*Journal of Oil Palm Research Vol. 35 (3) September 2023 p. 528-537*
*DOI: https://doi.org/10.21894/jopr.2022.0062*

# PROTEOMICS OF OIL PALM SOMATIC EMBRYOGENESIS REVEALS THE DIFFERENTIALLY EXPRESSED PROTEINS AS CANDIDATES FOR BIOMARKER DEVELOPMENT

**KAMOLWAN KHIANCHAIKHAN[1#]; SUVICHARK AROONLUK[1#]; NARUMON PHAONAKROP[2]; SITTIRUK ROYTRAKUL[2]; MANTIRA SUKSIRT[1]; PAPHAWARIN PINYOKHAM[1]; MYA THUZAR[3] and CHATCHAWAN JANTASURIYARAT[1,4*]**

## ABSTRACT

*Tissue culture through somatic embryogenesis is the best method to produce the true-to-type of an elite oil palm plantlet. However, the mechanism underlying this process in oil palm is still unknown. We aimed to identify differentially expressed proteins during oil palm somatic embryogenesis using embryogenic callus, somatic embryo maturation and plantlet stages. Total proteins were extracted followed by tryptic enzyme digestion. The tryptic digested peptides were examined by nano LC-MS/MS. Identified proteins were classified based on biological process, molecular function and cellular components. Twenty-seven differentially expressed proteins were validated at the transcript level using qRT-PCR. These proteins were involved in plant growth and development, gene regulation, signalling, hormone response and stress response. Moreover, the differentially expressed proteins were categorised and reported as the candidate proteins for the development of biomarkers, which could be used to differentiate the embryo's developmental stages. The information on identified proteins obtained from this study will serve as the foundation for the understanding of the oil palm somatic embryogenesis in the tissue culture process.*

**Keywords:** nano LC-MS/MS, proteome, tissue culture.

**Received:** 24 April 2022; **Accepted:** 29 August 2022; **Published online:** 28 September 2022.

## INTRODUCTION

Oil palm is one of the most efficient oil-bearing crops in the world. With the world's consumption of palm oil increasing, many oil palm-producing countries aim to expand their oil palm plantation area. Oil palm, as a cross-pollinated crop, is propagated through seed production, resulting in segregated yields. Therefore, tissue culture is the best method to produce true-to-type elite oil palm plantlets. The oil palm tissue culture is initiated by embryogenic callus formation and then undergoes somatic embryogenesis. Oil palm somatic embryogenesis usually takes 7-8 months (Thuzar *et al*., 2011). Very little information about somatic embryogenesis in

[1] Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand.

[2] National Center for Genetic Engineering, and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Pathum Thani 12120, Thailand.

[3] Rice Science Center, Kasetsart University, Kamphaeng Saen, Nakorn Pathom 73140, Thailand.

[4] Center for Advanced Studies in Tropical Natural Resources, National Research University-Kasetsart (CASTNAR, NRU-KU), Kasetsart University, Bangkok 10900, Thailand.

[#] Authors contributed equally to this work.

[*] *Corresponding author email: fscicwj@ku.ac.th*

oil palm is reported. Identifying genes and proteins involved in this process will facilitate a better understanding and help to improve this process in oil palm tissue culture.

Somatic embryogenesis is a process used for large-scale plant regeneration, in which somatic cell has the ability to dedifferentiate or re-differentiate and give rise to somatic embryos. Several genes including *Somatic embryogenesis receptor kinase* (*SERK*) (Jiménez-Guillen *et al*., 2019), *Wuschell* (*WUS*) (Bouchabké-Coussa *et al*., 2013), *Leafy cotyledon* (*LEC*) (Guo *et al*., 2013), and *Baby Boom* (*BBM*) (Florez *et al*., 2015) were reported to have important roles in this process.

A biomarker has been defined as a biological molecule such as genes, proteins *etc*. that have the potential to indicate a particular developmental stage, tissue or process. These genes/proteins are specifically exhibited during the process or specific tissues (Aroonluk *et al*., 2018; 2020; Sruthilaxmi and Babu, 2020). Recently, proteomic analysis has been used to identify proteins, which involve in somatic embryogenesis in many plant species including *Musa* spp. (Kumaravel *et al*., 2020), Brazilian pine (Borges Araujo *et al*., 2022), and winter triticale (Krzewska *et al*., 2021). The differentially expressed proteins during the early embryogenic stages including primary callus and pro-embryogenic callus obtained from zygotic embryos were identified and characterised in oil palm (Silva *et al*., 2014). To further identify and characterise proteins during oil palm somatic embryogenesis, we performed proteomic analysis of embryogenic callus, somatic embryo maturation including globular, torpedo and cotyledon stages, and plantlet stage during oil palm somatic embryogenesis to identify differentially expressed proteins and to validate their expression at mRNA level. These identified differentially expressed proteins are the important keys to help reveal the biological mechanism of oil palm somatic embryogenesis and could be used as biomarkers for the different developmental stages during oil palm somatic embryogenesis.

## MATERIALS AND METHODS

### Plant Materials

*Tenera*, a common commercial oil palm type, is produced by crossing between a *dura* female and a *pisifera* male. Fifteen-week-old *tenera* zygotic embryo explants from *Golden Tenera* (KB) variety were cultured on an N6 medium supplemented with 2.0 mg/L of 2,4-D to induce embryogenic callus under dark conditions until globular embryos appeared. After three months, embryogenic callus was transferred to an N6 medium supplemented with 0.1 mg/L of 2,4-D, 0.16 g/L of putrescine, 0.5 g/L of casein amino acids, 30 g/L of sucrose, 2 g/L of phytagel, and 2 g/L of activated charcoal under light condition for 16 hr photoperiod until somatic embryo started to come out. After six months, small shoots were then transferred to a modified N6 medium containing 0.5 g/L of activated charcoal and 30 g/L of sucrose under light conditions for 16 hr photoperiod to regenerate new plantlets. The embryogenic callus, globular, torpedo, cotyledon, and plantlet samples were collected for protein extraction (*Figure 1*).

