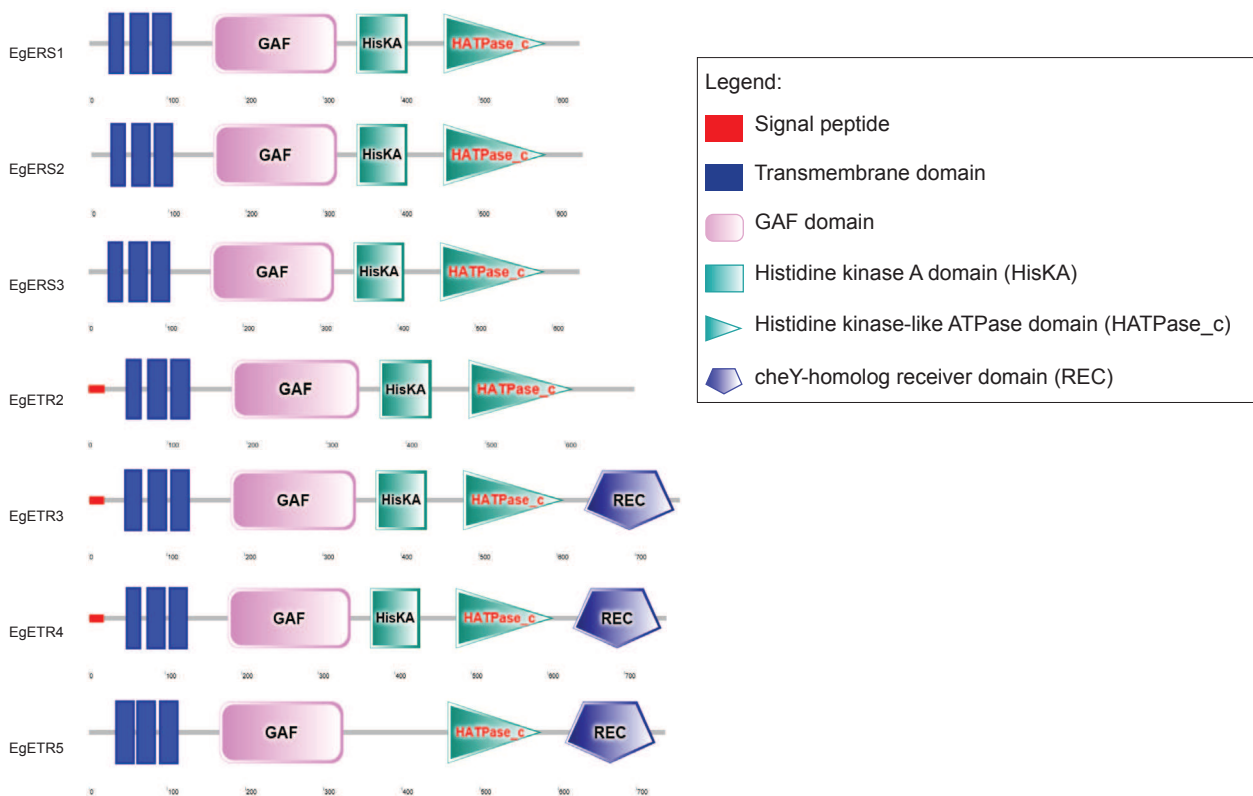


Figure 4. Primary protein structure of oil palm ethylene receptor proteins using SMART software (<http://smart.embl-heidelberg.de/>).

The REC domain is only found at the C-terminal of EgETR3, EgETR4 and EgETR5 (Figure 4). Multiple sequence alignment showed that the conserved residues important for the function of a receptor is present within this domain (Figure 6). Two aspartates, D and a lysine, K conserved residues that correspond to the conserved residues in the bacterial receiver domain are present within the REC domain in EgETR3, EgETR4 and EgETR5.

Protein sequence correlation between the seven ethylene receptors is shown in Table 5. Except EgETR2, the ethylene receptors without receiver domains at the C-terminal share higher similarities ranging from 82% to 88% with lowest similarity between EgERS1 and EgERS2 and highest between EgERS2 and EgERS3. In comparison, the ethylene receptors with receiver domains (EgETR3, EgETR4 and EgETR5) have lower similarities ranging from 45% to 72%. The highest similarity amongst the ethylene receptors containing a receiver domain is between EgETR3 and EgETR4 and the lowest similarity is between EgETR5-EgETR3 and EgETR5-EgETR4. Ethylene receptors without a receiver domain also have lower aa sequence similarity to those containing a receiver domain ranging from 33% to 45%.

