Agriculture and Natural Resources 52 (2018) 430–438

Contents lists available at ScienceDirect

Agriculture and Natural Resources

journal homepage: http://www.journals.elsevier.com/agriculture-and-natural-resources/

Original Article

Identification and characterization of glycoproteins during oil palm somatic embryogenesis

Suvichark Aroonluk,[a] Sittiruk Roytrakul,[b] Yodying Yingchutrakul,[b] Suthathip Kittisenachai,[b] Chatchawan Jantasuriyarat[a, c, *]

[a] Department of Genetics, Faculty of Science, Kasetsart University, Bangkok, 10900 Thailand
[b] National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), PathumThani, 12120, Thailand
[c] Center for Advanced Studies in Tropical Natural Resources, National Research University-Kasetsart University (CASTNAR, NRU-KU), Bangkok, 10900 Thailand

ARTICLE INFO

Article history:
Received 25 September 2017
Accepted 19 December 2017
Available online 22 November 2018

Keywords:
Glycoproteins
Nano-liquid chromatography-tandem mass spectrometry
Oil palm
Somatic embryogenesis

ABSTRACT

The objective of this experiment was to characterize differentially expressed glycoprotein in a somatic embryogenesis process during oil palm tissue culture. Embryogenic callus in the somatic embryo acquisition stages (globular, torpedo and cotyledonary) and oil palm plantlets were collected to extract total protein and isolated glycoprotein using a concanavalin A N-linked affinity column. Tryptic glycoproteins were isolated using nano-liquid chromatography-tandem mass spectrometry. In total, 383 glycoproteins were identified and analyzed. The percentages of existence of glycoprotein were observed, from highest to lowest amounts, in the metabolic process, binding and nuclear components, respectively. Different glycoproteins were involved in membrane trafficking machinery, signaling to stress and in hormonal and environmental response for plant growth and development. PREDICTED: protein TIC 40, chloroplastic-like was expressed in the globular, torpedo and plantlet stages. The protein can be further developed into a glycoprotein biomarker, as a biological indicator of the somatic embryo maturation stage. This study should assist deeper understanding of the important role of glycoprotein specificity in somatic embryogenesis that controls plant cell response and growth development during oil palm tissue culture.

Introduction

Oil palm (*Elaeis guineensis Jacq.*) belongs to the family Arecaceae and is a persistent monoecious crop that has separate male and female inflorescence on the same tree (Corley and Tinker, 2003). Oil palm is one of the most important oil-bearing crops that produces a higher yield of vegetable oil per hectare than soybean, and it is the largest source of edible vegetable oil in the world (Ndon, 2006). Most oil palm cultivars are a Dura × Pisifera hybrid with high oil yield hybrid vigor performance called the Tenera hybrid. Current development using conventional breeding and selection normally takes 10–15 yr to complete the cycle (Taeprayoon et al., 2015). A single growing apex and vegetative parts of oil palm cannot be used for multiple asexual reproduction and hence F_1 Tenera hybrid seeds are generally employed as a seed stock for the propagation of high yield traits (Barcelos et al., 2002). To increase the efficacy and rapid multiplication of the true type of elite oil palm, several studies have focused on the propagation process, nutrition, plant growth regulators and culture conditions to induce oil palm plantlet formation (Teixeira et al., 1995; Kanchanapoom et al., 2010; Romyanon et al., 2015; Corrêa et al., 2016). Thuzar et al. (2011) reported success in oil palm plantlet propagation through an indirect somatic embryogenesis process. Culture medium fortified with nutrients and plant growth regulator was also utilized in the propagation. This process takes at least 9–12 month to obtain a complete, new plant (Thuzar et al., 2011, 2012). Somatic embryogenesis has become the key method for the multiplication of oil palm elite genotypes. However, the functional mechanism to control somatic embryogenesis is poorly understood.

Plant somatic embryogenesis (SE) is a biological process that occurs naturally at the edges of leaves of several *Kalanchoë* species. Plant SE also appear as small bipolar structures that undergo dedifferentiation or redifferentiation to enter a new biological

* Corresponding author. Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900 Thailand.
E-mail address: fscicwj@ku.ac.th (C. Jantasuriyarat).

https://doi.org/10.1016/j.anres.2018.11.011

program that gives rise to somatic embryos (Loyola-Vargas and Ochoa-Alejo, 2016). The indirect somatic embryogenesis process consists of three differential steps. Step one is the transition of explant to embryogenic callus. Step two is the development of somatic embryo maturation including the globular, torpedo and cotyledonary stages. The last step is the maturation of the somatic embryo into a plantlet (Von Arnold et al., 2002). Several studies of the genes involved in somatic embryogenesis were reported in *LEAFY COTYLEDON (LEC)* (Braybrook and Harada, 2008; Ledwon and Gaj, 2011), *BABY BOOM (BBM)* (Boutilier et al., 2002), *WUSHEL (WUS)* genes (Arroyo-Herrera et al., 2008), *SOMATIC EMBRYOGENESIS RECEPTOR KINASE1* (SERK1) (Hecht et al., 2001) and the arabinogalactan protein (AGP) gene (Karami et al., 2009), which proposed the introduction of embryo formation in somatic cells.

The differentially expressed proteins during somatic embryogenesis are described as the proteins that are induced by oxidative stress and respond to higher cell division activity including ascorbate peroxidase, dehydroascorbate reductase, glutathione transferase and mitochondrial manganese superoxide dismutase (Takáč et al., 2011). In addition, Silva et al. (2014) revealed the identification of expressed proteins during acquisition of oil palm embryogenenic callus formation. Some proteins can be developed into biomarkers such as typeIIIa membrane protein cp-wap13, fructokinase and PR proteins. Furthermore, the classification of perspective proteomics represents a majority of protein function that is involved in metabolism and serves as an alternative energy source, and the synthesis of primary metabolites for somatic embryo development in secondary somatic embryogenesis in cassava (*Manihot esculenta*) (Baba et al., 2008) as well as in *Quercus suber* (Gomez-Garay et al., 2013). However, the understanding of protein post-translational modification is still unclear Protein glycosylation is one of the most common processes that occur in the post-translational modification and approximately 50% of cellular protein is glycosylated (VandenSteen et al., 1998). There are two types of protein glycosylation pathway: N and O-glycosylation. The N-glycosylation in plants is found in the consensus sequence of Ans-x-Ser/Thr (X≠ proline). This type of glycosylation is generally the attachment of various preassembled monosaccharide precursors in the lumen of the endoplasmic reticulum (ER) to asparagine residues. On the other hand, O-glycosylation varies among plant species, and the monosaccharide precursors are normally attached to serine residues from the extensins (EXTs) as well as other members of the hydroxyproline-rich glycoprotein (Hyp) family with a single derivative galactose (Saito et al., 2014; Strasser, 2016; Schoberer and Strasser, 2017). N-glycan and O-glycan of protein primarily occur in the secretory system (ER and Golgi) and can result in functional protein changes which are influenced by subcellular localization (Oxley et al., 2004; Zhou et al., 2005). Moreover, N-/O- glycoproteins are believed to play a vital role in protein stability, protein activity and protein function to control many aspects such as signaling for regulation of plant growth, cell polarity, morphogenesis and adaptation to biotic and abiotic stress (Baluška et al., 2003; Šamaj et al., 2003; Cannesan et al., 2012; Nguema-Ona et al., 2014). In previous study, glycoprotein was responsible for the inhibition and regulation of cell polarity in *Ustilago scitaminea* teliospores inoculated in sugarcane (Millanes et al., 2005). Xu et al. (2011) characterized hydroxyproline-rich glycoprotein during somatic embryogenesis of banana (*Musa spp. AAA*) using immunofluorescence labelling. The result suggested that this glycoprotein plays an important role especially in the process of embryo germination during plant regeneration via somatic embryogenesis of banana somatic embryos. However, profound insight into glycoproteomics during somatic embryogenesis is still lacking. The current study aimed to identify N-glycoproteomics, which may have important roles in response to somatic embryogenesis in oil palm tissue culture.