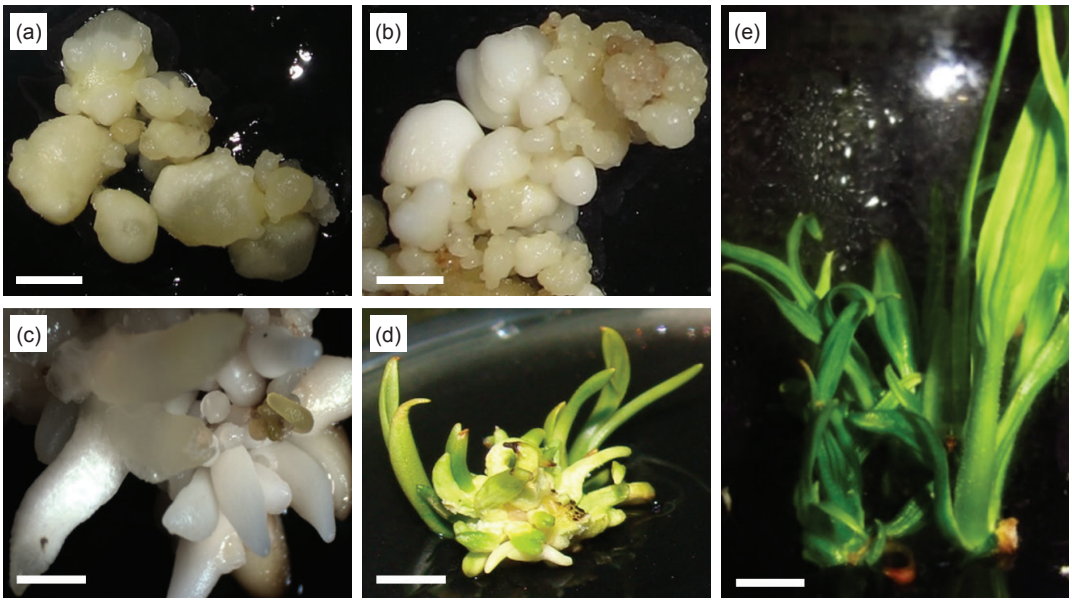


*Figure 1. The oil palm tissue samples at 5 stages; (a) embryogenic callus at 3 months on callus induction medium (bar = 5 mm), (b) globular at 1 month on somatic embryo maturation medium (bar = 5 mm), (c) torpedo at 3 months on somatic embryo maturation medium (bar = 3 mm), (d) cotyledon at 5 months on somatic embryo maturation medium (bar = 5 mm), and (e) plantlet at 1 month on plantlet regeneration medium (bar = 8 mm).*

**Protein Extraction, In-gel Digestion, and Protein Identification Using Nano LC-MS/MS**

The protein extracts from oil palm tissue samples were performed according to Kaweewong *et al*. (2013). Samples were ground in liquid nitrogen and dissolved in 0.5% SDS, mixed for 30 min at room temperature and then centrifuged at 13 201 g for 15 min at 4ºC. The supernatant was transferred to a new microcentrifuge tube then added an equal volume of cooled acetone and incubated for 2 hr at -20ºC. To collect the pellet, the samples were centrifuged at 13 201 g for 15 min at 4ºC. Washed the pellet three times with cooled acetone and centrifuged at 13 201 g for 15 min at 4ºC then dissolved in 0.5% SDS and desalted by ZebaTM Desalt spin column (ThermoFisher Scientific, USA) according to the manufacturer's instruction. The Lowry method was used to measure the concentration of soluble protein and stored at -80ºC. Total of 10 µg of protein samples were separated using 15% gel denaturing discontinuous polyacrylamide gel electrophoresis (SDS-PAGE) and stained with Pierce™ Silver Stain Kit (ThermoFisher Scientific, USA).

The protein gel was cut to gel plugs 1 x 1 $mm^3$, washed and dissolved with 200 µL of 100% acetonitrile and soaked with 50 µL of 10 mM dithiothreitol (DTT) in 10 mM ammonium bicarbonate for 1 hr to breakdown the disulfide bridge and block the reunion of the disulfide bridge with 50 µL of 100 mM iodoacetamide in 10 mM ammonium bicarbonate under the dark light for 1 hr. Protein gels were then dehydrated by 100% acetonitrile and dried at room temperature. Dried protein gels were digested by adding 10 ng of Trypsin in 50% acetonitrile and 10 mM ammonium bicarbonate and kept at room temperature for 20 min before adding 30 µL of 30% acetonitrile solution and incubated overnight. The digested peptide solution was re-extracted with 50% acetonitrile in 0.1% trifluoroacetic acid and 70% acetonitrile in 0.1% trifluoroacetic, respectively. The peptide solution was dried using a Speed Vacuum concentrator and resuspended in 0.1% formic acid (FA) for Nano LC-MS/MS analysis.

The digested peptide samples were subjected to HCT-Ultra PTM Discovery System (Bruker Daltonik, Bremen, Germany) coupled to the Ultimate 3000 LC system (Dionex, USA) using electrospray at the flow rate of 300 nL/min to a nanocolumn (PepSwift monolithic column 100 µm i.d. x 50 mm). Mobile phases of solvent A (0.1% formic acid) and solvent B (80% acetonitrile and 0.1% formic acid) were used to elute peptides with a linear gradient from 10%-70% of solvent A at 0-13 min (the time-point of retention time) following 90% of solvent B at 13-15 min, and final elution of 10% of solvent B at 15-20 min to remove remaining salt. The resolution of the MS step is 0.6 and the accuracy is 0.15 u (m/z). A total of three replicates were performed.

