Journal of Oil Palm Research Special Issue on Malaysia-MIT Biotechnology Partnership Programme: Volume 1-Oil Palm Tissue Culture (April 2008) p. 51-63

IDENTIFICATION OF GENES EXPRESSED IN THE EMBRYOID TISSUE OF OIL PALM (*Elaeis guineensis* Jacq.) TISSUE CULTURE VIA EXPRESSED SEQUENCE TAG ANALYSIS

ELYANA MD SHARIFF*; LESLIE LOW ENG TI*; HALIMAH ALIAS; CHEAH SUAN CHOO* and RAJINDER SINGH***

ABSTRACT

In vitro *propagation is the only means of vegetatively propagating oil palm. Despite this, understanding of the molecular basis of oil palm tissue culture, especially embryogenesis, remains poor. Therefore, this study used expressed sequence tags (ESTs) to investigate the genes expressed during embryogenesis of oil palm tissue culture. Total RNA was extracted from oil palm embryoid tissue. Spectrophotometry and agarose gel electrophoresis indicated that the RNA molecules were intact and suitable for complementary DNA (cDNA) library construction. Two cDNA libraries were constructed using mRNAs from oil palm embryoid tissue. The two cDNA libraries were Embryoid N (EN) and Embryoid O (EO). The primary titre for the EN cDNA library was 4.52 x 10^5 pfu, while the primary titre for the EO cDNA library was 1.20 x 10^6 pfu. The average insert sizes of the EN and EO cDNA libraries were 1.0 kb and 0.8 kb respectively. Subsequently, the EO cDNA library was chosen for generation of ESTs. Plasmid extraction was carried out for random recombinant clones picked from the cDNA library, and a total of 6535 recombinant clones were found suitable for DNA sequencing. A total of 5247 ESTs with PHRED score ≥ 20 and sequence length ≥ 100 bp were generated. Cluster analysis generated 3545 unique transcripts with 2692 singletons and 853 consensus sequences. Similarity search showed that 70% (2484/3545) of the unique transcripts had significant similarity (E value ≤ 10^{-10}) to sequences in GenBank. Gene function classification showed that the genes highly expressed are those involved in metabolism (16%), protein destination, modification and storage (10%), defence, development, ageing, disease and stress (9%). Among the genes identified are those that may potentially be involved in embryogenesis, such as the lipid transfer protein homolog (WBP1A), somatic embryogenesis receptor kinase 1 (SERK1) and defensin EGAD1. The results showed that the EST approach is an effective strategy in gene discovery and capable of generating important and useful information for gene expression studies in oil palm tissue culture.*

Keywords: EST, embryoid, tissue culture.

Date received: 31 July 2007; **Sent for revision:** 17 August 2007; **Received in final form:** 25 October 2007; **Accepted:** 7 November 2007.

* Malaysian Palm Oil Board,
P. O. Box 10620,
50720 Kuala Lumpur, Malaysia.
E-mail: rajinder@mpob.gov.my

** Malaysia Genome Institute,
UKM-MTDC Smart Technology Centre,
Universiti Kebangsaan Malaysia,
43600 UKM Bangi, Selangor, Malaysia.

INTRODUCTION

Malaysia is the world's largest producer and exporter of oil palm. In 2006, it produced 15.9 million tonnes (Oil World, 2006), exporting 14.0 million tonnes and earning RM 22 653.4 million in foreign exchange (MPOB, 2006). This shows that the palm oil industry is an important component of the

national economy, especially the agriculture sector. However, the oil palm industry faces stiff competition from other vegetable oils, such as soyabean and sunflower. Therefore, high quality planting materials are necessary to increase its yield to sustain its competitiveness.

Conventional oil palm breeding to produce seeds is time-consuming (Rohani *et al.*, 2000). Tissue culture, on the other hand, holds promise as an alternative way of propagating oil palm, producing more uniform planting materials with a 20% increase in oil yield potential (Meunier *et al.*, 1990). However, the rate of callogenesis in oil palm explants is low, only about 20% and the rate of embryogenesis too, only 6% (Wooi, 1995; Rohani *et al.*, 2000), making oil palm tissue culture rather inefficient. Little is known about the molecular mechanisms involved in the processes and the genes expressed during callogenesis and embryogenesis. Therefore, deciphering their molecular basis can help improve the oil palm tissue culture process.

This can be partly achieved by using expressed sequence tags (ESTs), which is an effective method in discovering novel genes and investigating gene expression in different organs and tissues (Adams *et al.*, 1991). ESTs are single-pass short sequences (200-700 bp) that represent the 5' or 3' ends of randomly selected cDNA clones. ESTs provide information on the transcribed region of a genome. Because cDNA libraries are typically generated from specific tissues, developmental stages or treatments, and are randomly selected for sequencing, ESTs representations provide a dynamic view of the genome content and expression (Lee *et al.*, 1995).

The success of EST programmes is largely due to the advances made in genomic and genome analyses. For example, the development of automated DNA sequencers has had a major impact as vast amount of ESTs can be generated faster and at a much reduced cost (Metzker, 2005). Furthermore, the advances in bioinformatics have aided the analysis of ESTs and helped establish databases for similarity searches and functional classification (Rhee *et al.*, 2006). These developments have had a tremendous impact on EST research, to the extent that ESTs form the largest component of public DNA sequence databases. As of January 2008, there are approximately 49 183 855 ESTs in the dbEST database which contains more than 16 million plant entries with the greatest number of entries for *Arabidopsis thaliana* (1 526 133 ESTs), rice (1 220 261 ESTs), maize (1 174 690 ESTs), wheat (1 051 196 ESTs) and *Brassica napus* (567 177 ESTs).

In oil palm, Jouannic *et al.* (2005) analysed 2411 ESTs generated from five cDNA libraries from male and female inflorescences, shoot apices and zygotic embryos. This study extends gene discovery into oil palm tissue culture, particularly for the genes expressed in embryoid tissue. It is envisaged that the generated ESTs will reflect the level and complexity of gene expression in embryoid tissues. Gaining knowledge on the genes expressed in this tissue is important, as the formation of embryoids is one of the factors that determine the success of oil palm tissue culture.

MATERIALS AND METHODS

RNA Extraction and cDNA Library Construction

Total RNA was extracted from oil palm embryoid tissues by aqueous phenol extraction as described by Rochester *et al.* (1986). Poly(A)$^+$ RNA was isolated by oligo-(dT)-cellulose chromatography. Double-stranded cDNAs were synthesized using the ZAP-cDNA® Gigapack III Gold Cloning synthesis kit (Stratagene, USA). Synthesis of the first strand cDNA was carried out with poly(A)$^+$ mRNA as template in a reverse transcription reaction using a poly(dT) oligonucleotide as primer. The product of the first-strand synthesis was then used as template in a nick-translation reaction with *RNase* H and DNA Polymerase I. The resulting cDNAs were treated with *pfu* DNA polymerase, ligated with *Eco*RI adaptors and digested with *Xho*I. The cDNAs were then size-fractionated, and the size-selected cDNAs ligated into Uni-Zap™XR vector. The ligation mixture was packaged using Gigapack® III Gold packaging extract (Stratagene, USA), to produce a cDNA library. Plating and amplification of the cDNA library were carried out as recommended by the manufacturer. Recombinant clones were screened based on blue/white selection as manifested by the reaction between *isopropyl-1-thio*-β-*galactopyranoside* (IPTG) and *5-Bromo-4-chloro-3*-β-*D-galactopyranoside* (X-gal). White colonies (*i.e.* those containing presumptive cDNA inserts) were selected randomly from the cDNA library for further analysis.