Materials and methods

Plant materials

Oil palm cv.Tenera zygotic embryos were used aged 15 wk after pollination from the Rice Science Unit, Kasetsart University, Kamphaengsaen campus, Thailand as the explants for tissue culture. The zygotic embryo explants were inoculated in N6 medium containing 1 mg/L of 2,4-dichlorophenoxyacetic acid (2,4-D) for callus induction under dark conditions until globular embryos appeared on primary callus. The embryo was then transferred to 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 and 2 g/L of activated charcoal under daylight conditions for a 16 h photoperiod until somatic the embryogenic maturation stages (globular, torpedo, cotyledon) were reached. Each small shoot was then transferred to a Modified N6 medium containing 0.5 g/L activated charcoal and 30 g/L sucrose to regenerate a complete new plant (Thuzar et al., 2011). Tissue samples at the callus stage and the embryo maturation stages (globular, torpedo, cotyledon) and the plantlet stage were collected for protein extraction (Fig. 1).

Protein extraction

Plant tissue samples were ground in liquid nitrogen and dissolved in 0.5% sodium dodecyl sulphate (SDS) with continuous vortexing for 30 min at room temperature. The pellet was precipitated using centrifugation at 12,000 revolutions per minute (rpm) for 15 min at 4 °C. The supernatant was transferred to a new microcentrifuge tube and mixed well with an equal volume of cooled acetone and incubated for 2 h at −20 °C. The mixture was then centrifuged at 12,000 rpm for 15 min at 4 °C to collect pelleted materials. The pellet was washed in cooled acetone and centrifuged three times at 12,000 rpm for 15 min at 4 °C. The pellet was dissolved again in 0.5% SDS and desalted in a Zeba™ Desalt spin column (Thermo Fisher Scientific; Madison, WI, USA) according to the manufacturer's instructions. The concentration of soluble protein was determined using the Lowry method (Lowry et al., 1951). The protein was stored at −80 °C until use in the glycoprotein isolation experiment.

Glycoprotein isolation and digestion

One microgram of protein from each sample was taken for glycoprotein isolation using a Pierce glycoprotein isolation kit, ConA and then desalted using the Zeba™ Desalt spin column according to manufacturer's protocol. To reduce disulfide bonding, 10 mM dithiothreitol in 10 mM ammonium bicarbonate was added to 5 μg of purified glycoproteins, and reformation of disulfide bonds in the glycoproteins was blocked by alkylation with 30 mM iodoacetamide in 10 mM ammonium bicarbonate. The glycoprotein samples were digested with 200 ng of sequencing grade porcine trypsin (Promega; Mannheim, Germany) for 16 h at 37 °C. The tryptic peptides were dried using a speed vacuum concentrator and resuspended in 0.1% formic acid for nano-liquid chromatography-tandem mass spectrometry (nanoLC-MS/MS) analysis.

Nano-liquid chromatography-mass spectrophotometry

Glycopeptide samples were injected in triplicate into an Impact II UHR-TOF MS system (Bruker Daltonics Ltd; Hamburg, Germany) which was coupled with a nanoLC system: UltiMate 3000 LC System (Thermo Fisher Scientific; Madison, WI, USA) as well as an electrospray at the flow rate of 300 nL/min to a nanocolumn (PepSwift monolithic column 100 μm internal diameter × 50 mm). A mobile phase of solvent A (0.1% formic acid) and solvent B

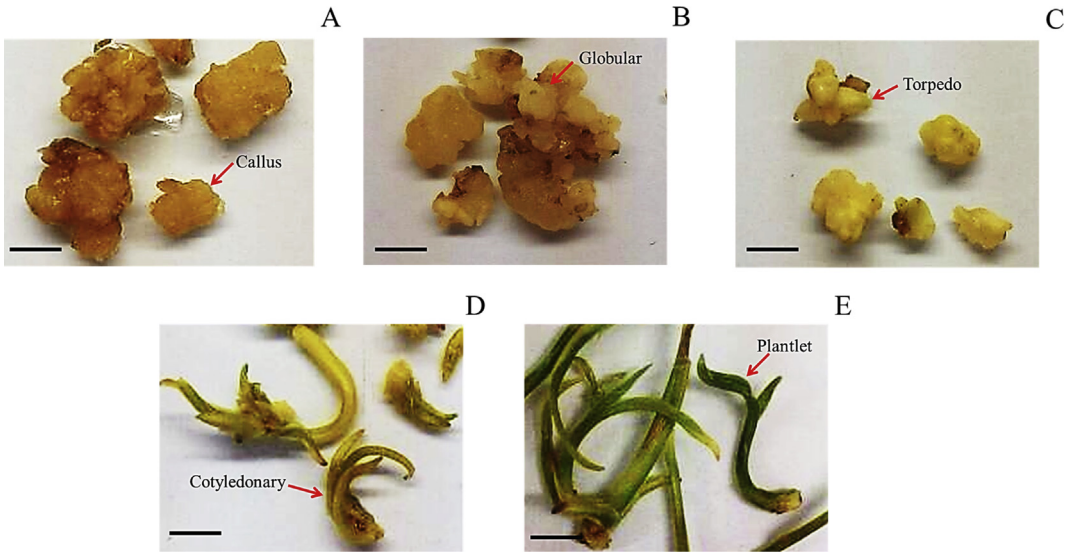

Fig. 1. Oil palm tissue samples: (A) callus stage; (B) globular stage; (C) torpedo stage; (D) cotyledonary stage; (E) oil palm plantlet, where scale bar = 0.5 cm.

(80% acetonitrile and 0.1% formic acid) were used to elute peptides using a linear gradient of 10–70% of solvent B at 0–13 min (the time-point of retention) followed by 90% B at 13–15 min to remove all peptides in the column. The final elution of 10% B at 15–20 min was carried out to remove any remaining salt. The resolution of the MS step was 0.6 and the accuracy was 0.15 U (m/z).

