

Transcriptome analysis of oil palm (*Elaeis guineensis* Jacq.) roots under waterlogging stress

Sunya Nuanlaong¹, Suwit Wuthisuthimethavee¹, Maruay Mekanawakul² and Potjamarn Suraninpong^{1,*}

¹School of Agricultural technology, Walailak University, Nakhon Si Thammarat, Thailand 80160

²School of Science, Walailak University, Nakhon Si Thammarat, Thailand 80160

*Correspondence: potjamas@hotmail.com

Abstract

Waterlogging seriously constrains growth and yields in oil palm. To date, the responsive molecular changes caused by waterlogging in oil palm remain elusive. To elucidate the molecular genetic mechanisms of waterlogging stress, two varieties of oil palm Deli x Lamé and Deli x Ghana were used. The transcriptome profiles of the roots under waterlogging stress and normal conditions were compared via Ion Torrent Sequencing. Four libraries (GNR, GSR, SNR, and SSR) of oil palm roots after 45 days of normal watering and waterlogging stress were constructed. Approximately 6.2 million sequenced reads per library were obtained, with 5.5 million mapped reads (88.64%) similar to the oil palm genome in the GenBank database. A comparison of GNR/GSR showed a high of 3,289 DEGs with most genes up-regulated (1,863 DEGs). The GO analysis revealed the distribution of the DEGs among various pathways, suggesting a wide spectrum of physiological processes impacted by waterlogging stress. Moreover, qRT-PCR showed strong expression of all selected RNA-seq genes in waterlogged Deli x Ghana (GSR), especially *GST*, *SAPK10* and *NAC29* that are reported for the time to respond to waterlogging stress. Thus, this study not only reveals the comprehensive mechanisms of waterlogging responsive transcription in oil palm, but also establishes Deli x Ghana as a highly-adaptable variety to waterlogging conditions.

Keywords: *Elaeis guineensis* Jacq., waterlogging stress, abiotic stress, transcriptome analysis, Ion Torrent Sequencing

Abbreviations: GNR_Deli x Ghana-control; SNR_Deli x Lamé-control; GSR_Deli x Ghana-waterlogged; SSR_Deli x Lamé-waterlogged; DEGs_differentially expressed genes; KEGG_Kyoto Encyclopedia of Genes and Genomes; ANPs_Anaerobically induced polypeptide; ROS_Relative oxygen species; NGS_Next-generation sequencing; TMAP_Torrent mapping alignment program; FPKM_fragments per kilobase of transcript per million fragments sequenced; GO_Gene ontology; qRT-PCR_Quantitative reverse transcription polymerase chain reaction.

Introduction

Waterlogging is a soil condition in which excess water causes inadequate gas exchange between the soil and the atmosphere. Soil pores that are normally gas-filled become water-filled, and are a major feature of waterlogging. Because the diffusion of oxygen in water is around 320,000 times less (Watanabe et al., 2013), the amount of oxygen available to the roots decreases (Lee et al., 2011), and photosynthesis and respiration are limited. Waterlogging affects plant growth and development by reducing oxygen supply (Voeselek and Bailey-Serres, 2015). The low oxygen conditions lead to lower adenosine triphosphate (ATP) production and have a negative impact on cellular energy status. Consequently, the acclimation to low oxygen will activate gene encoding proteins and enzymes for anaerobic fermentation, glycolysis, transcription factors, and signaling pathways in order to allow biological and physiological adjustments to the low oxygen conditions (Bailey-Serres et al., 2012). Waterlogging-tolerant plants have many morphological/biochemical responses to these stresses. A key response of many plants to low oxygen is the formation of adventitious roots (Steffens and Rasmussen, 2016), aerenchyma cells (Yamauchi et al., 2013) and radial oxygen loss (ROL) barriers (Nishiuchi et al., 2012) to ensure the

movement of oxygen from the well-aerated shoots to the roots, production of relative oxygen species (ROS) (Steffens and Rasmussen, 2016), and anaerobically induced polypeptides (ANPs) (Christianson et al., 2010). However, the response of plants to waterlogging conditions is complex with responses to low oxygen involving significant changes in approximately 5%-10% of all genes assayed (Qi et al., 2012). An increase in mRNA associated with several pathways such as glycolysis and fermentation, signal transduction, plant hormone, transcription factors, cell wall modification, nitrogen and sulfur metabolism in *Zea mays* (Rajhi et al., 2011), *Arabidopsis thaliana* (Narsai and Whelan, 2013), *Rumex palustris* and *Rumex acetosa* (van Veen et al., 2013), *Oryza sativa* (Narsai and Whelan, 2013), *Glycine max* (Tamang et al., 2014) and kiwifruit (Zhang et al., 2015) under waterlogging conditions have been reported.

Recently, the development of next-generation sequencing (NGS) technologies and associated bioinformatics tools has provided a new method for transcriptomic research of RNA-Sequences (Wang et al., 2009). Many transcriptomic reports on waterlogging currently address similar topics with regard to gene expression, but the response results have not been well studied in plant species. In fact, understanding of the

mechanism associated with the regulation of waterlogging tolerance remains a fundamental challenge. In oil palm (*Elaeis guineensis* Jacq.), there has been no large-scale report of gene expression analysis responsiveness to waterlogging. Oil palm, a plantation crop of major economic importance in Southeast Asia, is the predominant source of edible oil worldwide and its use in biodiesel is increasing. The primary yield-limiting factor of oil palm cultivation is water supply. However, excess water or floods can negatively impact oil palm yield by delaying planting, reducing vigor, altering development, and increasing susceptibility to disease in the oil palm roots (Rivera-Mendes et al., 2016). Although, oil palm is well adapted to a high-water table, it does not tolerate continuous waterlogging. Since, its roots are unable to carry out respiration when submerged under water, and if waterlogging is prolonged, the roots may die, water uptake may also be limited, and waterlogged oil palm often appear nitrogen deficient (Corley and Tinker, 2015). Thus, to gain a comprehensive insight into how oil palm responds to waterlogging at a molecular level, herein we report a detailed analysis of gene expression profiling between tolerance and susceptible variety of oil palm via Ion Torrent Sequencing. Our results reveal an understanding of the response of waterlogging-tolerance to soil waterlogging stress.

Results

Properties of the cDNA libraries sequenced

In four libraries, approximately 6.2 million total sequence reads per library with 574.2 million total base pairs were obtained. The reads had an average length of 142 bp. The SNR library showed the highest number of both total base pairs and reads number. The obtained sequencing reads when aligned against the combined GenBank database (accession No. PRJNA 192219) were mapped to approximately 88.64% of the oil palm genome (Table 1).

Genes detected per subset

Gene expression profiles during waterlogging stress in four libraries are drawn as Venn diagrams, with the intersections between expressed genes detected at each library shown in Fig. 1. A total of 18,518, 15,740, 15,941 and 15,121 transcripts were identified from GNR, GSR, SNR and SSR libraries, respectively. 13,284 out of 18,809 detected genes were expressed in all four libraries and 2,127 genes were expressed exclusively in the different varieties and conditions of oil palm roots. Of these, uniquely expressed genes, 1,540, 281, 62 and 244 genes were expressed exclusively in GNR, GSR, SNR and SSR libraries, respectively. The ratio of genes expressed exclusively in GSR and SSR libraries to that of the genes expressed in all libraries was 343/18,809, indicating significant expression changes during the development of oil palm roots under waterlogging conditions.