Template Preparation and DNA Sequencing

Plasmid clones were obtained by either single or mass *in vivo* excision of the pBluescript phagemid from the Uni-Zap™XR vector using the ExAssist helper phage with SOLR strain as recommended by the manufacturer (Strategene, USA) and selected on LB agar plates supplemented with ampicillin. Randomly-selected clones were grown overnight in LB broth and the plasmid DNA for sequencing was prepared using the Wizard® Plus Minipreps DNA Purification System Kit (Promega, USA). The plasmid DNA was used as template for a polymerase chain reaction (PCR) amplification of cDNA insert fragments using M13 Forward and Reverse primers in a GeneAMP® PCR System 9700 (Applied Biosystem, USA). Amplified insert fragments larger

than 250 base pairs (bp) were ethanol precipitated for purification.

Sequencing into the 5′ end of the cDNA was carried out with an SK oligonucleotide primer using the ABI PRISM™ Big Dye™ Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystem, USA). The sequencing reactions were resolved and analysed on the ABI 377 DNA sequencer (Applied Biosystem, USA).

Sequence Processing

Raw chromatogram reads were base-called using PHRED (Ewing *et al.*, 1998; Ewing and Green,1998). The vector sequences were screened using CROSS_MATCH (Green, 1999). Sequences with less than 100 non-vector good quality bases after trimming were eliminated from further analysis. The sequences were clustered and assembled using StackPACK (Miller *et al.*, 1999). The process included the sequential steps of masking, clustering assembly, alignment, analysis and consensus partitioning. Edited sequences were clustered based on their relative similarity (the default was ≥ 96% similarity over a window of 150 bases) as determined by D2_CLUSTER (Burke *et al.*, 1999), a word-based greedy clustering algorithm. The related but loose clusters were further aligned using PHRAP (Green, 1999) to improve alignment quality by generating particularly distinct sequences as singletons and highly related sequences as contigs. The aligned contigs were further analysed using CRAW (Chow and Burke, 1999). CRAW was used to analyse the contigs for error and alternative splice forms, create the final alignment and generate the best ranking consensus sequence for each assembled cluster of ESTs. The resulting singleton and consensus sequences were used as unique transcripts in this study.

Database Similarity Searches

Similarity searches were carried out against the National Centre for Biotechnology Information (NCBI) non-redundant (nr) protein database using blastx (Altschul *et al.*, 1997). Searches with an E value $\leq e^{-10}$ were considered to represent significant database matches.

Functional Classification

Unique transcripts with significant similarities to putative functions were then categorized based on the functional annotations available for *Arabidopsis thaliana* proteins following the Munich Information Centre for Protein Sequences (MIPS; http://mips.gsf.de) Functional Catalogue (FunCat) schema (Ruepp *et al.*, 2004).

RESULTS AND DISCUSSION

RNA Extraction and cDNA Library Construction

The yield and purity of the total RNA extracted from oil palm embryoid tissues are shown in *Table 1*. The yields of total RNA were 236.0-424.8 μg g^{-1} fresh weight. In all samples, the A_{260}/A_{280} ratio ranged from 1.8-1.9, indicating good quality RNA. The A_{260}/A_{230} ratio was 1.9, suggesting low level of polyphenol and polysaccharide contamination. To assess the RNA quality, agarose gel electrophoresis was carried out. The distinct 28S and 18S ribosomal RNA (rRNA) bands were visible in ultraviolet light (data not shown). The presence of these distinct bands on the gel indicated that the RNA obtained was intact.

Approximately 4 μg column purified poly$(A)^+$ mRNA were used to synthesize double-stranded cDNAs. Two cDNA libraries (EN and EO) were constructed. The characteristics of the libraries are shown in *Table 2*. The titre of the primary library was 4.5×10^5 pfu for the EN library and 1.2×10^6 pfu for the EO library. Preliminary screening of both the primary libraries, EN and EO, indicated that the percentages of recombinant clones were 96.4% and 95.3%, and their average insert sizes ranged from 0.7 kb to 2.3 kb and 0.5 kb to 2.3 kb, respectively.

Amplifications of both the libraries were carried out to produce large and stable quantities of cDNA clones. The titres of the amplified libraries increased to 2.9×10^9 pfu ml^{-1} and 2.2×10^9 pfu ml^{-1} for the EN and EO libraries, respectively. Preliminary screening of both the amplified libraries, EN and EO, indicated that the percentages of recombinant clones were 97.0% and 90.5% with average insert sizes ranged from 1.3 kb to 2.9 kb and 0.6 kb to 2.3 kb, respectively. The results showed that the quality of the cDNA library sufficient to identify the expressed genes in oil palm embryoid tissues. Nevertheless, the EO library was chosen to generate ESTs as the primary library had a higher titre than the EN library.

Generation of Sequence Data

Plasmids prepared from 6535 recombinants clones were selected for PCR amplification. Analysis of the PCR products on agarose gel showed that the size of the inserts ranged from 0.5 kb to 3.0 kb with an average of 0.8 kb.

Single-pass sequencing of the 5′-end of the cDNA clones was performed. All the 6535 recombinant clones were sequenced. Among them, 704 produced poor quality sequences and were excluded from further analysis. In total, 5831 EST sequences were generated. The quality of the chromatograms generated was analysed using the PHRED (Ewing *et al.*, 1998; Ewing and Green, 1998) programme. Edited sequences of less than 100 nucleotides and

TABLE 1. RESULTS OF THE RNA EXTRACTION FROM OIL PALM EMBRYOID TISSUES

Extraction	Concentration ($\mu g\ \mu l^{-1}$)	Yield ($\mu g\ g^{-1}$ of sample)	Purity	
			(A_{260}/A_{280})	(A_{260}/A_{230})
1	1.840	368.0	1.9	1.8
2	2.124	424.8	1.9	1.9
3	1.180	236.0	1.9	1.8

TABLE 2. CHARACTERISTICS OF EMBRYOID CDNA LIBRARIES

	cDNA library	
	EN	EO
Primary library		
Titre (pfu)	4.5×10^5	1.2×10^6
Percentage of recombinant	96.4	95.3
Insert size range (kb)	0.7-2.3	0.5-2.3
Average size (kb)	1.0	0.8
Amplified library		
Titre (pfu ml^{-1})	2.9×10^9	2.2×10^9
Percentage of recombinant	97.0	90.5
Insert size range (kb)	1.3-2.9	0.6-2.3
Average size (kb)	1.3	0.8

sequences with a PHRED score below 20 were not considered for further analysis. Following PHRED analysis and subsequent editing, 5247 good quality sequences were obtained with an average sequence length of 509 bp.