Bioinformatic and data analysis

The DeCyder MS 2.0 analysis software (GE Healthcare; Chicago, IL, USA) was used to measure the quantitative protein intensity based on peptide MS signal intensities of individually LC-MS analyzed data. To evaluate the peptide from each stage, a PepDetect module was used to produce ion peptides in the dataset as follows: 0.6 mass resolution, 0.3 typical peak width, 10,000 TOF resolution, 1 to 10 charge status and 0.1 U m/z shift tolerance. A PepMatch module was evaluated using signal intensity maps from each sample. The highest intensity sample was used as a control presenting the relative abundance of peptides as 2log intensities with mass tolerance set to 0.5 amu. An average abundance ratio more than 2-fold was determined as the overly expressed protein using a significance standard t-test and one-way analysis of variance (ANOVA; $p < 0.05$). All MS/MS spectra from Decyder MS analysis were analyzed by applying the global variable mode of carbamidomethyl (C), the variable mode of Hex(3)exNAc(1)Pent(1)(N), Hex(3)HexNAc(2)(N), Hex(3)HexNAc(2)-P(1)(N), Hex(3)HexNAc(4)(N), Hex(4)HexNAc(4)(N), Hex(5)HexNAc (2)(N), Hex(5)Hex-NAc(4)(N), Hex1HexNac1(S), Hex1HexNAc1(T), a peptide charge state (1+, 2+ and 3+) and an m/z tolerance 0.1 U. The data were searched against the NCBI FTP site of the *Elaeis guineensis* genome database (39,705 sequences; 17 February 2017) to identify matching peptides using the Mascot software (Matrix Science; London, UK). Identified proteins were filtered using one-way ANOVA ($p < 0.05$). In this experiment, 5 μg of BSA was used as an internal standard to normalize the protein intensities from each run. Then, a group of proteins from each sample was used to determine the similarity and differential protein expression using the Bioinformatics and Evolutionary Genomics tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) to collect candidate proteins to verify the expression level using real time quantitative polymerase chain reaction (PCR). Gene ontology was functionally characterized with identified glycoproteins using appropriate protein annotation shown in the uniprot protein database (http://www.uniprot.org/), PANTHER classification system (http://pantherdb.org/) and previously reported proteins.

Quantitative real-time reverse transcription-polymerase chain reaction

The total RNA was extracted from primary callus, somatic embryo maturation at the globular, torpedo and cotyledon stages and at the plantlet stage using Spin Plant RNA (STRATEC Molecular GmbH; Berlin, Germany) according to the manufacturer's protocol with minor modification using DNAaseI (Vivantis; Selangor Darul Ehsan, Malaysia.) to treat on an RNA binding column before eluting the total RNA. One microgram of total RNA was taken to synthesize first strand cDNA using a cDNA synthesis kit (biotechrabbit GmbH; Berlin, Germany) according to the manufacturer's instructions. The cDNA content of the *E. guineensis elongation factor alpha-1* gene (NCBI accession number XM_019850296) was used as a normalization of the expression level. The expression level of candidate proteins was examined with specifically designed primers: KAPA SYBR® FAST qPCR kit Mastermix (2x) universal (Kapa Biosystems; Wilmington, MA, USA) according the manufacturer's instructions on an Eppendorf Mastercycler ep Realplex (Thermo Fisher Scientific; Madison, WI, USA). The reaction mixture (10 μL) was filled with 2 μL of each 40-fold dilution cDNA sample, 100 ₙM of each primer and 1X KAPA SYBR® FAST qPCR kit Mastermix. The thermal cycling conditions were set as: 95 °C for 2 min; 40 cycles of 95 °C for 15 s, 58 °C for 15 s and 72 °C for 20 s following a melting curve at the final step according to instrument recommendation. The amplification for each gene was assessed for three replications and the transcription level for each gene was confirmed in three bioreplications. The double delta Ct calculation was used to determine fold change by comparison with the mRNA level of the callus stage as a calibrator.

Results

Glycoprotein isolation and nano-liquid chromatography-tandem mass spectrometry identification

To assess the glycoprotein change during the stages of oil palm tissue development, embryogenic callus, somatic embryo maturation and plantlet formation samples were collected for total protein extraction. Total proteins were isolated to obtain glycoprotein using

a lectin ConA affinity column for isolation of N-linked glycoprotein, and tryptic digestion into peptides coupled with nanoLC-MS/MS. The peptide sequences from nanoLC-MS/MS were searched against the NCBI FTP site of the *Elaeis guineensis* genome database and matched with 383 proteins. In total, 375, 375, 364, 374 and 368 proteins were presented in the callus, globular, torpedo, cotyledonary and plantlet stages, respectively, with 338 proteins shared in all stages. Nine proteins were commonly expressed in the callus, globular, cotyledonary and plantlet stages (except the torpedo stage), including PREDICTED: 26S proteasome non-ATPase regulatory subunit 14 homolog, PREDICTED: formamidopyrimidine-DNA glycosylase isoform X3, and PREDICTED: myb family transcription factor EFM. Six proteins were expressed in the callus, globular, torpedo and cotyledonary stages, excluding the plantlet stage including PREDICTED: proteasome assembly chaperone 4, PREDICTED: BTB/POZ domain and ankyrin repeat-containing protein NPR1 isoform X3 and PREDICTED: inactive poly [ADP-ribose] polymerase RCD1. Six proteins were expressed in the globular, torpedo, cotyledonary and plantlet stages but not in the callus stage including PREDICTED: rab9 effector protein with kelch motifs-like isoform X2, PREDICTED: cytochrome P450 704C1-like and PREDICTED: auxin-responsive protein SAUR50. Five proteins were shown in the callus, torpedo, cotyledonary and plantlet stages but not in the globular stage including PREDICTED: eukaryotic translation initiation factor 2 subunit alpha homolog, PREDICTED: 40S ribosomal protein S18-like and PREDICTED: SAP-like protein BP-73 isoform X2. Three proteins were particularly expressed in the callus, globular and torpedo stages but not in the cotyledonary and plantlet stages including PREDICTED: protein ENHANCED DISEASE RESISTANCE 2 isoform X4 and PREDICTED: FT-interacting protein 1-like. Three proteins were expressed in the callus cotyledon and plantlet stages but not in the globular and torpedo stages composed of PREDICTED: chalcone synthase 1-like, PREDICTED: zinc finger protein 1-like and PREDICTED: dentin sialophosphoprotein. Two proteins were particularly expressed in the callus, globular and cotyledonary stages but not in the torpedo and plantlet stages namely, PREDICTED: cleft lip and palate transmembrane protein 1 homolog and PREDICTED: CBL-interacting serine/threonine-protein kinase 11-like. PREDICTED: peroxisome biogenesis protein 22-like isoform X1 was expressed in the callus, globular, torpedo and plantlet stages but not in the cotyledonary stage. PREDICTED: uncharacterized protein LOC105038887 was expressed in the globular, cotyledonary and plantlet stages. PREDICTED: protein TIC 40, chloroplastic-like was particularly expressed in the globular torpedo and plantlet stages (Fig. 2). Three groups of expressed glycoproteins namely, protein expressed only in callus and globular (group A), proteins expressed only in globular, torpedo, cotyledon and plantlet stages (group B) and protein expressed in globular, torpedo and plantlet stages (group C) were collected for further quantitative real-time PCR analysis.