Identification of differentially expressed genes (DEGs) responding to waterlogging stress

From Venn diagram analysis, 27,122 genes in 16 chromosomes and scaffolds (Singh et al., 2013) were

obtained and continuously tested with the Cufflinks program for differentially expressed genes (DEGs) in 6 combination libraries. The results showed that 3,675 genes had p -values ≤ 0.05 , q -values ≤ 0.05 , and an estimated absolute $|\log_2 \text{Ratio}| \geq 1$ or ≤ -1 in at least one of the pairwise comparisons, which were used as a threshold to judge whether gene expression was significant (Table 2). GNR/GSR revealed that there were 3,287 DEGs, most of them up-regulated in the waterlogged system. Whereas SSR/GSR showed only 87 DEGs, with most of the genes down-regulated (78 genes). In contrast, SNR/SSR showed no ratio of DEGs. Among the DEGs, ethylene biosynthesis, starch metabolism, glycolysis and fermentation, signal transduction, protein kinases, cell wall modification/degradation, and transcription factors were the most common genes.

Functional annotation of DEGs

To further extend the molecular characterization of DEGs the 3,675 sequences of the oil palm roots were classified into several categories based on their allocated Gene Ontology (GO). The results revealed hits for 3,667 sequences in Blast, 3,229 sequences in the InterPro Scan database, with 2,377 sequences mapped and 2,659 sequences annotated (Fig. 2A). Statistical analysis of the top hits showed 40% of the homologous sequences were between 1.0E-150 to 1.0E-180 in the nr database, whereas 60% of the sequences had a threshold E -value with a homology of $>1.0E-150$ (Fig. 2B). The majority of the annotated sequences correspond to the recognized nucleotide sequences of plant species such as *Vitis vinifera* (22%), *Theobroma cacao* (7%), *Oryza sativa* Japonica group (7%), *Setaria italic* (6%), *Amborella trichopoda* (5%), and *Zea mays* (4%), respectively (Fig. 2C).

To gain a better understanding of the shift mechanism in oil palm roots in response to waterlogging, biological function analysis of DEGs was performed by GO enrichment using the whole transcriptome as the background. In total, 3,675 genes were categorized into three different GO trees of biological processes, cellular components, and molecular functions at level 3 (Fig. 3).

To identify the conserved domains or functional units within the protein query sequences InterPro Scan was used. A total of 3,229 sequences after annotation in InterPro Scan, generated 2,853 sequences of functional units. The 20 top-hit InterPro Scan IDs in the DEGs are summarized in Table. 3. The statistics showed that "P-loop containing nucleoside triphosphate hydrolase" was the most predominant conserved domain followed by "Zinc finger, RING/FYVE/PHD-type" and "Armadillo-type fold". These annotations provide a precious resource for exploring specific processes, functions and pathways during searching of the DEGs of oil palm roots under waterlogging conditions.

Functional classification using the KEGG database

A Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway-based analysis was performed to achieve a better understanding of the biological functions of the sequences. 673 out of the 3,675 sequences were significantly matched with the NCBI database and were assigned to 108 KEGG pathways. The most significant KEGG pathways were associated with metabolism, followed by genetic information processing, environmental information

Table 1. Summary of the Ion Torrent Sequencing data of oil palm root transcriptomes under waterlogging stress for 45 days.

	GNR	GSR	SNR	SSR	Mean
Number of bases	396,722,464	578,939,801	774,666,540	546,626,417	574,238,805
Total no. of reads with Q _≥ 20	6,030,662	5,961,601	6,506,821	6,323,331	6,205,604
Mean read length (bp)	142	146	139	141	142
Mapped reads (%)	5,289,612 (87.71)	5,250,198 (90.31)	5,872,346 (88.05)	5,595,390 (88.48)	5,501,886 (88.64)
Unmapped reads (%)	741,051 (12.29)	711,403 (11.69)	634,476 (9.69)	727,942 (11.52)	703,718 (11.37)

Note: GNR, Deli x Ghana-control; SNR, Deli x Lamé-control; GSR, Deli x Ghana-waterlogging treated; SSR, Deli x Lamé-waterlogging treated.

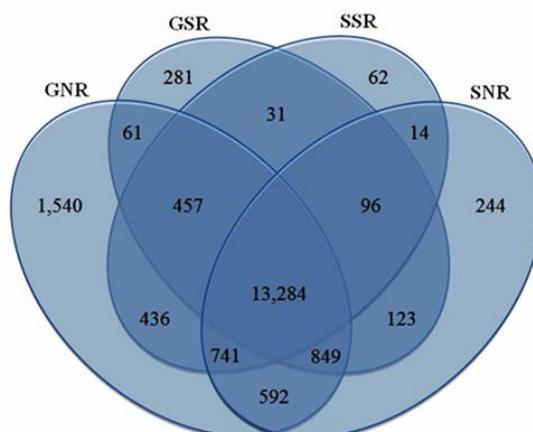


Fig 1. Venn diagram for the number of genes expressed in 4 libraries of oil palm roots under normal and waterlogging stress for 45 days. (GNR, Deli x Ghana-control; GSR, Deli x Ghana-waterlogged; SNR, Deli x Lamé-control; SSR, Deli x Lamé-waterlogged). Numbers in each intersection represent the number of genes detected with a least one read in these disjoint sets (intersections).

Table 2. Number of DEGs in different comparisons.

Pairwise Comparison of DEGs between 2 libraries	GNR/GSR ^a	SNR/SSR	SNR/GNR	SNR/GSR	SSR/GNR	SSR/GSR
Number of DEGs	3,289 ^b	<i>ns</i> ^c	230	561	88	87
Up-regulated	1,863	-	181	398	53	9
Down-regulated	1,426	-	49	163	35	78

Note: ^aGNR, Deli x Ghana-control; SNR, Deli x Lamé-control; GSR, Deli x Ghana-waterlogged; SSR, Deli x Lamé-waterlogged

^bAll the genes mapped to the oil palm genome sequences were examined for their expression differences across different libraries. Numbers of differentially expressed genes represent transcripts using threshold values q -values ≤ 0.05 and $|\log_2 \text{Ratio}| \geq 1$ for controlling false discovery rates.

^c*ns*, non-significant

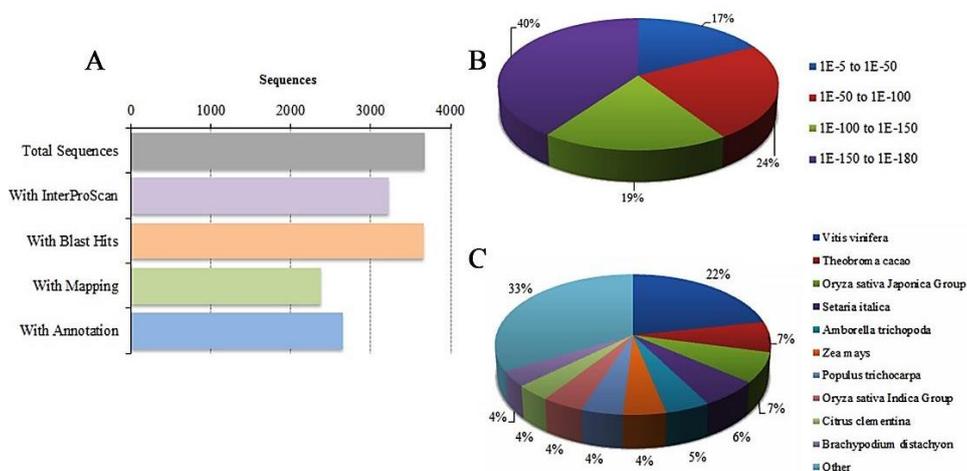


Fig 2. Characteristics of homology search of DEGs by Blast2GO program. A) Bar Graph illustrating the properties of the DEGs, B) E-value distribution of gene BLASTx hits in NCBI database, and C) species distribution.