### Phylogenetic Analysis

Phylogenetic relationship analysis using multiple sequence alignment of all seven oil palm ethylene receptors generated a tree with two branches (Figure 3). The first branch consists of EgERS1, EgERS2 and EgERS3 while the second branch consists of EgETR2, EgETR3, EgETR4 and EgETR5. EgETR5, however, is more distinct compared to EgETR2, EgETR3 and EgETR4. The phylogenetic tree was also constructed to analyse the relationship between all seven oil palm ethylene receptors and other plant ethylene receptors (Figure 7). The resulting tree classified the ethylene receptor proteins into two subfamilies, which conforms with the general pattern of ethylene receptor family in plants (Hua *et al.*, 1998). Within each subfamily, subgroups were also observed where each subfamily contains at least three subgroups and each subgroup has at least one clade (Figure 7). Each respective subgroup in both subfamilies contains ethylene receptors that are only present in either monocotyledons or dicotyledons. All seven ethylene receptors from oil palm are grouped in the same clade as the ethylene receptors from date palm. The sequence homology between ethylene receptors of both species is high, ranging from 87% to 94% identity.

### Gene Structure, Chromosomal and Genetic Map Localisation

Gene structure analysis based on the full-length ethylene receptor genes and their corresponding oil

palm genome scaffolds showed a similar grouping pattern with the phylogenetic analysis. Within the ORF, EgERS1, EgERS2 and EgERS3 contain five exons and four introns with the first and fourth exons having the largest (906 bp) and smallest (128 bp) size, respectively (Figure 3). On the other hand, EgETR2, EgETR3, EgETR4 and EgETR5 contain only two exons and one intron. The first exon represents the largest exon with sizes ranging from 1671 bp to 1743 bp. The similarity between members of each group is also higher (Table 5). Distribution of the exon-intron structure resembles members within the subfamilies of the ethylene receptor family in other plants. The subfamily I members normally contain at least four introns and are flanked by five or more exons whereas subfamily II members have only one intron within the ORFs (Hua *et al.*, 2000; Tieman *et al.*, 2000; Yau *et al.*, 2004). Therefore, EgERS1, EgERS2 and EgERS3 are categorised as subfamily I members of the oil palm ethylene receptor family while the remaining ethylene receptors belong to subfamily II.

Exonerate search on the P5 build showed that majority of the seven ethylene receptor genes are mapped on different chromosomes. EgETR5 and EgERS1 are found on chromosome 5 while the rest of the genes are confined to a single chromosome (Table 6). In addition, genetic markers were also used to locate the position of the ethylene receptor genes on T128 and P2 genetic linkage maps (Ting *et al.*, 2014). Results revealed that the distribution pattern of the seven ethylene receptor 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 (Table 6).