Functional classification of identified glycoproteins

To functionally classify the identified glycoproteins, a classification of protein was properly designed in terms of biological process, molecular function and cellular component based on proteins reported in the Uniprot tool and PANTHER-gene list analysis. Total identified glycoproteins were categorized into groups based on their percentage existence in biological process, from highest to lowest as: unknown (33.85), metabolic process (12.76), cellular process (10.42) and transcription (8.59). The molecular function of identified glycoproteins was described as percentage of unknown (37.24), binding (10.16), transcription factor (7.03) and transferase (5.99). The cellular component of identified

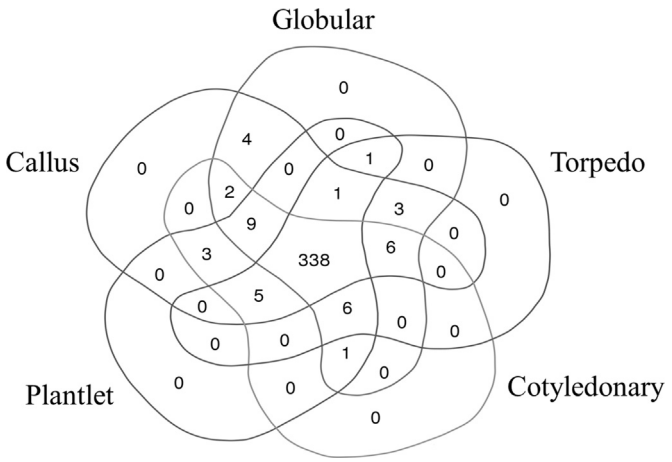

Fig. 2. Venn diagram of comparative differentially expressed glycoproteins showing number of glycoproteins identified in callus (blue), globular (pink), torpedo (brown), cotyledonary (yellow) and plantlet (green) stages.

glycoproteins had the highest percentage as unknown (40.36), nucleus (19.79), cytoplasm (12.76) and membrane (9.9) (Fig. 3).

Differentially expressed glycoproteins

For the expression analysis of glycoproteins involved in somatic embryogenesis, differentially expressed protein acquisition was identified during the somatic embryogenesis at the somatic embryo maturation and plantlet stages. The three groups of glycoproteins of interest were: group A, proteins expressed in the globular, torpedo and plantlet stages; group B, proteins expressed during embryogenic callus transition to the somatic embryo; and group C, proteins expressed during somatic embryo acquisition to the plantlet stage. PREDICTED: protein TIC 40, chloroplastic-like was suppressed in the globular stage but up regulated in the torpedo and plantlet stages. Four differentially expressed proteins involved in embryogenic callus transition to somatic embryos were observed, namely, PREDICTED: Werner Syndrome-like exonuclease isoform X2, PREDICTED: calcium-dependent protein kinase 8, PREDICTED: mitochondrial outer membrane protein porin 5 and PREDICTED: SRSF protein kinase 2. These four proteins were up regulated in the callus stage but suppressed in the globular stage. Six differentially expressed proteins involved in somatic embryo development to plantlets were gained: PREDICTED: rab9 effector protein with kelch motifs-like isoform X2, and PREDICTED: cytochrome P450 704C1-like that were up regulated in globular but slightly down regulated in the torpedo, cotyledonary and plantlet stages, PREDICTED: pto-interacting protein 1 isoform X1 and PREDICTED: LOW QUALITY PROTEIN: RNA-binding protein cabeza-like which were up regulated in both globular and plantlet stages but down regulated in torpedo and cotyledonary stages, PREDICTED: uncharacterized membrane protein At1g16860 and PREDICTED: auxin-responsive protein SAUR50 that were up regulated in both the cotyledonary and plantlet stages but slightly decreased in the globular and torpedo stages (Table 1).

Quantitative real-time reverse transcription-polymerase chain reaction analysis of differentially glycoproteins

To verify the differentially expressed glycoprotein during somatic embryogenesis development of oil palm, quantitative real-time reverse transcription-PCR (RT-qPCR) technique was used. The transcription level of 11 genes encoding differentially expressed

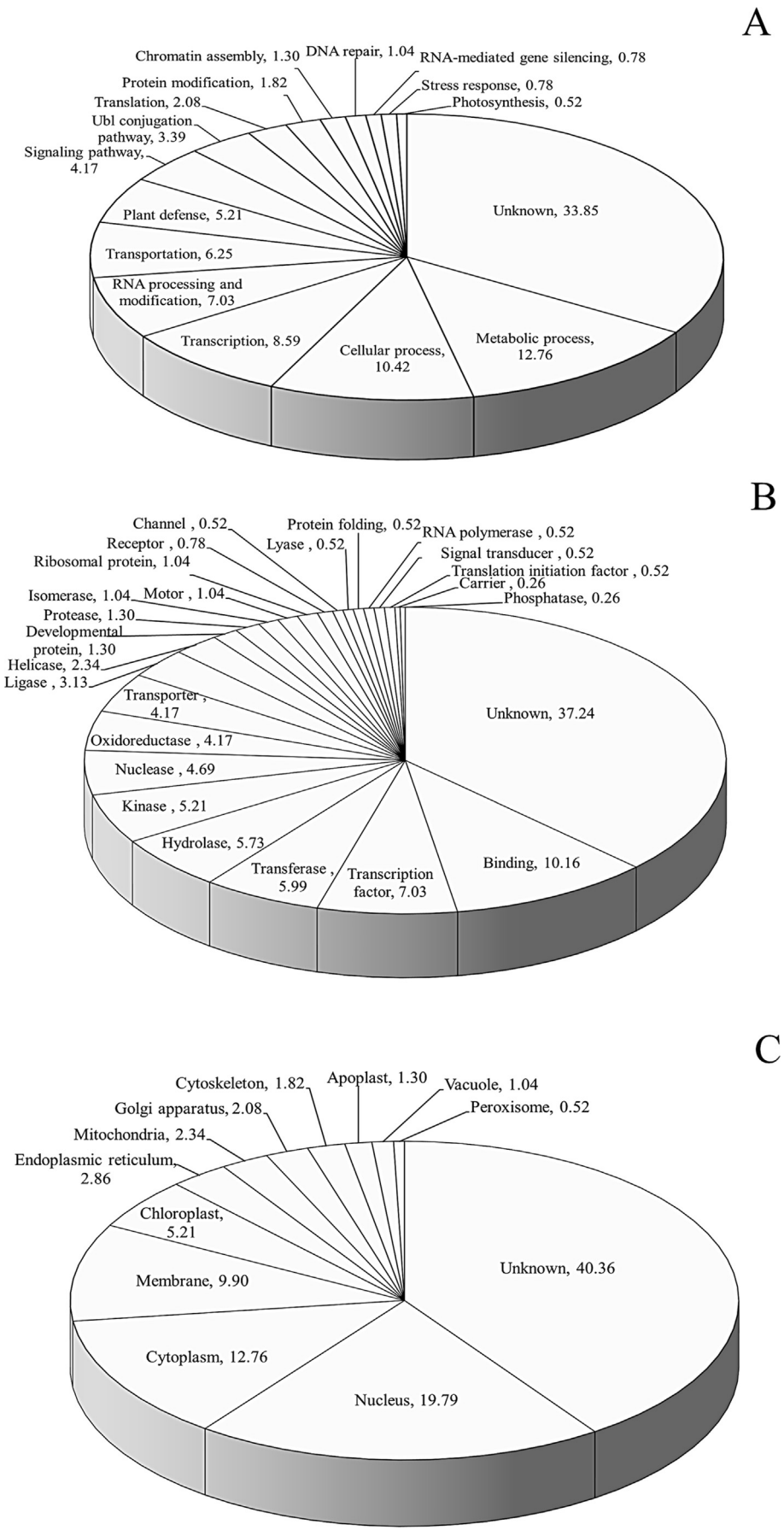

Fig. 3. Functional classification of oil palm glycoproteins based on: (A) biological process; (B) molecular function; (C) cellular component.