Table 3. InterPro Scan distributions of DEGs in oil palm root under waterlogging stress.

InterPro Scan IDs	Type	Description	Sequences
IPR027417	Domain	P-loop containing nucleoside triphosphate hydrolase	177
IPR013083	Domain	Zinc finger, RING/FYVE/PHD-type	108
IPR016024	Domain	Armadillo-type fold	102
IPR011989	Domain	Armadillo-like helical	98
IPR011009	Domain	Protein kinase-like domain	78
IPR011990	Domain	Tetratricopeptide-like helical domain	72
IPR016040	Domain	NAD(P)-binding domain	69
IPR000504	Domain	RNA recognition motif domain	68
IPR000719	Domain	Protein kinase domain	66
IPR009057	Domain	Homeobox domain-like	62
IPR011991	Domain	Winged helix-turn-helix DNA-binding domain	61
IPR015943	Domain	WD40/YVTN repeat-like-containing domain	60
IPR003593	Domain	AAA ⁺ ATPase domain	57
IPR001841	Domain	Zinc finger, RING-type	55
IPR032675	Domain	Leucine-rich repeat domain, L domain-like	53
IPR017986	Domain	WD40-repeat-containing domain	50
IPR001680	Repeat	WD40 repeat	45
IPR029058	Domain	Alpha/Beta hydrolase fold	44
IPR008271	Site	Serine/threonine-protein kinase, active site	44
IPR017853	Domain	Glycoside hydrolase superfamily	44

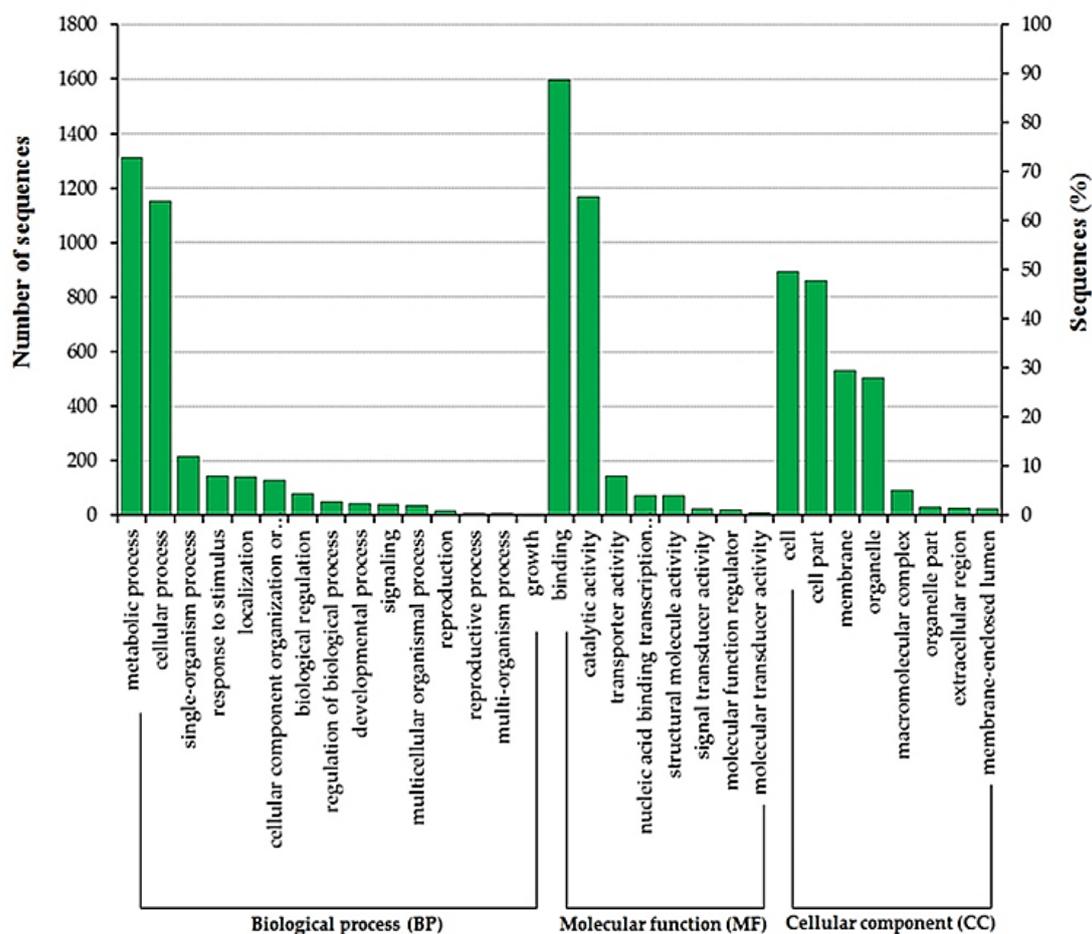


Fig 3. Histogram of GO functional classification of DEGs. Results were grouped into three main categories: BP, CC, and MF. The right y-axis indicated the number of sequences corresponding to each subcategory, and the left y-axis indicated the percentage of sequences involved in each specific subcategory.

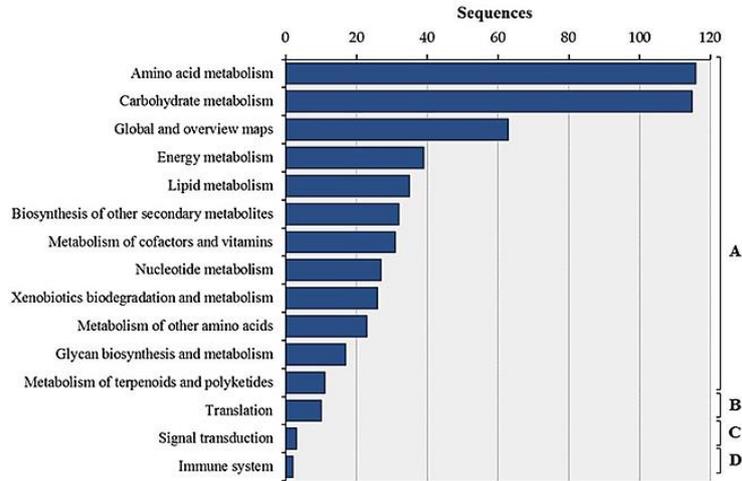


Fig 4. Bar Graph illustrating the pathway assignment based on KEGG distribution of DEGs. Pathways were assigned into four categories including A) Metabolism, B) Genetic information processing, C) Environmental information processing, and D) Organismal systems.