Cluster Construction

The StackPACK (Miller *et al.*, 1999) package was used to organize the embryoid transcripts. The sequences were organized either as singletons or consensi. The clustering process was divided into three main steps. In the first step, loose clusters were defined on the basis of sequence similarity using the D2_CLUSTER (Burke *et al.*, 1999) program, and the sequences grouped together if there was a 96% sequence similarity in a window of 150 bases. Some ESTs were left as single sequences as they did not match any other sequence. Subsequently, the loose clusters were aligned using PHRAP (Green, 1999), to verify the clusters and generate contigs, or to separate the genes into clusters if the differences in the sequences exceeded the set criteria. The alignment of the overlapping sequences inside each contig generates a new consensus sequence and if two or more contigs are in a cluster they would maintain a linked relationship being found in the same cluster. Lastly, the contigs were further refined using CRAW (Chow and Burke, 1999), which allows the identification of alternative spliced sequences of genes.

A total of 5247 ESTs were analysed through StackPACK, which led to the identification of 2692 singletons and 853 consensi. The StackPACK analysis of these 3545 unique transcripts from embryoid cDNA library is shown in *Table 3*. These numbers, however, could be an overestimate, since some of the unique ESTs could be non-overlapping sequences of the same transcript.

The frequency of the EST distribution is shown in *Figure 1*. The number of ESTs per consensus ranged between 1 (93 consensi) and 45 (one consensus).

TABLE 3. StackPACK ANALYSIS RESULTS

Analysis	No. of sequences
Total number of ESTs analysed	5 247
Singletons	2 692
Consensi	853
Unique transcripts (singletons + consensuses)	3 545

Database Similarity Searches

The 3545 unique transcripts assembled from embryoid ESTs with StackPACK were used to search against the NCBI nr protein database using blastx (Altschul *et al.*, 1997), and 2484 (70%) were found to have significant similarities.

The 30 most abundant genes in the EST collection are listed in *Table 4*. The number of ESTs representing these genes are 473, accounting for 9% of the total

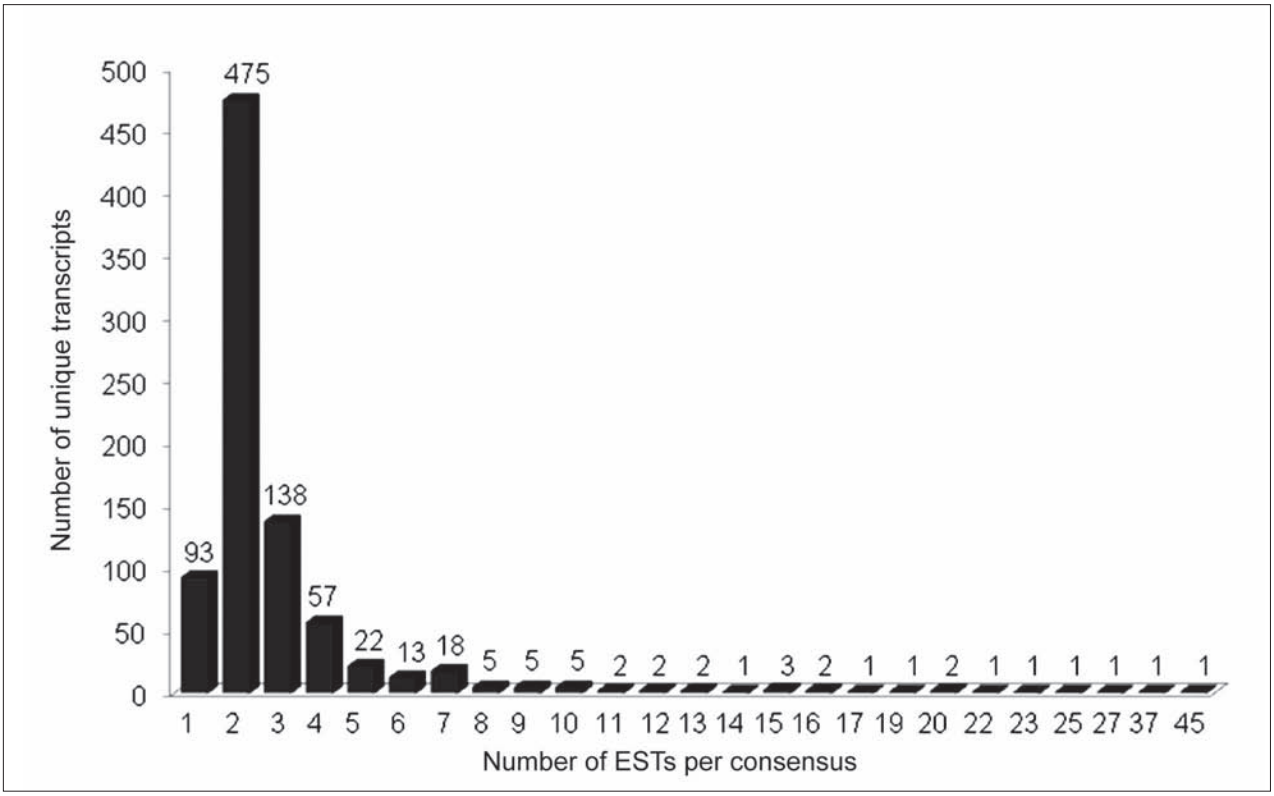

Figure 1. Expressed sequence tags (EST) distribution in the consensi of the embryoid cDNA library.

ESTs. Some of these ESTs, such as catalase 2, chitinase and metallothionin-like protein, have been shown to be functionally involved in plant defense reactions. As such, some of the most abundant ESTs in embryoid tissues correspond to the proteins involved in various defence mechanisms.

Interestingly, another group of abundant ESTs were genes without significant similarities or encoding for unknown proteins. The former could have corresponded to genes not well characterized in other species or as pointed out by Jouannic *et al.* (2005), may contain non-coding regions such as the 5' untranslated. Nonetheless, these genes are expressed in the embryoid tissues and are interesting candidates for further analysis. The list of highly expressed genes in the present collection differs to some extent with what was reported for oil palm by Jouannic *et al.* (2005). This is not surprising as the level of expressed genes differs depending on the tissue or the specific developmental stages being examined.

Functional Classifications

The significant putative functions for the unique transcripts were further categorized based on the functional annotations available for the *Arabidopsis thaliana* proteins following the Munich Information Centre for Protein Sequences (MIPS; http://mips.gsf.de) Functional Catalogue (FunCat) schema (Ruepp *et al.*, 2004). This approach assumes that the functionality is transferable based on sequence conservation. The 2484 unique transcripts with significant similarity to sequences in the databases were classified into 11 functional categories. The proportion of unique transcripts in each functional category is shown in *Figure 2*.

Of the 2484 unique transcripts analysed, the largest number (30%) was for unclassified proteins, mainly of unknown or hypothetical genes. The category 'metabolism' contained the highest numbers of identified genes (16%), followed by the categories 'protein destination, modification and storage' (10%) and 'defence, development, ageing and stress' (9%). These three functional categories were also reported as the major gene categories in *Arabidopsis* (Arabidopsis Genome Initiative, 2000) and rice (Goff *et al.*, 2002). The other categories included genes related to 'translation and transcription' (7%), 'cell division, cell cycle and DNA processing' (6%), 'cellular transportation' (6%), 'cellular communication and signal transduction' (4%), 'cellular component biogenesis and cellular organization' (4%) and energy (2%). Whether they are true homologues of the corresponding genes will require more experimental evidence to ascertain. Nonetheless, these results provide a glimpse of the diversity of genes in the oil palm embryoid cDNA collection.