**Bioinformatics and Data Analysis**

The DeCyder MS 2.0 analysis software (GE Healthcare, USA) was used to measure the quantitative protein intensity based on peptide MS signal intensities of individual LC-MS analysed data. The PepDetect module was used to produce ion peptides at the data set; the mass resolution is 0.6, typical peak width 0.5, TOF resolution 10 000, charge status from 1 to 10 and m/z 22 shift tolerance 0.1 u. The PepMatch module was used to evaluate the signal intensity maps. The highest intensity sample was used as a control presenting the relative abundance of peptides as 2 log intensities with mass tolerance set to 0.5 amu. The criteria of more than a 2-fold change of an average abundance ratio were used to determine the induced protein with a significant standard t-test and one-way ANOVA $p<0.05$. All MS/MS spectra from the Decyder MS analysis were performed by applying the global variable mode of carbamidomethyl (C), peptide charge state ($1^+$, $2^+$ and $3^+$), and m/z tolerance 0.1 u and searched against NCBI FTP site of *Elaeis guineensis* genome database to identify matching peptide using Mascot software (Matrix Science, UK). Identified proteins were filtered with a one-way ANOVA $p<0.05$. BSA was used as an internal standard to normalise protein intensities. Bioinformatics and Evolutionary Genomics Tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) was used to determine the similarity and differential protein expression. Gene ontology was used to characterise the function of the identified proteins using appropriate protein annotation that was reported in the UniProt protein database (http://www.uniprot.org/), PANTHER classification system (http://pantherdb.org/) and previously reported proteins. The heat map of differentially expressed proteins from each sample was generated by average linkage clustering using the online Heatmapper (http://www.heatmapper.ca/).

**Quantitative Real-Time PCR Verification**

Five stages of tissue development including embryogenic callus, globular, torpedo, cotyledon, and plantlet were used for total RNA extraction using Spin Plant RNA (STRATEC Molecular, Germany) according to the manufacturer's protocol with minor modification with DNase I (Vivantis, USA) treatment. The first strand of cDNA was synthesised using a cDNA synthesis kit (Biotechrabbit GmbH, Germany) according to the manufacturer's instructions. The cDNA content of *E. guineensis* elongation factor alpha-1 (NCBI accession number XM_019850296) was used to normalise the expression level of the candidate genes. The qRT-PCR was performed with the

specific primers using KAPA SYBR® FAST qPCR kit Mastermix (2x) Universal (Kapa Biosystems, USA) according to the manufacturer's instruction on Eppendorf Mastercycler ep Realplex (Thermo Fisher Scientific, USA). Three replications were conducted. The expression level was calculated as fold change using the double delta Ct values with the callus expression level as a calibrator.

## RESULTS AND DISCUSSION

Somatic embryogenesis is an important process for *in vitro* plant micropropagation. It is described as the differentiation of a somatic cell to form an embryogenic cell. In this study, we identified the differentially expressed proteins during the oil palm somatic embryogenesis. A total of 800 unique proteins were identified, in which 742, 722, 730, 755 and 756 were presented in callus, globular, torpedo, cotyledon and plantlet respectively (*Figure 2*). A total of 584 identified proteins were shared in all stages in which 216 identified proteins were specifically expressed in only certain stages and these proteins could be developed as the oil palm somatic embryogenesis biomarkers. Eight groups of identified proteins were classified based on their expression patterns (*Table 1*). Group 1 is composed of a serine/threonine-protein kinase Nek5, which was specifically expressed in callus and globular stages. Group 2 was composed of 11 identified proteins, which were specifically expressed in callus, globular and torpedo stages including auxin response factor 23-like, rust resistance kinase Lr10-like, importin subunit beta-1-like, vegetative cell wall protein gp1 *etc*. Group 3 is composed of a putative transporter arsB, which was specifically expressed in globular and torpedo stages. Group 4 is composed of a pentatricopeptide repeat-containing protein *At*5g14770, mitochondrial, which was specifically expressed in globular and cotyledon stages. Group 5 is composed of a transmembrane protein 184C-like, which was specifically expressed only in the torpedo stage. Group 6 is composed of three identified proteins, which were specifically expressed in globular, torpedo and cotyledon stages including phospholipase D delta-like isoform X2, solanesyl-diphosphate synthase 3, chloroplastic and LIM domain-containing serine/threonine-protein kinase DDB_G0286997. Group 7 is composed of two identified proteins, which were specifically expressed in globular, torpedo, and plantlet stages including a protein kinase PINOID-like and an uncharacterised protein. Group 8 is composed of seven identified proteins, which were specifically expressed in torpedo, cotyledon and plantlet stages including delta (8)-fatty-acid desaturase-like, RAN GTPase-activating protein, U4/U6 small nuclear ribonucleoprotein Prp3-like, fasciclin-like arabinogalactan protein 2, RING-H2 finger protein ATL13-like, metallothionein-like protein type 2 and one uncharacterised protein.

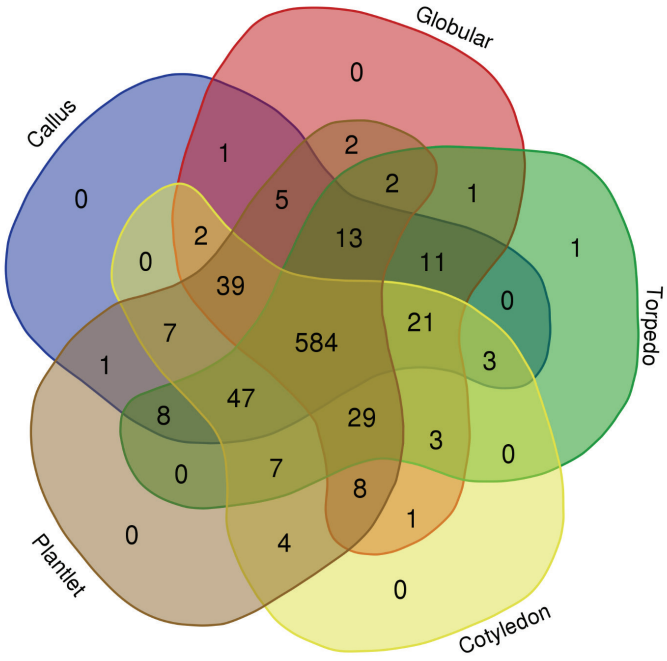


*Figure 2. Venn diagram of the differentially expressed proteins during oil palm somatic embryogenesis; callus (blue), globular (red), torpedo (green), cotyledon (yellow) and plantlet (brown). The number indicates the number of identified proteins.*

The 27 identified differentially expressed proteins from eight groups were selected for transcriptional expression verification using quantitative Real-Time RT-PCR. These proteins are potential candidates for biomarker development. The result showed that the mRNA expression level was consistence with the level of protein expression (*Table 1*). This result corresponded to a previous study of the genes associated with cDNA-AFLP of oil palm somatic embryogenesis (Pattarapimol *et al*., 2015). However, the correlation between protein and mRNA expression levels partially overlaps but is not identical (Aroonluk *et al*., 2020). This result indicated that many genes/proteins were induced in embryogenic callus to initiate somatic embryogenesis and these genes/proteins were differentially expressed at globular, torpedo, cotyledon and plantlet stages.