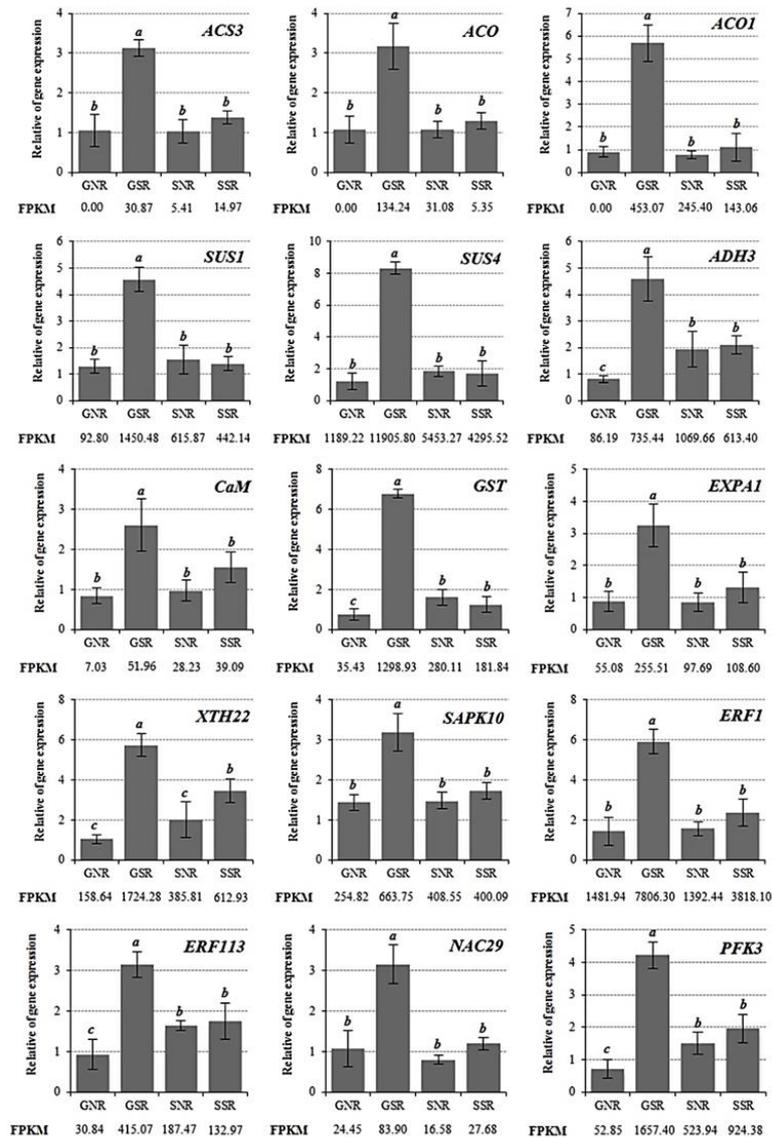


Fig 5. The relative expression levels of representative DEGs from GNR, GSR, SNR, and SSR samples. The *18S rRNA* gene was used as an internal control and values represent means \pm SE (n = 3) and the significance level of 0.05 was used for different letters above bars.

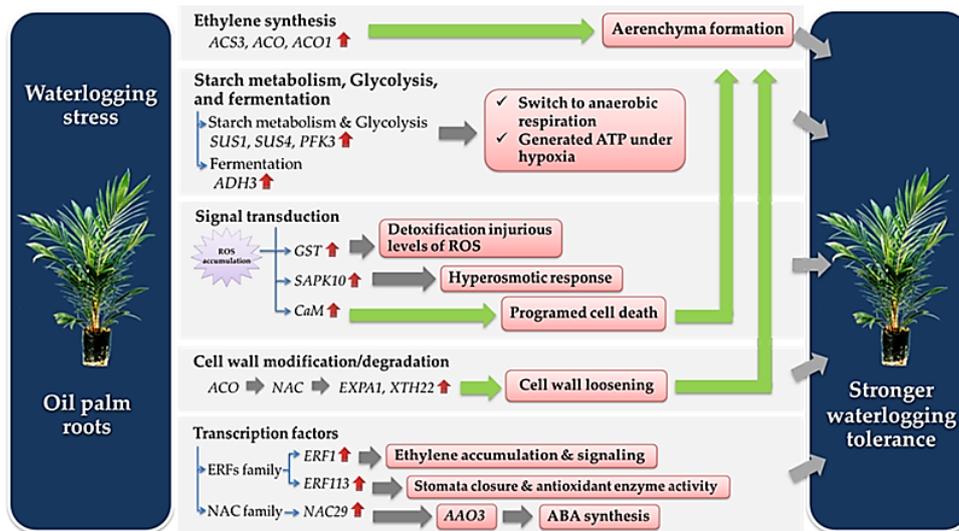


Fig 6. A hypothetical model of selected gene expression to waterlogging stress in oil palm roots.

and organismal systems. The five largest pathway groups in metabolism were amino acid metabolism (116 seqs in pathway), carbohydrate metabolism (115 seqs in pathway), Global and overview maps (63 seqs in pathway), Energy metabolism (39 seqs in pathway) and Lipid metabolism (35 seqs in pathway) (Fig. 4).

Verification of RNA-Seq data by real-time quantitative RT-PCR

Results from qRT-PCR revealed close similarity with the expression of DEGs obtained from the Ion Torrent Sequencing analysis, indicating the method used for determination of the DEGs is valid. The 15 selected genes showed differential expression between varieties and stress conditions as shown in Fig. 5.

Discussion

Transcriptome analysis of waterlogging stress in oil palm root

Since, waterlogging stress affects the growth and yield of oil palm, therefore identification of the waterlogging tolerance mechanism is an important for the improvement of oil palm. Comparison of a transcriptional response to waterlogging stress in the two oil palm varieties in this study exposed the highest number of DEGs in GNR/GSR together with a high number of up-regulated genes. Similar up-regulation of genes has been reported under the same conditions in various plants such as *Cucumis sativus* (Qi et al., 2012), *Brassica napus* (Zou et al., 2014), *Taxodium distichum* (Qi et al., 2014), *Glycine max* (Chen et al., 2016), and *Chrysanthemum morifolium* (Zhao et al., 2018). While waterlogging gives rise to a relatively large number of DEGs, the pathways associated with "metabolic process", "cellular process", and "single-organism process" were enhanced. This confirms current knowledge of waterlogging responsiveness in several plant species (Wei et al., 2016; Ren et al., 2017; Zhang et al., 2017). In addition, the InterPro

analysis found that the most predominantly conserved domains were involved in program cell death, disease, stress response, signal transduction, ubiquitination, and plant development (Pariyar et al., 2016; Sharma and Pendey, 2016), suggesting a critical role of oil palm tolerance in responding to waterlogging stress.

Ethylene biosynthesis as an affected by waterlogging

Ethylene is a critical hormone that has been reported for waterlogging stress induction (Sasidharan and Voesenek, 2015) in several plant species (Lee et al., 2011; van Veen et al., 2013; Sasidharan and Voesenek, 2015). In tomato roots, rapidly induced 1-aminocyclopropane-1-carboxylate (ACC) synthase (*LE-ACS3*) was detected (Olson et al., 1995). Similarly, Vriezen et al. (1999) found a more prominent role of ACC oxidase (*RP-ACO1*) in *Rumex palustris* during flooding. Our results complement these studies showing that *ACS3*, *ACO*, and *ACO1* were up-regulated in GSR (Deli x Ghana-waterlogged) and significantly different ($P \leq 0.05$) from each other. Moreover, ethylene is implicated in the induction of lysigenous aerenchyma formation (Evans, 2003; Joshi and Kumar, 2012) under waterlogging conditions. Lysigenous aerenchyma contributes oxygen from shoots to roots and ventilates gases (e.g. carbon dioxide and methane) from roots to shoots (Evans, 2003). Thus, the formation of lysigenous aerenchyma is essential to the survival and functional of plants subjected to waterlogging. A high relative gene expression of *ACS3*, *ACO*, and *ACO1* under waterlogging stress in Deli x Ghana root is associated with high lysigenous aerenchyma formation as reported in our previous study (Nuanlaong, 2018). Thus, the striking response of ethylene biosynthesis in oil palm might be a vital adaptation to cope with waterlogging stress.