### Expression Analysis via RT-PCR

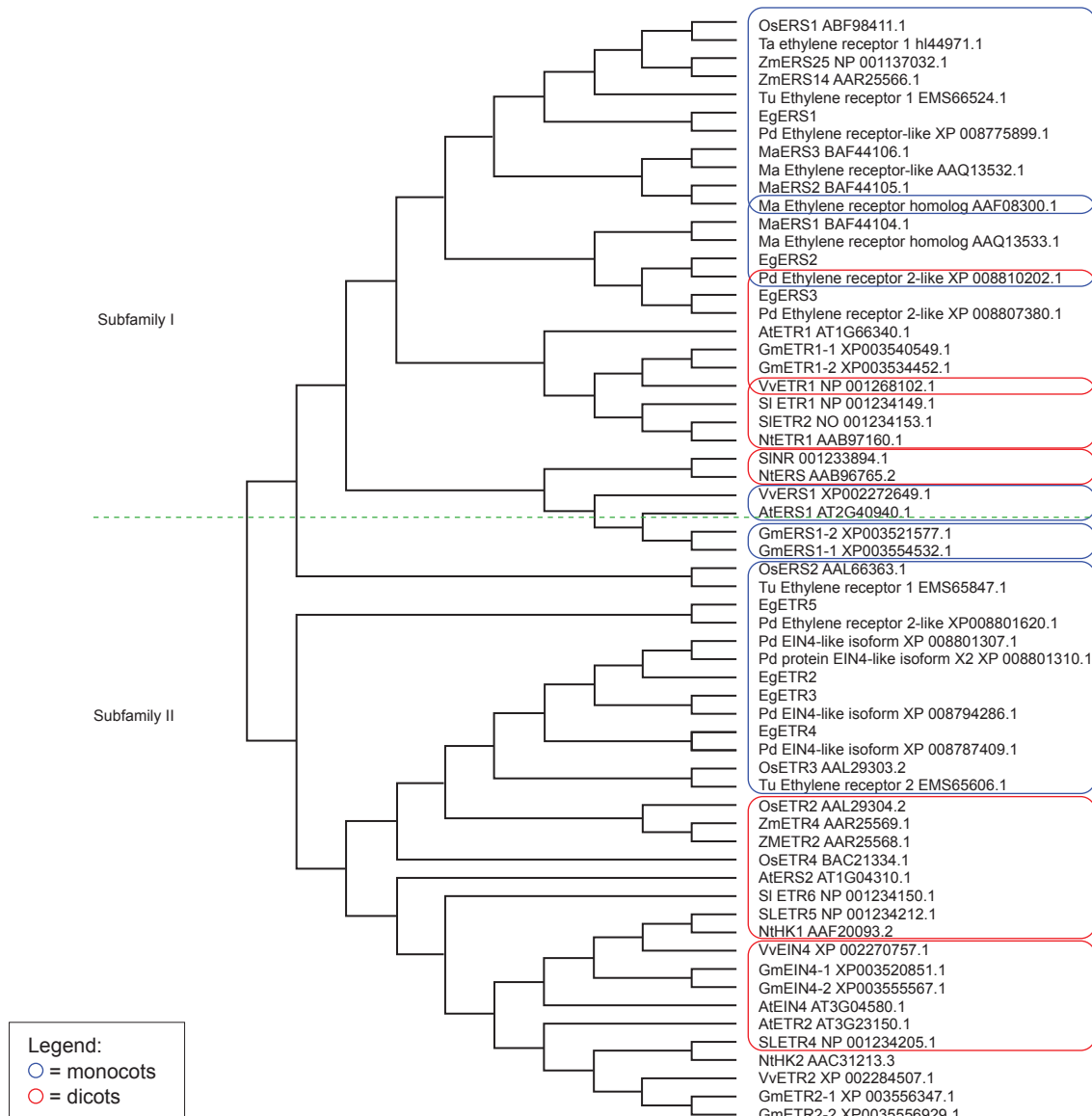
The expression pattern of the ethylene receptor genes was initially analysed in mesocarp tissues at various developmental stages. All seven ethylene receptor genes were found to be expressed in the mesocarp from young until the fruit ripens but the expression level varies (Figure 8). Four of the genes (EgETR2, EgETR3, EgERS1 and EgERS3) showed an increase in expression as the fruit develops, mature and ripens. In EgETR3, expression in young and developing fruits between 8 weeks after anthesis (WAA) to 12 WAA is low but the increase in expression is apparent at 15 WAA and is highest at 20 WAA. Similarly, expression pattern of EgETR2 and EgERS3 corresponds to that of EgETR3 except at 8 WAA, where the expression is slightly higher than at 10 and 12 WAA. For EgERS1, a slight increase in expression was detected in mesocarp tissues as the fruits develop from 8 WAA and reached its highest at 17 WAA followed by a decrease in expression at 20 WAA. In contrast, the remaining three genes (EgETR4, EgETR5 and EgERS2) exhibited a different





EgETR5	HL <b>S</b> IVSSRS <b>H</b> CLEMLY <b>L</b> KRNQ <b>F</b> HL <b>L</b> LID <b>L</b> QIF <b>E</b> EDRH <b>E</b> LSA <b>H</b> IK <b>N</b> ICTENR <b>P</b> L <b>I</b> VAL <b>T</b> PD	693
EgETR4	CVSSVSSGNQ <b>C</b> LR <b>C</b> LG <b>T</b> SG <b>T</b> PF <b>Q</b> LV <b>L</b> LD <b>I</b> NMPDMNGFEVAMR <b>I</b> Q <b>N</b> FRSGCWPL <b>T</b> VAL <b>T</b> AS	717
EgETR2	SVSSVASG <b>I</b> E <b>C</b> LN <b>S</b> F <b>W</b> QCCY <b>T</b> FP <b>T</b> R <b>N</b> HG-----P-----	687
EgETR3	SVSSV <b>T</b> SG <b>I</b> Q <b>C</b> LS <b>S</b> FG <b>T</b> SM <b>H</b> F <b>Q</b> LV <b>M</b> V <b>D</b> L <b>H</b> MP <b>K</b> MD <b>G</b> FEVAMR <b>I</b> R <b>K</b> FR <b>S</b> RC <b>W</b> PS <b>I</b> VAL <b>T</b> AS	719
	: * * : * . ** : * .	
EgETR5	THK <b>N</b> TRE <b>Q</b> CL <b>Q</b> DGM <b>H</b> GV <b>M</b> C <b>K</b> PV <b>I</b> L <b>Q</b> EM <b>V</b> DEL <b>Q</b> RIT <b>Q</b> GT <b>Q</b> SS <b>L</b> PL	737
EgETR4	VEED <b>T</b> WE <b>K</b> CV <b>Q</b> VG <b>M</b> SG <b>L</b> IR <b>K</b> PV <b>L</b> LL <b>H</b> VL <b>K</b> EEL <b>F</b> RV <b>L</b> Q <b>N</b> T-----	755
EgETR2	-----	687
EgETR3	AED <b>D</b> V <b>W</b> DR <b>C</b> L <b>Q</b> CG <b>I</b> NG <b>L</b> IR <b>K</b> PV <b>T</b> L <b>Q</b> T <b>L</b> A <b>E</b> EL <b>H</b> RV <b>L</b> Q <b>N</b> T-----	757

Figure 6. Amino acid multiple sequence alignment of ethylene receptor genes comprising of EgETR2, EgETR3, EgETR4 and EgETR5 of subfamily II from oil palm ethylene receptor family. 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 Omega, <https://www.ebi.ac.uk/Tools/msa/clustalo>). The four transmembrane domains at the amino terminal are highlighted in red and labelled I, II, III and IV. The GAF domain is boxed in black. The histidine kinase domain region is boxed in blue. The conserved motifs H and N in the histidine kinase domain found in bacteria protein kinase are highlighted in blue. The receiver domain is highlighted in green and the conserved residues (D, D and K) in this domain are mark in red.