**Functional Characterisation of Identified Proteins Based on Gene Ontology**

Identified proteins were functionally classified based on their gene ontology (GO); (a) biological process, (b) molecular function and (c) cellular component. For the biological process, 10.30% was classified in the metabolic process, 9.84% was classified in the transcription and 6.67% was classified in the transportation process. For the molecular function, 12.53% were classified as the binding proteins, 6.21% were classified as kinases and 5.50% were classified as transcription factors. For the cellular component, 21.55% was classified as nuclear proteins, 14.64% was classified as membrane proteins and 9.72% was classified as cytoplasmic proteins (*Figure 3*). However, the majority of identified proteins had an unknown function.

**Differentially Expressed Identified Proteins During Oil Palm Somatic Embryogenesis**

Identified proteins were hierarchical clustering based on their expression pattern (*Figure 4*). Five clusters were generated as clusters A to E. Cluster A is composed of 128 identified proteins, which were highly expressed in callus, globular, torpedo and plantlet stages but low expression in the cotyledon stage. Cluster B is composed of 300 identified proteins, which were highly expressed in callus, globular and torpedo stages but low expression in cotyledon and plantlet stages. Cluster C is composed of 108 identified proteins, which can be divided into 2 sub-clusters C1 and C2. Sub-cluster C1 was highly expressed in callus, globular and plantlet stages but low expression in torpedo and cotyledon stages. Sub-cluster C2 was highly expressed in all stages except in the torpedo stage. Cluster D is composed of 186 identified proteins, which can be divided into 2 sub-clusters D1 and D2. Sub-cluster D1 was highly expressed in cotyledon and plantlet stages but low expression in callus, globular and torpedo stages. Sub-cluster D2 was highly expressed in all stages except in the globular stage. Finally, cluster E is composed of 78 identified proteins, which were highly expressed in globular, torpedo, cotyledon, and plantlet stages but with low expression in the callus stage. These differentially expressed proteins could be developed as the biomarkers of the oil palm somatic embryogenesis.

Several proteins, identified in this study, were previously reported to be involved in somatic embryo development. Serine/threonine-protein kinase Nek5 is a member of the NIMA-related kinases (NEKs) family, involved largely in cell cycle control by regulating microtubule organisation during epidermal cell expansion (Motose *et al*., 2011). This protein contributed to controlling cell division for the somatic embryogenesis acquisition process. Receptor protein kinase-like protein ZAR1, a member of the RLK/Pelle kinase family, functions as a membrane integrator for extrinsic cues, $Ca^{2+}$ signal and G protein signalling to regulate the division of zygote and the cell fate of its daughter cells (Yu *et al*., 2016). Putative transporter arsB was specifically expressed in globular and torpedo stages and they could be used as protein biomarkers to determine the somatic embryo stage. This protein is one of the arsenite transporter genes, involved in arsenic and antimony transport pathways using vacuolar transporters of arsenic-phytochelatin complexes in plants (Maciaszczyk-Dziubinska *et al*., 2012). Fasciclin-like arabinogalactan protein 2 is a member of arabinogalactan proteins (AGPs), which have a significant role in the process of competence acquisition and the induction of shoot development (Johnson *et al*., 2003). RAN GTPase-activating protein 1-like is a member of proteins that bind to Ran (Ras-related nuclear small GTP-binding protein). It involves root growth and development and regulation of auxin-induced mitotic progression in root tips (Kim and Roux, 2003). These identified proteins can be used as biomarkers for oil palm somatic embryogenesis.

Moreover, two identified proteins have functions in cell walls to regulate cell shape and the direction of cell differentiation during somatic embryogenesis. First, vegetative cell wall protein gp1 is a constituent chaotropic-soluble glycoprotein in the cell wall and controls mechanisms of cell wall remodelling, signalling pathway from the apoplast to the cell wall, cell wall-sensing process and exchange of cell-cell information (Ringli, 2010; Voigt *et al*., 2009). The other, non-specific lipid-transfer protein is a small protein, involves in plant growth and development (Potocka *et al*., 2012).

Several transcription factors known to be involved in somatic embryogenesis were also identified in this study including the CCCH proteins,