Starch metabolism, glycolysis, and fermentation as an affected by waterlogging

Under waterlogging conditions, the change from aerobic to fermentative metabolism requires constant

supplementation of carbohydrates (soluble sugar). Anaerobic respiration includes glycolysis and fermentation, and plays an important role in responding to waterlogging stress (Hossain and Uddin, 2011; Juntawong et al., 2014), thus several genes that function in the metabolism of sugar and its derivatives were up-regulated (Cannarozzi et al., 2018). Sucrose synthase (SUS), a glycosyl transferase enzyme that plays a key role in sugar metabolism (Stein and Granot, 2019) was highly expressed in the present study, especially, *SUS1* and *SUS4*. Similar results are found for *Arabidopsis* where the *SUS1* and *SUS4* transcripts, as well as their relative proteins, increased significantly upon low oxygen both in roots and shoots (Baud et al., 2004). Moreover, ATP-dependent 6-phospho fructo kinase 3 isoform X1 (*PFK3*), which catalyzes the phosphorylation of D-fructose 6-phosphate to fructose 1,6-bisphosphate by ATP in the first step of glycolysis, was up-regulated in GSR. The appearance of *PFK3* for acceleration of the glycolysis pathway (Lasanthi-Kudahettige et al., 2007) was reported as a tolerance mechanism of plant to waterlogging stress (Qi et al., 2014). Moreover, this present study found the up-regulation of alcohol dehydrogenase 3 (*ADH3*) in GSR was significantly different ($p \leq 0.05$). ADH is a second enzyme that converts acetaldehyde to ethanol in the fermentation process. Under waterlogging conditions, oxygen deprivation acts as a primary signal in the response (Jackson and Colmer, 2005) and plant alcohol fermentation is activated under low-oxygen stress conditions. Thus, a higher rate of ethanol synthesis under waterlogging conditions may help the plant to generate ATP under low oxygen conditions (Ismail et al., 2009), which also acts as a tolerance mechanism.

Signal transduction as an affected by waterlogging

Activated production of reactive oxygen species (ROS) is a ubiquitous phenomenon under stress conditions and is a key factor in signal transduction stimulated by abiotic stress at the molecular level (Suzuki et al., 2011). Low-oxygen conditions that induced an increase in the redox potential of both plant roots and surrounding soil, is the ideal for ROS production (Yamauchi et al., 2017). Also, the high capacity of ROS generations requires strongly antioxidant (ROS-scavenging) defenses (Dalton et al., 2009). In oil palm, there are no previous reports of ROS-scavenging under waterlogging stress. This study was the first finding of *GST* that was encoded to glutathione s-transferase, strongly up-regulated in GSR. Normally, glutathione is an enzymatic antioxidant for detoxification of injurious levels of ROS, but it has been found highly expressed in flooding stress (Wu et al., 2017). So, the increased activity of ROS-scavenging enzymes, and subsequent lower accumulation of ROS, likely contribute to waterlogging tolerance in oil palm. Furthermore, this study found highly up-regulated *CaM* that was encoded to calmodulin in GSR more than in SSR. Calmodulin is a Ca^{2+} -sensing protein involved in the transduction of Ca^{2+} signals (Viridi et al., 2015). It promotes programmed cell death (PCD) in the formation of lysigenous aerenchyma under waterlogging stress (Rajhi et al., 2011; Yamauchi et al., 2017). The strong appearance of *CaM* in GSR is associated with lysigenous aerenchyma formation under waterlogging stress in Deli x Ghana variety as shown previously (Nuanlaong, 2018). Moreover, new expressed genes, Serine/threonine-protein kinase or osmotic

stress/ABA-activated protein kinase (*SAPK10*) were discovered in this study. *SAPK10* is activated by abscisic acid (ABA) and plays a role in signal transduction of the hyperosmotic response (Li et al., 2015), especially in stress responses (Kobayashi et al., 2005; Han et al., 2017). This study found differentially expressed *SAPK10* under waterlogging stress conditions, notably increased expression in GSR. The finding of *SAPK10* in oil palm has not been reported in plant responsiveness to waterlogging stress before.

Cell wall modification/degradation as an affected by waterlogging

Waterlogging stress induces the activity of cell-wall modification/degradation-related enzymes during lysigenous aerenchyma formation (Rajhi et al., 2011; Leite et al., 2017). In this study, the relative mRNA expressed xyloglucan endotransglucosylase/hydrolase 22 (*XTH22*) was up-regulated in GSR. Similarly, in maize and *Eragrostis tef*, high expression of *XTHs* gene under waterlogging/anaerobic conditions has been found (Takahashi et al., 2015; Cannarozzi et al., 2018). The activity of XTH has been found to be crucial during cell growth and differentiation in the construction, remodeling and disassembly of the xyloglucan/cellulose framework in primary cell walls (Hara et al., 2014; Tsuchiya et al., 2015). During lysigenous aerenchyma formation not only XTH, but also expansin (EXP) was found (Rajhi et al., 2011). Both of these genes interact to control cell wall-loosening processes during cell growth (Tenhaken, 2015). Thus, highly up-regulated *EXPA1* (expansin-A1) in GSR, is implicated in cell wall loosening in this study.

Transcription factors as an affected by waterlogging

Many studies have reported that transcription factors (TFs) families are involved in abiotic stress, and play an especially important role in waterlogging response, positively improved plant tolerance and generally act as key regulators of gene expression (Rauf et al., 2013). The ethylene response factors (ERFs) family is a large gene family of transcription factors and is part of the AP2/ERF superfamily, which also contains the AP2 and RAV families (Riechmann et al., 2000). ERFs were identified and implicated in many diverse functions in the development processes and responsiveness to both biotic and abiotic stresses (Zhang et al., 2012). In this study, two-ERFs, *ERF1* and *ERF113*, were noticeably induced in response to waterlogging, with especially high expression in GSR. *ERF1* is marked in the ethylene transduction pathway, together with *ERF2* that was induced during reoxygenation (Sasidharan et al., 2018). In addition, *ERF113* or *RAP2.6L* act as positive regulators of tolerance to waterlogging stress. These delay waterlogging-induced premature senescence by regulating stomatal closure and antioxidant enzyme activity (Liu et al., 2012). Thus, under waterlogging stress, *ERF113* was detected in the GSR root. Besides this, *NAC29* that was encoded to NAC transcription factor 29, acts as a transactivate promoter of abscisic aldehyde oxidase (Yang et al., 2014), was found to be highly mRNA expressed in GSR. NAC comprises a large family of transcription factors that play important roles in diverse physiological processes during development and response to

abiotic stress (Kou et al., 2016). Thus, *NAC29* has been reported for their up-regulation in polyethylene glycol, sodium chloride and cold treatments (Zhou et al., 2016). This study is the first time *NAC29* expression has been found in response to waterlogging stress. Therefore, further studies are necessary to clarify the regulatory mechanisms involving these genes in oil palm. From our study, we provided the first report of the transcriptome analysis of the oil palm root under waterlogging stress and identified genes-related to waterlogging stress using Ion Torrent Sequencing. Based on our results, waterlogging tolerance in oil palm plants at 45 d was observed in Deli x Ghana. A hypothetical model of the selected gene expression profile caused by waterlogging stress in oil palm roots is presented in Fig. 6. Our present research provides a better understanding of the response mechanism against waterlogging in oil palm roots at the molecular level.