Note: 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*.

Figure 7. 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.

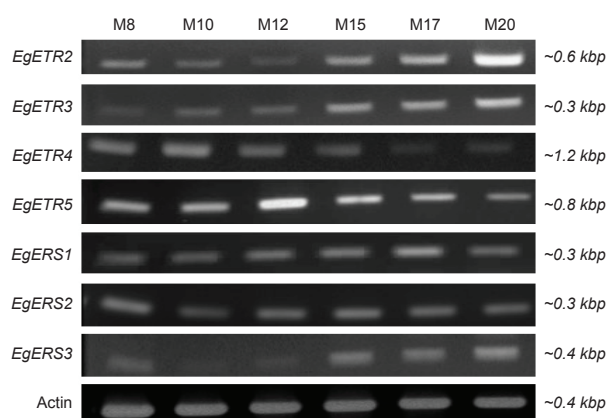


Figure 8. Expression profile of oil palm ethylene receptor genes throughout fruit development in oil palm mesocarp tissues via reverse transcription-polymerase chain reaction (RT-PCR). Actin was used to show approximate expression throughout the various tissues. M - mesocarp while numbers represent week after anthesis (WAA).

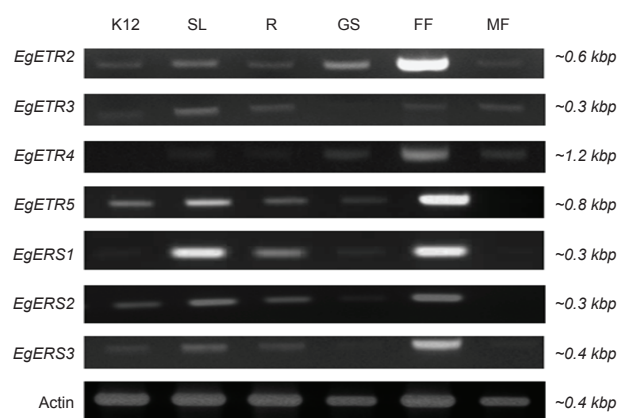


Figure 9. Expression profile of oil palm ethylene receptor genes in various oil palm tissues via reverse transcription-polymerase chain reaction (RT-PCR). Actin represents approximate expression throughout the various tissues. K12 - kernel at 12 weeks after anthesis, SL - spear leaves, GS - germinated seedlings, R - roots, FF - female flower and MF - male flower.

TABLE 6. LOCATION OF THE ETHYLENE RECEPTOR GENES ON OIL PALM CHROMOSOMES AND GENETIC LINKAGE GROUPS

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

expression pattern. *EgETR4* showed an opposite trend where the gene expression profile in mesocarp tissues decreased as the fruits develop, mature and ripens. *EgETR5* and *EgERS2*, on the other hand, maintained an almost consistent expression level in the all mesocarp tissues throughout fruit development from young until the ripening stages.

In addition to mesocarp tissues, expression of the ethylene receptor genes was also determined in other oil palm tissues. This included kernel at 12 WAA, spear leaves, roots, germinated seedlings, female and male flowers. As shown in Figure 5, the ethylene receptor genes showed differential gene expression profiles in the tested tissues. Overall, expression of the ethylene receptor genes was low, some of which was only at a detectable level. The highest expression was found in female flower for all genes except *EgETR3*. Three genes (*EgERS1*, *EgERS2* and *EgERS3*) from subfamily I had similar expression patterns but at varying levels. The genes were expressed in all the tested tissues except in the male flower. Subfamily II members (*EgETR2*, *EgETR3*, *EgETR4* and *EgETR5*) on the other hand, exhibited a different expression pattern with expression localised to specific tissues. *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.