**TABLE 1. LIST OF THE SPECIFICALLY EXPRESSED PROTEINS AND mRNA EXPRESSION VALIDATION**

| Accession No. | Protein name | MOWSE score | Protein abundance | | | | | mRNA expression level | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Callus | Globular | Torpedo | Cotyledon | Plantlet | Callus | Globular | Torpedo | Cotyledon | Plantlet |
| Group 1 | | | | | | | | | | | | |
| gi\|743858642 | PREDICTED: serine/threonine-protein kinase Nek5 | 5.69 | 16.07 | 15.40 | 0.00 | 0.00 | 0.00 | 1.00 | 0.63 | 2.56 | 1.42 | 56.49 |
| Group 2 | | | | | | | | | | | | |
| gi\|743858544 | PREDICTED: auxin response factor 23-like | 4.68 | 15.81 | 16.46 | 15.27 | 0.00 | 0.00 | 1.00 | 1.68 | 1.51 | 1.17 | 1.84 |
| gi\|743768063 | PREDICTED: rust resistance kinase Lr10-like | 5.24 | 13.06 | 14.52 | 12.99 | 0.00 | 0.00 | 1.00 | 0.21 | 4.77 | 3.45 | 4.95 |
| gi\|743842857 | PREDICTED: importin subunit beta-1-like | 5.07 | 14.45 | 18.89 | 16.48 | 0.00 | 0.00 | 1.00 | 2.46 | 3.01 | 1.69 | 4.36 |
| gi\|743858365 | PREDICTED: vegetative cell wall protein gp1 | 4.41 | 16.17 | 18.17 | 16.88 | 0.00 | 0.00 | 1.00 | 1.92 | 16.49 | 36.11 | 48.95 |
| gi\|743788936 | PREDICTED: probable LRR receptor-like serine/threonine-protein kinase *At*1g06840 isoform X1 | 3.92 | 11.23 | 15.57 | 17.37 | 0.00 | 0.00 | 1.00 | 2.14 | 8.66 | 7.27 | 3.08 |
| gi\|743835428 | PREDICTED: NAC domain-containing protein 78 isoform X1 | 3.49 | 17.25 | 13.73 | 16.82 | 0.00 | 0.00 | 1.00 | 1.05 | 4.89 | 2.43 | 4.83 |
| gi\|743761936 | PREDICTED: peptidyl-prolyl cis-trans isomerase CYP37, chloroplastic | 1.54 | 19.64 | 17.01 | 17.64 | 0.00 | 0.00 | 1.00 | 1.35 | 2.20 | 1.44 | 3.45 |
| gi\|743757543 | PREDICTED: receptor protein kinase-like protein ZAR1 | 1.40 | 17.11 | 16.19 | 18.29 | 0.00 | 0.00 | 1.00 | 10.17 | 3.29 | 3.56 | 7.42 |
| gi\|743801725 | PREDICTED: dr1-associated corepressor homolog isoform X1 | 3.67 | 21.96 | 20.08 | 21.35 | 0.00 | 0.00 | 1.00 | 0.89 | 2.38 | 1.78 | 4.89 |
| gi\|743872177 | PREDICTED: uncharacterised protein LOC105033396 isoform X3 | 1.14 | 18.11 | 16.95 | 19.87 | 0.00 | 0.00 | 1.00 | 9.07 | 7.43 | 4.69 | 7.59 |
| gi\|743779448 | PREDICTED: uncharacterised protein LOC105044092 | 3.81 | 16.17 | 18.17 | 16.88 | 0.00 | 0.00 | 1.00 | 1.64 | 1.77 | 1.43 | 1.65 |
| Group 3 | | | | | | | | | | | | |
| gi\|743776619 | PREDICTED: putative transporter arsB | 5.06 | 0.00 | 13.73 | 16.03 | 0.00 | 0.00 | 1.00 | 3.57 | 52.06 | 36.31 | 89.68 |
| Group 4 | | | | | | | | | | | | |
| gi\|743831750 | PREDICTED: pentatricopeptide repeat-containing protein *At5*g14770, mitochondrial | 4.72 | 0.00 | 17.43 | 0.00 | 19.14 | 0.00 | 1.00 | 2.11 | 7.40 | 4.56 | 4.40 |
| Group 5 | | | | | | | | | | | | |
| gi\|743761857 | transmembrane protein 184C-like | 1.52 | 0.00 | 0.00 | 13.94 | 0.00 | 0.00 | 1.00 | 1.05 | 6.76 | 2.93 | 9.49 |
| Group 6 | | | | | | | | | | | | |
| gi\|743763273 | PREDICTED: phospholipase D delta-like isoform X2 | 9.94 | 0.00 | 21.40 | 19.12 | 21.40 | 0.00 | 1.00 | 1.05 | 0.64 | 0.44 | 0.00 |
| gi\|743758416 | PREDICTED: probable solanesyl-diphosphate synthase 3, chloroplastic | 9.78 | 0.00 | 22.50 | 22.32 | 20.51 | 0.00 | 1.00 | 0.81 | 5.06 | 1.27 | 3.28 |
| gi\|743756281 | PREDICTED: probable LIM domain-containing serine/threonine-protein kinase DDB_G0286997 | 2.78 | 0.00 | 19.80 | 18.95 | 20.90 | 0.00 | 1.00 | 1.49 | 7.13 | 3.85 | 7.28 |
| Group 7 | | | | | | | | | | | | |
| gi\|743886489 | PREDICTED: protein kinase PINOID-like | 6.72 | 0.00 | 21.95 | 20.46 | 0.00 | 13.44 | 1.00 | 0.61 | 4.47 | 2.65 | 9.98 |
| gi\|743765954 | PREDICTED: uncharacterised protein LOC105038960 | 7.37 | 0.00 | 21.29 | 20.98 | 0.00 | 18.35 | 1.00 | 1.88 | 3.29 | 2.30 | 3.16 |
| Group 8 | | | | | | | | | | | | |
| gi\|743800192 | PREDICTED: delta (8)-fatty-acid desaturase-like | 3.98 | 0.00 | 0.00 | 17.26 | 19.19 | 19.19 | 1.00 | 1.87 | 2.41 | 1.44 | 33.69 |
| gi\|743839144 | PREDICTED: RAN GTPase-activating protein 1-like | 2.86 | 0.00 | 0.00 | 18.31 | 20.11 | 20.11 | 1.00 | 1.61 | 1.67 | 1.16 | 2.31 |
| gi\|743759713 | PREDICTED: U4/U6 small nuclear ribonucleoprotein Prp3-like | 2.21 | 0.00 | 0.00 | 16.33 | 13.63 | 13.63 | 1.00 | 1.70 | 4.48 | 2.28 | 4.88 |
| gi\|743861480 | PREDICTED: fasciclin-like arabinogalactan protein 2 | 7.50 | 0.00 | 0.00 | 21.23 | 14.26 | 14.26 | 1.00 | 1.40 | 1.70 | 1.28 | 4.32 |
| gi\|743811642 | PREDICTED: RING-H2 finger protein ATL13-like | 2.42 | 0.00 | 0.00 | 20.76 | 19.18 | 19.18 | 1.00 | 1.25 | 4.52 | 2.86 | 20.74 |
| gi\|743841808 | PREDICTED: metallothionein-like protein type 2 | 10.08 | 0.00 | 0.00 | 18.46 | 18.46 | 17.58 | 1.00 | 1.57 | 7.62 | 3.48 | 27.18 |
| gi\|743770889 | PREDICTED: uncharacterised protein *At5*g41620-like | 13.77 | 0.00 | 0.00 | 17.88 | 19.11 | 19.11 | 1.00 | 2.43 | 3.07 | 1.64 | 3.07 |