Materials and methods

Plant materials and waterlogging treatments

E. guineensis Jacq. var. Deli x Ghana and Deli x Lamé at 12 months-old were grown in 200 L fiberglass tanks containing topsoil and sandy soil (4:1, v/v) in a greenhouse with 12/12 h (34/24 °C) day/night temperature and 81.8-85.4%RH at Nakhon Si Thammarat, Thailand. The two varieties of oil palm plants were divided into two groups, thus four treatments were set up. Control was a normal water supply and waterlogging stress was imposed by ensuring the water level was at 10 cm above the soil surface. Each treatment included five replications. All of the treated roots both control and waterlogged at 45 days were collected and frozen in liquid nitrogen until used for RNA extraction.

RNA isolation

The total RNA isolated from the oil palm roots of Deli x Ghana-control, Deli x Lamé-control, Deli x Ghana-waterlogged and Deli x Lamé-waterlogged were termed GNR, SNR, GSR and SSR libraries, respectively. The total RNA was extracted from all treated frozen roots following a modified method of Corre et al. (1996). The RNA quality and quantity were assessed by a NanoDrop ND2000 Spectrophotometer (Thermo Fisher Scientific, USA) and by agarose gel electrophoresis. An RNA integrity number (RIN) of greater than 8.0 was used.

cDNA library preparation, transcriptome sequencing, and data processing

Library construction was carried out at the National Center for Genetic Engineering and Biotechnology (BIOTEC, Thailand). The messenger RNA was separated from total RNA using Absolutely mRNA Purification kit (Agilent Technologies, USA). The whole-transcriptome cDNA library was prepared using an Ion Total RNA-Seq kit v2 (Life Technologies, USA). Double-stranded cDNA was ligated to barcode-adapters and sequenced using an Ion PI Chip (Ion torrent, Life Technologies, USA). Torrent Suite Software 4.0 (Ion torrent, Life Technologies, USA) was used to process raw information, remove adapter sequences, base calling, and quality value calculations. Quality reads were obtained

by trimming the raw reads at a minimum PHRED score of $Q = 20$ (Shu et al., 2015). Then, the cleaned reads were mapped to the oil palm reference genome accession No. PRJNA 192219, GenBank (Singh et al., 2013) by Torrent Mapping Alignment Program (TMAP).

Differentially expressed gene analysis

Data was normalized by calculating the fragments per kilobase of transcript per million fragments sequenced (FPKM). Differentially expressed genes (DEGs) of GNR, SNR, GSR and SSR were identified using the Cufflinks program (Trapnell et al., 2012) and DEGs were compared within 6 combinations: GNR/GSR, SNR/SSR, SNR/GNR, SNR/GSR, SSR/GNR and SSR/GSR. The Pearson's Chi squared test to assess the lane effect was applied and the p -value was computed for each gene. The Benjamini-Hochberg false discovery rate (FDR) was applied to correct the results for the p - and q -value ≤ 0.05 , and further used as the threshold for judging significant difference in gene expression.

Functional annotation of DEGs

The Blast2GO software suite (www.blast2go.org/) (Conesa and Götzt, 2008) was used to predict Gene Ontology (GO) terms, assign the assembled sequences to the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Kanehisa et al., 2017), and identified protein domains, families, repeats and sites against the InterPro protein signature databases (Jones et al., 2014). Annotations using Blast2GO were conducted with $1.0E-5$ as the E -value hit filter. While, GO terms were classified into three categories at level 3 for biological process (BP), cellular component (CC), and molecular functions (MF).

Quantitative RT-PCR (qRT-PCR) analysis of candidate genes

To confirm RNA-Seq results, 15 randomly selected up-regulated genes were chosen for expression validation using qRT-PCR with gene specific primers (Supplementary Table. 1). The primers were designed with primer 3 plus software (Untergasser et al., 2007). Three biological replications with two samples of total RNA were used. Total RNA was treated with RNase free DNase. Reverse transcription of total RNA (1 μ g) was performed with iScriptTM Select cDNA synthesis Kit (Bio-Rad, USA). The determination of the expression level of fifteen genes by qRT-PCR was carried out using the Applied Biosystems 7300 Real-Time PCR System (Applied Biosystems, USA) with a 10- μ L reaction volume, containing 1 μ L tenfold diluted cDNA, 0.25 μ L (10 μ M) of each primer, 2 μ L HOT FIREPoI[®] EvaGreen[®] qPCR Mix Plus (ROX) (Solis BioDyne, Estonia), and 6.5 μ L Nuclear Free Water. The PCR conditions consisted of denaturation at 95 °C for 15 min, followed by 40 cycles of 95 °C for 30 s, 55 to 65 °C for 30 s, and 72 °C for 30 s. The specificity of the individual PCR amplification was monitored using a heat dissociation curve from 60 to 95 °C following the final PCR cycle. The expression level of genes relative to that of the *18S rRNA* control mRNA was analyzed using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). Data analyses were conducted using SPSS version 16.0 statistical software. For all analyses, the levels of significance between different samples were set at $p \leq 0.05$.

Conclusion

The transcript comparison of two varieties of oil palm under waterlogging stress using RNA-sequencing helped to explain the molecular basis of responsiveness of the remarkably waterlogging-tolerant oil palm. The DEGs between oil palm treated roots greatly varied at the transcription level, with highly differentially expressed genes of GNR/GSR comparison. The DEGs data demonstrated the most up-regulated genes including ethylene biosynthesis (*ASC3*, *ACO*, *ACO1*), starch metabolism, glycolysis, and fermentation (*SUS1*, *SUS4*, *PFK3*, *ADH3*), signal transduction (*CaM*, *GST*, *SAPK10*), cell wall modification/degradation (*EXPA1*, *XTH22*), and transcription factors (*ERF1*, *ERF113*, *NAC29*). All genes were highly expressed in GSR, especially, *GST*, *SAPK10* and *NAC29* which were found for the first time to be genes responding to waterlogging stress. The potential of waterlogging stress-related transcripts defined in the experiment provides important data for further knowledge of the molecular mechanisms of the waterlogging response in oil palm, as well as more efficient guidelines for the oil palm breeding program.

Acknowledgments

We are grateful for the support of Agricultural Research Development Agency, ARDA (Public Organization), Grant No. PRP5605021280, PRP5705021800 and PRP5905020320 and Walailak University Fund (Grant No.14/2557, 14/2560). Special thanks also to Assoc. Prof. Dr. David J. Harding, School of Science, Walailak University, for his kindness in the proofreading of this manuscript.