Table 1
Differentially expressed glycoproteins during oil palm tissue culture.

Code	Accession No	Protein name	Protein sequence	Identity score	Protein abundance (log_2)				
					Callus	Globular	Torpedo	Cotyledon	Plantlet
A1	gi|743764858	PREDICTED: protein TIC 40,chloroplastic-like	GGDGASQR	13.74	0.00	8.02	11.61	0.00	16.88
B1	gi|1130657759	PREDICTED: Werner Syndrome-like exonuclease isoform X2	IGGAPKK	7.33	15.21	13.32	0.00	0.00	0.00
B2	gi|743763830	PREDICTED: calcium-dependent protein kinase 8	EIEIMRGLPEH PNIVRLR	8.06	13.79	8.18	0.00	0.00	0.00
B3	gi|743846899	PREDICTED: mitochondrial outer membrane protein porin 5	FSTNENTFTVGG QYALDPLTTVK	7.20	13.73	7.99	0.00	0.00	0.00
B4	gi|743851757	PREDICTED: SRSF protein kinase 2	RGSMEGPLEWER	13.86	12.09	9.51	0.00	0.00	0.00
C1	gi|743844054	PREDICTED: rab9 effector protein with kelch motifs-like isoform X2	MFGNHWNQEK	17.86	0.00	16.85	16.04	12.84	12.23
C2	gi|1130639345	PREDICTED: cytochrome P450 704C1-like	IMKLLNIGSEA MLK	7.36	0.00	15.91	14.92	14.75	14.56
C3	gi|1130670691	PREDICTED: Uncharacterized membrane protein At1g16860	RSGELGKMFDIPV DNSK	13.97	0.00	14.93	13.53	15.59	15.75
C4	gi|743877450	PREDICTED: pto-interacting protein 1 isoform X1	EITKSFGDEALV GEGSFGR	13.68	0.00	10.00	9.83	6.72	9.52
C5	gi|743830897	PREDICTED: auxin-responsive protein SAUR50	MGSGAGK	14.39	0.00	9.12	14.08	19.16	17.58
C6	gi|1130671318	PREDICTED: LOW QUALITY PROTEIN: RNA-binding protein cabeza-like	GGGESLVGAR PPGGNR	19.33	0.00	9.10	8.85	6.61	9.99

glycoproteins was identified. All genes were calculated as fold change of the mRNA expression level by using the transcription level of callus as the reference. The results showed that genes were expressed in all stages (callus, globular, torpedo, cotyledonary and plantlet) but high levels of the mRNA expression rate were noticed in the torpedo, cotyledonary and plantlet stages (Fig. 4). The mRNA expression level was consistent with the level of glycoprotein identification.

Discussion

Glycoproteomics during oil palm tissues development

Somatic embryogenesis is the outstanding characteristic of plant stem cells, based on the totipotential cell ability that allows regeneration without gamete fusion. In this study, the samples were collected during the embryogenic callus, somatic embryo maturation and plantlet stages (Fig. 1) to identify and analyze differentially expressed glycoproteins. The isolated glycoproteins were digested using gel free tryptic digestion and the results revealed 383 identified N-linked glycoproteins. The results also suggested that a higher amount of identified protein may be gained when using gel free tryptic digestion than by using gel digestion due to the fact that there was approximately 10–15% peptides loss and the amount of glycoprotein depends on a variety of factors during peptide extraction, including the composition of the gel plugs and reaction conditions (Speicher et al., 2002; Von Hagen, 2008). The identified glycoproteins in this study were similar to the previous report by Ruiz-May et al. (2014) including PREDICTED: leucine-rich repeat receptor-like serine/threonine/tyrosine-protein kinase SOBIR1, PREDICTED: serine/threonine-protein kinase D6PKL2 isoform X1, PREDICTED: probable LRR receptor-like serine/threonine-protein kinase At3g47570 and PREDICTED: multicopper oxidase LPR1-like. Furthermore, the result showed that modified-glycoproteins were involved in somatic embryogenesis in oil palm including PREDICTED: pathogenesis-related protein 1 and PREDICTED: 40S ribosomal protein S18-like, which was also found by Silva et al. (2014) and in somatic embryos of *Theobroma cacao* for example, PREDICTED: 60S ribosomal protein L37-3, PREDICTED: peptidyl-prolyl cis-trans isomerase CYP40-like and PREDICTED: 26S proteasome non-ATPase regulatory subunit 14 homolog reported by Niemenak et al. (2015) and PREDICTED: probable glutathione peroxidase 5, previously reported by Takáč et al. (2011). These modified glycoproteins associated with somatic embryogenesis during plant development were also found in this study. From these findings, the post-translational modification of N-glycosylation may be effective in many aspects of protein targeting, enzymatic properties, protein folding that serves as signaling for regulation of growth, plant polarity, morphogenesis and adaptation to biotic and abiotic stress during the somatic embryogenesis and plantlet stages of oil palm (Baluška et al., 2003; Šamaj et al., 2003; Cannesan et al., 2012; Nguema-Ona et al., 2014).

Functional categorization of glycoproteins during oil palm tissues development

To characterize protein function, identified glycoproteins were classified based on previously reported proteins in the Uniprot tool and PANTHER-gene list analysis. Glycoproteins were categorized into groups using their percentage existence in the biological process, molecular function and cellular component as the criteria. In the biological process, the majority of identified glycoproteins were characterized based on metabolic process, cellular process and transcription, respectively. For the molecular function, the classification used the percentage existence of proteins from the highest to lowest level in binding, transcription factor and transferase, respectively. For the cellular component, identified glycoproteins showed the percentage existence from highest to lowest amount in the nucleus, cytoplasm and membrane, respectively. These results showed a similar pattern of gene ontology to the putative secreted N-glycoproteome of *Arabidopsis* (Song et al., 2011), Tomato fruit N-glycoproteome analysis (Ruiz-May et al., 2014) and cassava somatic embryos proteome analysis (Li et al., 2010) where all these showed a highly classified function in the metabolic process, nucleotide binding and membrane. The finding also revealed the highest biological process, molecular function and cellular component being in the metabolic process, binding and nuclear components. Moreover, this result indicated that identified glycoproteins may have an important role in cellular metabolism regulation during oil palm morphological alteration of tissues.