Note: mRNA expression level of the identified genes is normalised with *E. guineensis elongation factor alpha-1* (NCBI accession number XM_019850296). The callus stage is used as a reference expression level to calculate the fold changes. The blue colour represents the level of gene expression. The higher intensity of the blue colour indicates up-regulated expression and the lower intensity indicates down-regulated expression.

which are zinc finger families containing C3H-type motif proteins (Pi *et al*., 2018). B-box domain family proteins regulate embryo and hypocotyl elongation (Shalmani *et al*., 2019). In addition, a basic leucine zipper transcription factor has been reported to be upregulated in embryo/endosperm in the seed germination stage, and in embryo 25 days after seed development in barley (Pourabed *et al*., 2015).

Somatic embryogenesis can also be regulated by endogenous plant hormones. We identified several hormone-responsive proteins, which were expressed during somatic embryos as plant developmental requirements. For example, an auxin response factor 23-like has a major role in plant growth and developmental processes (Li *et al*., 2016). Moreover, NAC domain-containing protein 78, which plays an important role in the regulation of the transcriptional reprogramming associated with plant stress responses was also identified in this study (Nuruzzaman *et al*., 2013). In addition, two proteins, which were downregulated in the globular stage of somatic embryogenesis (sub-cluster D2), have been revealed to be the repressor proteins. Agamous-like MADs box protein AGL62 plays an important role in endosperm development. Its expression was detected in endosperm but not in the embryos as the result is downregulated during the early stage of somatic embryogenesis (Kang *et al*., 2008). B3 domain-containing protein *Os*07g0563300 is a member of a VAL family, which encodes the repressor of embryogenesis related-gene, *LEC1* (Schneider *et al*., 2016).

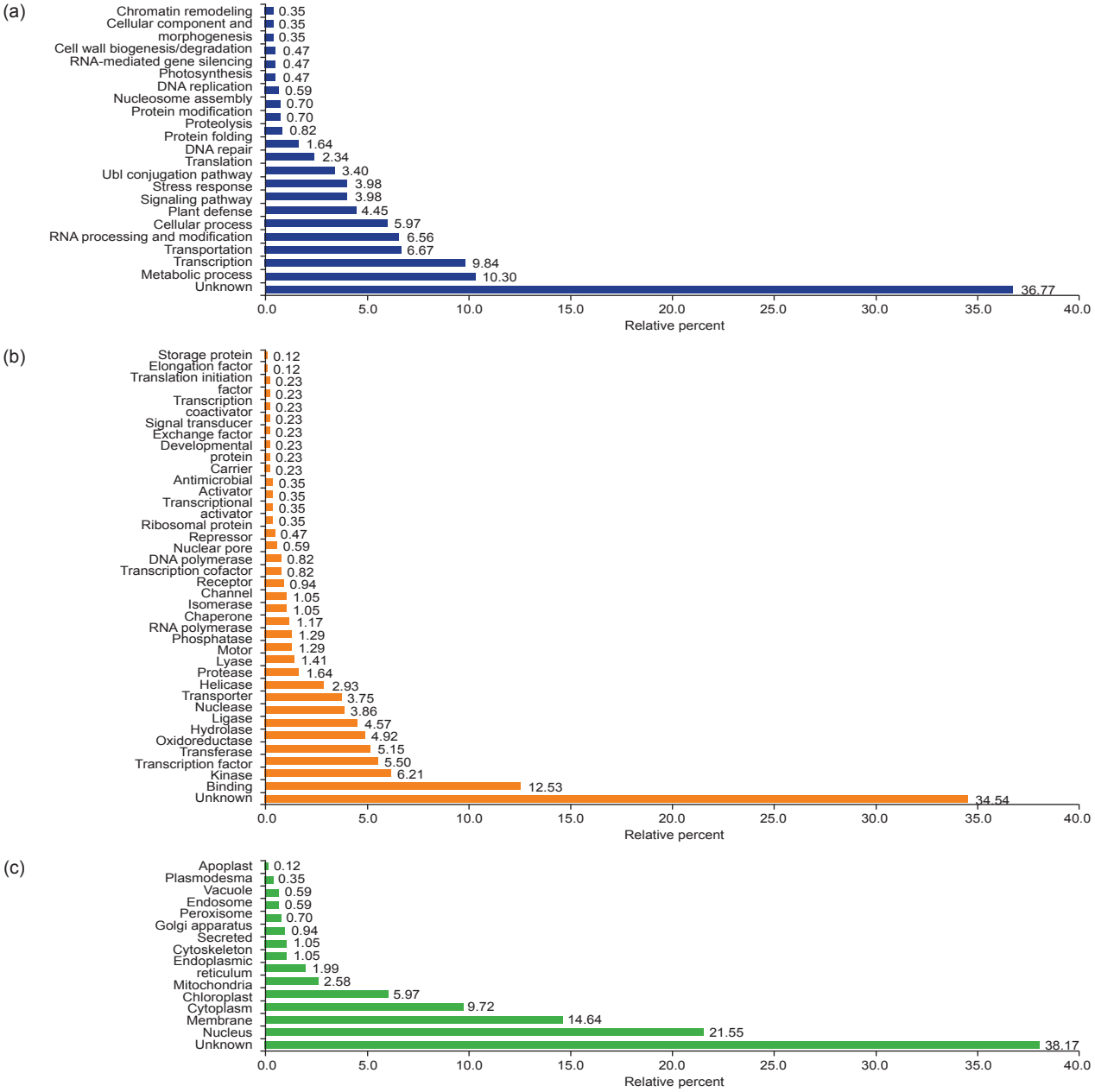


*Figure 3. Functional classification of identified proteins based on their gene ontology (GO); (a) biological process, (b) molecular function, and (c) cellular component.*