Most of the genes in the functional category 'metabolism' are enzymes from the glycolysis and fatty acid synthesis pathways, such as glyceraldehid-

TABLE 4. THE 30 MOST ABUNDANT GENES IN THE OIL PALM EMBRYOID EST COLLECTION

EST frequency	Unique transcripts ID	Putative identity	Organism	E value
45	cn849	Hypothetical protein	*Zea mays*	$9e^{-98}$
37	cn279	Lipid transfer protein homolog (WBP1A)	*Triticum aestivum*	$1e^{-25}$
27	cn787	Catalase 2	*Zantedeschia aethiopica*	0.0
25	cn328	PVR3-like protein	*Ananas comosus*	$2e^{-17}$
23	cn737	No significant similarity	-	-
22	cn518	Ribosomal protein S3	*Typha latifolia*	$1e^{-112}$
20	cn554	Hypothetical protein	*Oryza sativa*	$2e^{-49}$
20	cn773	Unknown protein	*Oryza sativa*	$2e^{-59}$
19	cn618	1-Cys peroxiredoxin	*Fagopyrum esculentum*	$4e^{-97}$
17	cn391	Extensin	*Nicotiana sylvestris*	$5e^{-29}$
16	cn706	Sucrose synthase 21	*Tulipa gesneriana*	$1e^{-107}$
16	cn794	Metallothionein-like protein (type 2)	*Typha latifolia*	$6e^{-20}$
15	cn273	No significant similarity	-	-
15	cn520	Mannose/glucose-specific lectin	*Parkia platycephala*	$1e^{-48}$
15	cn784	Extensin-like protein	*Capsicum chinense*	$2e^{-24}$
14	cn702	Proline-rich SAC51	*Brassica napus*	$1e^{-32}$
13	cn580	Class IV chitinase	*Vitis vinifera*	$1e^{-104}$
13	cn833	Mannose-binding lectin precursor	*Zingiber officinale*	$4e^{-47}$
12	cn777	No significant similarity	-	-
12	cn778	Cyclophilin	*Ricinus communis*	$2e^{-77}$
11	cn707	Blight-associated protein p12 precursor	*Oryza sativa*	$7e^{-34}$
11	cn729	Non-symbiotic hemoglobin class 1	*Malus domestica*	$2e^{-67}$
10	cn349	Ubiquitin-conjugating enzyme	*Oryza sativa*	$3e^{-79}$
10	cn390	Proline-rich-like protein	*Asparagus officinalis*	$6e^{-32}$
10	cn523	Early flowering protein 1	*Asparagus officinalis*	$3e^{-47}$
10	cn621	Pescadillo-like protein	*Oryza sativa*	$1e^{-134}$
10	cn806	Plasma membrane intrinsic protein	*Zea mays*	$1e^{-144}$
9	cn689	Metallothionein-like protein	*Elaeis guineensis*	$2e^{-34}$
9	cn776	No significant similarity	-	-
9	cn796	Defensin EGAD1	*Elaeis guineensis*	$4e^{-21}$

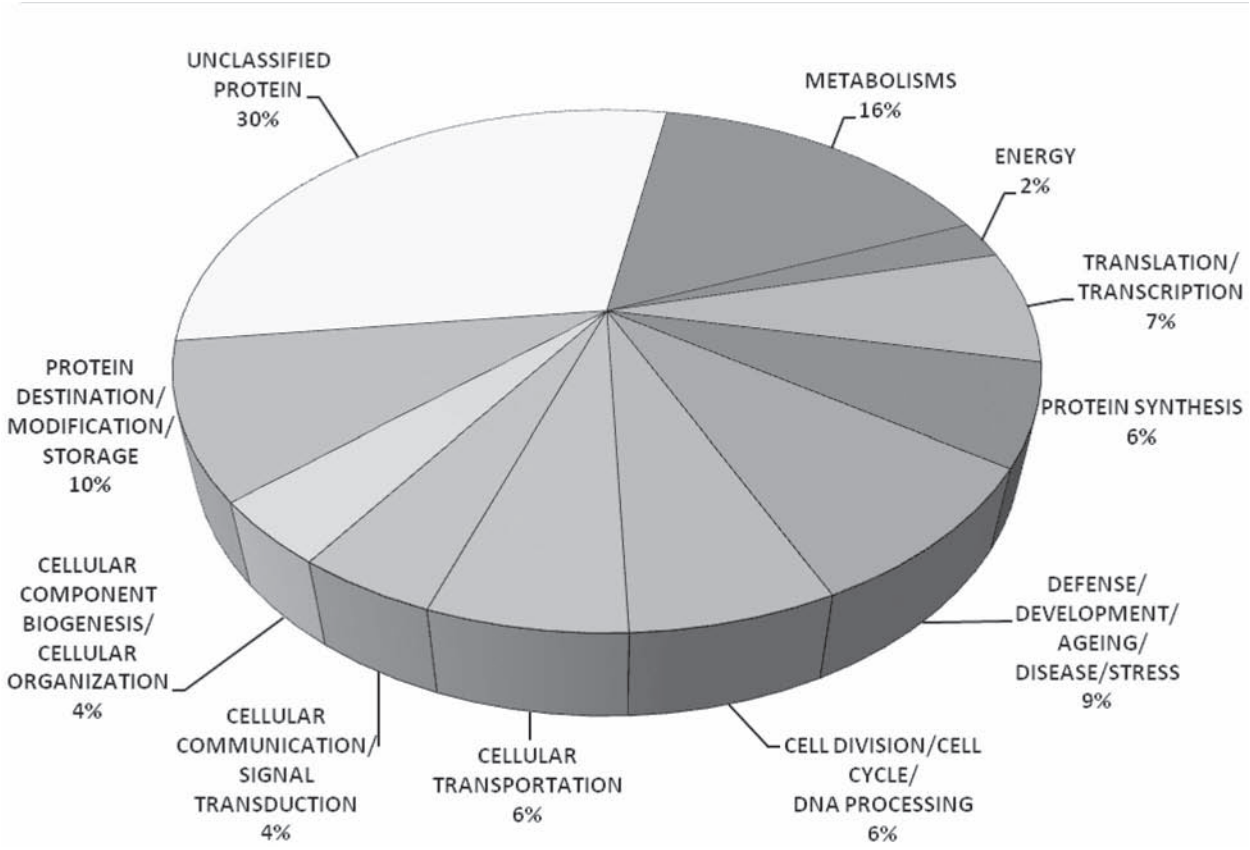

Figure 2. Functional classifications of the unique transcripts. Only unique transcripts with significant similarities were classified.

3-phosphate dehydrogenase, alcohol dehidrogenase, ACC oxidase and 3-ketoacyl-CoA thiolase.

About 10% of the unique transcripts with significant similarity to known proteins were categorized as genes involved in 'protein destination, modification and storage' (*Figure* 2). From the unique transcripts generated, globulin, extensin and profilin were the main storage proteins synthesized. Besides storage proteins, ubiquitin extension protein-like protein, ubiquitin fusion protein and ubiquitin fusion-degradation protein-like were also partially sequenced from the embryoid tissues of oil palm.

Plants are continuously subject to biotic stresses throughout their life, and their survival depends on the diversity of specific genetic systems and metabolic processes to maintain their integrity in the face of a challenging environment (Jackson and Taylor, 1996). As such, it is not surprising that genes involved in defence, development, ageing, disease and stress are represented by 9% of the unique transcripts in the embryoid tissues.