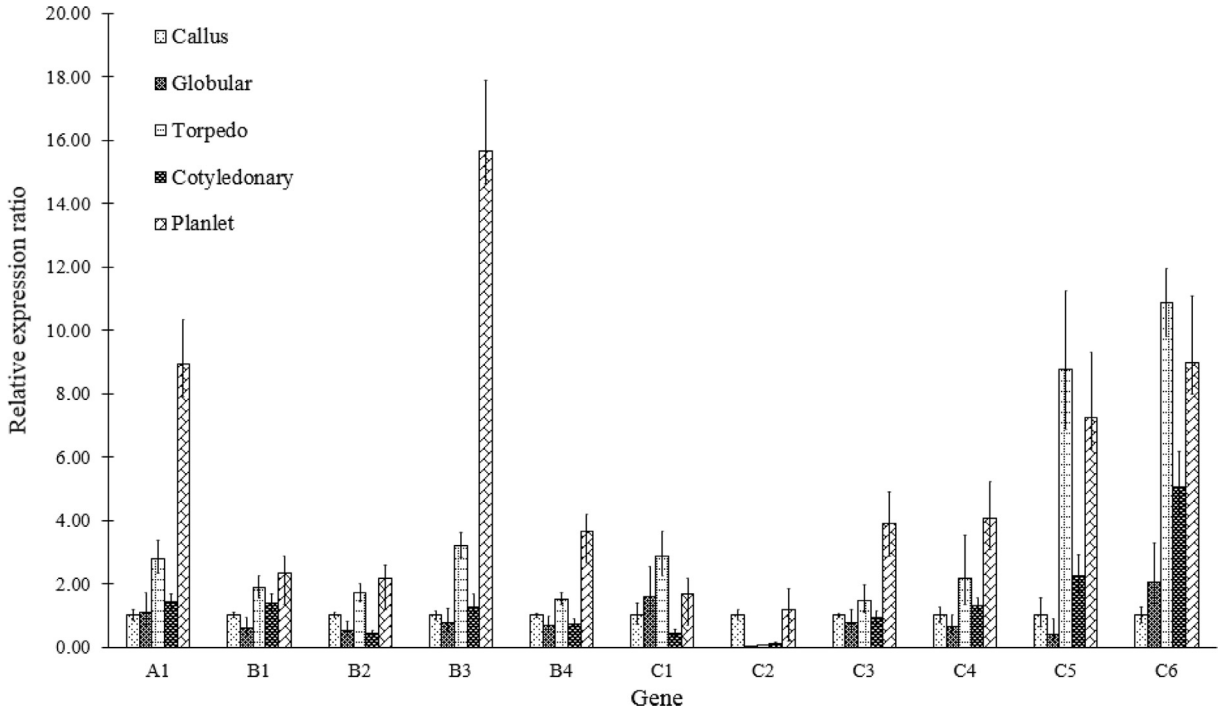

Fig. 4. Comparative mRNA expression levels of differential glycoproteins using quantitative real-time reverse transcription-polymerase chain reaction, where error bars represent ± SE.

Comparison of differentially expressed glycoproteins

To demonstrate the induction of somatic embryogenesis process, PREDICTED: protein TIC 40, chloroplastic-like (A1) was particularly expressed in the globular, torpedo and plantlet stages (Table 1). TIC40 protein is a chloroplast inner-membrane protein that functions as a co chaperone in the stromal chaperone complex that plays essential roles in lipid synthesis, metabolite transport and cellular signaling in plants (Chou et al., 2003; Chiu and Li, 2008; Singh et al., 2008). Since this protein is not expressed in callus but expressed during embryo development, it can be used as a biomarker for biological indication of the somatic embryo and plantlet stages.

Four glycoproteins were upregulated in the callus but suppressed in the globular stage (Table 1). Werner Syndrome-like exonuclease (B1) possesses DNA helicase and exonuclease activities which are involved in DNA metabolism and essential for genome stability (Plchova et al., 2003). Calcium-dependent protein kinase 8 (B2) are unimolecular calcium sensor/protein kinase effector proteins that directly bind calcium ions before substrate phosphorylation and are involved in metabolism, osmosis, the hormone response and stress signaling pathways (Schulz et al., 2013; Valmonte et al., 2014). Mitochondrial outer membrane protein porin 5 (B3) is responsible for the passage of small molecules between the mitochondrion and the cytosol and is involved in the regulation of organellar and cellular metabolism as well as being part of the mitochondrial structure (Bay et al., 2012; Salinas et al., 2014). SRSF protein kinase 2 (B4) is one of the receptor-like kinases family members that are involved in plant response to abiotic stresses, plant development, growth, hormone perception and response to pathogens (Goff and Ramonell, 2007; Kulik et al., 2011).

The abundant glycoproteins that were expressed only in somatic embryo transition to plantlets were determined. The six identified proteins were composed of five characterized proteins and one unknown protein (Table 1). The Rab9 effector protein with kelch motifs-like (C1) is one of the Rab protein family member that is a key mediator of vesicle transport and specificity and has an important role in the evolution of intracellular trafficking machinery in transport from late endosomes to the trans-golgi network (Mackiewicz and Wyroba, 2009; Brighouse et al., 2010). Cytochrome P450 704C1-like is a heme-containing protein that is involved in multiple metabolic pathways with distinct and complex functions with leading roles in a vast array of reactions (Pinot and Beisson, 2011; Xu et al., 2015). Pto-interacting protein 1 encodes a Ser/Thr protein kinase in rice (*Oryza sativa*), which is a negative regulator of innate immunity, including ion fluxes across plasma membranes, oxidative bursts, activation of mitogen-activated protein kinase cascades and transcriptional reprogramming of defense genes (Matsui et al., 2014). The auxin-responsive protein SAUR50 is one of the SMALL AUXIN UP RNAs (SAURs) auxins of the gene family that is involved in hormonal and environmental signals for plant growth regulation and development (Ren and Gray, 2015). The RNA-binding protein cabeza-like is one of the RNA-binding protein (RBP) group of the central regulatory factors, which controls post-transcriptional RNA metabolism during plant growth, development and stress responses (Lee and Kang, 2016).

Previous study of the differential N- glycoproteins involved in common bean drought stress was useful for further understanding drought response in this important legume (Zadražnik et al., 2017). Glycoproteins may be used to address importantly different biological questions and provide great assistance in enhancing insight into these biological processes (Song et al., 2011). Moreover, it was found that the differentially expressed glycoproteins during indirect somatic embryogenesis of oil palm could be used to develop a protein biological indicator. In addition, they can be utilized to determine the specific somatic embryo stage and are also useful for further understanding the important role of glycoprotein that is specific in taking less time in generating the oil palm plantlet via the somatic embryogenesis process.

Verification of differentially expressed glycoproteins

To confirm the mRNA expression level of differentially expressed glycoproteins, 11 proteins were verified using RT-qPCR. The transcription level of the 11 genes was expressed in all stages (Fig. 4). These results corresponded to previous study of the genes associated with cDNA AFLP of oil palm somatic embryogenesis that showed 82.33% polymorphism from each stage (callus, globular, torpedo, cotyledonary and plantlet) (Pattarapimol et al., 2015). However, the correlation between protein and mRNA expression levels partially overlaps but they are not identical (Payne, 2014). The transcription and translational control are composed of a complex biological mechanism that defines the abundance of mRNA and protein (Maier et al., 2009). This result indicated that the gene expression level at the callus stage may be induced for dedifferentiation and redifferentiation. The expression level also responds to the differential transcription level at the globular, torpedo, cotyledonary and plantlet stages and may be required to prepare and encourage itself for cellular process, metabolic process and tissue development during the oil palm somatic embryo and plantlet formation stages.