References

- Bailey-Serres J, Fukao T, Gibbs DJ, Holdworth MJ, Lee SC, Licausi F, Perata P, Voeselek LACJ, Dongen, JT (2012) Making sense of low oxygen sensing. *Trends Plant Sci.* 17: 129-138.
- Baud S, Vaultier MN, Rochat C (2004) Structure and expression profile of the sucrose synthase multigene family in Arabidopsis. *J Exp Bot.* 55: 397-409.
- Cannarozzi G, Weichert A, Schnell M, Ruiz C, Bossard S, Blösch R, Wüthrich SP, Chanyalew S, Assefa K, Tadele, Z (2018) Waterlogging affects plant morphology and the expression of key genes in *tef* (*Eragrostis tef*). *Plant Direct.* 1-22.
- Chen W, Yao Q, Patil GB, Agarwal G, Deshmukh RK, LIN L, Wang B, Wang Y, Prince SJ, Song L, Xu D, An YC, Valliyodan B, Varshney RK, Nguyen HT (2016) Identification and comparative analysis of differential gene expression in soybean leaf tissue under drought and flooding stress revealed by RNA-Seq. *Front Plant Sci.* 7: 1044.
- Christianson JA, Llewellyn DJ, Dennis ES, Wilson IW (2010) Global gene expression responses to waterlogging in roots and leaves of cotton (*Gossypium hirsutum* L.). *Plant Cell.* 51: 21-37.
- Conesa A, Götz S (2008) Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics.* 2008: 619832.
- Corley RHV, Tinker PB (2015) The oil palm, 5th ed. John Wiley & Sons, Chichester, UK, p 111.
- Corre F, Henry Y, Rode A, Hartmann C (1996) *Em* gene expression during somatic embryogenesis in the monocot *Triticum aestivum* L. *Plant Sci.* 117: 139-149.
- Dalton DA, Boniface C, Turner Z, Lindahl A, Kim HJ, Jelinek L, Govindarajulu M, Finger RE, Taylor CG (2009) Physiological roles of glutathione S-transferases in soybean root nodules. *Plant Physiol.* 150: 521-530.
- Evans DE (2003). Aerenchyma formation: Tansley review. *New Phytol.* 161: 35-49.
- Han S, Min MK, Lee SY, Lim CW, Bhatnagar N, Lee Y, Shin D, Chung KY, Lee SC, Kim BG, Lee S (2017) Modulation of ABA signaling by altering VxGL motif of PP2Cs in *Oryza sativa*. *Mol Plant.* 10: 1190-1205.
- Hara Y, Yokoyama R, Osakabe K, Toki S, Nishitani K (2014) Function of xyloglucan endotransglucosylase/hydrolases in rice. *Ann Bot.* 114(6):1309-1318.
- Hossain MA, Uddin SN (2011) Mechanisms of waterlogging tolerance in wheat: Morphological and metabolic adaptations under hypoxia or anoxia. *Aust J Crop Sci.* 5: 1094-1101.
- Ismail AM, Ella ES, Vergara GV, Mackill DJ (2009) Mechanisms associated with tolerance to flooding during germination and early seedling growth in rice (*Oryza sativa*). *Ann Bot.* 103(2): 197-209.
- Jackson MB, Colmer TD (2005) Response and adaptation by plant to flooding stress. *Ann Bot.* 96: 501-505.
- Jones P, Binns D, Chang HY, Fraser M, Li W, Mcanulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S (2014) InterProScan 5: Genome-scale protein function classification. *Bioinformatics.* 30: 1236-1240.
- Joshi R, Kumar P (2012) Lysigenous aerenchyma formation involves non-apoptotic programmed cell death in rice (*Oryza sativa* L.) roots. *Physiol Mol Biol Plants.* 18(1): 1-9.
- Juntawong P, Sirikhachornkit A, Pimjan R, Sonthirod C, Sangsrakru D, Yoocha T, Tangphatsornruang S, Srinives P (2014). Elucidation of the molecular responses to waterlogging in *Jatropha* roots by transcriptome profiling. *Front Plant Sci.* 5: 1-13.
- Kanehisa M, Furumichi M, Tanabe M, SATO Y, Morishima K (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45: D353-D361.
- Kobayashi Y, Murata M, Minami H, Yamamoto S, Kagaya Y, Hobo T, Yamamoto A, Hattori T (2005) Abscisic acid-activated SNRK2 protein kinase function in the gene-regulation pathway of ABA signal transduction by phosphorylating ABA response element-binding factor. *Plant J.* 44(6): 939-949.
- Kou X, Liu C, Han L, Wang S, Xue Z (2016) NAC transcription factors play an important role in ethylene biosynthesis, reception and signaling of tomato fruit ripening. *Mol Genet Genomics.* 291: 1205-1217.
- Lasanthi-kudahettige R, Magneschi L, Loreti E, Gonzali S, Licausi F, Novi G, Beretta O, Vitulli F, Alpi A, Perata P (2007) Transcript profiling of the anoxic rice coleoptile. *Plant Physiol.* 144: 218-231.
- Lee SC, Mustruph A, Sasidharan R, Vashisht D, Pederson O, Oosumi T, Voeselek LA, Bailey-Serres J (2011) Molecular characterization of the submergence response of the *Arabidopsis thaliana* ecotype Columbia. *New Phytol.* 190(2): 457-471.