## DISCUSSION

Previously, a member of the ethylene receptor gene family was isolated and characterised in oil palm (Nurniwalis *et al.*, 2018). The gene corresponds to an ERS-type ethylene receptor and was expressed in the mesocarp tissues from young until the ripening stages, indicating an important role of ethylene in regulating oil palm fruit development and ripening processes. With the accessibility to the oil palm genome, we have now isolated an additional six members of the oil palm ethylene receptor family. In total, seven (*EgETR2*, *EgETR3*, *EgETR4*, *EgETR5*, *EgERS1*, *EgERS2* and *EgERS3*) ethylene receptor genes were identified from the oil palm genome. These receptors were identified by homology search and a HMM profile build based on known and established ethylene receptors from various plant species as well as searches on the oil palm transcriptome data, gene models (Singh *et al.*, 2013a) and the previously isolated *EgERD3* gene (Nurniwalis *et al.*, 2018). Sequence alignment analysis showed that *EgERS1* match to *EgERD3* with 99.8% and 99.95% identity at the nucleotide and protein level. Subsequently, identification of the genes has led to the isolation and characterisation of their full-length sequence.

The oil palm ethylene receptor family consist of seven members which are classified into two subfamilies. Classification of the subfamilies was based on intron-exon structure of the coding sequences, protein conserved domains, their primary structure and phylogenetic analyses. Subfamily I consist of three members (EgERS1, EgERS2 and EgERS3) whereas subfamily II has four members (EgETR2, EgETR3, EgETR4 and EgETR5). Members in subfamily I have higher nucleotide and aa sequence similarities ranging from 80% to 90% and 82% to 88% compared to members of subfamily II. The highest similarity at the nucleotide and aa level in subfamily II members is 88% and 84% while the lowest is 53% and 45%. Between subfamily I and II, the aa sequence similarities ranged between 33% to 38%, respectively.

The oil palm ethylene receptor has a primary structure similar to the bacterial two-component receptor which consist of the transmembrane, GAF and histidine kinase domains. Three members of subfamily II contain a receiver domain at the C-terminal, which is the second domain of the two-component signaling system in bacteria (Parkinson and Kofoid, 1992). The most conserved protein domain in all the receptors is the transmembrane domain especially the three transmembrane segments at the N-terminal. These transmembrane segments contain aa residues that are important for ethylene-binding (Schaller and Bleeker, 1995). In contrast, the histidine kinase domain especially in subfamily II members, are more diverse in their sequence similarities to the bacterial histidine kinases (Parkinson and Kofoid, 1992). The primary structure of oil palm ethylene receptor subfamilies is also similar in other plants. Despite the similarity and diversity in their sequence within the subfamilies, mutation studies in *Arabidopsis* and tomato showed that all members of the ethylene receptor subfamilies play a role regulating ethylene responses (Chang *et al.*, 1993; Hua *et al.*, 1998; Sakai *et al.*, 1998; Tieman and Klee, 1999).

In plants, the most exhaustively studied ethylene receptor is *AtETR1* from *Arabidopsis* (Agarwal *et al.*, 2012; Gallie, 2015; Chang, 2016). *AtETR1* has a characteristic of a subfamily I member but it also contains a response regulator domain. This feature is not found in any of the oil palm ethylene receptors. Similarly, the ETR1-type ethylene receptor is also not found in other monocots (Yau *et al.*, 2004; Ma and Wang, 2003; Gallie and Young, 2004; Al-Mssallem *et al.*, 2013). Evolution studies of plant ethylene receptors demonstrated that subfamily I receptors in monocots, lack a receiver domain while subfamily II contains the receiver domain (Gallie, 2015). Phylogenetic relationship between oil palm and other plant ethylene receptors also demonstrated that both subfamily I and II members are divided into sub-groups where those from monocots are

segregated from dicots. The separation between the ethylene receptor gene family in monocots and dicots is likely as a result of receptor evolution from cyanobacteria into angiosperms that took place about 100 million years ago (Binder *et al.*, 2012). Subfamily I members are more conserved whereas subfamily II members are more diverse especially in the histidine kinase domain (Binder *et al.*, 2012; Gallie, 2015). In subfamily II, the existence of ethylene receptors without a receiver domain such as EgETR2 in oil palm and AcETR2 in pineapple (Li *et al.*, 2016) suggest that the ethylene receptor family is continuously evolving in a specific manner (Gallie, 2015).

Gene expression analyses of the seven ethylene receptors via RT-PCR indicated that the receptors are differentially expressed in various oil palm tissues including reproductive and vegetative tissues. Expression of the receptors appeared to be higher in reproductive tissues compared to vegetative tissues. Similar evidence has been reported in many plants (Tieman and Klee, 1999; Li *et al.*, 2016) where ethylene receptors are involved in the development of reproductive organs and promoting flower development. More ethylene is produced in female flowers than male flowers for proper flower development (Manzano *et al.*, 2014; 2011; Martinez *et al.*, 2013) where reduction in ethylene production in zucchini female flower can lead to the development of parthenocarpic fruits (Martinez *et al.*, 2013). In oil palm fruits, the ripening process involves changes in fruit colour, softening of tissue and high oil accumulation that occurs between 20-24 WWA depending on the genotype (Mohd Hudzairi *et al.*, 2012). The varying expression of the ethylene receptor genes in the mesocarp tissues suggest their involvement in the two-system control of ethylene production in climacteric fruits. The high expression of the receptors at 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). Since oil palm is a climacteric fruit, the ripening-related gene expression is likely to be the consequence of increased ethylene production in ripening mesocarp tissues (Henderson and Osborne, 1999; Tranbarger *et al.*, 2011; Nurniwalis, 2017). Expression of the oil palm ethylene receptors in vegetative tissues also demonstrated that ethylene is likely involved in almost every stage of oil palm's life cycle. The various expression profiles of the ethylene receptors in various oil palm tissues may also be due to the overlapping and non-overlapping roles of the receptors in modulating ethylene signals as shown in *Arabidopsis* and tomato (Shakeel *et al.*, 2013; Kevany *et al.*, 2008).