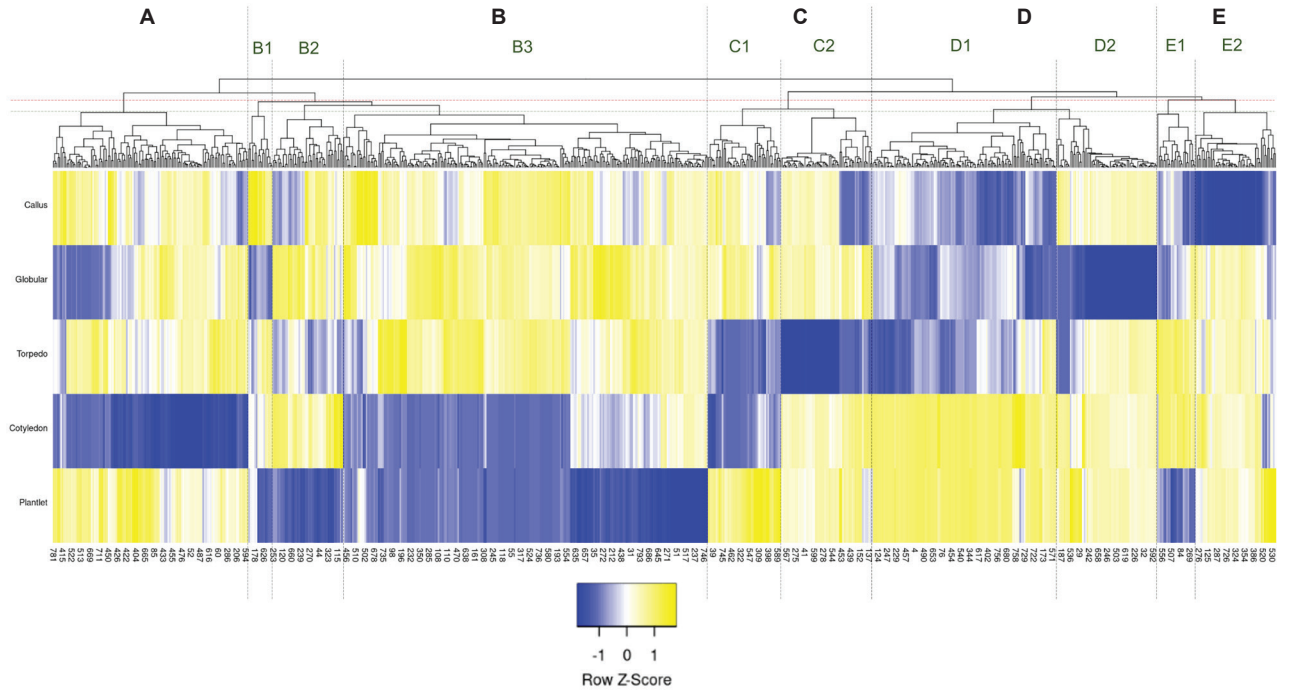


*Figure 4. Heat map of the differentially expressed proteins during somatic embryogenesis. Five clusters and sub-clusters were categorised based on the level of average normalised volume of each protein. The blue colour indicates a low level of expression and the yellow colour indicates a high level of expression. The level of expression is proportionate to the colour intensity.*

## CONCLUSION

In summary, we identified and characterised proteins that are expressed during oil palm somatic embryogenesis. The results revealed that several proteins are involved in plant growth and development, gene regulation, signalling, hormone response, and stress response. The identified proteins could be developed as biomarkers for the oil palm somatic embryogenesis. Furthermore, the information obtained from this study will facilitate the understanding of the regulation mechanism of oil palm somatic embryogenesis.

## ACKNOWLEDGEMENT

This research and innovation activity is funded by the National Research Council of Thailand (NRCT) as of the fiscal years 2017 and 2021, the National Research Council of Thailand (NRCT), and the Royal Golden Jubilee (RGJ) Ph.D. Program (Grant No. PHD/0001/2560), through the National Research Council of Thailand (NRCT), and Thailand Research Fund (TRF), Kasetsart University Research and Development Institute (KURDI).

## REFERENCES

Aroonluk, S; Roytrakul, S and Jantasuriyarat, C (2020). Identification and characterisation of phosphoproteins in somatic embryogenesis acquisition during oil palm tissue culture. *Plants (Basel)*, *9*: 36.

Aroonluk, S; Roytrakul, S; Yingchutrakul, Y; Kittisenachai, S and Jantasuriyarat, C (2018). Characterisation of glycoproteins during oil palm somatic embryogenesis. *Agric. Nat. Res.*, *52*: 430-438.

Borges Araujo, A J; Cerruti, G V; Zuccarelli, R; Rodriguez Ruiz, M; Freschi, L; Singh, R; Moerschbacher, B M; Floh, E I S and Wendt dos Santos, A L (2022). Proteomic analysis of S-Nitrosation sites during somatic embryogenesis in Brazilian pine, *Araucaria angustifolia* (Bertol.) Kuntze. *Front. Plant Sci.*, *13*: 902068.

Bouchabké-Coussa, O; Obellianne, M; Linderme, D; Montes, E; Maia-Grondard, A; Vilaine, F and Pannetier, C (2013). Wuschel overexpression promotes somatic embryogenesis and induces organogenesis in cotton (*Gossypium hirsutum* L.) tissues cultured *in vitro*. *Plant Cell Rep.*, *32*: 675-686.

Florez, S L; Erwin, R L; Maximova, S N; Guiltinan, M J and Curtis, W R (2015). Enhanced somatic embryogenesis in *Theobroma cacao* using the homologous BABY BOOM transcription factor. *BMC Plant Biol.*, *15*: 121.

Guo, F; Liu, C; Xia, H; Bi, Y; Zhao, C; Zhao, S; Hou, L; Li, F and Wang, X (2013). Induced expression of *AtLEC1* and *AtLEC2* differentially promotes

somatic embryogenesis in transgenic tobacco plants. *PLoS ONE*, *8*: e71714.