Genes Involved in Oil Palm Tissue Culture Embryogenesis

Embryogenesis is a complex developmental step in which somatic cells are induced to form totipotent embryogenic cells capable of becoming complete plants. The process of embryogenesis in culture starts with the induction of embryogenic cells with high concentrations of auxin. Although the morphology of plant embryogenesis has been extensively documented through microscopy (Mayer *et al.*, 1991), understanding of the molecular events that generate plant embryos is still limited.

Several studies have identified genes responsible for the various stages of embryogenesis. They include hormone responsive genes, homeobox genes, late embryogenesis abundant (LEA) genes, chitinases and kinases regulating somatic embryogenesis.

From the ESTs data generated in the study, the genes that can play important roles in embryogenesis were identified based on literature references and are listed in *Table 5*.

Among them, the lipid transfer protein homolog (WBP1A) is one of the genes of interest. The first lipid transfer protein homolog to be isolated and characterized was from carrot embryogenic cultures and was shown to be secreted extracellularly to transport lipids from their places of synthesis in the endoplasmic reticulum to various locations (Kader, 1997). Zeng *et al.* (2006) confirmed that the lipid transfer protein homolog is highly expressed in embryogenic callus of cotton. From the analysis shown, cn279 consists of 37 ESTs encoding a lipid

TABLE 5. LIST OF GENES THAT MAY BE REGULATED DURING EMBRYOGENESIS IN OIL PALM

Unique transcript ID	Putative identity	Organism	E value	*Genbank* accession No.	Reference
Metabolism					
cn116	Lanatoside 15'-O-acetylesterase-like	*Oryza sativa*	$2e^{-45}$	BAD61510	Kandzia *et al.* (1998)
cn662	Glutamine synthetase precursor	*Juglans nigra*	$6e^{-80}$	AAD49734	Rodriguez *et al.* (2006)
cn390	Proline-rich-like protein	*Asparagus officinalis*	$6e^{-32}$	CAA57810	Györgyey *et al.* (1997)
cn298	Glyceraldehyde 3-phosphate dehydrogenase	*Elaeis guineensis*	$2e^{-92}$	ABB72846	Carlson *et al.* (2006)
cn232	3-ketoacyl-CoA thiolase	*Oryza sativa*	$1e^{-108}$	XP_468412	Kato *et al.* (1996)
cn224	ACC oxidase	*Musa acuminata*	$2e^{-44}$	CAA11200	Reynolds & John (2000)
cn353	Lipoxygenase	*Fragaria* x *ananassa*	$1e^{-133}$	CAE17327	Liu *et al.* (1994)
cn378	Enolase	*Oryza sativa*	$1e^{-162}$	AAP94211	Eastmond & Rawsthorne (2000)
cn459	3-glucanase	*Oryza sativa*	$4e^{-54}$	XP_480764	Dong & Dunstan (2000)
cn476	4-alpha-glucanotransferase	*Oryza sativa*	$3e^{-46}$	NP_917928	Bresolin *et al.* (2006)
cn732	Sucrose synthase	*Oncidium* cv. *'Goldiana'*	$1e^{-47}$	AAM95943	Iraqi & Tremblay (2001)
pOP-EO07865	EMB1187; Kinase phosphotransferase	*Arabidopsis thaliana*	$2e^{-34}$	NP_180251	Tzafrir *et al.* (2004)
pOP-EO03370	EMB1075; Carboxy-lyase	*Arabidopsis thaliana*	$5e^{-31}$	NP_175036	Tzafrir *et al.* (2004)
pOP-EO03617	Xyloglucan endotransglycosylase XET1	*Asparagus officinalis*	$5e^{-45}$	AAF80590	Malinowski *et al.* (2004)
pOP-EO05601	Aminopeptidase M	*Oryza sativa*	$4e^{-45}$	XP_464667	Citharel & Garreau (1987)
Defence/development/ageing/diseases/stress					
cn580	Chitinase class IV	*Vitis vinifera*	$1e^{-104}$	AAG61140	Robinson *et al.* (1997)
cn511	Glutation S-transferase	*Oryza sativa*	$2e^{-94}$	XP_493844	Galland *et al.* (2001)
cn689	Metallothionein-like protein	*Elaeis guineensis*	$2e^{-34}$	CAB52585	Dong & Dunstan (1996)
cn708	QM-like protein	*Elaeis guineensis*	$1e^{-123}$	AAG27431	Chen *et al.* (2006)
cn819	Heat shock protein 82	*Oryza sativa*	$1e^{-137}$	BAD73668	Kitamiya *et al.* (2000)
cn313	Pathogenesis related protein	*Oryza sativa*	$1e^{-89}$	NP_916778	Helleboid *et al.* (2000)
cn596	Peroxidase	*Vigna angularis*	$1e^{-123}$	BAA01950	Takeda *et al.* (2003)
cn787	Catalase 2	*Zantedeschia aethiopica*	0	AAG61140	Lino-Neto *et al.* (2004)
cn796	Defensin EGAD1	*Elaeis guineensis*	$4e^{-21}$	AAN52490	Tregear *et al.* (2002)
pOP-EO05604	Embryonic flower 2	*Yucca filamentosa*	$2e^{-43}$	ABD85300	Yoshida *et al.* (2001)
pOP-EO02024	Manganese superoxide dismutase	*Avicennia marina*	$3e^{-57}$	AAN15216	Zhu & Scandalios (1994)

TABLE 5. LIST OF GENES THAT MAY BE REGULATED DURING EMBRYOGENESIS IN OIL PALM (Continued)

Cell division/cell cycle/DNA processing					
pOP-EO05757	SCARECROW-like protein	*Oryza sativa*	$1e^{-74}$	NP_915059	Heidstra *et al.* (2004)
pOP-EO02211	Late embryogenesis abundant protein (LEA)	*Oryza sativa*	$2e^{-50}$	XP_464870	Wilde *et al.* (1998)
pOP-EO06008	Cyclin-dependent protein kinase	*Arabidopsis thaliana*	$1e^{-32}$	NP_191614	Corellou *et al.* (2001)
pOP-EO03661	Embryogenic callus protein 98b	*Oryza sativa*	$4e^{-61}$	XP_470376	
Cellular component biogenesis/cellular organization					
cn751	Pollen Ole e 1 allergen and extensin	*Medicago truncatula*	$5e^{-23}$	ABD33344	Rodrigues (1994)
pOP-EO04436	Cell wall hydrolase	*Hyacinthus orientalis*	$2e^{-20}$	AAS20972	Malinowski & Filipecki (2002)
pOP-EO04860	Cellulose synthase	*Populus tremula* x *Populus tremuloides*	$5e^{-89}$	AAT09898	Goubet *et al.* (2003)
Protein destination/modification/storage					
cn318	Calcium-dependent protein kinase	*Oryza sativa*	$6e^{-89}$	AAT75264	Anil *et al.* (2000)
cn403	Germin A	*Oryza sativa*	$6e^{-79}$	XP_480463	Bishop-Hurley *et al.* (2003)
pOP-EO03418	Cysteine protease component of protease-inhibitor complex	*Zea mays*	$4e^{-66}$	BAA88898	He & Kermode (2003)
Cellular communication/signal transduction					
cn136	Somatic embryogenesis protein kinase 1	*Oryza sativa*	$1e^{-74}$	BAD68873	Hecht *et al.* (2001)
pOP-EO05078	GTP-binding protein RAB11G	*Oryza sativa*	$1e^{-69}$	XP_483418	Gonçalves *et al.* (2005)
pOP-EO03016	Gibberelin 2-oxidase	*Cucurbita maxima*	$1e^{-51}$	CAC85924	Tokuji & Kuriyama (2003)
pOP-EO04009	MAP kinase	*Papaver rhoeas*	$3e^{-51}$	CAH05025	Seguí-Simarro *et al.* (2005)
pOP-EO04749	Receptor-like protein kinase	*Elaeis guineensis*	$8e^{-94}$	AAO26312	Schmidt *et al.* (1997)
Cellular transportation					
cn279	Lipid transfer protein homolog (WBP1A)	*Triticum aestivum*	$1e^{-25}$	AAB32995	Zeng *et al.* (2006)
cn028	Auxin-induced lipid transfer protein	*Hyacinthus orientalis*	$2e^{-22}$	AAT08737	Basu *et al.* (2002)
pOP-EO02858	Calmodulin	*Glycine max*	$4e^{-37}$	AAA34015	Overvoorde & Grimes (1994)
Energy					
cn612	Cytochrome P450	*Oryza sativa*	$2e^{-93}$	NP_917788	Bishop-Hurley *et al.* (2003)
Translation/transcription					
cn597	Elongation factor 1-alpha 1	*Elaeis guineensis*	0	AAT45847	Kawahara *et al.* (1992)
pOP-EO03156	ZF-HD homeobox protein	*Oryza sativa*	$6e^{-15}$	BAD28899	Chugh & Khurana (2002)
pOP-EO05943	WRKY DNA-binding protein	*Nicotiana tabacum*	$1e^{-60}$	BAB61056	Lagace & Matton (2004)