Studies on modulating and altering ethylene sensitivity of plants using ethylene receptor have been reported (Czarny *et al.*, 2006; Agarwal *et al.*, 2012, Shakeel *et al.*, 2013). Several approaches for manipulating and regulating ethylene sensitivity for crop improvement have been described including genetic engineering (Agarwal *et al.*, 2012; Csukasi *et al.*, 2009). A cross-talk between ethylene and other plant hormones to regulate various plant growth and development processes is also evident (Kumar *et al.*, 2014; Iqbal *et al.*, 2017). Therefore, together with the oil palm ethylene receptor family, it is envisaged that manipulation of ethylene signaling pathway using molecular tools and/in combination with hormone regulation and interactions would be of great importance to the oil palm industry. The ability to control and modulate specific ethylene responses in oil palm would be desirable especially to help improve yield.

#### ACKNOWLEDGEMENT

We thank the Director-General of MPOB for permission to publish this article. We also would like to acknowledge MPOB for the Ph.D scholarship to Nurniwalis Abdul Wahab at the University of Nottingham Malaysia. Special thanks to Aminah Shuib, Dr Leslie Low, Dr Ting Ngoot Chin and the staff in the Gene Function Laboratory, MPOB for their technical help and assistance. We also acknowledge the Breeding and Tissue Culture Unit, MPOB for the supply of oil palm tissues, fruit bunches and total RNA. A special thanks are also due to Dr Ooi Siew Eng and Dr Abrizah Othman for constructive comments and suggestions for the article.

#### REFERENCES

- Abeles, F; Morgan, P and Saltveit, MJ (1992). *Ethylene in Plant Biology*, Academic Press, San Diego, CA. p. 1-414.
- Agarwal, G; Choudhary, D; Singh, V P and Arora, A (2012). Role of ethylene receptors during senescence and ripening in horticultural crops. *Plant Signal Behav.*, 7(7): 827-846.
- Alexander, L and Grierson, D (2002). Ethylene biosynthesis and action in tomato: A model for climacteric fruit ripening. *J. Exp. Bot.*, 53(377): 2039-2055.
- Al-Mssallem, I S; 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; Abdullah, M M B; Huang, D; Fang, Y; Alnakhli, Y O; Jia, S; Yin, A; Alhuzimi, E M; Alsaihati, B A; Al-Owayyed, S A; Zhao, D; Zhang, S; Al-Otaibi, N A; Sun, G; Majrashi, M A; Li, F; Wang, J; Yun, Q; Alnassar, N A; Wang, L; Yang, M; Al-Jelaify, R F; Liu, K; Gao, S; Chen, K; Alkhaldi, S R; Liu, G; Zhang, M; Guo, H and Yu, J (2013). Genome sequence of the date palm *Phoenix dactylifera* L. *Nat. Comm.*, 4: 2274.
- Altschul, S F; Madden, T M; Schaffer, A A; Zhang, J; Zhang, Z; Miller, W and Lipman, D J (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.*, 25(17): 3389-3402.
- Binder, B M; Chang, C and Schaller, G E (2012). Perception of ethylene by plants: Ethylene receptors. *Annual Plant Reviews Volume 44: The Plant Hormone Ethylene* (McManus, M T ed.). Wiley-Blackwell, Oxford, United Kingdom. p. 117-145.
- Breathnach, R and Chambon, P (1981). Organisation and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.*, 50(1): 349-383.
- Chan, K L; Tatarinova, T V; Rosli, R; Amiruddin, N; Azizi, N; Ab Halim, M A; Nik Mohd Sanusi, N S; Jayanthi, N; Ponomarenko, P; Triska, M; Solovyev, V; Raih, M F; Sambanthamurthi, S; Murphy, D and Low, L E T (2017). Evidence-based gene models for structural and functional annotations of the oil palm genome. *Biol Direct.*, 12: 21.
- Chang, C (2016). Q&A: How do plants respond to ethylene and what is its importance? *BMC Biol.*, 14: 7.
- Chang, C; Kwok, S F; Bleecker, A B and Meyerowitz, E M (1993). Arabidopsis ethylene-response gene ETR1: Similarity of product to two-component regulators. *Science*, 262(5133): 539-544.
- Csukasi, F; Merchante, C and Valpuesta, V (2009). Modification of plant hormone levels and signaling as a tool in plant biotechnology. *Biotechnol. J.*, 4: 1293-1304.
- Czarny, J C; Grichko, V P and Glick, B R (2006). Genetic modulation of ethylene biosynthesis and signaling in plants. *Biotechnol Adv.*, 24(4): 410-419.
- Eddy, S R (2011). Accelerated profile HMM searches. *PLoS Comput Biol.*, 7(10): e1002195.
- Gallie, D R and Young, T E (2004). The ethylene biosynthetic and perception machinery is differentially expressed during endosperm and embryo development in maize. *Mol. Genet. Genomics*, 271(3): 267-281.
- Gallie, D R (2015). Ethylene receptors in plants - Why so much complexity? *F1000Prime Reports*, 7: 1-12.