- Leite DCC, Grandis A, Tavares EQP, Piovezani AR, Pattathil S, Avci U, Rossini A, Cambler A, De Souza AP, Hahn MG, Buckeridge MS (2017) Cell wall changes during the formation of aerenchyma in sugarcane roots. *Ann Bot.* 120(5):693-708.
- Li C, Shen H, Wang T, Wang X (2015) ABA regulates subcellular redistribution of OsABI-LIKE2, a negative regulator in ABA signaling, to control root architecture and drought resistance in *Oryza sativa*. *Plant Cell Physiol.* 56: 2396-2408.
- Liu P, Sun F, Gao R, Dong H (2012) *RAP2.6L* overexpression delays waterlogging induced premature senescence by increasing stomatal closure more than antioxidant enzyme activity. *Plant Mol Biol.* 79: 609-622.
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real time quantitative PCR and the $2^{-\Delta\Delta C_q}$ method. *Methods.* 25: 402-408.
- Narsai R, Whelan J (2013) How unique is the low oxygen response? An analysis of the anaerobic response during germination and comparison with abiotic stress in rice and Arabidopsis. *Front Plant Sci.* 4: 349.
- Nishiuchi S, Yamauchi T, Takahashi H, Kotula L, Nakazono M (2012) Mechanisms for coping with submergence and waterlogging in rice. *Rice.* 5: 2.
- Nuanlaong S (2018) Gene identification and gene expression of lysigenous aerenchyma formation in oil palm (*Elaeis guineensis* Jacq.) roots under waterlogging condition. Ph.D thesis, Walailak University.
- Olson DC, Oetiker JH, Yang SF (1995) Analysis of *LE-ACS3*, a 1-aminocyclopropane-carboxylic acid synthase gene expressed during flooding in roots of tomato plants. *J Biol Chem.* 270(23): 14056-14061.
- Pariyar SR, Dababat AA, Sannemann W, Erginbas-Orakci G, Elashry A, Siddique S, Morgounov A, Leon J, Grundler FM (2016) Genome-wide association study in wheat identifies resistance to the cereal cyst nematode *Heterodera filipjevi*. *Phytopathology.* 106(10): 1128-1138.
- Qi B, Yang Y, Yin Y, Xu M, Li H (2014) *De novo* sequencing, assembly, and analysis of the *Taxodium 'Zhongshansa'* roots and shoots transcriptome in response to short-term waterlogging. *BMC Plant Biol.* 14(201): 1471-2229.
- Qi XH, Xu XW, Lin XJ, Zhang WJ, Chen XH (2012) Identification of differentially expressed gene in cucumber (*Cucumis sativas* L.) root under waterlogging stress by digital gene expression profile. *Genomics.* 99: 160-168.
- Rajhi I, Yamauchi T, Takahashi H, Nishiuchi S, Shiono K, Watanabe R, Mliki A, Nagamura Y, Tsutsumi N, Nishizawa NK, Nakazono M (2011) Identification of genes expressed in maize root cortical cells during lysigenous aerenchyma formation using laser microdissection and microarray analyses. *New Phytol.* 190: 351-368.
- Rauf M, Arif M, Fisahn J, Xue GP, Balazadeh S, Mueller-Roeber B (2013) NAC transcription factor SPEEDY HYPOPLASTIC GROWTH regulates flooding-induced leaf movement in Arabidopsis. *Plant Cell.* 25: 4941-4955.
- Ren CG, Kong CC, Yan K, Zhang H, Luo YM, Xie ZH (2017) Elucidation of the molecular responses to waterlogging in *Sesbania cannabina* roots by transcriptome profiling. *Sci Rep-UK.* 7: 1-12.
- Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu G (2000) Arabidopsis transcription factors: Genome-wide comparative analysis among eukaryotes. *Science.* 290(5499): 2105-2110.
- Rivera-Mendes YD, Cuenca JC, Romero HM (2016) Physiological responses of oil palm (*Elaeis guineensis* Jacq.) seedlings under different water soil conditions. *Agronomia Colombiana.* 34: 163-171.
- Sasidharan R, Hartman S, LIU Z, Martopawiro S, Sajeev N, Veen H, Elaine Y, Voesenek LACJ (2018) Signal dynamics and interactions during flooding stress. *Plant Physiol.* 176: 1106-1117.
- Sasidharan R, Voesenek LACJ (2015) Ethylene-mediated acclimations to flooding stress. *Plant Physiol.* 169: 3-12.
- Sharma M, Pandey GK (2016) Expansion and function of repeat domain proteins during stress and development in plants. *Front Plant Sci.* 6(1218): 1-15.
- Shu Y, Zhang J, Ao Y, Song L, Guo C (2015) Analysis of the *Thinopyrum elongatum* transcriptome under water deficit stress. *Int J Genomics.* 2015: 265791.
- Singh R, Ong-Abdullah M, Low ETL, Manaf MAA, Rosli R, Nookiah R, Ooi LCL, 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 (2013) Oil palm genome sequence reveals divergence of interfertile species in old and new worlds. *Nature.* 500: 335-339.
- Steffens B, Rasmussen A (2016) The physiology of adventitious roots. *Plant Physiol.* 170: 603-617.
- Stein O, Granot D (2019) An overview of sucrose synthases in plants. *Front Plant Sci.* 10: 95.
- Suzuki N, Miller G, Morales J, Shulaev V, Torres MA, Mittler R (2011) Respiratory burst oxidases: the engines of ROS signaling. *Curr Opin Plant Biol.* 14(6): 691-699.
- Takahashi H, Yamauchi T, Rajhi I, Nishizawa NK, Nakazono M (2015) Transcript profiles in cortical cells of maize primary root during ethylene-induced lysigenous aerenchyma formation under aerobic conditions. *Ann Bot.* 115(6): 879-894.
- Tamang BG, Magliozzi JO, Maroof MAS, Fukao J (2014) Physiological and transcriptomic characterization of submergence and reoxygenation responses in soybean seedlings. *Plant Cell Environ.* 37: 2340-2365.
- Tenhaken R (2015) Cell wall remodeling under abiotic stress. *Front Plant Sci.* 5: 771.
- Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn, JL, Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc.* 7(3): 562-578.
- Tsuchiya M, Satoh S, Iwai H (2015) Distribution of XTH, expansin, and secondary-wall-related *CesA* in floral and fruit abscission zones during fruit development in tomato (*Solanum lycopersicum*). *Front Plant Sci.* 6: 323.
- Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JA (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.* 35: W71-W74.
- van Veen H, Hustroph A, Barding GA, Vergeer-van Eijk M, Welschen-Evertman RAM, Pedersen O, Visser EJW, Larive CK, Pierik R, Bailey-Serres J (2013) Two *Rumex* species from contrasting hydrological niches regulate flooding tolerance through distinct mechanisms. *Plant Cell.* 25: 4691-4707.

- Virdi AS, Singh S, Singh P (2015) Abiotic stress responses in plants: roles of calmodulin-regulated proteins. *Front Plant Sci.* 6: 809.
- Voeselek LACJ, Bailey-Serres J (2015) Flood adaptive traits and processes: an overview. *New Phytol.* 206(1): 57-73.
- Vriezen WH, Hulzink R, Mariani C, Voeselek LACJ (1999) 1-amino cyclopropane-1-carboxylate oxidase activity limits ethylene biosynthesis in *Rumex palustris* during submergence. *Plant J.* 11: 1265-1271.
- Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: A revolutionary tool for transcriptomics. *Nat Rev Genet.* 10: 57-63.
- Watanabe K, Nishiuchi S, Kulichikhin K, Nakazono M (2013) Does suberin accumulation in plant roots contribute to waterlogging tolerance? *Front Plant Sci.* 4: 178.
- Wei J, Hai-lan L, Yuan-qi W, Su-Zhi Z, Jian L, Yan-li L, Qi-Lin T, Ting-Zhao R (2016) *De novo* assembly of *Zea nicaraquensis* root transcriptome identified 5261 full-length transcripts. *J Integr Agr.* 15(6): 1207-1217.
- Wu J, Zhao HB, Yu D, Xu X (2017) Transcriptome profiling of the floating-leaved aquatic plant *Nymphoides peltata* in response to flooding stress. *BMC Genomics.* 18: 119.
- Yamauchi T, Fukazawa A, Nakazono M (2017) METALLOTHIONEIN genes encoding ROS scavenging enzymes are down-regulated in the root cortex during inducible aerenchyma formation in rice. *Plant Signal Behav.* 12(11): e1388976.
- Yamauchi T, Shimamura S, Nakazono M, Mochizuki T (2013) Aerenchyma formation in crop species: A review. *Field Crop Res.* 152: 8-16.
- Yamauchi T, Yoshioka M, Fukazawa A, Mori H, Nishizawa NK, Tsutsumi N, Yoshioka H, Nakazono M (2017) An NADPH Oxidase RBOH functions in rice roots during lysigenous aerenchyma formation under oxygen-deficient conditions. *Plant Cell.* 29(4): 775-790.
- Yang J, Worley E, Udvardi M (2014) A NAP-AAO3 regulatory module promotes chlorophyll degradation via ABA biosynthesis in Arabidopsis leaves. *Plant Cell.* 26: 4862-4874.
- Zhang J-Y, Huang S-N, Mo Z-H, Xuan J-P, Jia X-D, Wang G, Guo Z-R (2015) *De novo* transcriptome sequencing and comparative analysis of differentially expressed genes in kiwifruit under waterlogging stress. *Mol Breeding.* 35: 208.
- Zhang J-Y, Wang Q-J, Guo Z-R (2012) Progresses on plant AP2/ERF transcription factor. *Hereditas.* 34(7): 835-847.
- Zhang P, Lyu D, He J, Qin S (2017) Physiological and *de novo* transcriptome analysis of the fermentation mechanism of *Cerasus sachalinensis* roots in response to short-term waterlogging. *BMC Genomics.* 18(649): 1-14.
- Zhao N, Li C, Yan Y, Cao W, Song A, Wang H, Chen S, Jiang J, Chen F (2018) Comparative transcriptome analysis of waterlogging-sensitive and waterlogging-tolerant *Chrysanthemum morifolium* cultivars under waterlogging stress and reoxygenation conditions. *Int J Mol Sci.* 19(5): pii: E1455.
- Zhou B, Zhang L, Ullah A, Jin X, Yang X, Zhang X (2016) Identification of multiple stress responsive genes by sequencing a normalized cDNA library from sea-land cotton (*Gossypium barbadense* L.). *PLOS ONE.* 11(3): e0152927.
- Zou X, Tan X, Hu C, Zeng L, Lu G, Fu G, Cheng Y, Zhang X (2014) The transcriptome of *Brassica napus* L. roots under waterlogging at the seedling stage. *Int J Mol Sci.* 14: 2637-2651.