transfer protein homolog (WBP1A). Therefore, the lipid transfer protein homolog may also play an active role in oil palm embryogenesis.

The second gene of interest is the somatic embryogenesis protein kinase 1 (SERK1). Hecht *et al.* (2001) reported that high levels of *SERK1* expression were detected when competent somatic cells underwent embryogenesis to become embryos. This gene is also known to play a crucial role in somatic embryogenesis of carrot (Schmidt *et al.*, 1997), where the transcripts are expressed transiently in zygotic embryos as well as in competent somatic cells, demonstrating an embryogenesis-specific signal transduction chain mediated by the *SERK 1* gene. The consensus 136 (cn136) is represented by 5 ESTs encoding the *SERK1* gene. Therefore, it would be interesting to examine the role of SERK 1 in oil palm embryogenesis.

And, finally, the third gene of interest is defensin EGAD1. Defensins are thought to play a role in pathogen defence and, in some cases, have been shown to exert anti-fungal action (Thevissen *et al.*, 1996). Defensin EGAD1 from oil palm was first isolated and characterized by Tregear *et al.* (2002) who determined that defensin EGAD1 transcript accumulated in significantly greater quantities in oil palm tissue culture samples. Consensus 796 (cn796) consists of 9 ESTs encoding a defensin EGAD1. Therefore, the expression of defensin EGAD1 in oil palm tissue culture, particularly during embryogenesis, is worth investigating.

For the time being, no definite conclusion can be drawn regarding the regulation of these genes during oil palm tissue culture. The expression profiles of these three genes particularly during the tissue culture process are being determined experimentally.

CONCLUSION

Two cDNA libraries were constructed using mRNAs from oil palm embryoid tissue. A total of 6535 recombinant clones were found suitable for DNA sequencing. A total of 5247 ESTs with PHRED score ≥ 20 and sequence length ≥ 100 bp were generated. Cluster analysis generated 3545 unique transcripts, distributed in 2692 singletons and 853 consensus sequences. Similarity searches showed that 70% (2484/3545) of the unique transcripts have significant similarity (E value ≤ 10^{-10}) to sequences in GenBank. Gene function classification showed that genes expressed in embryoid tissue are those involved in metabolism (16%), protein destination, modification and storage (10%), and defence, development, ageing, disease and stress (9%). Interestingly, the genes identified include those that may have a role in embryogenesis, such as the lipid transfer protein homolog (WBP1A), somatic embryogenesis receptor kinase 1 (SERK1) and defensin EGAD1. The EST approach proved to be an effective strategy in gene discovery, generating important and useful information on the genes expressed in the embryoid tissue of oil palm.

Research is underway to put these ESTs on a DNA microarray, as a means to understand their quantitative expression. Efforts are also ongoing to generate a larger number of ESTs to facilitate direct investigation into gene expression and as a valuable resource for oil palm tissue culture.

REFERENCES

ADAMS, M D; KELLEY, J M; GOCAYNE, J D; DUBNICK, M; POLYMEROPOULOS, M H; XIAO, H; MERRIL, C R; WU, A; OLDE, B; MORENO, R F; KERLAVAGE, A R; MCCOMBIE, W R and VENTER, J C (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. *Science, 252*: 1651-1656.

ALTSCHUL, S F; MADDEN, T L; SCHÄFFER, A A; ZHANG, J; ZHANG, Z; MILLER, W and LIPMAN, D J (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res., 25*: 3389-3402.

ANIL, V S; HARMON, A C and SANKARA-RAO, K (2000). Spatio-temporal accumulation and activity of calcium-dependent protein kinases during embryogenesis, seed development, and germination in sandalwood. *Plant Physiol., 122*: 1035-1041.

ARABIDOPSIS GENOME INITIATIVE (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature, 408*: 796-815.

BASU, S; SUN, H; BRIAN, L; QUATRANO, R L and MUDAY, G K (2002). Early embryo development in *Fucus distichus* is auxin sensitive. *Plant Physiol., 130*: 292-302.

BISHOP-HURLEY, S L; RICHARD C GARDNER, R C and WALTER, C (2003). Isolation and molecular characterization of genes expressed during somatic embryo development in *Pinus radiata*. *Plant Cell, Tissue and Organ Culture, 74*: 267-281.

BRESOLIN, N S; LI, Z; KOSAR-HASHEMI, B; TETLOW, I J; CHATTERJEE, M; RAHMAN, S; MORELL, M K and HOWITT, C A (2006). Characterisation of disproportionating enzyme from wheat endosperm. *Planta, 224(1)*: 20-31.

BURKE, J; DAVISON, D and HIDE, W (1999). D2_cluster: a validated method for clustering EST

and full-length cDNA sequences. *Genome Res., 9(11)*: 1135-1142.

CARLSON, J E; LEEBENS-MACK, J H; WALL, P K; ZAHN, L M; MUELLER, L A; LANDHERR, L L; HU, Y; ILUT, D C; ARRINGTON, J M; CHOIREAN, S; BECKER, A; FIELD, D; TANKSLEY, S D; MA, H and DEPAMPHILIS, W (2006). EST database for early flower development in California poppy (*Eschscholzia californica* Cham., Papaveraceae) tags over 6000 genes from a basal eudicot. *Plant Mol. Biol*., *62*: 351-369.

CHEN, C; WANDURAGALA, S; BECKER, D F and DICKMAN, M B (2006). Tomato QM-like protein protects *Saccharomyces cerevisiae* cells against oxidative stress by regulating intracellular proline levels. *Appl Environ Microbiol*., *72(6)*: 4001-4006.