In conclusion, N-linked glycosylation-post translational modification of proteins is involved in metabolic, transcription and cellular processes during indirect oil palm somatic embryogenesis. Furthermore, differentially identified glycoproteins are essential for genome stability, intracellular trafficking machinery and protein signaling in stress, as affected by hormonal and environmental factors with regard to plant growth and development. In addition, the differentially identified glycoproteins can be used as a glycoprotein biomarker to determine the specificity of the somatic embryo stage in oil palm. In this regard, RT-qPCR was carried out and specific primers were designed to verify the differential glycoproteins involved in oil palm somatic embryogenesis. This comprehensive study is beneficial for understanding the important role of glycoprotein which is specific in reducing the time for generating oil palm plantlets via the somatic embryogenesis process. This study can inform and deepen understanding of how glycoproteins work. However, further research and study are needed to characterize the molecular function of glycoprotein *in vivo* during oil palm tissue culture.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Acknowledgements

This research was supported by the Kasetsart University Research and Development Institute and a graduate scholarship was provided by the National Research Council of Thailand in the 2017 fiscal year. The Proteomic Research Unit, National Center for Genetic Engineering and Biotechnology (BIOTEC), Pathum Thani, Thailand provided shotgun proteomics analysis.

References

Arroyo-Herrera, A., Ku Gonzalez, A., Canche Moo, R., et al., 2008. Expression of WUSCHEL in *Coffea canephora* causes ectopic morphogenesis and increases somatic embryogenesis. Plant Cell Tiss 94, 171–180.

Baba, A.I., Nogueira, F.C.S., Pinheiro, C.B., et al., 2008. Proteome analysis of secondary somatic embryogenesis in cassava (*Manihot esculenta*). Plant Sci. 175, 717–723.

Baluśka, F., Šamaj, J., Wojtaszek, P., Volkmann, D., Menzel, D., 2003. Cytoskeleton plasma membrane-cell wall continuum in plants. Plant Physiol. 133, 482–491.

Barcelos, E., Amblard, P., Berthaud, J., Seguin, M., 2002. Genetic diversity and relationship in American and African oil palm as revealed by RFLP and AFLP molecular markers. Pesqui. Agropecu. Bras. 37, 1105–1114.

Bay, D.C., Hafez, M., Young, M.J., Court, D.A., 2012. Phylogenetic and coevolutionary analysis of the β-barrel protein family comprised of mitochondrial porin (VDAC) and Tom40. Biochim. Biophys. Acta 1818, 1502–1519.

Boutilier, K., Offringa, R., Sharma, V.K., et al., 2002. Ectopic expression of *BABY BOOM* triggers a conversion from vegetative to embryonic growth. Plant Cell 14, 1737–1749.

Braybrook, S.A., Harada, J.J., 2008. LECs go crazy in embryo development. Trends Plant Sci. 13, 624–630.

Brighouse, A., Dacks, J.B., Field, M.C., 2010. Rab protein evolution and the history of the eukaryotic endomembrane system. Cell. Mol. Life Sci. 67, 3449–3465.

Cannesan, M.A., Durand, C., Burel, C., et al., 2012. Effect of arabinogalactan proteins from the root caps of pea and *Brassica napus* on *Aphanomyces euteiches* zoospore chemotaxis and germination. Plant Physiol. 159, 1658–1670.

Chiu, C.C., Li, H.M., 2008. Tic40 is important for reinsertion of proteins from the chloroplast stroma into the inner membrane. Plant J. 56, 793–801.

Chou, M.L., Fitzpatrick, L.M., Tu, S.L., et al., 2003. Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. EMBO J. 16, 2970–2980.

Corley, R.H.V., Tinker, P.B., 2003. The Oil Palm, fourth ed. Blackwell Publishing, Oxford, UK.

Corrêa, T.R., Motoike, S.Y., Andrade, A.P.S., et al., 2016. Accelerated *in vitro* propagation of elite oil palm genotypes (*Elaeis guineensis* Jacq.) by substituting cytokinin with putrescine. Afr. J. Biotechnol. 15, 2767–2775.

Goff, K.E., Ramonell, K.M., 2007. The role and regulation of receptor-like kinases in plant defense. Gene Regul. Syst. Biol. 1, 167–175.

Gomez-Garay, A., Lopez, J.A., Camafeita, E., Bueno, M.A., Pintos, B., 2013. Proteomic perspective of *Quercus suber* somatic embryogenesis. J. Proteomics. 9, 314–325.

Hecht, V., Vielle-Calzada, J.P., Hartog, M.V., Schmidt, E.D., Boutilier, K., Grossniklaus, U., de Vries, S.C., 2001. The Arabidopsis *SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1* gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture. Plant Physiol. 127, 803–816.

Kanchanapoom, K., Phongdara, A., Kanchanapoom, K., 2010. The effect of chitosan on the organogenesis of oil palm embryo-derived callus. Not. Bot. Hort. Agrobot. 38, 213–217.

Karami, O., Aghavaisi, B., Aghil Mahmoudi, P., 2009. Molecular aspects of somatic-to-embryogenic transition in plants. J. Chem. Biol. 2, 177–190.

Kulik, A., Wawer, I., Krzywińska, E., Bucholc, M., Dobrowolska, G., 2011. SnRK2 protein kinases–key regulators of plant response to abiotic stresses. OMICS 15, 859–872.

Ledwon, A., Gaj, M.D., 2011. *LEAFY COTYLEDON1, FUSCA3* expression and auxin treatment in relation to somatic embryogenesis induction in Arabidopsis. Plant Growth Regul. 5, 157–167.

Lee, K., Kang, H., 2016. Emerging roles of RNA-binding proteins in plant growth, development, and stress responses. Mol. Cell. 39, 179–185.

Li, K., Zhu, W., Zeng, K., et al., 2010. Proteome characterization of cassava (*Manihot esculenta* Crantz) somatic embryos, plantlets and tuberous roots. Proteome Sci. 27, 10. doi.org/10.1186/1477-5956-8-10.

Lowry, O.H., Rosebrough, N.J., Farr, A.L., Randall, R.J., 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. 193, 265–275.

Loyola-Vargas, V.M., Ochoa-Alejo, N., 2016. Somatic Embryogenesis: Fundamental Aspects and Applications, first ed. Springer International Publishing, Basel, Switzerland.

Mackiewicz, P., Wyroba, E., 2009. Phylogeny and evolution of Rab7 and Rab9 proteins. BMC Evol. Biol. 14, 101. https://doi.org/10.1186/1471-2148-9-101.

Maier, T., Güell, M., Serrano, L., 2009. Correlation of mRNA and protein in complex biological samples. FEBS Lett. 583, 3966–3973.

Matsui, H., Fujiwara, M., Hamada, S., et al., 2014. Plasma membrane localization is essential for *Oryza sativa* Pto-interacting protein 1a-mediated negative regulation of immune signaling in rice. Plant Physiol. 166, 327–336.

Millanes, A.M., Fontaniella, B., Legaz, M.E., Vicente, C., 2005. Glycoproteins from sugarcane plants regulate cell polarity of *Ustilago scitaminea* teliospores. J. Plant Physiol. 162, 253–265.

Ndon, B.A., 2006. The Oil Palm. Concept Publication Ltd, Lagos, Nigeria.

Nguema-Ona, E., Vicré-Gibouin, M., Gotté, M., Plancot, B., Lerouge, P., Bardor, M., Driouich, A., 2014. Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function. Front. Plant Sci. 5, 499.