- Guruprasad, K 1; Reddy, B V and Pandit, M W (1990). Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting *in vivo* stability of a protein from its primary sequence. *Protein Eng.*, 4(2): 155-161.
- Henderson, J and Osborne, D J (1999). Ethylene as the initiator of the inter-tissue signaling and gene expression cascades in ripening and abscission of oil palm fruit. *Biology and Biotechnology of the Plant Hormone Ethylene II* (Kanellis, A K ed.). Kluwer Academic Publishers, Doedrecht, Netherlands. p. 129-136.
- Hua, J and Meyerowitz, E M (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, Q G; Bleecker, A B; Ecker, J R and Meyerowitz, E M (1998). EIN4 and ERS2 are members of the putative ethylene receptor gene family in *Arabidopsis*. *Plant Cell*, 10(8): 1321-1332.
- Iqbal, N; Khan, N A; Ferrante, A; Trivellini, A; Francini, A and Khan, M I R (2017). Ethylene role in plant growth, development and senescence: Interaction with other phytohormones. *Front. Plant Sci.*, 8: 475.
- Joshi, C P (1987). Putative polyadenylation signals in nuclear genes of higher plants: A compilation and analysis. *Nucleic Acids Res.*, 15(23): 9627-9640.
- Kevany, B M; Taylor, M G and Klee, H J (2008). Fruit-specific suppression of the ethylene receptor *LeETR4* results in early-ripening tomato fruit. *Plant J.*, 6(3): 295-300.
- Kumar, R; Khurana, A and Sharma, A K (2014). Role of plant hormones and their interplay in development and ripening of fleshy fruits. *J. Exp. Bot.*, 65(16): 4561-4575.
- Li, Y-H; Wu, Q-S; Huang, X; Liu, S-H; Zhang, H-N; Zhang, Z and Sun, G-M (2016). Molecular cloning and characterisation of four genes encoding ethylene receptors associated with pineapple (*Ananas comosus* L.) flowering. *Front. Plant Sci.*, 7: 710.
- Lin, Z; Zhong, S and Grierson, D (2009). Recent advances in ethylene research. *J. Exp. Bot.*, 60(12): 3311-3336.
- Ma, Q H and Wang, X M (2003). Characterisation of an ethylene receptor homologue from wheat and its expression during leaf senescence. *J. Exp. Bot.*, 54(386): 1489-1490.
- Manzano, S; Juan, C M; García, M; Megías, Z and Jamilena, M (2014). Involvement of ethylene in sex expression and female flower development in watermelon (*Citrullus lanatus*). *Plant Physiol. Biochem.*, 85: 96-104.
- Manzano, S; Martínez, C; Megías, Z; Gómez, P; Garrido, D and Jamilena, M (2011). The role of ethylene and brassinosteroids in the control of sex expression and flower development in *Cucurbita pepo*. *Plant Growth Regul.*, 65: 213-221.
- 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 and Bryant, S H (2013). CDD: Conserved domains and protein three-dimensional structure. *Nucleic Acids Res.*, 41(D1): 1-5.
- Martínez, C; Manzano, S; Megías, Z; Garrido, D; Picó, B and Jamilena, M (2013). Involvement of ethylene biosynthesis and signaling in fruit set and early fruit development in zucchini squash (*Cucurbita pepo* L.). *BMC Plant Biology*, 13: 139.
- Mohd Hudzairi, R; Somad, A S; Halim, M A and Roslan, S (2012). A review on crop plant production and ripeness forecasting. *Int. J. Agric. Crop Sci.*, 4(2): 54-63.
- Nakagawa, S; Niimura, Y; Gojobori, T; Tanaka, H and Miura, K I (2008). Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. *Nucleic Acids Res.*, 36(3): 861-871.
- Nurniwalis, A W; Zubaidah, R; Siti Nor Akmar, A; Suhaimi, N and Massawe, F (2018). Isolation and characterisation of an ethylene receptor (ERS-type) from oil palm (*Elaeis guineensis* Jacq.) mesocarp. *J. Oil Palm Res. Vol.* 30(2): 251-264.
- Nurniwalis, A W (2017). *Physiological, Biochemical and Molecular Analyses of Fruit Development in Oil Palm* (*Elaeis guineensis* Jacq.). Ph. D thesis. University of Nottingham Malaysia. 212 pp.
- Nurniwalis, A W; Suhaimi, N; Siti Nor Akmar, A; Aminah, S and Mohamad Arif, M A (2008). Gene discovery via expressed sequence tags from the oil palm (*Elaeis guineensis* Jacq.) mesocarp. *J. Oil Palm Res. Special Issue Vol.* 2: 87-96.
- Parkinson, J S and Kofoid, E C (1992). Communication modules in bacterial signaling proteins. *Annu. Rev. Genet.*, 26(1): 71-112.