CHOW, A and BURKE, J P (1999). CRAWview:for viewing splicing variation, gene families, and polymorphism in clusters of ESTs and full-length sequences. *Bioinformatics, 15(5)*: 376-381.

CHUGH, A and KHURANA, P (2002). Gene expression during somatic embryogenesis: recent advances. *Curr. Sci., 83*: 715-730.

CITHAREL, J and GARREAU, C (1987). Aminopeptidase from the cotyledons of *Vicia faba* L. Var. Minor seeds: partial purification and some properties. *New Phytologist*, *107(3)*: 499-506.

CORELLOU, F; BROWNLEE, C; DETIVAUD, L; KLOAREG, B and BOUGET, F (2001). Cell cycle in the fucus zygote parallels a somatic cell cycle but displays a unique translational regulation of cyclin-dependent kinases. *The Plant Cell, 13*: 585-598.

DONG, J and DUNSTAN, D I (2000). Expression of abundant mRNAs during somatic embryogenesis of white spruce. *Planta, 199*: 459-466.

EASTMOND, P J and RAWSTHORNE, S (2000). Coordinate changes in carbon partitioning and plastidial metabolism during the development of oilseed rape embryos. *Plant Physiol*., *122*: 767-774.

EWING, B; HILLIER, L A D; WENDL, M C and GREEN, P (1998). Base-calling of automated sequences traces using *Phred*. I. Accuracy assessment. *Genome Res., 8*: 175-185.

EWING, B and GREEN, P (1998). Base-calling of automated sequencer traces using *Phred.* II. Error probabilities. *Genome Res., 8*: 186-194.

GALLAND, R; RANDOUX, B; VASSEUR, J and HILBERT, J L (2001). A glutathione S-transferase cDNA identified by mRNA differential display is upregulated during somatic embryogenesis in *Cichorium*. *Biochem. Biophys. Acta, 1522*: 212-216.

GOFF, S A; RICKE, D; LAN, T H; PRESTING, G; WANG, R; DUNN, M; GLAZEBROOK, J; SESSIONS, A; OELLER, P and VARMA, H (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science, 296*: 92-100.

GONÇALVES, S; CAIRNEY, J; OLIVEIRA, M and MIGUEL, C (2005). Identification of differentially expressed genes during embryogenesis in maritime pine (*Pinus pinaster*). *Silva Lusitana, 13(2)*: 203-216.

GOUBET, F; MISRAHI, A; PARK, S K; ZHANG, Z; TWELL, D and DUPREE, P (2003). AtCSLA7, a cellulose synthase-like putative glycosyltransferase, is important for pollen tube growth and embryogenesis in *Arabidopsis*. *Plant Physiol*., *131*: 547-557.

GREEN, P (1999). SWAT/Crossmatch/PHRAP package. University of Washington. http://www.phrap.org.

GYÖRGYEY, J; NÉMETH, K; MAGYAR, Z; KELEMEN, Z; ALLIOTTE, T; INZÉ, D; DUDITS, D (1997). Expression of a novel-type small proline-rich protein gene of alfalfa is induced by 2,4-dichlorophenoxiacetic acid in differentiated callus cells. *Plant Mol. Biol., 34*: 593-602.

HE, X and KERMODE, A R (2003). Proteases associated with programmed cell death of megagametophyte cells after germination of white spruce (*Picea glauca*) seeds. *Plant Mol. Biol*., *52(4)*: 729-744.

HECHT, V; VIELLE-CALZADA, J P; HARTOG, M V; SCHMIDT, E D L; BOUTILIER, K; GROSSNIKLAUS, U and DE VRIES, S C (2001). The Arabidopsis somatic embryogenesis receptor Kinase 1 is expressed in developing ovules and embryos and enhances embryogenic competence in culture. *Plant Physiol., 127*: 803-816.

HEIDSTRA, R; WELCH, D and SCHERES, B (2004). Mosaic analyses using marked activation and deletion clones dissect Arabidopsis SCARECROW action in asymmetric cell division. *Genes Dev*., *18(16)*: 1964-1969.

HELLEBOID, S; HENDRICKS, T; BAUW, G; INZE, D; VASSEUR, J and HILBERT, J L (2000). Three major somatic embryogenesis related proteins in

Chichorium identified as PR proteins. *J. Exp. Bot., 51*: 1189-1200.

IRAQI, D and TREMBLAY, F M (2001). Analysis of carbohydrate metabolism enzymes and cellular contents of sugars and proteins during spruce somatic embryogenesis suggests a regulatory role of exogenous sucrose in embryo development. *J. Exp. Botany, 52*: 2301-2311.

JACKSON, A O and TAYLOR, C B (1996). Plant-microbe interactions: life and death at the interface. *Plant Cell, 8*: 1651-1668.

JOUANNIC, S; ARGOUT, X; LECHAUVE, F; FIZAMES, C; BORGEL, A; MORCILLO, F; ABERLENC-BERTOSSI, F; DUVAL, Y and TREGEAR, J (2005). Analysis of expressed sequence tags from oil palm. *FEBS Letters, 579 (12)*: 2709-2714.

KADER, J C (1997). Lipid-transfer proteins: a puzzling family of plant proteins. *Trends Plant Sci., 2*: 66-70.

KANDZIA, R; GRIMM, R; ECKERSKORN, C; LINDEMANN, P and LUCKNER, M (1998). Purification and characterization of lanatoside 15'-O-acetylesterase from *Digitalis lanata* Ehrh. *Planta, 204 (3)*: 383-389.

KAWAHARA, R; SUNABORI, S; FUKUDA, H and KOMAMINE, A (1992). A gene expressed preferentially in the globular stage of somatic embryogenesis encodes elongation factor 1-α 1 in carrot. *Eur. J. Biochem., 209*: 157-162.

KATO, A; HAYASHI, M; TAKEUCHI,Y and NISHIMURA, M (1996). cDNA cloning and expression of a gene for 3-ketoacyl-CoA thiolase in pumpkin cotyledons. *Plant Mol. Biol., 31(4)*: 843-852.

KITAMIYA, E; SUZUKI, S; SANO, T and NAGATA, T (2000). Isolation of two genes that were induced upon initiation of somatic embryogenesis on carrot hypocotyls by high concentrations of 2,4-D. *Plant Cell Rep., 19*: 551-557.

LAGACE, M and MATTON, D P (2004). Characterization of a WRKY transcription factor expressed in late torpedo-stage embryos of *Solanum chacoense*. *Planta., 219(1)*: 185-189.

LEE, N H; WEINSTOCK, K G; KIRKNESS, E F; EARLE-HUGHES, J A; FULDNER, R A; MARMAROS, S; GLODEK, A; GOCAYNE, J D; ADAMS, M D and KERLAVAGE, A R; FRASER, C M and VENTER, J C (1995). Comparative expressed sequence tags analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment. *Proc. Natl. Acad. Sci. USA, 92*. p. 8303-8307.

LINO-NETO,T; PIQUES, M C; BARBETA, C; SOUSA, M F; TAVARES, R M and PAIS, M S (2004). Identification of *Zantedeschia aethiopica* Cat1 and Cat2 catalase genes and their expression analysis during spathe senescence and regreening. *Plant Sci., 167 (4)*: 889-898.