Jiménez-Guillen, D; Pérez-Pascual, D; Souza-Perera, R and Zúñiga-Aguilar, J J (2019). Cloning and molecular characterisation of a putative Habanero pepper SERK1 cDNA expressed during somatic and zygotic embryogenesis. *Electron. J. Biotechnol.*, *41*: 48-55.

Johnson, K L; Jones, B J; Bacic, A and Schultz, C J (2003). The fasciclin-like arabinogalactan proteins of *Arabidopsis*. A multigene family of putative cell adhesion molecules. *Plant Physiol.*, *133*: 1911-1925.

Kang, I H; Steffen, J G; Portereiko, M F; Lloyd, A and Drews, G N (2008). The AGL62 MADS domain protein regulates cellularization during endosperm development in *Arabidopsis*. *Plant Cell*, *20*: 635-647.

Kaweewong, K; Garnjanagoonchorn, W; Jirapakkul, W and Roytrakul, S (2013). Solubilization and identification of hen eggshell membrane proteins during different times of chicken embryo development using the proteomic approach. *Protein J.*, *32*: 297-308.

Kim, S H and Roux, S J (2003). An *Arabidopsis* ran-binding protein, *AtRanBP1c*, is a co-activator of ran GTPase-activating protein and requires the C-terminus for its cytoplasmic localization. *Planta*, *216*: 1047-1052.

Krzewska, M; Dubas, E; Gołębiowska, G; Nowicka, A; Janas, A; Zieliński, K; Surówka, E; Kopeć, P; Mielczarek, P and Żur, I (2021). Comparative proteomic analysis provides new insights into regulation of microspore embryogenesis induction in winter triticale (×*Triticosecale* Wittm.) after 5-azacytidine treatment. *Sci. Rep.*, *11*: 22215.

Kumaravel, M; Uma, S; Backiyarani, S and Saraswathi, M S (2020). Proteomic analysis of somatic embryo development in *Musa* spp. cv. Grand Naine (AAA). *Sci. Rep.*, *10*: 4501.

Li, S B; Xie, Z Z; Hu, C G and Zhang, J Z (2016). A review of auxin response factors (ARFs) in plants. *Front. Plant Sci.*, *7*: 47.

Nuruzzaman, M; Sharoni, A M and Kikuchi, S (2013). Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. *Front. Microbiol.*, *4*: 248.

Maciaszczyk-Dziubinska, E; Wawrzycka, D and Wysocki, R (2012). Arsenic and antimony transporters in eukaryotes. *Int. J. Mol. Sci.*, *13*: 3527-3548.

Motose, H; Hamada, T; Yoshimoto, K; Murata, T; Hasebe, M; Watanabe, Y; Hashimoto, T; Sakai, T and Takahashi, T (2011). NIMA-related kinases 6, 4, and 5 interact with each other to regulate microtubule organisation during epidermal cell expansion in *Arabidopsis thaliana*. *Plant J.*, *67*: 993-1005.

Pattarapimol, T; Thuzar, M; Vanavichit, A; Tragoonrung, S; Roytrakul, S and Jantasuriyarat, C (2015). Identification of genes involved in somatic embryogenesis development in oil palm (*Elaeis guineensis* Jacq.) using cDNA AFLP. *J. Oil Palm Res.*, *27*: 1-11.

Pi, B; He, X; Ruan, Y; Jang, J C and Huang, Y (2018). Genome-wide analysis and stress-responsive expression of CCCH zinc finger family genes in *Brassica rapa*. *BMC Plant Biol.*, *18*: 373.

Potocka, I; Baldwin, T C and Kurczynska, E U (2012). Distribution of lipid transfer protein 1 (LTP1) epitopes associated with morphogenic events during somatic embryogenesis of *Arabidopsis thaliana*. *Plant Cell Rep.*, *31*: 2031-2045.

Pourabed, E; Ghane Golmohamadi, F; Soleymani Monfared, P; Razavi, S M and Shobbar, Z S (2015). Basic leucine zipper family in barley: Genome-wide Characterisation of members and expression analysis. *Mol. Biotechnol.*, *57*: 12-26.

Ringli, C (2010). Monitoring the outside: Cell wall-sensing mechanisms. *Plant Physiol.*, *53*: 1445-1452.

Schneider, A; Aghamirzaie, D; Elmarakeby, H; Poudel, A N; Koo, A J; Heath, L S; Grene, R and Collakova, E (2016). Potential targets of *VIVIPAROUS1/ABI3-LIKE1* (*VAL1*) repression in developing *Arabidopsis thaliana* embryos. *Plant J.*, *85*: 305-319.

Shalmani, A; Jing, X Q; Shi, Y; Muhammad, I; Zhou, M R; Wei, X Y; Chen, Q Q; Li, W Q; Liu, W T and Chen, K M (2019). Characterisation of B-BOX gene family and their expression profiles under hormonal, abiotic and metal stresses in Poaceae plants. *BMC Genom.*, *20*: 27.

Silva, R D C; Carmo, L S T; Luis, Z G; Silva, L P; Scherwinski-Pereira, J E and Mehta, A (2014). Proteomic identification of differentially expressed proteins during the acquisition of somatic embryogenesis in oil palm (*Elaeis guineensis* Jacq.). *J. Proteom.*, *104*: 112-127.

Sruthilaxmi, C B and Babu, S (2020). Proteome responses to individual pathogens and abiotic conditions in rice seedlings. *Phytopathology, 110*: 1326-1341.

Thuzar, M; Vanavichit, A; Tragoonrung, S and Jantasuriyarat, C (2011). Efficient and rapid plant regeneration of oil palm zygotic embryos cv. *tenera* through somatic embryogenesis. *Acta Physiol. Plantarum*, *33*: 123-128.

Voigt, J; Frank, R and Wöstemeyer, J (2009). The chaotrope-soluble glycoprotein GP1 is a constituent of the insoluble glycoprotein framework of the *Chlamydomonas* cell wall. *FEMS Microbiol. Lett.*, *291*: 209-215.

Yu, T Y; Shi, D Q; Jia, P F; Tang, J and Li, H J (2016). The Arabidopsis receptor kinase ZAR1 is required for zygote asymmetric division and its daughter cell fate. *PLoS Genet.*, *12*: e1005933.