- Pop, M (2009). Genome assembly reborn: Recent computational challenges. *Brief. Bioinform.*, 10(4): 354-366.
- Saitou, N and Neim, M (1987). The Neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, 4(4): 406-425.
- Sakai, H; Hua, J; Chen, Q G; Chang, C; Medrano, L J; Bleecker, A B and Meyerowitz, E M (1998). ETR2 is an ETR1-like gene involved in ethylene signaling in Arabidopsis. *PNAS.*, 95(10): 5812-5817.
- Sambanthamurthi, R; Sundram, K and Tan, Y (2000). Chemistry and biochemistry of palm oil. *Prog. Lipid Res.*, 39: 507-558.
- Schaller, G E and Bleecker, A B (1995). Ethylene-binding sites generated in yeast expressing the Arabidopsis *ETR1* gene. *Science*, 270: 1809-1811.
- Shakeel, S N; Wang, X; Binder, B M and Schaller, G E (2013). Mechanisms of signal transduction by ethylene: Overlapping and non-overlapping signaling roles in a receptor family. *AoB Plants*, 5: plt010.
- Singh, R; Ong-Abdullah, M; Low, L E T; Manaf, M A; Rosli, R; Nookiah, R; Ooi, L C; Ooi, S E; Chan, K L; Halim, M A; Azizi, N; Nagappan, J; Bacher, B; Lakey, N; Smith, S W; He, D; Hogan, M; Budiman, M A; Lee, E K; DeSalle, R; Kudrna, D; Goicoechea, J L; Wing, R A; Wilson, R K; Fulton, R S; Ordway, J M; Martienssen, R A and 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, L E T; Ooi, L C; Ong-Abdullah, M; Ting, N C; Nagappan, J; Nookiah, R; Amiruddin, M D; Rosli, R; Manaf, M A; Chan, K L; Halim, M A; Azizi, N; Lakey, N; Smith, S W; Budiman, M A; Hogan, M; Bacher, B; Van Brunt, A; Wang, C; Ordway, J M; Sambanthamurthi, R and Martienssen, R A (2013b). The oil palm SHELL gene controls oil yield and encodes a homologue of *SEEDSTICK*. *Nature*, 500(7462): 340-344.
- Slater, G S C and Birney, E (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, 6: 31.
- Tamura, K; Peterson, D; Peterson, N; Stecher, G; Nei, M and Kumar, S (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, maximum parsimony methods. *Mol. Biol. Evol.*, 28(10): 2731-2739.
- Thompson, J D; Higgins, D G and Gibson, T J (1994). Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 22(22): 4673-4680.
- Tieman, D M and Klee, H J (1999). Differential expression of two novel members of the tomato ethylene receptor family. *Plant Physiol.*, 120(95): 165-172.
- Tieman, D M; Taylor, M G; Ciardi, J A and Klee, H J (2000). The tomato ethylene receptors *NR* and *LeETR4* are negative regulators of ethylene response and exhibit functional compensation within a multigene family. *PNAS.*, 97(10): 5663-5668.
- Ting, N-C; Jansen, J; Mayes, S; Massawe, F; Sambanthamurthi, R; Ooi, C-L; Chin, C W; Arulandoo, X; Seng, T-Y; Sharifah, S R S A; Maizura, I and Singh, R (2014). High-density SNP and SSR-based genetic maps of two independent oil palm hybrids. *BMC Genomics*, 15: 309.
- Tranbarger, T J; Dussert, S; Joët, T; Argout, X; Summo, M; Champion, A; Cros, D; Omore, A; Nouy, B and Morcillo, F (2011). Regulatory mechanisms underlying oil palm fruit mesocarp maturation, ripening, functional specialisation in lipid and carotenoid metabolism. *Plant Physiol.*, 156(2): 564-584.
- Voorrips, R E (2002). Map chart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.*, 93(1): 77-78.
- Yau, C P; Wang, L; Yu, M; Sze, Y Z and Yip, W K (2004). Differential expression of three genes encoding an ethylene receptor in rice during development, in response to indole-3-acetic acid and silver ions. *J. Exp. Bot.*, 55(397): 547-556.