LIU, W; HILDEBRAND, D F; MOORE, P J and COLLINS, G B (1994). Expression of desiccation-induced and lipoxygenase genes during the transition from the maturation to the germination phases in soybean somatic embryos. *Planta, 194(1)*: 69-76.

MPOB (2006). Malaysian oil palm statistics. http://www.mpob.gov.my/statistics.html (Accessed on 8 Januari 2007.)

MALINOWSKI, R and FILIPECKI, M (2002). The role of cell wall in plant embryogenesis. *Cell Mol Biol Lett., 7(4)*: 1137-1151.

MALINOWSKI, R; FILIPECKI, M; TAGASHIRA, N; WISNIEWSKA, A; GAJ, P; PLADER, W and MALEPSZY, S (2004). Xyloglucan endo-transglucosylase/hydrolase genes in cucumber (*Cucumis sativus*) - differential expression during somatic embryogenesis. *Physiol Plantn, 120(4)*: 678-685.

MAYER, U; TORRES-RUIZ, R A; BERLETH, T; MISERA, S and Jurgens, G (1991). Mutations affecting body organization in the *Arabidopsis* embryo. *Nature, 353*: 402-407.

METZKER, M L (2005). Emerging technologies in DNA sequencing. *Genome Res., 15(12)*: 1767-1776.

MEUNIER, J; BAUDOUIN, L; NOUY, B and NOIRET, J M (1990). The expected value of oil palm clones. *Oleagineux, 43*: 195-200.

MILLER, R T; CHRISTOFFELS, A G; GOPALAKRISHNAN, C; BURKE, J; PTITSYN, A A; BROVEAK, T R and HIDE, W A (1999). A comprehensive approach to clustering of expressed human gene sequence: the sequence tag alignment and consensus knowledge base. *Genome Res., 9*: 1143-1155.

OIL WORLD (2006). *Oil World Annual*. ISTA Mielke GmbH.

OVERVOORDE, P J; GRIMES, H D (1994). The role of calcium and calmodulin in carrot embryogenesis. *Plant Cell Physiol., 35*: 135-144.

REYNOLDS, E A and JOHN, P (2000). ACC oxidase is found in seedlings of two (Coniferales, Gnetales) of the four gymnosperm orders. *Physiol. Plant.,110 (1)*: 38-41.

RHEE, S Y; DICKERSON, J and XU, D (2006). Bioinformatics and its applications in plant biology. *Annu. Rev. Plant Biol*., *57*: 335-360.

ROBINSON, S P; JACOBS, A K and DRY, I B (1997). A class IV chitinase is highly expressed in grape berries during ripening. *Plant Physiol*., *114*: 771-778.

ROCHESTER, D E; WINER, J and SHAH, D M (1986). The structure and expression of maize genes encoding the major heat shock protein, hsp 70. *Embo J., 5*: 451-458.

RODRIGUES, R (1994). Cloning and expression of Ole e I, the major allergen from olive tree pollen. *J. Biol. Chem*., *269*: 15217-15222.

RODRIGUEZ, M J; SUAREZ, M F; HEREDIA, R; AVILA, C; BRETON, D; TRONTIN, J F; FILONOVA, L; BOZHKOV, P; VON ARNOLD, S; HARVENGT, L and CANOVAS, F M (2006). Expression patterns of two glutamine synthetase genes in zygotic and somatic pine embryos support specific roles in nitrogen metabolism during embryogenesis. *New Phytologist*, *169(1)*: 35-44.

ROHANI, O; SHARIFAH, S A; MOHD RAFII, Y; ONG, M; TARMIZI, A H and ZAMZURI, I (2000). Tissue culture of oil palm. *Advances in Oil Palm Research* (Basiron, Y; Jalani, B S and Chan, K W eds.). MPOB, Bangi. p. 238-283.

RUEPP, A; ZOLLNER, A; MAIER, D; ALBERMANN, K; HANI, J; MOKREJS, M; TETKO, I; GULDENER, U; MANNHAUPT, G; MUNSTERKOTTER, M and MEWES, H W (2004). The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucl. Acids Res., 32(18)*: 5539-5545.

SCHMIDT, E D L; GUZZO, F; TOONEN, M A J and DE VRIES, S C (1997). A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. *Development, 124*: 2049-2062.

SEGUÍ-SIMARRO, J M; TESTILLANO, P S; JOUANNIC, S; HENRY, Y and Risueño, M C (2005). Mitogen-activated protein kinases are developmentally regulated during stress-induced microspore embryogenesis in *Brassica napus* L. *Histochem Cell Biol*., *123(4-5)*: 541-551.

TAKEDA, H; KOTAKE, T; NAKAGAWA, N; SAKURAI, N and NEVINS, D J (2003). Expression and function of cell wall-bound cationic peroxidase in asparagus somatic embryogenesis. *Plant Physiol*., *131(4)*: 1765-1774.

THEVISSEN, K; GHAZI, A; DE SAMBLANX, G W; BROWNLEE, C and BROEKAERT, W F (1996). Fungal membrane responses induced by plant defensins and thionins. *The J. Biological Chemistry, 271*: 15018-15025.

TOKUJI, Y and KURIYAMA, K (2003). Involvement of gibberellins and cytokinin in the formation of embryogenic cell clumps in carrot (*Daucus carota*). *J. Plant Physiol*., *160*: 133-141.

TREGEAR, J W; MORCILLO, F; RICHAUD, F; BERGER, A; SINGH, R; CHEAH, S C; HARTMANN, C; RIVAL, A and DUVAL, Y (2002). Characterization of a defensin gene expressed in oil palm inflorescences: induction during tissue culture and possible association with epigenetic somaclonal variation events. *J. Experimental Botany*, *53*: 1387-1396.

TZAFRIR, I; PENA-MURALLA, R; DICKERMAN, A; BERG, M; ROGERS, R; HUTCHENS, S; SWEENEY, T C; MCELVER, J; AUX, G; PATTON, D and DAVID MEINKE, D (2004). Identification of genes required for embryo development in *Arabidopsis*. *Plant Physiol*., *135*: 1206-1220.

WILDE, H D; NELSON, W S; BOOIJ, H; DE VIRIES, S C and THOMAS, T L (1998). Gene expression programs in embryogenic and non-embryogenic carrot cultures. *Planta, 176*: 205-211.

WOOI, K C (1995). Oil palm tissue culture - current practice and constraints. *Proc. of the 1993 ISOPB International Symposium on Recent Developments in Oil Palm Tissue Culture and Biotechnology* (Rao, V; Henson, I E and Rajanaidu, N eds.). PORIM, Bangi. p.1-32.

YOSHIDA, N; YANAI, Y; CHEN, L; KATO, Y; HIRATSUKA, J; MIWA, T; SUNG, Z R and TAKAHASHI, S (2001). EMBRYONIC FLOWER2, a novel polycomb group protein homolog, mediates shoot development and flowering in *Arabidopsis*. *Plant Cell*, *13(11)*: 2471-2481.

ZENG, F; ZHANG, X; ZHU, L; TU L; GUO, X and NIE, Y (2006). Isolation and characterization of genes associated to cotton somatic embryogenesis by suppression substractive hybridization and macroarray. *Plant Molecular Biol., 60*: 167-183.

ZHU, D and SCANDALIOS, J G (1994). Differential accumulation of manganese-superoxide dismutase transcripts in maize in response to abscisic acid and high osmoticum. *Plant Physiol*., *106(1)*: 173-178.