Niemenak, N., Kaiser, E., Maximova, S.N., Laremore, T., Guiltinan, M.J., 2015. Proteome analysis during pod, zygotic and somatic embryo maturation of *Theobroma cacao*. J. Plant Physiol. 180, 49–60.

Oxley, D., Currie, G.C., Bacic, A., 2004. Analysis of carbohydrate from glycoproteins in purifying proteins for proteomics: a laboratory manual. In: Simpson, R.J. (Ed.). Cold Spring Harbor Laboratory Press, New York, NY, USA, pp. 579–636.

Pattarapimol, T., Thuzar, M., Vanavichit, A., Tragoonrung, S., Roytrakul, S., 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.

Payne, S.H., 2014. The utility of protein and mRNA correlation. Trends Biochem. Sci. 40, 1–3.

Pinot, F., Beisson, F., 2011. Cytochrome P450 metabolizing fatty acids in plants: characterization and physiological roles. FEBS J. 278, 195–205.

Plchova, H., Hartung, F., Puchta, H., 2003. Biochemical characterization of an exonuclease from *Arabidopsis thaliana* reveals similarities to the DNA exonuclease of the human werner syndrome protein. J. Biol. Chem. 278, 44128–44138.

Ren, H., Gray, W.M., 2015. SAUR proteins as effectors of hormonal and environmental signals in plant growth. Mol. Plant 8, 1153–1164.

Romyanon, K., Mosaleeyanon, K., Kirdmanee, C., 2015. Direct-shoot organogenesis as an alternative protocol for in vitro regeneration of oil palm (*Elaeis guineensis* Jacq.). Sci. Hortic. 195, 1–7.

Ruiz-May, E., Hucko, S., Howe, K.J., Zhang, S., Sherwood, R.W., Thannhauser, T.W., Rose, J.K., 2014. A comparative study of lectin affinity based plant N-glycoproteome profiling using tomato fruit as a model. Mol. Cell. Proteomics 13, 566–579.

Saito, F., Suyama, A., Oka, T., Yoko-O, T., Matsuoka, K., Jigami, Y., Shimma, Y.I., 2014. Identification of novel peptidyl serine α-galactosyltransferase gene family in plants. J. Biol. Chem. 289, 20405–20420.

Salinas, T., El Farouk-Ameqrane, S., Ubrig, E., Sauter, C., Duchêne, A.M., Maréchal-Drouard, L., 2014. Molecular basis for the differential interaction of plant mitochondrial VDAC proteins with tRNAs. Nucleic Acids Res. 42, 9937–9948.

Šamaj, J., Baluška, F., Pretová, A., Volkmann, D., 2003. Auxin deprivation induces a developmental switch in maize somatic embryogenesis involving redistribution of microtubules and actin filaments from endoplasmic to cortical cytoskeletal arrays. Plant Cell Rep. 21, 940–945.

Schoberer, J., Strasser, R., 2017. Plant glyco-biotechnology. Semin. Cell Dev. Biol. 16, 30360–30373.

Schulz, P., Herde, M., Romeis, T., 2013. Calcium-dependent protein kinases: hubs in plant stress signaling and development. Plant Physiol. 163, 523–530.

Silva de Carvalho, C., Carmo, L.S., Luis, Z.G., Silva, L.P., Scherwinski-Pereira, J.E., Mehta, A., 2014. Proteomic identification of differentially expressed proteins during the acquisition of somatic embryogenesis in oil palm (*Elaeis guineensis* Jacq.). J. Proteomics. 104, 112–127.

Singh, N.D., Li, M., Lee, S.B., Schnell, D., Daniell, H., 2008. Arabidopsis Tic40 expression in tobacco chloroplasts results in massive proliferation of the inner envelope membrane and upregulation of associated proteins. Plant Cell 20, 3405–3417.

Song, W., Henquet, M.G., Mentink, R.A., et al., 2011. N-glycoproteomics in plants: perspectives and challenges. J. Proteomics 12, 1463–1474.

Speicher, K.D., Kolbas, O., Harper, S., Speicher, D.W., 2002. Systematic analysis of peptide recoveries from in-gel digestions for protein identifications in proteome studies. J. Biomol. Tech. 11, 74–86.

Strasser, R., 2016. Plant protein glycosylation. Glycobiology 26, 926–939.

Taeprayoon, P., Tanya, P., Lee, S.H., Srinives, P., 2015. Genetic background of three commercial oil palm breeding populations in Thailand revealed by SSR markers. AJCS 9, 281–288.

Takáč, T., Pechan, T., Samaj, J., 2011. Differential proteomics of plant development. J. Proteomics. 74, 577–588.

Teixeira, J.B., Stindahl, M.R., Nakamura, T., Kirby, E.G., 1995. Establishment of oil palm cell suspensions and plant regeneration. Plant Cell Tissue Organ Cult. 40, 105–111.

Thuzar, M., Vanavichit, A., Tragoonrung, S., Jantasuriyarat, C., 2011. Efficient and rapid plant regeneration of oil palm zygotic embryos cv. 'Tenera' through somatic embryogenesis. Acta Physiol. Plant. 33, 123–128.

Thuzar, M., Vanavichit, A., Tragoonrung, S., Jantasuriyarat, C., 2012. Recloning of regenerated plantlets from elite oil palm (*Elaeis guineensis* Jacq.) cv. Tenera. Afr. J. Biotechnol. 11, 14761–14770.

Valmonte, G.R., Arthur, K., Higgins, C.M., MacDiarmid, R.M., 2014. Calcium-dependent protein kinases in plants: evolution, expression and function. Plant Cell Physiol. 55, 551–569.

VandenSteen, P., Rudd, P.M., Dwek, R.A., Opdenakker, G., 1998. Concepts and principles of O-linked glycosylation. Crit. Rev. Biochem. Mol. Biol. 33, 151–208.

Von Arnold, S., Sabala, I., Bozhkov, P., Dyachok, J., Filonova, L., 2002. Developmental pathways of somatic embryogenesis. Plant Cell Tissue Organ Cult. 69, 233–249.

Von Hagen, J., 2008. Proteomics Sample Preparation, first ed. Blackwell Publishing, Oxford, UK.

Xu, C., Takáč, T., Burbach, C., Menzel, D., Šamaj, J., 2011. Developmental localization and the role of hydroxyproline rich glycoproteins during somatic embryogenesis of banana (*Musa* spp. AAA). BMC Plant Biol. 11, 38. https://doi.org/10.1186/1471-2229-11-38.

Xu, J., Wang, X., Guo, W., 2015. The cytochrome P450 superfamily: key players in plant development and defense. J. Integr. Agr. 14, 1673–1686.

Zadražnik, T., Egge-Jacobsen, W., Meglič, V., Šuštar-Vozlič, J., 2017. Proteomic analysis of common bean stem under drought stress using in-gel stable isotope labeling. J. Plant Physiol. https://doi.org/10.1016/j.jplph.2016.10.015.

Zhou, Q., Kyazike, J., Echelard, Y., Meade, H.M., Higgins, E., Cole, E.S., Edmunds, T., 2005. Effect of genetic background on glycosylation heterogeneity in human antithrombin produced in the mammary gland of transgenic goats. J. Biotechnol. 117, 57–72.